Prepared for submission to JHEP

Quantum Field Theory in Curved Spacetime

John Preskill

California Institute of Technology

Abstract: These lecture notes were prepared for a course taught at Caltech in the spring of 1990. The handwritten notes have been typeset by Nirmalya Kajuri. Preskill and Kajuri have made corrections and edits, but the content of the original lectures has not been updated.

The lectures were inspired by Kip Thorne's suggestion that a nine-week course devoted to quantum field theory in a curved spacetime background would be a Good Thing. The main objective is to explain the semiclassical theory of black hole radiance. Some familiarity with general relativity and quantum mechanics is assumed, but prior knowledge of quantum field theory is not essential. Unfortunately, quantum field theory in de Sitter space was not covered because time ran out.

Contents

1 Introduction

Typical first-year graduate students may take a year-long course on general relativity (GR) and a year-long course on quantum field theory (QFT). Most likely, they hear little about quantum theory in the GR course and little about gravity in the QFT course. Why?

There are practical reasons. Experts in GR are often not experts in QFT, and vice versa. Most interesting physics applications of GR concern phenomena on long-distance scales where QFT has little relevance, such as the structure and evolution of the universe, relativistic astrophysics, astrophysical black holes, or sources of gravitational radiation. Likewise, the exciting physics applications of QFT typically concern phenomena at very short distances, such as high-energy particle collisions. These applications fail to provide a context in which gravity and quantum theory need to be considered simultaneously.

In what sort of context are GR and QFT simultaneously relevant?

A first hint was provided in 1899 by Planck, who noticed that the fundamental constants $\hbar$, along with c and Newton's constant G, allow us to construct fundamental units of distance, time, energy, etc.

Distance: $L_P = \left(\frac{G\hbar}{c^3}\right)^{\frac{1}{2}} \sim 10^{-33}$ cm

Time: $T_P = \frac{L_P}{c} = \left(\frac{G\hbar}{c^5}\right)^{\frac{1}{2}} \sim 10^{-43}\,\mathrm{sec}$

Energy: $E_P = \frac{\hbar}{T_P} = \left(\frac{\hbar c^5}{G}\right)^{\frac{1}{2}} \sim 10^{19}$ GeV

Just on dimensional grounds, then, one expects that gravity and quantum theory are relevant for phenomena at distances $\sim L_P$ or energies $\sim E_P$.

To be more concrete, consider the gravitational contribution to hard (large-angle) scattering of relativistic particles. Merely on dimensional grounds, the cross section is $\sigma \sim G^2E^2$ (in units with $\hbar = c = 1$). This becomes appreciable (comparable to the "unitarity limit" $\sim 1/E^2$) for $E \sim E_P \sim 10^{19}$ GeV. At such energies quantum corrections higher order in $\hbar$ become important. These corrections turn out to be so sensitive to the nature of quantum fluctuations with frequency $\omega \geq E_P$ that the calculability of the theory breaks down. (Einstein's gravity is nonrenormalizable.) This breakdown of calculability at $E \sim E_P$ can indicate that Einstein gravity must be replaced by a more complete underlying theory (maybe superstring theory).

The breakdown of Einstein gravity at $L \sim L_P$ is more profound than, say, the breakdown of Fermi's weak interaction theory at $E \approx 300$ GeV. That's because the dynamical variable of quantum gravity is the spacetime metric itself. Large quantum gravitational fluctuations may mean that the whole notion of spacetime is breaking down. At the very least, it is obscure how to formulate microscopic causality for a theory in which the edge of the lightcone is not precisely defined.

Quite aside from the ambiguities that afflict quantum gravity at "short distances," there are deep conceptual questions concerning the interpretation of a quantum theory of gravity; these stem from the general covariance of the theory. In general relativity, the coordinates that parametrize spacetime have no invariant significance, and so we are free to reparametrize spacetime locally. When we formulate quantum mechanics, we must (*arbitrarily*) choose a time coordinate t and construct a Hamiltonian H that generates time evolution. Now we should distinguish two types of reparametrizations, which have quite distinct physical consequences:

1. We can reparametrize *each* time slice t = constant. This is an example of gauge invariance — it is a redundancy in how the metric ${}^3g$ of the time slice parametrizes the quantum states, like the redundancy in how the vector potential A_μ of electrodynamics parametrizes states. The "wave function" $\psi\left({}^3g\right)$ should be regarded as a functional of the 3-geometry represented by the metric, not the metric itself.

2. We can reparametrize time, or choose a new set of time slices. Because we have the freedom to move the time slice forward or backward locally, the quantum theory will not tolerate *any* time-dependence:

$$i\frac{\partial\psi}{\partial t} = H\psi = 0.$$

   This is the Wheeler-DeWitt equation. There are actually many such constraints, because we have the freedom to move the spacelike slice forward or backward however we please at each point on the slice.

So, naively, the theory has no dynamics, at least with respect to the extrinsic *parameter* time. To formulate dynamics, we must identify an *intrinsic* time. We ask not "What is $P(x,t)$, the probability that the electron is at x at time t?" but instead "What is $P(x|t)$, the conditional probability that the electron is at x when the clock says t?" And furthermore we should specify what we mean by being at position x without relying on any arbitrary choice of coordinates. Such changes are not trivial; for example if "time" is the reading of a clock that is itself a quantum-mechanical dynamical variable, then time itself is subject to quantum fluctuations.

Without a meaningful extrinsic time, it is subtle to construct a self-consistent probability interpretation of quantum mechanics. What we ordinarily do is choose a complete set of commuting observables and consider $|\lambda, t\rangle$ that are eigenstates of all these observables at time t. Then, $P(\lambda,t) = |\langle\lambda, t|\psi\rangle|^2$ is the probability distribution governing the outcomes if the observables are measured. It satisfies a normalization property:

$$\sum_\lambda P(\lambda,t) = \sum_\lambda \langle\psi|\lambda,t\rangle\langle\lambda,t|\psi\rangle = \langle\psi|\psi\rangle = 1$$

because the observables are complete. Because of the "problem of time" in general relativity, it is not so clear how to construct a complete set of observables that obey conservation of probability. There is a "semiclassical" limit in which fluctuations in time can be ignored and the normal theory applies, but we would like quantum gravity to have an interpretation beyond semiclassical theory. And what ensures that semiclassical behavior is ever attained?

Thus, because of the problem of time, it can be challenging to give quantum gravity a fully consistent probability interpretation. This is a serious conceptual problem, quite independent of the short-distance problem.

Since $E \sim E_P \sim 10^{19}$ GeV and $L \sim L_P \sim 10^{-33}$ cm are so remote from our experience, why should we care about such issues? The motivation to understand quantum gravity arises from:

1. The problem of initial conditions of the universe — "quantum cosmology."
2. What is the final state of an evaporating black hole? (Does quantum mechanics break down? Is causality flagrantly violated?)
3. The hope that a better understanding of quantum gravity will elucidate physics in unexpected ways, including consequences more directly related to experiments.

Although we have argued that understanding quantum gravity is important, we will not be quite so ambitious in this course. In the hope of avoiding the conceptual quagmire described above, we will mostly focus on the regime where $L \gg L_P$ and $E \ll E_P$. In this regime, it is an excellent approximation to do quantum field theory while treating gravity as classical.

For example, in Z^0 decay the process $Z \to \mu^+\mu^- +$ hard graviton gives a contribution to the total width of the Z resonance that is suppressed by $\sim Gm_Z^2 \sim 10^{-34}$ — gravity is negligibly weak at low energy. The nonrenormalizability of quantum gravity does not alter this conclusion. The large quantum fluctuations of $L \sim L_P$ *decouple* from the long-distance physics, except for effects that can be absorbed into renormalizing a few free parameters. This decoupling of short-wavelength physics from long-wavelength physics is crucial — without it, quantum field theory, and physics, would hardly be possible at all.

In this course, we will focus on the effect of *classical* gravity on the quantum theory of other fields — in other words, we will consider quantum field theory on a spacetime background with curvature that is non-zero, but small in Planck units. The curvature significantly affects the excitations and fluctuations of the fields on wavelengths comparable to the distance scale set by the curvature.

By considering this low-curvature limit, we will not avoid all subtleties and ambiguities. For example, the notion of a particle is difficult to make precise on a curved background. Correspondingly, it will not always be obvious how to define a "vacuum" state of the quantum fields, or how to construct a Hilbert space built on the vacuum. These are features not encountered in classical GR or in QFT on flat spacetime, and we will need some ingenuity and pragmatism to deal with them.

Further subtleties arise when we consider the backreaction of quantum fields on the spacetime metric. It is simply not consistent to say that the quantum fields fluctuate while spacetime geometry does not, because these fields are a *source* for gravity. In a zeroth-order approximation, we may regard $\langle\psi\,|T_{\mu\nu}|\,\psi\rangle$ as a source of backreaction (where $T_{\mu\nu}$ is the energy-momentum tensor), but this is reasonable only if the fluctuations in $T_{\mu\nu}$ are small enough so that they perturb the geometry only a little. And it is not easy to decide how the operator $T_{\mu\nu}$ appearing in the expectation value is to be quantized.

Our main objective in this course is to understand the classic applications of quantum field theory in curved spacetime:

- Black Hole Radiation [16] — The key to understanding the thermodynamics and entropy of black holes, as well as the (potential) loss of quantum coherence [17].
- Quantum Fluctuations in de Sitter Space [5] — Perhaps the origin of primordial density fluctuations in the early universe that seeded galaxy formation, in the inflationary universe scenario.

In attempting to grasp the above applications, it will also be enlightening to consider

- Rindler Space — The thermal radiation seen in the vacuum by a uniformly accelerated observer [24].

In all these cases, we will be able to discuss the essential features by considering *free* quantum fields with zero spin. A key attribute that all three examples have in common is an *event horizon.*

We will also consider, but only briefly, the backreaction — how the quantum fields act as a source for the spacetime geometry, and how the stress-energy tensor is to be defined as a quantum mechanical operator.

Textbooks by Birrell and Davies [3] and by Fulling [10] cover some of this same material.

2 Quantum Field Theory in Flat Spacetime

Let us consider how quantum theory is reconciled with special relativity. That is, we wish to construct a quantum theory such that:

(i) Physics is frame-independent (Lorentz invariant).

(ii) Relativistic causality is respected (no propagation of information at speed $v > c$).

We note to begin with that there is a certain tension between the principles of quantum theory and of relativity. The uncertainty principle causes a localized wave packet to spread quickly. What will prevent probability from leaking out of the light cone (which would allow information to propagate backward in time, in some frames)? It will turn out that causality requires a rather subtle conspiracy.

We are accustomed to the notion that symmetries in physics simplify physical problems. Thus, we might expect relativistic quantum theory – quantum theory constrained by Lorentz invariance – to be simpler than non-relativistic quantum mechanics. This expectation turns out to be a bit too naive. The reason is that, in a relativistic theory, particle production is possible, and, in fact, an indefinitely large number of particles can, in principle, be produced at sufficiently high energy. Therefore, relativistic quantum theory is inevitably a quantum theory with an infinite number of degrees of freedom. When we formulate perturbation theory, all possible states in the theory can appear as intermediate states. In a theory with an infinite number of degrees of freedom, we typically find:

– The perturbation theory is complicated.

– Perturbation theory suffers from (ultraviolet) divergences. Much of the subtlety of (relativistic) quantum field theory stems from these divergences. We need to understand their origin and how to deal with them.

2.1 Particles

We wish to construct a quantum theory of non-interacting relativistic (spinless) particles. How should we proceed? There are two basic strategies, complementary to one another.

(1) Begin with a Hilbert space of relativistic particle states. Then introduce *fields* (which create and annihilate particles) to construct observables that are localized in spacetime. This is called (for obscure reasons) "second quantization."

(2) Begin with a relativistic (Lorentz invariant) classical field theory. Construct a quantum mechanical Hilbert space by elevating the field to the status of an operator that obeys canonical commutation relations with its conjugate momentum. This is called "canonical quantization." One then finds relativistic particles in the spectrum of this theory.

Adopting either starting point, we are eventually led to the same theory. Strategy (1) is more logical and direct if our initial objective is to obtain a theory of particles. But (2) is reasonable if our goal is to construct a quantum theory with relativistically invariant (and causal) dynamics. Furthermore, when particles interact, the concept of a particle becomes less clear, and the advantages of (2) become more apparent.

Also, if spacetime is curved, the concept of a particle is ambiguous, leading us to favor (2). But we will see that (2) may also suffer from ambiguities in curved spacetime. These ambiguities in a spacetime that is curved, whether (1) or (2) is used, are central conceptual issues in the theory of quantum fields on curved spacetime.

In the case of a flat background, we will describe procedure (1) in some detail and then (2) somewhat more schematically. Then we will go on to try to apply these procedures to field theory on a nontrivial background.

Notation: We will usually set $\hbar = c = 1$. Our spacetime metric has the form

$$\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1).$$

To a particle theorist on a flat background spacetime (like me), a relativistic particle is a unitary irreducible representation of the Poincaré group. The Poincaré group is the semidirect product of the Lorentz group and translation group.

Lorentz transformation:

$$\Lambda : x^\mu \to \Lambda^\mu_\nu x^\nu,$$

$$\eta_{\mu\nu}\Lambda^\mu_\lambda \Lambda^\nu_\sigma = \eta_{\lambda\sigma}.$$

(We will always consider Λ to be a proper Lorentz transformation with $\Lambda^0_0 > 0$, and $\det\Lambda = 1$; these are Lorentz transformations that can be smoothly connected to the identity. That is, parity and time reversal are not considered.)

Translation:

$$a : x^\mu \to x^\mu + a^\mu.$$

The Poincaré transformation (Λ, a) acts as

$$(\Lambda, a) : x \to \Lambda x + a.$$

Thus, the composition law for Poincaré transformations is:

$$(\Lambda_1, a_1) \cdot (\Lambda_2, a_2) = (\Lambda_1\Lambda_2, \Lambda_1 a_2 + a_1),$$

$$x \to \Lambda_2 x + a_2 \to \Lambda_1\Lambda_2 x + \Lambda_1 a_2 + a_1.$$

Now consider a quantum theory that respects Poincaré invariance. It has a Hilbert space, denoted as $\mathcal{H}$, and unitary operators that represent Poincaré transformations acting on $\mathcal{H}$:

$$(\Lambda, a) : |\text{state}\rangle \longrightarrow U(\Lambda, a)|\text{state}\rangle.$$

U is required to be unitary to preserve the inner product: $\langle\psi|\chi\rangle = \langle U\psi|U\chi\rangle$. This is what it means for Poincaré transformations to be symmetries. Suppose we change reference frames twice:

$$\begin{aligned}|\text{state}\rangle &\to U(\Lambda_1, a_1)U(\Lambda_2, a_2)|\text{state}\rangle \\ &= U(\Lambda_1\Lambda_2, \Lambda_1 a_2 + a_1)|\text{state}\rangle.\end{aligned}$$

For consistency, this must be equivalent to:

$$U(\Lambda_1, a_1)U(\Lambda_2, a_2) = U\left((\Lambda_1, a_1) \cdot (\Lambda_2, a_2)\right)$$

(up to a possible physically irrelevant phase factor). This is what it means to say that the unitary operator U's provides a representation of the Poincaré group. Any such representation can be decomposed into irreducible components.

What is the structure of such a representation? First, consider translations $(\mathbb{1}, a)$. Translations are generated by the (Hermitian) momentum operator P:

$$U(a) = e^{iP\cdot a}.$$

The translation group is abelian, and its irreducible representations are one-dimensional. An irreducible representation acts on a state that is a simultaneous eigenstate of all components of P^μ:

$$P^\mu|k\rangle = k^\mu|k\rangle,$$

or

$$U(a)|k\rangle = e^{ik\cdot a}|k\rangle.$$

Here $|k\rangle$ is called a plane-wave state. Clearly, $U(a_1)U(a_2) = U(a_1 + a_2)$, so this is a representation. (If we wish to think of k^μ as the four-momentum of a particle, note that $\hbar$ has already implicitly entered the discussion in the relation between momentum and wave number.)

How shall Lorentz transformations be represented? Let's consider the simplest case, where particles have nonzero mass and a state with $\vec{k} = 0$ is invariant under rotations ("spin-0"). If we think of $|k\rangle$ as a particle with four-momentum k^μ, then we expect that the state $U(\Lambda)|k\rangle$ should have momentum Λk. Indeed, this is required by the Poincaré group structure:

$$\begin{aligned}&U(\Lambda^{-1})e^{iP\cdot a}U(\Lambda) = U(\mathbb{1}, \Lambda^{-1}a) \\ \Rightarrow\ &U(\Lambda^{-1})e^{iP\cdot a}U(\Lambda) = e^{iP\cdot(\Lambda^{-1}a)} = e^{i\Lambda P\cdot a} \\ \Rightarrow\ &U(\Lambda^{-1})PU(\Lambda) = \Lambda P.\end{aligned}$$

So $PU(\Lambda) = U(\Lambda) \cdot (\Lambda P)$, and therefore

$$PU(\Lambda)|k\rangle = (\Lambda k)U(\Lambda)|k\rangle.$$

Our representation of the translation group becomes a representation of the Poincaré group if we choose

$$U(\Lambda)|k\rangle = |\Lambda k\rangle$$

up to a normalization to be discussed in a moment.

The Lorentz transformations preserve the invariant

$$P^\mu P_\mu = m^2,$$

and for our representation to correspond to physical particles, we demand:

$$m^2 \geq 0 \text{ and } P^0 \geq 0.$$

Then, the states $|k\rangle$ with $k^2 = m^2$ and $k^0 > 0$ are the basis for an *irreducible* representation of the Poincaré group: Any k on the mass hyperboloid can be obtained from $k = (m, \vec{0})$ by applying a suitable Lorentz transformation.

The relative normalization of the state $|k\rangle$ for various values of k can be determined from the requirement that $U(\Lambda)$ defined by

$$U(\Lambda)|k\rangle = |\Lambda k\rangle$$

is a unitary operator. Because the states $|k\rangle$ form a complete basis for the representation space, we have:

$$\mathbb{1} = \int d\mu(k)|k\rangle\langle k|$$

for a suitable measure $d\mu$ defined on the hyperboloid. If U is unitary, then:

$$\begin{aligned}\mathbb{1} = U(\Lambda)\mathbb{1}U(\Lambda)^\dagger &= \int d\mu(k)U(\Lambda)|k\rangle\langle k|U(\Lambda)^\dagger \\ &= \int d\mu(k)|\Lambda k\rangle\langle\Lambda k| = \int d\mu\left(\Lambda^{-1}k\right)|k\rangle\langle k|.\end{aligned}$$

Thus, $d\mu(k)$ must be a *Lorentz invariant measure* satisfying:

$$d\mu(k) = d\mu\left(\Lambda^{-1}k\right).$$

The invariant measure is unique up to an overall multiplicative factor and can be written as:

$$d\mu(k) = \frac{d^4k}{(2\pi)^3}\delta\left(k^2 - m^2\right)\theta\left(k^0\right).$$

The measure d^4k is invariant because $\det\Lambda = 1$. Then, $\delta\left(k^2 - m^2\right)$ restricts this measure to the hyperboloid, and the $\theta\left(k^0\right)$ further restricts it to the positive energy hyperboloid. The $(2\pi)^{-3}$ is a convention that fixes the overall normalization.

Since k^0 can be trivially integrated, we may also write:

$$d\mu(k) = \frac{d^3k}{(2\pi)^3 2k^0}, \text{ where } k^0 = \sqrt{\vec{k}^2 + m^2}.$$

From

$$\mathbb{1} = \int \frac{d^3k}{(2\pi)^3 2k^0} |k\rangle\langle k|$$

and

$$\mathbb{1}\left|k'\right\rangle = \int \frac{d^3k}{(2\pi)^3 2k^0} |k\rangle \left\langle k \mid k'\right\rangle = \left|k'\right\rangle,$$

we find the relativistic normalization of states:

$$\left\langle k' \mid k\right\rangle = (2\pi)^3 2k^0 \delta^3(\vec{k'} - \vec{k}).$$

We have now completely specified the basis for the one-particle subspace $\mathcal{H}^{(1)}$ of Hilbert space in a theory of spinless relativistic particles and have defined the action of the Poincaré group on this space. It acts irreducibly. The states in $\mathcal{H}^{(1)}$ are wave packets that can be expanded in terms of the plane wave basis, e.g.:

$$|\tilde{f}(k)\rangle = \int \frac{d^3k}{(2\pi)^3 2k^0} \tilde{f}(k)|k\rangle$$

where

$$U(\Lambda)|\tilde{f}(k)\rangle = \int \frac{d^3k}{(2\pi)^3 2k^0} \tilde{f}(k)|\Lambda k\rangle = \left|\tilde{f}\left(\Lambda^{-1}k\right)\right\rangle$$

and

$$\langle \tilde{f}'|\tilde{f}\rangle = \int \frac{d^3k}{(2\pi)^3 2k^0} \tilde{f}'(k)^* \tilde{f}(k).$$

We have constructed a one-particle Hilbert space, but in anticipation of interactions that might change the number of particles, we will enlarge it to a many-particle Hilbert space. First, we need a vacuum – the zero-particle state. It is the unique state that is Poincaré invariant:

$$P^\mu|0\rangle = 0 \qquad U(\Lambda)|0\rangle = |0\rangle.$$

That is, the vacuum looks the same to all observers. It has the conventional normalization $\langle 0|0\rangle = 1$.

And we need many-particle states. We *assume* that the particles obey Bose statistics, e.g.:

$$|k_1, k_2\rangle = |k_2, k_1\rangle.$$

So, the normalization of the n-particle state $|k_1, k_2, \ldots| \, k_n\rangle$ must respect the permutation symmetry acting on the n momenta. In the n-particle Fock space, the completeness relation becomes:

$$(\mathbb{1})_{n \text{ particle}} = \frac{1}{n!} \int \frac{d^3k_1}{(2\pi)^3 2k_1^0} \cdots \frac{d^3k_n}{(2\pi)^3 2k_n^0} \left|k_1, \ldots, k_n\right\rangle \left\langle k_1, \ldots, k_n\right|,$$

where the $\frac{1}{n!}$ compensates for overcounting of states. The normalization is:

$$\langle k_1, \ldots, k_n | k'_1, \ldots, k'_n \rangle = n! \text{ terms },$$

e.g:

$$\langle k_1 k_2 \mid k'_1 k'_2 \rangle = (2\pi)^3 2k_1^0 (2\pi)^3 2k_2^0 \times \left[\delta^3\left(\vec{k}_1 - \vec{k}'_1\right)\delta^3\left(\vec{k}_2 - \vec{k}'_2\right) + \delta^3\left(\vec{k}_1 - \vec{k}'_2\right)\delta^3\left(\vec{k}_2 - \vec{k}'_1\right)\right].$$

The many-particle states transform as the (reducible) representation of the Poincaré group:

$$U(\Lambda, a) |k_1, \ldots, k_n\rangle = U(a)U(\Lambda) |k_1, \ldots k_n\rangle = e^{i\Lambda(k_1 + \cdots + k_n)\cdot a} |\Lambda k_1, \ldots, \Lambda k_n\rangle.$$

The full Hilbert space is a direct sum:

$$\mathcal{H} = \mathcal{H}^{(0)} \oplus \mathcal{H}^{(1)} \oplus \mathcal{H}^{(2)} \oplus \ldots$$

where $\mathcal{H}^{(0)}$is the zero-particle Hilbert space, $\mathcal{H}^{(1)}$ is the 1-particle Hilbert space, $\mathcal{H}^{(2)}$ is the 2-particle Hilbert space, etc. This (separable) Hilbert space is called *Fock space*.

2.2 Fields

Now we have the complete Hilbert space of the theory, and a representation of the Poincaré group acting on the space. Have we therefore completed the construction? No. This theory, so far, lacks *local* observables – operators with compact support in spacetime. Local observables are constructed from *fields*. A field, technically, is an "operator-valued distribution" — if $\phi(x)$ is a field, then

$$\int d^4x\, \phi(x) f(x)$$

is an operator $\mathcal{H} \to \mathcal{H}$, where f is a suitably smooth test function. The idea of field theory is that quantities that can be measured by an observer localized in spacetime can be modeled as functions of such so-called "smeared" fields.

Observers can communicate by emitting and absorbing particles, so the fields should be able to create or destroy particles — as operators, they mix up the different n-particle spaces $\mathcal{H}^{(n)}$. A further motivation for introducing fields comes from considering interactions among particles. In a consistent theory, particle interactions are typically local in spacetime and admit a natural description in the language of fields.

We will construct a fundamental field of the theory from which all local observables can be constructed. The field $\phi(x)$ is to be regarded as a Heisenberg operator in the theory. (Since fields depend on the spatial position $\vec{x}$, it is natural that they also depend on the time t in a relativistic theory.)

We demand that $\phi(x)$ have the following properties:

(i) ϕ creates or destroys one particle. Then, operators that create or destroy many particles are obtained from polynomials in ϕ.

(ii) $\phi = \phi^\dagger$. We want ϕ to be Hermitian so that smeared fields are observables.

(iii) $U(\Lambda, a)\phi(x)U(\Lambda, a)^{-1} = \phi(\Lambda x + a)$. ϕ transforms as a *scalar* under Poincaré transformations. It is convenient to formulate a relativistic theory in terms of fields that transform as simply as possible. Fields that transform this way will suffice for describing particles with spin-0.

These conditions almost completely fix $\phi(x)$.

To construct $\phi(x)$, it is helpful to first define operators with momentum space arguments that create and destroy particles. Define an operator $A(k)^\dagger$ as follows:

$$
\begin{aligned}
&A(k)^\dagger : \mathcal{H}^{(0)} \to \mathcal{H}^{(1)}, \\
&A(k)^\dagger|0\rangle = |k\rangle,
\end{aligned}
$$

Similarly, define the action of $A(k)^\dagger$ on $\mathcal{H}^{(n)}$ by:

$$
A(k)^\dagger \left|k_1, \ldots, k_n\right\rangle = \left|k, k_1, k_2, \ldots, k_n\right\rangle .
$$

This determines all matrix elements of $A(k)^\dagger$ between states in the Fock space $\mathcal{H}$, and hence also determines all matrix elements of its adjoint $A(k)$. For example:

$$
\begin{aligned}
&\left\langle 0 \left| A(k)^\dagger \right| \text{ arbitrary state} \right\rangle = 0 \Rightarrow A(k)|0\rangle = 0, \\
&\left\langle k' \left| A(k)^\dagger \right| 0 \right\rangle = \langle k'|k\rangle = (2\pi)^3 \left(2k^0\right) \delta^3 \left(\vec{k} - \vec{k}'\right) \\
&\Rightarrow A(k) \left|k'\right\rangle = (2\pi)^3 \left(2k^0\right) \delta^3 \left(\vec{k} - \vec{k}'\right) |0\rangle.
\end{aligned}
$$

Note that

$$
\langle 0| \left[A(k), A\left(k'\right)^\dagger\right] |0\rangle = \left\langle k \mid k'\right\rangle = (2\pi)^3 2k^0 \delta^3 \left(\vec{k} - \vec{k}'\right) \langle 0|0\rangle,
$$

and one can check that

$$
\langle \text{state}'| \left[A(k), A\left(k'\right)^\dagger\right] |\text{state}\rangle = (2\pi)^3 2k^0 \delta^3 \left(\vec{k} - \vec{k}'\right) \langle \text{state}'|\text{state}\rangle
$$

for arbitrary Fock space basis states. Therefore,

$$
\left[A(k), A\left(k'\right)^\dagger\right] | = (2\pi)^3 2k^0 \delta^3 \left(\vec{k} - \vec{k}'\right)
$$

is an operator identity on Fock space. Note also that

$$
|k_1, \ldots, k_n\rangle = A(k_1)^\dagger \ldots A(k_n)^\dagger |0\rangle.
$$

Thus, the assumed Bose symmetry of the many-particle states implies

$$\left[A(k)^\dagger, A\left(k'\right)^\dagger\right] = 0,$$

and hence

$$\left[A(k), A\left(k'\right)\right] = 0.$$

How do the $A(k)$ operators transform under Poincaré transformations? First we consider translations. Using translation invariance of the vacuum:

$$\begin{aligned} U(a)|k\rangle = e^{ik\cdot a}|k\rangle = U(a)A(k)^\dagger U(a)^{-1}U(a)|0\rangle \\ = U(a)A(k)^\dagger U(a)^{-1}|0\rangle = e^{ik\cdot a}A(k)^\dagger|0\rangle. \end{aligned}$$

So,

$$U(a)A(k)^\dagger U(a)^{-1} = e^{ik\cdot a}A(k)^\dagger$$

acting on the vacuum, and similarly acting on any state. Taking adjoints,

$$U(a)A(k)U(a)^{-1} = e^{-ik\cdot a}A(k).$$

Now consider Lorentz transformations, and use Lorentz invariance of the vacuum:

$$\begin{aligned} & U(\Lambda)|k\rangle = |\Lambda k\rangle, \text{ or} \\ & U(\Lambda)A(k)^\dagger U(\Lambda)^{-1}|0\rangle = A(\Lambda k)^\dagger|0\rangle \\ & \Rightarrow \quad U(\Lambda)A(k)^\dagger U(\Lambda)^{-1} = A(\Lambda k)^\dagger \\ & \text{and } U(\Lambda)A(k)U(\Lambda)^{-1} = A(\Lambda k). \end{aligned}$$

The general Hermitian operator that creates or destroys one particle is a linear combination of A's and $A^\dagger$'s:

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3 2k^0}\left[C(k,x)A(k) + C(k,x)^* A(k)^\dagger\right].$$

We would like $\phi(x)$ to transform as a scalar field:

$$\phi(\Lambda x) = U(\Lambda)\phi(x)U(\Lambda)^{-1} = \int \frac{d^3k}{(2\pi)^3 2k^0}\left[C(k,x)A(\Lambda k) + C(k,x)^* A(\Lambda k)^\dagger\right].$$

Using Lorentz invariance of the integration measure, we have

$$= \int \frac{d^3k}{(2\pi)^3 2k^0}\left[C\left(\Lambda^{-1}k, x\right)A(k) + C\left(\Lambda^{-1}k, x\right)^* A(k)^\dagger\right],$$

and we also require

$$\phi(\Lambda x) = \int \frac{d^3k}{(2\pi)^3 2k^0} \left[C(k,\Lambda x)A(k) + C(k,\Lambda x)^* A(k)^\dagger\right].$$

Therefore, $C(k,x)$ must be a Lorentz invariant function:

$$C\left(\Lambda^{-1}k, x\right) = C(k,\Lambda x) \text{ or } C(\Lambda k,\Lambda x) = C(k,x);$$

that is, C is a function of the Lorentz-invariant variables $m^2 = k^2$, x^2, and $k\cdot x$.

Now, under translations,

$$\phi(x+a) = U(a)\phi(x)U(a)^{-1} = \int \frac{d^3k}{(2\pi)^3 2k^0} \left[C(k,x)e^{-ik\cdot a}A(k) + C(k,x)^* e^{ik\cdot a}A(k)^\dagger\right].$$

Using

$$\phi(x+a) = \int \frac{d^3k}{(2\pi)^3 2k^0} \left[C(k,x+a)A(k) + C(k,x+a)^* A(k)^\dagger\right]$$

we see that C must satisfy

$$C(k,x+a) = e^{-ik\cdot a}C(k,x),$$

which determines the x dependence up to a multiplicative constant:

$$C(k,x) = \gamma e^{-ik\cdot x}.$$

In fact, the phase of γ is unphysical. We can absorb the phase by adjusting the phase of $A(k)$, or equivalently, by adopting a different phase convention for the states $|k\rangle$. Therefore, we can take γ to be real and obtain

$$\phi(x) = \gamma \int \frac{d^3k}{(2\pi)^3 2k^0} \left[e^{-ik\cdot x}A(k) + e^{ik\cdot x}A(k)^\dagger\right].$$

So we have determined $\phi(x)$ up to a real normalization constant. This constant quantifies the amplitude for $\phi(x)$ to create a one-particle state:

$$\langle k|\phi(x)|0\rangle = \gamma e^{ik\cdot x}.$$

We can choose $\gamma = 1$ by convention.

The states $\phi(x)|0\rangle$ provide a complete basis for $\mathcal{H}^{(1)}$ that is in a sense conjugate to the plane wave basis. The states

$$|0\rangle, \quad \phi(x_1)|0\rangle, \quad \phi(x_1)\phi(x_2)|0\rangle, \quad \cdots \quad \phi(x_1)\ldots,\phi(x_n)|0\rangle$$

span the subspace of $\mathcal{H}$ with no more than n particles ($\mathcal{H}^{(0)} \oplus \mathcal{H}^{(1)} \oplus \cdots \oplus \mathcal{H}^{(n)}$). Thus, polynomials in the fields, acting on the vacuum, span the Fock space. In this sense, the fields are a *complete* set of local observables.

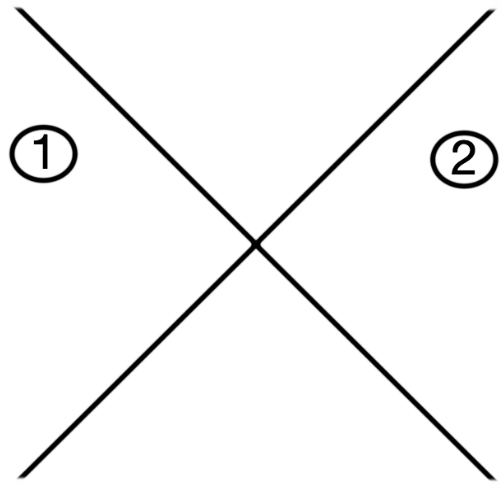

Figure 1. Regions 1 and 2 are spacelike separated.

2.3 Causality

Now we have a relativistic field theory with local observables, but there is one more thing to check — that the theory satisfies relativistic *causality*.

Causality requires that if measurements are performed in regions 1 and 2 that have spacelike separation (as in Figure 1), the measurements in 1 should not affect the outcome of measurements in region 2 and vice versa. In quantum mechanics, this means that observables measured in regions 1 and 2 must *commute*:

$$0 = [\mathcal{O}(1), \mathcal{O}(2)]$$

if regions 1 and 2 are spacelike separated. If all observables can be constructed from smeared fields, then it is necessary and sufficient that

$$[\phi(x), \phi(y)] = 0, \text{ if } (x-y)^2 < 0.$$

Is this true for the fields we constructed? Yes, and for a subtle reason.

Let's decompose $\phi(x)$ into a piece that annihilates particles and a piece that creates particles:

$$\begin{aligned}
\phi(x) &= \phi^{(-)} + \phi^{(+)}, \\
\phi^{(-)} &= \int \frac{d^3k}{(2\pi)^3 2k^0} e^{-ik\cdot x} A(k), \\
\phi^{(+)} &= \int \frac{d^3k}{(2\pi)^3 2k^0} e^{ik\cdot x} A(k)^\dagger.
\end{aligned}$$

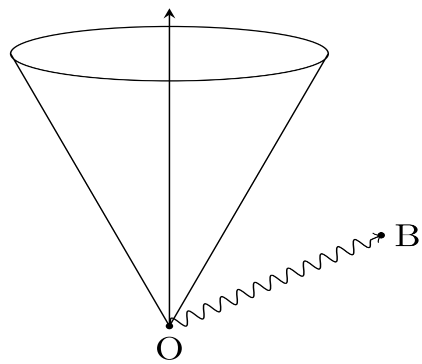

Figure 2. Propagation outside lightcone. OB is a spacelike interval.

Consider the function

$$
\begin{aligned}
G_+(x-y) &= [\phi(x), \phi(y)] = \langle 0|\phi^{(-)}(x)\phi^{(+)}(y)|0\rangle \\
&= \int \frac{d^3k}{(2\pi)^3 2k^0} e^{-ik\cdot(x-y)}.
\end{aligned}
$$

Note that, because of the Lorentz invariance of the measure, G_+ is a Lorentz-invariant (and translation-invariant) function:

$$G_+(\Lambda x) = G_+(x).$$

We could evaluate this integral (and express G_+ in terms of a modified Bessel function), but even without explicitly evaluating it, we can see that $G_+(x)$ does *not* have the property of vanishing for spacelike x. The reason is that G_+ is an analytic function and cannot vanish in an open set without vanishing throughout its domain of analyticity.

G_+ is analytic because of the positivity condition on the energy of a particle, $k^0 > 0$. If we express

$$G_+(x) = \int \frac{d^4k}{(2\pi)^3} \theta\left(k^0\right) \delta\left(k^2 - m^2\right) e^{-ik\cdot x}$$

we can see that G_+ is a sum of entire functions $e^{-ik\cdot x}$, and because k^0 is restricted to be positive, this sum converges for $\operatorname{Im} x^0 < 0$ — it is damped by the exponential. So real x^0 is at the very least on the boundary of the domain of analyticity. In fact, G_+ decays like $\exp\left(-m\sqrt{|x^2|}\right)$ for x outside the light cone, but it is not zero.

This sounds like a potentially serious breach of causality– since $G_+(x) = \langle 0\left|\phi^{(-)}(x)\phi^{(+)}(0)\right| 0\rangle \neq 0$ for x spacelike, the particle excitations localized at $x = 0$ seem to propagate outside the light cone, just as we feared. But can this "propagation" really be detected?

Let's consider a commutator of *observables*:

$$
\begin{aligned}
iG(x-y) = [\phi(x), \phi(y)] &= \left[\phi^{(-)}(x), \phi^{(+)}(y)\right] + \left[\phi^{(+)}(x), \phi^{(-)}(y)\right] \\
&= G_+(x-y) - G_+(y-x).
\end{aligned}
$$

Notice that, because $G_+(x)$ is Lorentz-invariant, it must be an even function, for spacelike x: $G_+(x) = G_+(-x)$. This holds because, when $x = (x^0, \vec{x})$ is spacelike, there is a proper Lorentz transformation that takes $x^0 \to -x^0$ and $\vec{x} \to -\vec{x}$. (Note that G_+ does not have this property for x timelike. It is not invariant under time reversal because only *positive* k^0 appears). Hence,

$$
[\phi(x), \phi(y)] = 0 \quad (x-y)^2 < 0
$$

The fields *are* causal observables.

Causality actually results from a subtle interference effect. There is (in a sense) propagation outside the light cone from x to y, but there is also propagation from y to x. Because the amplitudes for these processes interfere *destructively*, measurements at x and y do not influence one another.

Remarks:

- If we had considered a theory of particles that carry the value Q of a conserved charge, then the propagation of a particle with charge $-Q$ would be necessary to destructively interfere with the propagation of a charge Q particle outside the light cone. Together, causality and the positivity of the energy require, in a relativistic theory, the existence of *antiparticles*.

- Causality severely restricts the algebra of observables in a relativistic theory. With the phase convention that we adopted, we find

  $$
  \phi(x) = \int \frac{d^3k}{(2\pi)^3 2k^0} \left[e^{-ik\cdot x} A(k) + e^{ik\cdot x} A(k)^\dagger\right].
  $$

  But had we adopted a different convention, we could have had

  $$
  \phi^\theta(x) = \int \frac{d^3k}{(2\pi)^3 2k^0} \left[e^{-ik\cdot x} e^{-i\theta} A(k) + e^{ik\cdot x} e^{i\theta} A(k)^\dagger\right].
  $$

  Including both $\phi(x)$ and $\phi^\theta(x)$ as observables in the theory is worse than redundant because ϕ and ϕ^θ are not local relative to one another unless we have $e^{-i\theta} = \pm 1$. We have instead

  $$
  [\phi^\theta(x), \phi(y)] = e^{-i\theta} G_+(x-y) - e^{i\theta} G_+(y-x) = -2i \sin\theta \, G_+(x-y)
  $$

  for $(x-y)^2 < 0$. So we see that the algebra of observables is an exclusive club — no operator may join unless it commutes, at spacelike separation, with all operators that already belong. In particular, ϕ^θ (for $e^{-i\theta} \neq \pm 1$) cannot be admitted if ϕ is already a member.

- We note that

$$\langle 0|\phi(x)\phi(0)|0\rangle - \langle 0|\phi(x)|0\rangle\langle 0\phi(0)|0\rangle \neq 0$$

when x is spacelike This means that the field variables $\phi(x)$ and $\phi(0)$ are *correlated* in the (pure) vacuum state $|0\rangle$. This correlation indicates that the vacuum state exhibits quantum entanglement shared by regions with spacelike separation. However these correlations do not suffice to enable *communication* outside the light cone.

2.4 Some further properties of the field theory

2.4.1 Field equation

Observe that $\phi(x)$ satisfies a covariant wave equation, $\left(\partial^\mu\partial_\mu + m^2\right)\phi(x) = 0$, known as the *Klein-Gordon (KG) equation.* This wave equation is a consequence of the mass shell condition $P^\mu P_\mu = m^2$ satisfied by the particles. The equation has positive-frequency solutions $e^{-ik\cdot x}$ and negative-frequency solutions $e^{ik\cdot x}$, where $k^0 = \sqrt{\vec{k}^2 + m^2}$.

The negative-frequency solutions caused confusion when $\phi(x)$ was interpreted as a single-particle *wave function.* But we have seen that they have a simple and natural interpretation if $\phi(x)$ is an *operator* that creates and destroys particles.

2.4.2 4-momentum operator

The 4-momentum P^μ satisfying $P^\mu|k\rangle = k^\mu|k\rangle$ can be expressed in terms of A and $A^\dagger$. It is given by

$$P^\mu = \int \frac{d^3k}{(2\pi)^3 2k^0} k^\mu A(k)^\dagger A(k).$$

The properties $P^\mu|0\rangle = 0$ and $[P^\mu, A(k)] = -k^\mu A(k)$ (which follows from $e^{iP\cdot a}A(k)e^{-iP\cdot a} = e^{-ik\cdot a}A(k)$) are easily verified.

2.4.3 Conventional normalization

Although it is obviously convenient to expand

$$\phi = \int \frac{d^3k}{(2\pi)^3 2k^0}\left[e^{-ik\cdot x}A(k) + e^{ik\cdot x}A(k)^\dagger\right]$$

in terms of creation and annihilation operators for relativistically normalized states, it is more conventional to introduce

$$a(k) = \frac{A(k)}{(2\pi)^{3/2}(2k^0)^{1/2}}$$

so that

$$[a(k), a(k')^\dagger] = \delta^3(\vec{k} - \vec{k}').$$

Then we have

$$\phi(x) = \int \frac{d^3k}{(2\pi)^{3/2}(2k^0)^{1/2}}[e^{-ik\cdot x}a(k) + e^{ik\cdot x}a(k)^\dagger].$$

2.5 Alternative construction of the Hilbert space

The construction of the one-particle Hilbert space $\mathcal{H}^{(1)}$ can be described in alternative language, which is not needed now but will prove useful when we consider the theory of fields on a curved background.

Note first that when a local field is constructed, the overlap of an arbitrary state with $\phi(x)|0\rangle$, for all x, provides a natural $1-1$ map:

$$\mathcal{H}^{(1)} \simeq \left\{ \textit{positive} \text{ frequency solutions to the Klein-Gordon equation} \right\}.$$

The state

$$|\tilde{f}\rangle = \int \frac{d^3x}{(2\pi)^3 2k^0} \tilde{f}(k)|k\rangle.$$

is associated with the solution

$$\langle 0|\phi(x)|\tilde{f}\rangle = \int \frac{d^3k}{(2\pi)^3 2k^0} \tilde{f}(k) e^{-ik\cdot x}.$$

Equivalently,

$$\langle 0|\phi(x)|\tilde{f}\rangle^* = \langle \tilde{f}|\phi(x)|0\rangle$$

provides the $1-1$ map

$$\mathcal{H}^{(1)} \simeq \left\{ \textit{negative} \text{ frequency solutions to the Klein-Gordon equation} \right\}.$$

This observation can also serve as the starting point of the construction of the scalar quantum field theory. The classical field theory is defined by the classical Klein-Gordon field equation

$$\left(\partial^\mu \partial_\mu + m^2\right) \phi(x) = 0.$$

To obtain a quantum theory, we begin by constructing the "one-particle" Hilbert space $\mathcal{H}^{(1)}$. The general solution to the Klein-Gordon equation can be expanded in the basis

$$\begin{aligned} u_k(x) &= e^{-ik\cdot x} \quad \text{— positive frequency,} \\ u_k(x)^* &= e^{ik\cdot x} \quad \text{— negative frequency.} \end{aligned}$$

If Lorentz transformations act on the solutions according to

$$\Lambda : f(x) \to f\left(\Lambda^{-1}x\right)$$

(the *opposite* of how the quantum field ϕ transforms), then the positive frequency and negative frequency bases each transform irreducibly under the proper Lorentz transformations:

$$\begin{aligned} \Lambda &: u_k(x) \to u_k\left(\Lambda^{-1}x\right) = u_{\Lambda k}(x), \\ \Lambda &: u_k(x)^* \to u_k\left(\Lambda^{-1}x\right)^* = u_{\Lambda k}(x)^*. \end{aligned}$$

The proper Lorentz transformations do not mix up positive and negative frequency, so the positive (or negative) frequency solutions are a linear space on which Lorentz transformations (and translations, too) act irreducibly.

Solutions of the KG equation are also in one-to-one correspondence with initial value data: the solution to $\left(\partial^\mu \partial_\mu + m^2\right) f(x) = 0$ is uniquely determined by values of f and $\dot{f}$ on a surface $x^0 = t =$ constant. But for a solution with only positive (or negative) frequencies, the $\dot{f}$ initial data is not necessary to propagate f from the initial-value surface. There are $1-1$ correspondences:

$$\{\text{ positive frequency solutions }\} \simeq \{\text{ negative frequency solutions }\} \simeq \left\{\text{ functions on } \mathbb{R}^3\right\}.$$

To obtain a Hilbert space, we must specify an inner product. To define the inner product of two solutions, specify a time t and integrate over the time slice:

$$(f,g) = i \int_t d^3x \left[f^*(x)\partial_t g(x) - \partial_t f^*(x) g(x)\right].$$

Then, for the basis $u_k(x) = e^{-ik\cdot x}$, we have

$$\begin{aligned}
&(u_k, u_{k'}) = (2\pi)^3 2k^0 \delta^3(\vec{k} - \vec{k'}), \\
&(u_k, u^*_{k'}) = 0, \\
&(u^*_{k'}, u_k) = -(2\pi)^3 2k^0 \delta^3(\vec{k} - \vec{k'}).
\end{aligned}$$

This Klein-Gordon "inner product" is not positive definite on the space of all solutions. But it is positive definite on the space of positive frequency solutions. Furthermore, (f,g) is negative definite for negative frequency solutions, and the positive and negative frequency solutions are orthogonal to each other. There is a natural decomposition of the space of all solutions:

$$\{\text{solutions}\} = \{\text{positive frequency}\} \oplus \{\text{negative frequency}\},$$

such that the direct sum is a sum of spaces that are *orthogonal* with respect to the Klein-Gordon inner product.

The basis $u_k(x)$ for the positive frequency solutions has precisely the same normalization as does the plane wave basis for $\mathcal{H}^{(1)}$:

$$\langle k|k'\rangle = (2\pi)^3 2k^0 \delta^3(\vec{k} - \vec{k'}),$$

so the $u_k(x)$'s correspond to relativistically normalized one-particle plane-wave states.

In defining (f,g), we seem to have picked out a particular frame and a particular slice. But we can rewrite (f,g) in a form that is manifestly invariant under deformations of the spacelike surface. We write:

$$(f,g) = i \int_\Sigma d^3x \, n^\mu \left[f^* \partial_\mu g - \partial_\mu f^* g\right],$$

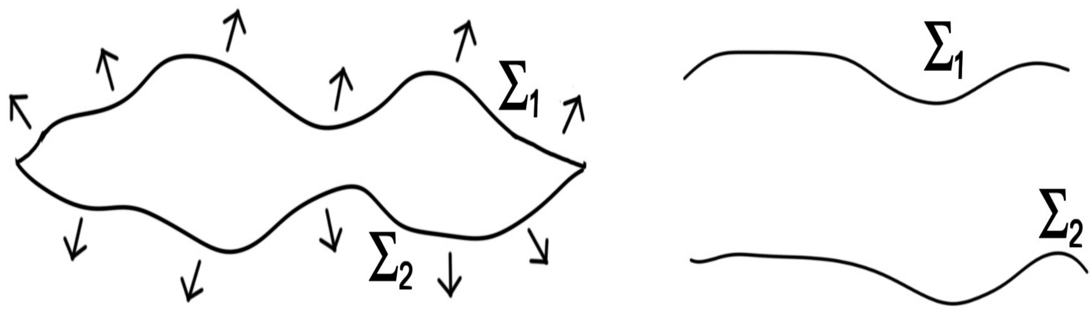

Figure 3. Left: Two hypersurfaces Σ_1, Σ_2 that can be deformed into one another. Right: Two hypersurfaces for which $\Sigma_1 - \Sigma_2$ is not a closed surface.

where we integrate over a spacelike 3-dimensional hypersurface Σ, and n^μ is the normalized unit (timelike) normal to Σ in the forward light cone.

If we distort the surface Σ_1 to a new surface Σ_2, we get:

$$\begin{aligned}(f,g)_{\Sigma_1} - (f,g)_{\Sigma_2} &= i \int_{\Sigma_1 - \Sigma_2} d^3x\, n^\mu \left(f^* \partial_\mu g - \partial_\mu f^* g\right) \\ &= i \int_\Omega d^4x\, \partial^\mu \left(f^* \partial_\mu g - \partial_\mu f^* g\right), \text{ where } \partial\Omega = \Sigma_1 - \Sigma_2\end{aligned}$$

using the divergence theorem. But since both f and g are solutions to the Klein-Gordon equation:

$$\begin{aligned}\partial^\mu \left(f^* \partial_\mu g - \partial_\mu f^* g\right) &= f^* \partial^2 g - \partial^2 f^* g \\ &= (m^2 - m^2) f^* g = 0.\end{aligned}$$

So, $(f,g)_{\Sigma_1} - (f,g)_{\Sigma_2} = 0$. This conclusion holds even if $\Sigma_1 - \Sigma_2$ is not a closed surface or if Σ_1 and Σ_2 are infinite spacelike slices, provided there is no contribution to the integral from spatial infinity (which is true for normalized solutions).

We have now described how, beginning with the classical field equation, we can construct $\mathcal{H}^{(1)}$ as a space spanned by positive frequency solutions. Once we have $\mathcal{H}^{(1)}$, we can obtain $\mathcal{H}^{(n)}$ as a symmetrized tensor product, for example:

$$\mathcal{H}^{(2)} = \mathcal{H}^{(1)} \otimes_s \mathcal{H}^{(1)}.$$

Then we proceed to construct $A(k)$ and $\phi(x)$ as before, obtaining:

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3 2k^0} \left[u_k(x) A(k) + u_k(x)^* A(k)^\dagger\right].$$

2.5.1 Relation to canonical methods

Recall that:

$$[\phi(x), \phi(y)] = G_+(x-y) - G_+(y-x)$$

and therefore

$$[\phi(x), \dot{\phi}(y)] = \int \frac{d^3k}{(2\pi)^3 2k^0}\left[ik^0 e^{-ik\cdot(x-y)} + ik^0 e^{ik\cdot(x-y)}\right],$$
$$[\dot{\phi}(x), \dot{\phi}(y)] = \int \frac{d^3k}{(2\pi)^3 2k^0}(k^0)^2\left[e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\right].$$

If we evaluate these at equal times, $x^0 - y^0 = 0$, we find:

$$[\phi(x), \phi(y)]_{\text{e.t.}} = 0$$
$$[\phi(x), \dot{\phi}(y)]_{\text{e.t.}} = i\delta^3(\vec{x}-\vec{y})$$
$$[\dot{\phi}(x), \dot{\phi}(y)]_{\text{e.t.}} = 0$$

Since $G_+(x)$ is a Lorentz-invariant function, these identities hold in all inertial reference frames. They remind us of the commutation relations

$$[q_i, q_j] = 0,$$
$$[q_i, p_j] = i\delta_{ij},$$
$$[p_i, p_j] = 0.$$

of a canonical quantum mechanical system (with $\hbar = 1$), except with continuum normalization. (The continuum normalization arises because the fields are *distributions*. We could obtain the discrete normalization by smearing the fields with a complete set of square-integrable functions.)

These canonical commutation relations are equivalent to the commutation relations satisfied by the $A(k)$'s and $A(k)^\dagger$'s. If we Fourier transform, we have:

$$\tilde{\phi}(t, \vec{k}) = \int d^3\vec{x} e^{-i\vec{k}\cdot\vec{x}}\phi(t, \vec{x}),$$
$$= \frac{1}{2k^0}\left[e^{-ik^0 t}A\left(k^0, \vec{k}\right) + e^{ik^0 t}A\left(k^0, -\vec{k}\right)^\dagger\right],$$
$$\dot{\tilde{\phi}}(t, \vec{k}) = -\frac{i}{2}\left[e^{-ik^0 t}A\left(k^0, \vec{k}\right) - e^{ik^0 t}A\left(k^0, -\vec{k}\right)^\dagger\right],$$

and hence

$$e^{-ik^0t}A(k) = k^0\tilde{\phi}(t,\vec{k}) + i\dot{\tilde{\phi}}(t,\vec{k}),$$
$$e^{ik^0t}A(k)^\dagger = k^0\tilde{\phi}(t,-\vec{k}) - i\dot{\tilde{\phi}}(t,-\vec{k}).$$

The equal-time commutation relations may be expressed as

$$[\tilde{\phi},\tilde{\phi}]_{\text{e.t.}} = [\dot{\tilde{\phi}},\dot{\tilde{\phi}}]_{\text{e.t.}} = 0$$

and

$$[\tilde{\phi}(t,\vec{k}),\dot{\tilde{\phi}}(t,-\vec{k}')] = i(2\pi)^3\delta^3(\vec{k}-\vec{k}').$$

So the commutation relations

$$\left[A(k), A\left(k'\right)\right] = 0 = \left[A(k)^\dagger, A\left(k'\right)^\dagger\right],$$
$$\left[A(k), A\left(k'\right)^\dagger\right] = (2\pi)^3 2k^0\delta^3\left(\vec{k}-\vec{k}'\right)$$

can evidently be recovered from the equal-time commutators of ϕ and $\dot{\phi}$ as well as vice versa.

To complete the specification of the canonical system, we need a Hamiltonian H expressed in terms of ϕ and its conjugate momentum $\pi = \dot{\phi}$. H is thus P^0, the generator of time evolution satisfying

$$\left[P^0, A(k)\right] = -k^0A(k),$$
$$\left[P^0, A(k)^\dagger\right] = k^0A(k)^\dagger,$$

and therefore

$$[H,\phi(x)] = -i\dot{\phi}(x),$$
$$[H,\pi(x)] = -i\dot{\pi}(x).$$

We have

$$H = \int \frac{d^3k}{(2\pi)^3 2k^0} k^0 A(k)^\dagger A(k)$$

where

$$A(k)^\dagger A(k) = \tilde{\pi}(t,\vec{k})\tilde{\pi}(t,-\vec{k}) + \left(k^0\right)^2\tilde{\phi}(t,\vec{k})\tilde{\phi}(t,-\vec{k}) + ik^0[\tilde{\phi}(t,-\vec{k})\tilde{\pi}(t,\vec{k}) - \tilde{\pi}(t,-\vec{k})\tilde{\phi}(t,\vec{k})],$$

and therefore

$$H = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}\left[\tilde{\pi}(t,\vec{k})\tilde{\pi}(t,-\vec{k}) + \left(k^0\right)^2\tilde{\phi}(t,\vec{k})\tilde{\phi}(t,-\vec{k}) - k^0(2\pi)^3\delta^3(0)\right].$$

Up to an additive constant, this is the Hamiltonian of an infinite set of uncoupled harmonic oscillators, where the oscillator labeled $\vec{k}$ has frequency $\omega_k = k^0 = \sqrt{\vec{k}^2+m^2}$.

To understand the origin of the additive constant, note that $(2\pi)^3\delta^3(0) = \int d^3x(1) = V$ is the spatial volume, and $d^3k/(2\pi)^3$ is the density of oscillator modes per unit volume; therefore the additive constant is

$$H_0 = \text{ constant } = -\sum_{\text{modes}}\left(\frac{1}{2}\omega_k\right)$$

The constant subtracts away the zero-point energy of all the oscillators. We make this subtraction so that $P^\mu|0\rangle = 0$. Otherwise, we would have to split P^μ up into two pieces, where one piece transforms as a four-vector, and the remainder, $\langle 0\,|P^\mu|\,0\rangle$, is invariant under Lorentz transformations.

2.5.2 Canonical quantization

The canonically quantized theory can be arrived at starting from a classical theory defined by an action principle.

To define a relativistic scalar field theory, we may specify that $\phi(x)$ is a scalar field transforming as

$$(\Lambda, a) : \phi(x) \to \phi(\Lambda x + a),$$

and construct the action

$$S = \int d^4x\, \mathcal{L}(\partial_\mu \phi(x), \phi(x))$$

We require:

- S is local and a functional of ϕ and its first derivatives, so that the initial value problem is well formulated.
- $\mathcal{L}$ is Poincaré invariant (frame-independent dynamics).
- $\mathcal{L}$ is quadratic in ϕ and $\partial^\mu\phi$ (linear equation of motion, i.e., free field theory).

The Lagrangian density satisfying these requirements is:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2.$$

This is unique (for $m^2 \neq 0$) – up to a linear redefinition of ϕ. We may remove any term linear in ϕ by making the replacement $\phi \to \phi + b$. Positivity of energy will require $m^2 \geq 0$ and a positive coefficient of the kinetic term $\frac{1}{2}\partial^\mu\phi\partial_\mu\phi$ term. We can fix the normalization of the kinetic term with a multiplicative rescaling $\phi \to c\phi$. For $m^2 = 0$, no linear term is allowed.

We obtain field equations from the action principle:

$$\delta S = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\delta\partial_\mu\phi\right] = \int d^4x\delta\phi \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\right] + \text{ surface term}$$
$$\Rightarrow \frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi} = 0 \text{ (if } \delta S = 0 \text{ for arbitrary variations that vanish at the boundary).}$$

For the above $\mathcal{L}$, we obtain $\left(\partial_\mu\partial^\mu + m^2\right)\phi(x) = 0$ — the Klein-Gordon equation.

We construct the canonical Hamiltonian

$$H = \sum_i \dot{q}_i p_i - L, \quad p_i = \frac{\partial L}{\partial\dot{q}_i},$$

which in the field theory becomes

$$H = \int d^3x(\dot{\phi}\pi - \mathcal{L}), \quad \pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}$$

where $\dot{\phi}$ is eliminated in favor of π.

For the free scalar field theory, we have:

$$\pi(x) = \dot{\phi}(x) \text{ and } H(\phi, \pi) = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\vec{\nabla}\phi)^2 + \frac{1}{2}m^2\phi^2\right].$$

We obtain a quantum field theory by requiring that ϕ and π obey canonical commutation relations at equal times (they are to be regarded as Heisenberg operators):

$$[\phi(x), \phi(y)]_{\text{e.t.}} = [\pi(x), \pi(y)]_{\text{e.t.}} = 0,$$
$$[\phi(x), \pi(y)]_{\text{e.t.}} = i\delta^3(\vec{x} - \vec{y}).$$

The field variables $\phi(t, \vec{x})$ at fixed t may be regarded as a complete set of commuting observables for the canonical system.

If we expand ϕ and $\dot{\phi} = \pi$ in terms of A and $A^\dagger$, we obtain again the quantum theory discussed previously. We chose a particular frame in which to canonically quantize, but the commutation relations are the same in all inertial frames, which is crucial for ensuring that the theory is causal ($\phi(x)$ and $\phi(y)$ commute at spacelike separation).

Since the theory is Poincaré invariant, we can construct conserved quantities by the Noether procedure. For instance, from translational invariance, we obtain:

$$\partial_\nu T^{\mu\nu} = 0, \text{ where } T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}$$

is the canonical energy-momentum tensor. We will not derive this here, as we will soon obtain $T^{\mu\nu}$ by another method.

The conserved four-momenta are then:

$$P^\mu = \int d^3x T^{\mu 0}$$

There is an ordering ambiguity in $P^0 = H$, which, as we noted before, can be resolved by demanding $\langle 0 | P^\mu | 0 \rangle = 0$. The corresponding ordering prescription is called *normal ordering.*

Remarks:

- Canonical quantization is a noncovariant procedure. We must choose a frame to define a Hamiltonian H. But we have seen that the theory that we obtain *is* covariant (admits a unitary representation of the Poincaré group).

- The canonical quantization method gives a quantum theory that agrees with a given classical theory in the classical limit. Hamilton's equations are satisfied at the operator level:

$$\begin{aligned}\dot{q} &= -i[q, H] = \frac{\partial H}{\partial p}, \quad \left(q \sim i\frac{\partial}{\partial p}\right), \\ \dot{p} &= -i[p, H] = -\frac{\partial H}{\partial q}, \quad \left(p \sim -i\frac{\partial}{\partial q}\right).\end{aligned}$$

  Thus, it is natural to apply this procedure to obtain a quantum version of a field that is observable classically — e.g., the electromagnetic field. Canonical quantization is less natural if we are trying to devise a relativistic theory of pions or electrons.

- Canonical quantization has an important advantage over other procedures — it is easily applied to an interacting (nonlinear) theory. In contrast, our construction of the Klein-Gordon inner product required that the field equation be linear (in order to be slice independent) and our Fock space construction assumed that the particles are non-interacting.

3 Quantum Field Theory in Curved Spacetime

Our experience with field theory on flat spacetime has prepared us to confront the problem of constructing quantum fields on a curved background.

We will rely on the idea that we usually invoke to promote flat-spacetime physics to generally covariant physics. Locally (at sufficiently short distances), spacetime is approximately flat. Our field theory on curved spacetime should reduce to flat-spacetime physics locally. Because we must consider local physics, it is essential that we describe quantum field theory in terms of local field variables.

If our flat spacetime theory is causal, then the theory on curved spacetime will be causal if it reduces to the flat theory locally. If information does not propagate outside the light cone *locally*, then it stays within the light cone *globally*.

The concept of a particle, which was fundamental in our discussion on flat spacetime, is less essential in the formulation of field theory on a curved background. A particle, as we defined it, is an irreducible representation of the Poincaré group. But a curved background will not be Poincaré invariant, nor will the quantum theory built on it admit a unitary representation of the Poincaré group. The notion of a particle is an approximate one, valid when the wavelength is much smaller than the length scale characteristic of the curvature. In practice, this limitation need not be serious. For example, the width of the Z^0 boson is not very sensitive to the Hubble constant H_0.

3.1 Classical scalar field in curved spacetime

To begin, we construct the Klein-Gordon classical field theory on a nontrivial background. The flat spacetime action

$$S = \int d^4x \, \frac{1}{2} \left[\eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - m^2 \phi^2 \right]$$

may be written as

$$S = \int d^4x \sqrt{g} \frac{1}{2} \left[g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - m^2 \phi^2 \right]$$

where $g = -\det(g_{\mu\nu})$, so that $d^4x\sqrt{g}$ is the invariant volume element. In this form, S is invariant under local coordinate transformations

$$x \to x'(x)$$

where ϕ is a scalar transforming as

$$\phi(x) \to \phi\left(x'\right).$$

From this action, we derive the Euler-Lagrange equation:

$$\begin{aligned}
&\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \\
&\text{where} \quad \mathcal{L} = \sqrt{g}\left(\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} m^2 \phi^2\right) \\
&\Rightarrow \quad \frac{1}{\sqrt{g}} \partial_\mu \left(\sqrt{g} g^{\mu\nu} \partial_\nu \phi\right) + m^2 \phi = 0.
\end{aligned}$$

We can put this equation in a more recognizable form by invoking the identity

$$\frac{1}{\sqrt{g}} \partial_\mu \sqrt{g} = \Gamma^\lambda_{\lambda\mu}$$

where Γ is the Christoffel symbol and the repeated index is summed. Thus, we have

$$\partial_\mu \partial^\mu \phi + \Gamma^\lambda_{\lambda\mu} \partial^\mu \phi + m^2 \phi = 0$$

or

$$\boxed{\left(\nabla_\mu \nabla^\mu + m^2\right) \phi = 0}$$

where the covariant derivative ∇_μ of a scalar is

$$\nabla_\mu \phi = \partial_\mu \phi$$

and of a 4-vector is

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda,$$

so that

$$\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\lambda_{\lambda\mu} V^\mu$$

is the covariant divergence. The boxed equation is the same as the flat space KG equation, but with derivatives replaced by covariant derivatives.

To derive the identity, recall that for any matrix M,

$$\begin{aligned}
\ln \det(M + \delta M) &= \operatorname{tr} \ln(M + \delta M) \\
&= \operatorname{tr}[\ln M + \ln(\mathbb{1} + M^{-1} \delta M)] \\
&= \operatorname{tr} \ln M + \operatorname{tr} M^{-1} \delta M + O(\delta M^2)
\end{aligned}$$

or

$$\begin{aligned}
\frac{\delta(\det M)}{\det M} &= \operatorname{tr} M^{-1} \delta M \\
\Rightarrow \delta \sqrt{\det M} &= \frac{1}{2} \sqrt{\det M} \operatorname{tr} M^{-1} \delta M;
\end{aligned}$$

thus

$$\delta \sqrt{g} = \frac{1}{2} \sqrt{g} g^{\mu\nu} \delta g_{\mu\nu},$$

and therefore

$$\frac{1}{\sqrt{g}}\partial_\mu\sqrt{g} = \frac{1}{2}g^{\lambda\nu}\partial_\mu g_{\lambda\nu}.$$

Recalling

$$\begin{aligned}\Gamma^\lambda_{\mu\nu} &= \frac{1}{2}g^{\lambda\sigma}[\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}] \\ \Rightarrow \Gamma^\lambda_{\lambda\nu} &= \frac{1}{2}g^{\lambda\sigma}[\partial_\lambda g_{\nu\sigma} + \partial_\nu g_{\lambda\sigma} - \partial_\sigma g_{\lambda\nu}] \\ &= \frac{1}{2}g^{\lambda\sigma}\partial_\nu g_{\lambda\sigma} \text{ (first and third terms cancel)},\end{aligned}$$

we have $\frac{1}{\sqrt{g}}\partial_\mu\sqrt{g} = \Gamma^\lambda_{\lambda\mu}$.

It is convenient to formulate the field theory in terms of an action principle because then we can easily extract the stress tensor that acts as a source in the Einstein equation. If the action for gravity coupled to matter is

$$S = S_{\text{grav}} + S_{\text{matter}}$$

where

$$S_{\text{grav}} = \frac{-1}{16\pi G}\int d^4x\sqrt{g}R$$

and we vary the metric

$$g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu},$$

then

$$\delta S_{\text{grav}} = \frac{-1}{16\pi G}\int d^4x\sqrt{g}G^{\mu\nu}\delta g_{\mu\nu},$$

where

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R$$

is the trace-reversed Ricci tensor, also called the Einstein tensor. If we now define a stress tensor $T^{\mu\nu}$ by

$$\delta S_{\text{matter}} = -\frac{1}{2}\int d^4x\sqrt{g}T^{\mu\nu}\delta g_{\mu\nu},$$

then the equation of motion is

$$G^{\mu\nu} = -8\pi G T^{\mu\nu}$$

— the *Einstein equation*. To derive $T^{\mu\nu}$ from the matter action of a free scalar field

$$S = \int d^4x\sqrt{g}\frac{1}{2}\left[g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - m^2\phi^2\right],$$

we use

$$\delta\sqrt{g} = \frac{1}{2}\sqrt{g}g^{\mu\nu}\delta g_{\mu\nu}$$

and

$$\delta g^{\sigma\rho} = -g^{\sigma\mu}g^{\rho\nu}\delta g_{\mu\nu}$$

which follows from $\delta M^{-1} = -M^{-1}\delta M M^{-1}$. Thus

$$T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \frac{1}{2} g^{\mu\nu} \left[g^{\lambda\sigma} \partial_\lambda \phi \partial_\sigma \phi - m^2 \phi^2 \right]$$

The coordinate invariance of S_{matter} implies that energy-momentum is covariantly conserved: $\nabla_\nu T^{\mu\nu} = 0$.

3.2 Quantization in curved spacetime

Now we want to construct a Hilbert space such that the fields acting on this space are causal. To do this, we will use the *solutions* to the classical KG equation, on which a natural "inner product" can be defined. Here the quotes remind us that this "inner product" is not positive definite, as we have already discussed for the case of a flat background.

Let f, g be solutions to

$$\left(\nabla_\mu \nabla^\mu + m^2 \right) f(x) = 0.$$

The Klein-Gordon inner product of two solutions is defined by choosing a spacelike surface Σ, and integrating:

$$(f, g) = i \int_\Sigma d^3x \sqrt{h} n^\mu \left[f^* \partial_\mu g - (\partial_\mu f^*) \, g \right]$$

where h_{ij} is the induced 3-metric on the surface Σ, and n^μ is the forward-pointing unit normal on Σ. This formula has the desirable property of being independent of the slice Σ. This follows from the divergence theorem, which can be written in the form

$$\int_{\mathcal{M}} d^4x \sqrt{g} \nabla^\mu V_\mu = \int_{\partial \mathcal{M}} d^3x \sqrt{h} n^\mu V_\mu$$

where V_μ is a 4-vector, and $\mathcal{M}$ is a four-manifold with boundary $\partial\mathcal{M}$. (Remember that $\nabla^\mu V_\mu = \frac{1}{\sqrt{g}} \partial^\mu \left(\sqrt{g} V_\mu \right)$ and that g reduces to h in an orthonormal coordinate system with $n^\mu = (1, 0, 0, 0)$.)

Hence, if Σ_1 and Σ_2 are two spacelike slices with the same boundary ($\Sigma_1 - \Sigma_2$ is a closed 3-surface),

$$\begin{aligned}
(f, g)_{\Sigma_1} - (f, g)_{\Sigma_2} &= \int_{\Sigma_1 - \Sigma_2} d^3x \sqrt{h} n^\mu \left(f^* \partial_\mu g - \partial_\mu f^* g \right) \\
&= \int_\Omega d^4x \sqrt{g} \nabla^\mu \left(f^* \nabla_\mu g - \nabla_\mu f^* g \right) \quad (\text{where } \partial\Omega = \Sigma_1 - \Sigma_2) \\
&= 0 \quad (\text{if } f,\ g \text{ solve the Klein-Gordon equation}).
\end{aligned}$$

This is just as in the flat space case. As in that case, the slice independence continues to hold even if $\Sigma_1 - \Sigma_1$ is not closed, provided that the contribution to the integral at spatial infinity is negligible.

Remarks

- It is, of course, implicit in this construction that the Klein-Gordon equation has solutions that are globally defined in spacetime. For this purpose, it suffices that the spacetime be "globally hyperbolic" — i.e., that the spacetime has a *Cauchy surface*.

  What is a "Cauchy surface"? It is a spacelike 3-surface Σ such that every timelike or null curve in the spacetime (without past or future endpoints) intersects Σ once and only once.

  This definition tells us that the initial data on Σ completely determines the physics in the future and past of Σ; that is, initial data on Σ is sufficient to determine the solution to the Klein-Gordon equation throughout the spacetime.

  In fact, it turns out that if the spacetime has one Cauchy surface, then there is a Cauchy surface through every point. Furthermore, we can choose a time coordinate t so that each $t =$ constant surface is Cauchy. (Technically, a "time coordinate" is a smooth scalar function $t(x)$ such that $(\partial_\mu t)\,(\partial^\mu t) > 0$.) We say that a spacetime with a Cauchy surface can be *foliated* by time slices.

  The globally hyperbolic spacetimes, then, share with flat space the property that

$$\begin{aligned}&\{\text{global solutions to the KG equation}\}\\ &\simeq \{\text{initial data on a ``time slice'' } \Sigma\}.\end{aligned}$$

  (Helpful references on global hyperbolicity include: Wald [28] chapters 8, 10, and Geroch and Horowitz [12].)

- The existence of a Klein-Gordon inner product that does not depend on the slice Σ also holds for a free field on a curved background coupled to external sources such that the field equation for ϕ is linear and homogeneous (the action S is quadratic). For example, it applies to the Klein-Gordon field coupled to an external electromagnetic field.

- Even restricting to a free scalar field on a nontrivial geometry, the action that we constructed is not the most general one that is quadratic in ϕ. For example, we could have included in the action the term

$$S' = \int d^4x \sqrt{g} \left(-\frac{1}{2} \xi R(x) \phi^2(x) \right),$$

  where R is the scalar curvature. Then the field equation would become

$$\left[\nabla^\mu \nabla_\mu + m^2 + \xi R(x) \right] \phi(x) = 0.$$

  More complicated terms in the field equation involving other invariants constructed from the curvature could also be included (but would be expected to be suppressed by powers of $G \sim M_{\text{Planck}}^{-2}$ and would be negligible for curvature small in Planck units). For any such field equation, we can construct a Klein-Gordon inner product. But for now, we will ignore any additional dependence on the background geometry and continue to consider the equation

$$\left(\nabla^\mu \nabla_\mu + m^2\right) \varphi = 0,$$

which already serves to illustrate some of the key features of QFT on a nontrivial curved background.

- In constructing a quantum theory, we will initially ignore the "backreaction" of the field ϕ on the geometry, expressed by $G^{\mu\nu} = -8\pi G T^{\mu\nu}$. We will consider the geometry to be a classical source that is not influenced by the (quantum) field ϕ. Later, we will briefly discuss some backreaction effects.

3.3 Canonical quantization

On a globally hyperbolic spacetime, which has a well-behaved globally defined time coordinate, we can perform canonical quantization of the classical free Klein-Gordon theory. We choose time slices and then impose

$$\begin{aligned} [\phi(x), \dot{\phi}(y)]_{\text{e.t.}} &= i\delta^3(\vec{x} - \vec{y}) \\ [\phi(x), \phi(y)]_{\text{e.t.}} &= 0 = [\dot{\phi}(x), \dot{\phi}(y)]_{\text{e.t.}} \end{aligned}$$

We may put this in a more covariant-looking form by denoting a Cauchy surface by Σ, and letting n^μ be the normalized forward-pointing normal to Σ. Then

$$\begin{aligned} [\phi(x), n^\mu \partial_\mu \phi(y)]_\Sigma &= i\frac{1}{\sqrt{h}}\delta^3(\vec{x} - \vec{y}) \\ [\phi(x), \phi(y)]_\Sigma &= 0 \\ [n^\mu \partial_\mu \phi(x), n^\nu \partial_\nu \phi(y)]_\Sigma &= 0 \end{aligned}$$

(Here all commutators are evaluated for two fields on the slice Σ, h is the induced 3-metric on Σ, and $\frac{1}{\sqrt{h}}\delta^3(\vec{x} - \vec{y})$ is the appropriate δ-function normalization for integration against the invariant induced volume element $\mathrm{d}^3 x\sqrt{h}$.)

In fact, if we impose the canonical commutation relations (CCR) on any spacelike Σ, then they are automatically satisfied on every other spacelike surface, provided that $\phi(x)$ satisfies the KG equation. To see this, we first show that ϕ satisfies that CCR on Σ if and only if $[(f,\phi)_\Sigma, (g,\phi)_\Sigma] = -(f, g^*)_\Sigma$ for any two solutions f and g to the KG equation. To show the only if part of this statement, recall

$$(f,g)_\Sigma \equiv i\int_\Sigma \mathrm{d}^3 x\sqrt{h} n^\mu \left(f^* \partial_\mu g - \partial_\mu f^* g\right),$$

and use the CCR to evaluate

$$
\begin{aligned}
&[(f,\phi)_\Sigma,(g,\phi)_\Sigma] \\
&= \left[i\int_\Sigma d^3x\sqrt{h}n^\mu \left(f^*\partial_\mu\phi - \partial_\mu f^*\phi\right), i\int_\Sigma d^3x'\sqrt{h'}n^\nu \left(g^*\partial_\nu\phi - \partial_\nu g^*\phi\right)\right] \\
&= -\int_\Sigma d^3x\sqrt{h}\int_\Sigma d^3x'\sqrt{h'}f^*(x)(-n^\nu\partial_\nu g^*(x'))[n^\mu\partial_\mu(\phi(x),\phi(x')] \\
&\quad - n^\mu\partial_\mu f^*(x)g^*\left(x'\right)\left[\phi(x),n^\nu\partial_\nu\phi(x')\right] \\
&= -i\int_\Sigma d^3x\sqrt{h}n^\mu\left(\partial_\mu f(x)^* g^*(x) - f^*(x)\partial_\mu g^*(x)\right) \\
&= -(f,g^*)_\Sigma .
\end{aligned}
$$

To see the if part, we note that we can choose f, g and their normal derivatives to be arbitrary functions on Σ. From this initial data we obtain solutions to KG equation that are globally defined in spacetime, if the spacetime is globally hyperbolic. Now, if ϕ satisfies the KG equation, we can use the slice independence of the inner product on Σ and show that ϕ satisfies CCR on Σ if and only if ϕ satisfies CCR on Σ', where Σ and Σ' are any two spacelike slices.

Thus for a field that satisfies the KG equation, we find that:

(i) If the CCR are imposed on any spacelike slice, they hold on all slices. The quantum field theory does not depend on how we slice the spacetime.

(ii) The theory is causal in the sense that

$$[\phi(x),\phi(y)] = 0$$

whenever there is a spacelike slice that contains both x and y. (For globally hyperbolic spacetimes, this slice presumably exists if there are no timelike or null curves connecting x and y.)

Now we wish to proceed with the construction of the Fock space, the Hilbert space of this theory. As was emphasized in the discussion of the flat spacetime case, the construction of the Fock space does not require that the notion of a particle apply. It is enough to be able to divide the space of solutions of the KG equation into subspaces with positive-definite and negative-definite KG norm.

The existence of a *complete* basis for the solutions to the KG equation whose inner product on Σ satisfies

$$
\begin{aligned}
(u_i,u_j)_\Sigma &= \delta_{ij} \\
\left(u_i,u_j^*\right)_\Sigma &= 0 \\
\left(u_i^*,u_j^*\right)_\Sigma &= -\delta_{ij}
\end{aligned}
$$

follows from the property that initial data on Σ determines a unique solution, because arbitrary initial data can be expanded in such a basis. (Note that we are now using a schematic notation in which discrete normalization of the solutions is assumed.) This is particularly easy to see in the case where the slice Σ has the topology of $\mathbb{R}^3$. In that case, we can smoothly deform the geometry so that it is flat in the sufficiently distant past of Σ. Then we can choose such a basis in the flat region and propagate it ahead to Σ to get the desired basis in the vicinity of Σ. (The KG inner product is basis independent and depends only on the solution and its first derivative on the slice.)

Now, if we expand the field ϕ in terms of this basis

$$\phi = \sum_i \left(u_i a_i + u_i^* a_i^\dagger \right)$$

we have

$$\begin{aligned} (u_i, \phi) &= a_i, \\ -\left(u_i{}^*, \phi\right) &= a_i{}^\dagger. \end{aligned}$$

evaluated on any slice. Therefore, the CCR in the form

$$[(f, \phi), (g, \phi)] = -\left(f, g^*\right)$$

implies

$$\begin{aligned} [a_i, a_j] &= -\left(u_i, u_j^*\right) = 0, \\ [a_i^\dagger, a_j^\dagger] &= -\left(u_i^*, u_j\right) = 0, \\ [a_i, a_j^\dagger] &= -\left(u_i, -u_j\right) = \delta_{ij}. \end{aligned}$$

That is, the $a_i, a_i^\dagger$ are conventionally normalized creation and annihilation operators. And since $[(f, \phi), (g, \phi)] = -\left(f, g^*\right)$ for a complete basis of solutions implies the CCR on any spacelike Σ, we could just as well assume the a, $a^\dagger$ commutators and then infer the canonical commutators.

We see that an alternative to canonical quantization is to choose a complete basis $\{u, u^*\}$ of solutions of the KG equation, and construct $\mathcal{H}^{(1)}$, the Hilbert space spanned by the u_i. Then extend to the Fock space $\mathcal{H}$ and define a_i on $\mathcal{H}$. Finally, we can construct the field ϕ on $\mathcal{H}$, which we have shown is local, i.e. satisfies $[\phi(x), \phi(y)] = 0$ if x, y lie in the same spacelike slice.

In the construction of the Fock space basis, there is, however, an ambiguity, because there are many ways to choose the basis $\{u_i\}$ of solutions with positive Klein-Gordon norm. For example, if u satisfies

$$(u, u) = 1, \quad (u, u^*) = 0, \quad (u^*, u^*) = -1$$

then

$$\begin{aligned} u' &= \cosh\theta\, u + \sinh\theta\, u^* \\ u'^* &= \sinh\theta\, u + \cosh\theta\, u^* \end{aligned}$$

satisfies

$$\big(u', u'\big) = \cosh^2\theta - \sinh^2\theta = 1$$
$$\big(u', u'^*\big) = \cosh\theta\sinh\theta - \sinh\theta\cosh\theta = 0$$

This linear combination of positive and negative norm solutions is just as acceptable as a basis as u.

There is a *natural* way to decompose the space of solutions to the KG equation into subspaces on which the KG inner product is positive definite and negative definite, respectively, only in the special case of a stationary spacetime (like flat space). We say that the spacetime is stationary if the time coordinate t can be chosen so that the metric is t-independent. Then $\frac{\partial}{\partial t}$ generates a *symmetry* of the geometry and is said to be a timelike "Killing vector" (the spacetime is invariant under time translations).

Since time translation is a symmetry of the equation, if $u(t,x)$ is a solution to the KG equation then so is

$$u(t+dt,\vec{x}) = u(t,\vec{x}) + dt\frac{\partial}{\partial t}u(t,\vec{x}).$$

Therefore, the solutions transform as a representation of the time translation group, and we may decompose this representation into one-dimensional *irreducible* subrepresentations. In other words, the operator

$$H = i\frac{\partial}{\partial t}$$

preserves the space of solutions, and we can diagonalize H on this space:

$$Hu_k = \omega_k u_k,$$
$$Hu_k^* = -\omega_k u_k^*,$$

where $\omega_k \geq 0$ is the frequency of the solution. The solutions then have the form $u_k(t,x) = e^{-i\omega_k t}v_k(\vec{x})$.

Suppose f and g are two solutions. From $(f,g) = i\int d^3x\sqrt{h(t)}\Big(f^*(t)\dot{g}(t) - \dot{f}^*(t)g(t)\Big)$, we see that, since the inner product is independent of t and we have for $\dot{h}(t)=0$,

$$\begin{aligned}(f(t),g(t))_t &= (f(t+dt),g(t+dt))_t\\ &= (f,g) + dt[(\dot{f},g)+(f,\dot{g})]\\ &\Rightarrow (\dot{f},g)+(f,\dot{g}) = 0.\end{aligned}$$

So for solutions of definite frequency, $\dot{u}_k = -i\omega_k u_k$, we have

$$i(\omega_k - \omega_j)(u_k,u_j) = 0.$$

This shows that solutions of distinct frequency are *orthogonal* in the KG inner product. Since the inner product

$$(u_k,u_j)_{t=0} = (\omega_k+\omega_j)\int_{t=0} d^3x\, v_k(\vec{x})^* v_j(\vec{x})$$

is evidently positive definite for positive frequency solutions, we have found a basis (the positive frequency basis) such that

$$(u_i, u_j) = \delta_{ij},$$
$$(u_i, u_j^*) = 0,$$
$$(u_i^*, u_j^*) = -\delta_{ij}.$$

Acting on this basis, the operator H is nonnegative. If we define operators $\{a_i\}$ and a state $|0\rangle$ such that

$$a_i|0\rangle = 0, \quad |i\rangle = a_i^\dagger|0\rangle \text{ where } \langle i|j\rangle = \delta_{ij},$$

we have

$$[a_i, a_j^\dagger] = \delta_{ij}, \quad [a_i, a_j] = 0, \quad [a_i^\dagger, a_j^\dagger] = 0$$

acting on Fock space

$$\mathcal{H}^{(0)} \oplus \mathcal{H}^{(1)} \oplus \mathcal{H}^{(2)} \oplus \cdots$$

and the Hamiltonian H can be diagonalized as

$$H = \sum_k \omega_k a_k^\dagger a_k.$$

But in general, in the case of a nonstationary spacetime, there is no natural choice for the positive norm subspace of the space of KG solutions, and hence no natural Fock space vacuum. The "correct" choice would then be motivated not by mathematics but by the *physical* question we are trying to answer in a particular context.

In the case of a black hole, the geometry is stationary, but it is stationary only outside the event horizon. (Inside the horizon, the light cones tip inward, and the Killing vector becomes spacelike.) In this case, the choice of a vacuum for the stationary region outside the horizon amounts to imposing appropriate boundary conditions on the fields at the horizon. We will discuss this point at length in Chapter 5.

3.4 The S-matrix

A situation in which the spacetime is not stationary, but there are (two) natural choices for a Fock space vacuum is the case of a spacetime that becomes asymptotically stationary in the past or future (or both). This is shown in Figure 4.

Consider in particular the case where the spacetime becomes flat in the asymptotic past and future. Then if $\{u_i, u_i^*\}$ is our standard basis for the solutions to the KG equation on flat space, we may choose solutions to the exact KG equation such that

$$p_i \to u_i \text{ in the } \textit{past.}$$

Then $(p_i, p_j) = \delta_{ij}, (p_i, p_j^*) = 0 \quad (p_i^*, p_j^*) = -\delta_{ij}$ on any slice, since this is true for a slice in the asymptotic past, and the inner product is independent of the slice. Similarly we may choose a basis

$$f_i \to u_i \text{ in the } \textit{future}$$

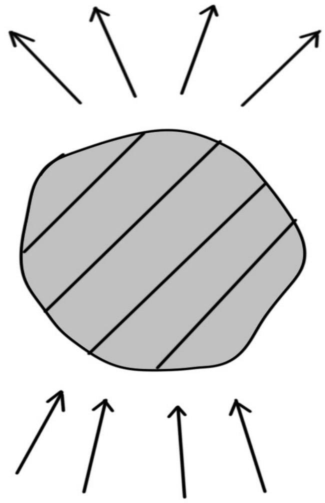

Figure 4. A situation where the spacetime is stationary in the past and future.

that becomes positive frequency in the future. These two bases need not coincide. A basis of positive frequency solutions in the past will propagate to a *positive norm* basis in the future, but not necessarily to a basis of *positive frequency* solutions in the future. Hence, the incoming Fock space vacuum may evolve to a state that is not a vacuum in the future — wiggles of the geometry at intermediate times can create particles.

In this situation, where the spacetime is asymptotically stationary (in particular, flat) in the past and future, an S matrix can be defined that relates the past and future Fock space basis, and hence gives the amplitude for an incoming particle state to "scatter" off the geometry, or for particles to be created or annihilated in the time-dependent background.

Let $|\psi\rangle$ denote a Fock space state for QFT on flat spacetime. Then we denote by $|\psi_{\rm in}\,\rangle, |\psi_{\rm out}\,\rangle$ the states of the theory on a nontrivial background that asymptotically approach $|\psi\rangle$ in the past and future respectively. We define the S-matrix by

$$S|\psi_{\rm out}\rangle = |\psi_{\rm in}\rangle.$$

S will then be a unitary operator (it preserves the inner product — if something comes in, than something goes out). Thus $|\psi_{\rm out}\,\rangle = S^{-1}\,|\psi_{\rm in}\rangle \Rightarrow \langle\psi_{\rm out}\,| = \langle\psi_{\rm in}\,|\left(S^{-1}\right)^\dagger = \langle\psi_{\rm in}\,|\,S$. Then we have

$$\begin{aligned}\langle\chi_{\rm out}\,|\psi_{\rm in}\rangle &= \langle\chi_{\rm out}\,|S|\psi_{\rm out}\,\rangle \\ &= \langle\chi_{\rm in}\,|S|\psi_{\rm in}\,\rangle\,.\end{aligned}$$

Let $\{|i\rangle\}$ denote a complete orthonormal basis for the Fock space of the theory on a trivial background (not just for $\mathcal{H}^{(1)}$, but for all of Fock space). Because time evolution is unitary, the bases $\{|i,\text{in}\rangle\}$ and $\{|i,\text{out}\rangle\}$ are also orthonormal. Then the S-matrix has the representation:

$$S = \sum_i |i,\text{in}\rangle\langle i,\text{out}|$$

(This ensures $S|i,\text{out}\rangle = |i,\text{in}\rangle$ acting on the basis). Therefore,

$$S^{-1} = S^\dagger = \sum_i |i,\text{out}\rangle\langle i,\text{in}|.$$

If $\mathcal{O}^{\text{in}}$ is any operator, then

$$\langle i,\text{in}|\mathcal{O}^{\text{in}}|j,\text{in}\rangle = \langle i,\text{out}|S^{-1}\mathcal{O}^{\text{in}}S|j,\text{out}\rangle;$$

that is,

$$\mathcal{O}^{\text{out}} = S^{-1}\mathcal{O}^{\text{in}}S$$

has the same matrix elements between out-states as $\mathcal{O}^{\text{in}}$ does between in-states.

We can calculate S by solving the (linear) KG equation on the nontrivial background. The solutions f, p solve the flat space equation in the future and past; since $\{u_i, u_i^*\}$ form a complete basis for the flat-space solutions, we have:

$$\begin{aligned} &f_i \to u_i \text{ in the future} \\ \Rightarrow &f_i \to \alpha_{ij}u_j + \beta_{ij}u_j^* \text{ in the past} \end{aligned}$$

for appropriate matrices α, β (repeated indices are summed). Similarly,

$$\begin{aligned} &p_i \to u_i \text{ in the past} \\ \Rightarrow &p_i \to \gamma_{ij}u_j + \delta_{ij}u_j^* \text{ in the future.} \end{aligned}$$

Because the KG equation is linear and homogeneous even when the background is nontrivial, we have:

$$\begin{aligned} f_i &= \alpha_{ij}p_j + \beta_{ij}p_j^*, \\ p_i &= \gamma_{ij}f_j + \delta_{ij}f_j^*. \end{aligned}$$

By expanding ϕ in terms of both bases, we find:

$$\begin{aligned} \phi &= \sum_i \left(p_i a_i^{\text{in}} + p_i^* a_i^{\text{in}\dagger}\right) \\ &= \sum_j \left(f_i a_i^{\text{out}} + f_i^* a_i^{\text{out}\dagger}\right), \end{aligned}$$

where now $a_i^{\text{in}}, a_i^{\text{in}\dagger}$ act on the Fock states in the distant past in the same way as the usual annihilation and creation operators, and $a_i^{\text{out}}, a_i^{\text{out}\dagger}$ act on the Fock states in the distance future in the same way as the usual annihilation and creation operators. We may then solve for S by demanding:

$$\begin{aligned} S^{-1}a^{\text{in}}S &= a^{\text{out}}, \\ S^{-1}a^{\text{in}\dagger}S &= a^{\text{out}\dagger}. \end{aligned}$$

3.5 Bogoliubov transformation

To see how this goes, let's consider the somewhat more general problem of relating two Fock space bases obtained by expanding the field in terms of two different orthonormal bases for the KG solutions (that are linearly related):

$$\begin{aligned} u_i' &= \alpha_{ij} u_j + \beta_{ij} u_j^*, \\ u_i'^* &= \beta_{ij}^* u_j + \alpha_{ij}^* u_j^*. \end{aligned}$$

In matrix notation:

$$\begin{pmatrix} u' \\ u'^* \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \begin{pmatrix} u \\ u^* \end{pmatrix}.$$

If both bases are normalized so that

$$\begin{aligned} (u_i, u_j) = \delta_{ij}, \quad (u_i^*, u_j) = 0, \quad \left(u_i^*, u_j^*\right) = -\delta_{ij}, \\ \left(u_i', u_j'\right) = \delta_{ij}', \quad \left(u_i'^*, u_j'\right) = 0, \quad \left(u_i'^*, u_j'^*\right) = -\delta_{ij}, \end{aligned}$$

then we have

$$\begin{aligned} \delta_{ij} = \left(u_i', u_j'\right) &= \alpha_{ik}^* \alpha_{jl} \delta_{kl} - \beta_{ik}^* \beta_{jl} \delta_{kl} \\ &= \left(\alpha \alpha^\dagger - \beta \beta^\dagger\right)_{ji}, \end{aligned}$$

Thus,

$$\boxed{\alpha \alpha^\dagger - \beta \beta^\dagger = \mathbb{1}.}$$

Also,

$$\begin{aligned} 0 = (u_i'^*, u_j) &= \beta_{ik} \alpha_{jl} \delta_{kl} - \alpha_{ik} \beta_{jl} \delta_{kl} \\ &\Rightarrow \boxed{\alpha \beta^T - \beta \alpha^T = 0.} \end{aligned}$$

From these identities, it follows that the right inverse is

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}^{-1} = \begin{pmatrix} \alpha^\dagger & -\beta^T \\ -\beta^\dagger & \alpha^T \end{pmatrix}.$$

(The inverse is unique if it exists. And this is the left inverse as well as the right inverse, as we will see from additional identities shortly.) Therefore,

$$\begin{pmatrix} u \\ u^* \end{pmatrix} = \begin{pmatrix} \alpha^\dagger & -\beta^T \\ -\beta^\dagger & \alpha^T \end{pmatrix} \begin{pmatrix} u' \\ u'^* \end{pmatrix}.$$

Compare now the two mode expansions:

$$\begin{aligned}\phi &= \sum_i \left(u_i a_i + u_i^* a_i^\dagger\right) = \left(a\ a^\dagger\right)\begin{pmatrix} u \\ u^* \end{pmatrix} \\ &= \sum_i \left(u_i' a_i' + u_i'^* a_i^\dagger\right) = \left(a'\ a'^\dagger\right)\begin{pmatrix} u' \\ u'^* \end{pmatrix}.\end{aligned}$$

Substituting for u', u'^* in terms of u, u^* and vice versa gives

$$\begin{aligned}\left(a\ a^\dagger\right)\begin{pmatrix} u \\ u^* \end{pmatrix} &= \left(a'\ a'^\dagger\right)\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}\begin{pmatrix} u \\ u^* \end{pmatrix}, \\ \left(a'^+ a'^\dagger\right)\begin{pmatrix} u' \\ u'^* \end{pmatrix} &= \left(a\ a^\dagger\right)\begin{pmatrix} \alpha^\dagger & -\beta^T \\ -\beta^\dagger & \alpha^T \end{pmatrix}\begin{pmatrix} u' \\ u'^* \end{pmatrix}.\end{aligned}$$

Therefore, since the bases are complete and orthonormal (e.g., we can isolate coefficients by taking the KG inner product), we have

$$\begin{aligned}\left(a\ a^\dagger\right) &= \left(a'\ a'^\dagger\right)\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}, \\ \left(a' a'^\dagger\right) &= \left(a\ a^\dagger\right)\begin{pmatrix} \alpha^\dagger & -\beta^T \\ -\beta^\dagger & \alpha^T \end{pmatrix},\end{aligned}$$

and taking transposes yields

$$\boxed{\begin{pmatrix} a' \\ a'^\dagger \end{pmatrix} = \begin{pmatrix} \alpha^* & -\beta^* \\ -\beta & \alpha \end{pmatrix}\begin{pmatrix} a \\ a^\dagger \end{pmatrix}, \quad \begin{pmatrix} a \\ a^\dagger \end{pmatrix} = \begin{pmatrix} \alpha^T & \beta^\dagger \\ \beta^T & \alpha^\dagger \end{pmatrix}\begin{pmatrix} a' \\ a'^\dagger \end{pmatrix}.}$$

In addition, we now see that

$$\begin{pmatrix} \alpha^T & \beta^\dagger \\ \beta^T & \alpha^\dagger \end{pmatrix} = \begin{pmatrix} \alpha^* & -\beta^* \\ -\beta & \alpha \end{pmatrix}^{-1},$$

and we therefore have the further identities:

$$\begin{aligned}\alpha^T \alpha^* - \beta^\dagger \beta &= \mathbb{1}, \\ \alpha^T \beta^* - \beta^\dagger \alpha &= 0.\end{aligned}$$

Since the u, u^*'s and u', u'^*'s are normalized, both the $a, a^\dagger$ and $a', a'^\dagger$ satisfy the standard commutation relations. The transformation $(a, a^\dagger) \to (a', a'^\dagger)$ is a *canonical transformation*, a change of variables that preserves the commutation relations. A canonical transformation such that a' and $a'^\dagger$ are linear combinations of a and $a^\dagger$ is called a Bogoliubov transformation.

For $\beta \neq 0$, this transformation does not preserve the positive-frequency subspace of the space of solutions to the KG equation. Nonetheless, it can be expressed as a unitary transformation acting on the Hilbert space — this is a general feature of canonical transformations. (Actually, for a system with an infinite number of degrees of freedom, this is true only subject to certain conditions that will emerge from the calculation below.) That is, there is a unitary transformation U with the properties

$$U^{-1}aU = a'$$
$$U^{-1}a^\dagger U = a'^\dagger$$

From these equations and the Bogoliubov transformation, we can solve for U, expressing it in terms of either a or a'. (In fact, U is given by the same expression whether written in terms of a or a'. This reflects the observation we made earlier about the S-matrix, that matrix elements of the operator $\mathcal{O}^{\rm in}$ between in-states coincide with the matrix elements of the operator $\mathcal{O}^{\rm out}$ between out-states.)

3.6 Normal ordering

It is convenient to express U as a normal-ordered function of a and $a^\dagger$:

$$U =: U(a, a^\dagger) :$$

(where the double dots denote normal ordering). Normal-ordered means that all a operators lie to the right of all $a^\dagger$ operators. To solve for U that realizes the Bogoliubov transformation,

$$aU = Ua' = U(\alpha^* a - \beta^* a^\dagger),$$
$$a^\dagger U = Ua'^\dagger = U(-\beta a + \alpha a^\dagger),$$

we note that operators $a, a^\dagger$ act on normal-ordered functions as:

$$a_i : \mathcal{O} : \; =: (a_i + \frac{\partial}{\partial a_i^\dagger})\mathcal{O} :$$
$$a_i^\dagger : \mathcal{O} : =: a_i^\dagger \mathcal{O} :$$
$$: \mathcal{O} : a_i \; =: a_i \mathcal{O} :$$
$$: \mathcal{O} : a_i^\dagger \; =: (a_i^\dagger + \frac{\partial}{\partial a_i})\mathcal{O} :$$

because

$$[a_i, f(a^\dagger)] = \frac{\partial}{\partial a_i^\dagger} f(a^\dagger),$$
$$[f(a), a_i^\dagger] = \frac{\partial}{\partial a_i} f(a).$$

So we must find a function $U(a, a^\dagger)$ satisfying:

(1) $\left(a + \frac{\partial}{\partial a^\dagger}\right) U = \left[\alpha^* a - \beta^* \left(a^\dagger + \frac{\partial}{\partial a}\right)\right] U,$

(2) $a^\dagger U = \left[-\beta a + \alpha \left(a^\dagger + \frac{\partial}{\partial a}\right)\right] U.$

(Normal ordering is now understood and not explicitly indicated.)

We solve this by means of the ansatz:

$$\begin{aligned} U &= C \exp\left[\frac{1}{2}\begin{pmatrix} a \, a^\dagger \end{pmatrix}\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}\begin{pmatrix} a \\ a^\dagger \end{pmatrix}\right] \\ &= C \exp\left[\frac{1}{2}\left(aM_{11}a + aM_{12}a^\dagger + a^\dagger M_{21}a + a^\dagger M_{22}a^\dagger\right)\right] \end{aligned}$$

where we may assume without loss of generality:

$$M_{11} = M_{11}^T, \quad M_{22} = M_{22}^T, \quad M_{12} = M_{21}^T.$$

Then:

$$\begin{aligned} \frac{\partial}{\partial a} U &= \left(M_{11}a + M_{12}a^\dagger\right) U, \\ \frac{\partial}{\partial a^\dagger} U &= \left(M_{21}a + M_{22}a^\dagger\right) U. \end{aligned}$$

and we have the algebraic equations:

(i) $\left(a + M_{21}a + M_{22}a^\dagger\right) = \alpha^* a - \beta^* \left(a^\dagger + M_{11}a + M_{12}a^\dagger\right),$

(ii) $a^\dagger = -\beta a + \alpha \left(a^\dagger + M_{11}a + M_{12}a^\dagger\right).$

Since the coefficients of a and $a^\dagger$ must match between the two sides, (ii) implies:

$$\begin{aligned} \mathbb{1} &= \alpha + \alpha M_{12} \\ 0 &= -\beta + \alpha M_{11} \end{aligned} \Rightarrow \boxed{\begin{aligned} M_{12} &= \alpha^{-1} - \mathbb{1} \\ M_{11} &= \alpha^{-1}\beta \end{aligned}}$$

How do we know that α is invertible? This follows from the identity:

$$\alpha\alpha^\dagger = \mathbb{1} + \beta\beta^\dagger$$

which shows that $\alpha\alpha^\dagger$ has no zero eigenvalue. Thus, $\langle\psi|\alpha\alpha^\dagger|\psi\rangle = \|\alpha^\dagger|\psi\rangle\|^2 > 0$, and therefore $\alpha^\dagger$ has a trivial kernel, which means that the range of α is the entire space. Furthermore, $\left(\alpha^\dagger\alpha\right)^* = \mathbb{1} + \beta^\dagger\beta$ shows that the kernel of α is trivial — so α has a left and right inverse.

From (i), we find:

$$\begin{aligned} M_{22} &= -\beta^* - \beta^* M_{12} = -\beta^* \alpha^{-1}, \\ \mathbb{1} + M_{21} &= \alpha^* - \beta^* M_{11} = \alpha^* - \beta^* \alpha^{-1}\beta, \end{aligned}$$

or

$$\boxed{M_{22} = -\beta^* \alpha^{-1}, \quad M_{21} = -\mathbb{1} + \alpha^* - \beta^* \alpha^{-1} \beta}.$$

Now, we must check that this solution is consistent with the assumptions:

$$M_{11} = M_{11}^T, \quad M_{22} = M_{22}^T, \quad M_{21} = M_{12}^T.$$

These relations actually follow from the identities:

$$\begin{array}{cc} \alpha\alpha^\dagger - \beta\beta^\dagger = \mathbb{1} & \alpha^T\alpha^* - \beta^\dagger\beta = \mathbb{1} \\ \beta\alpha^T = \alpha\beta^T & \beta^\dagger\alpha = \alpha^T\beta^* \end{array}$$

E.g.,

$$\begin{aligned} &\beta\alpha^T = \alpha\beta^T \Rightarrow M_{11} = \alpha^{-1}\beta = \beta^T(\alpha^T)^{-1} = \left(\alpha^{-1}\beta\right)^T = M_{11}^T, \\ &\beta^\dagger\alpha = \alpha^T\beta^* \Rightarrow -M_{22} = \beta^*\alpha^{-1} = \left(\alpha^T\right)^{-1}\beta^\dagger = \left(\beta^*\alpha^{-1}\right)^T = -M_{22}^T. \end{aligned}$$

And from

$$\mathbb{1} = \alpha^T\alpha^* - \beta^\dagger\beta = \alpha^T\alpha^* - \left(\alpha^T\beta^*\alpha^{-1}\right)\beta,$$

we have

$$\left(\alpha^T\right)^{-1} = \alpha^* - \beta^*\alpha^{-1}\beta,$$

so that $M_{12}^T = M_{21}$. We have now found:

$$\boxed{U = C : \exp\left[\frac{1}{2}a\left(\alpha^{-1}\beta\right)a + a\left(\alpha^{-1} - \mathbb{1}\right)a^\dagger + \frac{1}{2}a^\dagger\left(-\beta^*\alpha^{-1}\right)a^\dagger\right] :}$$

To complete the calculation of U, we must find the constant C. We may determine C up to a phase by requiring that U is unitary, so:

$$\langle 0|UU^\dagger|0\rangle = 1$$

where $|0\rangle$ is the vacuum satisfying $a_i|0\rangle = 0$. Since the adjoint of a normal-ordered operator is also normal-ordered, we have:

$$U^\dagger = C^* : \exp\left[\frac{1}{2}a^\dagger\left(\alpha^{-1}\beta\right)^* a^\dagger + \cdots\right] :$$

and hence $U^\dagger|0\rangle = C^* \exp\left[\frac{1}{2}a^\dagger\left(\alpha^{-1}\beta\right)^* a^\dagger\right]|0\rangle$.

Now, we invoke the identity:

$$\left\langle 0\middle|\exp\left(\frac{1}{2}aMa\right)\exp\left(\frac{1}{2}a^\dagger M^* a^\dagger\right)\middle|0\right\rangle$$
$$= \left[\det\left(\mathbb{1} - MM^*\right)\right]^{-\frac{1}{2}}, \quad \text{where } M = M^T.$$

(We will return to the derivation of this identity shortly.) Thus:

$$1 = |C|^2\left(\det\left[\mathbb{1} - \alpha^{-1}\beta\left(\alpha^{-1}\beta\right)^*\right]\right)^{-\frac{1}{2}}.$$

We can simplify the determinant by recalling:

$$\alpha^\dagger\alpha - \beta^T\beta^* = \mathbb{1}, \quad \alpha^\dagger\beta = \beta^T\alpha^*,$$
$$\Rightarrow \mathbb{1} = \alpha^\dagger\alpha - \alpha^\dagger\beta\left(\alpha^*\right)^{-1}\beta^*$$
$$\Rightarrow \left(\alpha^\dagger\alpha\right)^{-1} = \mathbb{1} - \alpha^{-1}\beta\left(\alpha^{-1}\beta\right)^*.$$

So we have $|C|^2 = (\det(\alpha^\dagger\alpha))^{-\frac{1}{2}}$, and we have determined C up to a phase:

$$|C| = (\det\alpha^\dagger\alpha)^{-\frac{1}{4}}.$$

We will not attempt to determine the phase of C. This phase is of little relevance. Moreover, it is subject to ambiguities concerning how the renormalized energy-momentum $T^{\mu\nu}$ is defined, as we will discuss later.

3.7 Computation of the S-matrix

In the case of an asymptotically flat spacetime, we have found the S-matrix:

$$\boxed{S = (\text{phase})(\det\alpha^\dagger\alpha)^{-\frac{1}{4}} : \exp\left[\frac{1}{2}a^{\text{in}}(\alpha^{-1}\beta)a^{\text{in}} + a^{\text{in}}(\alpha^{-1} - \mathbb{1})a^{\text{in}\dagger} + \frac{1}{2}a^{\text{in}\dagger}(-\beta^*\alpha^{-1})a^{\text{in}\dagger}\right]:}$$

It has the same form when expressed in terms of a^{out}, $a^{\text{out}\dagger}$, since:

$$\langle\psi_{\text{in}}|S|\chi_{\text{in}}\rangle = \langle\psi_{\text{out}}|S|\chi_{\text{out}}\rangle\rangle.$$

In particular, we conclude that:

$$|\langle 0_{\text{out}}|0_{\text{in}}\rangle|^2 = \langle 0_{\text{in}}|S|0_{\text{in}}\rangle|^2 = \left(\det\alpha^\dagger\alpha\right)^{-\frac{1}{2}}.$$

Note that, for this construction of a unitary S-matrix to make sense, we must have $(\alpha^\dagger\alpha) < \infty$, so $|0_{\text{ out}}\rangle$ must have a nonzero overlap with $|0_{\text{ in}}\rangle$. We can state this criterion in a somewhat more physical way by noting that:

$$
\begin{aligned}
N_i &= \langle 0_{\text{in}}| \underbrace{a_i^{\text{out}\dagger} a_i^{\text{out}}}_{\text{not summed}} |0_{\text{in}}\rangle \\
&= \langle 0_{\text{in}}| \left(-\beta_{ij} a_j^{\text{in}}\right) \left(-\beta_{ik}^* a_k^{\text{in}\dagger}\right) |0_{\text{in}}\rangle \\
&= \left(\beta\beta^\dagger\right)_{ii} = \sum_j |\beta_{ij}|^2 \,. \qquad (3.1)
\end{aligned}
$$

This is the expectation value of the number of particles of type i produced by the time-varying geometry, when the incoming state is the vacuum. The sum over i gives the total particle number, $N = \sum_i N_i = \text{Tr}(\beta\beta^\dagger)$. And since $\alpha\alpha^\dagger = \mathbb{1} + \beta\beta^\dagger$, the condition that $\det(\alpha\alpha^\dagger) = \det\left(\mathbb{1} + \beta\beta^\dagger\right) < \infty$ is precisely the same as $N < \infty$. The canonical transformation $\left(a^{\text{in}}, a^{\text{in}\dagger}\right) \to \left(a^{\text{out}}, a^{\text{out}\dagger}\right)$ *cannot* be implemented by a unitary operator if the $|0\rangle_{\text{in}}$ vacuum is a state with an indefinite number of particles when expanded in terms of the Fock basis in the distant future.

(This may happen in principle if there are massless particles and many soft particles are produced by the fluctuating geometry. It does not happen if the (smooth) spacetime is flat outside a compact region, because fluctuating geometry in a compact region does not produce arbitrarily soft particles. For further discussion, see Wald [26, 27].)

Some other S-matrix elements are (see Figure 5):

$$
\langle i_{\text{out}}|j_{\text{in}}\rangle = \langle 0_{\text{out}}|0_{\text{in}}\rangle \left(\mathbb{1} + \underbrace{\alpha^{-1} - \mathbb{1}}_{\text{“connected part”}}\right)_{ij} = \alpha_{ij}^{-1}.
$$

Thus an incoming particle may scatter and emerge with its momentum changed (if $\alpha \neq \mathbb{1}$). We also have

$$
\langle ij_{\text{out}}|0_{\text{in}}\rangle = \langle 0_{\text{out}}|0_{\text{in}}\rangle \left(-\beta^*\alpha^{-1}\right)_{ij} ,
$$

$$
\langle 0_{\text{out}}|ij_{\text{in}}\rangle = \langle 0_{\text{out}}|0_{\text{in}}\rangle \left(\alpha^{-1}\beta\right)_{ij} .
$$

We conclude that particles are created or annihilated only if $\beta \neq 0$ (in which case the Bogoliubov transformation mixes positive and negative frequencies). If $\beta = 0$, then α is unitary, and $\langle 0_{\text{out}}|0_{\text{in}}\rangle$ is a phase.

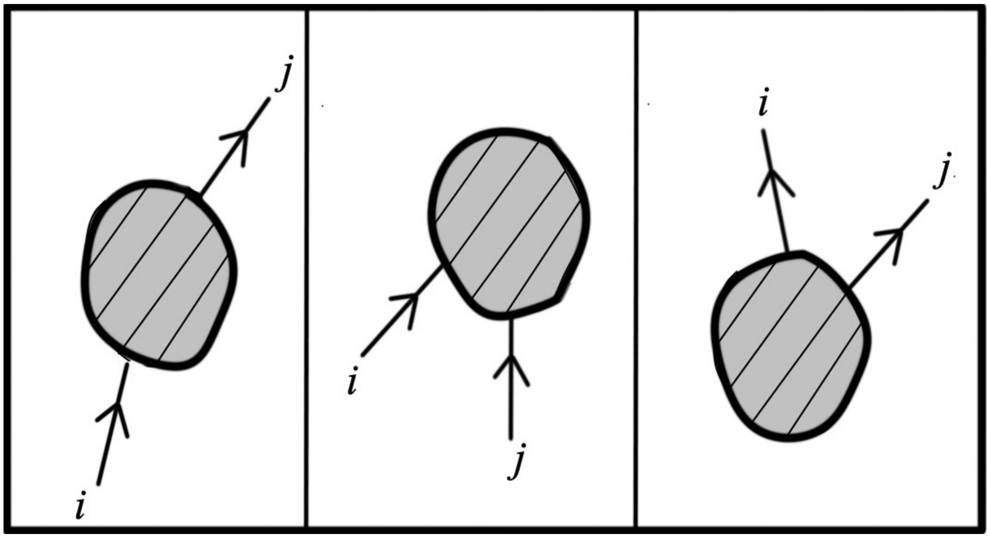

Figure 5. Left: Particle scattering off geometry. Center: Particle annihilation by geometry. Right: Particle creation by geometry.

Particles are not created or annihilated singly, but only in pairs. (These could be particle-antiparticle pairs in the case of a complex scalar field.). It is also easy to extract from our expression for S the many-particle $\to$ many-particle amplitudes. Consider, for example,

$$\langle i_1 \ldots i_{2n\ \text{out}} | 0_{\text{in}} \rangle = \langle i_1 \ldots i_{2n} | S | 0 \rangle = \langle 0_{\text{out}} | 0_{\text{in}} \rangle \left\langle i_1 \ldots i_{2n} \left| \frac{1}{n!} \left(\frac{1}{2} a^\dagger V a^\dagger \right)^n \right| 0 \right\rangle$$

where $V = -\beta^* \alpha^{-1}$. Allowing the $2n$ $a^\dagger$ operators to annihilate (to the left) the $2n$ particles $i_1 \ldots i_{2n}$ generates $(2n)!$ terms, each term giving a product $V.V \ldots V$ of n matrix elements of the symmetric matrix V. This product depends only on how the $2n$ indices are paired, and each pairing occurs $n!2^n$ times; this is the number of the $(2n)!$ permutations of indices that leave the pairing unchanged. Thus, the $\frac{1}{n!2^n}$ gets cancelled, and we have:

$$\frac{\langle i_1 \ldots i_{2n\ \text{out}}\ | 0_{\text{in}} \rangle}{\langle 0_{\text{out}} | 0_{\text{in}} \rangle} = V_{i_1,i_2} V_{i_3,i_4} \ldots V_{i_{2n-1,1} i_n} + \text{ all other pairings,}$$

there being $(2n)!/(n!2^n)$ terms in the sum. E.g.,

$$\langle i_1 \ldots i_{4\ \text{out}}\ | 0_{\text{in}} \rangle = (V_{i_1,i_2} V_{i_3,i_4} + V_{i_1,i_3} V_{i_2,i_4} + V_{i_1,i_4} V_{i_2,i_3}) \langle 0_{\text{out}} | 0_{\text{in}} \rangle.$$

3.8 Proof of the identity using Gaussian integrals

We return now to the derivation of the identity:

$$\left\langle 0 \left| \exp\left(\frac{1}{2} a M a \right) \exp\left(\frac{1}{2} a^\dagger M^* a^\dagger \right) \right| 0 \right\rangle = \left[\det\left(\mathbb{1} - M M^* \right) \right]^{-\frac{1}{2}} \text{ (where } M = M^T).$$

The slick way to evaluate such matrix elements is to convert them to Gaussian integrals.

An arbitrary Fock space state can be represented as a function of $a^\dagger$ acting on $|0\rangle$:

$$|\psi\rangle = \psi(a^\dagger)|0\rangle = \sum_{n=0}^{\infty} \sum_{i_1 \ldots i_n} \psi^{(n)}_{i_1 \ldots i_n} a^\dagger_{i_1} \ldots a^\dagger_{i_n} |0\rangle,$$

where $\psi^{(n)}_{i_1 \ldots i_n}$ is symmetric under the exchange of indices.

We claim:

$$\langle \chi | \psi \rangle = \frac{\int da da^\dagger \chi^*(a) \psi\left(a^\dagger\right) e^{-a^\dagger a}}{\int da da^\dagger e^{-a^\dagger a}}.$$

Here, on the RHS, $a_i, a^\dagger_i$ are complex c-numbers, and

$$\int da da^\dagger = \int \Pi_i da_i da^\dagger_i$$

is an integral $\int_{-\infty}^{\infty}$ over real and imaginary parts of each a_i.

To verify the claim, it is enough to show (since we can expand in powers of $a^\dagger$):

$$\left\langle 0 \left| a_{i_1} \ldots a_{i_n} a^\dagger_{j_1} \ldots a^\dagger_{j_m} \right| 0 \right\rangle = \frac{\int da da^\dagger a_{i_1} \ldots a_{i_n} a^\dagger_{j_1} \ldots a^\dagger_{j_m} e^{-a^\dagger a}}{\int da da^\dagger e^{-a^\dagger a}} \quad .$$

We can derive this identity by evaluating both sides. Commuting a's through $a^\dagger$'s, we find:

$$\left\langle 0 \middle| a_{i_1} \ldots a_{i_n} a^\dagger_{j_1} \ldots a^\dagger_{j_m} \middle| 0 \right\rangle = \begin{cases} 0 & \text{if } m \neq n, \\ \delta_{i_1, j_1} \ldots \delta_{i_n j_n} + \text{permutations } (n! \text{ terms altogether}) & \text{if } m = n. \end{cases}$$

To evaluate the integral on the LHS, construct the generating function

$$Z[J, \overline{J}] = \int da da^\dagger e^{\overline{J} a} e^{a^\dagger J} e^{-a^\dagger a}.$$

We complete the square

$$Z[J, \overline{J}] = \int da da^\dagger \exp\left[- \left(a^\dagger - \overline{J} \right) (a - J) + \overline{J} J \right]$$

and shift the integral:

$$= e^{\overline{J} J} Z[0, 0].$$

Now

$$\begin{aligned}
&\frac{\int dada^\dagger a_{i_1}\ldots a_{i_n}a^\dagger_{j_1}\ldots a^\dagger_{j_m}e^{-a^\dagger a}}{\int dada^\dagger e^{-a^\dagger a}}\\
=&\frac{1}{Z[0,0]}\frac{\partial}{\partial \overline{J}_{i_1}}\cdots\frac{\partial}{\partial \overline{J}_{i_n}}\frac{\partial}{\partial J_{j_1}}\cdots\frac{\partial}{\partial J_{j_m}}Z[J,\overline{J}]\bigg|_{J=\overline{J}=0}\\
=&\frac{\partial}{\partial \overline{J}_{i_1}}\cdots\frac{\partial}{\partial J_{j_m}}e^{\overline{J}J}\bigg|_{J=\overline{J}=0}
\end{aligned}$$

This evidently vanishes for $n \neq m$. For $n = m$, we have:

$$\begin{aligned}
&=\frac{\partial}{\partial J_{j_1}}\cdots\frac{\partial}{\partial J_{j_n}}\left(J_{i_1}\ldots J_{i_n}\right)\\
&=\delta_{i_1 j_1}\ldots\delta_{i_n j_n}+\text{permutations.}
\end{aligned}$$

By means of this trick, we have:

$$\left\langle 0\left|\exp\left(\frac{1}{2}aMa\right)\exp\left(\frac{1}{2}a^\dagger M^* a^\dagger\right)\right|0\right\rangle = \frac{\int dada^\dagger \exp\left[-\frac{1}{2}\left(a\,a^\dagger\right)\begin{pmatrix}-M & \mathbb{1}\\ \mathbb{1} & -M^*\end{pmatrix}\begin{pmatrix}a\\ a^\dagger\end{pmatrix}\right]}{\int dada^\dagger \exp\left(a^\dagger a\right)} \quad .$$

You can evaluate such an integral by changing variables to real $X = \frac{1}{\sqrt{2}}(a + a^\dagger)$ and $Y = \frac{1}{i\sqrt{2}}(a - a^\dagger)$. A real Gaussian integral

$$\int dX e^{-\frac{1}{2}XAX} \qquad \left(A = A^T\right)$$

can be evaluated if A has an orthonormal basis of eigenvectors:

$$Ae_n = \lambda_n e_n.$$

By expanding $X = \sum_n X_n e_n$ in this basis, we find:

$$\begin{aligned}
\frac{\int dXe^{-\frac{1}{2}X\cdot AX}}{\int dXe^{-\frac{1}{2}X\cdot X}} &= \Pi_n\left[\frac{\int dX_n e^{-\frac{1}{2}X_n\lambda_n X_n}}{\int dX_n e^{-\frac{1}{2}X_n^2}}\right]\\
&= \Pi_n\lambda_n^{-\frac{1}{2}} = (\det A)^{-\frac{1}{2}}.
\end{aligned}$$

(The integrals converge for Re $\lambda_n > 0$, and may be defined by analytic continuation otherwise.)

Writing the integral in terms of real variables and using elementary properties of determinants, we find (Exercise 6.4):

$$\left\langle 0 \left| \exp\left(\frac{1}{2} a M a\right) \exp\left(\frac{1}{2} a^\dagger M^* a^\dagger\right) \right| 0 \right\rangle = \left[\det \begin{pmatrix} \mathbb{1} & -M^* \\ -M & \mathbb{1} \end{pmatrix} \right]^{-\frac{1}{2}}$$

and

$$\begin{aligned} \det \begin{pmatrix} \mathbb{1} & -M^* \\ -M & \mathbb{1} \end{pmatrix} &= \det \begin{pmatrix} \mathbb{1} & 0 \\ M & \mathbb{1} \end{pmatrix} \det \begin{pmatrix} \mathbb{1} & -M^* \\ -M & \mathbb{1} \end{pmatrix} \\ &= \det \begin{pmatrix} \mathbb{1} & -M^* \\ 0 & \mathbb{1} - M M^* \end{pmatrix} = \det\left(\mathbb{1} - M M^*\right) \end{aligned},$$

which was to be shown.

4 Quantum Field Theory in Rindler Space

Our next topic is: "How does the Minkowski space vacuum look to a uniformly accelerated observer?" After all the fuss about curved spacetime, it may seem disappointing to return to flat spacetime, but there is ample motivation.

- Quantum field fluctuations that have positive frequency according to a clock of an inertial observer can have both positive and negative frequency as measured by a non-inertial observer. The uniformly accelerated observer provides us with a nontrivial and calculable example of a Bogoliubov transformation.
- The principle of equivalence asserts that physics as seen by a uniformly accelerated observer is the same as physics in a uniform gravitational field, so we will be investigating a gravitational effect on quantum fields.
- This example provides a first glimpse of a profound connection among gravitation, quantum mechanics, and thermodynamics (for the accelerated observer sees a thermal radiation bath). An even deeper connection of this kind will emerge when we study black holes.
- Invoking analogies with the simpler Rindler spacetime can guide our understanding of more subtle spacetimes (e.g., Schwarzschild or de Sitter).

4.1 Vacuum fluctuations in an inertial frame

Before considering the effect of quantum field fluctuations on an accelerated observer, let's first consider some features of field fluctuations in an inertial frame. We will continue to consider a theory of a real free scalar field, and to keep things as simple as possible, we'll suppose that the field is massless $\big(m^2 = 0\big)$.

As an exercise (Exercise 6.2), you may evaluate

$$G_+(x) = \int \frac{d^3k}{(2\pi)^3 2k^0} e^{-ik\cdot x} = \langle 0|\phi(x)\phi(0)|0\rangle$$

where $m^2 = 0 \Rightarrow k^0 = |\vec{k}|$, finding

$$G_+(x) = \frac{1}{8\pi^2 r}\left(\frac{1}{r-(t-i\epsilon)} + \frac{1}{r+(t-i\epsilon)}\right)$$

(where $t = x^0, r = |\vec{x}|$)

$$= \frac{-1}{4\pi^2\left(x^2 - 2i\epsilon x^0\right)}.$$

Here we have invoked the observation that $G_+(x)$ is analytic for $\operatorname{Im} t < 0$, and so have defined the integral by giving a small imaginary part $-i\epsilon$ to t. Using

$$\lim_{\epsilon\to 0}\left(\frac{1}{a+i\epsilon} - \frac{1}{a-i\epsilon}\right) = -2\pi i\delta(a),$$

you can also evaluate

$$iG(x) = [\phi(x), \phi(0)] = G_+(x) - G_+(-x)$$
$$= \frac{-i}{4\pi r}[\delta(r-t) - \delta(r+t)].$$

The field commutator, in the $m^2 = 0$ case, has support only on the light cone $t = \pm r$. This makes sense, as massless particles propagate information at *only* the speed of light $c = 1$.

The field commutation relations reflect limitations on how accurately the field can be measured. We may attempt to measure (smeared) fields in the vicinity of x by observing the response of a test charge that couples to $\phi(x)$ (this actually measures $\vec{\nabla}\phi(x)$). We obtain a precise measurement of the field gradient by precisely measuring the *impulse* received by the charge, but because we measure the charge's momentum, the position of the charge undergoes uncontrolled quantum fluctuations. Hence, the charge accelerates, and an accelerated charge *radiates*. The field commutator is just a difference of retarded and advanced Green functions, describing the $1/r$ decay of the ϕ wave front propagating along the lightcone, which interferes with a subsequent measurement of $\phi(x)$ there.

The function

$$G_+(x) = \langle 0|\phi(x)\phi(0)|0\rangle = \frac{1}{4\pi^2\left(x^2 - i\epsilon x^0\right)}$$

blows up as $x^2 \to 0$, as does the expectation value $\langle\phi(x)\phi(0)\rangle$ in any Fock state. This divergence cautions us again that $\phi(x)$ is not itself an observable; observables are obtained by *smearing* fields — i.e., integrating them against smooth functions. $\phi(x)$ cannot be measured because, although its mean value is finite, its variance is infinite.

If we measure the field smeared with a test function f_L whose support has a linear size of order L, then

$$\langle 0|\phi(f_L)^2|0\rangle \sim \frac{1}{L^2}$$

So the typical value of $\phi(f_L)$ is $\sim 1/L$. Correspondingly, the typical value of $|\nabla\phi(f_L)|$ is $1/L^2$ and the typical value of the energy due to field fluctuations in a smearing volume L^3 is $E_L = \int d^3x \frac{1}{2}(\nabla\phi)^2 \sim \frac{1}{L}$.

These fluctuations of the fields in the vacuum are not surprising — the position of a harmonic oscillator also fluctuates in the ground state (zero-point oscillation), and the (free) field is just a system of many uncoupled oscillators.

The position of a one-dimensional harmonic oscillator is

$$x(t) = \frac{1}{\sqrt{2\omega}}\left(e^{-i\omega t}a + e^{i\omega t}a^\dagger\right),$$

and therefore

$$\langle 0|x(t)x|0\rangle = \frac{1}{2\omega}e^{-i\omega t}.$$

Similarly, if we Fourier analyze the field,

$$\tilde{\phi}(t,\vec{k}) = \int d^3x e^{-ik\cdot x}\phi(t,\vec{x})$$
$$= \frac{1}{2k^0}\left[e^{-ik^0t}A\left(k^0,\vec{k}\right) + e^{ik^0t}A\left(k^0,-\vec{k}\right)^\dagger\right],$$

then

$$\left\langle 0\left|\tilde{\phi}(t,\vec{k})\tilde{\phi}\left(0,\vec{k}'\right)\right|0\right\rangle = \frac{1}{2k^0}e^{-ik^0t}(2\pi)^3\delta^3\left(\vec{k}+\vec{k}'\right).$$

Because of translation invariance, the fluctuations are diagonal in the $\vec{k}$ basis; each mode of the field fluctuates independently (no correlations between different momentum modes).

Now consider the vacuum fluctuations in a region of linear size L at a fixed time $t = 0$. Suppose we smear the field with a smooth real-valued function $f_L(\vec{x})$ supported on the region:

$$\phi(f_L) = \int d^3x f_L(\vec{x})\phi(0,\vec{x}) = \int \frac{d^3k}{(2\pi)^3}\tilde{f}_L(-\vec{k})\tilde{\phi}(0,\vec{k}).$$

The fluctuations are characterized by

$$\langle 0|\phi(f)^2|0\rangle = \int \frac{d^3k}{(2\pi)^3}\tilde{f}_L(-\vec{k})\int \frac{d^3k'}{(2\pi)^3}\tilde{f}_L(-\vec{k}')\langle 0|\tilde{\phi}(0,\vec{k})\tilde{\phi}(0,\vec{k}')|0\rangle = \int \frac{d^3k}{(2\pi)^3 2k^0}|f_L(\vec{k})|^2.$$

The $\vec{k}$ integral is dominated by wave number $k \sim 2\pi/L$, yielding

$$\langle 0|\phi(f)^2|0\rangle \approx \frac{1}{L^2}.$$

The field admits a "classical" description only if the fluctuations in ϕ are small compared to the mean value of ϕ.

Classical fields and field quanta are "complementary" concepts. Measurement of the field with high spatial resolution typically produces many quanta. And for the spatially averaged field to behave classically, the number of field quanta per (wavelength)^3 must be large. (Vacuum fluctuations cause $(\Delta E)_L \sim \frac{1}{L}$, and E(wavelength L) $\sim \frac{n}{L}$, so $\Delta E/E \sim \frac{1}{n}$ if there are n quanta.)

4.2 Response of an inertial detector

To describe measurements in more detail, consider a "particle detector" that can absorb and emit field quanta. Idealize the detector as pointlike and suppose it travels along a worldline $x^\mu(\tau)$, parameterized by τ, the proper time along its world line. The Hamiltonian for the coupled detector-field system is

$$H = (H_0)_{\text{detector}} + (H_0)_{\text{field}} + \lambda M\phi(x(\tau)).$$

This is the Hamiltonian in the instantaneous reference frame of the detector. This Hamiltonian acts on a Hilbert space

$$\mathcal{H} = \mathcal{H}_{\text{detector}} \otimes \mathcal{H}_{\text{field}}.$$

M is a perturbation of $(H_0)_{\text{detector}}$ that is capable of causing a transition in the detector, and λ is a coupling constant, assumed small so that we can use perturbation theory.

The unperturbed detector has various energy eigenstates, and the coupling of the detector to the field allows the detector to become excited by absorbing a field quantum or to de-excite by emitting a quantum. To lowest order in λ, the amplitude for a transition is given by

$$A = \langle \text{final}| - i\lambda \int_{-\infty}^{\infty} d\tau M_I(\tau)\phi(x(\tau))\, |\text{initial}\rangle,$$

where $M_I(\tau) = e^{i(H_0)_{\text{det}}\tau} M^{-i(H_0)_{\text{det}}\tau}$ is the detector perturbation in the interaction picture. (It is the detector's proper time that appears here because M is the perturbation of $(H_0)_{\text{detector}}$ in the rest frame of the detector. We assume that the detector is sufficiently robust that it is not affected by the noninertial motion.)

If the initial state of the detector is an energy eigenstate with energy E, and its final state is an energy eigenstate with energy $E + \omega$, we have

$$A = (-i\lambda)\langle E + \omega|M|E\rangle \int_{-\infty}^{\infty} d\tau e^{i\omega\tau} \langle \text{final field}|\phi(x(\tau))|\text{initial field}\rangle.$$

Now suppose we observe only the change in the state of the detector. The probability for this change is given by squaring the amplitude and summing over the final states of the field:

$$\begin{aligned} &\text{Prob}(E \to E + \omega) \\ &= \lambda^2 |\langle E + \omega|M|E\rangle|^2 \int_{-\infty}^{\infty} d\tau e^{i\omega\tau} \int_{-\infty}^{\infty} d\tau' e^{-i\omega\tau'} \sum_{\text{fin}} \langle \text{init}|\phi(x(\tau'))|\text{fin}\rangle\langle \text{fin}|\phi(x(\tau))|\text{init}\rangle \\ &\Rightarrow \text{Prob}(E \to E + \omega) = \lambda^2 \left|\langle M\rangle\right|^2 \int d\tau d\tau' e^{-i\omega(\tau'-\tau)} \langle \text{init}|\phi\left(x(\tau')\phi(x(\tau))\right)|\text{init}\rangle. \end{aligned}$$

If $|\text{init}\rangle$ is a momentum eigenstate with 4-momentum P, then $U(a)|\text{init}\rangle = e^{iP\cdot a}|\text{init}\rangle$, where a is a spacetime translation. Recalling that $U(a)\phi(x)U(a)^{-1} = \phi(x + a)$, we can insert a translation to see that

$$\langle \text{init}|\phi\left(x(\tau')\phi(x(\tau))\right)|\text{init}\rangle = \left\langle \text{init}|\phi\left(x\left(\tau'\right) - x(\tau)\right)\phi(0)|\text{init}\right\rangle.$$

Suppose now that the detector moves *inertially*. Then $x^\mu\left(\tau'\right) - x^\mu(\tau) = u^\mu(\tau' - \tau)$, where u^μ is the detector's 4-velocity. Under this assumption, one of the τ integrals is trivial and

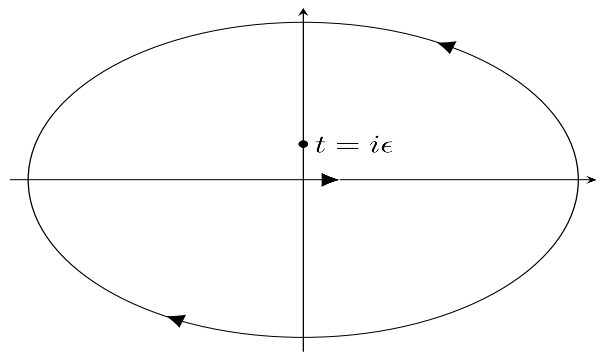

Figure 6. Integration contour for evaluating the particle detection rate. For negative frequency ω we complete the contour in the upper half plane, enclosing the pole.

gives a factor of elapsed proper time. Hence we obtain a rate (probability per unit proper time) at which the detector becomes excited:

$$\text{Rate} = \frac{\text{Prob}(E \to E+\omega)}{\text{Proper time}} = \lambda^2 \left|\langle M \rangle\right|^2 \int d\tau e^{-i\omega\tau} \langle \text{init}|\phi(x(\tau))\phi(0)|\text{init}\rangle.$$

And if the field is initially in its vacuum state, we have

$$\text{Rate} = \lambda^2 \left|\langle M \rangle\right|^2 \int d\tau e^{-i\omega\tau} G_+(x(\tau)).$$

This is most easily evaluated in the rest frame of the detector. Recalling

$$G_+(x) = \frac{-1}{4\pi^2\left[(t-i\epsilon)^2 - \vec{x}^2\right]},$$

we readily evaluate the Fourier integral by contour integration. For $\omega < 0$, we close the contour in the upper half-plane (UHP) (enclosing the pole), and for $\omega > 0$, we close it in the lower half-plane (LHP) (as shown in Figure 6):

$$\int dt e^{-i\omega t} \frac{-1}{4\pi^2 (t-i\epsilon)^2} = \begin{cases} 0 & \text{if } \omega > 0, \\ \frac{-1}{4\pi^2}(2\pi i)\frac{d}{dt} e^{-i\omega t}\Big|_{t=0} = -\frac{\omega}{2\pi} & \text{if } \omega < 0 \end{cases}$$
$$= \theta(-\omega)\frac{-\omega}{2\pi} = \theta(-\omega)\frac{|\omega|}{2\pi}.$$

We find then, that the rate vanishes for $\omega > 0$. That makes sense — a detector moving inertially in the vacuum state should not get excited. But the detector can deexcite by emitting a quantum:

$$\text{Rate} = \lambda^2 |\langle M \rangle|^2 \frac{|\omega|}{2\pi} \text{ for } \omega < 0.$$

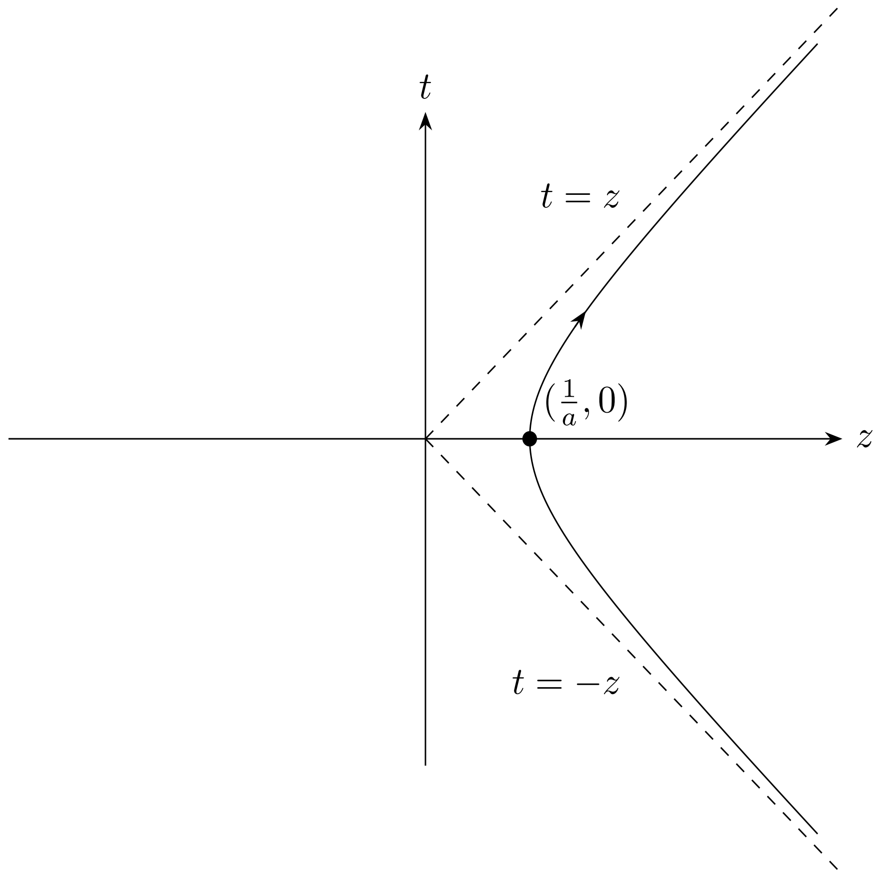

Figure 7. Trajectory of an accelerated observer

This rate, computed in lowest-order perturbation theory, agrees with Fermi's golden rule

$$\text{Rate} = 2\pi\, |\langle \text{perturbation} \rangle|^2 \times \Big(\text{density of final states per unit energy} \Big),$$

since the relativistic density of states is

$$\frac{d^3k}{(2\pi)^3 2k^0} = \frac{4\pi k^2 dk}{(2\pi)^3 2k^0} = \frac{\omega d\omega}{4\pi^2} \quad (m^2 = 0 \Rightarrow \omega = |\vec{k}|).$$

4.3 Response of an accelerated detector

We turn now to the case of a *uniformly accelerated* detector, i.e., a detector moving with constant proper acceleration as measured in its instantaneous inertial frame.

An observer accelerating uniformly in the $+\hat{z}$ direction moves on a world line

$$z^2 - t^2 = \frac{1}{a^2}$$

— a hyperbola in the z - t plane, where a is the proper acceleration (such motion is hence sometimes called "hyperbolic"). This is shown in Figure 7. Parametrized by proper time τ, the world line is

$$z(\tau) = \frac{1}{a}\cosh(a\tau),$$
$$t(\tau) = \frac{1}{a}\sinh(a\tau).$$

From

$$dz = (\sinh a\tau)d\tau,$$
$$dt = (\cosh a\tau)d\tau,$$

we verify $d\tau^2 = dt^2 - dz^2$. So τ is indeed the proper time.

The instantaneous velocity is

$$v(\tau) = \frac{dz}{dt} = \tanh(a\tau).$$

Recall that relativistic velocities add like tanh's:

$$\left.\begin{matrix} v_1 = \tanh\theta_1 \\ v_2 = \tanh\theta_2 \end{matrix}\right\} \Rightarrow v = \tanh\left(\theta_1 + \theta_2\right).$$

So as $\tau \to \tau + d\tau$, the velocity as measured in the rest frame of the moving object changes by

$$dv = \tanh(ad\tau) = ad\tau.$$

This shows that a is indeed the proper acceleration. (All inertial observers who watch the accelerating object agree on the value of this proper acceleration.)

Note that the world line, the hyperbola, is preserved by a Lorentz boost along the $\hat{z}$ direction. In fact, the effect of a boost is

$$\begin{pmatrix} t \\ z \end{pmatrix} \to \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} \begin{pmatrix} \frac{1}{a}\sinh a\tau \\ \frac{1}{a}\cosh a\tau \end{pmatrix} = \begin{pmatrix} \frac{1}{a}\sinh(a\tau+\theta) \\ \frac{1}{a}\cosh(a\tau+\theta) \end{pmatrix},$$

which is equivalent to a shift in τ by a constant. In other words, the tangent to the world line is just a Lorentz boost generator:

$$\frac{1}{a}\frac{\partial}{\partial\tau} = \text{ boost generator.}$$

Return now to the formula

$$\text{Prob}(E \to E + \omega) = \lambda^2 \left|\langle M\rangle\right|^2 \int d\tau d\tau' e^{-i\omega(\tau'-\tau)} \langle 0|\phi\left(x(\tau')\right)\phi\left(x(\tau)\right)|0\rangle$$

(where the field is initially in the vacuum state). For a uniformly accelerated observer

$$x^\mu(\tau) = \Lambda(\theta = a\tau)x^\mu(\tau = 0),$$

where $\Lambda(\theta = a\tau)$ denotes a boost along $\hat{z}$. Since ϕ is a scalar field, $\phi(\Lambda x) = U(\Lambda)\phi(x)U\left(\Lambda^{-1}\right)$, and because the vacuum is Lorentz invariant, $U(\Lambda)|0\rangle = |0\rangle$, we have

$$\begin{aligned}
&\left\langle 0\left|\phi(x(\tau'))\phi(x(\tau))\right|0\right\rangle \\
&= \langle 0|\phi(x(0))U\left(\Lambda\left(\theta'\right)\right)^{-1}U(\Lambda(\theta))\phi(x(0))|0\rangle \\
&= \langle 0|\phi(x(0))U\left(\Lambda\left(\theta-\theta'\right)\right)\phi'(x(0))|0\rangle \quad \left(\text{using } \Lambda\left(\theta'\right)^{-1}\Lambda(\theta) = \Lambda\left(\theta-\theta'\right)\right).
\end{aligned}$$

As for the case of an inertial observer, this matrix element is a function of $\tau' - \tau$, and we obtain a constant rate per unit proper time:

$$\begin{aligned}
\text{Rate}(E \to E+\omega) &= \lambda^2|\langle M\rangle|^2 \int d\tau e^{-i\omega\tau}\langle 0|\phi(x(\tau))\phi(x(0))|0\rangle \\
&= \lambda^2|\langle M\rangle|^2 \int d\tau e^{-i\omega\tau} G_+(x(\tau)-x(0)).
\end{aligned}$$

But now we calculate

$$G_+(x) = \frac{-1}{4\pi^2\left[(t-i\epsilon)^2 - \vec{x}^2\right]}$$

along the hyperbola

$$x(\tau) - x(0) = \left[\frac{1}{a}\sinh(a\tau), \frac{1}{a}\left(\cosh(a\tau)-1\right)\right],$$

and so

$$\begin{aligned}
(t-i\epsilon)^2 - z^2 &= \frac{1}{a^2}\left[\sinh^2(a\tau) - \cosh^2(a\tau) - 1 + 2\cosh(a\tau) - i\epsilon\sinh a\tau\right] \\
&= \frac{4}{a^2}\sinh^2\left(\frac{a\tau}{2}\right) - \frac{i\epsilon}{a^2}\sinh(a\tau) \simeq \frac{4}{a^2}\sinh^2\left(\frac{a\tau}{2} - i\epsilon\right)
\end{aligned}$$

or

$$\text{Rate} = \lambda^2|\langle M\rangle|^2 \int_{-\infty}^{\infty} d\tau e^{-i\omega\tau}\frac{-a^2}{16\pi^2\sinh^2\left(\frac{a\tau}{2} - i\epsilon\right)}.$$

To evaluate

$$\pi(\omega) = \int_{-\infty}^{\infty} d\tau e^{-i\omega\tau}\frac{-a^2}{16\pi^2\sinh^2\left(\frac{a\tau}{2} - i\epsilon\right)},$$

consider the contour shown in Fig 8.

Since $\sinh^2(x)$ is periodic with period $i\pi$ and the $\sinh^{-2}$ prevents the contour at $\pm\infty$ from contributing, we find

$$\begin{aligned}
\pi(\omega)(1 - e^{2\pi\omega/a}) &= 2\pi i \times (\text{ Residue at } \tau = 0) \\
&= 2\pi i\left(\frac{-a^2}{16\pi^2}\right)\left(\frac{d}{d\tau}e^{-i\omega\tau}\Big|_{\tau=0}\right)\frac{4}{a^2} \\
&= -\frac{\omega}{2\pi}.
\end{aligned}$$

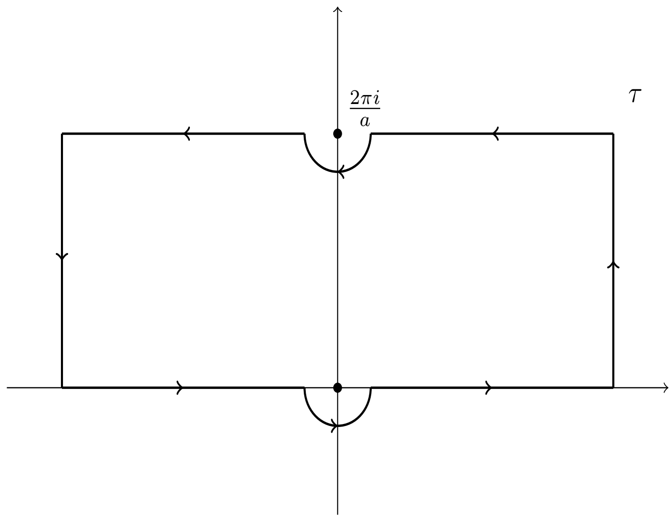

Figure 8. Contour of integration for evaluating particle detection rate for the accelerated observer.

Hence

$$\pi(\omega) = \frac{\omega/2\pi}{e^{2\pi\omega/a} - 1}$$

or

$$\boxed{\pi(\omega) = \frac{|\omega|}{2\pi} \begin{cases} \frac{e^{2\pi|\omega|/a}}{e^{2\pi|\omega|/a}-1} & (\omega < 0), \\ \frac{1}{e^{2\pi\omega/a}-1} & (\omega > 0). \end{cases}}$$

(You can check that this agrees with our earlier result for an inertial detector in the limit $a \to 0$.)

This function $\pi(\omega)$ does not vanish for $\omega > 0$. The function G_+ is strictly positive frequency according to an inertial clock, but it contains both positive and negative frequencies when Fourier analyzed by a uniformly accelerated clock.

If we compare the rates for excitation and de-excitation, we have

$$\frac{\text{Rate}(E \to E + \omega)}{\text{Rate}(E + \omega \to E)} = e^{-2\pi\omega/a} \quad (\text{where } \omega > 0\ ) \ .$$

In order to be in equilibrium with the fluctuating radiation field, the detector must occupy the states of energy E and $E + \omega$ with a relative probability

$$\frac{\text{prob}(E + \omega)}{\text{prob}(E)} = e^{-2\pi\omega/a}.$$

This is just the Boltzmann distribution, with an effective temperature given by

$$T = \frac{a}{2\pi} \quad \left(\text{ or } k_B T = \frac{\hbar}{2\pi c} a \Rightarrow T = 4 \times 10^{-21} \text{ K} \times \left(a \text{ in m/s}^2\right)\right).$$

The detector behaves as though in contact with a thermal bath of radiation at this temperature.

A uniformly accelerated observer perceives a thermal radiation bath — the *Unruh effect* [24]. Why aren't you blinded by thermal photons when you floor the accelerator of your Maserati? It is because the Unruh radiation has an extraordinarily long wavelength for any acceleration that we normally experience. Roughly, the wavelength is the distance you would travel starting from rest before reaching a velocity comparable to the speed of light. For $a \approx 1g$, the wavelength is approximately 1 light year.

To further appreciate the sense in which the uniformly accelerated observer perceives the vacuum fluctuations as thermal fluctuations, consider thermal fluctuations of a harmonic oscillator

$$\begin{aligned} x(t) &= \frac{1}{\sqrt{2\omega}}\left(ae^{-i\omega t} + a^\dagger e^{i\omega t}\right) \\ \Rightarrow \langle x(t)x(0)\rangle_\beta &= \frac{1}{2\omega}\left\langle\left(e^{-i\omega t}aa^\dagger + e^{i\omega t}a^\dagger a\right)\right\rangle_\beta \\ &= \frac{1}{2\omega}\left[e^{-i\omega t}\langle n+1\rangle_\beta + e^{i\omega t}\langle n\rangle_\beta\right] \\ &= \frac{1}{2\omega}\left[\frac{e^{\beta\omega}}{e^{\beta\omega}-1}e^{-i\omega t} + \frac{1}{e^{\beta\omega}-1}e^{i\omega t}\right], \end{aligned}$$

using the Planck distribution for an oscillator at inverse temperature β.

The function $\pi(\omega)$ we derived has a similar form, times a density of states factor. Thus,

$$\langle 0|\phi(x(\tau))|\phi(x(0))\rangle|0\rangle = \langle\phi(t,\vec{0})\phi(0)\rangle_\beta \quad (\beta = 2\pi/a).$$

The two-point time correlations of the field as perceived by the uniformly accelerated observer are the same as those seen by an observer at rest in a thermal state of the field, where each oscillator mode of the field has an occupation number given in expectation value by its thermal value. The higher moments also coincide, for the $2n$-point correlations in both cases are Gaussian — they are expressible as products of two-point correlators. The stochastic properties of the radiation seen by the uniformly accelerated detector are precisely those of a thermal bath. We'll discuss this point in greater depth later in this chapter.

Remarks

- This conclusion illustrates the idea discussed in Chapter 3, that positive and negative frequency are observer-dependent notions. Vacuum fluctuations cannot excite the inertial detector, because they are purely positive frequency. But these fluctuations have a negative frequency component for the accelerated observer, and so can excite the detector carried by that observer.

- Note that this thermal radiation does not get redshifted by a boost along $\hat{z}$; it depends on proper acceleration, not velocity. For the accelerated observer, the boost is instead a *time* translation $\tau \to \tau +$ const.

4.4 Rindler coordinates

In order to appreciate more clearly what these effects have to do with field theory in curved spacetime, we will rederive the thermal radiation seen by an accelerated observer using Bogoliubov gymnastics. That is, we'll find the Bogoliubov transformation that relates the solutions that are positive frequency with respect to Minkowski time, and those that are positive (and negative) frequency with respect to the accelerated observer's clock ("Rindler time").

The coordinate system that is natural for an accelerated (Rindler) observer is one for which the world line

$$z = \frac{1}{a}\cosh a\tau,$$
$$t = \frac{1}{a}\sinh a\tau,$$

is at a fixed spatial position, and time is (proportional to) the observer's proper time τ. So, we replace z, t by a ξ, η defined by:

$$t := \xi \sinh\eta \Rightarrow \ dt = d\xi \sinh\eta + d\eta\xi \cosh\eta,$$
$$z := \xi \cosh\eta \Rightarrow \ dz = d\xi \cosh\eta + d\eta\xi \sinh\eta.$$

The invariant interval can now be expressed as

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 = \xi^2 d\eta^2 - d\xi^2 - dx^2 - dy^2.$$

In this coordinate system, $\frac{\partial}{\partial\eta}$ is a timelike Killing vector (the metric is η independent) — which we know is the generator of a $\hat{z}$ *boost* (a symmetry of Minkowski space). And in fact, the metric is static (the η =constant surfaces are orthogonal to the Killing vector; i.e., $g_{\eta i} = 0$).

These are the *Rindler coordinates*. Another way to express the coordinate transformation is in terms of light-cone coordinates:

$$u = t - z = -\xi e^{-\eta} = -e^{-\eta+\ln\xi} = -e^{-U},$$
$$v = t + z = \xi e^{\eta} = e^{\eta+\ln\xi} = e^{V},$$

where

$$U = \eta - \ln\xi,$$
$$V = \eta + \ln\xi,$$

and $\xi > 0$. Here U and V are light-cone coordinates for the Rindler metric, which can be written as

$$ds^2 = du dv - dx^2 - dy^2 = \xi^2\left(d\eta^2 - [d(\ln\xi)]^2\right) - dx^2 - dy^2 = \xi^2 dU dV - dx^2 - dy^2.$$

Notice that the Rindler coordinates do not cover all of Minkowski space but only the region $u < 0, v > 0$, denoted region R in the figure.

– IMAGE –

R stands for "right" and we refer to this region as the "right Rindler wedge;" likewise, the region L in the figure $(u > 0, v < 0)$ is called the "left Rindler wedge," while regions F $(u > 0, v > 0)$ and P $(u < 0, v < 0)$ are the "future" and "past" Rindler wedges.

There are coordinate singularities at $u = 0$ and $v = 0$ for a good reason. We have insisted that $\frac{\partial}{\partial\eta}$ is a boost generator. But boosts propagate spacetime points along hyperbolas as shown in the figure; the tangent to the hyperbola is timelike in Region R and L and spacelike in F and P. So, if $\frac{\partial}{\partial\eta}$ is a boost generator, then η must cross from being timelike to spacelike on the curves $u = 0$ and $v = 0$.

To cover Minkowski space completely, we need to use 4 sets of Rindler coordinates in the 4 regions R, L, F, P:

$$\begin{array}{ccc} F: & t = \xi\cosh\eta & u = t - z = e^{-\eta+\ln\xi}, \\ (u, v > 0) & z = \xi\sinh\eta & v = t + z = e^{\eta+\ln\xi}, \end{array}$$

$$\begin{array}{ccc} L: & t = -\xi\sinh\eta & u = e^{-\eta+\ln\xi}, \\ (u > 0, v < 0) & z = -\xi\cosh\eta & v = -e^{\eta+\ln\xi}, \\ P: & t = -\xi\cosh\eta & u = -e^{-\eta+\ln\xi}, \\ (u, v < 0) & z = -\xi\sinh\eta & v = -e^{-\eta+\ln\xi}. \end{array}$$

In regions F and P, the Rindler metric becomes

$$ds^2 = -\xi^2 d\eta^2 + d\xi^2 - dx^2 - dy^2.$$

Now consider the *causal structure* in this coordinate system.

– IMAGE –

An observer in region R who is static in Rindler coordinates, and hence traveling along the hyperbola ξ = constant, cannot receive any signal emitted in regions L or F (the portion of Minkowski spacetime with $u > 0$); hence $u = 0$ is a (future) *event horizon* for this observer. From a Minkowski space viewpont, there are null geodesics that never "catch up" with the accelerated observer. Similarly, no signal emitted from R can enter L or P, so $v = 0$ is a (past) event horizon for the Rindler observer.

– IMAGE –

If two Rindler observers at ξ_1 and ξ_2 exchange signals, the signal is redshifted (or blueshifted) when received. The Minkowski observer attributes the shift to the ordinary special relativistic Doppler effect. For example, we can choose an inertial frame in which the Rindler observer at ξ_1 is instantaneously at rest when he emits the signal. When the signal is received at $\xi_2 > \xi_1$, it is redshifted because the receiver is moving away from the emitter. On the other hand, in the frame where the receiver is instantaneously at rest when receiving a signal at ξ_1, the signal is blueshifted when received from an emitter at $\xi_2 > \xi_1$, because the emitter was moving toward the receiver at the time of emission.

– IMAGE –

To be more quantitative, consider a light signal emitted from $(t, z) = (0, \xi_1)$, which crosses the hyperbola $(t, z) = (\xi_2 \sinh \eta, \xi_2 \cosh \eta)$ at a Rindler time η given by

$$z = \xi_2 \cosh \eta = \xi_1 + t = \xi_1 + \xi_2 \sinh \eta \Rightarrow \xi_1 / \xi_2 = \cosh \eta \, (1 - \tanh \eta) \, .$$

As measured by a Minkowski observer, an object moving on the hyperbola has velocity $v = \frac{dz}{dt} = \tanh \eta$. Therefore, the Doppler shift of the signal when received is given by

$$\frac{\nu_2}{\nu_1} = \gamma(1 - v) = \frac{\xi_1}{\xi_2}.$$

That is, a signal emitted at Rindler position ξ_1 and received at Rindler position ξ_2 is redshifted for $\xi_2 > \xi_1$ and blueshifted for $\xi_2 < \xi_1$.

To Rindler observers, though, the hyperbolas $\xi = \xi_1$ and $\xi = \xi_2$ are static positions. For a signal emitted at ξ_1 and received at ξ_2, they attribute the redshift (or blueshift) to time dilation in a gravitational field. For a Rindler static clock, the proper time τ and the Rindler time η are related by

$$\frac{d\tau}{d\eta} = \sqrt{g_{00}(\xi)} = \xi,$$

resulting in the gravitational redshift

$$\frac{\nu_2}{\nu_1} = \left[\frac{g_{00}\left(\xi_2\right)}{g_{00}\left(\xi_1\right)} \right]^{-\frac{1}{2}} = \frac{\xi_1}{\xi_2}.$$

The gravitational field becomes very strong, and the gravitational time dilation diverges, at the past and future event horizons where $g_{00} = \xi^2 \to 0$.

Remarks

- We have already observed that at a given Rindler time η all static Rindler observers, irrespective of their Rindler position ξ, move with the same velocity $x = \tanh \eta$ in the Minkowski coordinate system. That makes sense. The Minkowski and Rindler time slices coincide when $t = \eta = 0$ — on that slice all Rindler observers are simultaneously at rest in both coordinate systems. Furthermore, we have seen that the Killing vector $\partial/\partial\eta$ of the Rindler metric is a generator of a Lorentz boost in Minkowski space. This means that for any value of the Rindler time η there is a Lorentz frame such that all the static Rindler observers are at rest. This is possible only if all the Rindler observers on a Rindler time slice (fixed η) are moving at the same velocity in the Minkowski coordinate system.

- A related observation is that Rindler time η has been chosen such that the motion of uniformly accelerated observers is always orthogonal to surfaces with $\eta =$ constant. It is instructive to check this claim by noting that an observer with Rindler spacetime position (η, ξ) has Minkowski spacetime position $(t, z) = (\xi \sinh\eta, \xi \cosh\eta)$, and that this observer's velocity in Minkowski space is

$$\left(\frac{\partial t}{\partial\eta}, \frac{\partial z}{\partial\eta}\right) = (\xi \cosh\eta, \xi \sinh\eta),$$

which is indeed orthogonal to (t, z) in the Minkowski metric.

- Because the Rindler coordinate system is static, the proper distance between positions ξ_1 and ξ_2 is independent of the Rindler time η.

– IMAGE –

To verify this, note that the invariant interval between Minkowski spacetime points $(t_1 = \xi_1 \sinh\eta, z_1 = \xi_1 \cosh\eta)$ and $(t_2 = \xi_2 \sinh\eta, z_2 = \xi_2 \sinh\eta)$ is

$$(t_1 - t_2)^2 - (z_1 - z_2)^2 = (\xi_1 - \xi_2)^2 \left(\cosh^2\eta - \sinh^2\eta\right) = (\xi_1 - \xi_2)^2 ,$$

which is independent of η.

4.5 Extending the coordinates beyond the Rindler horizon

Consider a family of accelerated observers who are forever confined to the right Rindler wedge R. Would they be able to infer the existence of regions P, L, and F, even though they are never able to visit there? Contemplating this puzzle will prepare us to confront a similar conundrum in the case of the black hole geometry.

Suppose a Rindler physicist ponders the behavior of a baseball moving inertially in region R. Choose a frame in which the ball is instantaneously at rest in Rindler coordinates at $\eta = 0$, and therefore always at rest in Minkowski coordinates. At Rindler time η, then, $z = \xi \cosh\eta =$ constant or

$$\xi = \xi_0 / \cosh\eta.$$

– IMAGE –

To the Rindler observer, the ball pops through the "floor" at $\xi = 0$ at Rindler time $\eta = -\infty$, reaches its maximum height ξ_0 at $\eta = 0$, and then falls, reaching the floor at time $\eta = \infty$. To her, the ball is moving in a gravitational field that becomes so powerful at the $\xi = 0$ floor that no light signal can escape from below the floor. The floor is an event horizon:

Our Rindler physicist recognizes, however, that the ball travels from $\xi = 0$ to $\xi = \xi_0$ in elapsed *proper* time $\tau = \xi_0$ (this is obvious to the Minkowski observer).

$$
\begin{aligned}
&d\tau^2 = \xi^2 d\eta^2 - d\xi^2 \text{ and } d\xi = -\frac{\xi_0 \sinh \eta}{\cosh^2 \eta} d\eta = -\xi \tanh \eta \, d\eta \\
&\Rightarrow d\tau^2 = \xi^2 \left(1 - \tanh^2 \eta\right) d\eta^2 = \xi_0^2 d\eta^2 / \cosh^4 \eta \\
&\Rightarrow \int d\tau = \xi_0 \int_0^\infty \frac{d\eta}{\cosh^2 \eta} = \xi_0 \tanh \eta |_0^\infty = \xi_0.
\end{aligned}
$$

She therefore perceives the need to extend the Rindler coordinate system beyond Rindler time $\eta = \infty$, and "before" Rindler time $\eta = -\infty$.

She thus discovers that there are actually two horizons, one in her past and one in her future. To account for the ball's peculiar motion, she constructs spacetime region P where the ball came from, and she constructs spacetime region F where the ball went to.

Thinking further, she realizes that there can be balls in the F region that are going up (which had to come from somewhere) and balls in the P region going down (which have to go somewhere). She therefore discovers that if her spacetime is *geodesically complete*, then it must contain a region L. She knows that L must exist, even though she can neither receive signals from L nor send signals to L, as long as she remains in R.

– IMAGE –

She might discover that there are alternative coordinate systems, like the Minkowski coordinates (t, z), that cover the spacetime smoothly and completely, without the need for any spurious coordinate singularities. But as a denizen of a uniformly accelerated world, she has a natural preference for Rindler coordinates, and she notices that these can be analytically extended beyond the right Rindler wedge. To her surprise, she finds that the Rindler time η in region R, when extended to region L, must run *backward* relative to Minkowski time in that region.

The reason is that η becomes a spacelike coordinate in region F, such that η decreases as a ball moves "downward" in that region and η increases as a ball moves "upward." The result is that when timelike geodesics in region F are extended back to R and to L, η runs in opposite directions in regions R and L. That is, η increases as Minkowski time moves forward in R, but η decreases as Minkowski time moves forward in L. This may seem like an arcane technical point, but it will be useful to bear in mind in what follows.

– IMAGE –

4.6 Field theory in Rindler space

Now, we want to do quantum field theory on Rindler spacetime. Specifically, we will want to express the ordinary Minkowski vacuum in terms of modes that are positive and negative frequency with respect to Rindler time, in order to determine what particles the uniformly accelerated observer sees in the Minkowski vacuum.

We should observe first that the "Rindler wedge" R is a globally hyperbolic spacetime in its own right. That is, the "constant time" surface $\eta =$ constant is Cauchy in region R (all timelike or null paths through points in R cross this surface). Therefore initial data on this surface completely determines a solution to the Klein-Gordon equation throughout R. It does *not* suffice, of course, to determine the solution in L, F, P, but this is of no consequence to Rindler observers, as those regions are beyond their event horizon and cannot be detected. The same remark applies to region L. Furthermore, a Cauchy surface in R together with one in L is Cauchy in all of Minkowski space.

We can find a set of solutions to the Klein-Gordon equation that are complete in R (or L) and are positive frequency with respect to Rindler time η. In Rindler coordinates, the KG equation (assumed massless for simplicity) is

$$\partial_\mu(\sqrt{g} g^{\mu\nu}\partial_\nu u) = 0,$$

where $g_{\mu\nu} = (\xi^2, -1, -1, -1)$, $\sqrt{g} = \xi$, $g^{\mu\nu} = (\xi^{-2}, -1, -1, -1)$. Therefore,

$$\begin{aligned}
&\left[\partial_\eta \xi^{-1}\partial_\eta - \partial_\xi \xi \partial_\xi - \xi\left(\partial_x^2 + \partial_y^2\right)\right] u = 0 \\
\Rightarrow &\left[\partial_\eta^2 - \xi\partial_\xi \xi \partial_\xi - \xi^2\left(\partial_x^2 + \partial_y^2\right)\right] u = 0.
\end{aligned}$$

This equation is easily separable, with the positive frequency solutions of the form

$$u = e^{-i\omega\eta} e^{i(k_x x + k_y y)} f_{k,\omega}(\xi),$$

where

$$0 = \left[\xi\partial_\xi \xi \partial \xi - \left(k_x^2 + k_y^2\right)\xi^2 + \omega^2\right] f_{k,\omega}(\xi).$$

This is Bessel's equation (for imaginary argument), and the normalizable solutions (up to normalization) are

$$f_{k,\omega}(\xi) = K_{i\omega}(k\xi) \quad \left(k^2 = k_x^2 + k_y^2\right).$$

By normalizing these solutions, we find a complete set of positive frequency solutions $u^{\text{Rin}}_{R,j}$ to the KG equation in the wedge R. Similarly, we can construct a complete set of positive frequency solutions $u^{\text{Rin}}_{L,j}$ in the wedge L. An important difference, though, is that

$$u^{\text{Rin}}_{L,j} \propto e^{i\omega\eta}$$

if η in region L is the coordinate obtained by analytically extending η in region R, given by $t/z = \tanh\eta$.

Now the u_R's and u_L's together are a complete basis for solutions to the KG equation in Minkowski space, so normalized plane waves can be expanded in terms of them. We have:

$$\begin{pmatrix} u^{\text{Min}} \\ u^{\text{Min}*} \end{pmatrix} = \begin{pmatrix} \alpha^\dagger & -\beta^T \\ -\beta^\dagger & \alpha^T \end{pmatrix} \begin{pmatrix} u_R^{\text{Rin}} \\ u_L^{\text{Rin}} \\ u_R^{\text{Rin}*} \\ u_L^{\text{Rin}*} \end{pmatrix},$$

which is shorthand for

$$u_i^{\text{Min}} = \left(\alpha_R^\dagger\right)_{ij} u_{R_j}^{\text{Rin}} + \left(\alpha_L^\dagger\right)_{ij} u_{Lj}^{\text{Rin}} + \left(-\beta_R^T\right)_{ij} u_{R_j}^{\text{Rin}*} + \left(-\beta_L^T\right)_{ij} u_{L_j}^{\text{Rin}*},$$

etc. We can then evaluate Bogoliubov coefficients by calculating KG inner products:

$$\begin{aligned} \left(\alpha_R^\dagger\right)_{ij} &= \left(u_i^{\text{Min}}, u_{R_j}^{\text{Rin}}\right), \\ \left(-\beta_R^T\right)_{ij} &= -\left(u_i^{\text{Min}}, u_{Rj}^{\text{Rin}*}\right), \end{aligned}$$

etc. That is, we can Fourier transform the solutions u_R^{Rin} and u_L^{Rin}.

Fortunately, there is an easy way to get the answer (following Unruh [24]). Let us first note that the Bogoliubov transformation is nontrivial – the positive frequency Rindler solutions are not strictly positive frequency as Minkowski solutions. To see this, recall that an arbitrary solution that is positive frequency with respect to Minkowski time has the expansion

$$\int \frac{d^3k}{(2\pi)^3 2k^0} \tilde{f}(k) e^{-ik\cdot x}$$

and so is analytic and bounded for all $\operatorname{Im} x^0 < 0$. In light-cone coordinates, the Minkowski positive frequency solutions are

$$e^{-i(\omega t - k_3 z)} = e^{-\frac{i}{2}(\omega + k_3)u} e^{-i\frac{1}{2}(\omega - k_3)v},$$

so we can just as well say that positive frequency solutions are analytic for $\operatorname{Im} u < 0, \operatorname{Im} v < 0$.

Now, in region R ($u < 0, v > 0$),

$$\begin{aligned} \ln(-u) &= -\eta + \ln\xi \\ \ln v &= \eta + \ln\xi \end{aligned} \Rightarrow \begin{aligned} \eta &= \tfrac{1}{2}[\ln v - \ln(-u)] \\ \ln\xi &= \tfrac{1}{2}[\ln v + \ln(-u)] \end{aligned},$$

and in region L ($u > 0, v < 0$),

$$\begin{aligned} \ln(u) &= -\eta + \ln\xi \\ \ln(-v) &= \eta + \ln\xi \end{aligned} \Rightarrow \begin{aligned} \eta &= \tfrac{1}{2}[\ln(-v) - \ln(u)] \\ \ln\xi &= \tfrac{1}{2}[\ln(-v) + \ln(u)] \end{aligned}.$$

We can continue from region R to region L, remaining in the domain of analyticity of Minkowski positive frequency solutions, by giving u and v small negative imaginary parts to avoid the branch point of the logarithm. But then

$$R : e^{-i\eta} \text{ continues to } \quad L : e^{-i\eta};$$

positive frequency in R continues to negative frequency in L with respect to Rindler time, recalling that Rindler time runs backward in region L. Since Minkowski time runs forward in both R and L, this shows that solutions that are positive frequency with respect to Minkowski time must have a negative frequency component with respect to Rindler time.

This observation also suggests how we can construct the Bogoliubov transformatioon. The positive frequency Minkowski solutions are the linear commbinations of Rindler R and L solutions that are analytic in u (and v) in the lower half plane. On the hyperbolas with constant ξ, we have

$$\begin{aligned} \eta &\sim -\ln(-u) \quad \text{(Region } R\text{)}, \\ \eta &\sim -\ln(u) \quad \text{(Region } L\text{)}. \end{aligned}$$

(Analytic continuation of the v dependence gives the same result as below.) Our positive frequency Rindler solutions have the behavior:

$$u^{\mathrm{Rin}}_{R,\omega} \sim \begin{cases} e^{-i\omega\eta} \sim e^{i\omega\ln(-u)} & \text{(Region } R\text{)}, \\ 0 & \text{(Region } L\text{)}, \end{cases}$$

$$u^{\mathrm{Rin}}_{L,\omega} \sim \begin{cases} 0 & \text{(Region } R\text{)}, \\ e^{i\omega\eta} \sim e^{-i\omega\ln(u)} & \text{(Region} L\text{)}. \end{cases}$$

To get a positive frequency Minkowski solution, we need to match these analytically across $u = v = 0$, where regions R and L meet; in particular we require analyticity for $\operatorname{Im} u < 0$.

If we analytically continue the region R solution$\sim e^{i\omega\ln(-u)}$,(where $u < 0$) to Region L, we get $e^{i\omega\ln(-u)}$, where u is now positive.

– IMAGE –

But the principle that a positive frequency Minkowski solution is the boundary value of a function analytic in the region $\operatorname{Im} u < 0$, $\operatorname{Im}(-u) > 0$, tells us to evaluate the log *above* the cut, where $\ln(-u) = \ln u + i\pi$.

Thus, a positive frequency Minkowski solution is

$$u^{\mathrm{Min}}_{1,\omega} \propto \begin{cases} e^{i\omega\ln(-u)} & \text{(Region } R\text{)} \\ e^{-\pi\omega} e^{i\omega\ln(u)} & \text{(Region } L\text{)} \end{cases}$$

$$\boxed{\Rightarrow u^{\mathrm{Min}}_{1,\omega} = \left(u^{\mathrm{Rin}}_{R,\omega} + e^{-\pi\omega} u^{\mathrm{Rin}*}_{L,\omega}\right) / \left(1 - e^{-2\pi\omega}\right)^{\frac{1}{2}},}$$

where we have included the denominator to normalize the solution.

Similarly, if we continue $e^{-i\omega \ln u}$ in region L to region R, we evaluate the log below the cut in the u plane. obtaining $\ln u - i\pi$. Thus, the positive frequency Minkowski solution becomes $e^{-i\omega(\ln u - i\pi)}$ in region R:

$$\boxed{u^{\rm Min}_{2,\omega} = \left(u^{\rm Rin}_{L,\omega} + e^{-\pi\omega} u^{\rm Rin*}_{R,\omega}\right) / \left(1 - e^{-2\pi\omega}\right)^{\frac{1}{2}}.}$$

Notice that when the Rindler frequency ω is large, the Minkowski solution $u^{\rm Min}_{1,\omega}$ is mostly localized in region R, while the Minkowski solution $u^{\rm Min}_{2,\omega}$ is mostly localized in region L.

We have found that the Bogoliubov coefficients for the solutions with *Rindler* frequency ω are:

$$\begin{pmatrix} u^{\rm Min}_{1,\omega} \\ u^{\rm Min}_{2,\omega} \end{pmatrix} = N_\omega \left(\alpha^\dagger_\omega, -\beta^T\right) \begin{pmatrix} u^{\rm Rin}_{R,\omega} \\ u^{\rm Rin}_{L,\omega} \\ u^{\rm Rin*}_{R,\omega} \\ u^{\rm Rin*}_{L,\omega} \end{pmatrix},$$

$$\alpha^\dagger_\omega = N_\omega \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad -\beta^T = N_\omega e^{-\pi\omega} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad N_\omega = \left(1 - e^{-2\pi\omega}\right)^{-\frac{1}{2}}.$$

Now, we recall our expression for the S-matrix:

$$|0, {\rm Min}\rangle = ({\rm phase})\, ({\rm det}\alpha)^{-\frac{1}{2}} \exp\left(\frac{1}{2} a^{\rm Rin\dagger} \left(-\beta^* \alpha^{-1}\right) a^{\rm Rin\dagger}\right) |0, {\rm Rin}\rangle,$$

and we use

$$-\beta^*_\omega \alpha^{-1}_\omega = e^{-\pi\omega} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \det\, \alpha_\omega = N^2_\omega = \frac{1}{1 - e^{-2\pi\omega}}.$$

to obtain (up to a phase):

$$|0, {\rm Min}\rangle = \prod_j \left(1 - e^{-2\pi\omega_j}\right)^{\frac{1}{2}} \exp\left(\sum_j e^{-\pi\omega_j} a^{\rm Rin\dagger}_{R,j} a^{\rm Rin\dagger}_{L,j}\right) |0, {\rm Rin}\rangle\ .$$

And since

$$\frac{1}{n!} a^{\rm Rin\dagger}_{R,j} a^{\rm Rin\dagger}_{L,j} |0, {\rm Rin}\rangle = |n_j, R\rangle \otimes |n_j, L\rangle$$

we have

$$|0, {\rm Min}\rangle = \prod_j \left[\left(1 - e^{-2\pi\omega_j}\right)^{\frac{1}{2}} \sum_{n_{j=0}}^{\infty} e^{-\pi\omega_j n_j} |n_j, R\rangle \otimes |n_j, L\rangle\right],$$

where n_j denotes the occupation number of Rindler mode j.

An observer who is static with respect to Rindler time in the wedge R, and is therefore unable to observe anything in L which lies beyond his event horizon, will describe this situation in terms of a density matrix for region R. Tracing out the unobserved particles in wedge L we obtain

$$\boxed{\rho_R = \mathrm{Tr}_L\left(|0,\mathrm{Min}\rangle\langle 0,\mathrm{Min}|\right) = \prod_j \left[\left(1-e^{-2\pi\omega_j}\right)\sum_{n_j} e^{-2\pi\omega_j n_j}\,|n_j,R\rangle\,\langle n_j,R|\right].}$$

This is precisely the density matrix (normalized to $\mathrm{Tr}\,\rho = 1$) for a thermal ensemble of Rindler particles. To extract a temperature, we recall that ω is the frequency with respect to the Rindler time coordinate η, and since

$$\begin{aligned} t &= \xi \sinh\eta, \\ z &= \xi \cosh\eta, \end{aligned}$$

while t and z are related to the proper time τ and proper acceleration a of a Rindler observer by

$$\left.\begin{aligned} t &= \frac{1}{a}\sinh a\tau \\ z &= \frac{1}{a}\cosh a\tau \end{aligned}\right. \qquad \Rightarrow \eta = a\tau,$$

So, with respect to proper time, the frequency ω is replaced by ω/a, and the Boltzmann factor is

$$e^{-2\pi\omega/a} = e^{-\omega/T} \text{ where } T = \frac{a}{2\pi}.$$

This expression for the temperature T agrees with our earlier computation of the response of a uniformly accelerated particle detector, as it should.

Remarks

- To reiterate, we have described two different ways of defining "positive frequency" in Minkowski spacetime, and two corresponding ways of defining a vacuum. These two notions of positive frequency invoke two different timelike Killing vectors of the spacetime metric (the time-translator generator $\partial/\partial t$ and the boost generator $\partial/\partial\eta$). We have seen that the two vacuum states are not equivalent. This happens because the two sets of positive frequency solutions are not related by merely analytically continuing from one time coordinate to the other.

  One way to understand why the two sets of positive frequency solutions are not connected by analytic continuation is that Rindler spacetime has a horizon. Because the Rindler Killing vector tips over from timelike to spacelike and then to (reversed) timelike, the positive frequency solutions with respect to η in region R extend to negative frequency in region L; this is the origin of the nontrivial Bogoliubov transformation.

- We streamlined the derivation of the Bogoliubov transformation by thinking through the consequences of analyticity for the solutions that are positive frequency with respect to

Minkowski time. Alternatively, we could use a more conventional method, and without doing a whole lot of work. From $u^{\rm Rin} = \alpha u^{\rm Min} + \beta u^{\rm Min^*}$, we have

$$\alpha_{ij} = \left(u_i^{\rm Rin}, u^{\rm Min}\right) \quad \beta_{ij} = -\left(u_i^{\rm Rin}, u_j^{\rm Min*}\right)$$

and evaluating these Klein-Gordon inner products amounts to Fourier transforming the Rindler solutions. These are related to Bessel functions, as noted previously.

We can simplify the integration by a felicitous choice of the surface on which we evaluate the Klein-Gordon inner product. It is convenient to choose a surface lying close to the (past) horizon $v = 0$. Close to $\xi = 0$, the Klein Gordon equation

$$\left[(\xi\partial_\varepsilon)^2 + \omega^2 - k_\perp^2\xi^2\right] f_{k,\omega}(\xi) = 0$$

simplifies, because the $k_\perp^2\xi^2$ term can be neglected, and the positive-frequency KG solutions have the form (up to normalization)

$$\begin{aligned} &e^{-i\omega\eta}e^{\pm i\ln\xi}e^{ik_\perp\cdot(\vec{x},\vec{y})} \\ &= \left\{\begin{matrix} e^{-i\omega U} \\ e^{-i\omega V}\end{matrix}\right\} e^{ik_\perp\cdot(\vec{x},\vec{y})} \qquad \text{(Region } R\text{)}, \end{aligned}$$

where

$$\begin{aligned} u &= t - z = -e^{-U}, \\ v &= t + z = e^{V}. \end{aligned}$$

What happens for small ξ is that the KG solutions of definite *Rindler* frequency become strongly blueshifted as they approach the event horizon, so that the transverse wave numbers become negligible, and the problem reduces to an effectively two-dimensional one.

On the other hand, the solutions with definite frequency with respect to Minkowski time are

$$\left\{\begin{matrix} e^{-i\omega u} \\ e^{-i\omega v}\end{matrix}\right\} e^{ik_\perp\cdot(\vec{x},\vec{y})}.$$

In general, the solutions are functions of u and v in Minkowski coordinates and they are functions of U and V in Rindler coordinates. But on the null surface with $v = 0$, they become functions of u alone in Minkowski coordinates and functions of U alone in Rindler coordinates. That makes our task easier, because we can derive the Bogoliubov transformation by Fourier analyzing a function of u. Written in terms of the Minkovski coordinate u, the solution that is positive frequency on the $v = 0$ surface at the boundary of region R is

$$f(u) = \begin{cases} e^{i\omega\ln(-u)} & \text{if } u < 0, \\ 0 & \text{if } u > 0, \end{cases}$$

which has Fourier transform $\tilde{f}(\sigma)$ where

$$f(u) = \int_{-\infty}^{\infty} \frac{d\sigma}{2\pi} e^{-i\sigma u} \widetilde{f}(\sigma).$$

As an exercise (Exercise 6.7), you may show that

$$\widetilde{f}(-\sigma) = -e^{-\pi\omega} \widetilde{f}(\sigma) \text{ if } \sigma > 0.$$

Thus,

$$u^{\text{Rin}}_{R,\omega} = N_\omega (u^{\text{Min}} - e^{-\pi\omega} u^{\text{Min}*});$$

we have derived the same Bogoliubov transformation as before.

- It is frequency ω with respect to Rindler time η that appears in the Boltzmann factor $e^{-2\pi\omega}$. This needs to be translated into frequency with respect to proper time of an observer fixed in Rindler coordinates, in order to determine what energy state of a uniformly accelerated detector gets excited. Since

  $$d\tau^2 = \xi^2 d\eta^2 \Rightarrow \frac{d}{d\tau} = \frac{1}{\xi} \frac{d}{d\eta},$$

  the proper-time frequency is ω/ξ. Hence we have $e^{-2\pi\xi\omega_{\text{proper}}}$, so $T = 1/2\pi\xi$. Because the proper acceleration of an observer at Rindler position ξ is $a = 1/\xi$, this reproduces our statement that temperature and acceleration are related by $T = a/2\pi$.

  An observer who gets closer and closer to the event horizon at $\xi = 0$ encounters radiation that gets hotter and hotter. Conversely, as thermal radiation pulls away from the horizon, propagating from ξ_1 to $\xi_2 > \xi_1$, it experiences a gravitational redshift:

  $$\frac{T_2}{T_1} = \left[\frac{g_{00}(\xi_2)}{g_{00}(\xi_1)} \right]^{-\frac{1}{2}} = \frac{\xi_1}{\xi_2}.$$

4.7 Unruh effect from a Minkowski perspective

What does the Minkowski observer see when the accelerating detector absorbs a quantum from the Rindler thermal gas? Since absorption changes the state of the radiation field, and the state is initially the Minkowski vacuum, she must see the *emission* of a Minkowski radiation quantum by the detector. If we invoke the Bogoliubov transformation, we have

$$a^{\text{Rin}}_R = \left(\alpha^*_R a^{\text{Min}} - \beta^*_R a^{\text{Min}\dagger} \right) = \frac{1}{(1 - e^{-2\pi\omega})^{\frac{1}{2}}} \left(a^{\text{Min}}_{1,\omega} + e^{-\pi\omega} a^{\text{Min}\dagger}_{2,\omega} \right).$$

(Remember that the Minkowski quantum created by $a^{\text{Min}}_{2,\omega}$ does not have definite Minkowski frequency). In the Minkowski Fock basis, then, absorption of the Rindler quantum excites the Minkowski vacuum:

$$|0,\mathrm{Min}\rangle \to a_R^{R\mathrm{in}}\,|0,\mathrm{Min}\rangle = \frac{e^{-\pi\omega}}{(1-e^{-2\pi\omega})^{\frac{1}{2}}} a_{2,\omega}^{\mathrm{Min}\dagger}\;\;|0,\mathrm{Min}\rangle\,.$$

Oddly, the mode function of this radiation quantum that is created when a detector absorbs a Rindler quantum in region R is

$$u_{2,\omega}^{\mathrm{Min}} = \frac{1}{(1-e^{-2\pi\omega})^{\frac{1}{2}}}\left(e^{-\pi\omega}u_{R,\omega}^{\mathrm{Rin}*} + u_{L,\omega}^{\mathrm{Rin}}\right),$$

which is predominantly localized in region L.

The energy of the absorbed quantum is seen by the Rindler observer to be

$$\omega/\xi = \omega_{\mathrm{proper}} = \omega a$$

where a is the acceleration of the detector. The Rindler energy is given by

$$E_{\mathrm{Rin}} = \int_\eta n^\mu n^\nu T_{\mu\nu}\, d^3x$$

where the integral is taken over a surface of constant Rindler time η, and n^μ is the unit normal to that surface. This unit vector (up to the rescaling by $\frac{1}{\xi}$) is the generator $\frac{\partial}{\partial\eta}$ of a Rindler time translation.

Both the Minkowski and Rindler observers can calculate how the Rindler energy E_{Rin} stored in the scalar field changes when the particle detector becomes excited, although of course the Rindler observer can see $T_{\mu\nu}$ only in region R. First let's compute the change in Rindler energy in region R from the Minkowski observer's viewpoint. Making use of the isomorphism which relates the single-particle Hilbert space to KG solutions, we have

$$\delta E_{\mathrm{Rin}} = \left(u_{2,\omega}^{\mathrm{Min}}, \frac{1}{\xi} i\frac{\partial}{\partial\eta}\Big|_{\mathrm{Region}\ R} u_{2,\omega}^{\mathrm{Min}}\right) / \left(u_{2,\omega}^{\mathrm{Min}}, u_{2,\omega}^{\mathrm{Min}}\right).$$

Rewriting $u_{2,\omega}^{\mathrm{Min}}$ in terms of Rindler solutions, this becomes

$$\delta E_{\mathrm{Rin}} = \frac{e^{-2\pi\omega}}{1-e^{-2\pi\omega}}\left(u_{R,\omega}^{\mathrm{Rin}\,*}, \frac{1}{\xi} i\frac{\partial}{\partial\eta} u_{R,\omega}^{\mathrm{Rin}\,*}\right) = \frac{e^{-2\pi\omega}}{1-e^{-2\pi\omega}}\left(\frac{\omega}{\xi}\right).$$

Restating this expression in terms of the inverse temperature $\beta = 2\pi\xi$ seen by the Rindler observer at position ξ, and the Rindler energy $E = \omega/\xi$ of the detected Rindler quantum, we have

$$\delta E_{\mathrm{Rin}} = E\frac{e^{-\beta E}}{1-e^{-\beta E}} = \frac{E}{e^{\beta E}-1}.$$

Consistency requires the Rindler observer to agree with the Minkowski observer that the adsorption of the Rindler radiation quantum has actually *increased* the energy stored in the radiation bath. At first that sounds wrong — shouldn't absorption of a quantum reduce the energy in the bath?

Let's think about this more carefully. For a radiation mode with energy E, the probability distribution governing the number of quanta n occupying the mode is

$$P_n = \left(1 - e^{-\beta E}\right) e^{-\beta n E},$$

and the mean occupation number is

$$\langle n(E)\rangle_\beta = \frac{1}{e^{\beta E} - 1}.$$

If a particle occupying this mode is detected, then we must update this probability distribution in light of the knowledge we have gained by detecting the particle, because the likelihood of detection is proportional to the occupation number. After detecting a particle (which reduces the occupation number by 1), the updated probability distribution is

$$P_{n-1} \propto n e^{-\beta n E}$$

The mean occupation number resulting from this updated distribution is

$$\langle n\rangle_{\text{after detection}} = \sum_n n P_n = \frac{2}{e^{\beta E} - 1} = 2\langle n(E)\rangle_\beta$$

(Exercise 6.8a). Contrary to our naive expectation, detecting a particle increases the expected number of particles in the mode rather than decreasing it. Indeed, the resulting increase in energy $E\langle n(E)\rangle_\beta$ agrees exactly with the Minkowski observer's conclusion. The Minkowski observer thinks that the energy in the field went up because the accelerating detector emitted a quantum in the Minkowski vacuum. From her perspective, this energy was provided by the agent who provided the force that caused the detector to accelerate. The Rindler observer thinks the energy went up (by the same amount) because the detector absorbed a quantum from the thermal Rindler bath, increasing the expected occupation number of the mode.

This clever insight is due to Unruh and Wald [25].

If $\langle E\rangle$ actually goes up when a particle is removed from the thermal bath, can we continue to extract energy indefinitely? Surely not, but why not?

Suppose the Rindler observer periodically checks whether her detector has become excited. Sometimes the answer is yes, and then the expected energy in the bath goes up as we just explained. But sometimes the answer is no, and in that case, too, she should update her probability distribution taking into account this new information. When she doesn't see a particle, the updated distribution results in a decrease in the expected occupation number, not an increase. On average, she does not gain energy.

Suppose for example that the detector is weakly coupled to the radiation, so that when the occupation number is n the probability of detection is εn (where $\varepsilon \ll 1$), and the probability of no detection is $1-\varepsilon n$. When no particle is detected, the updated probability distribution is

$$P_n \propto (1-\varepsilon n)e^{-n\beta E}.$$

As an exercise (Exercise 6.8b) you can check that when the possibility of no detection is taken into account, measurements do not, on average, increase $\langle E\rangle$.

4.8 Time correlations in thermal equilibrium

The Minkowski vacuum state is a pure state, but to the Rindler observer it appears to be a highly mixed state. This in itself is not shocking. The Minkowski vacuum exhibits correlations between the right and left Rindler wedges — we say that the regions R and L are entangled. In particular, we observed that, in the Minkowski vacuum, if a Rindler mode in R with a given frequency has occupation number n, then the corresponding mode in L with the same frequency is guaranteed to have occupation number n as well. Because of these vacuum correlations, the marginal density operator for region R, obtained by tracing out region L is a mixed state, even though the joint state of R and L is pure.

But the density operator seen by the uniformly accelerated observer is not merely mixed; it is a very special type of mixed state, a thermal state. Why should there be a connection between uniform acceleration and thermodynamics? We can understand this better by probing more deeply into the characteristic properties of time correlations in thermal equilibrium.

To get started, we consider (again) the *two-point correlation function* of a one-dimensional harmonic oscillator in thermal equilibrium. We have

$$x(t) = \frac{1}{\sqrt{2\omega}}\left(e^{-i\omega t}a + e^{i\omega t}a^\dagger\right),$$

and therefore

$$\begin{aligned}
G_+^\beta(t) &= \langle x(t)x(0)\rangle_\beta \\
&= \frac{1}{2\omega}\left\langle e^{-i\omega t}aa^\dagger + e^{i\omega t}a^\dagger a\right\rangle_\beta \\
&= \frac{1}{2\omega}\left[\langle n+1\rangle_\beta e^{-i\omega t} + \langle n\rangle_\beta e^{i\omega t}\right] \\
&= \frac{e^{\beta\omega}}{2\omega\left(e^{\beta\omega}-1\right)}\left(e^{-i\omega t} + e^{-\beta\omega}e^{i\omega t}\right).
\end{aligned}$$

Notice that the function

$$f_+^\beta(t) = e^{-i\omega t} + e^{-\beta\omega}e^{i\omega t}$$

has the property

$$f_+^\beta(t-i\beta) = e^{-\beta\omega}e^{-i\omega t} + e^{i\omega t} = f_+^\beta(-t).$$

So, the correlation function $G_+^\beta(t)$, when analytically continued to the complex t plane, has the property

$$G_+^\beta(t-i\beta) = G_+^\beta(-t)\Big) \equiv G_-^\beta(t) \quad \text{where} \quad G_-^\beta(t) = \langle x(0)x(t)\rangle_\beta.$$

This property (and its generalizations) is called the Kubo-Martin-Schwinger (KMS) condition; it is a fundamental property of correlation functions in statistical mechanics.

The KMS property is much more general than this example. It can be extended to arbitrary operators and Hamiltonians. Consider

$$\langle A(t)B(0)\rangle_\beta$$

where $A(t) = e^{iHt}A(0)e^{-iHt} \equiv e^{iHt}Ae^{-iHt}$ is an operator in the Heisenberg representation, and

$$\langle \mathcal{O}\rangle_\beta = \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H}\mathcal{O}\right), \qquad \text{where } Z = \mathrm{Tr}(e^{-\beta H})$$

is the thermal expectation value. Note first that

$$\langle A(t)B(0)\rangle_\beta = \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H}e^{iHt}Ae^{-itH}B\right) = \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H}Ae^{-itH}Be^{iHt}\right) = \langle A(0)B(-t)\rangle_\beta,$$

since the trace is cyclic, and $e^{-\beta H}$ commutes with e^{itH}. More generally,

$$\langle A(t)B(s)\rangle_\beta = \langle A(t+c)B(s+c)\rangle_\beta;$$

thermal equilibrium is time-translation invariant. In addition, if we extend the correlation function to complex time, and again use the cyclicity of the trace, we have

$$\begin{aligned}
\langle A(t-i\beta)B(0)\rangle_\beta &= \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H}e^{i(t-i\beta)H}Ae^{-i(t-i\beta)H}B\right)\\
&= \frac{1}{Z}\mathrm{Tr}\left(e^{itH}Ae^{-\beta H}e^{-itH}B\right)\\
&= \frac{1}{Z}\mathrm{Tr}\left(e^{-\beta H}e^{-itH}Be^{itH}A\right)\\
&= \langle B(-t)A(0)\rangle_\beta.
\end{aligned}$$

This is the KMS condition.

Now let's consider the special case of a free scalar field and pursue the analytic continuation into the complex t plane in greater detail. The free field is just a superposition of harmonic oscillators; so, the thermal correlation function

$$G_+^\beta(t,\vec{x}) = \langle\phi(t,\vec{x})\phi(0)\rangle_\beta$$

is given by a sum over the modes of the scalar field. Considering it as a function of t for $\vec{x}$ fixed, we have

$$G_+^\beta(t) \propto \sum_{\text{modes } k}\left(e^{-i\omega_k t} + e^{-\beta\omega_k}e^{i\omega_k t}\right).$$

For $\beta\to\infty$ (zero temperature), this turns into the positive frequency function

$$G_+(t) \propto \sum_k e^{-i\omega_k t}$$

that we have considered before; the sum converges, and $G_+(t)$ is an analytic function, for $\mathrm{Im}\, t < 0$. There is also a function

$$G_-^\beta(t,\vec{x}) = \langle \phi(0)\phi(t,\vec{x})\rangle_\beta = G_+^\beta(-t,-\vec{x}) = G_+^\beta(-t,\vec{x})$$

(using time-translation invariance and rotational invariance), which is purely negative frequency for $\beta \to \infty$:

$$G_-(t) \propto \sum_k e^{i\omega_k t};$$

this is analytic for $\mathrm{Im}\, t > 0$. For finite β, the analytic structure of these functions is modified; the sum

$$G_+^\beta \propto \sum_k \left(e^{-i\omega_k t} + e^{-\beta\omega_k} e^{i\omega_k t}\right)$$

is a convergent only on a strip

$$-\beta < \mathrm{Im}\, t < 0,$$

and G_-^β converges only for

$$0 < \mathrm{Im}\, t < \beta.$$

Because scalar fields $\phi(x)$ and $\phi(y)$ commute when x and y are spacelike separated, there is an interval

$$-|\vec{x}| < t < |\vec{x}|$$

on the real t axis where $G_-^\beta(t) = G_+^\beta(t)$.

– IMAGE –

This means that $G_-^\beta(t)$, which is analytic on the upper strip (above the real axis) can be viewed as the (unique) analytic continuation of $G_+^\beta(t)$, which is analytic on the lower strip. Evidently, this function has a branch cut beginning at $t = |\vec{x}|$ (and a similar branch cut beginning at $t = -|\vec{x}|$), where

$$G_+^\beta(t,\vec{x}) - G_-^\beta(t,\vec{x}) = \langle[\phi(x),\phi(0)]\rangle_\beta = iG^{\beta=0}(x) \neq 0.$$

We may now interpret $G_+(t)$ and $G_-(t)$ as values for a single analytic function in the complex t plane, where G_+ is obtained as we approach a cut on the real axis from below, G_- is obtained as we approach the cut from above, and iG is the discontinuity across the cut. (Recall that the field commutator is a number, not an operator, so its expectation value does not depend on β. Strictly speaking, the singularity at $t = |\vec{x}|$ is a branch cut only if $m^2 \neq 0$; in the massless theory, the commutator has support only on the light cone, and the singularity becomes a pole.)

Now recall the KMS condition, which in this case tells us that

$$G_+^\beta(t - i\beta) = G_-^\beta(t).$$

From this and the property noted above (that $G_-^\beta(t)$ is the analytic continuation of $G_+^\beta(t)$ around the branch points), we can see how to analytically continue G_+ beyond the lower strip, to the whole complex t plane. The KMS condition says that our analytic function is periodic with period $i\beta$, and so it can be periodically extended to all values of $\operatorname{Im} t$.

We find then that thermal correlation functions are given by boundary values of an analytic function in the complex t plane that is periodic in imaginary time with period β. The result is evidently quite general and applies not just to the free scalar field but instead to arbitrary local field operators in an interacting theory; the KMS condition is satisfied for the correlators of such fields, and we can continue across the real t axis because the fields are local observables that commute at spacelike separation.

One may also wish to consider the time-ordered thermal Green function

$$iG_F^\beta(t) = \theta(t)G_+^\beta(t) + \theta(-t)G_-^\beta(t).$$

This may be viewed as a boundary value of the same analytic function in the complex t plane, but where now we approach the cut on the real t axis from below for $t > 0$ and from above for $t < 0$. More generally, since the Riemann surface of this function has multiple sheets, in discussing its periodicity properties we need to specify a particular sheet. In particular, when we analytically continue the function beyond the strip, we'll get a different result if we continue around the branch points (from below the cut to above the cut) than if we continue across the cut.

4.9 Klein-Gordon Green functions in flat spacetime

In the case of a free scalar field theory, the analytic function we have constructed, for either zero or nonzero temperature, can be regarded as an analytically continued Green function for the Klein-Gordon equation. Let's first consider the zero-temperature case, and denote by $G(x)$ the analytic function that matches $G_+(x)$ for Im $t < 0$ and matches $G_-(x)$ for Im $t > 0$. For $|\vec{x}| > 0$, this function has no singularities on the imaginary axis $t = i\tau$. Specifying G for imaginary time determines it throughout the cut t plane, because of the uniqueness of analytic continuation.

In fact, on flat space $G(x)$ for $t = i\tau$ is the unique Green function for the *Euclidean* Klein-Gordon equation

$$\left(\Box_E - m^2\right)\phi(\tau,\vec{x}) = \left(\frac{\partial^2}{\partial\tau^2} + \vec{\nabla}^2 - m^2\right)\phi(\tau,\vec{x}) = 0.$$

Because $\phi = 0$ is the unique normalizable solution to this equation, the KG operator $\Box_E - m^2$ is *invertible* (unlike the Minkowski KG operator). The unique decaying solution to

$$\left(\Box_E - m^2\right)G_E(\tau,\vec{x}) = -\delta(\tau)\delta^3(\vec{x})$$

may be expressed as

$$G_E(\tau, \vec{x}) = \int \frac{d^4k}{(2\pi)^4} e^{ik^0\tau} e^{i\vec{k}\cdot\vec{x}} \frac{1}{(k^0)^2 + \vec{k}^2 + m^2}.$$

We can evaluate the k^0 integral by completing the contour in the upper half plane for $\tau > 0$ or the lower half plane for $\tau < 0$, enclosing a pole at $k^0 == \pm i\omega_k = \pm i\sqrt{\vec{k}^2 + m^2}$, thus obtaining

$$G_E(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{i\vec{k}\cdot\vec{x}} e^{-\omega_k|\tau|}.$$

This Euclidean Green function can be analytically continued to imaginary $\tau = it$, or equivalently to real time t. The result for positive τ is

$$G_+(t, \vec{x}) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{i\vec{k}\cdot\vec{x}} e^{-i\omega_k t}$$

and for negative τ we obtain

$$G_-(t, \vec{x}) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{i\vec{k}\cdot\vec{x}} e^{i\omega_k t}$$

As claimed, this agrees with the two-point correlation function computed earlier — we get G_+ if we approach the real t axis from below and G_- if we approach from above.

A similar construction allows us to relate the finite-temperature correlation function G^β to a Euclidean KG Green function, except that now the Green function obeys a *cylinder* boundary condition; that is, it is a periodic function of τ with period β. This Green function inverts the KG operator on a cylinder, where τ is identified with $\tau + \beta$.

On the cylinder, the unique decaying solution to

$$\left(\Box_E - m^2\right) G_E^\beta(\tau, \vec{x}) = -\delta(\tau)\delta^3(\vec{x})$$

may be expressed as

$$G_E^\beta(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\beta} \sum_{n=-\infty}^{\infty} e^{i\vec{k}\cdot\vec{x}} e^{i\frac{2\pi}{\beta}n\tau} \frac{1}{\left(\frac{2\pi}{\beta}\right)^2 + \vec{k}^2 + m^2},$$

since $\frac{1}{\beta}\sum_{n=-\infty}^{\infty} e^{i\frac{2\pi}{\beta}n\tau} = \delta(\tau)$ on the interval $\tau \in [0, \beta]$. By evaluating the sum over n explicitly, we can verify that this Euclidean Green function agrees with the thermal correlators we obtained earlier, G_+^β on the lower strip, and G_-^β on the upper strip. This ensures that the thermal correlators are the unique analytic continuation of the Euclidean Green function off the imaginary-time axis.

There is a trick, the Sommerfeld-Watson transform, for doing such sums. We can express the sum as a contour integral, where summing over the residues of the poles enclosed by the contour reproduces the sum of interest. This trick is useful if we are able to simplify the evaluation of the integral by distorting the contour.

In particular, we may invoke the identity

$$\sum_{n=-\infty}^{\infty} f\left(\frac{2\pi}{\beta}n\right) = \frac{1}{2\pi i}\int_{-\infty-i\epsilon}^{\infty-i\epsilon} dz f(z)\frac{\beta}{2}\cot\left(\frac{\beta z}{2}\right) + \frac{1}{2\pi i}\int_{\infty+i\epsilon}^{-\infty+i\epsilon} dz f(z)\frac{\beta}{2}\cot\left(\frac{\beta z}{2}\right).$$

This works because $\frac{\beta}{2}\cot\left(\frac{\beta z}{2}\right)$ has poles at $z = \frac{2\pi}{\beta}n$, each with residue $= 1$.

Because $\cot\frac{\beta z}{2} - i$ decays rapidly in the lower half-plane (LHP), and $\cot\frac{\beta z}{2} + i$ decays rapidly in the upper half-plane (UHP), it is convenient to rewrite this expression as

$$\begin{aligned}\sum_{n=-\infty}^{\infty} f\left(\frac{2\pi}{\beta}n\right) = \frac{\beta}{2\pi}\int_{-\infty}^{\infty} dz f(z) &+ \frac{1}{2\pi i}\int_{-\infty-i\epsilon}^{\infty-i\epsilon} dz f(z)\frac{\beta}{2}\left(\cot\frac{\beta z}{2} - i\right) \\ &+ \frac{1}{2\pi i}\int_{\infty+i\epsilon}^{-\infty+i\epsilon} dz f(z)\frac{\beta}{2}\left(\cot\frac{\beta z}{2} + i\right)\end{aligned}$$

The integral below the real axis, accompanied by completion of the contour, encloses poles of $f(z)$ in the LHP, and the integral above the real axis encloses poles of $f(z)$ in the UHP.

In the evaluation of $G_E^\beta(\tau, \vec{x})$, we have

$$f(z) = \frac{1}{z^2 + \omega_k^2},$$

which has poles at $z = \pm i\omega_k$. To evaluate the residues of these poles we observe that

$$\cot(\pm i\beta\omega/2) = \pm i\frac{e^{-\beta\omega/2} + e^{\beta\omega/2}}{e^{-\beta\omega/2} - e^{\beta\omega/2}} = \mp 2\left(\frac{1}{2} + \frac{1}{e^{\beta\omega} - 1}\right).$$

As an exercise (Exercise 6.9) you can check that, assuming $-\beta < \tau < \beta$,

$$G_E^\beta(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{i\vec{k}\cdot\vec{x}} \frac{e^{\beta\omega_k}}{e^{\beta\omega_k} - 1}\left(e^{-\omega_k|\tau|} + e^{\omega_k(|\tau|-\beta)}\right).$$

The expression

$$\frac{e^{\beta\omega_k}}{e^{\beta\omega_k} - 1}\left(e^{-\omega_k|\tau|} + e^{\omega_k(|\tau|-\beta)}\right)$$

agrees with $G_+^\beta \sim e^{-i\omega_k t} + e^{-\beta\omega_k}e^{i\omega_k t}$ when continued to $t = i\tau$ for $\tau < 0$, and agrees with $G_-^\beta \sim e^{i\omega_k t} + e^{-\beta\omega_k}e^{-i\omega_k t}$ when continued to $t = i\tau$ for $\tau > 0$. So, the analytically continued cylinder Green function agrees with G_+^β on the strip below the real axis, and with G_-^β on the strip above the real axis; therefore it matches perfectly with the function we constructed earlier.

4.10 Klein-Gordon Green functions in curved spacetime

These observations concerning the two-point thermal correlation function can be extended to more general spacetimes. Consider the KG equation in the background metric $g_{\mu\nu}$, namely

$$\left(\frac{1}{\sqrt{g}}\partial_\mu\sqrt{g}g^{\mu\nu}\partial_\nu + m^2\right)\phi(x) = 0.$$

If the spacetime is static, then the metric is time independent and

$$g_{0i} = 0,$$

so we have

$$\left(g^{00}\frac{\partial^2}{\partial t^2} + \frac{1}{\sqrt{g}}\partial_i\sqrt{g}g^{ij}\partial_j + m^2\right)\phi(x) = 0$$

or

$$\left(\frac{\partial^2}{\partial t^2} + g_{00}\left(\frac{1}{\sqrt{g}}\partial_i\sqrt{g}g^{ij}\partial_j + m^2\right)\right)\phi(x) = 0.$$

In shorthand we may write this as $\left(\frac{\partial^2}{\partial t^2} + K\right)\phi = 0$ where K is a positive differential operator that acts on the $\vec{x}$ dependence of $\phi(\vec{x},t)$ ($K = -\vec{\nabla}^2 + m^2$ for flat spacetime).

If we diagonalize K,

$$Ku_i(\vec{x}) = \omega_i^2 u_i(\vec{x}),$$

and normalize the eigenmodes so that

$$\int d^3x\sqrt{g^{00}h}u_i(\vec{x})^* u_j(\vec{x}) = \delta_{ij}$$

then positive frequency solutions normalized with respect to KG inner product are

$$\frac{1}{\sqrt{2\omega_i}}u_i(\vec{x})e^{-i\omega_i t}.$$

Here $d^3x\sqrt{h}$ is the induced spatial volume element on a time slice, and $\sqrt{g^{00}}$ appears because $n^\mu\partial_\mu = \sqrt{g^{00}}\partial_t$.

Expanding fields in terms of these modes, we have

$$\phi(t,\vec{x}) = \sum_i \frac{1}{\sqrt{2\omega_i}}\left[u_i(\vec{x})e^{-i\omega_i t}a_i + u_i(\vec{x})^* e^{i\omega_i t}a_i^\dagger\right],$$

and if we define a vacuum state $|0\rangle$ by demanding $a_i|0\rangle = 0$, we then have the zero-temperature two-point function

$$G_+(t,\vec{x},\vec{y}) = \langle 0|\phi(t,\vec{x})\phi(0,\vec{y})|0\rangle = \sum_i \frac{1}{2\omega_i}u_i(\vec{x})u_i(\vec{y})^* e^{-i\omega_i t}.$$

If we evaluate the expectation value in the thermal ensemble with inverse temperature β rather than the vacuum we have instead

$$G_+^\beta(t,\vec{x},\vec{y}) = \langle \phi(t,\vec{x})\phi(0,\vec{y})\rangle_\beta = \sum_i \frac{1}{2\omega_i} u_i(\vec{x})u_i(\vec{y})^* \left(\left\langle a_i a_i^\dagger \right\rangle_\beta e^{-i\omega_i t} + \left\langle a_i^\dagger a_i \right\rangle_\beta e^{i\omega_i t} \right)$$
$$= \sum_i \frac{1}{2\omega_i} u_i(\vec{x})u_i(\vec{y})^* \frac{e^{\beta\omega_i}}{e^{\beta\omega_i}-1} \left(e^{-i\omega_i t} + e^{-\beta\omega_i} e^{i\omega_i t} \right]).$$

To derive this we used that for each frequency ω, the sum over mode functions with that frequency obeys $\sum_{\omega_i=\omega} u_i(\vec{x})u_i(\vec{y})^* = \sum_{\omega_i=\omega} u_i(\vec{x})^* u_i(\vec{y})$. We can ensure this is true by choosing the mode functions appropriately. Note that if $u_i(\vec{x})e^{-i\omega_i t}$ is a solution with positive frequency ω_i, then $u_i(\vec{x})^* e^{-i\omega_i t}$ is also a solution with the same frequency. (For example, $e^{\pm i\vec{k}\cdot\vec{x}} e^{-i\omega t}$ are linearly independent solutions in the flat-space case.)

We can continue $G_+, G_-, G_+^\beta, G_-^\beta$ away from the real-t axis following the same argument we used in the case of flat space. In particular, we can show that these continued functions become the Euclidean Green functions of the KG equation on the imaginary-t axis. The Euclidean Green function satisfies

$$g^{00}\left(\frac{\partial^2}{\partial\tau^2} + K\right) G_E(t,\vec{x},\vec{y}) = \frac{-1}{\sqrt{g}} \delta(t)\delta^3(\vec{x}-\vec{y})$$
$$\Rightarrow \left(\frac{\partial^2}{\partial\tau^2} + K\right) G_E(t,\vec{x},\vec{y}) = \frac{-1}{\sqrt{g^{00}h}} \delta(t)\delta^3(\vec{x}-\vec{y}).$$

Here the four-dimensional δ-function is suitably normalized for integration against the volume element $d^4x\sqrt{g}$, and we have used $g = g_{00}h = \left(g^{00}\right)^{-1} h$.

Using the completeness relation satisfied by the solutions,

$$\sum_i u_i(\vec{x})u_i(\vec{y})^* = \frac{1}{\sqrt{g^{00}h}} \delta^3(\vec{x}-\vec{y}),$$

the zero-temperature Green function may be expressed as

$$G_E(\tau,\vec{x},\vec{y}) = \int \frac{dk^0}{2\pi} \sum_i u_i(\vec{x})u_i(\vec{y})^* e^{ik^0\tau} \frac{1}{(k^0)^2 + \omega_i^2}.$$

As previously discussed, we can do the k^0 integral by completing the contour in the upper half place for $\tau > 0$ and in the lower half plane for $\tau < 0$; then the continuation of the Green function to real time t agrees with $G_+(t,\vec{x},\vec{y})$ when the real-t axis is approached from below, and agrees with $G_-(t,\vec{x},\vec{y})$ when the real-t axis is approached from above.

For finite inverse temperature β, the Euclidean thermal Green function is obtained by imposing the cylinder boundary condition, identifying τ with $\tau + \beta$. Then the Green function is given by the sum

$$G_E^\beta(\tau,\vec{x},\vec{y}) = \frac{1}{\beta} \sum_{n=-\infty}^{\infty} \sum_i u_i(\vec{x})u_i(\vec{y})^* e^{i\frac{2\pi}{\beta}n\tau} \frac{1}{\left(\frac{2\pi n}{\beta}\right)^2 + \omega_i^2}.$$

Manipulating the sum using the Sommerfeld-Watson transform as we did in the flat-space case, we can infer that the analytic continuation of this Euclidean Green function agrees with $G_-^\beta(t)$ in the upper strip with $\beta > \operatorname{Im} t > 0$, and with $G_+^\beta(t)$ in the lower strip with $0 > \operatorname{Im} t > -\beta$.

To summarize, we have shown that, for any static spacetime, the two-point correlation function of a free scalar field in the thermal ensemble can be obtained from the Klein-Gordon Green function on the "Euclidean section" of the spacetime (assumed to be periodic in Euclidean time with period β), by analytically continuing the Green function to real time.

4.11 Rindler space revisited

Now consider the special case of Rindler spacetime, recalling that in the right Rindler wedge R, the Minkowski coordinates (t, z) are related to the Rindler coordinates (η, ξ) by

$$t = \xi \sinh \eta,$$
$$z = \xi \cosh \eta.$$

If we continue to imaginary time, where $t = i\tau$ and $\eta = i\eta_E$, this becomes

$$\tau = \xi \sin \eta_E,$$
$$z = \xi \cos \eta_E.$$

The Euclidean coordinates (τ, z) are periodic functions of the Euclidean Rindler time τ with period 2π. A peculiar feature of this transformation relating the Euclidean Minkowski coordinates to the Euclidean Rindler coordinates is that, although the right Rindler wedge R has horizons that can be crossed by geodesics, the region $\xi \in [0, \infty]$ and $\eta_E \in [0, 2\pi)$ covers the entire (τ, z) plane. The Rindler horizon at $\xi = 0$ becomes the coordinate singularity at the origin of a polar coordinate system in the Euclidean Rindler spacetime.

This periodicity in imaginary Rindler time provides a fresh perspective on why the quantum fluctuations in the Minkowski vacuum state look like thermal fluctuations to the uniformly accelerated Rindler observer. We have already seen that the zero-temperature correlation function in Minkowski space,

$$\langle 0, \text{Mink}|\phi(t, \vec{x})\phi(0, \vec{x}')|0, \text{Mink}\rangle,$$

can be obtained by analytically continuing the Euclidean Klein-Gordon Green function $G_E(\tau, \vec{x}, \vec{x}')$ to real time. After transforming to Rindler coordinates, this same function on Eucldean spacetime solves the Euclidean Rindler space KG equation

$$\left[\xi^{-2}\partial_{\eta_E}^2 + \frac{1}{\xi}\partial_\xi \xi \partial_\xi + \partial_x^2 + \partial_y^2\right] G_E(\eta_E, \xi, x, y; \xi', x', y') = \frac{1}{\xi}\delta(\eta_E)\delta(\xi - \xi')\delta(x - x')\delta(y - y'),$$

and furthermore is periodic in η_E with period 2π. If we now continue this Green function to real Rindler time, we obtain a thermal correlator with inverse temperature $\beta = 2\pi$. Because the Rindler time $d\eta$ corresponds to proper time $\xi d\eta$ for an observer at Rindler position ξ, we conclude that this static Rindler observer (with proper acceleration $a = 1/\xi$) sees a thermal bath with inverse temperature $\beta = 2\pi\xi = 2\pi/a$. This agrees with our previous analysis of the uniformly accelerated observer in the Minkowski vacuum.

So far we have been considering the special case of a free scalar field, but the conclusion that the Rindler observer sees a thermal bath actually follows from quite general properties of correlation functions, and can therefore be extended to interacting fields on Rindler spacetime. We have already seen that the KMS condition is a general property of thermal correlators; this condition, together with the requirement that fields commute at spacelike separation, sufficed to tell us that the

$$\langle \phi(t,\vec{x})\phi(0,\vec{y})\rangle_\beta$$

is the boundary value of a function $G(t,\vec{x},\vec{y})$ that is analytic in the complex t plane, periodic in t with period $i\beta$, with branch cuts only at $\mathrm{Im}\, t = \beta n$ where n is an integer. Furthermore, the discontinuity across the cut on the real axis is such that

$$G(t,\vec{x},\vec{y})|_{\text{above cut}} = G(-t,\vec{y},\vec{x})|_{\text{below cut}}.$$

These properties of $G(t,\vec{x},\vec{y})$ hold not only for Minkowski spacetime and Rindler spacetime, but for any static spacetime.

Now let's revisit our analysis of how a detector responds to field fluctuations. As discussed in §4.2 and §4.3, the response of a detector held at position $\vec{x}$ is governed by a response function

$$\tilde{G}(\omega,\vec{x},\vec{x}) = \int dt e^{-i\omega t} G(t,\vec{x},\vec{x})|_{\text{below cut}} \,.$$

Because $\tilde{G}(t,\vec{x},\vec{x})$ is analytic in the lower strip $0 > \mathrm{Im}\, t > -\beta$, we can distort the contour from the top to the bottom of that strip, obtaining

$$\begin{aligned}\tilde{G}(\omega,\vec{x},\vec{x}) =& \int dt e^{-i\omega(t-i\beta)} G(t-i\beta,\vec{x},\vec{x})|_{\text{above cut}} \\ =& \int dt e^{-i\omega(t-i\beta)} G(t,\vec{x},\vec{x})|_{\text{above cut}} \\ =& e^{-\beta\omega} \int dt e^{-i\omega t} G(-t,\vec{x},\vec{x})|_{\text{below cut}} \\ =& e^{-\beta\omega} \int dt e^{i\omega t} G(t,\vec{x},\vec{x})|_{\text{below cut}} \; = e^{-\beta\omega}\tilde{G}(-\omega,\vec{x},\vec{x}).\end{aligned}$$

where in the first step we slide down the contour across the lower strip, in the second step we use the periodicity of $G(t)$ in t with period $i\beta$, in the third step we flip the sign of t by crossing the cut, and in the fourth step we replace $t \to -t$. This relation between $\tilde{G}(\omega)$ and $\tilde{G}(-\omega)$ holds for both positive and negative ω. Either way, we conclude that the positive frequency response is suppressed compared to the negative frequency response by a Boltzmann factor $e^{-\beta\omega}$. A detector coupled to a mode of the field will be thermally excited with inverse temperature β when in equilibrium with these field fluctuations.

The assumed properties of $G(t, \vec{x}, \vec{x})$ are satisfied by interacting quantum fields in the Minkowski vacuum state, if x denotes the Rindler position and t the Rindler time. Hence, the vacuum appears to be a thermal state to a uniformly accelerated observer, with temperature $T = a/2\pi$, even if the fields are not free.

Remarks

- This idea of sliding down the contour was used in a related context by Hartle and Hawking [15].
- More details about the analytic properties of thermal Green functions can be found in Fulling and Ruijsenaars [11].

5 Black Hole Radiance

5.1 Schwarzschild geometry

In 1916, K. Schwarzschild found that the empty space Einstein equation

$$R_{\mu\nu} = 0$$

admits the spherically symmetric solution

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \frac{1}{1 - \frac{2M}{r}} dr^2 - r^2 \left(d\theta^2 + \sin^2\theta d\phi^2\right).$$

(Here $G = 1$; otherwise, replace $M \to GM$). We may *define* M as the total mass of the gravitating body at the origin; this then agrees with gravitational mass M in Newtonian theory for large r.

It was later shown by G. Birkhoff (1923) that the above solution is the *unique* spherically symmetric solution of $R_{\mu\nu} = 0$. ("Uniqueness" meaning, of course, up to reparametrization.) Thus, in particular, spherical symmetry implies the geometry is *static*. "Birkhoff's theorem" says that a pulsating spherical body cannot radiate gravitational waves.

According to Birkhoff's theorem, the Schwarzschild metric describes geometry exterior to a spherical star, even as the star undergoes (spherically symmetric) gravitational collapse. When the surface of the star collapses inside the (coordinate) singularity of the metric at $r = 2M$, we say that a (static) black hole has formed. The surface with $r = 2M$ is called the event horizon of the black hole.

To understand this geometry better, consider the radial null geodesics. For this purpose, it is convenient to rewrite the metric as

$$ds^2 = \left(1 - \frac{2M}{r}\right) \left(dt^2 - \left[\frac{dr}{1 - \frac{2M}{r}}\right]^2\right) - r^2 d\Omega^2$$

and make a coordinate transformation

$$dr_* = \frac{r dr}{r - 2M} \Rightarrow r_* = \int \left(1 + \frac{2M}{r - 2M}\right) dr = r + 2M \ln\left(\frac{r - 2M}{2M}\right).$$

The $2M$ in the denominator is the arbitrary constant of integration. Now we have

$$ds^2 = \left(1 - \frac{2M}{r}\right) \left(dt^2 - dr_*^2\right) - r^2 d\Omega^2$$

and hence the radial null geodesics can be expressed as $r_* = \pm t + \text{constant}$.

This "tortoise coordinate" r_* goes to $r_* = -\infty$ as the sphere $r = 2M$ is approached asymptotically — it is the radial coordinate apportioned according to how much time an infalling photon spends as it approaches $r = 2M$, as perceived by a fiducial observer (a.k.a. "FIDO") at a fixed position in Schwarzschild coordinates. To this FIDO, the photon does not reach $r = 2M$ until $t = +\infty$; in this sense the sphere at $r = 2M$ seems to be "infinitely far away." Similarly a photon moving radially outward seems to emerge from the $r = 2M$ surface at $t = -\infty$.

This happens because of the very strong gravitational field which causes the gravitational time delay

$$dt = \frac{d\tau}{\sqrt{1 - \frac{2M}{r}}}$$

of a static clock to diverge at $r = 2M$. That a signal sent by a FIDO does not reach the horizon until $t = \infty$, and that a signal received by a FIDO emerges from the horizon at $t = -\infty$, is reminiscent of Rindler spacetime.

Let us pursue this analogy further. Recall the Rindler metric

$$ds^2 = \xi^2(d\eta^2 - d\xi_*^2)$$

where $\xi_* = \ln \xi$ and the $-dx^2 - dy^2$ part of the metric is suppressed. The coordinate singularity at $\xi = 0$ is an artifact of the accelerated motion of Rindler observers and is not seen by freely falling observers. To appreciate that nothing special happens to geodesics at the horizon, consider the affine parameter of a radially propagating photon. This parameter λ may be defined by $d\lambda = dt/\omega$ where dt is time as measured by an observer who perceives the frequency of the photon to be ω. (Then $d\lambda$ is frame-independent, although dt and ω are not, just as the proper time along a trajectory is a frame-independent parameter for a massive freely falling object.)

For geodesics in Rindler spacetime, there is a conservation law

$$k_0 = g_{00}k^0 = E = \text{ constant } .$$

For a null geodesic with $x, y =$ constant in Rindler spacetime, this becomes

$$\xi^2 \frac{d\eta}{d\lambda} = E \text{ and } 0 = g_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{d\lambda} = \xi^2 \left[\left(\frac{d\eta}{d\lambda} \right)^2 - \left(\frac{d\xi_*}{d\lambda} \right)^2 \right] \quad .$$

It is convenient to introduce null coordinates

$$u = \eta - \xi_*, \quad v = \eta + \xi_*$$

in terms of which the metric is

$$ds^2 = e^{v-u} du dv$$

and the affine parameter λ satisfies

$$e^{v-u} \frac{1}{2} \frac{d}{d\lambda}(u + v) = E = \text{ constant}$$

From $\frac{du}{d\lambda}\frac{dv}{d\lambda} = 0$ we see that either $u =$ constant or $v =$ constant (geodesics move either up or down). For $u =$ const, we have

$$d\lambda = \frac{1}{2E} e^{-u} \int e^v dv \Rightarrow \lambda = \text{ const } + \frac{1}{2E} e^{v-u} \text{ or } \lambda = A + Be^v$$

For $v =$ const, we have

$$d\lambda = \frac{1}{2E} e^{v} \int e^{-u} du \Rightarrow \lambda = \text{ const } - \frac{1}{2E} e^{v-u} \text{ or } \lambda = A' - B' e^{-u}.$$

The photons propagating downward ($v =$ const) reach the horizon $u = \infty$ at a finite value of the affine parameter λ. Similarly, photons propagating upward ($u =$ const) emerged from the horizon $v = -\infty$ at a finite value of λ.

It is natural to use affine parameters of null geodesics as new coordinates and so define

$$V = e^v, \quad U = -e^{-u} \Rightarrow ds^2 = dU dV.$$

This is just the flat Minkowski metric

$$ds^2 = dt^2 - dz^2$$

which is easily extended beyond the Rindler horizon.

Now let's apply this lesson to the Schwarzschild geometry

$$ds^2 = \left(1 - \frac{2M}{r}\right)\left(dt^2 - dr_*^2\right) - r^2 d\Omega^2$$

For radial null geodesics, we have

$$k_0 = g_{00} k^0 = \left(1 - \frac{2M}{r}\right) \frac{dt}{d\lambda} = \text{ constant.}$$

Introducing null coordinates

$$u = t - r_*, \quad v = t + r_*,$$

we have

$$e^{(v-u)/4M} = e^{r_*/2M} = \left(\frac{r}{2M} - 1\right) e^{r/2M}$$
$$\Rightarrow ds^2 = \frac{2M}{r} e^{-r/2M} e^{(v-u)/4M} du dv,$$

where we have suppressed the metric on the sphere $r^2 d\Omega^2$.

Near the horizon, $r \simeq 2M$, we have precisely the Rindler metric, up to constant factors. Thus, we transform as before, defining new coordinates

$$U = -e^{-u/4M}, \quad V = e^{v/4M}.$$

These may be interpreted as *affine* parameters of radial null geodesics *at the horizon.* Since

$$dUdV = \frac{1}{16M^2} e^{v-u} du dv,$$

we have

$$ds^2 = \frac{32M^3}{r} e^{-r/2M} dUdV - r^2 d\Omega^2$$

This is the metric for Schwarzschild geometry expressed in terms of *Kruskal coordinates*, which readily admits the *Kruskal extension* across the past and future horizons. The coordinate singularity at $r = 2M$ in the Schwarzschild metric should be contrasted with the singularity at $r = 0$; the latter is a genuine singularity in the curvature tensor, an intrinsic boundary of the spacetime beyond which geodesics cannot be extended.

In the region of Schwarzschild geometry exterior to the horizon, $r > 2M$, we have

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

and so our coordinate transformation may be expressed

$$U = -\left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left(\frac{r-t}{4M}\right),$$
$$V = \left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} \exp\left(\frac{r+t}{4M}\right).$$

These coordinates, cover the region $U < 0$, $V > 0$, which we will refer to as "Region I;" they may be smoothly extended to the other quadrants of U-V plane. Note that it is rather natural to adopt $4M$ as a unit of length, in which case we may write

$$ds^2 = \frac{1}{2r} e^{-2r} dUdV - r^2 d\Omega^2,$$

which looks a little nicer. We may also, if we wish, define T, X by

$$U = T - X, \quad V = T + X,$$

so that

$$ds^2 = \frac{1}{2r} e^{-2r} \left(dT^2 - dX^2\right) - r^2 d\Omega^2;$$

Then radial null geodesics are tilted 45° from vertical in T, X plane. If Schwarzschild t, r are analogous to Rindler η, ξ, then T, X are analogous to Minkowski t, z.

The causal structure of the spacetime may be summarized by a "Kruskal diagram" on the T-X plane. Each point on the diagram represents a 2-sphere; alternatively, we may think of the diagram as representing a $\theta, \phi =$ constant slice through the spacetime.

– IMAGE –

Region I of the diagram is the Schwarzschild geometry exterior to the horizon. In Region I, $U/V = -e^{-t/2M}$, so surfaces of constant Schwarzschild time t are straight lines through the origin, and

$$UV = -\left(\frac{r}{2M} - 1\right) e^{r/2M},$$

so surfaces of constant Schwarzschild radius r are hyperbolas.

Now we can extend these lines and hyperbolas into the other quadrants of the U-V plane, just as we analytically extended the Rindler metric.

Region F $(U > 0, V > 0)$:

$$U = \left(1 - \frac{r}{2M}\right)^{\frac{1}{2}} \exp[(r+t)/(4M)],$$
$$V = \left(1 - \frac{r}{2M}\right)^{\frac{1}{2}} \exp[(r-t)/(4M)].$$

Region II $(U > 0, V < 0)$:

$$U = \left(\tfrac{r}{2M} - 1\right)^{\frac{1}{2}} \exp[(r-t)/(4M)],$$
$$V = -\left(\tfrac{r}{2M} - 1\right)^{\frac{1}{2}} \exp[(r+t)/(4M)].$$

Region P $(U < 0, V < 0)$:

$$U = -\left(1 - \tfrac{r}{2M}\right)^{\frac{1}{2}} \exp[(r+t)/(4M)],$$
$$V = -\left(1 - \tfrac{r}{2M}\right)^{\frac{1}{2}} \exp[(v-t)/(4M)].$$

The logic behind this extension is the same as for Rindler spacetime. Since geodesics move upward through the past horizon at $t = -\infty$ and downward through the future horizon of $t = +\infty$, we need a region "before" $t = -\infty$ and "after" $t = +\infty$ to extend these geodesics. And since there are also upcoming geodesics in Region F and downgoing geodesics in Region P, we learn of the need for Region II to continue these geodesics, even though I and II are causally disconnected.

The new feature, not found in Rindler spacetime, is the singularity at $r = 0$ in Schwarzschild coordinates, or in Kruskal coordinates at

$$UV = 1 \text{ in Regions } F \text{ and } P.$$

This is as far as our analytic extension of Schwarzschild geometry can take us. An observer who falls past the future event horizon inevitably meets this singularity (Note that $r = 0$ is a spacelike surface, for r is a spacelike coordinate inside horizon). Similarly, an observer who enters Region I from region P must have come from the past singularity that bounds that region.

Because of the singularity, the horizon has a more fundamental significance for the Schwarzschild geometry than for the Rindler geometry. All observers agree, regardless of their state of motion, that once the future horizon is passed, it is impossible to escape again to $r = \infty$, and furthermore that the dreaded singularity at $r = 0$ cannot be avoided, bringing any trajectory that crosses the future horizon to an end in a finite amount of proper time and in a region of infinite tidal force.

The constant-r hyperbolas (UV = constant) are invariant under boosts in the T-X plane. This is because

$$UV = T^2 - X^2 = -\left(\frac{r}{2M} - 1\right) e^{r/2M} \quad \text{in Region I,}$$

and similar relations in other regions, define r as a function of $T^2 - X^2$, which is invariant under boosts. All points on the hyperbola have the same proper acceleration relative to a local inertial frame, since proper acceleration is frame-independent. All of the constant t surfaces in Region I are related by boosts, as

$$\frac{T}{X} = \frac{\frac{1}{2}(U+V)}{\frac{1}{2}(V-U)} = \tanh(t/4M),$$

and a boost in the T-X plane takes

$$\tanh(t/4M) \to \tanh(t/4M + \theta).$$

Similarly, the constant-t lines that connect the past singularity in Region P with the future singularity in region F are related by Lorentz boosts (and so all have the same proper length). This length (really a *proper time*) is

$$2\int_0^{2M} \frac{dr}{\sqrt{\frac{2M}{r} - 1}} = 4M \int_0^1 dx \sqrt{\frac{x}{1-x}}$$

Let $x = \sin^2\theta$; then $dx = 2\sin\theta\cos\theta d\theta$ and the integral gives:

$$8M \int_0^{\pi/2} d\theta \ \sin\theta\cos\theta \frac{\sin\theta}{\cos\theta} = 2\pi M$$

Thus πM is the largest possible proper time for falling from the horizon to the singularity.

It may seem somewhat surprising at first that the extended Schwarzschild geometry has *two* spacelike singularities. Geodesics that pass through past horizon cannot be extended indefinitely back in time, just as those that pass through the future horizon cannot be extended indefinitely forward in time. To better appreciate the structure of spacetime, note first of all that, while the Schwarzschild geometry exterior to horizon is static ($\frac{\partial}{\partial t}$ is a timelike killing vector), the interior region is not static, as $\frac{\partial}{\partial t}$ becomes spacelike there. Thus, we may regard the geometry as dynamical, at least inside the horizon. The way we describe this dynamics depends on how we "foliate" the spacetime with spacelike surfaces.

For example, suppose we choose surfaces of constant T, as shown.

– IMAGES –

Then, the "initial" spacetime for sufficiently "early" T, consists of two disconnected asymptotically flat components, each containing a singularity (clothed by a past horizon). Then, at some time, the two singularities join together and smooth out, forming a neck or "wormhole" that connects the two components. The neck widens, reaching a maximal proper radius $r = 2M$, at which point the neck is instantaneously static, and the event horizons of Regions I and II instantaneously join. Then the neck recontracts, eventually pinching off as the two singularities reappear, and the spacetimes disconnect.

As is clear from the Kruskal diagram, the process of wormhole formation and recollapse occurs so rapidly that it is impossible to traverse the wormhole and travel from Region I to Region II . Whoever does try inevitably encounters the dreaded singularity.

Just to emphasize how the situation dynamics depends on the foliation, consider the alternative shown.

– IMAGES –

Here the spacetime comes into being as a spacelike past singularity, flares out, reaches a maximal girth, and recollapses to a spacelike future singularity.

In representing causal structure in a spacetime, it is is often common to use coordinates that parameterize spacelike and temporal infinity by finite values of coordinates. Consider for example the flat spacetime

$$ds^2 = dt^2 - dr^2 - r^2 d\Omega^2.$$

Expressing the metric in terms of new coordinates u', v', we find

$$\begin{aligned} t - r \equiv u &= \tan u'/2 \\ t + r \equiv v &= \tan v'/2 \\ \Rightarrow dt^2 - dr^2 = dudv &= \frac{1}{4}\sec^2\left(\frac{u'}{2}\right)\sec^2\left(\frac{v'}{2}\right)du'dv' \end{aligned}$$

Radial null geodesics are given by $u =$ const or $v =$ const, and hence $u' =$ const or $v' =$ const. But now $u = \pm\infty$ is at $u' = \pm\pi$ and $v = \pm\infty$ is at $v' = \pm\pi$. We may view $u' = \pm\pi$, $v' = \pm\pi$ as defining the boundary of the spacetime. Though this boundary resides at finite values of the u', v' coordinates, the boundary is actually an infinite proper distance away because the metric diverges close to the boundary.

Going further, we may conformally rescale the metric:

$$g_{\mu\nu} \to \Omega^2 g_{\mu\nu}.$$

This transformation deforms the geometry, but it preserves null geodesics, so the rescaled geometry may usefully capture the causal structure of the original geometry. In fact we are free to choose Ω^2 to be a function on spacetime that maps the original boundary to a new boundary which is a finite proper distance from the origin. After this convenient conformal rescaling, the new finite spacetime geometry can be represented by a "conformal diagram" or "Penrose diagram," from which the causal structure can be easily gleaned.

The Penrose diagram for Minkowski space

– IMAGE –

has five asymptotic regions which are labeled in the figure. These are:

i^+: $u = \infty, v = \infty$. This is future timelike ∞, reached if t approaches ∞ faster than r does. All timelike paths arrive at i^+ in the future.

i^-: $u = -\infty, v = -\infty$. This is past timelike ∞, reached if t approaches $-\infty$ faster than $-r$ does. All timelike paths originate at i^- in the past.

i^0: $u = -\infty, v = \infty$. This is spacelike ∞, reached if r approaches ∞ faster than t does. All constant timeslices contain i^0.

$\mathscr{I}^+$: $v = \infty, u =$ finite. This is future null infinity, $r \sim t +$ const. Outgoing null geodesics arrive at $\mathscr{I}^+$ in the future.

$\mathscr{I}^-$: $u = -\infty, v =$ finite. This is past null infinity, $r \sim -t +$ const. Incoming null geodesics originate at $\mathscr{I}^-$ in the past.

Also shown in the Penrose diagram is $r = 0 : u = v$. This is where a radial geodesic "crosses over" from incoming to outgoing.

We may apply a similar transformation to the Kruskal coordinates of the extended Schwarzschild geometry:

$$U = \tan\left(U'/2\right),$$
$$V = \tan\left(V'/2\right).$$

After conformally rescaling, we obtain the Penrose diagram

– IMAGE –

This diagram shows the four regions reminiscent of the regions I, II, F, and P. As explained earlier, the analytically extended geometry actually describes two black holes, each with a past and future event horizon, and a wormhole connecting the black holes behind their horizons. Regions I and II are exteriors of the two black holes, each with its own asymptotic i^-, $\mathscr{I}^-$, i^0, $\mathscr{I}^+$, i^+. Region P is the shared black hole interior behind the past horizons and Region F is the shared interior behind the future horizons.

5.2 Spherically Symmetric Collapse

The Kruskal metric describes an "eternal" static black hole such that the singularity is present in the asymptotic past. It does not describe the "realistic" case of a black hole that forms from a collapsing star — the geometry of a collapsing star has no past horizon or past singularity. (Regions P and II are not part of this geometry.)

It is instructive to describe this geometry using yet another coordinate system. To arrive at this description, recall that the Schwarzschild metric is

$$ds^2 = \left(1 - \frac{2M}{r}\right)(dt + dr_*)(dt - dr_*) + \dots$$

in terms of the tortoise coordinate with $dr_* = dr/(1 - 2M/r)$, and that null coordinates are $u = t - r_*$, $v = t + r*$. An outgoing radial null geodesic ($u =$ constant) is conveniently parametrized by v. Writing $dt - dr_*$ as $dv - 2dr_*$, the metric is

$$ds^2 = \left(1 - \frac{2M}{r}\right) dv^2 - 2dvdr,$$

and we see that the outgoing null geodesic satisfies

$$\left(\frac{dr}{dv}\right)_{\text{outgoing}} = \frac{1}{2}\left(1 - \frac{2M}{r}\right).$$

Similarly, we have

$$ds^2 = \left(1 - \frac{2M}{r}\right) du^2 - 2dudr,$$

so that the incoming null geodesic ($v =$ constant) satisfies

$$\left(\frac{dr}{du}\right)_{\text{incoming}} = -\frac{1}{2}\left(1 - \frac{2M}{r}\right).$$

In these coordinates, a null geodesic seems to get "stuck" at the horizon $r = 2M$ when we follow it going either forward or backward in time.

Suppose we choose instead the so-called "incoming Eddington-Finkelstein time" coordinate

$$\tilde{t} = v - r = t + r_* - r.$$

Then for an incoming geodesic with $v =$ constant we have

$$\left(\frac{dr}{d\tilde{t}}\right)_{\text{incoming}} = -1;$$

the geodesic zooms past the horizon unimpeded. In contrast, for an outgoing geodesic $dt = dr_*$, we have

$$d\tilde{t} = dt + dr_* - dr = 2dr_* - dr = \frac{2 - (1 - 2M/r)}{1 - 2M/r} dr = \frac{1 + 2M/r}{1 - 2M/r} dr,$$

and hence

$$\left(\frac{dr}{d\tilde{t}}\right)_{\text{outgoing}} = -\frac{1-\frac{2M}{r}}{1+\frac{2M}{r}};$$

outgoing geodesics freeze at the event horizon.

– IMAGE –

An Eddington-Finkelstein diagram captures the geometry of spherically symmetric collapse when expressed using this time coordinate. (There is an analytic expression for the metric if we assume for simplicity that the collapsing object implodes at light speed, the Vaidya metric, but we won't need this explicit form.) There is no past horizon at $t = -\infty$, but there are null geodesics that pass through the center of a collapsing star just before the formation of the future horizon and are long delayed close to the horizon before finally escaping. (We assume here that "photons" are influenced only by the geometry, and do not interact with the collapsing matter.) A photon that arrives at the center simultaneously with the formation of the horizon remains frozen in the future at $r = 2M$, unable to pull away.

The photons that eventually emerge after being stuck close to the event horizon for a long time are very strongly redshifted relative to their incoming frequency. This picture will be useful when we investigate the radiation emitted from a collapsed star using quantum field theory. We can apply the field theory in Region I of the Kruskal spacetime, with appropriate boundary conditions on the past horizon that are appropriate for a black hole formed by gravitational collapse.

For the geometry of the collapsing star, as for the static geometry, the causal structure is clearly conveyed by its Penrose diagram.

– IMAGE –

Outgoing radial null geodesics that pass through the origin before the formation of a horizon eventually arrive at outgoing null infinity $\mathscr{I}^+$, while those that pass through the origin too late are trapped inside the horizon and reach the singularity at a finite value of the affine parameter.

5.3 Two "easy" derivations of black hole radiance

5.3.1 From Unruh radiation to Hawking radiation

This argument builds on our observation that a uniformly accelerated observer in the Minkowski-space vacuum perceives a thermal bath. The key idea is that an observer who stays in place at a fixed position r in Schwarzschild coordinates (a Fiducial Observer = FIDO) has a large proper acceleration compared to a Freely Falling Observer (FFO). The FFO is analogous to the inertial observer in Minkowski space; nothing special happens at the horizon, so fluctuations of quantum fields near the horizon as seen by the FFO should resemble field fluctuations in the Minkowski space vacuum as seen by the inertial observer. The FIDO is analogous to the uniformly accelerated Rindler observer; vacuum fluctuations look to the FIDO like thermal fluctuations with a high temperature. Some of these thermal quanta escape to infinity; by computing their redshifted temperature we can determine the temperature of the Hawking radiation detected by an observer who is far away from the black hole.

To make this argument quantitative we need to compute the proper acceleration of the FIDO. For this purpose, we first derive the trajectory followed by an object freely falling in the radial direction. To make the logic clear, it is helpful to consider a more general static and spherically symmetric metric

$$ds^2 = e^{2\Phi}dt^2 - e^{2\Lambda}dr^2 - r^2 d\Omega^2$$

where Φ and Λ are functions of r only. Because $\partial/\partial t$ is a timelike Killing vector, we have the conservation law

$$\text{constant} = \tilde{E} = p_0/m = g_{00}\frac{dt}{d\tau} = e^{2\Phi}\frac{dt}{d\tau}.$$

In addition, the invariant mass of the falling object is

$$m^2 = g_{\mu\nu}p^\mu p^\nu = m^2\left(e^{-2\Phi}\tilde{E}^2 - e^{2\Lambda}\left(\frac{dr}{d\tau}\right)^2\right) \Rightarrow u^2 = e^{2\Lambda}\left(\frac{dr}{d\tau}\right)^2 = e^{-2\Phi}\tilde{E}^2 - 1,$$

where u is the velocity of the falling object, expressed in terms of proper distance traveled per unit of the object's proper time.

Differentiating with respect to the proper time τ we obtain

$$2u\frac{du}{d\tau} = \left(\frac{dr}{d\tau}\right)\tilde{E}^2\frac{d}{dr}\left(e^{-2\Phi}\right),$$

which, using $u = e^{\Lambda}dr/d\tau$, yields

$$\frac{du}{d\tau} = \frac{1}{2}\tilde{E}^2 e^{-\Lambda}\frac{d}{dr}\left(e^{-2\Phi}\right).$$

To find the proper acceleration of the FFO relative to the FIDO (and hence the proper acceleration of the FIDO relative to the FFO), we evaluate this acceleration $du/d\tau$ in the reference frame where the FFO is instantaneously at rest; setting $u = 0$ we find $\tilde{E}^2 = e^{2\Phi}$. We therefore find the proper acceleration to be

$$a_{\text{proper}} = \left(\frac{du}{d\tau}\right)_{u=0} = \frac{1}{2}e^{-\Lambda}e^{2\Phi}\frac{d}{dr}\left(e^{-2\Phi}\right).$$

In the case of the Schwarzschild metric,

$$e^{2\Phi} = 1 - \frac{2M}{r}, \quad e^{2\Lambda} = \left(1 - \frac{2M}{r}\right)^{-1},$$

this yields

$$a_{\text{proper}}(r) = \left(1 - \frac{2M}{r}\right)^{-1/2} \frac{M}{r^2} \approx \frac{1}{4M} \left(1 - \frac{2M}{r}\right)^{-1/2}$$

close to the horizon ($r \approx 2M$). Therefore, a FIDO close to the horizon observes a thermal bath with temperature

$$T(r) = \frac{1}{2\pi} a_{\text{proper}}(r) \approx \frac{1}{8\pi M} \left(1 - \frac{2M}{r}\right)^{-1/2}$$

When this radiation propagates to $r = \infty$, it is redshifted by the factor

$$\left[\frac{g_{00}(r)}{g_{00}(\infty)}\right]^{\frac{1}{2}} = \left(1 - \frac{2M}{r}\right)^{1/2};$$

thus, a distant observer sees radiation with temperature

$$T(\infty) = \frac{1}{8\pi M}.$$

5.3.2 Periodicity in imaginary time

This alternative derivation leverages our observation that field fluctuations at temperature T, when analytically continued to imaginary time, are periodic with period $\beta = 1/T$. The transformation from Schwarzschild to Kruskal coordinates, in all four quadrants, is *periodic* in Schwarzschild time t with period $8\pi M i$. For example, in Region I:

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

thus U and V are unchanged if $2\pi i$ is added to or subtracted from $t/4M$. Now consider a "state" of the quantum fields such that the two-point correlation function $\langle \phi(x)\phi(y) \rangle$ is analytic on the "Euclidean section" (continuation to imaginary time) of Schwarzschild (Kruskal) geometry, and matches the Euclidean Kruskal Klein-Gordon Green function $G_E(x, y)$ there. (Euclidean Schwarzschild has no singularity, as we will discuss later)

This Kruskal Green function is also a Schwarzschild Green function, but periodic in $\tau = it$ with a period $\beta = 8\pi M$. Thus, field correlations measured by an observer are thermal, with temperature

$$T = \frac{1}{8\pi M}.$$

Here temperature, which has the dimensions of frequency, is expressed in units of Schwarzschild time, and corresponds to a locally measured temperature

$$T(r) = \frac{1}{8\pi M} \left[g_{00}(r)\right]^{-\frac{1}{2}}$$

in agreement with our previous computation. We will revisit this argument and provide further details in §5.9.

Remark: Both arguments suggest a black hole in contact with a thermal bath, rather than a black hole emitting radiation into surrounding empty space. To better appreciate what this "state" is, to see how to describe a black hole radiating into empty space, and to gain a firmer grasp of what is going on, we now proceed to more detailed study of quantum field theory in the black hole background.

5.4 Quantum field theory in Kruskal geometry

To describe the "realistic" case of a black hole that forms by gravitational collapse, we will actually need to do QFT on the background of a collapsing star. But first we will consider the case of an "eternal" black hole — the extended Kruskal geometry — with both past and future horizons.

– IMAGE –

Note that Region I of the Kruskal extension is itself a globally hyperbolic spacetime (though not geodesically complete) — it has Cauchy surfaces, just like the right Rindler wedge. Imagine distorting the Cauchy surface in I, so that it lies close to $H^-, \mathscr{I}^-$ (the past horizon and past null infinity) or $H^+, \mathscr{I}^+$ (the future horizon and future null infinity).

In fact, in the massless case, it is legitimate to regard $H^- \cup \mathscr{I}^-$ or $H^+ \cup \mathscr{I}^+$ to be a Cauchy surface for the equation$\nabla_\mu \nabla^\mu \phi = 0$. In the massive case, we would need to specify data at timelike infinity as well. But in the massless case, any wave packet, when propagated forward in time, eventually reaches $H^+ \cup \mathscr{I}^+$ (rather than i^+). And any wave packet, when propagated backward in time, eventually reaches $H^- \cup \mathscr{I}^-$ (rather than i^-) We will mainly consider the massless case because of this simplifying feature that $H^- \cup \mathscr{I}^-$ or $H^+ \cup \mathscr{I}^+$ may be regarded as Cauchy.

The Klein-Gordon equation

$$\left[\frac{1}{\sqrt{g}}\partial_\mu \sqrt{g} g^{\mu\nu} \partial_\nu + m^2\right] u(x) = 0$$

has solutions that can be expanded in terms of spherical harmonics because of the spherical symmetry of Schwarzchild geometry. Writing

$$u = \sum_{l,m} f_{lm}\left(r_*, t\right) \frac{1}{r} Y_{lm}(\theta, \phi),$$

one finds (Exercise 6.11) that the Klein-Gordon equation above implies

$$\left[\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial r_*^2} + V_l\left(r_*\right)\right] f_l\left(r_*, t\right) = 0,$$

where the "effective potential" is

$$V_l(r_*) = \left(1 - \frac{2M}{r}\right)\left[\frac{l(l+1)}{r^2} + \frac{2M}{r^3} + m^2\right]$$

Recalling that $r_* = r + 2M \ln\left(\frac{r}{2M} - 1\right)$, we have

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

Along null geodesics ($r_* = -t +$ constant), the effective potential turns off very rapidly as the horizon is approached ($r \to 2M \Rightarrow r_* \to -\infty$). The prefactor $e^{-t/2M}$ in V_l arises because a wave approaching the horizon is very strongly blueshifted — when the wavelength is very short and frequency very high, the mass, centrifugal barrier, and curvature perturb the propagating wave very little and it obeys the (1+1)-dimensional wave equation in the r_*, t coordinates. For $r_* \to \infty$, the effective potential is

$$V_l\left(r_*\right) \sim m^2\left(1 - \frac{2M}{r}\right) + \frac{l(l+1)}{r^2} + \cdots$$

Thus in the massless case $V_l \to 0$ for both $r_* \to -\infty$ and $r_* \to \infty$.

– IMAGE –

Consider now the waves of definite Schwarzschild frequency

$$f\left(r_*, t\right) = e^{\pm i\omega t} f_{l,\omega}\left(r_*\right)$$
$$\Rightarrow \left[-\frac{\partial^2}{\partial r_*^2} - \omega^2 + V_l\left(r_*\right)\right] f_{l,\omega}\left(r_*\right) = 0$$

For waves with specified angular momentum l, the Klein-Gordon equation has just the form of a 1-dimensional Schrödinger equation. By solving it, we can relate partial waves that are incoming from $r_* = -\infty$ (the horizon) or $r_* = \infty$ to outgoing waves that reach $r_* = -\infty$ or $r_* = \infty$ after being transmitted or reflected by the effective potential.

In the $r_* \to -\infty$ region, the solution has the general form

$$f_\omega\left(r_*\right) = A_{\text{in}}^{(-)} e^{i\omega r_*} + A_{\text{out}}^{(-)} e^{-i\omega r_*}$$

and for $r_* \to \infty$ the solution is

$$f_\omega\left(r_*\right) = A_{\text{out}}^{(+)} e^{i\omega r_*} + A_{\text{in}}^{(+)} e^{-i\omega r_*}.$$

Since the equation is linear, the $r_* \to +\infty$ solution depends linearly on the $r_* \to -\infty$ solution. We may define an "S-matrix" for each partial wave l.

$$\begin{pmatrix} A_{\text{out}}^{(+)} \\ A_{\text{out}}^{(-)} \end{pmatrix} = S \begin{pmatrix} A_{\text{in}}^{(-)} \\ A_{\text{in}}^{(+)} \end{pmatrix}.$$

This has been defined so that S is the identity matrix for a trivial potential ($V = 0$). In that case, a wave escaping from the horizon with amplitude $A_{\text{in}}^{(-)}$ evolves without being disturbed to become a wave propagating outward with amplitude $A_{\text{out}}^{(+)}$, and a wave approaching the black hole with amplitude $A_{\text{in}}^{(+)}$ evolves without being disturbed to become a wave propagating toward the horizon with amplitude $A_{\text{out}}^{(-)}$.

Because the "probability flux" is conserved,

$$\frac{d}{dr_*}\left(f^* \frac{d}{dr_*} f - \left(\frac{d}{dr_*} f^*\right) f\right) = 0,$$

we see that

$$\left|A_{\text{in}}^{(-)}\right|^2 - \left|A_{\text{out}}^{(-)}\right|^2 = \left|A_{\text{out}}^{(+)}\right|^2 - \left|A_{\text{in}}^{(+)}\right|^2$$
$$\Rightarrow \left|A_{\text{in}}^{(-)}\right|^2 + \left|A_{\text{in}}^{(+)}\right|^2 = \left|A_{\text{out}}^{(+)}\right|^2 + \left|A_{\text{out}}^{(-)}\right|^2.$$

This implies S is a unitary matrix.

A further property of S follows from the "time-reversal invariance" of the KG equation. If $f(r_*)$ solves the above equation, then so does $f(r_*)^*$. Therefore we still have a solution after the replacements

$$\begin{aligned} A_{\text{in}}^{(-)} &\to A_{\text{out}}^{(-)*} \\ A_{\text{out}}^{(-)} &\to A_{\text{in}}^{(-)*} \\ A_{\text{out}}^{(+)} &\to A_{\text{in}}^{(+)*} \\ A_{\text{in}}^{(+)} &\to A_{\text{out}}^{(+)*} \end{aligned}$$

and hence

$$\begin{pmatrix} A_{\text{in}}^{(+)*} \\ A_{\text{in}}^{(-)*} \end{pmatrix} = S \begin{pmatrix} A_{\text{out}}^{(-)*} \\ A_{\text{out}}^{(+)*} \end{pmatrix}$$

or

$$\begin{pmatrix} A_{\text{out}}^{(-)} \\ A_{\text{out}}^{(+)} \end{pmatrix} = (S^*)^{-1} \begin{pmatrix} A_{\text{in}}^{(+)} \\ A_{\text{in}}^{(-)} \end{pmatrix}.$$

Since $S^{-1} = S^\dagger$, this is equivalent to $S^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} S \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, or

$$\begin{pmatrix} S_{11} & S_{21} \\ S_{12} & S_{22} \end{pmatrix} = \begin{pmatrix} S_{22} & S_{21} \\ S_{12} & S_{11} \end{pmatrix} \Rightarrow S_{11} = S_{22}.$$

For waves propagating outward from the horizon, we may define transmission and reflection coefficients by

$$\begin{array}{ll} A_{\rm in}^{(-)} = 1, & \Rightarrow \quad A_{\rm out}^{(+)} = t. \\ A_{\rm in}^{(+)} = 0, & \quad A_{\rm out}^{(-)} = r. \end{array}$$

Then

$$\begin{pmatrix} t \\ r \end{pmatrix} = S \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow \quad \begin{array}{l} S_{11} = t, \\ S_{21} = r, \end{array}$$

and thus time reversal implies $S_{22} = t$, so we have

$$S = \begin{pmatrix} t & S_{12} \\ r & t \end{pmatrix},$$

and unitarity implies $S_{12} = -\frac{t}{t^*} r^*$

$$\implies S = \begin{pmatrix} t & -\left(\frac{t}{t^*}\right) r^* \\ r & t \end{pmatrix}.$$

Having completely characterized the S-matrix, we can how predict what happens to waves that propatage inward from spatial infinity.

$$\begin{array}{l} A_{\rm in}^{(-)} = 0 \\ A_{\rm in}^{(+)} = 1 \end{array} \Rightarrow \begin{pmatrix} A_{\rm out}^{(+)} \\ A_{\rm out}^{(-)} \end{pmatrix} = S \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} S_{12} \\ S_{22} \end{pmatrix} = \left(\frac{t}{t^*}\right) \begin{pmatrix} -r^* \\ t^* \end{pmatrix}.$$

We can change the phase convention for the incoming wave so that $A_{\rm in}^{(+)} = (t^*/t)$, and then reflected and transmitted waves are given by

$$\begin{array}{l} A_{\rm out}^{(-)} = t^* \\ A_{\rm out}^{(+)} = -r^* \end{array}$$

We will use this property shortly.

Let's consider further the idea that $\mathscr{I}^- \cup H^-$ can be regarded as a Cauchy surface for massless KG equation. Consider first the simpler case of flat space. Then for

$$u(x) = f_{lm}(r,t) \frac{1}{r} Y_{lm}(\theta, \phi),$$

we have

$$\partial^\mu \partial_\mu u = 0 \Rightarrow \left[\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial r^2} + V_l(r) \right] f_{lm}(r,t) = 0$$

where $V_l(r) = \frac{l(l+1)}{r^2}$ (this is the $M \to 0$ limit of the Schwarzschild wave equation). For e.g. the $l = 0$ wave, the solution is

$$f_{00}(r,t) = g(u) + h(v)$$

where $u = t - r$, $v = t + r$.

The boundary condition $u =$ finite at $r = 0$ becomes $f = 0$ at $r = 0$ or

$$g(t) + h(t) = 0 \Rightarrow f_0(r,t) = g(u) - g(v)$$

Thus, for the $l = 0$ partial wave (and in fact for all partial waves), specifying f as a function of v on $\mathscr{I}^-$ ($u = -\infty$) completely determines f. In particular, there is no need to specify a "normal derivative;" the "normal" actually lies on the same null slice. If we pull this null surface inward to a finite value of u it will no longer be Cauchy, but it becomes Cauchy if augmented by a constant-v slice (the dotted line in the figure). With this choice, specifying the v dependence of the incoming wave on a slice of constant u, and well as the u dependence of the outgoing wave on the corresponding slice of constant v, determines the solution completely.

– IMAGE –

Now recall that we can build on our understanding of the solutions to the KG equation for the purpose of constructing the free scalar field theory on a background geometry, except now we want to consider defining the massless theory making use of solutions determined by initial data on null surfaces. Before we constructed the theory's Hilbert space using the KG inner product defined on a spacelike slice:

$$(f, g) = i \int_\Sigma d^3x \sqrt{h} n^\mu \left(f^* \partial_\mu g - \partial_\mu f^* g \right),$$

where n^μ is the timelike vector orthogonal to Σ and normalized by $n^\mu n_\mu = 1$. In flat space, consider how the KG inner product behaves as the surface Σ on which it is evaluated tips toward the light cone: As the surface Σ tips toward null, n^μ tends toward Σ, and when Σ becomes null, n^μ blows up (due to the normalization $n^\mu n_\mu = 1$), but the divergence is canceled by the zero of $\sqrt{h}$ so that the integral remains well defined. Expanding in terms of spherical harmonics

$$f = \sum_{l,m} f_{lm}(r,t) \frac{1}{r} Y_{lm}(\theta, \phi),$$

$$g = \sum_{l,m} g_{lm}(r,t) \frac{1}{r} Y_{lm}(\theta, \phi),$$

we can evaluate the inner product on the $u = -\infty$ null surface, obtaining

$$(f, g) = \sum_{l,m} \quad i \int dv \left(f^*_{l,m} \partial_v g_{l,m} - \partial_v f^*_{l,m} g_{l,m} \right).$$

– IMAGE –

We can, in fact, distort $\mathscr{I}^-$ through the sequence of "Cauchy" null surfaces (e.g. the dotted line in the figure) that eventually ends at $\mathscr{I}^+$. Any of these slices can be used to construct a basis for the positive frequency solutions to the KG equation. A particular choice would be a basis of solutions $f_{lm} \sim e^{-i\omega v}$ defined on $\mathscr{I}^-$ or a basis $f_{lm} \sim e^{-i\omega u}$ defined on $\mathscr{I}^+$. In other words, the basis for the solutions is

$$\{\text{positive frequency solutions}\} \simeq \{\text{incoming on } \mathscr{I}^-\} \simeq \{\text{outgoing on } \mathscr{I}^+\}.$$

For Region I of the Kruskal extension, though, specifying incoming waves on $\mathscr{I}^-$ does not suffice; initial data must also be specified on the past horizon H^-. Instead the basis for solutions has the form

$$\begin{aligned}\{\text{positive frequency solutions}\} &\simeq \{\text{incoming on } \mathscr{I}^-\} \oplus \{\text{incoming on } H^-\}, \\ &\simeq \{\text{outgoing on } \mathscr{I}^+\} \oplus \{\text{outgoing on } H^+\},\end{aligned}$$

and we can construct a corresponding tensor product of two Fock spaces

$$\mathcal{H} \simeq \mathcal{H}_{\mathscr{I}^-} \otimes \mathcal{H}_{H^-}$$

or

$$\mathcal{H} = \mathcal{H}_{\mathscr{I}^+} \otimes \mathcal{H}_{H^+}.$$

While there is a natural way to choose a basis of solutions which have positive KG norm on $\mathscr{I}^-$ ($\sim e^{-i\omega v}$) or $\mathscr{I}^+$ ($\sim e^{-i\omega u}$), it is not completely obvious how to choose a basis on H^- or H^+ in that the Killing vector $\frac{\partial}{\partial t}$ of the Schwarzschild geometry, which is timelike in Region I, becomes null at the horizon. Arguably, defining positive frequency with respect to U, V at the horizon is more more sensible than defining it with respect to u, v, at least from the viewpoint of FFOs, since, as we have seen U, V are the affine parameters of the radial null geodesics.

One might argue that, from the perspective of an observer in Region I, the ambiguity in the construction of $\mathcal{H}_{H^+}$ is of little consequence. The quanta in the Fock space $\mathcal{H}_{H^+}$ are quanta that cross the future horizon into Region F (and meet the singularity), thus escaping detection in Region I. For the purpose of describing measurements of observables localized in Region I, we may construct a density matrix by tracing out the $\mathcal{H}_{H^+}$ degrees of freedom, and this density matrix does not depend on the $\mathcal{H}_{H^+}$ basis. So there is no need to resolve the ambiguity in the construction of $\mathcal{H}_{H^+}$.

On the other hand, how we define the vacuum state of the $\mathcal{H}_{H^-}$ Fock space *does* affect what a Region I observer sees. To describe what she sees, we must resolve the ambiguity and pick out a particular $\mathcal{H}_{H^-}$ state. If our goal is to describe the physics outside a black hole that formed in gravitational collapse, rather than the the physics outside the eternal black hole described by the extended Kruskal geometry, we must understand what state is preferred in the gravitational collapse scenario, where there is a future event horizon but no past event horizon.

5.5 Hawking radiation emitted by a collapsing star

5.5.1 Bogoliubov transformation

In the case of a black hole formed from collapse, there is no horizon (and no singularity) in the indefinite past. Instead, all radial null geodesics, continued back in time, arrive at $\mathscr{I}^-$. Thus, we may regard $\mathscr{I}^-$ as a Cauchy surface for massless KG eqn. At $\mathscr{I}^-$, the effective potential can be neglected, and a basis for solutions is

$$u_{\omega lm} = \frac{1}{\sqrt{4\pi\omega}} e^{-i\omega v} \frac{1}{r} Y_{lm}(\theta, \phi)$$

and $u^*_{\omega lm}$. These have inner product

$$\begin{aligned} (u_{\omega lm}, u_{\omega l'm'}) &= \delta_{ll'}\delta_{mm'}\delta\left(\omega - \omega'\right), \\ (w, w^*), &= 0 \end{aligned}$$

This basis is chosen to have definite frequency with respect to the Schwarzschild Killing vector $\frac{\partial}{\partial t}$. The corresponding vacuum state $|0_{\rm in}\rangle$ such that

$$a_{\omega lm} |0_{\rm in}\rangle = 0$$

is the state such that no quanta are incoming from $\mathscr{I}^-$ toward the collapsing object.

We want to understand what outgoing quanta an observer sees after the star collapses to form a black hole — that is, what quanta arrive at $\mathscr{I}^+$? For this purpose we are to expand the Fock state $|0_{\rm in}\rangle$ in terms of the Fock space

$$\mathcal{H}_{H^+} \otimes \mathcal{H}_{\mathscr{I}^+}.$$

Then by tracing over the degrees of freedom that propagate to the future horizon, we can obtain a density matrix that describes the properties of the outgoing radiation that arrives at $\mathscr{I}^+$.

The geometry of the collapsing star is certainly not static, so we have reason to expect particle production to occur. Of course, outside the surface of the star, by Birkhoff's theorem, the geometry is spherically symmetric collapsing star *is* static. A wave encounters time-dependent geometry only as it propagates *through* the collapsing star. This propagation through the star scatters an incoming wave in a manner that does not preserve its frequency, and in particular mixes up waves that are positive and negative frequency with respect to $\frac{\partial}{\partial t}$.

In considering propagation through the star, it is actually much more convenient to use an alternative basis — (radial) wave packet states that are localized in u. That way, we can consider waves to arrive at the center $r = 0$ of the collapsing star at different times. We construct the basis by superposing periodic waveforms with frequencies in a narrow band of width E, and ranging from jE to $(j + t)E$, with j a non-negative integer.

– IMAGES –

For each partial wave, then, in place of

$$u_\omega \sim \frac{1}{(4\pi\omega)^{1/2}} e^{-i\omega v}.$$

we have

$$u_{jn}^{(E)} = \frac{1}{\sqrt{E}} \int_{jE}^{(j+1)E} e^{2\pi i n\omega/E} u_\omega d\omega.$$

The wave packet has width $\Delta v = 2\pi/E$ and is centered where the phase is stationary:

$$v = \frac{2\pi}{E} n$$

This wave packet basis is obviously complete and orthonormal. We choose E to be small so the wave is nearly monochromatic.

Now we need to propagate these positive frequency wave packets through the collapsing body and out to $\mathscr{I}^+$, where they can be Fourier analyzed with respect to the basis

$$\frac{1}{(4\pi\omega)^{\frac{1}{2}}} \begin{cases} e^{-i\omega u} \\ e^{+i\omega u} \end{cases}$$

for positive and negative frequency (with respect to $\frac{\partial}{\partial t}$) solutions of the KG equation at $\mathscr{I}^+$. This sounds hard, and it seems that the results should depend on the details in the collapse — that is, on the time-dependent radial profile of the spherically-symmetric collapsing star.

This would indeed be the case were it not for the following crucial insight due to Hawking. It is indeed true that wave packets that propagate through the collapsing star long before the formation of the horizon are perturbed in a way that depends on the details of the collapse. But such wave packets have nothing to do with the radiance of the black hole long after it forms.

– IMAGE –

For this purpose, we are instead interested in field modes that reached the centre of the star just prior to the formation of the horizon, because these modes are delayed arbitrarily long (in Schwarzschild time) near the horizon. They are wave packets that propagate away from the vicinity of the horizon at late times

Furthermore, these long-delayed modes experience a corresponding large red shift. So a mode of given frequency ω at $\mathscr{I}^+$ has an enormously blueshifted frequency at $\mathscr{I}^-$, if it emerges from the vicinity of the horizon very late. For this reason, propagation from $\mathscr{I}^-$ through the star and out to vicinity of the horizon can be treated with very good accuracy in a "geometrical optics" approximation. In this approximation, the details of the collapse do not matter. Here "geometrical optics" (also called the WKB approximation) means that the wavelength is very short compared to the characteristic length scale of the potential.

Then light waves can be treated as rays. The solution to KG is $u \sim e^{i\theta}$ where the constant-θ surfaces are null surfaces (the rays propagate like particles at $v = c$). We are interested in waves very close to the radial geodesic (by convention, $v = 0$ at $\mathscr{I}^-$) that gets caught at the horizon. We want to know the relation between incoming wave $\sim e^{-i\omega v}$ and outgoing wave that has just passed through the collapsing star and is still near the horizon.

– IMAGE –

Consider the null geodesic $v = 0$ and another nearby one at $v < 0$. These geodesics are both surfaces of constant phase of an incoming wave $f(v)$. Curvature may cause these null geodesics to deviate from one another but if the geodesics are sufficiently close the deviation is linear. Geodesic deviation can be described as follows: Consider null surfaces that cross the incoming rays. A vector in this surface points from the $v = 0$ geodesic to the neighboring one. Because the geodesic deviation is linear, the affine parameter length of this deviation (a null vector) changes linearly as a function of the affine parameter length of the incoming (v = constant) geodesics.

In geometric optics, the wave f is just a function of the incoming null geodesics; its argument is hence a linear function of the affine parameter separation between the geodesics. So a wave $f(v)$ on $\mathscr{I}^-$ becomes $f(a\lambda + b)$ as it propagates in.

– IMAGE –

But after the horizon has formed, as the wave is pulling away from the horizon and is still close to the horizon, the affine parameter along the deviation vector is the Kruskal null coordinate U; hence we have

$$\text{wave } \sim f(aU + b);$$

that is, the wave $e^{-i\omega v}$ incoming from $\mathscr{I}^-$ evolves as $\sim e^{-i\omega aU}$ as it propagates outward from the horizon.

– IMAGE –

We conclude that, for the purpose of describing the emission at late times from the black hole, an incoming wave that is positive frequency at $\mathscr{I}^-$ with respect to $\frac{\partial}{\partial t}$, and propagates through the collapsing star, is equivalent to a wave incoming from the past horizon on Region I of Kruskal geometry that is positive frequency with respect to $\frac{\partial}{\partial U}$. This is Unruh's way of rephrasing the central conceptual point in Hawking's argument [24]. With it in hand, it is now reasonably straightforward to proceed with the theory of black hole radiance by computing the density matrix of the outgoing radiation. Using the appropriate positive frequency bases on H^- and $\mathscr{I}^-$, we define a vacuum state $|0_{\rm in}\rangle$ by

$$a^{H^-}|0_{\rm in}\rangle = a^{\mathscr{I}^-}|0_{\rm in}\rangle = 0,$$

and calculate the Bogoliubov coefficients to express $|0_{\rm in}\rangle$ in terms of Fock basis of

$$\mathcal{H}_{H^+} \otimes \mathcal{H}_{\mathscr{I}^+}.$$

Finally, we trace out the $\mathcal{H}_{H^+}$ degrees of freedom to obtain a density matrix that describes outgoing radiation detectable at $\mathscr{I}^+$. For this computation, it makes no difference how we choose the basis for the Fock space $\mathcal{H}_{H^+}$, but we can make the calculation easier and the interpretation of the result cleaner by making a sensible choice of basis.

In the description below, I will drop the l, m indices on the solutions and creation and annihilation operators. It will be understood that we can carry out a similar analysis for each partial wave. We will use the obvious bases for the solutions to the KG equation, positive frequency with respect to $\frac{\partial}{\partial t}$, on $\mathscr{I}^\pm$

$$u_\omega^{\rm in} \sim e^{-i\omega v} \text{ on } \mathscr{I}^-,$$
$$u_\omega^{\rm out} \sim e^{-i\omega u} \text{ on } \mathscr{I}^+.$$

On H^-, in accord with Unruh's observation, we choose a basis of solutions that are positive frequency wrt $\frac{\partial}{\partial U}$: $e^{-i\omega' U}$. On H^+, where we are free to make any convenient choice, we may choose our basis set to be positive frequency with respect to $\frac{\partial}{\partial t}$: $e^{-i\omega v}$.

To save work, let's make this calculation as similar as possible to the calculation in Rindler spacetime from §4.6. In the Kruskal geometry, the past horizon in Region I, when extended, becomes the past horizon of Region II (recalling that Schwarzschild time runs backwards in Region II).

– IMAGE –

Solutions that are positive frequency with respect to Schwarzschild time on this past horizon in Regions I and II are

$$u_\omega^{I,H^-} \sim e^{-i\omega u}\theta(-U) \quad \text{ on } H^-,$$
$$u_\omega^{II,H^-} \sim e^{i\omega u}\theta(U) \quad \text{ on } H^-;$$

note that $e^{i\omega u}$ is positive frequency with respect to $\frac{\partial}{\partial t}$ in Region II because Schwarzschild time runs backwards there. We can construct a linear combination of these solutions that is positive frequency with respect to Kruskal time, by demanding that this combination is analytic as a function of U in the lower half plane. The functions we seek are

$$u_{1,\omega}^{\rm Kruskal} = N_\omega \left(u_\omega^{I,H^-} + e^{-4\pi M\omega} (u_\omega^{II,H^-})^* \right),$$
$$u_{2,\omega}^{\rm Kruskal} = N_\omega \left(u_\omega^{II,H^-} + e^{-4\pi M_\omega} (u_\omega^{I,H^-})^* \right),$$

where $\quad N_\omega = \left(1 - e^{-8\pi M\omega}\right)^{-\frac{1}{2}}$. To verify analyticity, recall that

$$\begin{aligned} u &= -4M\ln(-U) \qquad \text{in Region I ,} \\ u &= 4M\ln(U) \qquad \text{in Region II .} \end{aligned}$$

Following the same logic as in §4.6, we continue around the cut of the logarithm at $U = 0$ by giving U a small negative imaginary part. This means that when we analytically continue from negative U (Region I) to positive U (Region II), u picks up an additive term $-4M(i\pi)$, which maps $e^{-i\omega u}$ to $e^{-4\pi M\omega}(e^{-i\omega u})^*$.

Our basis for the "in" solutions to the KG equation, then, is $\{u^{\text{Kruskal}}_{1,\omega}, u^{\text{Kruskal}}_{2,\omega}, u^{\mathscr{I}^-}_\omega\}$, and the corresponding "in" vacuum satisfies

$$0 = a^{\text{Kruskal}}_{1,\omega}|0_{\text{in}}\rangle = a^{\text{Kruskal}}_{2,\omega}|0_{\text{in}}\rangle = a^{\mathscr{I}^-}_\omega|0_{\text{in}}\rangle.$$

We are imposing a condition on the state at H^- in Region II, but there is no harm in doing so because this condition has no effect on what is seen in Region I. For our "out" solutions, we will use $\{u^{\mathscr{I}^+}_\omega, u^{H^+}_\omega, u^{II,H^-}_\omega\}$. (If no superscript I or II is shown, Region I is understood.) The last two are arbitrary choices, which do not effect the density matrix of the emitted Hawking radiation.

Now, to derive the Bogoliubov transformation, we must take into account the scattering off the potential barrier. Defining transmission and reflection coefficients by

$$u^{I,H^-}_\omega = t_\omega u^{\mathscr{I}^+}_\omega + r_\omega u^{H^+}_\omega,$$

unitarity implies

$$|t_\omega|^2 + |r_\omega|^2 = 1,$$

and as explained in §5.4 by time reversal we have

$$u^{\mathscr{I}^-}_\omega = -r^*_\omega u^{\mathscr{I}^+}_\omega + t^*_\omega u^{H^+}_\omega$$

(up to an unimportant overall phase). Now we can express

$$u^{\text{in}}_\omega = \begin{pmatrix} u^{\text{Kruskal}}_{1,\omega} \\ u^{\text{Kruskal}}_{2,\omega} \\ u^{\mathscr{I}^-}_\omega \end{pmatrix} \text{ in terms of } u^{\text{out}}_\omega = \begin{pmatrix} u^{\mathscr{I}^+}_\omega \\ u^{H^+}_\omega \\ u^{II,H^-}_\omega \end{pmatrix}.$$

We have

$$\begin{aligned} u^{\text{Kruskal}}_{1,\omega} &= N_\omega\left[t_\omega u^{\mathscr{I}^+}_\omega + r_\omega u^{H^+}_\omega + e^{-4\pi M\omega}(u^{II,H^-}_\omega)^*\right] \\ u^{\text{Kruskal}}_{2,\omega} &= N_\omega\left[(u^{II,H^-}_\omega + e^{-4\pi M\omega}\left([t^*_\omega(u^{\mathscr{I}^+}_\omega)^* + r^*_\omega(u^{H^+}_\omega)^*\right)\right] \\ u^{\mathscr{I}^-}_\omega &= \left(-r^*_\omega u^{\mathscr{I}^+}_\omega + t^*_\omega u^{H^+}_\omega\right), \end{aligned}$$

which implies

$$u_\omega^{\rm in} = \alpha_\omega u_\omega^{\rm out} + \beta_\omega u_\omega^{out*}$$

where

$$\alpha_\omega = \begin{bmatrix} N_\omega t_\omega & N_\omega r_\omega & 0 \\ 0 & 0 & N_\omega \\ -r_\omega^* & t_\omega^* & 0 \end{bmatrix},$$
$$\beta_\omega = \begin{bmatrix} 0 & 0 & N_\omega e^{-4\pi M\omega} \\ N_\omega e^{-4\pi M\omega} t_\omega^* & N_\omega e^{-4\pi M\omega} r_\omega^* & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and therefore

$$\alpha_\omega^{-1} = \begin{bmatrix} N_\omega^{-1} t_\omega^* & 0 & r_\omega \\ N_\omega^{-1} r_\omega^* & 0 & t_\omega \\ 0 & N_\omega^{-1} & 0 \end{bmatrix},$$
$$\left(\alpha^{-1}\beta\right)^* = \begin{bmatrix} 0 & 0 & e^{-4\pi M\omega} t_\omega \\ 0 & 0 & e^{-4\pi M\omega} r_\omega \\ e^{-4\pi M\omega} t_\omega & e^{-4\pi M\omega} r_\omega & 0 \end{bmatrix}.$$

Note that $\alpha^{-1}\beta$ is a symmetric matrix, in accord with general arguments in §3.6.

Drawing on the formulas in §3.6,

$$\begin{pmatrix} u' \\ u'^* \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \begin{pmatrix} u \\ u^* \end{pmatrix}$$

implies

$$\left|0'\right\rangle = {\rm phase} |\det\alpha|^{-\frac{1}{2}} \exp\left[\frac{1}{2} a^\dagger \left(\alpha^{-1}\beta\right)^* a^\dagger\right] |0\rangle$$

from which we obtain

$$\begin{aligned} |0_{\rm in}\rangle &= \prod_\omega N_\omega^{-1} \exp\left(e^{-4\pi M\omega}\left(a_\omega^{II,H^-}\right)^\dagger \left[t_\omega \left(a_\omega^{\mathscr{I}^+}\right)^\dagger + r_\omega \left(a_\omega^{H^+}\right)^\dagger\right]\right) |0_{\rm out}\rangle \\ &= \prod_\omega N_\omega^{-1} \sum_{n_\omega} \frac{1}{\sqrt{n_\omega!}} e^{-4\pi M\omega n_\omega} |n_{\omega,{\rm out},II}\rangle \otimes \left[t_\omega \left(a_\omega^{\mathscr{I}^+}\right)^\dagger + r_\omega \left(a_\omega^{H^+}\right)^\dagger\right] |0_{\rm out,I}\rangle. \end{aligned}$$

Here $\det\alpha_\omega = -N_\omega^2$, and recall there is also a product over l, m which has been suppressed in this formula.

To find the density operator for Region I, we trace out the II, H^- Fock space, obtaining

$$\rho_I = \prod_\omega \frac{1}{n_\omega!}\left(1-e^{-8\pi M\omega}\right)\sum_{n_\omega} e^{-8\pi M\omega n_\omega}$$
$$\left[t_\omega \left(a_\omega^{\mathscr{I}^+}\right)^\dagger + r_\omega \left(a_\omega^{H^+}\right)^\dagger\right]^{n_\omega} |0_{\mathrm{out},I}\rangle\langle 0_{\mathrm{out},I}| \left[t_\omega^* \left(a_\omega^{\mathscr{I}^+}\right) + r_\omega^* \left(a_\omega^{H^+}\right)\right]^{n_\omega}. \tag{5.1}$$

We can proceed to trace over the states in $\mathcal{H}_{H^+}$ to find $\rho_I{}^{\mathscr{I}^+}$, which describes the black hole radiation seen at ∞. But the interpretation of the result is particularly clear in its current form. If $t=1$ and $r=0$, the modes have precisely a thermal distribution. When transmission and reflection are taken into account, we see that each mode emerges from the past horizon thermally populated - with an amplitude t_ω of being transmitted through the barrier to $\mathscr{I}^+$, and amplitude r_ω of being reflected back H^+.

If we take the trace with respect to $\mathcal{H}_{H^+}$ degrees of freedom (Exercise 6.12), we find

$$\rho_I{}^{\mathscr{I}^+} = \prod_\omega \left[\sum_{n_\omega} \left|n_\omega, \mathscr{I}^+\right\rangle P_{n,\omega}^{(\beta)} \left\langle n_\omega, \mathscr{I}^+\right|\right]$$

where

$$P_{n,\omega}^{(\beta)} = \left(1-e^{-\beta\omega}\right)\frac{\left(e^{-\beta\omega} \mid t_\omega \mid^2\right)^n}{\left(1- \mid r_\omega \mid^2 e^{-\beta\omega}\right)^{n+1}}, \quad \beta = 8\pi M.$$

For perfect transmission ($|t_\omega|^2 = 1$), this is a thermal density matrix with temperature $T = \frac{1}{8\pi M}$. By summing a geometric series, you can check that the density matrix is properly normalized even when $|r_\omega|$ is nonzero:

$$\sum_n P_{n,\omega}^{(\beta)} = 1.$$

We may also compute the mean occupation number of each mode at $\mathscr{I}^+$ using formulas from §3.7. There we saw

$$u^{\mathrm{in}} = \alpha u^{\mathrm{out}} + \beta u^{\mathrm{out}*}$$

implies

$$a^{\mathrm{out}} = \alpha^T a^{\mathrm{in}} + \beta^\dagger a^{\mathrm{in}\dagger}$$
$$a^{\mathrm{out}\dagger} = \beta^T a^{\mathrm{in}} + \alpha^\dagger a^{\mathrm{in}\dagger}$$

and therefore

$$\begin{aligned} N_{\omega,i} &= \left\langle 0_{\mathrm{in}} \left| a_{\omega,i}^{\mathrm{out}\dagger} a_{\omega,i}^{\mathrm{out}} \right| 0_{\mathrm{in}} \right\rangle \\ &= \sum_{k,j} \beta_{\omega,ij}^\dagger \beta_{\omega,ik}^T \langle 0_{\mathrm{in}}|\, a_{\omega,k}^{\mathrm{in}} a_{\omega,j}^{\mathrm{in}\dagger} |0_{\mathrm{in}}\rangle \\ &= \sum_j \beta_{\omega,ji}^* \beta_{\omega,ji} = \sum_j |\beta_{\omega,ji}|^2. \end{aligned}$$

(The indices i, j label particle species with frequency ω; these labels are reversed here compared to equation (3.1), because in and out have been interchanged.)

From the 1st column of the β_ω matrix above, we find

$$N_{\omega,\mathscr{I}^+} = |t_\omega|^2 \frac{1}{e^{\beta\omega} - 1},$$

the thermal occupation number modified by the transmission probability.

Up until now we have suppressed factors of k, $\hbar$, and c. Restoring these factors by dimensional analysis obtain the temperature of the emitted radiation:

$$kT = \frac{\hbar c^3}{8\pi GM}.$$

This remarkably beautiful formula combines thermodynamics, relativity, and quantum mechanics into a single equation. Numerically, the result expressed in Kelvin is

$$T = 6 \times 10^{-8} \left(\frac{M_\odot}{M} \right) \text{ K}$$

where $M_\odot \cong 2 \times 10^{30}$ kg is a solar mass. A solar-mass black hole has Schwarzschild radius

$$R = \frac{2GM_\odot}{c^2} \approx 3 \text{ km}.$$

The typical wavelength of emitted radiation received at $\mathscr{I}^+$ is somewhat longer than this.

5.5.2 Black hole evaporation

Since no radiation is coming in from $\mathscr{I}^-$, and a steady flux is received at $\mathscr{I}^+$, the black hole is evidently radiating (even long after the formation at the horizon). To compute the radiation flux, we need only integrate the occupation numbers against a density of states factor. To facilitate state counting, imagine putting the black hole in a large spherical cavity with radius $R_c \gg M$. For each l, m, consider the *outgoing* spherical waves $\sim e^{-i\omega t} e^{ikr}$. If walls of the cavity are perfectly reflecting, radial modes are $\sin\left(\frac{\pi n r}{R_c}\right)$ where n is an integer (we can ignore the effective potential if cavity is very large), and the sum over n can be approximated by an integral over the radial wave number $k = \pi n / R_c$: $\sum_n \sim \frac{R_c}{\pi} \int dk$. But only half of these modes are outgoing, so

$$\sum_{\text{outgoing}} \sim \frac{R_c}{2\pi} \int dk$$

If the modes travel at velocity v_k, we obtain outgoing flux $= \sum_{l,m} \int \frac{v_k}{R_c} \frac{R_c}{2\pi} dk N_{\omega lm}$. For massless modes, with $v = c$ and $k = \omega$, by summing m over $2\ell + 1$ values and substituting our computed occupation number for $N_{\omega\ell m}$ yields

$$\frac{\text{number outgoing}}{\text{time}} = \sum_l (2l+1) \int \frac{d\omega}{2\pi} \frac{|t_{\omega l}|^2}{e^{\beta\omega} - 1}.$$

To find the total luminosity of the emitted energy, we weight the modes by the energy ω of the quanta, so

$$L = \frac{\text{energy emitted}}{\text{time}} = \sum_l (2l+1) \int \frac{d\omega}{2\pi} \omega \frac{|t_{\omega l}|^2}{e^{\beta\omega} - 1}. \tag{5.2}$$

To compute the luminosity, we must find the $t_{\omega l}$ by numerically solving the radial KG equation.

Note that the black hole is a rather unconventional black body, in that the typical wavelength of the emitted radiation is comparable to the size of the body. Therefore it is not really so black — a typical thermal quantum incident on the (centrifugal and curvature) barrier surrounding the black hole may reflect off the barrier and hence fail to be absorbed. Correspondingly, the emission of the black hole is suppressed, relative to an ideal black body with the same temperature and surface area. But if the frequency ω of the incoming quantum is high compared to $1/M$, then propagation is well described by geometric optics and the absorption cross section can be computed by considering null geodesics in Schwarzschild geometry.

Geodesics of massless particles in the black hole geometry are described in MTW [20], page 674. The dependence of the radial coordinate r on the affine parameter λ is governed by

$$\left(\frac{d}{d\lambda}\left(\frac{1}{r}\right)\right)^2 = b^{-2} - B^{-2}(r),$$

where $B^{-2}(r) = r^{-2}\left(1 - \frac{2M}{r}\right)$ is an effective potential and b is the impact parameter. B^{-2} is maximized at $r = 3M$, where it attains the value $1/27M^2$. Therefore, if b is larger than $b_{\max} = 3\sqrt{3}M$, the massless particle will bounce off the potential barrier and fail to reach the horizon at $r = 2M$. But if b is smaller than $b_{\max}$, the particle will sail over the barrier and be absorbed by the black hole. Therefore, in the approximation of geometrical optics, the black hole has a capture cross section

$$\sigma_{\text{capture}} = \pi b_{\max}^2 = 27\pi M^2.$$

– IMAGE –

For large ω, then, where geometrical optics applies, the differential luminosity of the black hole should agree with that of an ideal black body with area $4\pi b_{\max}^2 = 4\sigma_{\text{capture}}$ and with inverse temperature $\beta = 8\pi M$, namely

$$\frac{dL}{d\omega} = (\text{Area}) \frac{1}{2\pi^2} \frac{\omega^3}{e^{\beta\omega} - 1} \quad (\omega M \gg 1).$$

But the differential luminosity is suppressed relative to the emission of an ideal black body at smaller frequency ω due to the "gray-body" factors arising from the effective potential.

Although not precisely correct, the ideal black body curve provides qualitative guidance concerning the scaling of the total luminosity L as a function of the black hole mass M:

$$L \sim \text{ (Area) } T^4 \sim M^2 M^{-4} \sim M^{-2}.$$

In contrast to a conventional thermal system, the black hole heats up instead of cooling down as it loses energy — one says it has a *negative specific heat*. In fact, as it sheds mass, the reduction in area is more than compensated by the increase in temperature, so that the total luminosity increases, resulting in an explosion.

Integrating, we find

$$\begin{aligned} \frac{dM}{dt} = -CM^{-2} &\Rightarrow \frac{1}{3}\left(M^3 - M_0^3\right) = -C\,(t - t_0) \\ &\Rightarrow M = \left[M_0^3 - 3C\,(t - t_0)\right]^{\frac{1}{3}}. \end{aligned}$$

The black hole disappears completely in time

$$t_{\text{life}} = M^3/3C.$$

– IMAGE –

To calculate C requires numerical computations (e.g., by Page [21, 22] — see also chapter 8 of *The Membrane Paradigm* [23]). For $T_{BH} \ll M_{\text{electron}}$ (or $M_{BH} \gg 10^{14}$ kilograms) only light species are emitted, which are gravitons, photons and neutrinos. Assuming 3 species of effectively massless neutrinos, Page's calculation gives

$$t_{\text{life}} = 10^{10} \text{ yr } \left(\frac{M}{5 \times 10^{11}\text{kg}}\right)^3$$

"Primordial" black holes in the mass range $M \sim 5 \times 10^{11}$kg would be disappearing today. Radiation emitted by primordial black holes has not been seen.

When the temperature of the radiation is large compared to neutrino masses, the neutrinos actually dominate the emission; in that case t_{life} is longer by a factor of about 7 to 8 compared to the case where only gravitons and photons are emitted.

Expressed in terms of solar mass units, the lifetime is

$$t_{\text{life}} \sim 10^{66}\text{yr}\,(M/M_\odot)^3\,.$$

Solar mass black holes will not evaporate any time soon. Of course, astrophysical black holes have a Hawking temperature which is much lower than the 2.7 K cosmic background radiation, so an isolated solar mass black hole would actually gain mass by accreting cosmic microwave photons more rapidly than it sheds mass by emitting Hawking radiation.

5.5.3 Why Hawking radiation?

Let's consider further *why* a black hole emits a steady flux of quanta. It should be evident that the *horizon* is an absolutely essential ingredient. For example, suppose we consider the collapse to form a (zero-temperature) neutron star — a cold static ball somewhat larger than its Schwarzschild radius. During collapse, the geometry inside the ball is not static, and we would not be surprised to see (scalar) quanta emitted (No gravitons, though, in spherically symmetric collapse according to Birkhoff's theorem). But once the star settles down, there will be no emission at late times. In the black-hole case, there are outgoing field modes that are delayed arbitrarily long by the horizon, and *these* give rise to the emission even at very late times.

If we consider the spacetime of a static neutron star, it is globally static, both outside and inside the star. So $\partial/\partial t$ is a globally timelike Killing vector, unlike the black-holes case, where $\partial/\partial t$ tips over and becomes spacelike at the horizon. So there is a natural vacuum state

$$|0_{\rm in}\rangle = |0_{\rm out}\rangle.$$

That is, the Fock space state with no incoming quanta at $\mathscr{I}^-$ and no outgoing quanta at $\mathscr{I}^+$ are the same state. No particle production occurs in the background of a neutron star (no past horizon provides a "source" of outgoing particles that reach $\mathscr{I}^+$).

We can imagine squeezing the neutron star until its radius reaches its Schwarzschild radius, and it becomes a black hole. In this way, we can, by quasistatically perturbing the quantum field vacuum of the neutron star background, produce a quantum state in the black-hole background that is pure vacuum outside the horizon, with no quanta emerging from the past horizon. Formally, this is the state in Kruskal geometry obtained by using a basis for the KG equation that is positive frequency with respect to Schwarzschild time t both on $\mathscr{I}^-$ and on H^-. It is called the "Boulware vacuum" (pronounced "bowler") [4].

There is clearly something unphysical about this state, the vacuum as seen by an observer with infinite proper acceleration at the surface of the star. If we could *really* squeeze a neutron star as described, it would be unable to support itself as its radius approached the Schwarzschild radius, and would fall through the horizon. The quasistatic approximation would break down, and we would be back to considering a black hole that forms from a collapsing object. We will appreciate better the unphysical nature of the Boulware vacuum later, when we compute in §5.10.6 the (renormalized) energy-momentum tensor and find that it diverges at the horizon even as measured by *freely-falling observers*.

There is another funny thing about our picture of black hole radiation from a collapsed star. The outgoing flux long after collapse is alleged to be due to field modes that spend a long time near the horizon, and become strongly redshifted before escaping. If we take this picture seriously, then outgoing quanta with $\omega \sim T$ must be occupying incoming modes that have truly incredible frequencies on $\mathscr{I}^-$.

– IMAGE

To estimate this effect, consider two photons at radial coordinates r_1, r_2 around the time the horizon forms, with $r_{1,2} \sim 2M$. If on outgoing radial null geodesics, these photons propagate as

$$r_* = t + \text{constant},$$

and arrive at $r \gg 2M$ redshifted by the factor

$$(g_{00})^{\frac{1}{2}} = \left(1 - \frac{2M}{r}\right)^{\frac{1}{2}} \sim e^{r_*/4M}$$

If the initial frequencies close are $\omega_{1,0}, \omega_{2,0}$, then the redshifted frequencies are

$$\frac{\omega_1(\infty)}{\omega_2(\infty)} = \frac{\omega_{1,0} e^{r_{*,1}/4M}}{\omega_{2,0} e^{r_{*,2}/4M}} = \frac{\omega_{1,0}}{\omega_{2,0}} e^{\delta t/4M},$$

where δt is the time interval between receipt of photons 1 and 2 by a FIDO far from the black hole. So if $\omega_1(\infty) \sim \omega_2(\infty)$, then

$$\omega_{2,0} \sim \omega_{1,0} e^{\delta t/4M};$$

the photons that arrive late must have had *exponentially* large frequencies coming in.

These enormous frequencies seem fishy, but it is reassuring to note that in this picture there are no *physical* photons with enormous frequencies. These very high frequency modes at $\mathscr{I}^-$ are *unoccupied*, as the initial state was the vacuum at $\mathscr{I}^-$. The point of our calculation of the Bogoliubov coefficients is that these modes are *not* purely positive frequency with respect to u on $\mathscr{I}^+$ if they *are* positive frequency with respect to v on $\mathscr{I}^-$.

Still, there is something disquieting about considering frequencies that are enormous, even compared to the Planck scale, at $\mathscr{I}^-$. It should not be *necessary* to do this. We ought to be able to impose a *cutoff* on the frequency of the vacuum field fluctuations: $\omega \leq \Lambda$. And this ought not to drastically modify the calculation of black hole radiance as long as Λ is much larger than the temperature of the radiation that arrives at $\mathscr{I}^+$.

The whole idea of *renormalization* in quantum field theory is that, for the purpose of discussing physics at low energies (e.g., emitted quanta with temperature $T \ll M_{\text{Planck}}$) we can "integrate out" very short wavelength field fluctuations, and incorporate the effects of these fluctuations into the parameters of an "effective field theory" with a smaller cutoff; these parameters correspond to the quantities that can be measured *directly* at low-energy ("renormalized" as opposed to "bare" quantities). We speak of "decoupling" of the short-wavelength degrees of freedom, meaning that to good accuracy the effects of the short-wavelength fluctuations can be completely absorbed into such renormalizations, and that the number of resulting renormalized parameters is small and manageable.

This is, after all, what makes physics possible — we would be in trouble if we needed to know all about Planck-scale physics to predict e^+e^- annihilation at 100 GeV. Conversely, we cannot expect to learn much about Planck scale physics by measuring low-energy processes. Likewise, there ought to be an effective field theory description of black hole radiance, in which frequencies exceeding Λ are integrated out and need not be considered.

Something like this is achieved by the "membrane paradigm" viewpoint of Thorne et al. [23]. In effect, they integrate out physics very close to the horizon, and incorporate this short distance physics into a modification of the boundary conditions satisfied on a "stretched horizon" that hovers above the actual horizon. We might choose the (past) stretched horizon so that the local temperature (measured by a FIDO) there is $T_{\rm local} \leq \Lambda$. This membrane then provides an inexhaustible supply of upward propagating quanta occupying the field modes that are positive frequency with respect to u on $\mathscr{I}^+$.

Admittedly, a fully satisfactory "effective field theory" picture of a black hole that forms in gravitational collapse is complicated to formulate in detail if we wish to describe the dynamics of the quantum fields during the collapse. Fortunately, after the (spherically symmetric) collapse, spacetime outside the horizon rapidly approaches the static Schwarzschild geometry. As we have already noted, quantum field theory on this exterior static geometry is easier to contemplate, if we know how to choose a suitable boundary condition at (or close to) the horizon. *The proper boundary condition is dictated by the principle that an FFO who crosses the horizon will see field modes that are unoccupied.* It follows that a FIDO very close to the horizon will see thermal radiation, as in our analysis of Rindler spacetime.

Guided by the Rindler analogy, what should we expect FIDO's and FFO's to see when close to the black hole horizon? The first important point is that the temperature close to the horizon as measured by a FIDO is

$$\begin{aligned} T(r) &= \frac{1}{8\pi M}\left(1-\frac{2M}{r}\right)^{-\frac{1}{2}} \\ &= \frac{1}{2\pi S_{\rm proper}} \end{aligned}$$

where $S_{\rm proper}$ is the proper distance to the horizon (this is just as for a Rindler observer in the Minkowski vacuum). Hence, for this observer the wavelength of the thermal radiation is comparable to the proper distance to the horizon and is not accurately described on that distance scale by geometrical optics. The FIDO very close to the horizon is not being blasted by a flux of radiation emanating from the horizon.

In fact, to an approximation that becomes better and better for a FIDO closer and closer to the horizon, the FIDO sees a (nearly) isotropic thermal bath. This is because nearly all of the radiation is unable to penetrate the effective potential

$$V_l = \left(1-\frac{2M}{r}\right)\left(\frac{l(l+1)}{r^2}+\frac{2M}{r^3}\right);$$

it gets reflected by the potential back to the (future) horizon.

– IMAGE –

We gain a heuristic understanding in this picture by considering the propagation of the outgoing wave in geometrical optics. An outgoing (nonradial) null geodesic manages to escape from the black hole only if directed outward in a narrow "escape cone" that gets narrower and narrower closer and closer to the horizon. As measured by a FIDO, the half-opening angle of this cone is

$$\delta \approx \frac{3\sqrt{3}}{2}\left(1-\frac{2M}{r}\right)^{\frac{1}{2}}$$

for $r \approx 2M$. Only the null geodesics inside this cone can escape. (This is calculated in MTW [20], page 675.) In terms of wave propagation, there are propagating modes for given ω and all l as $r^* \to -\infty$, but almost all are reflected back by the centrifugal barrier, just as almost all l-modes "miss" the black hole when incoming from $r = \infty$ (the time-reversed process).

The reflected quanta rain back down onto the FIDO near the horizon, endowing the black hole with what Thorne calls a "thermal atmosphere" [23]. Thus our analogy with a Rindler observer in the Minkowski vacuum becomes very accurate close to the horizon. And since FIDOs very close to the horizon see a thermal bath, it follows that FFOs crossing the horizon see *vacuum*.

5.5.4 Massive Particles

How is the analysis modified if our scalar field has a mass?

Close to the horizon, because the (FIDO-measured) temperature $T(r) = \frac{1}{8\pi M}\left(1-\frac{2M}{r}\right)^{-\frac{1}{2}}$ of the thermal bath is much larger than the mass, the mass has a negligible effect there. The only difference between massless and massive particles is their ability to surmount the potential barrier.

For a massive scalar field, the effective potential is

$$V_l = \left(1-\frac{2M}{r}\right)\left[\frac{l(l+1)}{r^2}+\frac{2M}{r^3}+m^2\right]$$

It tapers off close to the horizon, where frequencies are highly blueshifted. Far from the horizon it is dominated by the last term:

$$V_l \to m^2 \text{ for } r \to \infty.$$

Modes with $\omega^2 < m^2$ are unable to propagate and are completely reflected. The modes that escape (to i^+) have a (redshifted) thermal distribution, modified by an transmission probability $|t|^2$.

Our analysis for the massless case in §5.4 is slightly modified; frequencies at $r_* \to \pm\infty$ do not match, because the effective potential has a different value for large negative and positive r_*. but momentum conservation and time reversal invariance still hold, and it remains true that outgoing and incoming waves are reflected with the same probability.

5.6 Kerr black hole

5.6.1 Kerr geometry

Now let's consider how our calculation of black hole radiance is modified in the (more realistic) case of a rotating Kerr black hole.

The Kerr black hole is the end result of non-spherical gravitational collapse. When spherical symmetry is not assumed, the exterior geometry of a collapsing object may be complicated (e.g., many gravitational multipoles, etc.). However, once a horizon forms, we may invoke the remarkable black hole uniqueness theorems, which tell us that the only *stationary* black hole solutions to the (vacuum) Einstein equation are the Kerr solutions. (Here we are actually making an additional assumption — the cosmic censorship hypothesis — that the geometry exterior to the horizon is globally hyperbolic, for this is a necessary hypothesis in the no-hair theorems.) Now the idea is that, after collapse, the object that forms loses all its "hair" by radiating it to infinity or through the (future) horizon, and settles down to a stationary configuration. Thus the exterior geometry becomes extremely well approximated by Kerr geometry (to exponential accuracy, as time passes).

The Kerr solutions are a two-parameter family of solutions, parameterized by mass (M) and angular momentum per unit mass (a). In the Boyer-Lindquist coordinate system:

$$ds^2 = \frac{\Delta}{\rho^2}\left(dt - a\sin^2\theta d\phi\right)^2 - \frac{\sin^2\theta}{\rho^2}\left[\left(r^2+a^2\right)d\phi - adt\right]^2 - \frac{\rho^2}{\Delta}dr^2 - \rho^2 d\theta^2,$$

where

$$\rho^2 = r^2 + a^2\cos^2\theta,$$
$$\Delta = r^2 - 2Mr + a^2;$$

$a = 0$ is the Schwarzschild solution. This metric is stationary (independent of t) but is not static ($g_{0i} \neq 0$). It is invariant not under $t \to -t$, but has instead the symmetry

$$t \to -t,$$
$$\phi \to -\phi,$$

as time reversal changes the black hole's sense of rotation.

In the Schwarzschild case, we were able to invoke the Rindler analogy. We considered a FIDO close to horizon, and, since she has proper acceleration

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

she should see radiation with temperature

$$T(r) = \frac{a(r)}{2\pi}.$$

To invoke the Rindler analogy in the Kerr case, we need to decide who the FIDO's are. This is not so obvious, because the rotating black hole drags locally inertial frames along as it spins. Continuing with the coordinates t, r, θ, ϕ, we can rewrite the Kerr metric in the form

$$ds^2 = \alpha^2 dt - g_{ij}\left(dx^i + \beta^i dt\right)\left(dx^j + \beta^j dt\right)$$

Here α is called "lapse," and β^i is called "shift." The world line of a FIDO is chosen so that she perceives the geometry to be *static* — independent of time — *and* with constant-time surfaces orthogonal to her motion. Thus, we choose

$$\left(\frac{d\vec{x}}{dt}\right)_{FIDO} = -\vec{\beta},$$

and then

$$\left(\frac{d\tau}{dt}\right)_{FIDO} = \alpha$$

is the redshift factor that specifies the rate at which a FIDO's clock runs. (See *The Membrane Paradigm* [23], p. 67 ff.)

The horizon is the surface where $\alpha = 0$. This occurs for

$$r = r_H = M + \sqrt{M^2 - a^2}$$

($a < M$, according to cosmic censorship). The FIDOs spin around the BH, with an angular velocity that, as the horizon is approached, becomes

$$\omega_{FIDO} \to \Omega_H = \frac{a}{2Mr_H}, \quad \text{as } r \to r_H.$$

In fact, due to the freezing of her motion (in Boyer-Lindquist "universal" time t) as the horizon is approached, all timelike and null paths corotate at angular frequency Ω_H at the horizon.

Now, the proper acceleration, relative to an FFO, of a FIDO very close to the horizon is

$$a_{FIDO} \approx \frac{1}{\alpha}\kappa$$

where κ is the "surface gravity;" it turns out to be a constant on the horizon, given by

$$\kappa = \frac{r_H - M}{2Mr_H};$$

κ is acceleration measured per unit t, rather than per unit of FIDO proper time. We can interpret the surface gravity this way: If a mass m is hovering over the horizon, at rest in the FIDO frame, a distant observer supporting the mass at the end of a long string must exert a force $F = m\ \kappa$.

– IMAGE –

If we now invoke the Rindler analogy for these FIDOs, we conclude that they detect radiation with the locally measured temperature

$$T(r,\theta) = \frac{1}{\alpha(r,\theta)}\left(\frac{\kappa}{2\pi}\right),$$

which becomes

$$T_\infty = \frac{\kappa}{2\pi} \quad \text{at } r = \infty.$$

So $\frac{1}{4M}$ is replaced by κ when we generalize our Schwarzschild analysis to Kerr geometry.

5.6.2 Hawking emission from a Kerr black hole

If we repeat our computation using the Bogoliubov transformation for a Kerr black hole, the most significant difference is that

$$\frac{\partial}{\partial t} + \Omega_H \frac{\partial}{\partial \phi}$$

is the Killing vector, in the Kerr case, that characterizes the horizon — i.e., it generates the world line of an outgoing photon that is stuck at the horizon.

In the Schwarzschild case, the key to the derivation of the Bogoliubov transformation was the relation between the Killing parameter u and the affine parameter U of a null geodesic at the horizon:

$$U = -e^{-u/4M}.$$

In the Kerr case, this becomes, for an outgoing geodesic of the horizon:

$$(\text{Affine Parameter}) = -\exp[-\kappa(\text{Killing parameter})],$$

with κ replacing $\frac{1}{4M}$. In addition, since an outgoing wave $u \sim e^{-i\omega t}e^{im\phi}$ has eigenvalue $\omega - \Omega_H m$ with respect to the Killing vector $\frac{\partial}{\partial t} + \Omega_H \frac{\partial}{\partial \phi}$ that generates the horizon, when we compute the Bogoliubov coefficients the factor $e^{-8\pi M\omega}$ becomes generalized to $e^{-\frac{2\pi}{\kappa}(\omega - \Omega_H m)}$. The occupation numbers of the outgoing modes, therefore, are

$$n_{\omega m} = |t_{\omega m}|^2 \frac{1}{e^{\beta(\omega - \Omega_H m)} - 1}$$

with $\beta = 2\pi/\kappa$. Thus, $\beta\Omega_H$ enters the Gibbs distribution like a chemical potential for angular momentum, as expected for a rotating thermal reservoir.

A peculiar feature of the result is that for $\omega - \Omega_H m < 0$, we have

$$n = |t|^2(\text{negative occupation number}).$$

To interpret this formula, $|t|^2$ should be replaced by $1 - |r|^2$, and for $\omega - \Omega_H m < 0$ we have

$$1 - |r_{\omega m}|^2 < 0.$$

This phenomenon is called "superradiance." The amplitude of the wave that is reflected by the potential barrier is actually larger than the amplitude of the incoming wave. In the scattering process, the reflected wave picks up some angular momentum from the black hole. In quantum-mechanical language, this classical field transfer process can be interpreted as "stimulated emission." In addition to this stimulated emission, there is also a spontaneous emission process whereby the black hole reduces its angular momentum by emitting a quantum with $m\Omega_H > \omega$ (this is *not* Hawking radiation, but a logically independent phenomenon associated with the *ergosphere* of the Kerr black hole.) The rate for this emission process is enhanced by incoming quanta. (More detail concerning the mode expansion in a Kerr background can be found in DeWitt [8].)

Since time-reversal invariance is broken by the Kerr background, the relation between the reflection of modes incoming from $\mathscr{I}^-$ and from H^- is somewhat different than for the Schwarzschild background. The symmetry of Kerr is $t \to -t$ and $\phi \to -\phi$, which relates $\omega \to -\omega$ to $m \to -m$. That is, the symmetry relates reflection of modes with angular momentum m coming from H^- to reflection of modes with angular momentum $-m$ coming from $\mathscr{I}^-$. Hence, for $\Omega_H > 0$, we have

$$\begin{cases} m > 0 & \text{modes preferentially emitted,} \\ m < 0 & \text{modes preferentially absorbed.} \end{cases}$$

Both phenomena cause the black hole to spin down.

Angular momentum is radiated away efficiently, so that a/M tends to decrease. (Calculations of Page [21] — see *The Membrane Paradigm* [23]). Note also, that because of this tendency to spin down, a Kerr black hole cannot actually be in equilibrium with a thermal bath, unless the bath is *also* spinning (with a "chemical potential" $\mu = \beta\Omega_H$).

5.7 Black hole thermodynamics

Naively, a black hole seems to provide a mechanism for violating the second law of thermodynamics. We can dump some hot gas into the hole; when it crosses the horizon, we lose access to the gas and to its entropy.

But the thermal emission from the black hole suggests that black holes do not supercede the second law, but rather require a modification in how it is stated: We must define a new entropy

$$S' = S_{\text{everything else}} + S_{\text{black hole}};$$

this is the quantity that is nondecreasing; the loss of entropy of the gas is (at least) compensated by the gain of entropy of the black hole.

Using the calculated temperature $T = \frac{1}{8\pi M}$, we can compute the intrinsic entropy of the black hole by considering a black hole in equilibrium with a radiation bath. Our calculation showed that a black hole surrounded by radiation at this temperature will emit and absorb at the same rate (the same transmission probability $|t|^2$ enters into both processes). Let us demand that the generalized entropy S' is *stationary* for this state. (We leave aside, for the moment, the question whether S' is at a local maximum; i.e., whether the equilibrium is stable.)

We have:

$$\Delta S' = 0 = \left(\frac{\partial S}{\partial E}\right)_{\text{rad}} (\Delta E)_{\text{rad}} + \left(\frac{\partial S}{\partial E}\right)_{BH} (\Delta E)_{BH}$$

and energy conservation implies $(\Delta E)_{\text{rad}} + (\Delta E)_{BH} = 0$. Therefore $\left(\frac{dS}{dE}\right)_{BH} = \frac{1}{T_{\text{rad}}} = 8\pi M_{BH}$. Integrating, we find:

$$dS_{BH} = 8\pi M_{BH} dM_{BH}$$
$$\implies S_{BH} = 4\pi M_{BH}^2 + \text{constant}$$

It is appropriate to ignore the additive constant of integration in S if one believes that the entropy of a Planck-sized black hole should be of the order of one, and $M \gg M_{\text{Planck}}$; thus

$$S_{BH} = \frac{1}{4} A,$$

where A is the area of the event horizon. In this form, the formula also applies to the Kerr black hole, where now the area is $A = 4\pi(r_H^2 + a^2)$. Here we've used units with $\hbar = c = G = k_B = 1$. Recalling that entropy has the units of Boltzmann's constant k_B, we can use dimensional analysis to restore the missing factors, obtaining.

$$\frac{S}{k_B} = \frac{1}{4}\frac{c^3}{G\hbar} A = \frac{1}{4}\frac{A}{L_{\text{Planck}}^2}.$$

Aside from a factor of $1/4$, the entropy is the area of the event horizon in Planck units.

This intrinsic entropy of the black hole is enormous. For a solar mass black hole S/k_B is of order 10^{78}, about 20 orders of magnitude larger than the entropy of the sun. Black holes in galactic centers can have masses larger by factors of tens of billions, and hence entropy larger by a factor 10^{21} or more. This means that the entropy of a single black hole can far exceed the entropy of the 2.7K cosmic background radiation in the entire visible universe.

Why should a black hole have this enormous intrinsic entropy? Perhaps because, once a horizon forms, (almost) all information about how it was made becomes inaccessible, even in principle, to an observer outside the horizon. (Only energy, angular momentum, and charge, if any, can be measured—-as for a thermal reservoir when we are completely ignorant of its microscopic state.) In some sense,

$$e^{S_{BH}/k_B} = e^{A/4}$$

represents the number of "microscopic states" accessible to a black hole. It is as though there is a membrane at the horizon, one Planck length deep, with about one bit of information residing in each Planck volume of the membrane.

This idea that the black hole has huge intrinsic entropy was anticipated by Bekenstein before Hawking's discovery of the thermal emission [1, 2]. He was inspired by the theorem that no *classical* process can violate the area law of the horizon (a result also due to Hawking), which bears an obvious resemblance to the second law of thermodynamics. Bekenstein suggested that S_{BH} is of the same order as the horizon area in Planck units, but the numerical coefficient could not be determined until $T = (8\pi M)^{-1}$ was computed.

5.8 Black hole in a radiation cavity

5.8.1 Local and global stability

The equilibrium between a black hole and a very large radiation bath is actually unstable, because specific heat of a black hole is *negative*. For most thermal bodies, the temperature decreases as the body loses heat. For a black hole it is the opposite — its temperature *increases* as it radiates heat away.

The entropy and energy of a black hole at temperature T are:

$$S_{BH} = \frac{1}{16\pi G T^2}, \quad E_{BH} = \frac{1}{8\pi G T};$$

both decrease as the temperature rises. Contrast this behavior with that of a box filled with a gas of massless quanta with temperature T:

$$E = aVT^4, \quad a = \frac{\pi^2}{30}\left(B + \frac{7}{8}F\right), \quad S = \int \frac{dE}{T} = \frac{4}{3}aVT^3.$$

Here B is the number of bosonic helicity states and F is the number of fermionic helicity states.

If we put a black hole (BH) with temperature T in a box of radiation at temperature T, where the box is very large, the equilibrium is unstable. If the BH accretes radiation, it cools down and accretes more radiation. If the BH loses mass by emitting radiation, it heats up and emits more.

But if the box is sufficiently small, the equilibrium will be stable. The criterion for *local* stability is that when energy is transferred from the BH to the radiation bath, the radiation heats up more than the BH does; likewise, when energy is transferred from the bath to the BH, the radiation cools down even more than the BH does. Since

$$T_R \sim E_R^{\frac{1}{4}} \Rightarrow dT_R = \frac{1}{4}T_R\frac{dE_R}{E_R},$$
$$T_{BH} \sim E_{BH}^{-1} \Rightarrow dT_{BH} = -T_{BH}\frac{dE_{BH}}{E_{BH}},$$

the criterion for *local* stability is $dT_R > dT_{BH}$ when $T_R = T_{BH}$ and $dE_R = -dE_{BH}$, or

$$E_R < \frac{1}{4}E_{\rm BH} = \frac{1}{4}M_{\rm BH}.$$

If this criterion is not met (if the box is too big or the black hole is too small), the BH will evaporate completely.

Now consider the condition for *global* stability. Given a box with volume V containing energy E, how should E be apportioned between the BH and the radiation gas to maximize the total entropy? Note first that it will never pay to have more than one BH. Consider two black holes of mass M_1 and M_2; how does the entropy change if they merge to form a single black hole with mass $M_1 + M_2$? Since $S_{BH} = 4\pi M_{BH}^2$ (for a Schwarzschild BH), the entropy increases by the ratio

$$\frac{(M_1 + M_2)^2}{M_1^2 + M_2^2} > 1.$$

The merger is entropically favored.

Now suppose a single BH with mass M_{BH} is in thermal contact with a radiation bath with energy E_R; hence the total energy is $E_R + M_{BH}$, and the total entropy is

$$S = 4\pi\frac{M_{BH}^2}{M_{\rm Pl}^2} + \frac{4}{3}(aV)^{\frac{1}{4}}E_R^{\frac{3}{4}},$$

which we can rewrite as

$$S = 4\pi\frac{E^2}{M_{\rm Pl}^2}f(x), \quad \text{where} \quad f(x) = x^2 + c(1-x)^{\frac{3}{4}} \quad x = \frac{M_{BH}}{E}, \quad c = \frac{M_{\rm Pl}^2}{3\pi}\left(\frac{aV}{E^5}\right)^{\frac{1}{4}}.$$

The entropy is maximized when x maximizes the function $f(x)$ in the interval $[0, 1]$.

For c large (large box volume or small total energy) the maximum occurs at $x = 0$; the preferred configuration is pure radiation with no black hole. As more energy is pumped into a box of fixed volume V (or the volume decreases with the total energy fixed), eventually a BH in equilibrium with radiation becomes locally stable, but the global maximum of the entropy still occurs at $x = 0$, until E is somewhat higher.

The temperature T in the box, as a function of the total energy E, behaves as shown.

– IMAGE –

The solid line is the temperature of the globally stable configurations, and the dotted line indicates locally stable configurations.

Exercise 6.13a: For a black hole with temperature 300K, what is the maximum volume V of a cavity filled with gravitons at 300K such that the black hole is locally stable?

Remark: The cavity is large, because the energy of a black hole ($E_{BH} = M_{BH}c^2$) with $T_{BH} = 300\text{K}$ is enormous.

Exercise 6.13b: 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 attained when there is barely enough energy in the box for a black hole to nucleate that is cool enough to be in equilibrium with the remaining radiation. If the energy increases beyond that point, the black hole grows larger and the temperature decreases. The total energy required for the black hole to be stable is huge, and therefore the maximum temperature is quite high.

5.8.2 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 phase 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 black hole with temperature T; then nucleation of a black hole with temperature T will lower the radiation temperature T by a negligible amount. If a black hole with temperature higher than the radiation temperature T occurs as a thermal fluctuation, it will quickly evaporate. But if a black hole with temperature less than T appears, it will continue to accrete radiation (cooling the box) until the black hole has absorbed a large portion of the total energy in the box.

– IMAGE –

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_{BH} is given (in Planck units) by

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

This is maximized at $E_{BH} = \frac{\beta}{8\pi}$, where the black hole's temperature matches the radiation temperature. The maximum value of the free energy is

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

Thermal fluctuations producing black holes large enough (and hence cool enough) to accrete and grow must surmount this free energy barrier; such fluctuations occur at a rate

$$(\text{Probability of nucleation per unit time and volume}) \propto \exp\left(-\beta F_{BH}|_{\max}\right) = \exp\left(-\frac{1}{16\pi T^2}\right).$$

The prefactor in front of this exponential has also been computed to leading order in $\hbar$; see Gross, Perry, and Yaffe [14],

Of course, it is the temperature in Planck units that appears in the exponential. For, say, $T = 300\text{K}$, we have

$$\text{Rate} \sim \exp\left[-10^{59}\right].$$

So nucleation of a black hole takes quite a while!

5.9 Periodicity in imaginary time revisited

Now that we are considering a black hole in equilibrium with radiation in a cavity, let's revisit the "easy" derivation of black hole radiance in §5.3.2 based on the periodicity in Euclidean time of the analytically continued Schwarzschild geometry.

In terms of Euclidean time τ defined by $t = i\tau$, the Schwarzchild metric is

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

This geometry is easier to grasp if we make a coordinate transformation:

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

in terms of which the metric becomes

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

This geometry is perfectly smooth at $R = 0$ ($r = 2M$) if we regard τ as an angular coordinate in the R–τ plane: $d\tau/4M = d\theta$, where θ has period 2π. Hence, we should regard τ as a coordinate that is periodic modulo $8\pi M$, and the Euclidean Schwarzschild geometry elucidates the periodicity of the Schwarzschild-Kruskal transformation noted in §5.3.2. We have found that the Schwarzschild coordinates *exterior* to the horizon, when continued to imaginary time, cover *all* of a Euclidean manifold with the topology

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

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

As described in §4.10, we may construct the unique Euclidean Green's function that satisfies

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

on this Euclidean geometry and decays for $R \to \infty$. It is analytic on Euclidean Schwarzschild and when continued back to real Schwarzschild time t, is just the *thermal* Green function in Region I, with (redshifted) $\beta = 8\pi M$, because it is the boundary value of a Green function that is periodic in t with period $i\beta$ and is analytic in the strip. (References: Hartle and Hawking (1976), Gibbons and Perry (1976)).

Recall that, as discussed in §4.11, the thermal Green function in Rindler spacetime has an alternative interpretation as a zero-temperature Green function in Minkowski spacetime. It is natural to ask whether this thermal Green function on Region I of the Schwarzschild geommetry has a similar interpretation. Is it $\langle\psi|T[\phi(x)\phi(x')]|\psi\rangle$ for some state $|\psi\rangle$? To identify such a state, we consider the boundary conditions satisfied by the Green function.

For this purpose, we contemplate the analyticity properties of the Green function $G(x, x')$. The thermal Green function in Region I is analytic as a function of t in the lower strip $-\beta < \mathrm{Im}\, t < 0$, where $\beta = 8\pi M$. What does this imply about its behavior as a function of the Kruskal coordinates U and V? Recall that in Region I the Kruskal and Schwarzschild coordinates are related by

$$
\begin{aligned}
U &= -\left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} e^{(r-t)/4M} = -e^{-(t-r_*)/4M} = -e^{-u/4M},\\
V &= \left(\frac{r}{2M} - 1\right)^{\frac{1}{2}} e^{(r+t)/4M} = e^{(t+r_*)/4M} = e^{v/4M};
\end{aligned}
$$

in particular U is an entire function of u and V is an entire function of v. If we give t an imaginary part $t \to t - i\sigma$, then (in Region I) we have

$$
\begin{aligned}
U &\to -e^{-u/4M} e^{i\sigma/4M},\\
V &\to e^{v/4M} e^{-i\sigma/4M};
\end{aligned}
$$

As σ sweeps halfway through the lower strip, $\sigma/4M \in [0, \pi]$, and u, v range over the real line, U covers the lower half plane, except for the point $U = 0$; similarly V also covers the lower half plane, except for the point $V = 0$. Hence the Green function in Region I may be regarded as the boundary value of an analytic function of U and V as both arguments vary in the lower half plane. Specifically, for $V \to 0$ (the past horizon H^-), the Green function is analytic in the lower half U-plane, and for $U \to 0$ (the future horizon H^+, the Green function is analytic in the lower half V-plane. These analyticity properties imply that the Green function is positive frequency with respect to U on the past horizon, and positive frequency with respect to V on the future horizon.

The "right-moving" solutions to the Klein-Gordon equation $\{u_{i,\mathrm{out}}\}$ that are positive frequency with respect to U on the past horizon H^-, together the the "left-moving" solutions to the Klein-Gordon equation $\{u_{i,\mathrm{in}}\}$ that are positive frequency with respect to V on the future horizon H^+, are a complete set of solutions. Consider the corresponding annihilation operators $\{a_{i,\mathrm{out}}\}$, $\{a_{i,\mathrm{in}}\}$ which occur when we expand the field in terms of these solutions.

What we have found is that the thermal Green function obtained by analytically continuing the Klein-Gordon on periodically identified Euclidean space to real time can be equivalently described as the "vacuum" state $|H\rangle$ such that

$$a_{i,\text{out}}|H\rangle = a_{i,\text{in}}|H\rangle = 0.$$

This state $|H\rangle$ is called the "Hartle-Hawking vacuum" [15].

$|H\rangle$ is one of the three vacuum states that are frequently discussed on the black hole background. Another is the "Boulware vacuum" $|B\rangle$ [4], where operators that annihilate the state are associated with solutions that are positive frequency with respect to u on H^- and positive frequency with respect to v on $\mathscr{I}^-$. In this state, FIDOs in Region I detect no particles. Yet another is the "Unruh vacuum" $|U\rangle$ [24], where operators that annihilate the state are associated with solutions that are positive frequency with respect to U on H^- and positive frequency with respect to v on $\mathscr{I}^-$. This state arises in the realistic collapse scenarion where there are no particles infalling from $r = \infty$, and a thermal flux emanates from the (past) horizon.

The observation that the Euclidean Schwarzian geometry is periodic in imaginary time with period $\beta = 8\pi M$ provides a simple and elegant explanation for the thermal emission from the black hole horizon. Importantly, this argument also applies to interacting quantum field theories, for which Green functions have analyticity properties similar to those in the free field theory [13].

The disadvantage of the argument is that it leaves unexplained why the Hartle-Hawking state $|H\rangle$ is the correct one to consider, rather than, for example, the Boulware state $|B\rangle$. (And in fact, it is the Unruh vacuum $|U\rangle$ that really corresponds to a black hole evaporating into empty space; $|H\rangle$ describes a black hole in a thermal bath.) To better appreciate why $|H\rangle$ is preferred to $|B\rangle$, it is useful to consider how the (renormalized) energy-momentum tensor behaves near the horizon in these two states, which will be our next topic.

5.10 The renormalized stress-tensor and Hawking radiation

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 preferred 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 positive and negative frequency, but in terms of local observables $\mathcal{O}(x)$ that can be constructed from fields and are in principle measurable. Since $\langle\phi(x)\rangle = 0$ in a state with definite number of quanta, we might use, e.g., $\phi^2(x)$, or other "composite operators" constructed from $\phi(x)$.

A particular local phenomenon that we might be interested in is backreaction — how the quantum state of the fields feeds back and influences the spacetime geometry. In quantum field theory the Einstein equation

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

becomes an operator equation; hence if the matter fields undergo quantum fluctuations, then the metric fluctuates also. A fully correct discussion of black hole evaporation, then, requires us to go beyond quantized matter fields on a fixed background geometry and to properly quantize gravity.

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

$$G_{\mu\nu} = -8\pi \langle T_{\mu\nu} \rangle \quad \text{("semi-classical gravity").}$$

(This is *not* just the expectation value of the operator equation above, because $\langle G_{\mu\nu} \rangle \neq (G_{\mu\nu}(\langle g \rangle)$ since $G_{\mu\nu}$ depends nonlinearly on g.) This semiclassical Einstein equation is not precisely correct, but might be a reasonable approximation if the fluctuations in the operator $T_{\mu\nu}$ about its mean value are small enough to neglect.

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

$$\begin{aligned} &\langle T_{\mu\nu} \rangle \, u^\mu u^\nu \quad \text{(energy density)}, \\ &\langle T_{\mu\nu} \rangle \, u^\mu n^\nu \quad \text{(flux)}, \end{aligned}$$

where u^μ is the observer's 4-velocity ($u^2 = 1$) and $u \cdot n = 0$, $n^2 = -1$. For instance, in the case of an evaporating black hole, FIDOs near the horizon should measure an energy flux into the hole that is *negative*, since the black hole 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 quantum field theory in curved spacetime concerns $\langle T_{\mu\nu} \rangle$, and some of this literature is rather technical (see Chapter 6 of Birrell and Davies [3], and the book by Fulling [10].) To discuss this, one must consider an aspect of quantum field theory 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 some basic concepts.

5.10.1 Composite operators

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

For example, recall

$$G_+(x-y) = \langle 0|\phi(x)\phi(y)|0\rangle = \frac{-1}{4\pi^2(x-y)^2} \tag{5.3}$$

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 $\left\langle \psi \left| \int d^4x \mathcal{O}(x) f(x) \right| \chi \right\rangle$ finite. That is, a prescription that defines a renormalized operator $\mathcal{O}_{\text{ren}}$.

For example, we can define $\phi(x)^2$ by the point-splitting procedure:

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

In this prescription, we subtract away a divergent c-number (the second term on the RHS). In general, in taking the limit $x \to y$ we may need to average over orientation of $(x-y)^\mu$, with some suitable measure, to ensure that $\mathcal{O}_{\text{ren}}$ transforms as a tensor of the desired type.

The renormalized operator that we have constructed is said to be normal ordered. We have in effect, by making the subtraction, moved all $a(k)$'s to the right and $a(k)^\dagger$'s to the left. The resulting renormalized operator has the property

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

and has finite matrix elements between Fock space states.

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

- Regulate the operator: We modify how $\mathcal{O}$ behaves "at short distances" so that $(\mathcal{O})_{\text{reg}}$ has finite matrix elements — e.g. by point splitting in the above example. Regularization alone is an unsuitable way to define operators because:
    1. Matrix elements of $(\mathcal{O})_{\text{reg}}$ are very sensitive to the artificial cutoff, e.g., $\epsilon = x - y$ in the point splitting case.
    2. Regulation spoils tensor properties of the operator, e.g., a scalar becomes a "biscalar" in the example 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-number, 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 numbers as $\mathcal{O}$ (perhaps including $\mathcal{O}$ itself).
- Remove the regulator. E.g., we take $x \to y$ in the point-splitting method after the divergent part is subtracted away.

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

5.10.2 Stress tensor in curved spacetime

For a massless free scalar field in flat spacetime, the "canonical" stress tensor is (see §3.1):

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \frac{1}{2}\eta_{\mu\nu}\partial_\lambda \phi \partial^\lambda \phi. \tag{5.4}$$

It is a composite object and requires renormalization.

It is natural to normal order it, subtracting so that

$$\langle 0 | (T_{\mu\nu})_{\text{ren}} | 0 \rangle = 0;$$

that is, so that stress-energy vanishes in the vacuum. But subtractions that make $(T_{\mu\nu})_{\text{ren}}$ finite in flat spacetime might not suffice in a general spacetime.

The subtractions that must be made in a general spacetime are "c-number" 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, for conservation of stress-energy to be maintained, whatever terms are subtracted must be conserved tensors. And these subtracted terms are local in spacetime, because they arise from very short-distance fluctuations that probe only the local structure.

We can see the form of subtractions 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 $T_{\mu\nu}$ for ϵ small has the form:

$$\langle T_{\mu\nu} \rangle_{\text{reg}} \sim \underbrace{\frac{1}{\epsilon^4} + \frac{1}{\epsilon^2} + \ln \epsilon}_{\text{infinite part}} + \underbrace{\mathcal{O}\left(\epsilon^2\right)}_{\text{finite part}} .$$

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) $= G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$, involving two derivatives of the metric.

$\ln \epsilon$ term $\propto$ (conserved tensor of dimension 4) $=$ two independent terms each with four derivatives of the metric. These terms are (page 161 in Birrell and Davies):

$$\begin{aligned} H^{(1)}_{\mu\nu} &= 2R_{;\mu\nu} - 2g_{\mu\nu}\Box R - \frac{1}{2}g_{\mu\nu}R^2 + 2RR_{\mu\nu}, \\ H^{(2)}_{\mu\nu} &= 2R^{\alpha}_{\mu;\nu\alpha} - \Box R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\Box R + 2R_\mu{}^\alpha R_{\alpha\nu} - \frac{1}{2}g_{\mu\nu}R^{\alpha\beta}R_{\alpha\beta}. \end{aligned}$$

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

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

Λ_0, G_0 are the "bare" values of cosmological constant and Newton's constant, before renormalization. The RHS equals:

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

If we move the "infinite" part to the left-hand side of the equation, the Einstein equation has the form:

$$G_{\mu\nu} + \Lambda_{\rm ren} g_{\mu\nu} + \text{(4-derivative terms)} = -8\pi G_{\rm ren}\langle (T_{\mu\nu})_{\rm ren}\rangle.$$

Here $\Lambda_{\rm ren}$ and $G_{\rm ren}$ are "renormalized" parameters that are in principle measurable. The new 4-derivative term, which did not appear in the bare Einstein equation, is induced by renormalization.

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

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

where

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

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 the renormalization of the parameters of the theory. So the same low-energy physics can be described by a renormalized action:

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

where

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

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

$$\delta\left(\frac{\Lambda}{8\pi G}\right) \sim M^4,$$
$$\delta\left(\frac{1}{16\pi G}\right) \sim M^2,$$
$$\delta(\text{4-derivative coupling}) \sim \ln\left(\frac{M}{\mu}\right).$$

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

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

while in fact we know from observing the expansion of the universe:

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

This is a spectacular disagreement between theory and experiment — the "Cosmological Constant Problem," which is still an unresolved mystery. The Planck mass, i.e., $\frac{1}{16\pi G}$, also gets substantially renormalized due to field fluctations with wavelength comparable to M_{Pl}^{-1}.

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 equation of motion. We need not consider

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

because

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

is the integral of a *total* derivative (a topological invariant) and therefore does not contribute to the 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 the renormalizations of Λ and G are. The logarithmic dependence arises because many length scales contribute democratically to the renormalization, so that the leading contribution to the renormalization comes from fluctuations at length scales larger than M^{-1}. (The renormalized 4-derivative couplings "run" with μ; more about this below, in connection with the conformal anomaly.)

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 $\left(\frac{L_{\rm Pl}}{L}\right)^2$ relative to the leading terms, where L is the length scale characterizing the curvature.

As noted above, different renormalization schemes for $(T_{\mu\nu})_{\rm ren}$ 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 simply moving a term from $T_{\mu\nu}$ on the RHS of the Einstein equation over to the LHS. This reshuffling has no effect on any physical predictions.

5.10.3 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.

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

is the object satisfying $\partial_\mu T^\mu_\nu = 0$ that is derived from translation invariance by the Noether procedure. The corresponding 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:

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

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

$$\begin{aligned}\left(T^\mu_\mu\right)_{\rm new} &= -\partial_\lambda\phi\partial^\lambda\phi + 3\xi\partial^2\phi^2\\ &= (-1+6\xi)\partial_\lambda\phi\partial^\lambda\phi\end{aligned}$$

(using the equation of motion $\partial^2\phi = 0$). So that $(T^\mu_\mu)_{\rm new} = 0$ for $\xi = \frac{1}{6}$. This choice of ξ 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 for a massless scalar field is

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

where ξ is a free parameter. The Klein-Gordon equation then becomes

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

We can extract the flat-space $T_{\mu\nu}$ by varying S_{matter} with respect to $g_{\mu\nu}$ and then setting $g_{\mu\nu} = \eta_{\mu\nu}$, inferring from

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

that

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

We see that the free parameter in $T_{\mu\nu}$ arises from the freedom to choose how the scalar field couples to the curvature.

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

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

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

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

The coupling to matter is conformally invariant if $T^\mu_\mu = 0$. From the perspective of flat-space quantum field theory, the significance of $T^\mu_\mu = 0$ is that

$$D_\mu = x^\nu \left(T_{\mu\nu} \right)_{\text{new}}$$

is then conserved, where $\int d^3x D_0$ is the operator that generates global scale transformations.

5.10.4 Conformal anomaly

Even if we choose $\xi = \frac{1}{6}$ in coupling the massless scalar to a 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 must introduce an implicit mass scale when we remove the logarithmic divergence proportional to $\ln \epsilon$ (where ϵ is the small short-distance regulator).

Formally, when we make a change of length scale

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

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

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

and

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

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

This is the conformal anomaly: The trace of the renormalized stress tensor is nonzero due to the scale dependence introduced by renormalization. In other words, if the logarithmic renormalization is

$$S = \int d^4x \sqrt{g} \left[\cdots + \underbrace{(\ln \mu) \alpha^{(4)} \mathcal{O}^{(4)}}_{\substack{\text{term arising from} \\ \text{logarithmic} \\ \text{divergence}}} + \cdots \right],$$

then

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

where $\mathcal{O}^{(4)}$ is the term with 4 derivatives of the metric.

5.10.5 Vacuum polarization in Rindler spacetime

As usual, Rindler spacetime provides a good warm-up for the black hole case. We wish to compare how back reaction behaves near the horizon in the Rindler vacuum and the Minkowski vacuum.

Consider

$$\left[\langle 0, \text{Rin} \,| T_{\mu\nu} |\, 0, \text{Rin} \rangle - \langle 0, \text{Min} \,| T_{\mu\nu} |\, 0, \text{Min} \rangle \right] u^{\mu} u^{\nu}$$

where u^{μ} is 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 [7]), but we know the result without doing a new computation, because we know that a Rindler observer sees no quanta in $|0, \text{Rin}\rangle$, and sees a thermal spectrum in $|0, \text{Min}\rangle$.

Hence the above quantity equals:

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

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

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

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

Thus

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

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" contribution that is large and negative.

On the other hand, the Rindler observer in the *Rindler* vacuum sees only 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

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

also blows up at the horizon.

The divergent vacuum polarization may be interpreted as follows: Because of the infinite redshift 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 desperately want to fluctuate at the horizon, but they are prevented from doing so, and therefore exert a strong force on the horizon.

5.10.6 Vacuum polarization in black hole spacetime

Since the Boulware state $|B\rangle$ is analogous to $|0, \mathrm{Rin}\rangle$ and the Hawking-Hartle state $|H\rangle$ is analogous to $|0, \mathrm{Min}\rangle$, it is natural to consider

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

which is 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 [6]).

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 the above quantity is

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

This is the thermal energy density at the local temperature $T = (8\pi M)^{-1}(1-\frac{2M}{r})^{-\frac{1}{2}}$ seen by a static observer at Schwarzschild radius r, 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,

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

while the leading behavior in the Boulware state is

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

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 FFO, who sees no thermal radiation and a smoothly behaving vacuum polarization at the horizon.

On the other hand, 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 net *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 black hole 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 initially at 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 black hole, after it has descended below the stretched horizon. Similarly, a field quantum with redshifted frequency ω_∞ adds mass ω_∞ to the black hole when it is absorbed.

For each mode of the field, the locally measured energy at the horizon is finite, and hence the redshifted energy is zero, if the mode is thermally occupied. So the "energy at infinity" carried by a mode of redshifted frequency ω_∞ is

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

where n is the mode's occupation number and we may view the n_{thermal} term as the vacuum polarization correction near the horizon.

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

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

The occupation number of modes that are escaping downward through the future horizon 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. Therefore

$$n_\downarrow = |r|^2 n_\uparrow + (1 - |r|^2) n_{\text{in}}, \tag{5.5}$$

where $|r|^2$ is the reflection probability and $(1 - |r|^2)$ is the transmission probability.

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 ∞:

$$\begin{aligned}(E_\infty)_\downarrow &= (n_\downarrow - n_{\text{thermal}})\,\omega_\infty \\ &= \left(1 - |r|^2\right)(n_{\text{in}} - n_{\text{thermal}})\,\omega_\infty\end{aligned}$$

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

– IMAGE –

To do the mode counting, imagine putting two concentric spherical shells around the black hole, both very close to the horizon. Consider the radiation modes in the cavity between the two shells. The radial behavior of the modes is $\sim e^{-i\omega(t - r_*)}$, where r_* denotes the tortoise coordinate, and so, as the inner shell is pushed closer to the horizon ($(r_*)_{\text{bottom}} \to -\infty$), we can replace the sum over radial modes by an integral:

$$\sum_{\text{modes}} \to \frac{L}{2\pi} \sum_{lm} \int d\omega$$

where $L = (r_*)_{\text{top}} - (r_*)_{\text{bottom}}$. Since a wave packet propagates from the top to the bottom of the cavity in Schwarzschild time $\Delta t = L$, the rate per unit Schwarzschild time at which energy escapes is

$$\frac{\Delta M}{\Delta t} = \sum_l (2l+1) \int \frac{d\omega}{2\pi} \left(1 - |r_{\omega l}|^2\right) \left(n_{\omega l}^{\text{in}} - n_{\omega l}^{\text{thermal}}\right).$$

In the Hartle-Hawking state, we have $n^{\text{in}} = n^{\text{thermal}}$, and so no mass loss occurs. But in the Unruh state, $n^{\text{in}} = 0$, and this expression coincides exactly with the black hole luminosity in (5.2). 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 modes propagating outward get reflected back. This endows the black hole in the Unruh state with a "thermal atmosphere" as described in the §5.5.3. 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 $n_\uparrow = n_{\text{thermal}}$ and $n_\downarrow = 0$. So the Unruh state has a divergent negative energy density on H^-. This is similar to how the Boulware state behaves near H^-, except that in the Unruh state only half of the modes, rather than all, are empty.

(This description of the energy flux at H^+ in the Unruh state follows *The Membrane Paradigm* [23] and Frolov and Thorne [9])

5.11 The final state of the evaporating black hole

As a black hole evaporates and loses mass, it will eventually reduce its mass to $M \sim M_{\text{Planck}}$. At this point its size is of order L_{Planck}, and its temperature, as computed semiclassically, is of the 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 mass $M \sim M_{\text{Planck}}$ remains, an exotic, stable "elementary" particle.

If the black hole disappears completely, what are the implications?

5.11.1 Breakdown of global conservation laws

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 formed from a collapsing matter star evaporate in the same way as black holes formed from a collapsing antimatter star; in the semiclassical theory, both produce baryons and antibaryons in equal abundance. Even if baryonic and antibaryonic black holes behave differently when $M \sim M_{\text{Planck}}$, by then it is too late to rescue the law of baryon conservation, because a solar mass black hole made from 10^{57} baryons 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 black hole remnants, each with a different value of B, and all, presumably, with $M \sim M_{\text{Planck}}$. That does not sound plausible.

If black hole physics can change baryon number, there is no reason to expect that *virtual* black holes (quantum gravity) cannot induce the decay of the proton. One can reasonably expect that a process

$$\text{quark} + \text{quark} \to \text{antiquark} + \text{antilepton}$$

(the simplest B-changing process allowed by the gauge symmetries of the standard model) would then have an *amplitude*, on dimensional grounds, of $1/M_{\text{Planck}}^2$. Thus, in order of magnitude,

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

To see one decay a year, one would need to watch about $10^6\,\text{km}^3$ 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.

5.11.2 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 ρ describing the mixed quantum state on $\mathscr{I}^+$. In a sense, this was a matter of convenience, since the observer had no access to information about the quantum state of the 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. Our computation of the radiation's state indicates that a pure initial quantum state at $\mathscr{I}^-$ can evolve into a mixed final state at $\mathscr{I}^+$. There is an intrinsic, unavoidable loss of phase information concerning the initial quantum state, and a corresponding intrinsic generation of entropy. This phase information cannot be recovered 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 black hole 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 black hole 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 on could leave an imprint on the black hole that could influence the radiation emitted later.

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 $\exp\left(\frac{1}{4}A_{\text{initial}}\right)$, where A_{initial} is the area of the horizon when the black hole 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 [17]; that this conclusion might be too hasty has been stressed by 't Hooft [19].)

5.11.3 Topology change in quantum gravity

– IMAGE –

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 [18].

The idea seems to be 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 net change in quantum numbers is carried away by a "baby universe" with quantum numbers $\alpha\overline{\beta}$ (e.g., the baby universe 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, explaining the intrinsic loss of information. Since a closed universe has 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 black hole event $\alpha_{\text{in}} \to \beta_{\text{out}}$ is accompanied by the *absorption* by our universe of a baby universe with quantum numbers $\beta\overline{\alpha}$. Putting together

$\alpha_{\rm in} \to \beta_{\rm out}$ with the time-reversed process $\beta_{\rm in} \to \alpha_{\rm out}$, 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 appears 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.

6 Exercises

1. **Free scalar field**

   The free scalar field $\phi(x)$ can be decomposed as

$$\phi(x) = \phi^{(-)}(x) + \phi^{(+)}(x), \tag{6.1}$$

   where

$$\phi^{(-)}(x) = \int \frac{\mathrm{d}^3k}{(2\pi)^3\, 2E}\, e^{-ik\cdot x}\, A(k), \qquad \phi^{(+)}(x) = \phi^{(-)}(x)^\dagger, \tag{6.2}$$

   and $A(k)$ is the relativistically normalised annihilation operator.
   Express $\phi^{(-)}(x)$ as a functional of ϕ and $\dot\phi = \partial_0\phi$.

2. **Two-point functions**

   For the *massless* ($m^2 = 0$) free scalar field theory, calculate

   (i) the Wightman function

$$G_+(x) = \langle 0|\, \phi(x)\, \phi(0)\, |0\rangle, \tag{6.3}$$

   regulating the expression by giving x^0 a small negative imaginary part ($x^0 \to x^0 - i\varepsilon$);

   (ii) the commutator function

$$i\, G(x) = \big[\phi(x), \phi(0)\big]. \tag{6.4}$$

3. **Ground state projection**

   Express the projection operator $\widehat{P}_0 = |0\rangle\langle 0|$ as a normal-ordered function of a and $a^\dagger$:

$$\widehat{P}_0 = :P_0(a, a^\dagger):\ . \tag{6.5}$$

   Hint: Find the differential equations satisfied by $P_0(a, a^\dagger)$ and solve them.

4. **A useful formula**

   Complete the derivation (outlined in §3.8) of the identity

$$\langle 0|\exp\Big[\tfrac{1}{2}\, aMa\Big]\ \exp\Big[\tfrac{1}{2}\, a^\dagger M^* a^\dagger\Big]|0\rangle = \big(\det[\mathbb{1} - MM^*]\big)^{-1/2}, \tag{6.6}$$

   where $M = M^T$.

5. **Bogoliubov transformation as a unitary operator**

   Solve the coupled Bogoliubov equations

$$Ua'U^{-1} = a = \alpha^T a' + \beta^\dagger a'^\dagger, \tag{6.7}$$

$$Ua'^\dagger U^{-1} = a^\dagger = \beta^T a' + \alpha^\dagger a'^\dagger, \tag{6.8}$$

to find the unitary operator U expressed in terms of a' and $a'^\dagger$. Show that it has the same functional form as the expression for U in terms of a and $a^\dagger$ derived in §3.6.

6. **Static metric in flat spacetime**

   The purpose of this exercise is to show that the most general *static* metric on flat spacetime has either the Rindler or Minkowski form.

   (a) Consider a static metric in two–dimensional spacetime. "Static" means there exists a time coordinate η such that the metric is η–independent and that the $\eta = \text{const}$ slices are orthogonal to ∂_η. Show that the coordinates (η, ξ) can be chosen so that the metric takes the form
   $$ds^2 = A(\xi)\, d\eta^2 - d\xi^2.$$

   (b) Coordinates U, V in two–dimensional spacetime are called *null coordinates* if $U =$ const and $V =$ const are null (lightlike) curves. Show that if the spacetime is flat, the most general metric written in null coordinates can be expressed as
   $$ds^2 = B(U)\, C(V)\, dU\, dV.$$

   (c) Using part (b), show that for flat spacetime the function $A(\xi)$ obtained in part (a) must be either
   $$A(\xi) = \text{const} \quad \text{or} \quad A(\xi) = C\big(\xi - \xi_0\big)^2.$$
   Hence, after suitable rescaling of η and translating ξ, the metric can be written in one of the two canonical forms
   $$ds^2 = d\eta^2 - d\xi^2 \quad \text{(Minkowski)}, \qquad \text{or} \qquad ds^2 = \xi^2\, d\eta^2 - d\xi^2 \quad \text{(Rindler)}.$$

7. **A Fourier-transform identity**

   Consider the function
   $$f(u) = \begin{cases} e^{i\omega \ln(-u)}, & u < 0, \\ 0, & u > 0. \end{cases}$$
   If $\tilde{f}(\sigma)$ is the Fourier transform of $f(u)$ defined by
   $$f(u) = \int_{-\infty}^{\infty} \frac{d\sigma}{2\pi}\, e^{-i\sigma u}\, \tilde{f}(\sigma),$$
   show that
   $$\tilde{f}(-\sigma) = -e^{-\pi\omega}\, \tilde{f}(\sigma), \qquad \text{for } \sigma > 0.$$
   Explain the connection of this relation with the Rindler–Minkowski Bogoliubov transformation.

8. **How detecting a quantum alters a thermal state**

A harmonic oscillator of frequency ω is in contact with a thermal reservoir, so that the probability to occupy a state of energy $E = n\omega$ is $P(E) \propto e^{-\beta E}$. Consequently

$$\langle E \rangle_\beta = \frac{\omega}{e^{\beta\omega} - 1}$$

is the expectation value of the oscillator's energy. A "detector" is briefly brought into contact with the oscillator. With probability ε (to lowest order in $\varepsilon \ll 1$) the detector lowers the oscillator's energy by ω, and with probability $1 - \varepsilon$ it leaves the oscillator state unchanged.

(a) If the detector *does* succeed in removing energy from the oscillator, what is the expectation value $\langle E \rangle$ of the oscillator *after detection*?

(b) How, on average, is $\langle E \rangle$ changed by the detector (to leading order in ε), taking into account that the detector may or may not succeed in removing energy?

9. **Green functions on the cyllinder**

Using the contour-integration method described in §4.9, complete the argument that relates Euclidean Green functions on the cylinder and thermal correlation functions by showing

$$\frac{1}{\beta} \sum_{n=-\infty}^{\infty} \frac{e^{2\pi i n \tau/\beta}}{\left(\frac{2\pi n}{\beta}\right)^2 + \omega^2} = \frac{1}{2\omega} \frac{1}{1 - e^{-\beta\omega}} \left(e^{-\omega|\tau|} + e^{-\beta\omega} e^{\omega|\tau|} \right), \qquad -\beta < \tau < \beta. \tag{6.9}$$

10. **A special case of the no-hair theorem**

Prove that the only static (time-independent) solution to the Klein–Gordon equation in the $r > 2M$ region of the Schwarzschild geometry that satisfies

(i) $u \to 0$ as $r \to \infty$,

(ii) u remains *finite* as $r \to 2M$,

is the trivial solution $u = 0$.
Hint: Write the static field equation in the form $0 = Ku$ and consider the positive-definite quantity $\displaystyle\int_{r>2M} u^* K u$. Then perform an integration by parts.

11. **Schwarzschild effective potential**

Show that the Klein–Gordon equation

$$\left[\frac{1}{\sqrt{g}} \partial_\mu \left(\sqrt{g}\, g^{\mu\nu} \partial_\nu \right) + m^2 \right] u(x) = 0 \tag{6.10}$$

in the $r > 2M$ region of Schwarzschild geometry admits solutions of the form

$$u_\ell(t, r, \theta, \phi) = e^{-i\omega t}\, \frac{1}{r}\, R_\ell(r)\, Y_{\ell m}(\theta, \phi), \tag{6.11}$$

where the radial function $R_\ell(r)$ obeys

$$\Big[-\partial_{r_*}^2 - \omega^2 + V_\ell(r_*)\Big] R_\ell(r) = 0, \tag{6.12}$$

and determine the effective potential $V_\ell(r_*)$. Here the tortoise coordinate is

$$r_* = r + 2M \ln\!\left(\frac{r}{2M} - 1\right).$$

12. **Quantum state of Hawking radiation**

    (a) Starting from the expression for ρ_I in eq. (5.1), perform the partial trace over the modes on the future horizon H^+ to find the density matrix ρ_{out} that describes radiation arriving at future null infinity.

    (b) Interpret the result in terms of a thermal distribution with probability $|t_{\omega\ell}|^2$ of reaching $\mathscr{I}^+$.

13. **Black hole in a radiation cavity**

    (a) For a black hole with temperature 300 K, what is the maximum volume V of a cavity filled with gravitons at 300 K such that the black hole is (locally) stable?

    (b) What is the *maximum temperature* of a globally stable configuration in a cavity with volume $1\,\text{cm}^3$?

14. **A collapsing shell**

    The point of this problem is to motivate the analogy between a black hole forming in gravitational collapse and a "moving mirror" that accelerates away from the asymptotic region.

– IMAGE–

For simplicity, assume the collapsing object is a thin shell of massless particles. The shell has negligible thickness and moves on an ingoing radial null geodesic. Because the shell is spherically symmetric, Birkhoff's theorem implies that the geometry is

flat inside the shell:	$ds^2 = dU dV - r^2 d\Omega^2$,
Schwarzschild outside:	$ds^2 = \left(1 - \frac{2M}{r}\right) du dv - r^2 d\Omega^2$.

The two coordinate systems are glued together by demanding that the area $4\pi r^2$ of the shell agrees in both descriptions.

We wish to extend the (u, v) coordinates used by asymptotic observers into the interior by introducing a reparametrization

$$U = U(u), \qquad V = V(u).$$

Since the two coordinate systems coincide as $r \to \infty$, we may choose $V(u) = v$. The function $U(u)$ is then determined by matching the value of r at position of the shell.

(a) Suppose the shell follows $v = v_{\text{shell}} = \text{const}$. By matching the value of r at $v = v_{\text{shell}}$ in the two coordinate systems, determine $u = u(U)$. (Note: U tends to a finite constant as $u \to \infty$.)

(b) For $u \to \infty$, invert this relation to obtain $U(u)$.

(c) In Minkowski coordinates the origin satisfies $V - U = 2r = 0$ (i.e. $V = U$). Find the trajectory $v = v_B(u)$ of the origin in the (u, v) coordinate system.

(d) For solutions of the Klein–Gordon equation of the form $\phi = f_\ell(u, v)\, \frac{1}{r} Y_{\ell m}(\theta, \phi)$ we require $f_\ell \to 0$ as $r \to 0$, so the origin behaves like a perfectly reflecting boundary. If an incoming wave has $f_\ell \sim e^{-i\omega u}$ near the boundary, what is the form of the reflected wave?

References

[1] Jacob D. Bekenstein. Black holes and entropy. *Physical Review D*, 7(8):2333, 1973.

[2] Jacob D. Bekenstein. Generalized second law of thermodynamics in black-hole physics. *Physical Review D*, 9(12):3292, 1974.

[3] Nicholas David Birrell and Paul Charles William Davies. *Quantum fields in curved space*. Cambridge University Press, 1984.

[4] David G. Boulware. Quantum field theory in Schwarzschild and Rindler spaces. *Physical Review D*, 11(6):1404, 1975.

[5] Timothy S. Bunch and Paul C. W. Davies. Quantum field theory in de Sitter space: renormalization by point-splitting. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 360(1700):117–134, 1978.

[6] Philip Candelas. Vacuum polarization in Schwarzschild spacetime. *Physical Review D*, 21(8):2185, 1980.

[7] Philip Candelas and David Deutsch. On the vacuum stress induced by uniform acceleration or supporting the ether. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 354(1676):79–99, 1977.

[8] Bryce S. DeWitt. Quantum field theory in curved spacetime. *Physics Reports*, 19(6):295–357, 1975.

[9] Valery P. Frolov and Kip S. Thorne. Renormalized stress-energy tensor near the horizon of a slowly evolving, rotating black hole. *Physical Review D*, 39(8):2125, 1989.

[10] Stephen A. Fulling. *Aspects of quantum field theory in curved spacetime*. Cambridge University Press, 1989.

[11] Stephen A. Fulling and Simon Norbertus Maria Ruijsenaars. Temperature, periodicity and horizons. *Physics Reports*, 152(3):135–176, 1987.

[12] Robert Geroch and Gary T. Horowitz. Global structure of spacetimes. *General Relativity; an Einstein Centenary Survey*, pages 212–293, 1979.

[13] Gary W. Gibbons and Malcolm J. Perry. Black holes in thermal equilibrium. *Physical Review Letters*, 36(16):985, 1976.

[14] David J. Gross, Malcolm J. Perry, and Laurence G. Yaffe. Instability of flat space at finite temperature. *Physical Review D*, 25(2):330, 1982.

[15] James B. Hartle and Stephen W. Hawking. Path-integral derivation of black-hole radiance. *Physical Review D*, 13(8):2188, 1976.

[16] Stephen W. Hawking. Particle creation by black holes. *Communications in Mathematical Physics*, 43(3):199–220, 1975.

[17] Stephen W. Hawking. Breakdown of predictability in gravitational collapse. *Physical Review D*, 14(10):2460, 1976.

[18] Stephen W. Hawking. Wormholes in spacetime. *Physical Review D*, 37(4):904, 1988.

[19] Gerard 't Hooft. On the quantum structure of a black hole. *Nuclear Physics B*, 256:727–745, 1985.

[20] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. Freeman San Francisco, 1973.

[21] Don N. Page. Particle emission rates from a black hole. II. massless particles from a rotating hole. *Physical Review D*, 14(12):3260, 1976.

[22] Don N. Page. Particle emission rates from a black hole. III. charged leptons from a nonrotating hole. *Physical Review D*, 16(8):2402, 1977.

[23] Kip S. Thorne, Richard H. Price, and Douglas A. MacDonald. *Black holes: the membrane paradigm*. Yale University Press, 1986.

[24] William G. Unruh. Notes on black-hole evaporation. *Physical Review D*, 14(4):870, 1976.

[25] William G. Unruh and Robert M. Wald. What happens when an accelerating observer detects a rindler particle. *Physical Review D*, 29(6):1047, 1984.

[26] Robert M. Wald. On particle creation by black holes. *Communications in Mathematical Physics*, 45(1):9–34, 1975.

[27] Robert M. Wald. Existence of the S-matrix in quantum field theory in curved space-time. *Annals of Physics*, 118(2):490–510, 1979.

[28] Robert M. Wald. *General Relativity*. University of Chicago Press, 1984.