(45)

Quantum Gravity and Topological Field Theory

by

Gilad Lifschytz

B.Sc in Physics
Tel-Aviv University
(1986)

Submitted to the Department of Physics
in Partial Fulfillment of
the requirement for the Degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1995

Signature of the Author ______________________________
Department of Physics
May, 1995

Certified by ______________________________
Professor Roger Brooks
Thesis Supervisor

Accepted by ______________________________
Professor George F. Koster
Chairman, Graduate Committee

Quantum Gravity and Topological Field Theory

by

Gilad Lifschytz

Submitted to the Department of Physics
in Partial Fulfillment of the requirement for the
Degree of Doctor of Philosophy in Physics
at the Massachusetts Institute of Technology
May 1995

ABSTRACT

In this thesis I pursued two paths. In the first part a new construction for getting normalized wave functional and corelation densities is proposed. This scheme is independent of the dimension of space time and hence should in principle work in four dimensions. It is applied to the three dimensional case and it is shown how one can apply it to four dimensions. We believe this construction may give some hints on constructing the full Hilbert space of quantum gravity.

In the second part an investigation of quantum effects in black holes is presented. First a careful analysis of the semiclassical approximation to quantum gravity (quantum matter propagating on a fixed background) is given, which clarifies some confusion in the literature. Then we investigate, in detail, the semiclassical properties of matter on a $2+1$ dimensional black hole background. In chapter 5 we look closer at the semiclassical approximation in the case of black holes and we show that in the toy model of dilaton gravity, the approximation of neglecting geometry fluctuations breaks down on certain hypersurfaces near the black hole horizon. This puts Hawking's conclusion about information loss in doubt, and lends support to the idea that the information can come out in the Hawking radiation. The analysis is done in a two dimensional dilaton gravity model. In the last section it is shown that this effect transcends to four dimensional black holes. We propose an effective theory for describing the interaction between matter and gravity, and explore its properties.

Thesis Supervisor: Professor Roger Brooks

Acknowledgments

I would like to thank my advisor Roger Brooks for our collaboration, for many enlightening discussions and for helping me navigate through the treacherous waters of high energy physics.

I thank my collaborators Esko Keski-Vakkuri and Samir Mathur for a fruitful collaboration. I would like especially to thank Miguel Ortiz for a long and fun collaboration and many philosophical discussions from which I have learned a lot.

I would also like to thank the members of the CTP for financially supporting me with a research assistantship, and for creating a pleasant environment to work in. Special thanks to Jeffrey Goldstone, Alan Guth, Roman Jackiw and Barton Zwiebach for teaching valuable courses and having the patience to answer all my questions.

Thanks are also due to Oren Bergman, Roben Levi, DongSu Bak Esko Keski-Vakkuri and Catherina Riconda for many discussions on physics and life in general.

Last but not least I thank my wife Einat for here support during the last five years and my son Yotam for reminding me on a day to day basis the significance of Gravity.

To My Parents

Contents

Chapter 1

Introduction

A theory of quantum gravity, that is a theory based on the principles of quantum mechanics and of general relativity, has been the goal of physicists for many years. One is faced with numerous conceptual and technical problems when trying to formulate such a theory. Problems like: identifying a time variable, finding physical observables and wavefunctionals, understanding the emergence of our semiclassical world and what physical notions (like unitarity, inner-product, etc.) do actually exist in quantum gravity.

Further there is the issue of the black hole paradox [1] which is hoped to play the same role for quantum gravity as the hydrogen atom played in quantum mechanics.

This thesis presents some investigations on these issues.

A brief introduction to canonical quantization of gravity, topological field theory, quantum field theory on curved space time and black holes is given below.

1.1 Quantum Gravity

In this section we will briefly present the canonical quantization of the Einstein-Hilbert action (for a review see [2]). One starts with the Einstein-Hilbert action ($\hbar = 16\pi G = c = 1$)

$$S = \int_{\mathcal{M}} \sqrt{-g} R$$

where $g = det(g_{\mu\nu})$, $g_{\mu\nu}$ is the 4-metric and R is the curvature scalar. Assuming $\mathcal{M}$ has the topology of $\Sigma \times R$ one can canonically quantize the theory (we will assume Σ is compact). The 3-metric on Σ (g_{ij}) plays the role of the configuration variable. The rate of change of g_{ij} with respect to the label time t pulled back from the foliation is related to the extrinsic curvature by

$$K_{ij}(x,t) = \frac{1}{N(x,t)}\{-\dot{g}_{ij}(x,t) + L_{\vec{N}} g_{ij}(x,t)\}.$$

where N is the lapse function, $\vec{N}$ are the shift functions and $L_{\vec{N}}$ is the Lie derivative along the vector field $\vec{N}$. Then the action takes the form of

$$S = 2\int dt \int d^3x N\sqrt{-g}\{K_{ij}K^{ij} - K_i^i + R\}.$$

The momentum conjugate to g_{ij} is

$$p^{ij} = -2\sqrt{-g}(K^{ij} - g^{ij}K_l^l).$$

From this the hamiltonian can be calculated and is a sum of first class constraints.

$$H = \int_\Sigma d^3x(N\mathcal{H} + N^i\mathcal{H}_i).$$

$\mathcal{H}$ is called the super-Hamiltonian constraint and $\mathcal{H}_i$ are called the super-momentum constraints. One can show that these constraints are equivalent to the vacuum Einstein equations.

The algebra of the super-momentum constraints is the spatial diffeomorphism algebra. The super-Hamiltonian constraint is interpreted as generating deformation of the hypersurface normal to itself as embedded in $\mathcal{M}$. However, the algebra of all the constraints does not generate the diffeomorphism algebra of $\mathcal{M}$, and in fact is not a Lie algebra at all. The canonical quantization follows from the Dirac procedure, that is imposing the constraints on the wave functions

$$\mathcal{H}\Psi[g_{ij}] = 0.$$

This equation is known as the Wheeler-De Witt equation, and is the central equation in canonical quantum gravity. The other equations read

$$\mathcal{H}_i\Psi[g_{ij}] = 0.$$

Further the physical operators $\mathcal{O}$ must weakly commute with the constraints (and hence with the Hamiltonian).

As the Hamiltonian is constrained to vanish on the physical Hilbert space one is faced with the problem of having no time evolution in quantum gravity. This is known as the problem of time.

In the previous discussion I have ignored problems of the normal ordering of the constraint and the possibility of anomalies in the algebra of the constraints.

1.2 Topological Field Theory

Topological field theories are a class of field theories whose physical correlation functions are topological invariants of space time (for a review see [3]). A very important subclass are topological quantum field theories (TQFT). These are field theories with a classical Lagrangian being zero or a topological invariant of the space time manifold. The initial information consists only of the field content and the classical Lagrangian has an enormous set of symmetries which should be gauged fixed. The BRST operator of this gauge fixing will of course satisfy $\mathcal{L} = \{Q, V\}$ where $\mathcal{L}$ is the total gauge fixed Lagrangian. Although at first sight these theories appear trivial they have a rich and interesting behavior. Each TQFT is associated with a moduli space of fields and the geometry of that space is represented in the TQFT by operators and fields. For example the exterior derivative on the moduli space is represented by a BRST operator Q. Other fields represent the cotangent vectors and the curvature two-form of the moduli space. The physical operators in these theories are in the cohomology of the Q operator. The expectation value of physical operators in the specific example of Donaldson-Witten theory turn out to be the Donaldson invariants of four dimensional manifold.

Topological field theories are relevant to physics. Two and three dimensional quantum gravity are topological field theories of the BF type. Further, it was suggested by Witten that certain TQFT's represent an unbroken diffeomorphism invariant phase of quantum gravity.

In this thesis we will present another application which is the construction of a limited class of correlation densities and normalized wave functionals in quantum gravity, which in general should work in any dimension.

1.3 QFT in curved space time

In this section we will briefly describe how to quantize a scalar field on curved space time (for a review see [4]). Treating the matter field as quantized and gravity as a classical background is called the semiclassical approximation. It is believed to be valid in regions of low curvature.

We will be dealing with a free scalar field whose Lagrangian is (the signature is $(+---)$):

$$L = \frac{1}{2}\sqrt{g}\{g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - [m^2 + \xi R]\phi^2\}.$$

The equation of motions are

$$\{\Box + m^2 + \xi R\}\phi = 0.$$

The solution to these equation form the basis of the mode expansion analogous to the modes $e^{i(wt-kx)}$ in Minkowski space.

The Klein-Gordon inner product is

$$(\phi_1, \phi_2) = -i\int_\Sigma \phi_1 \overleftrightarrow{\partial}_\mu \phi_2^* [-g_\Sigma(x)]^{\frac{1}{2}} d\Sigma^\mu$$

where Σ is a spacelike Cauchy surface. The different mode solutions ϕ_n are orthogonal with respect to this inner product.

i.e. $(\phi_m, \phi_{m'}) = \delta_{mm'}$, $(\phi_m, \phi_m^*) = 0$, and $(\phi_m^*, \phi_{m'}^*) = -\delta_{mm'}$. As usual the field operator is expanded in these modes

$$\phi = \sum_m \phi_m a_m + \phi_m^* a_m^+$$

so that a, a^+ destroy and create particles. Define the vacuum state $|0\rangle$, by $a_m|0\rangle = 0 \;\; \forall m$.

The two point function (Wightman function) is defined as

$$G^{+}(x,x') = \langle 0|\phi(x)\,\phi(x')|0\rangle = \sum_{m} \phi_m(x)\phi_m^*(x').$$

Now in curved space time there is in general no canonical choice of positive frequency modes (ϕ_n) which amounts to the fact that there is no canonical choice for what we call a particle, or a vacuum state.

So in general one could define different sets of modes which will correspond to a different definition the vacuum. However all these modes are in the same Hilbert space hence one can write

$$a_i = \sum_j (\alpha_{ji}\tilde{a}_j + \beta^*_{ji}\tilde{a}^\dagger_j)$$

Where $\tilde{a}$ are the annihilation modes in a differently prescribed vacuum. α, β are called Bogolubov coefficients. One finds that with respect to the original vacuum the tilde vacuum has expectation value for the original particle operator

$$\langle\tilde{0}|N_i|\tilde{0}\rangle = \sum_j |\beta_{ji}|^2.$$

Of course in the presence of a time like Killing vector one could define a canonical set of modes by

$$\mathcal{L}_\zeta \phi_m = -iw\phi_m$$

Where $\mathcal{L}_\zeta$ is the Lie derivative with respect to the time like Killing vector ζ.

1.4 Black holes

Black holes have many interesting features both classically and quantum mechanically. Black hole solutions to the classical equation of motion are found in almost all theories of gravity (Einstein, dilaton ,strings). The best known black hole is called the Schwarzschild black hole. The metric of a mass M Schwarzschild black hole is

$$ds^2 = -(1-\frac{2M}{r})dt^2 + (1-\frac{2M}{r})^{-1}dr^2 + r^2 d\Omega^2.$$

In classical Einstein gravity it was found in the seventies by Bardeen Carter and Hawking that black hole solutions are characterised by mass (M), charge (Q) and angular momentum (J). Now charged, rotating black holes obey an equation which is very suggestive

$$M = \frac{\kappa}{8\pi}A - 2\Phi Q + \Omega J$$

where A the area of the event horizon, κ the surface gravity at the horizon, Φ the electric potential at the horizon and Ω the angular velocity at the horizon. Further it was found that in all classical processes the horizon area always grows. This inspired Bekenstein to interpret A as an entropy and κ as the temperature of the black hole. This did not seem to be all too consistent as classically black holes do not emit any thing and hence can not be in thermal equilibrium with a heat bath. This was later understood when Hawking found that quantum mechanically black holes do emit particles with a thermal distribution with temperature $T = \frac{\kappa}{2\pi}$ [5]. The Hawking radiation is due to negative frequency matter modes becoming positive frequency modes when propagated through the collapsing matter.

It is possible to reproduce Hawking's result as a calculation on an eternal black hole background via a choice of vacuum states. There are three natural vacuum states to be considered on a black hole background.

1. Boulware vacuum. The vacuum is defined with respect to Schwarzschild modes on the past horizon and at $\mathcal{I}^-$. However the energy momentum tensor diverges at both future and past horizons.

2. Hartle Hawking vacuum. The vacuum is defined with respect to Kruskal modes at both future and past horizons. This represents a black hole in thermal equilibrium with a heat bath of particles.

3. Unruh vacuum. The vacuum is defined with respect to Kruskal modes on the past horizon and Schwarzschild modes at $\mathcal{I}^-$. This represents a black hole with a thermal particle flux at future infinity.

Despite this level of understanding there is still no satisfactory picture of the microstates associated with the black hole entropy or a proof that the generalized second law of thermodynamics (GSL) is valid (the GSL is just the usual law taking into account the black hole entropy).

1.4.1 Black hole paradox

If black holes emit particles they will gradually shrink and then disappear leaving behind, according to Hawking's calculation, just thermal radiation. This final sate does not depend on how or from what the black hole formed. Given the initial state one can deduce the final state but not vice versa. This led Hawking to suggest that the black hole evaporation process is not unitary and that pure states can evolve into mixed states and that the usual rules of physics should be changed to accommodate that [1]. This point of view raises some problems and over the years there appeared three possible resolutions to this problem, although none is completely satisfactory (for a review see [6]).

1. Hawking is right. Information is lost or effectively lost to some baby universe.

2. The black hole shrinks down to around the Planck mass and then stops, leaving a remnant which contains all information of the initial state.

3. The Hawking radiation carries with it the information of the initial state. This can happen only if the semiclassical approximation breaks down, and this means that some new interesting physics is involved.

The main objection to the first suggestion is that all known models having the property of converting pure states into mixed states have unwanted features like non conservation of energy. Further, particle physicists find it hard to be separated from friendly notions like unitarity.

The objection to the remnant idea is that you basically need an infinite number of them in order to keep track of all the possible initial states that formed the black hole. It is then unclear how one would get rid of the problem of having those remnants invade the low energy physics through infinite production rates.

The last suggestion has a clear disadvantage. The difference between the first two ideas can only be settled theoretically once we have a full theory of gravity that can be trusted near the singularity. For the last idea to work quantum gravity effects must become large in a region of very low curvature where it is believed they can be neglected (the horizon). Even further it involves abandoning the usual connection (in general relativity) between observation made by different observers.

Chapter 2

Quantum Gravity and Equivariant Cohomology

On Monday, when the sun is hot
I wonder to myself a lot:
"Now is it true, or is it not,
That what is which and which is what?"
(The World of Pooh by, A. A. Milne)

Abstract

A procedure for obtaining correlation function densities and wavefunctionals for quantum gravity from the Donaldson polynomial invariants of topological quantum field theories, is given. We illustrate how our procedure may be applied to three and four dimensional quantum gravity. Detailed expressions, derived from super-BF gauge theory, are given in the three dimensional case. A procedure for normalizing these wavefunctionals is proposed.

This chapter is based on [7]

2.1 Introduction

Topological invariants on a manifold are a subset of diffeomorphism invariants. Thus we expect that elements of the set of topological invariants should be a subset of the quantum gravity observables. Additionally, it is generally believed that observables, which are elements of the BRST complex, may be used to construct vertex operators or wavefunctionals for the theory. Consequently, should we succeed in constructing observables for quantum gravity, we might also be able to construct wavefunctionals. These statements form the nexus for the present work. The puzzle is how to find representations of topological invariants in quantum gravity theories in sufficient generality so as not to explicitly exploit the topological nature of low-dimensional gravitational theories. In this paper, we will give a formal procedure for constructing operators which have the interpretation as the densities of correlation functions of observables and which lead to wavefunctionals, in this fashion.

Loop observables, which are constructed from Wilson loops, have been proposed [8, 9] for four dimensional canonical gravity in the Ashtekar formalism [10] via the loop representation [11]. In this way, observables which measure the areas of surfaces and volumes of regions have been constructed [9]. These are intricate constructions and we wonder if they may be placed in a different context via appealing to the geometry of the space of solutions to the constraints. From observables, we expect to be able to find states, and, perhaps, their wavefunctionals. Put into focus, our quest for a geometrical interpretation for general quantum gravity is a hope that we may be able to exploit the geometry to directly construct wavefunctionals. This is not to mean that we are diminishing the importance of observables.

Indeed, the geometry which underlies gauge field theories suggests another way of representing wavefunctionals; this will be the focal point of our exploration in this work. In particular, as both three dimensional gravity [12, 10] and the super-BF gauge theory [3] of flat $SO(2,1)$ or $SO(3)$ connections (in which the geometry of the space of connections is explicit) share the same moduli space, these theories are natural choices for experimentation on this idea. We will find an interesting relation between the polynomial topological invariants of three dimensional flat connection bundles, which are the analogs of Donaldson's invariants [13] for self-dual connections in four dimensional Yang-Mills gauge theory, and correlation densities of three dimensional quantum gravity. This does not mean that we will find correlation densities of new observables. We expect that the ones we will obtain may be decomposed in terms of Wilson loops. Further pursuit of our ideas then lead us to expressions for canonical and Hartle-Hawking wavefunctionals which satisfy the constraints of three dimensional gravity. By exploiting previous work on four dimensional topological gravity, we are also able to sketch how our approach works in this physical dimension. Due to the fact that much more is known about the associated three dimensional topological quantum field theories (TQFT's) [3] than four dimensional topological gravity, we are presently unable to give expressions which are as detailed as those for three dimensional quantum gravity. We should point out that while the correlation densities and wavefunctionals which may be constructed via our approach for three dimensional gravity are likely to span the full space of such quantities, we do not expect this to be the case for four dimensional gravity. The reason is simply that the phase space of the former theory and of TQFT's is finite dimensional while that of the latter is not.

Our work is relatively formal as our objective is to establish an approach to solving some of these long standing problems of quantum gravity. In particular, we give expressions in terms of path integrals which, in principle, may be computed exactly. These path integrals appear as those of topological quantum field theories which are strongly believed to be, at worst, renormalizable [3]. This allows us to make use of BRST analysis techniques in order to establish our results. A related approach for the computation of scattering amplitudes in string theory was undertaken by one of us in ref. [14].

Commencing, we establish the framework of our approach while attempting to be as general as possible, in the next section. Implementation of the approach is carried out for three dimensional BF-gauge theory, in general, and 3D quantum gravity, in particular, in section 2.3. Expressions for correlation densities are given in sub-section 2.3.1 while wavefunctionals may be found in sub-section 2.3.2. The four dimensional case is sketched in section 2.4. Our conclusions may be found following that section. In addition, appendices summarizing BF-gauge theories and super-BF gauge theories are given. In appendix C we suggest the possible existence of polynomial invariants in pure three quantum gravity, before applying our approach. Our global notations are given in appendix D.

2.2 The Heuristic Construction

As was discussed in the introduction, our approach is to first find correlation densities and then extract the wavefunctionals from them. Thus, in this section, we first concentrate on our general approach to obtaining the correlation densities. Then, we will discuss how to obtain the wavefunctionals from them, at the end of this section.

2.2.1 Correlation Densities

Given a field theory, one is interested in its physical states and the observables; *i.e*, functionals and functions of the fields which obey the constraint of the theory. One reason why observables are important is that from them physical correlation functions can be constructed. However, it is not necessary to find observables in order to construct physical correlation functions of the fields. As an example, given a function, $\hat{\mathcal{O}}$, of the fields, we will only demand that the vacuum expectation value $\frac{\delta\langle\hat{\mathcal{O}}\rangle}{\delta g_{\mu\nu}}$ vanishes, where $g_{\mu\nu}$ is some background metric. This allows for $\frac{\delta\hat{\mathcal{O}}}{\delta g_{\mu\nu}} \neq 0$. The $\hat{\mathcal{O}}_i$ we will construct will have the property that generally $\frac{\delta\langle\prod_{i=1}^{n}\hat{\mathcal{O}}\rangle}{\delta g_{\mu\nu}} \neq 0$ for $n \geq 2$. Thus they are really physical correlation function densities. In this section, we will describe how we can use TQFT's in order to construct, a set of the $\hat{\mathcal{O}}$'s for a general field theory (GFT). Our focus will be on quantum gravity for which topological observables are of interest.

Take a GFT for fields, X, which are sections of a bundle over a manifold, M, and whose space of physical fields is called $\mathcal{N}$. Construct [15] a TQFT which describes the geometry of a subspace of $\mathcal{N}$, which we call $\mathcal{M}$ (the dimension of $\mathcal{M}$ is finite). In this way, we have projected the GFT onto the TQFT. Expectation values of observables in the TQFT (which we generally know how to write), are topological invariants of M. Now if the TQFT has the constraints, $\mathcal{G}$, of the GFT as a subset of it's own constraints then we can construct physical correlation functions and wavefunctionals of the GFT, with the use of the TQFT. We now describe two different ways of doing this.

First, suppose we are given a particular GFT for fields X and are able to construct a TQFT with fields X and Y. Let us require that this TQFT has the same Lagrangian as the GFT plus additional terms which are also invariant under the local symmetries of the GFT*. Furthermore, we require that a subalgebra of the constraints of our TQFT is isomorphic to the constraint algebra of the GFT. In particular, the action of this subset on the X, in the TQFT is the same as the action of the GFT's constraints on X. As an example, take the GFT to be BF-gauge theory and the TQFT to be super-BF gauge theory.

Now take a set of observables in the TQFT and *almost* compute their correlation function. By this we mean the following. Integrate the path integral over all the fields that are present in the TQFT but not in the GFT; that is, over Y. We then get an expression (which is typically non-local), $\hat{\mathcal{O}}$, in terms of the fields X. The expectation value of $\hat{\mathcal{O}}$, in the GFT, is a topological invariant of M. More precisely,

$$\left\langle \hat{\mathcal{O}} \right\rangle_{GFT} = \int [dX] e^{-S_{GFT}} \hat{\mathcal{O}}(X) = \int [dX][dY] e^{-S_{TQFT}} \mathcal{O}(X,Y) \quad , \qquad (2.2.1.1)$$

where $\mathcal{O}$ is a product of observables in the TQFT, S_{GFT} is the action of the GFT and $\hat{\mathcal{O}}$ is a gauge invariant, non-local expression in terms of the original fields. Really what we are doing is taking the original theory and coupling special "matter" to it, and using the matter part to construct physical correlation functions. However, using TQFT's has additional rewards. First, these expressions are computable as the theories are, at worst, renormalizable. Second, we will see that we will be able to write expressions without integrating over the entire spacetime manifold, which satisfy the constraints of quantum gravity.

Although it is tempting to call the correlation densities observables, this can only be done with qualification as they do not have one of the important properties we associate with observables. That is, generally the product of two or more of them is not an observables in the sense that this product's vacuum expectation value will not be diffeomorphism invariant.

---

*Y may be thought of as the supersymmetric partners of X and the additional terms in the action as supersymmetric completion.

To place the above arguments in a geometrical setting, let us look at the geometry of the space of connections [16]. This argument applies, in principle, to any gauge theory built from a Yang-Mills fields space. Let $P(M, G)$ be a G-bundle over the spacetime manifold, M, and $\mathcal{A}$ be the space of its connections. Forms on the space $P \times \mathcal{A}$, $\Omega^{(p,q)}(P \times \mathcal{A})$, will be bi-graded inheriting degrees p from M and q from $\mathcal{A}$. A connection $\mathbf{A} \equiv A + c$ may be introduced on the bundle $P \times \mathcal{A}$ along with an exterior derivative $\mathbf{d} \equiv d + Q$ where $d(Q)$ is the exterior derivative on $M(\mathcal{A})$. The object, c is the ghost field of the Yang-Mills gauge theory. The total form degree of $\mathbf{A}$ is one and is given by the sum of the degree on M and ghost number. The curvature of the connection $\mathbf{A}$ is $\mathbf{F} = \mathbf{dA} + \frac{1}{2}[\mathbf{A}, \mathbf{A}] = F + \psi + \phi$, where the $(2,0)$ form $F = dA + A \wedge A$ is the usual curvature of P, $\psi = QA + d_A c$ is a $(1,1)$ form and $\phi = Qc + \frac{1}{2}[c, c]$ is a $(0,2)$ form. Gauge invariant and metric independent operators may be constructed out of these objects. They are the Donaldson invariants written in a field theoretic language. Thus we will be attempting to recover these geometrical objects which already exist, but are hidden, in physical gauge theories.

2.2.2 Wavefunctionals

We can also obtain wavefunctionals of the GFT's fields, X, which satisfy the latter theories constraints. A general method will be described first, then another prescription which we will later see works for three dimensional gravity, but which is not guaranteed to work in general, will be given. In the following we will use the term geometrical sector to refer to those fields which are realized as the curvature components for the geometry of the universal bundle over X. For example, these would be (A, ψ, ϕ) in a theory defined over a Yang-Mills field space.

First, take the TQFT to be defined over a spacetime manifold M with a boundary, ∂M, which is homeomorphic to the surface, Σ, on which we wish to quantize the GFT. As for the GFT, let the phase space of the TQFT be even-dimensional. Note that M need not be diffeomorphic to $\Sigma \times \boldsymbol{R}$. Form the correlation function of a set of observables in this TQFT. Choose a polarization and functionally integrate over the X and Y sets of fields in the TQFT with boundary conditions on Σ. Then, the correlation function will yield a functional of the boundary values of half of the Cauchy data for the X fields, call that set $X|_\Sigma$, and half of the Y fields, call that set $Y|_\Sigma$. By construction this is a Hartle-Hawking wavefunctional for the TQFT which is guaranteed to be computable since, at worst, TQFT's are renormalizable:

$$\Psi[X|_\Sigma, Y|_\Sigma] \;=\; \int [dX][dY] e^{-S_{TQFT}} \mathcal{O}(X, Y) \quad . \tag{2.2.2.1}$$

Here S_{TQFT} is the TQFT action on the manifold with boundary, Σ. The wavefunctional, $\Psi[X|_\Sigma, Y|_\Sigma]$ is diffeomorphism invariant due to the properties of TQFT's. For the particular TQFT, any fields which appear in $\Psi[X|_\Sigma, Y|_\Sigma]$ and which are not in the geometrical sector, should be integrated out. Then all the Y fields which remain in $\Psi[X|_\Sigma, Y|_\Sigma]$ may be replaced by non-local expressions involving X and $\frac{\partial X}{\partial m}$. This is an idiosyncrasy of TQFT's. Then since $\frac{\partial X}{\partial m}$ is a function of X, we obtain

$$\Psi[X|_\Sigma, Y|_\Sigma] \implies \Psi[X|_\Sigma] \ . \qquad (2.2.2.2)$$

In practice, we find that those $Y|_\Sigma$ fields which appear in $\Psi[X|_\Sigma, Y|_\Sigma]$ are Grassmann-odd and the projection to $\Psi[X|_\Sigma]$ stated above is performed by first choosing a basis for $T^*\mathcal{M}$, expanding those $Y|_\Sigma$ in this basis and then expanding $\Psi[X|_\Sigma, Y|_\Sigma]$ as a superfield whose components are wavefunctionals, $\Psi[X|_\Sigma]$.

This approach leads us to the following ansatz for a normalization procedure which stems from the axiomatic approach [17] to TQFT's. Given two wavefunctionals, Ψ_1 and Ψ_2, defined on diffeomorphic boundaries, ∂M_1 and ∂M_2, we might try defining the inner product by gluing the two manifolds together. This will result in a path integral of some observable of the TQFT defined on the glued manifold. As these expressions are finite this gives a possible normalization procedure. We defer the exact construction to future work [18].

A second approach to constructing the wavefunctionals stems from the observation that, in the above, we took the wavefunctionals of the TQFT and projected onto the X subspace to obtain the wavefunctionals of the GFT. Thus it is suggestive to simply construct the wavefunctionals of the TQFT by any means possible and then apply the projection. Thus we need not restrict ourselves to Hartle-Hawking wavefunctionals but might also consider those obtained by directly analyzing the constraints of the canonically quantized TQFT.

Now let us specialize to a certain set of GFT's. For certain theories, such as three dimensional gravity, we may construct such wavefunctionals by building TQFT's which are defined in a background which solves the constraints of the GFT. We will call such TQFT's, servant theories. That is, in the GFT, we solve the constraints first and then quantize. The quantization then demands that we find wavefunctionals which have support only on the constraints' solutions. Realizing this, we construct correlation functions in a servant TQFT which is defined over a certain background. As we will see in the next section, this works when the servant TFT is of the Schwarz [19] type. As the servant TFT's must be topological, this approach restricts the background; *i. e.*, those $X|_\Sigma$ which solve the constraints, to be non-propagating fields or global data. Thus we expect that this approach will only work for certain sectors of four dimensional gravity.

Having given a cursory discussion of our procedures for obtaining observables and wavefunctionals, let us now turn to some specific applications. Three dimensional quantum gravity and BF-gauge theories, in general, are first.

2.3 Application to 3D BF Gauge Theories

As BF-gauge theories are TFT's, they are the logical choice for the first application of the ideas discussed in the previous section. Although our analysis below may be carried out in arbitrary dimensions, we will focus on 2+1 dimensional manifolds. In this dimension, BF-gauge theories are of more than a passing interest; as with gauge group $G = SO(2,1)$, they are known [12, 10, 20] to be theories of quantum gravity. In subsection 2.3.1, we will study the construction of correlation densities in the covariant quantization of BF-gauge theories based on the geometry of the universal bundle. Then in subsection 2.3.2 we will give formal expressions for canonical and Hartle-Hawking wavefunctionals of BF-gauge theories again based on the geometry of the universal bundle. Where appropriate, we will make allusions to three dimensional quantum gravity

Before proceeding we would like to be further explain the rationale for choosing BF-gauge theories (see appendix 2.6) as a first application of our constructions. There are cohomological field theories (or TQFT's), called super-BF gauge theories, which share the same moduli space. As quantum field theories, they are very closely related [21] and the manifest appearance of the geometry of the constraint space of BF-gauge theories in the super-BF gauge theories will be most useful. These two facets make the construction of observables and wavefunctionals for BF-gauge theories from super-BF gauge theories highly suggestive and, as we will find, possible.

2.3.1 Pulling Back $H^*(\mathcal{M})$ to BF-gauge theories

Define $\mathcal{N}$ to be the restriction of $\mathcal{A}$ to flat connections: $\mathcal{N} \leftrightarrow \mathcal{A}|_{F=0}$ and $\mathcal{M} = \mathcal{N}/G$ to be the moduli space of flat connections. Let $m^I,\ I = 1, \cdots, \dim \mathcal{M}$ be local coordinates on $\mathcal{M}$. Flat connections are then parameterized as $A(m)$. Given two nearby flat connections as $A(m)$ and $A(m+dm)$, we expand the latter to see that the condition for it to also be a flat connection is that

$$d_A \frac{\partial A}{\partial m^I} dm^I \;=\; 0 \ . \tag{2.3.1.1}$$

By definition, the zero-mode of the $(1,1)$ curvature component on $P \times \mathcal{A}$, $\psi^{(0)}$, satisfies the equation

$$d_A \psi^{(0)} \;=\; 0 \ , \tag{2.3.1.2}$$

where A is a flat connection. Thus we immediately find a basis from which $\psi^{(0)} = \psi_I^{(0)} dm^I$ may be constructed; namely, $\psi_I^{(0)} = \frac{\partial A}{\partial m^I}$.

We seek observables in the BF-gauge theory which we can formally write in terms of $\frac{\partial A}{\partial m^I}$ assuming we have chosen a coordinate patch on $\mathcal{N}$. In order for them to be observables they must be gauge invariant and diffeomorphism invariant. These conditions are related as we will soon see. Let us now turn to their construction.

For a homology two-cycle, Γ, on M, we define

$$\mathcal{O}_{IJ} \;\equiv\; \frac{1}{2}\int_\Gamma Tr(\frac{\partial A}{\partial m^I} \wedge \frac{\partial A}{\partial m^J}) \ . \tag{2.3.1.3}$$

Under a gauge transformation, $A \to A^g$, $\frac{\partial A}{\partial m^I}$ transforms into $\frac{\partial A^g}{\partial m^I}$. Then for an infinitesimal gauge transformation, with parameter ϵ,

$$\delta_\epsilon \mathcal{O}_{IJ} \;=\; \int_\Gamma Tr(\frac{\partial \epsilon}{\partial m^{[I}} \frac{\partial F}{\partial m^{J]}}) \ . \tag{2.3.1.4}$$

Thus we see that $\mathcal{O}_{IJ}$ is gauge invariant if A is a flat connection. Hence it is a possible observable in BF-gauge theories.

A check of diffeomorphism invariance remains to be done. Diffeomorphisms of the manifold, M, by the vector field, K, are generated by the Lie derivative $\mathcal{L}_K = di_K + i_K d$. By direct computation,

$$\mathcal{L}_K \frac{\partial A}{\partial m^I} = i_K d_A \frac{\partial A}{\partial m^I} + [\frac{\partial A}{\partial m^I}, \alpha(K)] + d_A(i_K \frac{\partial A}{\partial m^I}) , \tag{2.3.1.5}$$

where $\alpha(K) \equiv i_K A$. If A is a flat connection, the first term in the right-hand-side of this expression vanishes. The second term is a gauge transformation. Although the last term is inhomogeneous the fact that it is a total derivative means that after integrating by parts and imposing the flat connection condition, its contribution vanishes. It then follows that $\mathcal{O}_{IJ}$ is an example of a gauge invariant operator whose correlation functions in the BF-gauge theory are diffeomorphism invariants.

Having convinced ourselves, by the example above, of the existence of operators in BF-gauge theories which lead to diffeomorphism invariant correlation functions, we must now establish a procedure for constructing such quantities. This will be done by implementing the ideas in section 2.2; namely, almost compute the topological invariants from the super-BF gauge theory theory. By almost, we mean integrating over all of the fields in the functional integrals except for the gauge connection and the field B. This will leave us with a functional integral expression over the space of fields in the ordinary BF theory but with operator insertions at various points on the manifold. Now we know that we are in fact computing topological invariants. Then it follows that these operators, which will appear as functionals of the connection will be physical correlation densities[†] in the BF-gauge theory whose expectation values are topological invariants.

To illustrate the procedure, let us write the generic Donaldson polynomials as $\mathcal{O}_i(\phi, \psi, F; \mathcal{C}_i))$ where $\mathcal{C}_i$ is the cycle the observable is integrated over. Then we have to compute

$$\begin{aligned} \left\langle \prod_i \mathcal{O}_i(\phi, \psi, F; \mathcal{C}_i) \right\rangle_{SBF} &= (Z_{SBF})^{-1} \int [d\mu]_{SBF}\, e^{-S_{SBF}} \times \\ &\quad \times \prod_i \mathcal{O}_i(\phi, \psi, F; \mathcal{C}_i) \ , \end{aligned} \tag{2.3.1.6}$$

where Z_{SBF} is the partition function of the super-BF theory and $[d\mu]_{SBF}$ (see appendices 2.7 and 2.9) is the measure for the path integral over the fields χ, ψ, etc. with the χ zero-modes inserted. It is known (see appendix 2.7) that certain classes of operators $\mathcal{O}_i$ exist for which these correlation functions are topological invariants.

The integral over λ may be performed leading to the delta function $\delta(\Delta_A \phi + [\psi, {}^\star\psi])$. This means that the ϕ^a field is replaced by $\langle \phi^a(x) \rangle = -\int_M G^{ab}(x,y)[\psi, {}^\star\psi]_b$, where $G^{ab}(x,y)$ is the Green's function of the scalar covariant laplacian (Δ_A), in the computation of the correlation function of the observables. Since the functional

[†] The relation of these quantities to what we normally expect observables to be is discussed in section 2.2.

integral has support only on flat connections, if there are no B-fields in the observables (as is the case for our $\mathcal{O}$'s), the correlation densities reduce effectively to functions of ψ and the flat connections only. In order for correlation functions to be non-zero, the product of the observables – reduced in this way – must include all ψ zero-modes. For a genus $g \geq 2$ handlebody, this number is $(g-1)\dim(G)$. Let us now look at the various classes of correlation functions.

The vacuum expectation value of a single $\mathcal{O}$ is a topological invariant in the super-BF gauge theory. Hence, the gauge invariant operator (S_S is defined in appendix 2.7)

$$\hat{\mathcal{O}}(A;\mathcal{C}) \equiv \int [d\mu]_S e^{-S_S} \mathcal{O}(\phi,\psi,F;\mathcal{C}) \ , \tag{2.3.1.7}$$

has the property that its vacuum expectation value in the BF-gauge theory is a topological invariant. It is important to note that in general, $\hat{\mathcal{O}}$ depends on the background metric on M. Furthermore, although it is gauge invariant, it is not in the cohomology of the, $Q_{SBF} = Q^H + Q_{YM}$, total BRST charge (see appendix 2.7 for a discussion on Q^H), where Q_{YM} is the Yang-Mills BRST charge. Hence, its correlation functions will not be independent of the background metric, in general. Additionally, the $\hat{\mathcal{O}}$ will be non-local operators in general. Although these last two points may be viewed as drawbacks of this approach, there is one important lesson to be learned here. This construction clearly demonstrated that the three dimensional analogs of Donaldson invariants give rise to operators in the BF-gauge theories whose vacuum expectation values, in the latter theories, are themselves topological and are physical in three dimensional quantum gravity. It should also be noted that although the fields B, c, $\bar{c}$, c', $\bar{c}'$ appear in S_S, they do not survive the $[d\mu]_S$ integration due to ψ zero-mode saturation.

Until this point, we have only looked at the vacuum expectation values of the $\hat{\mathcal{O}}$'s. Now, we would like to investigate the expectation value of $\hat{\mathcal{O}}$ in any physical state of the BF-gauge theory. In particular, we would like to see whether or not such an expression is independent of the background metric, $g_{\mu\nu}$, used in forming the gauge fixed action. Let us suppose that such a state may be constructed out of the action of Wilson loop operators on the vacuum. Alternatively, we can ask whether or not the correlation function of the $\hat{\mathcal{O}}$'s with Wilson loop operators, $W[R,\gamma] = Tr_R P \exp\left(\oint_\gamma A\right)$, is background metric dependent. Hence we are led to study the functional integral

$$\begin{aligned} \mathcal{E}(R,\gamma,\mathcal{C}) &= \int [d\mu]_{SBF} e^{-S_{SBF}} \mathcal{O}(\phi,\psi,F;\mathcal{C}) W[R,\gamma] \\ &= \int [d\mu]_{BF} e^{-S_{BF}} \hat{\mathcal{O}}(A;\mathcal{C}) W[R,\gamma] \ . \end{aligned} \tag{2.3.1.8}$$

Functionally differentiating $\mathcal{E}(R,\gamma,\mathcal{C})$ with respect to the inverse metric, $g^{\mu\nu}$, we find

$$\frac{\delta\mathcal{E}(R,\gamma,\mathcal{C})}{\delta g^{\mu\nu}} = \int [d\mu]_{SBF} e^{-S_{SBF}} \Lambda_{\mu\nu} \mathcal{O}(\phi,\psi,F;\mathcal{C}) Tr_R P\left(\oint_\gamma \psi e^{\oint_\gamma A}\right) \ , \tag{2.3.1.9}$$

after use of the properties of S_{SBF} and where $\frac{\delta S_{SBF}}{\delta g^{\mu\nu}} = \{Q, \Lambda_{\mu\nu}\}$ with

$$\Lambda_{\mu\nu} = \frac{\delta}{\delta g^{\mu\nu}} \int_M (\lambda d_A^\star \psi + \lambda' d_A^\star \chi + \bar{c}' d_A^\star B + \bar{c}\delta A) \ . \tag{2.3.1.10}$$

We notice that the integral over ϕ' yields $\delta({\Delta_A}^{(0)}\lambda')$. Since we assume that ${\Delta_A}^{(0)}$ does not have any zero-modes then this restricts λ' to be zero. As a result of this, the only appearance of χ left is in the action. Integrating over this field we find $\delta(d_A\psi - {}^*d_A\eta)$. Now, the integrability condition for this restriction is $[F,\psi] = {\nabla_A}^{(0)}\eta$. However, as the integral over B can be seen to enforce $F=0$, we find that $\eta = 0$, hence $d_A\psi = 0$. This means that all ψ's in the path integral are now restricted to be zero-modes. For all but the first term in $\Lambda_{\mu\nu}$, the λ integration can be performed and it restricts each ϕ in the $\mathcal{O}$ to be replaced by an expression (see below) which depends on two ψ zero-modes. This then means that the path integrals involving each of the last three terms in $\Lambda_{\mu\nu}$ is saturated by ψ zero-modes due to the presence of $\mathcal{O}$. Thus, we see that the extra $\oint_\gamma \psi$ due to the Wilson loop makes those expressions vanish. We are then left with the first contribution for $\Lambda_{\mu\nu}$. If $\mathcal{O}$ depends on ϕ this will not be zero. Thus we deduce that $\frac{\delta\mathcal{E}(R,\gamma,\mathcal{C})}{\delta g^{\mu\nu}} = 0$, in general, only if the $\mathcal{O}$ does not depend on ϕ; otherwise, the only restriction on $\mathcal{O}$ is that it saturates the number of ψ zero-modes. Additionally, the result will not be altered if we included more than one Wilson loop in $\mathcal{E}$. Thus we conclude that the correlation functions of those $\hat{\mathcal{O}}$ operators whose ancestors – $\mathcal{O}$ – saturated the number of fermion zero-modes and are independent of ϕ, with Wilson loops is independent of the background metric.

Observables in the BF-gauge theory which depend on B have been constructed in the literature [22, 3]. An immediate observation is that if we compute correlation function of quantities which depend on B then the path integral is not restricted to $\mathcal{N}$. This invalidates the proof above. However, if we restrict to $M = \Sigma \times R$ (i.e canonical quantization), then there will be only dependence on $B|_\Sigma$ in the observables and the restriction to $F|_\Sigma = 0$ survives. In this case correlation functions involving A, B and $\hat{\mathcal{O}}(\psi)$ are gauge invariant and metric independent.

Haven given formal expressions for physical correlation functions in BF-gauge theories, we would now like to trace our steps back to the analysis at the beginning of this section and see how it might arise directly from super-BF gauge theories. Let us choose quantum gravity on a genus three handle body as a specific theory; thus, $g = 3$ and $G = SO(2,1)$. Six ψ zero modes are needed so we pick three homology 2-cycles which we label as Γ_i. Then we compute the correlation function $\left\langle \prod_{i=1}^{3} \int_{\Gamma_i} Tr(\psi\wedge\psi)\right\rangle_{SBF}$, in the super-BF gauge theory also with $g = 3$ and $G = SO(2,1)$. After integrating out the Y-fields, we obtain[‡]

$$\left\langle \prod_{i=1}^{3} \int_{\Gamma_i} Tr(\psi\wedge\psi)\right\rangle_{SBF} = (Z_{SBF})^{-1} \int [d\mu]_{BF,\alpha_1\ldots\alpha_6}\, e^{-(S_{BF}+S_{BF,gf})}\hat{\mathcal{O}}(A) \quad , \tag{2.3.1.11}$$

where

[‡]In general, the path integrations over the bosonic zero-modes are understood to drop out due to the division by Z_{SBF} in the expressions for the correlation densities.

$$
\begin{aligned}
\hat{O}^{\alpha_1\cdots\alpha_6}(A) \;=\; & T(A)\int_{\Gamma_1} Tr(\Upsilon^{[\alpha_1}(A)\wedge\Upsilon^{\alpha_2}(A)) \;\times \\
& \times \int_{\Gamma_2} Tr(\Upsilon^{\alpha_3}(A)\wedge\Upsilon^{\alpha_4}(A)) \;\times \\
& \times \int_{\Gamma_3} Tr(\Upsilon^{\alpha_5}(A)\wedge\Upsilon^{\alpha_6]}(A)) \;. \qquad (2.3.1.12)
\end{aligned}
$$

Here, the $\Upsilon^{\alpha_i}(A)$ form a six-dimensional basis for $H^1(M,G)$. By $[d\mu]_{BF,\alpha_1\ldots\alpha_6}$ we mean $[d\mu]_{BF}$ with the functional measure over flat connections, $A^{(0)}$ given by $[dA^{(0)}_{\alpha_1}]\cdots[dA^{(0)}_{\alpha_6}]$. The $A^{(0)}_{\alpha_i}$ and $\Upsilon^{\alpha_i}(A)$ are chosen to form a canonical basis for $T^*\mathcal{M}$ as in ref. [23]. As this expression was derived directly from the super-BF theory, the result is independent of the choice of basis for the fermionic zero-modes. The quantity $T(A)$ arises from the non-zero mode integration in $[d\mu]_S$. The remaining functional integral has support only on flat connections, hence $T(A)$ is ostensibly the Ray-Singer (R-S) torsion [24]. We then identify the $\Upsilon(A)$ as $\frac{\partial A}{\partial m}$. Notice that in our analysis of the BF-gauge theory at the beginning of this section, it was not evident that the R-S torsion appears as part of the observable's definition.

Now we realize that

$$
\left\langle \prod_{i=1}^{3}\int_{\Gamma_i} Tr(\psi\wedge\psi)\right\rangle_{SBF} \;=\; \left\langle \hat{O}(A)\right\rangle_{BF} \;. \qquad (2.3.1.13)
$$

Then interpreting $\hat{O}(A)$ as a correlation density in the BF-gauge theory we continue the computation to find

$$
\left\langle \hat{O}(A)\right\rangle \;=\; (Z_{SBF})^{-1}\int_{\mathcal{N}}\prod_{i=1}^{3}\int_{\Gamma_i} Tr(\Upsilon(A^{(0)})\wedge\Upsilon(A^{(0)})) \;, \qquad (2.3.1.14)
$$

where $\Upsilon(A^{(0)})$ is a form on the space, $\mathcal{N}$, of connections. The functional integral, $\int_{\mathcal{N}}$ over $\mathcal{N}$ is done with a wedge product of the Υ's, on that space, understood.

As a second example, we construct a correlation density in quantum gravity which is considerably less obvious in the BF theory than the prior example. We start with $\int_\gamma Tr(\phi\psi)$, here γ is a one-cycle. It carries ghost number three. Thus we construct a correlation density in quantum gravity on a genus $g\geq 2$ handle-body given as

$$
\hat{O}(A;\gamma_i) \;=\; (Z_{SBF})^{-1}\int [d\mu]_S e^{-S_S}\prod_i\int_{\gamma_i} Tr(\phi\psi) \;. \qquad (2.3.1.15)
$$

Integrating over λ we find that at the expense of a factor $\det^{-1}(\triangle^{(0)}_A)$, we should replace $\phi(x)$ by $-\int_{M_y} G_A(x,y)[\psi(y),{}^{\star}\psi(y)]$, where G_A is the Greens' function of the scalar covariant laplacian. Then functionally integrating over ψ we obtain

$$
\begin{aligned}
\hat{O}(A,\gamma_i) \;=\; & -(Z_{SBF})^{-1}T(A) \;\times \\
& \times \prod_i\{\int_{\gamma_i} Tr\{\int_{M_y} G_A(x_i,y)[\Upsilon(A(y)),{}^{\star}\Upsilon(A(y))]\Upsilon(A(x_i))\}\} \\
& \qquad (2.3.1.16)
\end{aligned}
$$

to be another correlation density in quantum gravity. In this expression, the Υ's appear anti-symmetrized as in (2.3.1.12).

Concluding this sub-section, we note one more point about the correlation densities we have been writing down. Unlike observables, our expressions are, in addition to being non-local in the BF-gauge theory, given in terms of path integrals. These functional integrals are best computed in perturbation theory. However, by invoking BRST theorems we were able to obtain some expressions non-perturbatively, in the above. It is safe to say that one lesson we have learned from this sub-section is that for diffeomorphism invariant theories, quantum gravity in particular, we must enlarge our scope of what an observable is. Here, we have used the geometry of the universal bundle and more directly the de Rham complex on moduli space to guide us. Presumably, this direction is worth a try in four dimensions also. We will turn to the latter in the next section. However, before that, we would like to discuss some even more profitable results; namely, expressions for wavefunctionals based on the universal bundle geometry.

2.3.2 Wavefunctionals

The physical Hilbert space of a super-BF gauge theory consists of L^2-functions on the moduli space, $\mathcal{M}$, of flat connections. In principle, quantization of this field theory is then reduced to quantum mechanics on $\mathcal{M}$. However, the pragmatism of such a program is limited as, a priori, it becomes unwieldy to pull such wavefunctions back into wavefunctional of the connection. In this sub-section we will demonstrate how this problem may be obviated. To be precise, we will write down expressions for the functionals of the connections on the G-bundle which are annihilated by the constraints of the theory.

Correlation functions of observables in TQFT's are equal to the integral over moduli space of a top form on that space [23]. Typically, such a top form is wedge product of forms of lesser degree:

$$\left\langle \prod_i^d \mathcal{O}_1 \right\rangle = \int_{\mathcal{M}} \Psi \quad , \qquad \Psi = f_1 \wedge f_2 \wedge \cdots \wedge f_d \quad , \tag{2.3.2.1}$$

where the forms, f_i, are obtained after integrating over the non-zero modes and fermionic zero-modes in the path integral. Now, let us assume that a metric exists on $\mathcal{M}$ so that we can define the Hodge dual map which we denote by the tilde symbol. Then $\tilde{\Psi}$ is a scalar function on moduli space. Let us now give representations for Ψ. All we seek is Ψ's which are gauge invariant and have support only on flat connections, ω, on Σ_g: $\Psi[\omega]$.

Clearly [25], a delta function, $\delta(F)$, where F is the curvature of a G-bundle over Σ_g satisfies our criteria. However, as it is highly unlikely to be normalizable. Regardless, we realize that it might be worthwhile to look at diffeomorphism invariant theories on Σ_g which are defined in a flat connection background. Considering the two-dimensional BF-gauge theory we see that the analog of B, φ, is an $ad(G)$-valued zero form and the action is $S^{2D}_{BF} = \int_{\Sigma_g} Tr(\varphi F)$. The delta function arises from the integral over φ. If we constructed the analog of S_{SBF}, we would find that it shares many of the terms which appear in the three dimensional action. However, here the analog of χ (we will call this ξ below) is a zero-form. What is more, there are no primed fields due to the degree of ξ. Considering this, we introduce the functional $\int_{\Sigma_g} Tr(\xi d_\omega \psi)$, for $ad(G)$-valued, Grassmann-odd zero- (ξ) and one- (ψ) forms defined in a flat connection background, ω, It is invariant under the local symmetry, $\delta\psi = \epsilon d_\omega \phi$ and upon gauge fixing it we obtain the quantum functional

$$S_\omega \;=\; \int_{\Sigma_g} Tr(\xi d_\omega \psi \;-\; \eta d_\omega^\star \psi \;-\; \lambda^\star \Delta_\omega \phi) \;\; . \tag{2.3.2.2}$$

The partition function for this action is metric independent as the part of S_ω which is metric dependent is exact with respect to the BRST charge for the gauge fixing of the symmetry just discussed. Furthermore, it is simple enough to compute exactly and is found to be equal to the Ray-Singer torsion of the G-bundle with flat connection, ω. In fact, the action S_ω is recognized as the action for a two-dimensional Grassmann-odd BF field theory in a flat connection background. As was the case with the super-BF gauge theory, the correlation functions of quantities such as $\frac{1}{2}Tr(\phi^2(x))$, etc., are topological invariants. This is seen to be due to the transformation given by the BRST charge: $\{Q, \psi\} = d_\omega \phi$. The partition function has support only on solutions of those ψ which are in $\ker(d_\omega)$. Hence, they span the cotangent space of $\mathcal{M}(\Sigma_g)$ whose dimension is $(2g-2)\dim(G)$.

As before, let us focus on three dimensional quantum gravity taking $G = SO(2,1)$. Our first example of a wavefunctional is found by taking the $(3g-3)$ times product of $\int_{\Gamma_i} Tr(\psi \wedge \psi)$ where the Γ_i are homology 2-cycles in Σ_g:

$$\Psi[\omega] \;=\; \int [d\xi][d\psi][d\eta][d\lambda][d\phi] e^{-S_\omega} \prod_{i=1}^{(3g-3)} \int_{\Gamma_i} Tr(\psi \wedge \psi) \;\; , \tag{2.3.2.3}$$

defined over the two-dimensional super-BF theory. The generic form of the wavefunctionals obtained by this construction is

$$\begin{aligned} \Psi_{\vec{n}}[\omega] \;\;&=\;\; \int [d\xi][d\psi][d\eta][d\lambda][d\phi] e^{-S_\omega} \prod_{i=1}^{n_4} Tr(\phi^2(x_i)) \;\times \\ &\times \prod_{j=1}^{n_3} \oint_{\gamma_j} Tr(\phi\psi) \prod_{k=1}^{n_2} \int_{\Gamma_k} Tr(\psi \wedge \psi) \; , \end{aligned} \tag{2.3.2.4}$$

subject to the condition $4n_4 + 3n_3 + 2n_2 = \dim \mathcal{M}(\Sigma_g, G)$. If ω is not an irreducible connection, then there are no ϕ zero-modes and the only non-zero $\Psi_{\vec{n}}[\omega]$ are those for which $n_4 = n_3 = 0$.

In the preceding, we have not used the full power of the two-dimensional super-BF gauge theory. As a matter of fact, we did not use it at all. The transition from S_ω to the super-BF gauge theory on Σ_g is straightforward. Its action is

$$S^{2D}_{SBF} \;=\; \int_{\Sigma_g} Tr(\varphi F) \;-\; S_\omega \;+\; \int_{\Sigma_g} Tr(\lambda[\psi, {}^\star\psi]) \;\; , \tag{2.3.2.5}$$

where φ is a zero-form which imposes the flat connection condition on ω and the rest of the action is reminiscent of the three dimensional theory but without the primed fields. Unlike the pure S_ω theory, the absence of ϕ zero-modes does not imply $\phi = 0$, but $\phi(x) = -\int_{\Sigma_{g,y}} G_A(x,y)[\psi(y),^{\star}\psi(y)]$ as we saw in the previous sub-section. Thus, more wavefunctionals result from this theory. They are of the same form as $\Psi_{\vec{n}}$ except that the functional measure must be enlarged to include all the fields in S^{2D}_{SBF}. Additionally, S_ω is replaced by S^{2D}_{SBF}. We find the general form of these wavefunctionals to be

$$\begin{aligned}\Psi^S_{\vec{n}}[\omega] &= (-)^{n_3}\prod_i^{n_4} Tr((\int_{\Sigma_{g,y}} G_\omega(x_i,y)[q(y),^{\star}q(y)])^2) \times \\ &\times \prod_j^{n_3}\oint_{\gamma_j} Tr(\int_{\Sigma_{g,y}} G_\omega(x_j,y)[q(y),^{\star}q(y)]q(x_j)) \times \\ &\times \prod_k^{n_2}\int_{\Gamma_k} Tr(q\wedge q)\ ,\end{aligned} \tag{2.3.2.6}$$

again with $n_4+n_3+n_2 = \dim(\mathcal{M}(\Sigma_g,G))$ and where the $q(\omega)$ form a $(2g-2)\dim(G)$ dimensional basis for $H^1(\Sigma_g,G)$. As was the case in eqn. (2.3.1.12), the q's appear in a totally anti-symmetric combination.

Our philosophy thus far has been to identify a Riemann surface (which is homeomorphic to the hypersurface of the foliated three-dimensional BF-gauge theory) and construct a servant partition function[§] for fields in a background which solves the constraints of the three-dimensional BF-gauge theory. Having done this we then identified operators which yield diffeomorphism invariant observables in the two-dimensional topological "theory". We assume that we can solve the equation which defines the constraints (as though they were classical equations) and parametrize them by the coordinates on moduli space. For example, the ω which defines the background above is really $\omega(x;m)$. This means that the wavefunctionals are not simply defined at one point in moduli space, but rather on all of $\mathcal{M}(\Sigma_g,G)$. We advocate this as a very robust approach to constructing quantum gravity wavefunctionals as (1) we need only solve the constraints classically and (2), thanks to our experience with TFT's it is rather straightforward to at least formally construct the servant partition functions and correlation functions in such a parametrized background.

Now, the fact that, in the previous sub-section, we were successful in formally constructing correlation densities leads to another possible approach to constructing wavefunctionals. If those correlation densities can be written, as operators, as $\hat{\mathcal{O}} = \hat{O}^\dagger\hat{O}$ for some operator $\hat{O}$ and adjoint, $\dagger$, then we would have $\left\langle\hat{\mathcal{O}}\right\rangle = \langle 0|\hat{O}^\dagger\hat{O}|0\rangle$. Interpreting this as the norm of a state $\hat{O}|0\rangle$, it is suggestive that the wavefunctional of such a state may be formed from the path integral expression for the correlation density. The manner in which we see this state arising is analogous to sewing in the super-BF gauge theory. Hence, we expect to be able to form the corresponding wavefunctional by surgery in the super-BF gauge theory. Although we delay detailed

[§]We will call these servant partition functions to distinguish them from the partition functions of the theories we are constructing the wavefunctionals for.

investigation of such an approach until a future publication, we would like to point out here that normalized wavefunctionals are expected. In the rest of this sub-section, we will show how to construct wavefunctionals from a super-BF gauge theory on a three manifold whose boundary is Σ_g.

We start with a super-BF gauge theory for a G-bundle over a three-dimensional manifold M with boundary Σ_g. Then we insert the pertinent operators as was done in the previous sub-section. Having done this, we choose a polarization (for which the fields in the geometrical sector appear as "position" variables) and perform all functional integrals with appropriate boundary conditions. This gives a wavefunctional for the super-BF gauge theory which is annihilated by the full BRST operator. It is also gauge and diffeomorphism invariant [23, 26]. In general, there may be fields which do not depend on the boundary values of the geometrical sector. Starting with the wavefunctional of the super-BF gauge theory, we integrate over their boundary values. This leads to a functional, $\Psi[\omega, \varpi]$, where ω is a flat connection on Σ_g, ϖ denotes a zero-mode of ψ and is a solution of $d_\omega \varpi = 0$ on Σ_g, and we have replaced the boundary value of ϕ with the appropriate expression in terms of ω and ϖ. Ψ also depends on the boundary values of c and c'; however, for notational simplicity they will be omitted.

Focusing our attention on genus-g handle-bodies, $\partial M = \Sigma_g$, we compute the correlation functions for the topological invariants fixing the boundary value of the connection to be a flat connection on Σ_g. This can be done by inserting a delta function, $\delta(U_I(A) - g_I(\omega))$ for each longitude, l_I in M. Here, $U_I(A)$ is the holonomy of the connection along the longitude, l_I, and $g_I(\omega)$ is the holonomy of a parametrized flat connection on the cycle, b_I, on Σ_g which (in the handlebody) is homotopic to l_I. For the ψ field, we insert delta functions $\delta(\oint_{l_I} \psi - \oint_{b_I} \varpi)$ where l_I and b_I are as before. Consequently, we arrive at our generic ansatz for wavefunctionals of super-BF gauge theories:

$$
\begin{aligned}
\Psi[\omega, \varpi] \;\; = \;\; & \int [d\mu]_{SBF} \prod_{I=1}^{(g-1)\dim(G)} \delta(U_I(A) - g_I(\omega)) \prod_{J=1}^{(g-1)\dim(G)} \delta(\oint_{l_J} \psi - \oint_{b_J} \varpi) \times \\
& \times \prod_i \mathcal{O}_i(\phi, \psi, F; \mathcal{C}_i)\; e^{-S_{SBF}} \quad . \qquad (2.3.2.7)
\end{aligned}
$$

In this expression, the product of polynomials, $\prod_i \mathcal{O}_i$ is such that it saturates the number of ψ zero-modes.

Now we must project out ϖ. This we do by treating $\Psi[\omega, \varpi]$ as a superfield and obtaining the wavefunctional of ω as a component via superfield projection. Choose a $(2g-2)\dim(G)$-dimensional basis, $q_\alpha(\omega)$, for $H^1(\Sigma_g, G)$ and expand the one-form field, ϖ, in it as $\varpi \equiv \sum_\alpha \theta^\alpha q_\alpha$, where the Grassmann-odd coefficients θ^α are the quantum mechanical oscillators. As the wavefunctional depends on $H^1(M, G)$ only $(g-1)\dim(G)$ of the θ^α will be non-zero. Then, Ψ is formally $\Psi[\omega, \theta]$ and we write,

$$
\Psi_{\alpha_1, \ldots, \alpha_n}[\omega] \;\; = \;\; \frac{\partial^n}{\partial\theta^{\alpha_n} \cdots \partial\theta^{\alpha_1}} \Psi[\omega, \theta]| \quad , \qquad (2.3.2.8)
$$

where the slash means setting $\theta = 0$ after differentiating. Each $\Psi_{\alpha_1, \ldots, \alpha_n}$ for $n = 1, \ldots, (g-1)\dim(G)$, is a wavefunctional in the BF-gauge theory in that it satisfies the constraints of the latter theory. We adopt this procedure as it is closest to the sewing/surgery procedure, is of geometrical significance (see below) and it incorporates the two naive guesses: setting $\varpi = 0$ or integrating out ϖ.

The closest analogy of these $\Psi_{\alpha_1,\ldots,\alpha_n}[\omega]$ is to Hartle-Hawking wavefunctionals [27]. Notice that unlike the previous wavefunctionals, $\Psi^S_{\vec{n}}$, in eqn. (2.3.2.6), which are analogous to canonical wavefunctionals, these are functionals only of half of the flat connections on Σ_g. This is due to the fact that the meridians of the handlebody are contractible in M.

Based on our discussion at the beginning of this sub-section, we realize that the $\Psi_{\alpha_1,\ldots,\alpha_n}$ are n-forms on moduli space. This returns us to our earlier discussion of the wavefunctionals of quantum gravity as being L^2-functions on moduli space. In writing down the $\Psi_{\alpha_1,\ldots,\alpha_n}$, we have given formal expressions for the pertinent functions on the moduli space in terms of the physical variable in the problem; namely, the connection.

The question remains which of them is normalizable. Although we do not have much to say about this question in this work, we would like to bring to the fore a possible strategy for normalizing wavefunctionals constructed in this way. Any closed piecewise linear three dimensional manifold, N, may we formed via the Heegaard splitting, $N = M_1 \cup_h M_2$, where M_1 and M_2 are two handlebodies whose boundaries are homeomorphic (with map, h) to each other [28]. Making use of this, we view [18] the norm of $\Psi[\omega, \theta]$ to be a functional integral on N, $\Psi[\omega, \theta]$ itself to be the functional integral on M_1 and its adjoint to be the functional integral on M_2. Reversing, we start with the functional integral for the correlation functions of polynomial invariants on N, perform surgery and then identify that component of Ψ which appears in the form $\Psi[\omega]^\dagger \Psi[\omega]$ after integrating $[d\mu]_S$. Obtaining the adjoint (†) is interpreted as arising from the surgery/sewing process.

2.4 Application to 4D Quantum Gravity

In this section, we sketch how the above construction can be realized in four dimension quantum gravity. The main ingredient that we have to supply is a TQFT whose action starts off as an Einstein-Hilbert action, or rather when some of the fields are put to zero one gets the usual action for gravity. Generally, we expect that there is more than one such action. Unlike the three dimensional case where quantum gravity was already defined over a finite dimensional phase space (flat connections modulo gauge transformation), in four dimension there are propagating fields and we can then try projecting the theory onto many different moduli spaces. In a way, the Einstein-Hilbert action is a good example for this construction, Whereas it is non-renormalizable, the topological projection takes us to a renormalizable theory in which to do calculations on physical observables for quantum gravity. In addition, we can construct formal, diffeomorphism invariant expressions¶ without integration

¶We remind the reader that in constructing observables, we work in covariant – not canonical – quantization.

over the whole manifold.

Our first task is to select a TQFT. The logical candidates are topological gravity theories. Four dimensional topological gravity theories for which the pure metric part of the action is given by the square of the Weyl tensor, not the Einstein-Hilbert action, are known [29]. As mentioned above, we seek a topological gravity theory whose pure metric action is the Einstein-Hilbert action. Now, TQFT's may be obtained from supersymmetric theories via twisting [23]. The fact that the four dimensional gravity theories which were first constructed were conformal is apparently correlated with the fact that N=2 supergravity in four dimensions has this feature. A Poincaré supergravity theory was proposed sometime ago by de Wit|| [30]. Thus we expect that a twisted version of this should exist as a topological gravity theory.

In ref. [31], a topological gravity theory with Einstein-Hilbert action as is pure metric part was obtained by twisting a N=2 supergravity theory. Here, we will simply use the results of this work. The twisting procedure defined a Lorentz scalar, Grassmann-odd (BRST) charge, Q which is nilpotent. As it turns out this topological gravity theory is seen to be the TQFT for the projection of the spin-connection form (in the second order formulation) to be self-dual:

$$w^{-ab}(e) \ = \ 0 \ , \qquad (2.4.0.1)$$

where a etc. are Lorentz indices.

The observables are constructed from the cohomology of Q. After some re-definitions of the fields one ends up with a BRST charge whose action upon the geometrical sector of the theory is

$$\begin{aligned}
&Q : e^a \rightarrow \psi^a \ - \ \mathcal{D}\epsilon^a \ + \ \epsilon^{ab} \wedge e_b \ , \\
&Q : w^{ab} \rightarrow \chi^{ab} \ - \ \mathcal{D}\epsilon^{ab} \ , \\
&Q : \psi^a \rightarrow -\mathcal{D}\phi^a \ - \ \eta^{ab} \wedge e_b \ - \ \chi^{ab} \wedge \epsilon_b \ + \ \epsilon^{ab} \wedge \psi_b \ , \\
&Q : \phi^a \rightarrow \epsilon^{ab} \wedge \phi_b \rightarrow \eta^{ab} \wedge \epsilon_b \ , \\
&Q : \chi^{ab} \rightarrow -\mathcal{D}\eta^{ab} \ + \ \epsilon^{ac} \wedge \chi_c^{\ b} \ - \ \chi^{ac} \wedge \epsilon_c^{\ b} \ , \\
&Q : \eta^{ab} \rightarrow \epsilon^{ac} \wedge \eta_c^{\ b} \ - \ \eta^{ac} \wedge \epsilon_c^{\ b} \ , \\
&Q : \epsilon^a \rightarrow \phi^a \ + \ \epsilon^{ab} \wedge \epsilon_b \ , \\
&Q : \epsilon^{ab} \rightarrow \eta^{ab} \ + \ \epsilon_c^{\ a} \wedge \epsilon^{cb} \ , \qquad (2.4.0.2)
\end{aligned}$$

where ϵ^{ab} and ϵ^a are the ghosts for Lorentz and diffeomorphism symmetries, respectively. As was mentioned above all this is in second order formalism. Although in the BRST transformations all the fields look independent, this is not the case. However, according to [31], these transformations are consistent with the conditions of the second order formalism : $w^{ab} \wedge e_b = de^a$, $\chi^{ab} \wedge e_b = -\mathcal{D}\psi^a - R^{ab} \wedge \epsilon_b$, etc...

The cohomology is constructed exactly as in Donaldson-Witten theory [3] and observables are found. For example,

$$\mathcal{O}^{(4)} \ = \ \frac{1}{2} Tr(\eta^2) \ ,$$

||We thank S. J. Gates, Jr. for bring this to our attention.

$$
\begin{aligned}
\mathcal{O}^{(3)} &= \int_f Tr(\eta\chi) \ , \\
\mathcal{O}^{(2)} &= \int_\Gamma Tr(\eta R + \frac{1}{2}\chi\wedge\chi) \ , \\
\mathcal{O}^{(0)} &= \int_M Tr(R\wedge R) \ . \qquad (2.4.0.3)
\end{aligned}
$$

The geometrical meaning of each field is in pure analogy with those in Donaldson-Witten theory. As in the three dimensional case, the correlation functions of these observables become, after integration over the non-zero-mode parts of the fields, functions of the zero modes of χ^{ab} and the veirbein, e^a. The zero-modes of χ are a basis for the cotangent space to the moduli space of $w^{-ab}(e) = 0$.

Now that we have described the results of [31] we would like to indicate how the constructions of the previous sections can be applied here. To construct the observables we insert, in the path integral a combination of the TQFT observables which saturates the χ zero-modes and then integrate over all the fields except e, ϵ^a and ϵ^{ab}. This will result in a non-local expression in terms of the veirbein whose expectation value is a topological invariant of space time. In this way, we obtain $\hat{\mathcal{O}}(e)$ from the $\mathcal{O}^{(k)}$ above.

The wavefunctional construction is also very similar. However, at present, we can only construct the Hartle-Hawking type wavefunctionals, as unlike the 3D case, we do not have the corresponding servant action. Defining the action to be on an four dimensional manifold, M, with boundary $\partial M = \Sigma$, we obtain wavefunctionals by following the same steps we took in the three dimensional case. This results in a functional $\Psi[e|_\Sigma]$. The proof that the wavefunctional so constructed satisfies the constraints of quantum gravity is now the same as that of Hartle-Hawking [27]. We differ the exact results, in particular, a presentation of the normalization procedure, to a future work [18].

2.5 Conclusions

In this work we have indicated a possible way of using TQFT's to construct wavefunctionals and physical correlation functions in three and four dimensional quantum gravity. We gave explicit results in the three dimension case and laid the building blocks for the construction in four dimensions. Along the way we have shown that in quantum gravity, there are functions of the fields whose vacuum expectation values are not only diffeomorphism invariants of spacetime, but also of geometrical significance on moduli space. A possible definition of an inner product was mentioned and will be elaborated in [18]. This work also indicates that it might be useful to consider quantum gravity in a larger geometrical setting than usual.

In concluding, it is tempting to speculate that pursuit along the lines advocated in this work may lead to possible field theoretic relations between intersection numbers on a manifold and the intersection numbers on the moduli space of field configurations for sections of bundles over that manifold. In particular, we know that the Wilson loop observables in BF-gauge theories [32] and the loop observables [9] construct knot invariants. Well, we have found the projections of Donaldson-like polynomial invariants into three and four dimensional quantum gravity. From a purely field theoretical point-of-view, we then expect to find a relation between these two sets of operators.

Appendices

2.6 A. 3D BF Gauge Theory

Recall [3] that the actions of BF-gauge theories are topological and of the form

$$S_{BF} \;=\; \int_M Tr(B \wedge F) \quad , \tag{2.6.4}$$

where B is an $ad(G)$-valued $(n-2)$-form on the oriented, closed n-manifold, M and F is the curvature of the G-bundle whose connection is A. Path integrals for these theories with the insertion of any operators except those composed of B have support only on flat connections. The wavefunctionals for these theories reduce to L^2-functions on the moduli space, $\mathcal{M}$, of flat connections. For purposes of path integral quantization, the partition function of the BF-gauge theoryis

$$Z_{BF} \;=\; \int [dA][dB][d\bar{c}][dc][d\bar{c}'][dc'][db][db']e^{-[S_{BF}+S_{BF,gf}]} \quad , \tag{2.6.5}$$

where

$$S_{BF,gf} \;=\; \int_M Tr(bd^\star A \;+\; b'd^\star_A B \;-\; \bar{c}d^\star d_A c \;-\; \bar{c}'d^\star d_A c') \quad . \tag{2.6.6}$$

The last action represents the projection of the connection into the Lorentz gauge and the removal of the covariantly exact part of B; all done by means of the symmetries of the BF-gauge theory. In this gauge fixing, the $c(c')$ and $\bar{c}(\bar{c}')$ fields are the zero-form, anti-commuting ghosts and anti-ghosts for the $A(B)$ projections, respectively.

Canonical quantization on $M = \Sigma \times \boldsymbol{R}$ immediately leads to the constraints [33],

$$^\star d_A B \;\approx\; 0 \quad , \qquad ^\star F \;\approx\; 0 \quad , \tag{2.6.7}$$

where the Hodge dual here is defined on Σ and is induced from that on M. The first of these constraints is Gauss's law enforcing the gauge invariance of physical states and the second requires that these states have support only on flat connections. In this special case of 2+1 dimensional quantum gravity, it can be shown [20, 12] that on physical states $Diff(\Sigma)$ is equivalent to these constraints.

2.7 B. 3D Super-BF Gauge Theory

The action for super-BF gauge theory [3, 34] is

$$\begin{aligned} S_{SBF} \;=\; \int_M Tr\{B \wedge F \;&-\; \chi \wedge d_A \psi \\ &+\; \eta d^\star_A \psi \;+\; \lambda^\star \triangle_A \phi \;+\; \lambda[\psi, {}^\star\psi] \\ &+\; \eta' d^\star_A \chi \;+\; \lambda'^\star \triangle_A \phi' \;+\; \lambda'[\psi, {}^\star\chi]\;\} \quad . \end{aligned} \tag{2.7.8}$$

All fields are $ad(G)$ valued and their form degree, Grassmann-parity and fermion/ghost number are listed in the following table:

FIELD	DEGREE	G-PARITY	GHOST #
B	1	even	0
A	1	even	0
χ	2	odd	-1
ψ	1	odd	1
η	0	odd	-1
η'	0	odd	1
λ	0	even	-2
ϕ	0	even	2
λ'	0	even	0
ϕ'	0	even	0

Placing this set of fields in the context of section 2.2, the BF-gauge theory is the GFT and super-BF gauge theory is the TQFT. The sets of fields are represented by $X = (B, A)$ with Y being the rest of the fields in this table.

The (Yang-Mills) gauge invariant action (2.7.8) may be obtained by starting with the zero lagrangian and gauge fixing the topological symmetry $\delta A = \epsilon\psi$ to the flat connection condition, $F = 0$. It is invariant under the horizontal BRST transformations

$$\begin{aligned}
&Q^H : A \to \psi \ , \qquad Q^H : \psi \to d_A\phi \ , \\
&Q^H : \chi \to B \ + \ d_A\phi' \ , \qquad Q^H : B \to [\chi, \phi] \ + \ [\phi', \psi] \ , \\
&Q^H : \lambda \to \eta \ , \qquad Q^H : \eta \to [\lambda, \phi] \ , \\
&Q^H : \lambda' \to \eta' \ , \qquad Q^H : \eta' \to [\lambda', \phi] \ .
\end{aligned} \tag{2.7.9}$$

Additionally, it is invariant under the one-form symmetry $\delta B = d_A\Lambda$. The gauge fixing of these symmetries introduces the usual ghost "kinetic" terms plus some new terms which involve Yukawa-like couplings with ψ and B. We will return to these later. It is worthwhile to note that the action, S_{SBF} may be written as the action of a BF-gauge theory plus "supersymmetric" completion term as**:

$$S_{SBF} \ = \ S_{BF} \ + \ S_S \ . \tag{2.7.10}$$

The partition function for super-BF gauge theory is (see appendix 2.9 for our notation)

$$Z(M) \ = \ \int [d\mu]_{SBF} e^{-S_{SBF}} \ . \tag{2.7.11}$$

Integrating over B we see that this partition function has support only on flat connections as is the case with the BF-gauge theories. However, ψ and ϕ have made their appearances. The observables [3] of this theory are elements of the Q^H-equivariant cohomology and are maps from $H_*(M)$ to $H^*(\mathcal{M})$. For rank two groups they are constructed as polynomials of the following homology cycle integrals:

$$\mathcal{O}^{(4)} \ = \ \frac{1}{2} Tr(\phi^2) \ ,$$

**We use S_S heavily in the text.

$$\begin{aligned}
\mathcal{O}^{(3)} &= \int_\gamma Tr(\phi\psi) \ , \\
\mathcal{O}^{(2)} &= \int_\Gamma Tr(\phi F + \frac{1}{2}\psi\wedge\psi) \ , \\
\mathcal{O}^{(1)} &= \int_M Tr(\psi\wedge F) \ .
\end{aligned} \tag{2.7.12}$$

In these expressions, the index (k) represents the fact that $\mathcal{O}^{(k)}$ is a k-form on moduli space or carries ghost number k in the BRST language. These are the three dimensional analogs of the Donaldson invariants which may be constructed in four dimensional topological Yang-Mills theory [23].

2.8 C. Special Topology for BF: $\Sigma_g \times S^1$

In this appendix, we would like to suggest the possible existence of polynomial invariants in the pure BF-gauge theory if the three-dimensional manifold is taken to be the Lens space $S^2 \times S^1$ or $\Sigma_g \times S^1$, in general††. In what follows we will assume the temporal gauge. However, this is not completely possible due to the holonomy of the gauge field in the S^1 direction. It is for this reason that our discussion is only suggestive. Nevertheless, see ref. [36] in which an explicit demonstration of the relation between Chern-Simons theory and G/G WZW theory on Σ_g is given.

Expand all of the fields in the harmonics of S^1, $e^{in\theta}$, (where θ is the coordinate on S^1 and n is an integer) symbolically as $\Phi(\Sigma_g \times S^1) \equiv \sum_n \Phi_{(n)}(\Sigma_g)e^{in\theta}$. Then choose the "temporal gauge" so that the connection in the S^1 direction vanishes leading to the action‡‡

$$\begin{aligned}
S_{BF} = \int_{\Sigma_g} Tr\Big(& \sum_n \phi_{(n)} dA_{(-n)} + \sum_{n,m} \phi_{(n)} A_{(m)} \wedge A_{(-n-m)} \\
& - i\sum_n nB_{(n)}A_{(-n)} + \sum_n n\bar{c}_{(n)}c_{(-n)} \Big) \ ,
\end{aligned} \tag{2.8.13}$$

where $\phi_{(n)}$ are the components of the original B field in the S^1 direction. Realizing that the $B_{(n)}$ for $n \neq 0$ does not appear in a term with derivatives, we integrate it out of the action finding

$$S_{BF} = \int_{\Sigma_g} Tr\Big(\phi F + \sum_n n\bar{c}_{(n)}c_{(-n)}\Big) \ , \tag{2.8.14}$$

where $\phi \equiv \phi_{(0)}$ and F is the curvature on Σ_g constructed out of $A_{(0)}$. With the exception of the completely decoupled fermionic term, this is the action for two dimensional BF theory. It has been recently studied quite extensively [35, 26]. Notice that we did not obtain it via compactifying the S^1 direction. That direction simply decouples due to the first order and off-diagonal nature of the gauge theory. Extending the theory

††Here, as in the text, Σ_g is a genus g Riemann surface.

‡‡In this appendix, ϕ is not the scalar field in the super-BF gauge theory considered in the body of the paper.

to incorporate an equivariant cohomology is possible, however, we will not need this in order to obtain our results.

From the partition function for this action, it is easy to show that gauge invariant functions of ϕ will be invariant under diffeomorphisms. As $\mathcal{L}_K\phi = i_k d_A\phi + [\phi, \alpha(K)]$, we must show that $\langle d\mathcal{O}(\phi)\rangle = 0$, where $\mathcal{O}$ is some gauge invariant function constructed only out of ϕ. This equality follows from the statement that $d_A\phi$ is obtained by varying the action with respect to the connection, thus its expectation value and/or correlation function with any other functions of ϕ is a total functional derivative on $\mathcal{A}$; hence it vanishes. Another way to see this result is that a symmetry of the action (2.8.13) exists in which ϕ may be shifted into a $\bar{c}_n$ (or c_n). This symmetry, however, does not affect the ϕ zero-mode (solution of $d_A\phi = 0$) as it does not appear in the action. The exponent in definition of $\mathcal{B}_I$ is Q^H exact and since it is constructed to be gauge invariant it is also Q^H closed.

2.9 D. Notation

The symbol, G is used to denote a semi-simple Lie group. The space of gauge connections is written as $\mathcal{A}$. Our generic notation for spacetime manifolds is M while we use the symbol, $\mathcal{M}(M,G)$ for the moduli space of specific (e.g., flat) connections for the G-bundle, P, over M. The exterior derivative on M is d while the covariant exterior derivative with respect to the connection A is d_A. Coordinates are M are written as x, y, etc. while coordinates on $\mathcal{M}(M,G)$ denoted by m. The gauge covariant laplacians on k-forms are written as $\triangle_A^{(k)}$. The genus of a handlebody/Riemann surface is g. Any metrics which appear explicitly are written with indices or otherwise in an obvious manner. While γ denotes a homology one-cycle, Γ stands for a homology two-cycle. The functional measures used are defined in the following table:

NOTATION	MEASURE	ACTION
$[d\mu]_{BF}$	$[dA][dB][d\bar{c}][dc][d\bar{c}'][dc'][db][db']$	$S_{BF} + S_{BF,gf}$
$[d\mu]_S$	$[d\chi][d\psi][d\eta][d\eta'][d\phi][d\lambda][d\phi'][d\lambda']$	S_S
$[d\mu]_{SBF}$	$[d\mu]_{BF}[d\mu]_S$	S_{SBF}

All other notations are established in the text. Except note that our field notations for the TQFT's are not the same in three and four dimensions (see section 2.4).

Chapter 3

A Note on the Semi-Classical Approximation in Quantum Gravity

On Tuesday, when it hails and snows,
The feeling on me grows and grows
That hardly anybody knows
If those are these or these are those.
(The World of Pooh by, A. A. Milne)

Abstract:

We re-examine the semiclassical approximation to quantum gravity in the canonical formulation. It is shown that the usual interpretation of a WKB state does not give an adequate semiclassical description of both matter and gravity degrees of freedom. A state for the gravitational field is proposed which has the necessary properties to describe quantum field theory on a background spacetime with small quantum fluctuations. Its connection with WKB states is clarified using a reduced phase space formalism. This state is used to give a qualitative analysis of the effects of geometry fluctuations, which can be related to the breakdown of the semiclassical approximation near a black hole horizon.

This chapter is based on [37]

3.1 Introduction

Although we do not yet have a consistent quantum theory of gravity, there is a simple requirement that can be placed on any sensible theory: It should be capable of describing quantum matter fields interacting with an essentially classical gravitational field in some limit. This is known as the semiclassical limit of quantum gravity. In some formulations of quantum gravity, such as perturbation theory for the linearized field around a given background, quantum field theory in curved spacetime is built in. In non-perturbative approaches, however, where spacetime is only a derived concept, it is useful to see how this limit can be obtained.

Some considerable work towards understanding the semiclassical limit has been done in the canonical approach to quantum gravity using the ADM formulation [38]. In this approach there is no background spacetime, since the dynamical variables are the 3-metrics of spacelike hypersurfaces, plus the matter fields on these hypersurfaces. Employing the Dirac procedure, physical states must be annihilated by the momentum and Hamiltonian constraints. The momentum constraints reduce the phase space to the space of all 3-geometries (the equivalence classes of 3-metrics under spatial diffeomorphisms). The Hamiltonian constraint is imposed by the Wheeler–DeWitt equation which has the effect of factoring out translations in the time direction, and is a direct analogue of the Klein-Gordon equation in the quantized relativistic particle.

The semiclassical limit is obtained by expanding the Wheeler–DeWitt equation in powers of the gravitational coupling constant G. To first order, a perturbative solution yields a WKB approximation

$$e^{iS_{HJ}[h_{ij}]/\hbar G}$$

to the gravitational Wheeler–DeWitt equation (where $S_{HJ}[h_{ij}]$ is a Hamilton–Jacobi function for general relativity). As was shown by Lapchinski and Rubakov, and later by Banks [39], the next order approximation is obtained by solving the functional Schrödinger equation for a matter state on a set of eikonal trajectories corresponding to the set of solutions of Einstein's equations given by $S_{HJ}[h_{ij}]$. The approach of Refs. [39] focussed on deriving the Schrödinger equation for matter variables propagating on a classical background, without giving a careful treatment of the gravitational degrees of freedom. More recently, attempts have been made to interpret the entire state, including both matter and gravity [40, 41]. It has been suggested that a solution of the gravitational WKB equation could be interpreted as describing a statistical ensemble of classical spacetimes, each of which acts as a background for quantum matter fields (see [2] for a review of this point). Some reservations have been expressed about this interpretation [42], and it seems that simple WKB eikonals may not be adequate. We add simple arguments explaining why a first order WKB state cannot be used to describe the state of a classical gravitational field of the kind we observe, and why it is not consistent to regard the ensemble as statistical in any sense. The question we then have to address is whether it is possible to find a modification of a first order WKB state that can consistently describe a *single* classical spacetime, but that still leads to the functional Schrödinger equation.

It is useful to recall how a state representing a single spacetime can arise from a superposition of WKB type states, a point of view expressed originally by Gerlach [43], but which appears to have been discarded in the standard semiclassical WKB interpretation. The importance of this is most clearly understood in a reduced phase space language, which can be derived from standard Hamilton–Jacobi theory for the gravitational field. Using this language, it is evident that a state which admits a classical interpretation must be well localized in the constants of the motion that classify different classical spacetimes. The operators corresponding to these variables occur in canonically conjugate pairs, and so do not all commute. Thus Gerlach's states can be understood as coherent states with respect to these observables. The important result is that superpositions of this kind can be approximated by gravitational WKB states with a judicious choice of the WKB prefactor. As a consequence, the functional Schrödinger equation for the matter state still appears in the second order expansion of the Wheeler–DeWitt equation with this choice of state to describe the gravitational field. It follows that a consistent semiclassical approximation exists that describes quantum fields propagating on a single fixed background.

Identifying the correct superposition of first order WKB states that leads to a consistent semiclassical interpretation is essential in order to discuss the Wheeler–DeWitt equation beyond the semiclassical approximation. A coherent superposition can be used to demonstrate the absence of quantum gravity effects in macroscopic physics. It can also be used to predict the situations in which the semiclassical approximation breaks down, an example being the breakdown of the semiclassical picture close to a black hole horizon described in chapter 5.

The study of the semiclassical limit of quantum gravity is of interest for a variety of reasons. The first question that one should ask in quantum gravity is how the 3+1 dimensional spacetime emerges from the wavefunction on 3-geometries. The usual approach is then to ask how matter states are to be defined, how these form a Hilbert space, and what ought to be an inner product on such states. In that case it is essential to identify a single background spacetime on which a many fingered time function $\tau(x)$ is defined; the fluctuations of gravity around this mean metric are then included as normal degrees of freedom (gravitons) on the same footing as the matter degrees of freedom.

Here we shall consider a somewhat different construction, focussing on what kind of a solution to the Wheeler–DeWitt equation best represents an approximately classical spacetime with matter propagating on it. It is unclear how one would determine what wavefunction the Universe (or a region thereof) would have from first principles. However, one usually imagines that the process of decoherence leads to an effective description where semiclassical gravity is valid. Thus for the matter observables of interest one imagines that out of all 3 geometries on which the wavefunctional has support, only those close to a mean 3+1 dimensional spacetime effectively contribute. Thus it makes sense to construct a wavefunction that directly describes such a semiclassical limit. Here not only must the gravity look classical, but the matter must evolve to a good approximation by the Schrödinger equation on this metric. We discuss in this paper how such a wavefunction should look.

In Sec. 2 we present a review of the expansion of the Wheeler–DeWitt equation in powers of the Planck mass, and discuss the pros and cons of standard interpretational framework that accompanies this expansion. In Sec. 3, we discuss the WKB state in quantum gravity in terms of physical (non-gauge) degrees of freedom, in order to show why a coherent superposition of first order WKB states is the quantum state of the gravitational field that exhibits classical behaviour. In Sec. 4, we show how a superposition of WKB states is approximated by a single WKB state, where information about superpositions is contained in the WKB prefactor. In this way a superposition of WKB states leads to the functional Schrödinger equation within the usual perturbation expansion. Sec. 5 contains a discussion of higher order corrections to the semiclassical approximation, with the aim of understanding when and how this approximation can break down. Finally, in Sec. 6, we review some results from a 1+1 dimensional model which serve to to illustrate the rather formal discussion of Sec. 3.

3.2 The semi-classical approximation

Expanding the Wheeler–DeWitt and momentum constraint equations order by order in the gravitational coupling constant G leads to a WKB approximation. The aim of this approximation is to find a quantum state that represents both a classical background spacetime and a quantum field propagating on that background. In this section, we shall give a brief review of some of the large amount of work on this subject [39, 40, 41].

We ignore the details of the momentum constraint in the following discussion, and assume that spatial diffeomorphism invariance is imposed at all orders. We do not need to specify whether we are working with open or closed spatial topology. Although we shall generally assume that spacelike hypersurfaces are compact, this discussion can be easily generalized to spacetimes with well-defined asymptotics (see for example Sec. 6.2).

The Wheeler-DeWitt equation reads

$$-16\pi G\hbar^2 G_{ijkl}\frac{\delta^2\Psi[f,h]}{\delta h_{ij}\delta h_{kl}} - \frac{\sqrt{h}R}{16\pi G}\Psi[f,h] + H_{matter}\Psi[f,h] = 0 \tag{3.2.1}$$

taking $c = 1$. Consider expanding the state Ψ as

$$\Psi = e^{i(S_0/G+S_1+GS_2+\cdots)/\hbar} \tag{3.2.2}$$

where each S_i is assumed to be of the same order. Eq. (3.2.1) can then be expanded perturbatively in G. The first order equation simply states that $S_0[f,h] = S_0[h]$ is independent of the matter degrees of freedom. At the next order, we find that $S_0[h]$ must be a solution of the general relativistic Hamilton–Jacobi equation [43, 46]

$$\frac{1}{2}G_{ijkl}\frac{\delta S_0}{\delta h_{ij}}\frac{\delta S_0}{\delta h_{kl}} - 2\sqrt{h}R = 0. \tag{3.2.3}$$

There is not a unique solution to the Hamilton–Jacobi equation (3.2.3), so there is some freedom in the choice of $S_0[h]$, as we shall discuss in detail below.

At the next order, we obtain an equation for $S_1[f,h]$. It is convenient to split $S_1[f,h]$ into two functionals $\chi[f,h]$ and $D[h]$, so that

$$\frac{1}{D[h]}e^{iS_0[h]/\hbar G} \tag{3.2.4}$$

is the next order full WKB approximation to the purely gravitational Wheeler–DeWitt equation. The equation for the WKB prefactor $D[h]$ is

$$G_{ijkl}\frac{\delta S_0[h]}{\delta h_{ij}}\frac{\delta D[h]}{\delta h_{kl}} - \frac{1}{2}G_{ijkl}\frac{\delta^2 S_0[h]}{\delta h_{ij}\delta h_{kl}}D[h] = 0. \tag{3.2.5}$$

The remaining condition on $\chi[f,h]$,

$$i\hbar G_{ijkl}\frac{\delta S_0}{\delta h_{ij}}\frac{\delta\chi[f,h]}{\delta h_{kl}} = H_{matter}\chi[f,h] \tag{3.2.6}$$

is an evolution equation for the functional $\chi[f,h]$ on the whole of superspace (the space of all 3-geometries). Its solution requires initial data for $\chi[f,h]$ on a surface in superspace.

Eq. (3.2.6) is closely related to the functional Schrödinger equation. Having specified the initial data, it can be solved by the method of characteristics, by restricting to eikonal tracks on superspace. The tracks are specified by solving the classical equations

$$\pi^{ij} = \delta S_0[h]/\delta h_{ij} \qquad \text{or} \qquad \frac{dh_{ij}(\mathbf{x},\tau)}{d\tau} = -2N(\mathbf{x},\tau)K_{ij}(\mathbf{x},\tau) + \nabla_{(i}N_{j)}(\mathbf{x},\tau) \tag{3.2.7}$$

which give the family of solutions of Einstein's equations defined by the Hamilton–Jacobi function $S_0[h]$. The solution of Eq. (3.2.7) requires a choice of integration constants and a choice of lapse and shift functions $N(\mathbf{x},\tau)$ and $N_i(\mathbf{x},\tau)$. The integration constants specify different classical spacetimes while the lapse and shift are just choices of co-ordinates on each of these spacetimes. Along each characteristic, Eq. (3.2.6) becomes the functional Schrödinger equation

$$i\hbar\frac{\delta\chi[f,h]}{\delta\tau} = H_{matter}\chi[f,h] \tag{3.2.8}$$

where τ is the time parameter corresponding to the chosen foliation.

There are two potential integrability conditions to worry about when solving Eq. (3.2.6) using the method of characteristics. Firstly, Eq. (3.2.7) can be integrated using different lapse and shift functions, corresponding to using different co-ordinates on the background spacetime. We expect (3.2.8) to be covariant under changes of co-ordinates, but this is not always the case. Integration with different lapse and shift functions can lead to ambiguities in the definition of $\chi[f,h]$, as has been discussed by various authors [105, 48, 49]. We shall ignore this problem here. Secondly, there is the question of different integration constants in the solution of Eq. (3.2.7), corresponding to integration of (3.2.6) along different classical spacetimes. Formally this causes no problems, since in general there is at most one solution to Einstein's equations that passes through any point in superspace and is compatible with a given Hamilton-Jacobi function $S_0[h]$. However, as we shall see in Sec. 5, this is an important issue when considering corrections to the semiclassical approximation.

It has been argued that although (3.2.4) has support throughout superspace, its Wigner function is peaked around configurations where $\pi^{ij} = \delta S_0/\delta h_{ij}$ (which defines the ensemble of classical solutions (3.2.7)) [40, 41]. This has led to the suggestion that the WKB state

$$\Psi_1 = \frac{1}{D[h]} e^{iS_0[h]/\hbar G} \chi[f, h] \tag{3.2.9}$$

describes an ensemble of semiclassical spacetimes, where the WKB prefactor $D[h]$ gives a measure on the ensemble of solutions, suggesting a statistical interpretation.

Since each eikonal trajectory defined by $S_0[h]$ is exactly classical, it does not make sense to regard (3.2.9) as describing an ensemble. Even if some form of decoherence is invoked, it is impossible for a quantum state to describe an ensemble of strictly classical spacetimes. It might at first sight appear that fixing an initial 3-geometry will select one characteristic from this ensemble and then any subsequent question about the matter state depends only on integration along that characteristic. However, the observation of an initial 3-geometry is not compatible with a state that has support on the whole of superspace since it cannot leave that state unperturbed†. In order for a wavefunctional to describe a classical background, it should have expectation values for observables associated with classical quantities that are well defined with small quantum spreads. Observables associated with picking out a particular characteristic certainly fall into this category. This suggests that a quasiclassical state for the gravitational field should have support on a narrow tube in superspace around a classical spacetime, much like a quasiclassical state in quantum mechanics.

3.3 The WKB state in quantum gravity

In this section, we shall focus on the gravitational WKB state. We present a review of Hamilton-Jacobi theory as applied to the gravitational field, which clarifies the interpretation of WKB states.

3.3.1 Hamilton–Jacobi theory and WKB states

Let us begin by considering the general relativistic Hamilton-Jacobi equation (3.2.3). In order to specify a solution of (3.2.3), it is necessary to supply a series of constants of integration which are usually called α–parameters in Hamilton–Jacobi theory (see for example Ref. [50]). Any solution S takes the form $S[h_{ij}(x), \alpha]$, where α represents an infinite number of integration constants (equivalent to two field theory degrees of

†The idea that any observation dramatically perturbs the gravitational state is somewhat out of line with a semiclassical interpretation. After all, if this were to happen, the interesting question would then be to identify the perturbed state, which would be the one giving semiclassical behaviour.

freedom in 3+1 dimensions [2]). Given a Hamilton–Jacobi functional $S[h,\alpha]$, the relation

$$\pi^{ij} = \frac{\delta S[h_{ij},\alpha]}{\delta h_{ij}} \tag{3.3.10}$$

gives the momenta conjugate to h in terms of h and α. Eqs. (3.3.10) are a set of first order differential equations (c.f. Eqs. (3.2.7)), that yield solutions to Einstein's equations, but require a further set of integration constants to pick out a particular solution. Alternatively, a classical solution can be fixed by defining a set of constants

$$\beta = \frac{\delta S[h_{ij},\alpha]}{\delta \alpha}, \tag{3.3.11}$$

which are precisely the integration constants for (3.3.10), and then solving for h_{ij} (the set of all h_{ij} that satisfy this equation form a track in superspace defining a solution of Einstein's equations). From either (3.3.10) or (3.3.11) it follows that a single solution of Einstein's equations requires a choice of values for both the α and β parameters.

Although this is so far just standard Hamilton–Jacobi theory, it helps to understand WKB states of the form $e^{iS[h_{ij},\alpha]}$. It is clear that a WKB state supplies the values of the α parameters, but that defines a family of solutions to Einstein's equations with fixed α and arbitrary β. It is in this sense that the WKB state contains information about an ensemble of spacetimes. Given a set of α parameters (*i.e.* a Hamilton–Jacobi function), a 3-geometry h_{ij} fixes a unique value of β for which the α,β spacetime contains h_{ij}.

We can use Hamilton–Jacobi theory to obtain a gauge-invariant description of WKB states [51]. Eq. (3.3.10) can be turned around to give a set of *functionals* $\alpha[h_{ij},\pi^{ij}]$ which are constants of the motion – that is they have vanishing Poisson bracket with the Hamiltonian constraint. It is also possible to define functionals $\beta[h_{ij},\pi^{ij}]$ which are conjugate to the $\alpha[h_{ij},\pi^{ij}]$, and are similarly constants of the motion. They are defined in terms of h_{ij} and π^{ij} by the equations

$$\beta[h_{ij},\pi^{ij}] = \left.\frac{\delta S\left[h_{ij},\alpha\right]}{\delta \alpha}\right|_{\alpha=\alpha[h_{ij},\pi^{ij}]}. \tag{3.3.12}$$

These definitions provide a canonical transformation between the $h_{ij}(\mathbf{x})$ and $\pi^{ij}(\mathbf{x})$ and the α and β, so that the Hamiltonian vanishes in the new co-ordinates. We are free to write our theory in terms of these constants of the motion. The α and β are co-ordinates and momenta on the physical phase space‡ (if we assume that we have solved the momentum constraint) and so are the correct variables to use for quantization according to the Dirac procedure. They can be though of as parametrizing classical solutions of the Einstein equations [52], in the sense that fixing the values of $\alpha[h_{ij},\pi^{ij}]$ and $\beta[h_{ij},\pi^{ij}]$ yields a classical solution simply by solving the equations

$$\alpha[h_{ij},\pi^{ij}] = \alpha_0, \qquad \beta[h_{ij},\pi^{ij}] = \beta_0. \tag{3.3.13}$$

‡Of course the implicit equations (3.3.10) are extremely difficult to solve in four dimensions, and so this discussion should be regarded as somewhat formal in this sense.

Let us imagine promoting the Poisson bracket algebra

$$\{\alpha_I, \beta_J\} = \delta_{IJ}, \qquad \{\mathcal{H}, \alpha\} = \{\mathcal{H}, \beta\} = 0 \tag{3.3.14}$$

to an operator algebra in the space of functionals $\Psi[h_{ij}(\mathbf{x})]$, ignoring any anomalies or ordering ambiguities. Although the Hamiltonian vanishes in the α or β representation, so that any state $\Psi[\alpha]$ or $\Psi[\beta]$ is automatically a physical state, we continue to work in the metric representation since this makes the interpretation of states somewhat easier.

The first order WKB state $\Psi_0[h_{ij}, \alpha_0] = e^{iS[h_{ij}, \alpha_0]/\hbar G}$ is an approximate eigenstate of the operator $\hat{\alpha}[\hat{h}_{ij}, \hat{\pi}^{ij}]$ with eigenvalue α_0, in the sense that:

$$\hat{\alpha}\Psi_0[h_{ij}, \alpha_0] = \alpha_0 \Psi_0[h_{ij}, \alpha_0] + o(\hbar G). \tag{3.3.15}$$

It is easy to see why this is the case. Let us write $\alpha[h_{ij}, \pi^{ij}]$ as an operator by replacing the π^{ij} by $i\hbar\delta/\delta h_{ij}$. Then the leading order contribution to the rhs comes when all derivatives bring down the exponent with its accompanying powers of $1/\hbar G$. In this leading term, the derivatives are replaced by $\pi_0^{ij}[h_{ij}, \alpha_0] \equiv \delta S[h_{ij}, \alpha_0]/\delta h_{ij}$ which has the obvious property that $\alpha[h_{ij}, \pi_0^{ij}] = \alpha_0$. An eigenstate of $\hat{\alpha}$ can be expected to have maximal uncertainty in β: As a functional of h_{ij}, $\Psi_0[h_{ij}, \alpha_0]$ is damped where h_{ij} is not found within any classical solution defined by α_0 and β for any β.

The above discussion shows that first order WKB states in the metric representation are closely tied to the α parameters. Since there is no known physical principle that prefers the α parameters to the β's, for example, it seems rather odd that they play such an important role in the semiclassical approximation. It would be more satisfactory if some mechanism restored a symmetry between the α's and the β's. Eq. (3.3.15) for the first order WKB states suggests thinking about the general properties of eigenstates of gauge invariant operators. A single spacetime is defined by a pair α_0 and β_0 of gauge invariant quantities, and any sensible observation within that spacetime implies a knowledge of both α and β to some degree of accuracy. For example any sequence of observations of h_{ij} that is made within a classical spacetime must yield values for both α and β to a reasonable accuracy. This shows that the quantum state corresponding to a classical spacetime should have expectation values for both $\hat{\alpha}$ and $\hat{\beta}$ with small quantum spreads, rather than being an eigenstate of one set of operators. This can be achieved by taking a gaussian superposition of eigenstates of $\hat{\alpha}$.

3.4 WKB superpositions and the semiclassical approximation

So far we have only concentrated on the leading order term $\Psi_0[h_{ij}, \alpha_0]$ of the WKB approximation. We have seen that this is closely related to an eigenstate of $\hat{\alpha}$. At any subsequent order, there is a freedom in how we choose to solve the Wheeler–DeWitt equation, which allows us either to construct successive approximations to an exact eigenstate of $\hat{\alpha}$, or to construct other states. These other states must be some subset of superpositions of $\hat{\alpha}$ eigenstates since eigenstates of $\hat{\alpha}$ form a basis for the set of all physical states.

Let us write exact eigenstates of $\hat{\alpha}$ which are also exact physical states as $\Psi_{\alpha_0}[h_{ij}]$. A general physical state is a superposition of eigenstates of $\hat{\alpha}$. The discussion in the previous section suggests that a quasi-classical state in quantum gravity should be a coherent superposition of $\hat{\alpha}$ eigenstates $\Psi_\alpha[h_{ij}]$:

$$|\Psi\rangle = \int d\alpha\, \omega(\alpha)|\alpha\rangle \qquad \text{or} \qquad \Psi[h_{ij}] = \int d\alpha\, \omega(\alpha)\Psi_\alpha[h_{ij}] \tag{3.4.16}$$

where $\omega(\alpha)$ is a distribution that ensures a close to minimal uncertainty in both sets of observables α and β. It has support only on a restricted region of superspace centered around a classical trajectory, and is compatible with weak observations of a classical spacetime which effectively measure the gauge invariant quantities α and β.

From the fact that a first order WKB state approximates an eigenstate of $\hat{\alpha}$, we can obtain an expansion in $\hbar G$ for the superposition (3.4.16). The important conclusion we reach is that an approximation to (3.4.16) can be obtained in the usual perturbative expansion in powers of $\hbar G$ of the gravitational Wheeler–DeWitt equation, by appropriate choice of integration constants for the terms of higher order than $S_0[h]$.

A simple example of this is given by the WKB approximation of the free relativistic particle. Although the first order WKB wavefunction solves the Klein–Gordon equation exactly, a WKB prefactor that is not constant approximates superpositions of the exact WKB states. As was first noted in Ref. [53], the fact that the Klein–Gordon equation is second order implies the presence of degrees of freedom that can give rise to the kind of superpositions we are considering.

In the case of gravity, let us write a state which approximates a classical spacetime with parameters α_0 and β_0 as

$$|\Psi\rangle = \int d\alpha\, e^{-i(\alpha-\alpha_0)\beta_0/\hbar G} e^{-(\alpha-\alpha_0)^2/\hbar G}|\alpha\rangle$$

so that $\omega(\alpha) = e^{-i(\alpha-\alpha_0)\beta_0/\hbar G} e^{-(\alpha-\alpha_0)^2/\hbar G}$. Here we are working with α and β normalized so that they have the same dimensions and that $[\alpha,\beta] = \hbar G$. The choice of $\omega(\alpha)$ ensures that α and β are localized to within $\sqrt{\hbar G}$ of their mean values α_0 and β_0. We also assume that some α and β are large compared to $\sqrt{\hbar G}$ so that there is a large dimensionless parameter with respect to which we can perform the expansion. This is related to the physical criterion that fluctuations should be small compared to the characteristic scale of the solution. For example, a cosmology with a maximal size of the order of the Planck scale (see Sec. 6 for an example) should not be considered classical.

In the metric representation we have

$$\langle h_{ij}|\alpha\rangle \cong e^{iS[h_{ij},\alpha]/\hbar G}$$

to first order, so that

$$\Psi_G[h_{ij}] \cong \int d\alpha\, e^{-i(\alpha-\alpha_0)\beta_0/\hbar G} e^{-(\alpha-\alpha_0)^2/\hbar G} e^{iS[h_{ij},\alpha]/\hbar G}. \tag{3.4.17}$$

The integration in (3.4.17) can be performed after expanding $e^{iS[h_{ij},\alpha]}$ in powers of $\alpha - \alpha_0$. Keeping the terms contributing to S_0 and D of Sec. 1, the result is

$$\begin{aligned}\Psi_G[h_{ij}] \;&\cong\; e^{iS[\alpha_0]/\hbar G}\frac{(2\hbar G)^{1/2}}{\left(4+(S''[\alpha_0])^2\right)^{1/4}}e^{-(S'[\alpha_0]-\beta_0)^2/\hbar G\left(4+(S''[\alpha_0])^2\right)} \\ &\quad e^{i\tan^{-1}[S''[\alpha_0]/2]/\hbar G}e^{-i(S'[\alpha_0]-\beta_0)^2 S''[\alpha_0]/2\hbar G\left(4+(S''[\alpha_0])^2\right)} \qquad (3.4.18)\end{aligned}$$

where $S[\alpha] = S[h_{ij}, \alpha]$, $S'[\alpha] = \delta S[h_{ij}, \alpha]/\delta\alpha$ and $S''[\alpha] = \delta^2 S[h_{ij}, \alpha]/\delta\alpha^2$. The first term in (3.4.18) is the first order WKB approximation for $\alpha = \alpha_0$, the center of the gaussian, and is the only rapidly oscillating term of order $1/\hbar G$. The other terms all belong in the next order correction, S_1. The last two exponentials are corrections to $S[\alpha_0]$ and provide some order $\hbar$ corrections to Einstein's equations. The remaining terms are real and make up part of the WKB prefactor $D[h_{ij}]$. The important term is the exponential, which damps 3-geometries h_{ij} which are not compatible with $\beta = \beta_0$. Although the α and β damping in this representation occur in different ways, the resulting state is damped away from a narrow tube surrounding the mean spacetime given by α_0 and β_0. This is precisely what one expects for a gaussian superposition (3.4.16).

The definition of a semiclassical state is not limited to the case where $\omega(\alpha)$ is an exact gaussian. A general $\omega(\alpha)$ in (3.4.16) will do equally well provided that it is peaked around some α_0 and that its fourier transform $\tilde{\omega}(\beta)$ is peaked around some β_0 so that $\Delta\alpha, \Delta\beta \sim \sqrt{\hbar G}$. Under these conditions one can write (3.4.16) to a good approximation as

$$\Psi[h_{ij}] = e^{iS[h_{ij},\alpha_0]}\tilde{\omega}\left(\frac{S'(\alpha_0) - \beta_0}{\sqrt{\hbar G}}\right)$$

where $\tilde{\omega}$ contributes the damping in β, and both it and its derivatives belong to S_1 or lower order terms.

We conclude that any quasiclassical superposition of pure WKB states is approximately of the WKB form (3.2.4) if we take the prefactor into account. Thus a coherent superposition fits into the expansion scheme described in Sec. 2, and the matter portion of the WKB state, $\chi[f, h]$, is still given by solving the functional Schrödinger equation along characteristics, now restricted to lie in a narrow tube around the α_0, β_0 classical solution. The characteristics are only those for α_0 and for all β's within the tube, so that the semiclassical approximation still looks asymmetric with respect to α and β. However, the differences between evolving $\chi[f, h]$ on any of the characteristics generically belong to lower order corrections because of the narrowness of the tube. In this sense the asymmetry has been removed and we can think only of solving the functional Schrödinger equation on a mean spacetime defined by α_0 and β_0.

3.5 Beyond the semi-classical approximation

The fact that a semiclassical wave functional for gravity (3.4.18) has support on a tube in superspace rather than a single eikonal track indicates the presence of corrections beyond the simple picture of quantum field theory on a single background spacetime. These corrections should be compared with those discussed by Kiefer and Singh [54]. In general, they should be related to Planck scale physics, as was first explained by Wheeler [55].

When one wants to talk about quantum gravity beyond the semi-classical approximation, the lack of a background spacetime, and of the notion of matter fields living on that background makes life difficult. It is unclear what is meant by a unitary theory and how to define an inner product under these circumstances. It is likely that some or all of these concepts make sense only to the same order as the semiclassical approximation, but nonetheless they form the basis of our description of nature. The semiclassical approximation using gaussian states is very much in agreement with Wheeler's picture, since the Planck scale uncertainties in α and β can be related to fluctuations in the underlying spacetimes which are generically on the Planck scale (see chapter 5 and some of Wheeler's original arguments [55]).

Lets examine what we can learn qualitatively about corrections arising from the geometry fluctuations. We want to look at the state of matter at some hypersurface

$$\Psi_{h^0_{ij}}[f] \equiv \Psi[f, h^0_{ij}].$$

To the order of the semiclassical approximation (equations (3.2.5) and (3.2.6)), this is given by taking h_{ij} to be embedded only in the spacetimes labeled by α_0 and some β. Then one finds that $\Psi_{h^0_{ij}}[f]$ is a solution to the functional Schrödinger equation on a fixed background.

Corrections to this approximation come from taking into account contributions from all the possible spacetimes labeled by α and β which are not damped in the gaussian state. A simple way to get qualitative information about these corrections is to consider solving the functional Schrödinger equation on all of these spacetimes (not just those with $\alpha = \alpha_0$) and comparing the properties of the solutions. In order to solve the functional Schrödinger equation for the set of spacetimes defined by (3.4.18), it is necessary to give initial data on each of them (that is on a surface in superspace transverse to the tube). This initial data should presumably be arranged to make the corrections to the semiclassical approximation as small as possible.

If there are to be only Planck scale corrections to the semiclassical approximation, the difference between Schrödinger evolution of matter states on each of the spacetimes should be small, except at the Planck scale. If this is not the case, the results obtained to the order of the semiclassical approximation are not consistent. It is clear that one situation in which the semiclassical approximation breaks down unexpectedly is if there is chaotic behavior of the matter state with respect to the α and β parameters.

An example of the comparison of matter states is provided by comparing the properties of matter correlation functions in all the different spacetimes. An insertion point (*i.e.* an event) is only defined by its position on a 3-geometry (since the basic variable in canonical quantum gravity is a 3-geometry). To define a correlation function one needs to pick a 3-geometry in the mean spacetime given by α_0 and β_0 that contains all the insertion points. The insertion points are defined by the 3-geometry and their locations within it. Next one needs to look at how the given 3-geometry is embedded in the other spacetimes. This then identifies the insertion points in each of the spacetimes, allowing a comparison of the correlation functions to be made [45].

The location of the insertion points in other spacetimes depends upon the choice of 3-geometry containing the points in the α_0, β_0 spacetime. It follows that the size of the corrections depends on the choice of 3-geometry and thus on how one chooses to foliate the mean spacetime. However small the dependence on the choice of foliation, this seems to say that coordinate invariance is lost[§], which looks puzzling since we started off with the Wheeler–DeWitt equation which is supposed to impose that symmetry. Recall, however, that the familiar notion of coordinate invariance comes from the semiclassical expansion of the Wheeler–DeWitt equation, which gives a covariant equation for matter evolution on a fixed background spacetime. To this order, observations are independent of a choice of foliations of the mean background spacetime. It is this notion which breaks down when one takes into account the geometry fluctuations which are higher order corrections. This is because the meaning of the Wheeler–DeWitt equation is different at this next order, since the notion of diffeomorphism invariance is now a property of the combined matter–gravity system, not just of matter on a fixed background.

What makes this discussion particularly relevant is that quantum gravity effects that normally occur at the Planck scale are magnified to the classical scale by the apparently chaotic behaviour of functional Schrödinger evolution on certain hypersurfaces close to a black hole horizon. In the 1+1 dimensional CGHS model [56], it will be shown in chapter 5 that, for a particular choice of hypersurface, an insertion point located near the horizon in a mean spacetime of mass M is identified with points almost all over the horizon in spacetimes of masses differing from M on the Planck scale. This is enough to show that corrections to the semiclassical approximation can be large. Also, since the result depends sensitively on the choice of hypersurface, it indicates a breakdown in coordinate invariance. This probably means that in an effective description, the results of certain sets of observations near the black hole horizon are not covariant, a conclusion similar to those of 't Hooft [57] and Susskind [58].

3.6 A two-dimensional example

The use of Hamilton–Jacobi theory to reduce to the physical degrees of freedom was discussed rather abstractly in Sec. 3. It is instructive to illustrate this using a simple 1+1 dimensional dilaton gravity model, which also allows some brief comments about open spacetimes. The model we shall consider was discussed by Louis-Martinez *et al* [59], and we shall make extensive use of their results. Related work on open and closed spacetimes can be found in Ref. [60].

[§]This foliation dependence is independent of the anomalies discussed in Sec. 2 and in Refs. [105, 48, 49].

3.6.1 Classical theory: closed universe

Let us focus primarily on the case of a closed cosmology. Consider the action,

$$S = \int_{\mathcal{M}} d^2x\sqrt{-g}\,[\phi R - V(\phi)]$$

where x is a periodic co-ordinate (with period 2π), so that $\mathcal{M} = S^1 \times R$. A constant potential $V(\phi) = -4\lambda^2$ gives the closed universe version [61] of the CGHS model [56].

Working with the parameterization

$$ds^2 = e^{2\rho}\left(-N^2dt^2 + (dx + N_\perp dt)^2\right)$$

for the metric $g_{\mu\nu}$, the canonical variables are $\rho(x)$ and $\phi(x)$, with conjugate momenta

$$\Pi_\phi = \frac{2}{N}(N_\perp\rho' + N_\perp' - \dot{\rho}) \tag{3.6.19}$$

$$\Pi_\rho = -\frac{2}{N}\left(\dot{\phi} - N_\perp\phi'\right) \tag{3.6.20}$$

while the lapse and shift functions N and $N_\perp$ are Lagrange multipliers. As usual, the Hamiltonian is just a sum of the Hamiltonian and diffeomorphism constraints

$$H = \int dx\,[N\mathcal{H} + N_\perp\mathcal{H}_\perp]$$

where

$$\mathcal{H} = 2\phi'' - 2\phi'\rho' - \frac{1}{2}\Pi_\rho\Pi_\phi + e^{2\rho}V(\phi), \tag{3.6.21}$$

$$\mathcal{H}_\perp = \rho'\Pi_\rho + \phi'\Pi_\phi - \Pi_\rho'. \tag{3.6.22}$$

The Hamilton-Jacobi equation reads

$$g[\phi,\rho] + \frac{\delta S}{\delta\phi}\frac{\delta S}{\delta\rho} = 0$$

where

$$g[\phi,\rho] = -4\phi'' + 4\phi'\rho' - 2e^{2\rho}V(\phi).$$

This is solved by the functional

$$S[\phi,\rho,C] = 2\int dx\left\{Q_C + \phi'\ln\left[\frac{2\phi' - Q_C}{2\phi' + Q_C}\right]\right\} \tag{3.6.23}$$

where C is a constant,

$$Q_C[\phi,\rho] = 2\left[(\phi')^2 + (C + j(\phi))\,e^{2\rho}\right]^{\frac{1}{2}}$$

and

$$\frac{dj(\phi)}{d\phi} = V(\phi).$$

(3.6.23) is also invariant under spatial diffeomorphisms.

In the Hamilton–Jacobi function (3.6.23), we see the presence of a parameter C which is the α parameter in this problem. To deduce C as a functional of ϕ, ρ and their conjugate momenta, we must invert the relations

$$\frac{\delta S}{\delta \phi} = \frac{g[\phi, \rho]}{Q_C[\phi, \rho]}, \qquad \frac{\delta S}{\delta \rho} = Q_C[\phi, \rho].$$

These lead to the definition

$$C = e^{-2\rho}\left(\frac{1}{4}\Pi_\rho^2 - (\phi')^2\right) - j(\phi) \tag{3.6.24}$$

Similarly we can define the quantity β, we shall call P following [59], as

$$P = \frac{\delta S}{\delta C} = -\int dx \frac{2e^{2\rho}\Pi_\rho}{\Pi_\rho^2 - 4(\phi')^2}. \tag{3.6.25}$$

It is easy to check that C and P are conjugate and that they have vanishing Poisson brackets with the constraints.

From Eq. (3.6.23) for the Hamilton–Jacobi functional, we can solve the classical equations of motion, using the relations

$$\Pi_\phi = \frac{\delta S}{\delta \phi}, \qquad \Pi_\rho = \frac{\delta S}{\delta \rho}$$

and Eqs. (3.6.19) and (3.6.20).

For a constant potential $V(\phi) = -4\lambda^2$, and taking $\sigma = 1$ and $M = 0$, there are homogeneous solutions

$$ds^2 = \frac{\lambda^2 P^2}{\pi^2} e^{-2\lambda^2 Pt/\pi}(-dt^2 + dx^2)$$

and

$$\phi = \frac{C}{4\lambda^2} - e^{-2\lambda^2 Pt/\pi}$$

for all values of C and P.

3.6.2 Classical theory: spacetimes with boundary

The case of an open universe has been studied by various authors [60]. It has been shown that the variable C is related to the ADM mass of the spacetime, while P, integrated throughout a spacelike slice, is related to the time at infinity (or more precisely, to the synchronization between times at infinity). These results are in keeping with the much earlier work of Regge and Teitelboim [62] on conserved charges in canonical quantum gravity in open universes.

It is interesting to note that while P is associated with a constant of the motion as described in Refs. [60], a closely related quantity provides a local geometric definition of time for different hypersurfaces within a static spacetime. Consider any 1-geometry associated with a hypersurface in a static classical solution. It can be intrinsically described by the function $\phi(s)$, where s is the proper distance along the hypersurface measured from some base point B at infinity, and ϕ is the value of the dilaton field. Let $\phi_0(s)$ be the function defining a constant time surface $t = t_1$ passing through $B = \phi_0(0)$. Consider now a set of hypersurfaces passing through B that are defined by $\phi_i(s)$ which differ from $\phi_0(s)$ only in some finite interval $0 < s < s_0$. For $s > s_0$, $\phi_i(s)$ also define constant time hypersurfaces but at some time $t_i = t_0 + \Delta t_i$.

Using (3.6.24) to give Π_ρ in terms of C and ϕ, we can define a quantity

$$T_i(S) = -\int_0^S ds \frac{\left[\left(\frac{d\phi_i}{ds}\right)^2 + C + j(\phi_i)\right]^{1/2}}{(C + j(\phi_i))}$$

closely related to P. Here $S > s_0$ so that the integration extends well into the region where $\phi_i(s) = \phi_0(s)$.

Now in a static coordinate system

$$ds^2 = e^{2\rho}(-dt^2 + dx^2)$$

we can compute the change in the time coordinate along any hypersurface using

$$\left(\frac{dt}{dx}\right)^2 = 1 - \frac{e^{-2\rho}\left(\frac{d\phi_i}{dx}\right)^2}{\left(\frac{d\phi_i}{ds}\right)^2}.$$

Since in the static coordinate system $\Pi_\rho = 0$, it follows that

$$\left(\frac{dt}{dx}\right)^2 = 1 + \frac{C + j(\phi_i)}{\left(\frac{d\phi_i}{ds}\right)^2}.$$

From this expression we deduce that

$$\Delta t_i = \int dx \frac{\left[\left(\frac{d\phi_i}{ds}\right)^2 + C + j(\phi_i)\right]^{1/2}}{\left(\frac{d\phi_i}{ds}\right)} = \int_0^S ds \frac{\left[\left(\frac{d\phi_i}{ds}\right)^2 + C + j(\phi_i)\right]^{1/2}}{e^\rho \left(\frac{d\phi_i}{ds}\right)}.$$

By definition Δt is zero for ϕ_0, but is non-zero for any $\phi_i(s)$ over the region $0 < s < s_0$.

The connection between Δt_i and $T_i(S)$ emerges by noting that any static solution has

$$e^\rho = c_0 \left(\frac{d\phi}{ds}\right)$$

where c_0 is some constant of proportionality (this can be shown using Eqs. (3.6.21) and (3.6.24)). From this it follows that

$$\Delta t_i(S) = \frac{T_i(S)}{c_0}.$$

3.7 Conclusions

We have shown that the appropriate state to consider as a semiclassical state for quantum gravity is a superposition of WKB type states which is peaked around some values of the reduced phase space variables, with close to minimal uncertainty in the reduced phase space variables. A pure first order WKB state on the other hand is an approximate eigenstate of half of these variables and hence not adequate. When matter is present the correct ansatz for the gravity part is still a gaussian superposition, since this is perfectly compatible with the derivation of the Schrödinger equation from the Wheeler–DeWitt constraint.

Using a superposition of WKB states, we were able to give a heuristic treatment of higher order effects due to the quantum nature of the background geometry. This allowed us to identify certain situations in which the semiclassical approximation is inconsistent because of the sensitivity of matter propagation to small fluctuations in the background geometry. An example of this type of situation is given by the breakdown of a semiclassical description of matter propagating on a black hole spacetime [44, 45]. We also showed that as a consequence of quantum fluctuations, coordinate invariance is lost on the Planck scale, and in certain cases, such as near the black hole, this extends to macroscopic scales.

We have not discussed in this paper how it is that a system described by the Wheeler–DeWitt equation comes to find itself in the particular state that exhibits semiclassical behaviour. There is in principle no dynamical or kinematical reason to prefer this state over any other. Perhaps the most likely answer to this question is that decoherence effectively drives any initial state to a configuration in which observations are equivalent to those within the gaussian state. It is important to note, however, that decoherence cannot drive a state towards any configuration for which the background spacetime is more classical than the one we have described.

Chapter 4

Scalar field quantization on the 2+1 Dimensional Black Hole Background

On Wednesday, when the sky is blue,
And I have nothing else to do,
I sometimes wonder if it's true
That who is what and what is who.
(The World of Pooh by, A. A. Milne)

Abstract

The quantization of a massless conformally coupled scalar field on the 2+1 dimensional Anti de Sitter black hole background is presented. The Green's function is calculated, using the fact that the black hole is Anti de Sitter space with points identified, and taking into account the fact that the black hole spacetime is not globally hyperbolic. It is shown that the Green's function calculated in this way is the Hartle-Hawking Green's function. The Green's function is used to compute $\langle T^{\mu}_{\nu} \rangle$, which is regular on the black hole horizon, and diverges at the singularity. A particle detector response function outside the horizon is also calculated and shown to be a fermi type distribution. The back-reaction from $\langle T_{\mu\nu} \rangle$ is calculated exactly and is shown to give rise to a curvature singularity at $r = 0$ and to shift the horizon outwards. For $M = 0$ a horizon develops, shielding the singularity. Some speculations about the endpoint of evaporation are discussed.

This chapter is based on [63]

4.1 Introduction

The study of black hole physics is complicated by the many technical and conceptual problems associated with quantum field theory in curved spacetime. One serious difficulty is that exact calculations are almost impossible in 3+1 dimensions. In this paper we shall instead study some aspects of quantum field theory on a 2+1 dimensional black hole background. This enables us to obtain an exact expression for the Green's function of a massless, conformally coupled scalar field in the Hartle-Hawking vacuum [64]. We use this Green's function to study particle creation by the black hole, back-reaction and the endpoint of evaporation.

We shall work with the 2+1 dimensional black hole solution found by Bañados, Teitelboim and Zanelli (BTZ) [65]. It had long been thought that black holes cannot exist in 2+1 dimensions for the simple reason that there is no gravitational attraction, and therefore no mechanism for confining large densities of matter. This difficulty has been circumvented in the BTZ spacetime†, but not surprisingly, their solution has some features that we do not normally associate with black holes in other dimensions, such as the absence of a curvature singularity. It is interesting to ask whether this spacetime behaves quantum mechanically in a way consistent with more familiar back holes.

The spinless BTZ spacetime has a metric [65]

$$ds^2 = -N^2 dt^2 + N^{-2} dr^2 + r^2 d\phi^2$$

where

$$N^2 = \frac{r^2 - r_+^2}{\ell^2}, \qquad r_+^2 = M\ell^2.$$

Here M is the mass of the black hole. The metric is a solution to Einstein's equations with a negative cosmological constant, $\Lambda = -\ell^{-2}$, and the curvature of the black hole solution is constant everywhere. As a result there is no curvature singularity as $r \to 0$. A Penrose diagram of the spacetime is given in Fig. 1.

An important feature of the BTZ solutions is that the solution with $M = 0$ (which we refer to as the vacuum solution), is not AdS_3. Rather, it is a solution that is not globally Anti de Sitter invariant. It has no horizon, but does have an infinitely long throat for small $r > 0$, which is reminiscent of the extreme Reissner-Nordström solution in 3+1 dimensions. It is worth noting that there are other similarities between the spinless BTZ black holes, $M \geq 0$, and the Reissner-Nordström solutions for $M \geq Q$. In particular, the temperature associated with the Euclidean continuation of the BTZ black holes has been computed in [65], and it was found to increase with the mass, and to decrease to zero as $M \to 0$. Thus, if we carry over the usual notions from four dimensional black holes, the $M = 0$ solution appears to be a stable endpoint of evaporation.

†A charged black hole solution in 2+1 dimensions had previously been found in Ref. [66]. For further discussions on the BTZ black hole, see Refs. [67].

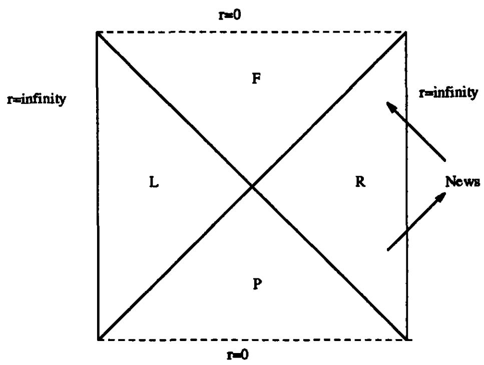

Figure 1: A Penrose diagram of the $M \neq 0$ black hole. Information can leak through spatial infinity, unless we impose boundary conditions at $r = \infty$.

A feature of the BTZ solution that we shall make use of, is that the solution arises from identifying points in AdS_3, using the orbits of a spacelike Killing vector field. It is this property that is the starting point of our derivation of a Green's function on the black hole spacetime. We construct a Green's function on AdS_3, and this translates to a Green's function on the black hole via the method of images.

A Green's function constructed in this way is only interesting if we can identify the vacuum with respect to which it is defined. We prove that our construction gives the Hartle-Hawking Green's function. It is worth noting that for the BTZ black hole, there is a limited choice of vacua. Quantization on AdS_3 is hampered by the fact that AdS_3 is not globally hyperbolic, and this necessitates the use of boundary conditions at spatial infinity [68] (see Fig. 2), as discussed in Appendix A. This problem carries over to the black hole solution, and as a result, the value of the field at spatial infinity is governed by either Dirichlet or Neumann type boundary conditions. Thus a Cauchy surface for the region R of the BTZ black hole is either the past or the future horizon only. With this knowledge, the natural definition of the Hartle-Hawking vacuum is with respect to Kruskal modes on either horizon, whereas there does not appear to be a natural definition of an Unruh vacuum (see [69, 70] for a discussion of the various eternal black hole vacua). The definition of an Unruh vacuum might be possible given a description of the formation of a BTZ black hole from the vacuum via some sort of infalling matter, but as far as we are aware, no such construction has been found.

Having an explicit expression for the Hartle-Hawking Green's function, we are able to obtain a number of exact results. As a check, we show that it satisfies the KMS thermality condition [71]. We then compute the expectation value of the energy-momentum tensor and the response of a particle detector for both nonzero M and for the vacuum solution. For nonzero M we address the issue of whether the response of the particle detector can be interpreted as radiation emitted from the black hole, although a clear picture does not emerge.

For the $M = 0$ solution, we find a non-zero energy-momentum tensor, although the corresponding Green's function is at zero temperature, and there is no particle detector response. We interpret this as a sort of Casimir energy. Classically, the vacuum solution appears to be similar to the extremal Reissner-Nordström solution, in the sense that we expect that if any matter is thrown in, a horizon develops. Quantum mechanically, the $M = 0$ solution appears to be unstable to the formation of a horizon, when the back-reaction caused by the Casimir energy is taken into account. This suggests that the endpoint of evaporation may not look like the classical $M = 0$ solution.

The paper is organized as follows. In section 4.2 we study the 1+1 dimensional solution which arises from a dimensional reduction of the BTZ black hole [72], and show that the vacuum defined by the Anti de Sitter (AdS) modes is the same as that defined by the Kruskal modes; with this encouraging result we tackle the 2+1 case. Section 4.3 contains a review of the essential features of the geometry of the BTZ black hole. In section 4.4 we construct the Wightman Green's functions on the black hole spacetime from the AdS_3 Wightman function, using the method of images. We then show that the Green's function coincides with the Hartle-Hawking Green's function [64], by showing that it is analytic and bounded in the lower half of the complex $\bar{V}$ plane on the past horizon ($\bar{U} = 0$), where $\bar{V}, \bar{U}$ are the Kruskal null coordinates. We also compute the Wightman function for the $M = 0$ solution, and compare this to the $M \to 0$ limit of the results for $M \neq 0$. Section 4.5 contains a calculation of $\langle T_{\mu\nu} \rangle$ for all $M \geq 0$. For the black hole solutions, it is regular on the horizon, and for all M it is singular as $r \to 0$. In Section 4.6 the response function of a stationary particle detector outside the horizon is calculated and shown to be of a fermi type distribution. A discussion is given of how this response might be interpreted. In Section 4.7 we calculate the back reaction induced by $\langle T_{\mu\nu} \rangle$, and show that the spacetime develops a curvature singularity and a larger horizon for a given M. Throughout this paper we use metric signature $(- + +)$, and natural units in which $8G = \hbar = c = 1$.

4.2 2-D Black Hole

Let us begin by looking at quantum field theory on the region of Anti de Sitter spacetime in 1+1 dimensions described by the metric

$$ds^2 = -\left(\frac{r^2 - M\ell^2}{\ell^2}\right) dt^2 + \left(\frac{r^2 - M\ell^2}{\ell^2}\right)^{-1} dr^2 \qquad 0 < r < \infty \qquad -\infty < t < \infty,$$

where M is the mass of the solution. This metric was discussed in [72] as the dimensional reduction of the spinless BTZ black hole, and can be thought of as being a region of AdS_2 in Rindler-type co-ordinates. Under the change of co-ordinates

$$r = \sqrt{M\ell^2} \sec\rho \cos\lambda, \qquad \tanh\left(\frac{\sqrt{M}t}{\ell}\right) = \frac{\sin\rho}{\sin\lambda},$$

where we shall call (λ, ρ) AdS co-ordinates, the metric becomes

$$ds^2 = \ell^2 \sec^2\rho(-d\lambda^2 + d\rho^2)$$

which for $-\frac{\pi}{2} \leq \rho \leq \frac{\pi}{2}$ and $-\infty < \lambda < \infty$ is just AdS_2 [73].

It is possible to define Kruskal-like co-ordinates for this black hole, which do not coincide with the usual AdS co-ordinates. For $r > M\ell^2$, they are:

$$U = \left(\frac{r - \sqrt{M}\ell}{r + \sqrt{M}\ell}\right)^{\frac{1}{2}} \cosh \frac{\sqrt{M}}{\ell} t \tag{4.2.1}$$

$$V = \left(\frac{r - \sqrt{M}\ell}{r + \sqrt{M}\ell}\right)^{\frac{1}{2}} \sinh \frac{\sqrt{M}}{\ell} t. \tag{4.2.2}$$

The metric then takes the form

$$ds^2 = \frac{-2\ell^2}{1 + \bar{U}\bar{V}} d\bar{U} d\bar{V}$$

where $\bar{U} = V + U$, $\bar{V} = V - U$, and the transformation between Kruskal and AdS co-ordinates is

$$\bar{U} = \tan\left(\frac{\rho + \lambda}{2}\right) \qquad \bar{V} = \tan\left(\frac{\rho - \lambda}{2}\right)$$

which is valid over all the Kruskal manifold. The Kruskal co-ordinates cover only the part of AdS_2 with

$$-\frac{\pi}{2} \leq \rho \leq \frac{\pi}{2} \qquad -\frac{\pi}{2} < \lambda < \frac{\pi}{2}.$$

We shall now show that the notion of positive frequency in (λ, ρ) (AdS) modes coincides with that defined in (U, V) (Kruskal) modes.

The AdS modes for a conformally coupled scalar field are normalized solutions of $\Box\psi = 0$, subject to the boundary conditions

$$\phi\left(\rho = \frac{\pi}{2}\right) = \phi\left(\rho = \frac{-\pi}{2}\right) = 0.$$

The positive frequency modes are then

$$\phi_m = \frac{1}{\sqrt{m\pi}} e^{-im\lambda} \sin m\rho \qquad m \text{ even } \geq 0$$

$$\phi_m = \frac{1}{\sqrt{m\pi}} e^{-im\lambda} \cos m\rho \qquad m \text{ odd } \geq 0$$

and these define a vacuum state $|0\rangle_A$ in the usual way.

The Kruskal modes are solutions of $\Box\psi = 0$ with the boundary condition $\psi(\bar{U}\bar{V} = -1) = 0$. Positive frequency solutions are given by

$$\psi_\omega = N_\omega \left(e^{-i\omega\bar{U}} - e^{i\omega/\bar{V}}\right) \qquad \omega > 0$$

where $N_\omega = (8\pi\omega)^{-1/2}$, and these define $|0\rangle_K$. These modes are analytic and bounded in the lower half of the complex $\bar{U}, \bar{V}$ plane. In order to show equivalence of the two vacua $|0\rangle_A$ and $|0\rangle_K$, it is enough to show that the positive frequency AdS modes can be written as a sum of only positive frequency Kruskal modes. Because of the analyticity properties of the Kruskal modes, it is enough to show that the AdS modes are analytic and bounded in the lower half of the complex $\bar{U}, \bar{V}$ plane [69, 4]. Changing co-ordinates, we have

$$\phi_m = \frac{1}{\sqrt{\pi m 2i}}\left(e^{-2im\arctan\bar{V}} - e^{-2im\arctan\bar{U}}\right) \qquad m \text{ even} \tag{4.2.3}$$

$$\phi_m = \frac{1}{\sqrt{\pi m 2i}}\left(e^{-2im\arctan\bar{V}} + e^{-2im\arctan\bar{U}}\right) \qquad m \text{ odd.} \tag{4.2.4}$$

Using the definition $\arctan z = \frac{1}{2i}\ln\frac{1+iz}{1-iz}$ [77], (4.2.3) and (4.2.4) become

$$\phi_m = \frac{1}{2\sqrt{\pi m i}}\left[\left(\frac{1-i\bar{V}}{1+i\bar{V}}\right)^m \mp \left(\frac{1-i\bar{U}}{1+i\bar{U}}\right)^m\right]$$

where $\pm$ is for m odd or even. These modes can easily be seen to be bounded and analytic in the lower half of the complex $\bar{U}, \bar{V}$ plane for all m. This establishes that the vacuum defined by the AdS modes is the same as that defined by the Kruskal modes. Thus a Green's function defined on this spacetime using AdS co-ordinates (λ, ρ) corresponds to a Hartle-Hawking Green's function, in the sense discussed in the Introduction.

4.3 The Geometry of the 2+1 Dimensional Black Hole

In this paper, we shall be working only with the spinless black hole solution in 2+1 dimensions

$$ds^2 = -N^2dt^2 + N^{-2}dr^2 + r^2d\phi^2 \tag{4.3.5}$$

where

$$N^2 = \frac{r^2 - r_+^2}{\ell^2}, \qquad r_+^2 = M\ell^2.$$

As was shown in [65], this metric has constant curvature, and is a portion of three dimensional Anti de Sitter space with points identified. The identification is made using a particular killing vector ξ, by identifying all points $x_n = e^{2\pi n i \xi}x$. In order to see this most clearly, it is useful to introduce different sets of co-ordinates on AdS_3.

AdS_3 can be defined as the surface $-v^2 - u^2 + x^2 + y^2 = -\ell^2$ embedded in R^4 with metric $ds^2 = -du^2 - dv^2 + dx^2 + dy^2$. A co-ordinate system (λ, ρ, θ) which covers this space, and which we shall refer to as AdS co-ordinates, is defined by [68]

$$\begin{aligned} u &= \ell\cos\lambda\sec\rho \qquad v = \ell\sin\lambda\sec\rho \\ x &= \ell\tan\rho\cos\theta \qquad y = \ell\tan\rho\sin\theta \end{aligned}$$

where $0 \le \rho \le \frac{\pi}{2}$, $0 < \theta \le 2\pi$, and $0 < \lambda < 2\pi$. In these co-ordinates, the AdS_3 metric becomes

$$ds^2 = \ell^2\sec^2\rho(-d\lambda^2 + d\rho^2 + \sin^2\rho\, d\theta^2).$$

AdS_3 has topology S^1 (time) $\times R^2$ (space) and hence contains closed timelike curves. The angle λ can be unwrapped to form the covering space of AdS_3, with $-\infty < \lambda < \infty$, which does not contain any closed timelike curves. Throughout this paper we work with this covering space, and this is what we henceforth refer to as AdS_3. As mentioned in the Introduction, even this covering space presents difficulties since it is not globally hyperbolic (see the discussion in Appendix A).

The identification taking AdS_3 into the black hole (4.3.5) is most easily expressed in terms of co-ordinates (t, r, ϕ), related in an obvious way to those used above, and defined on AdS_3 by

$$\begin{aligned} u &= \sqrt{A(r)} \cosh\left(\tfrac{r_+}{\ell}\phi\right) \\ x &= \sqrt{A(r)} \sinh\left(\tfrac{r_+}{\ell}\phi\right) \\ y &= \sqrt{B(r)} \cosh\left(\tfrac{r_+}{\ell^2}t\right) \\ v &= \sqrt{B(r)} \cosh\left(\tfrac{r_+}{\ell^2}t\right) \end{aligned} \qquad r > r_+$$

$$\begin{aligned} u &= \sqrt{A(r)} \cosh\left(\tfrac{r_+}{\ell}\phi\right) \\ x &= \sqrt{A(r)} \sinh\left(\tfrac{r_+}{\ell}\phi\right) \\ y &= -\sqrt{-B(r)} \cosh\left(\tfrac{r_+}{\ell^2}t\right) \\ v &= \sqrt{-B(r)} \cosh\left(\tfrac{r_+}{\ell^2}t\right) \end{aligned} \qquad 0 > r > r_+.$$

Note that $-\infty < \phi < \infty$. Under the identification $\phi \to \phi + 2\pi n$, where n is an integer, these regions of AdS_3 become regions R and F of the black hole. Regions P and L are defined in an analogous way [65] (see Fig. 1 for a definition of regions F (future), P (past), R (right) and L (left)). The $r = 0$ line is a line of fixed points under this identification, and hence there is a singularity there in the black hole spacetime of the Taub-NUT type [65, 73].

Finally, it is possible to define Kruskal co-ordinates on the black hole. The relation between the Kruskal co-ordinates V and U and the black hole co-ordinates t and r is precisely as in (4.2.1) and (4.2.2). U, V and an unbounded ϕ cover the region of AdS_3 which becomes the black hole after the identification.

4.4 Green's Functions on the 2+1 Dimensional Black Hole

In this section we derive a Green's function on the black hole spacetime, by using the method of images on a Green's function on AdS_3. We then show that the resulting Green's function is thermal, in that it obeys a KMS condition [71]. Using the analyticity properties discussed in the Introduction, the Green's function is also shown to be defined with respect to a vacuum state corresponding to Kruskal co-ordinates on both the past and future horizons of the black hole. We therefore interpret it as a Hartle-Hawking Green's function. Finally we derive the Green's function for the $M = 0$ solution directly from a mode sum, and compare it with the $M \to 0$ limit of the black hole Green's function.

4.4.1 Deriving the Green's Functions

Since the black hole spacetime is given by identifying points on AdS_3 using a spacelike Killing vector field, we can use the method of images to derive the two point function on the black hole spacetime. Given the two point function $G_A^+(x,x')$ on AdS_3,

$$G^+_{\mathrm{BH}}(x,x';\delta)=\sum_n e^{-i\delta n}G^+_{\mathrm{A}}(x,x'_n)$$

Here x'_n are the images of x' and $0\le\delta<\pi$ can be chosen arbitrarily. For a general δ the modes of the scalar field on the black hole background will satisfy $\phi_m(e^{2\pi n\xi}x)=e^{-i\delta n}\phi_m(x)$. $\delta=0$ for normal scalar fields and $\delta=\pi$ for twisted fields. From now on we will restrict ourselves to $\delta=0$.

This definition of the Green's function on the black hole spacetime means that when summing over paths to compute the Feynman Green's function $G_F(x,x')$, we sum over all paths in AdS_3. Hence paths that cross and recross the singularities must be taken into account (compare this with the results of Hartle and Hawking [64]).

As explained in Appendix A, boundary conditions at infinity must be imposed on any Green's function on AdS_3 in order to deal with the fact that AdS_3 is not globally hyperbolic. From Appendix A, we have

$$G_A^+=G_{A1}^+\pm G_{A2}^+ \tag{4.4.6}$$

where $+(-)$ corresponds to Neumann (Dirichlet) boundary conditions (from now on, it should be assumed that the upper (lower) sign is always for Neumann (Dirichlet) boundary conditions unless otherwise stated). The individual terms in (4.4.6) are given by

$$\begin{aligned}
G_{A1}^+(x,x') &= \frac{1}{4\sqrt{2}\pi\ell}\left(\cos(\Delta\lambda-i\epsilon)\sec\rho'\sec\rho-1-\tan\rho\tan\rho'\cos\Delta\theta\right)^{-\frac{1}{2}}\\
G_{A2}^+(x,x') &= \frac{1}{4\sqrt{2}\pi\ell}\left(\cos(\Delta\lambda-i\epsilon)\sec\rho'\sec\rho+1-\tan\rho\tan\rho'\cos\Delta\theta\right)^{-\frac{1}{2}}.
\end{aligned}$$

$\Delta\lambda$ is defined as $\lambda-\lambda'$, and similarly for all other co-ordinates.

The sign of the $i\epsilon$ is proportional to $\mathrm{sign}(\sin\Delta\lambda)$. It is only important for timelike separated points, for which the argument of the square root is negative. In the three dimensional Kruskal co-ordinates on AdS_3, the identification is only in the angular direction. For timelike separated points, $\mathrm{sign}\Delta\lambda=\mathrm{sign}\Delta V$, where V is the Kruskal time. It follows that for all identified points the sign of $i\epsilon$ in $G(x,x'_n)$ is the same.

We now work in the black hole co-ordinates (t,r,ϕ), so that the identification taking AdS_3 into the black hole spacetime is given by $\phi\to\phi+2\pi n$. Under this identification, the two point function on the black hole background becomes

$$G^+(x,x')=\frac{1}{4\sqrt{2}\pi\ell}\left[G_1^+(x,x')\pm G_2^+(x,x')\right]$$

where for $x,x'\in$ region R

$$G_1^+(x,x')=\sum_{n=-\infty}^{\infty}\left[\frac{rr'}{r_+^2}\cosh\frac{r_+(\Delta\phi+2\pi n)}{\ell}-1-\frac{(r^2-r_+^2)^{\frac{1}{2}}(r'^2-r_+^2)^{\frac{1}{2}}}{r_+^2}\cosh\frac{r_+(\Delta t-i\epsilon)}{\ell^2}\right]^{-\frac{1}{2}} \tag{4.4.7}$$

$$G_2^+(x,x') = \sum_{n=-\infty}^{\infty} \left[\frac{rr'}{r_+^2} \cosh \frac{r_+(\Delta\phi + 2\pi n)}{\ell} + 1 - \frac{(r^2 - r_+^2)^{\frac{1}{2}}(r'^2 - r_+^2)^{\frac{1}{2}}}{r_+^2} \cosh \frac{r_+(\Delta t - i\epsilon)}{\ell^2} \right]^{-\frac{1}{2}} . \tag{4.4.8}$$

For $x, x' \in$ region F,

$$G_1^+(x,x') = \sum_{n=-\infty}^{\infty} \left[\frac{rr'}{r_+^2} \cosh \frac{r_+(\Delta\phi + 2\pi n)}{\ell} - 1 + \frac{(r_+^2 - r^2)^{\frac{1}{2}}(r_+^2 - r'^2)^{\frac{1}{2}}}{r_+^2} \cosh \frac{r_+\Delta t}{\ell^2} + i\bar{\epsilon} \right]^{-\frac{1}{2}}$$

$$G_2^+(x,x') = \sum_{n=-\infty}^{\infty} \left[\frac{rr'}{r_+^2} \cosh \frac{r_+(\Delta\phi + 2\pi n)}{\ell} + 1 + \frac{(r_+^2 - r^2)^{\frac{1}{2}}(r_+^2 - r'^2)^{\frac{1}{2}}}{r_+^2} \cosh \frac{r_+\Delta t}{\ell^2} + i\bar{\epsilon} \right]^{-\frac{1}{2}} .$$

Where $\bar{\epsilon} = \epsilon$ sign ΔV and of course in this region sign $\Delta V \neq$ sign Δt. For $x \in$ region R and $x' \in$ region F, we have

$$G_1^+(x,x') = \sum_{n=-\infty}^{\infty} \left[\frac{rr'}{r_+^2} \cosh \frac{r_+(\Delta\phi + 2\pi n)}{\ell} - 1 - \frac{(r^2 - r_+^2)^{\frac{1}{2}}(r_+^2 - r'^2)^{\frac{1}{2}}}{r_+^2} \sinh \frac{r_+\Delta t}{\ell^2} + i\bar{\epsilon} \right]^{-\frac{1}{2}}$$

$$G_2^+(x,x') = \sum_{n=-\infty}^{\infty} \left[\frac{rr'}{r_+^2} \cosh \frac{r_+(\Delta\phi + 2\pi n)}{\ell} + 1 - \frac{(r^2 - r_+^2)^{\frac{1}{2}}(r_+^2 - r'^2)^{\frac{1}{2}}}{r_+^2} \sinh \frac{r_+\Delta t}{\ell^2} + i\bar{\epsilon} \right]^{-\frac{1}{2}} .$$

In all of these expressions, $G^-(x,x') = \langle 0|\phi(x')\phi(x)|0\rangle$ is obtained by reversing the sign of $i\epsilon$. All of these expressions are uniformly convergent for x, x' real, and $r, r' > \delta$, $\delta > 0$. Notice that as $M \to 0$ ($r_+ \to 0$), $G(x,x')$ will diverge like $\sum \frac{1}{n}$ unless we take the Dirichlet boundary conditions.

From these expressions the Feynman Green's function can easily be constructed and in fact has exactly the same form, but with the sign of $i\epsilon$ being strictly positive. It should be noted that none of these Green's functions are invariant under Anti de Sitter transformations, as the Killing vector field defining the identification does not commute with all the generators of the AdS group.

4.4.2 KMS condition

A thermal noise satisfies a skew periodicity in imaginary time called Kubo-Martin-Schwinger (KMS) condition [71]

$$g_\beta(\Delta\tau - \frac{i}{T}) = g_\beta(-\Delta\tau)$$

where $g_\beta(\Delta\tau) = G_\beta^+(x(\tau), x(\tau'))$ and $G_\beta^+ = \langle 0_\beta|\phi(x)\,\phi(x')|0_\beta\rangle\big|_{\mathrm{Im}\ x^0=-\epsilon}$ with the world line $x(\tau)$ taken to be the one at rest with respect to the heat bath (for a more extensive discussion of the KMS condition, see [74]). We will show that $g(\Delta\tau) = G_A^+(x(\tau), x(\tau'))$ with $x(\tau) = (\frac{\tau}{b}, r, \phi)$ and $b = (r^2 - r_+^2)^{1/2}/\ell$, satisfies this condition outside the horizon, with a local temperature $T = \frac{r_+}{2\pi\ell(r^2 - r_+^2)^{1/2}}$, which agrees with the Tolman relation [75] $T = (g_{00})^{-1/2} T_0$, with $T_0 = r_+/2\pi\ell^2$ the temperature of the black hole.

$g(\Delta\tau)$ is defined as

$$g(\Delta\tau) = \frac{1}{4\sqrt{2}\pi\ell}(g_1(\Delta\tau) \pm g_2(\Delta\tau))$$

where

$$g_1(\Delta\tau) = \sum_{n=-\infty}^{\infty}\left[\frac{r^2}{r_+^2}\cosh\frac{2\pi n r_+}{\ell} - 1 - \frac{(r^2 - r_+^2)}{r_+^2}\cosh\left(\frac{r_+}{\ell^2}(\frac{\Delta\tau}{b} - i\epsilon)\right)\right]^{-\frac{1}{2}}$$
$$g_2(\Delta\tau) = \sum_{n=-\infty}^{\infty}\left[\frac{r^2}{r_+^2}\cosh\frac{2\pi n r_+}{\ell} + 1 - \frac{(r^2 - r_+^2)}{r_+^2}\cosh\left(\frac{r_+}{\ell^2}(\frac{\Delta\tau}{b} - i\epsilon)\right)\right]^{-\frac{1}{2}}.$$

We will demonstrate the KMS property for each term in these sums.

Take a typical term in the sum. It has singularities at (p is an integer).

$$\Delta\tau_n = \pm\Delta\tau_n^0 + \frac{i}{T}p + i\epsilon$$

These singularities are square root branch points and the branch cuts go from $\left(\Delta\tau_n^0 + 2\pi\frac{i}{T}p + i\epsilon \to \infty + \frac{i}{T}p + i\epsilon\right)$ and $\left(-\Delta\tau_n^0 + 2\pi\frac{i}{T}p + i\epsilon \to -\infty - 2\pi\frac{i}{T}p + i\epsilon\right)$. In any region without the branch cuts, g_1 and g_2 are analytic. Going around a branch point gives an additional minus sign. Now for a given n, if $\Delta\tau$ is such that the expression inside the square root is positive, then $|\text{Real }\Delta\tau| < \Delta\tau_n$. In this region, g_1^n and g_2^n are analytic and periodic in $\frac{i}{T}$. What's more $g^n(-\Delta\tau) = g^n(\Delta\tau)$ as $\epsilon \to 0$. If on the other hand the expression inside the square root is negative, then because of the branch cuts, $g^n(\Delta\tau - \frac{i}{T}) = -g^n(\Delta\tau)$. As $g^n(\Delta\tau) = (-A + i\epsilon\text{ sign }\Delta\tau)^{-\frac{1}{2}}$ and because our definition of the square root is with a branch cut along the negative real axis, we see that $g^n(\Delta\tau) = -g^n(-\Delta\tau)$ (A is only a function of $|\Delta\tau|$). This shows that the KMS condition is satisfied, and hence that G^+ is a thermal Green's function.

4.4.3 Identifying the Vacuum State

In the region R where $r > r_+$, the Kruskal co-ordinates are defined as

$$U = \left(\frac{r - r_+}{r + r_+}\right)^{\frac{1}{2}}\cosh\frac{r_+}{\ell^2}t$$
$$V = \left(\frac{r - r_+}{r + r_+}\right)^{\frac{1}{2}}\sinh\frac{r_+}{\ell^2}t.$$

Defining $\bar{V} = V - U$ and $\bar{U} = V + U$, r is given by

$$\frac{r}{r_+} = \frac{1 - \bar{U}\bar{V}}{1 + \bar{U}\bar{V}}.$$

In these co-ordinates the two point function becomes

$$
\begin{aligned}
G_J^{+}(x,x') &= \frac{1}{\sqrt{2}4\pi\ell}\sum_n \left\{ \frac{1}{(1+\bar{U}\bar{V})(1+\bar{U}'\bar{V}')} \times \right. \\
&\left[(1-\bar{U}\bar{V})(1-\bar{U}'\bar{V}')\cosh\left(\frac{r_+}{\ell}(\Delta\phi+2\pi n)\right) \mp (1+\bar{U}\bar{V})(1+\bar{U}'\bar{V}') \right. \\
&+ \left.\left. 2(\bar{V}\bar{U}'+\bar{U}\bar{V}') + 2i\epsilon \,\mathrm{sign}\, \Delta V \right] \right\}^{-\frac{1}{2}}
\end{aligned}
$$

where $\mp$ is for $J=1,2$. Here the sign of $i\epsilon$ is the same as $\mathrm{sign}\,\Delta V$ which is the same as $\mathrm{sign}\Delta\lambda$ for timelike separated points. For $x, x' \in$ R, this is just $\mathrm{sign}\Delta t = \mathrm{sign}\,(\bar{V}\bar{U}' - \bar{U}\bar{V}')$. This expression is valid all over the Kruskal manifold.

As discussed in the Introduction, the Hartle-Hawking Green's function is defined to be analytic and bounded in the lower half complex plane of $\bar{V}$ on the past horizon ($\bar{U}=0$), when $\bar{U}', \bar{V}', \phi, \phi'$ are real, or in the lower half plane of $\bar{U}$ on the future horizon ($\bar{V}=0$).

On the past horizon $\bar{U}=0$ we have

$$
\begin{aligned}
G_J^{+} = \frac{1}{\sqrt{2}4\pi\ell}\sum_n \Bigg\{ \frac{1}{1+\bar{U}'\bar{V}'} \Big[&(1-\bar{U}'\bar{V}')\cosh\left(\frac{r_+}{\ell}(\Delta\phi+2\pi n)\right) \\
&\mp(1+\bar{U}'\bar{V}') + 2\bar{V}\bar{U}' + 2i\epsilon\Delta V \Big] \Bigg\}^{-\frac{1}{2}}.
\end{aligned}
$$

In order to prove analyticity and boundedness we will show that the singularities occur in the upper half plane of $\bar{V}$. Hence every term in the sum is a holomorphic function in the lower half plane. We will then use Wierstrass's Theorem on sums of holomorphic functions in order to prove that the G_J are analytic in the lower half of the complex $\bar{V}$ plane.

G_J^{+} has singularities when

$$
\bar{V} = \frac{\pm(1+\bar{U}'\bar{V}') - (1-\bar{U}'\bar{V}')\cosh\left(\frac{r_+}{\ell}(\Delta\phi+2\pi n)\right)}{2(1+i\epsilon)\bar{U}'},
$$

Now suppose that $x' \in$ R, then $-1 \le \bar{U}'\bar{V}' < 0$ and $\bar{U}' > 0$. Defining $\bar{U}\bar{V}' = -a$, $(1 > a > 0)$, the singularity occurs at

$$
\bar{V} = \frac{\pm(1-a) - (1+a)\cosh\left(\frac{r_+}{\ell}(\Delta\phi+2\pi n)\right)}{2(1+i\epsilon)\bar{U}'}.
$$

We see that when $\epsilon \to 0$, $\bar{V}$ is real and negative. Hence $\bar{V} = \frac{-A}{1+i\epsilon} \simeq -A + i\epsilon$ with $A > 0$, so that the singularities are in the upper half plane. Similarly, for the future horizon $\bar{V}=0$, there are singularities when

$$
\bar{U} = \frac{\pm(1+\bar{U}'\bar{V}') - (1-\bar{U}'\bar{V}')\cosh(\Delta\phi+2\pi n)}{(1-i\epsilon)\bar{V}'}.
$$

For $x' \in \mathrm{R}$, then $-1 < \bar{U}'\bar{V}' < 0$, and $\bar{V}' < 0$, so that $\bar{U} = \frac{A}{1-i\epsilon} = A + i\epsilon$, with $A > 0$, so the singularities are in the upper half plane of $\bar{U}$. At this point it should be noted that for G_J^- we get singularities in the lower half plane of $\bar{U}$ on the surface $\bar{V} = 0$, and singularities in the lower half plane of $\bar{V}$ on the surface $\bar{U} = 0$.

For $x' \in \mathrm{F}$, if $\bar{U} = 0$ and x and x' connected by a null geodesic, then $\Delta V < 0$. This is the case because for timelike and null separations, $\operatorname{sign}\Delta V = -\operatorname{sign}\Delta r$ (r is a timelike co-ordinate in F) and Δr is always positive if x is on the horizon. Then it can be checked that the singularities are again in the upper half plane of either $\bar{U}$ or $\bar{V}$.

Now that we have established that each term in the infinite sum is holomorphic in the lower half plane of $\bar{V}$ on the past horizon (and in $\bar{U}$ on the future horizon) we will use Weierstrass's Theorem. This states [76] that if a series with analytic terms

$$f(z) = f_1(z) + f_2(z) + \cdots$$

converges uniformly on every compact subset of a region Ω, then the sum $f(z)$ is analytic in Ω, and the series can be differentiated term by term. It is easily seen that unless $\bar{U}'\bar{V}' = 1$, *i.e.* x' is at the singularity, the sum converges uniformly on every compact subset of the lower half plane. For $\bar{U}'\bar{V}' = 1$ the sum diverges and the Green's function becomes singular at $r = 0$. This is because $r = 0$ is a fixed point of the identification.

To conclude we have shown that our Green's function is analytic on the past horizon in the lower half $\bar{V}$ complex plane, and similarly on the future horizon in the lower half $\bar{U}$ complex plane. Its singularities occur when x, x' can be connected by a null geodesics either directly or after reflection at infinity (see Appendix A and Ref. [70]). We conclude that the Green's function we have constructed is the Hartle-Hawking Green's function as defined in the Introduction, for both Neumann and Dirichlet boundary conditions.

4.4.4 The $M = 0$ Green's function

The black hole solution as $M \to 0$ is the spacetime with metric [65]

$$ds^2 = -\left(\frac{r}{\ell}\right)^2 dt^2 + \left(\frac{\ell}{r}\right)^2 dr^2 + r^2 d\phi^2$$

with $r > 0$, and t and ϕ as in (4.3.5). Defining $z = \frac{\ell}{r}$ and $\gamma = \frac{t}{\ell}$ the metric becomes

$$ds^2 = \frac{\ell^2}{z^2}(-d\gamma^2 + dz^2 + d\phi^2). \tag{4.4.9}$$

The modes for a massless conformally coupled scalar field are solutions of the equation

$$\Box\phi - \frac{1}{8}R\phi = 0$$

where again $R = -6\ell^{-2}$, and are given by

$$\phi_{km} = N_\omega \sqrt{\frac{z}{l}}\, e^{-i\omega t} e^{im\phi} e^{ikz}$$

where $\omega^2 = k^2 + m^2$, m is an integer, and $N_\omega = (8\pi^2\omega)^{\frac{1}{2}}$ is a normalization constant such that $(\psi_{mk}, \psi_{m'k'}) = \delta_{mm'}\delta(k-k')$.

As in quantization on AdS$_3$, care must be taken at the boundary $z = 0$, which is at spatial infinity. The metric (4.4.9) is conformal to Minkowski spacetime with one spatial co-ordinate periodic and the other restricted to be greater than zero. As in the case of AdS$_3$, we impose the boundary conditions

$$\frac{1}{\sqrt{z}}\psi = 0 \qquad \text{or} \qquad \frac{\partial}{\partial z}(\frac{1}{\sqrt{z}}\psi(z)) = 0$$

at $z = 0$, corresponding to Dirichlet or Neumann boundary conditions in the conformal Minkowski metric. Our approach will be to first calculate the Green's function without boundary conditions and them use the method of images to impose them.

Summing modes, we obtain the two point function

$$\begin{aligned}
\tilde{G}(x,x') &= \frac{1}{8\pi^2\omega}\frac{\sqrt{zz'}}{\ell}\sum_m \int_k e^{-i\omega\Delta\gamma}\, e^{im\Delta\phi}\, e^{ik\Delta z}\, dk \\
&= \frac{1}{2\pi}\frac{\sqrt{zz'}}{\ell}\sum_m e^{im\Delta\phi}\; G_2(y,y',m)
\end{aligned}$$

where $G_2(y,y',m)$ is the massive 1+1 dimensional Green's function and $y = (\gamma, z)$.

Now [4],

$$\begin{aligned}
G_2(y,y',m) &= \frac{1}{2\pi}K_0(|m|\,d) \qquad m \neq 0 \\
&= -\frac{1}{2\pi}\log d + \lim_{n\to 2}\frac{\Gamma(\frac{n}{2}-1)}{n\pi^{\frac{n}{2}}} \qquad m = 0
\end{aligned}$$

where $d = \epsilon + i\Delta\ \mathrm{sign}\,\Delta t$, with $\Delta = ((\Delta\gamma)^2 - (\Delta z)^2)^{\frac{1}{2}}$ for timelike separation, and $d = ((\Delta z)^2 - (\Delta\gamma - i\epsilon)^2)^{\frac{1}{2}}$ for spacelike separation. Here K_0 is a modified Bessel function. It follows that

$$\tilde{G}(x,x') = \frac{1}{4\pi^2}\frac{\sqrt{zz'}}{\ell}\left[2\sum_{m>0}\cos(m\Delta\beta)K_0(md) - \log d\right]$$

where the infinite constant in the $m = 0$ expression was dropped to regularize the infrared divergencies of the 1+1 dimensional Green's function. Using [77]

$$\sum_{m=1}^{\infty} K_0(mx)\cos(mxt) = \frac{1}{2}(c+\ln\frac{x}{4\pi}) + \frac{\pi}{2\sqrt{x^2+(xt)^2}} + \frac{\pi}{2}\sum_{l\neq 0}\left[\left(x^2 + (2\ell\pi - tx)^2\right)^{-\frac{1}{2}} - \frac{1}{2\ell\pi}\right].$$

Here c is the Euler constant. This expression is valid for $x > 0$ and real t, and gives the following expression for spacelike separated x and x':

$$\begin{aligned}\tilde{G}(x,x') &= \frac{1}{4\pi}\frac{\sqrt{zz'}}{\ell}\left[\sum_n\left[(\Delta z)^2+(\Delta\phi+2\pi n)^2-(\Delta\gamma-i\epsilon)^2\right]^{-\frac{1}{2}}-\sum_{n\neq 0}\frac{1}{2\pi n}+c_1\right]\\ &= \frac{1}{4\pi}\frac{\sqrt{zz'}}{\ell}F(x,x')\end{aligned}$$

Here $c_1 = c - \ln 4\pi$. Although the above formula was only true for $x > 0$ and real t, the result is analytic for every real Δz, $\Delta\gamma$ and $\Delta\phi$. Hence it is also correct for timelike separated points.

This is just what is expected from the conformality to the Minkowski space, other than the $\sum \frac{1}{2\pi\ell} - c_1$ which regularizes $\tilde{G}$. Now the boundary condition can be easily put in by writing

$$G^{+}(x,x') = \frac{1}{4\pi}\frac{\sqrt{zz'}}{\ell}\left(F(x,x') \pm F(x,\bar{x}')\right)$$

where $\bar{x}' = (\gamma', -z', \phi)$. Notice that for Dirichlet boundary conditions, this agrees with the $M \to 0$ limit of (4.4.7) and (4.4.8).

Going to the (t, r, ϕ) co-ordinates we have

$$G^{+}_{M=0} = \frac{1}{4\pi}(rr')^{-\frac{1}{2}}(G^{+}_{1} \pm G^{+}_{2})$$

with

$$\begin{aligned}G_1^{+} &= \sum_n\left[\ell^2\left(\frac{r'-r}{rr'}\right)^2+(\Delta\phi+2\pi n)^2-\left(\frac{\Delta t-i\epsilon}{\ell}\right)^2\right]^{-\frac{1}{2}}-\sum_{n\neq 0}\frac{1}{2\pi n}+c_1\\ G_2^{+} &= \sum_n\left[\ell^2\left(\frac{r'+r}{rr'}\right)^2+(\Delta\phi+2\pi n)^2-\left(\frac{\Delta t-i\epsilon}{\ell}\right)^2\right]^{-\frac{1}{2}}-\sum_{n\neq 0}\frac{1}{2\pi n}+c_1\end{aligned}$$

4.4.5 Computation of $\langle\phi^2\rangle$

$\langle\phi^2\rangle$ is defined as $\langle\phi^2\rangle = \lim_{x\to x'}\frac{1}{2}G_{\mathrm{Reg}}(x,x')$ where $G = G^{+} + G^{-}$ is the symmetric Green's function. In order to compute $\langle\phi^2\rangle$, we need to regularize G. Now only the $n = 0$ term in G_1^{+} is infinite and is just a Green's function on AdS_3. Hence, we can use the Hadamard development in AdS_3 to regularize G [4]:

$$G_{\mathrm{Had}} = \frac{-i}{2\sqrt{2}\pi}\frac{\Delta^{\frac{1}{2}}}{\sigma^{\frac{1}{2}}}$$

where

$$\sigma = \frac{\ell^2}{2}\left[ar\cos Z\right]^2, \qquad \Delta^{-\frac{1}{2}} = \frac{\sin\left(\frac{2\sigma}{\ell^2}\right)^{\frac{1}{2}}}{\left(\frac{2\sigma}{\ell^2}\right)^{\frac{1}{2}}} \qquad \text{and} \qquad Z = \frac{\cos\Delta\lambda - \sin\rho\sin\rho'\cos\Delta\theta}{\cos\rho\cos\rho'}$$

(here Δ is the Van Veleck determinant). Defining

$$G_{\mathrm{Reg}}(x,x') = G_{\mathrm{BH}}(x,x') - G_{\mathrm{Had}}(x,x')$$

we get

$$\langle\phi^2\rangle = \frac{1}{4\sqrt{2}\pi\ell}\frac{r_+}{r}\left[\sum_{n\neq 0}\left(\cosh\left(\frac{r_+}{\ell}2\pi n\right)-1\right)^{-\frac{1}{2}} \pm \sum_{n}\left(\cosh\left(\frac{r_+}{\ell}2\pi n\right)-1+2\left(\frac{r_+}{r}\right)^2\right)^{-\frac{1}{2}}\right]$$

which, for Dirichlet boundary conditions, can be seen to be regular as $M \to 0$ (that is $r_+ \to 0$), and to coincide in this limit with the $M = 0$ result for Dirichlet boundary conditions.

4.5 The Energy-Momentum Tensor

The energy-momentum tensor for a massless conformally coupled scalar field in AdS$_3$ is given by the expression

$$T_{\mu\nu}(x) = \frac{3}{4}\partial_\mu\phi(x)\partial_\nu\phi(x) - \frac{1}{4}g_{\mu\nu}g^{\rho\sigma}\partial_\rho\phi(x)\partial_\sigma\phi(x) - \frac{1}{4}\nabla_\mu\partial_\nu\phi(x)\phi(x) + \frac{1}{96}g_{\mu\nu}R\phi^2(x)$$

where $R = -6\ell^{-2}$. In order to compute $\langle T_{\mu\nu}\rangle$ one differentiates the symmetric two-point function $G = \langle 0|\phi(x)\,\phi(x') + \phi(x')\,\phi(x)|0\rangle$ [4], and then takes the coincident point limit. This makes $\langle T_{\mu\nu}\rangle$ divergent and regularization is needed. A look at our Green's function reveals that only the $n = 0$ term in G_1 diverges as $x \to x'$, so only the $\langle T_{\mu\nu}\rangle$ derived from it should be regularized.

The $n = 0$ term is just the Green's function in AdS$_3$ in accelerating co-ordinates. The vacuum in which this Green's function is derived is symmetric under the Anti de Sitter group and AdS$_3$ is a maximally symmetric space. Hence [78] $\langle T_{\mu\nu}\rangle = \frac{1}{3}g_{\mu\nu}\langle T\rangle$ where $\langle T\rangle = g^{\mu\nu}\langle T_{\mu\nu}\rangle$. For a conformally coupled massless scalar field $\langle T\rangle = 0$ (there is no conformal anomaly in 2+1 dimensions) so $\langle T^{\mathrm{AdS}}_{\mu\nu}\rangle = 0$.

Having shown that we may drop the $n = 0$ term in G_1, after a somewhat lengthy calculation we arrive at the result for $M \neq 0$,

$$\langle T^\nu_\mu(x)\rangle = \frac{1}{16\pi\ell^3 r^3}\sum_{n>0}\left\{\left[\frac{r_+^2 f_n^{-1}}{2}\left[1 \pm \left(1+(f_n r)^{-2}\right)^{-\frac{3}{2}}\right] + f_n^{-3}\right]\ \mathrm{diag}(1,1,-2)\right.$$
$$\left.\pm\frac{3}{2}\left(1-\frac{r_+^2}{r^2}\right)f_n^{-3}\left(1+(f_n r)^{-2}\right)^{-\frac{5}{2}}\ \mathrm{diag}(1,0,-1)\right\} \tag{4.5.10}$$

where $f_n = \sinh(\frac{r_+}{\ell}\pi n)/r_+$. diag$\,(a,b,c)$ is in (t,r,ϕ) co-ordinates. As expected the $n = 0$ term from G_2 did not contribute.

For $M = 0$ we get from Sec. 4.4.4

$$\langle T^\mu_\nu(x)\rangle = \frac{1}{16\pi r^3}\sum_{n>0}\left\{\frac{1}{(n\pi)^3}\ \mathrm{diag}\,(1,1,-2)\right.$$
$$\left.\pm\frac{3}{2(n\pi)^3}\left(1+(f_n r)^{-2}\right)^{-\frac{5}{2}}\ \mathrm{diag}\,(1,0,-1)\right\} \tag{4.5.11}$$

where now $f_n = \pi n/\ell$. Note that the $M = 0$ result agrees with the $M \to 0$ limit of (4.5.10).

Some properties of $\langle T^{\nu}_{\mu} \rangle$ are:

- As we can see from (4.5.10), far away from the black hole, $\langle T^{\mu}_{\nu} \rangle$ obeys the strong energy condition [73] only for the Dirichlet boundary conditions, while for the Neumann boundary conditions, the energy density is negative in this limit.

- For Dirichlet boundary conditions, as M decreases, although the temperature decreases, the energy density increases; just the opposite occurs for Neumann boundary conditions.

- In the limit $M \to \infty$, $\langle T^{\mu}_{\nu} \rangle \to 0$ for both sets of boundary conditions, which suggests the presence of a Casimir effect.

- On the horizon, $\langle T^{\mu}_{\nu} \rangle$ is regular, and hence in the semiclassical approximation, the horizon is stable to quantum fluctuations; on the other hand, at $r = 0$, $\langle T^{\mu}_{\nu} \rangle$ diverges.

- Our Green's function was thermal in (t, r, ϕ) co-ordinates, but although $\langle T^{\mu}_{\nu} \rangle \sim T^3_{\text{loc}}$ for large r, it is not of a thermal type [75].

4.6 The response of a Particle detector

In this section we calculate the response of a particle detector which is stationary in the black hole co-ordinates (t, r, ϕ), and outside the black hole. The simplest particle detector can be described by an idealized point monopole coupled to the quantum field through an interaction described by $\mathcal{L}_{\text{int}} = cm(\tau)\,\phi[x(\tau)]$ where τ is the detector's proper time, and $c \ll 1$. The probability per unit time for the detector to undergo a transition from energy E_1 to E_2 [4] is $R(E_1/E_2) = c^2\,|\langle E_2|m(0)|E_1\rangle|^2\,F(E_2 - E_1)$ to lowest order in perturbation theory, where

$$F(\omega) = \lim_{s\to 0}\lim_{\tau_0\to\infty}\frac{1}{2\tau_0}\int_{-\tau_0}^{\tau_0} d\tau \int_{-\tau_0}^{\tau_0} d\tau'\, e^{-i\omega(\tau-\tau')-S|\tau|-S|\tau'|}\; g(\tau,\tau').$$

$g(\tau,\tau') = G^{+}(x(\tau), x(\tau'))$ and $x(\tau)$ is the detector trajectory.

$F(\omega)$ is called the response function. It represents the bath of particles that the detector sees during its motion [79]. We take $x(\tau) = (\frac{\tau}{b}, r, \phi)$ where $b = \left(\frac{r^2 - r_+^2}{\ell^2}\right)^{\frac{1}{2}}$. Because $g(\tau,\tau') = g(\Delta\tau)$, then

$$F(\omega) = \int_{-\infty}^{\infty} e^{-i\omega\Delta\tau}\, g(\Delta\tau)d(\Delta\tau) \tag{4.6.12}$$

where $g(\Delta\tau) = g_1(\Delta\tau) \pm g_2(\Delta\tau)$. Here

$$g_i(\Delta\tau) = \frac{r_+}{\sqrt{24}\pi\ell}(r^2 - r_+^2)^{-\frac{1}{2}} \sum_n \left[\frac{r_+^2}{r^2 - r_+^2}\left(\frac{r^2}{r_+^2}\cosh\left(\frac{r_+}{\ell}2\pi n\right) \mp 1\right) - \cosh\frac{r_+}{\ell^2}\left(\frac{\Delta\tau}{b} - i\epsilon\right)\right]^{-\frac{1}{2}}. \tag{4.6.13}$$

In this expression $-(+)$ is for $g_1(g_2)$.

Defining

$$\cosh\alpha_n = \frac{r_+^2}{r^2-r_+^2}\left(\frac{r^2}{r_+^2}\cosh\left(\frac{r_+}{\ell}2\pi n\right)-1\right)$$

and

$$\cosh\beta_n = \frac{r_+^2}{r^2-r_+^2}\left(\frac{r^2}{r_+^2}\cosh\left(\frac{r_+}{\ell}2\pi n\right)+1\right)$$

we have from Appendix B that

$$F(\omega) = \frac{1}{2}\,\frac{1}{e^{\omega/T}+1}\sum_n\left(P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n) \pm P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\beta_n)\right)$$

where $T = \frac{r_+}{2\pi\ell(r^2-r_+^2)^{\frac{1}{2}}}$ is the local temperature. This looks like a fermion distribution with zero chemical potential and a density of states

$$D(\omega) = \frac{\omega}{2\pi}\sum_n\left(P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n) \pm P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\beta_n)\right).$$

Notice that $F(\omega)$ is finite on the horizon in contrast with black holes in two and four dimensions (see [80]). This seems to be a consequence of the Fermi type distribution. Statistical inversion in odd-dimensional flat spacetime was first noted in Ref. [74].

If the mass of the black hole satisfies $e^{2\pi\sqrt{M}} \gg 1$ and $2\ell^{-1} > \omega \gg T$, then far from the horizon, $r \gg r_+$, we can sum the series, and for Dirichlet boundary conditions we obtain

$$F(\omega) \simeq 2\pi^2\ell^2T^2\frac{1}{e^{\omega/T}+1}\left[\left(\frac{\omega}{2\pi T}\right)^2+\frac{1}{4}+\frac{8e^{-3\pi\sqrt{M}}}{\pi}\left(\frac{\omega}{T}\right)^{\frac{1}{2}}\left(\sin\frac{w\sqrt{M}}{T}-\cos\frac{w\sqrt{M}}{T}\right)\right].$$

A similar result holds for Neumann boundary conditions at large r,

$$F(\omega) \simeq \frac{1}{e^{\omega/T}+1}\left[1+4\left(\frac{\omega}{T}\right)^{-\frac{1}{2}}e^{-\pi\sqrt{M}}\left(\sin\frac{\omega\sqrt{M}}{T}+\cos\frac{\omega\sqrt{M}}{T}\right)\right]$$

where the approximation improves for large M as before.

It seems clear that the particle detector response will consist of a Rindler-type effect [4], and, if present, a response due to Hawking radiation (real particles). The former is due to the fact that a stationary particle detector is actually accelerating, even when $r \to \infty$ (there is no asymptotically flat region). This is reflected for instance in the fact that for some range of ω, the behaviour of $F(\omega)$ for $r \gg r_+$ and $M \gg 1$ is governed by the $n = 0$ term in G^+, which is AdS invariant. Hence all observers connected by an AdS transformation (a subgroup of the asymptotic symmetry group) register the same response, even though they might be in relative motion; this means that $F(\omega)$ as a whole cannot be interpreted as real particles (see [81, 82] for a discussion of this point). Unfortunately, one cannot filter out these effects in a simple way, and further work is needed in order to find the spectrum of the Hawking radiation.

Finally, for $M = 0$, we may again define

$$F_i(\omega) = \int_{-\infty}^{\infty} e^{-i\omega\Delta\tau} g_i(\Delta\tau) d\Delta\tau.$$

Now, however, g_1 and g_2 are analytic in the lower half complex plane of $\Delta\tau$. Hence for $\omega > 0$ we can close the integral in an infinite semicircle in the lower half plane and by Cauchy's theorem $F_i(\omega) = 0$ for $\omega > 0$ so that no particles are detected by a stationary particle detector.

4.7 Back-reaction

In this section we shall discuss the back-reaction on the BTZ solutions due to quantum fluctuations, using the energy momentum tensor $\langle T^\mu_\nu \rangle$ derived in Sec. 4.5. We shall show that for all M, including $M = 0$, divergencies in the energy momentum tensor cause the curvature scalar $R_{\mu\nu}R^{\mu\nu}$ to blow up at $r = 0$ (note that since the energy momentum tensor is traceless, R does not blow up). It is also interesting to consider the effects of back-reaction on the location of the horizon, even thought this is only an order $\hbar$ effect. It is possible to show for all $M \neq 0$ that the horizon shifts outwards under the effect of quantum fluctuations. For $M = 0$, the effect is that a horizon develops at a radius of order $\hbar$, but where we may still be able to trust the semi-classical approximation.

We compute the back-reaction in the usual way by inserting the expectation value of the energy-momentum tensor (4.5.10) or (4.5.11), into Einstein's equations,

$$G_{\mu\nu} = \ell^{-2} g_{\mu\nu} + \pi \langle T_{\mu\nu} \rangle.$$

The first thing to note is that although the external solution is of constant curvature everywhere, the perturbed solution is not, and the curvature scalar $R_{\mu\nu}R^{\mu\nu}$ diverges at the origin, $r = 0$. Einstein's equations give

$$\begin{aligned} R_{\mu\nu}R^{\mu\nu} &= (\pi\langle T^\mu_\nu \rangle - 2\ell^{-2}\delta^\mu_\nu)(\pi\langle T^\nu_\mu \rangle - 2\ell^{-2}\delta^\nu_\mu) \\ &= \pi^2 \langle T^\mu_\nu \rangle \langle T^\nu_\mu \rangle + 12\ell^{-4} \\ &> \pi^2 \langle T^r_r \rangle^2 + 12\ell^{-4} = 12\ell^{-4} + \frac{1}{64\ell^6 r^6} \left(\sum_{n>0} f_n^{-3} \right)^2 \end{aligned}$$

The sum in the last expression is a constant depending only on M. In the limit as $M \to 0$, the curvature still diverges as $1/r^6$. Although the divergence in the curvature scalar occurs precisely where the semi-classical approximation is unreliable, the result does say that we must go beyond semi-classical physics in order to describe the region near $r = 0$. This seems to be the natural notion of a singularity at the semi-classical level.

Having shown that the back-reacted metric becomes singular, it remains to look at horizons. We begin with a general static, spherically symmetric metric, which we take to be

$$ds^2 = -N^2\, dt^2 + \frac{dr^2}{N^2} + e^{2A}\, d\phi^2 \quad ,$$

where N and A are functions of r only. A linear combination of Einstein's equations implies that

$$(N^2)'' = 2\ell^{-2} + 2\pi \langle T^\phi_\phi \rangle. \tag{4.7.14}$$

Integrating Eq. (4.7.14) once, and inserting (4.5.10), we obtain the result

$$(N^2)' = \frac{2r}{\ell^2} + \frac{1}{8\ell^3 r^2} \sum_{n>0} \left\{ \frac{r_+^2 f_n^{-1}}{2} \left[1 \pm \left(1 + (f_n r)^{-2}\right)^{-\frac{3}{2}} \right] \right.$$
$$\left. + f_n^{-3} \left[1 \pm \frac{f_n^2 r^2}{2} \left(1 - \left(1 + (f_n r)^{-2}\right)^{-\frac{3}{2}} \right) \right] \right\} \tag{4.7.15}$$

where an integration constant has been included to make the result finite. A second integration gives

$$N^2 = \frac{r^2}{\ell^2} - M - \frac{1}{8\ell^3 r} \sum_{n>0} \left\{ \frac{r_+^2 f_n^{-1}}{2} \left[1 \pm \left(1 + (f_n r)^{-2}\right)^{-\frac{1}{2}} \right] \right.$$
$$\left. + f_n^{-3} \left[1 \pm \frac{1}{2} \left(f_n^2 r^2 \left[\left(1 + (f_n r)^{-2}\right)^{\frac{1}{2}} - 1 \right] + \left(1 + (f_n r)^{-2}\right)^{-\frac{1}{2}} \right) \right] \right\} \tag{4.7.16}$$

where the second integration constant has been set to M, and is the ADM mass of the solution [65]. The two integration constants ensure that $N^2 \to r^2/\ell^2 - M + o(\frac{1}{r})$ as $r \to \infty$.

Having obtained an expression for N, it is also necessary to look at the $g_{\phi\phi}$ component given by A. A is given in terms of N by the equation

$$A' = \frac{16\ell r^3 + \sum_{n>0} \left[\frac{r_+^2 f_n^{-1}}{2} \left[1 \pm \left(1 + (f_n r)^{-2}\right)^{-\frac{3}{2}} \right] + f_n^{-3} \right]}{8\ell^3 r^3 (N^2)'}$$

which we shall not attempt to integrate, although it is easy to see that as $r \to \infty$, $A \to \ln r$. The important thing to notice is that A' diverges only at $r = 0$ or where $(N^2)' = 0$. If the singularity at $r = 0$ is to be taken seriously, it is important that $(N^2)'$ should not vanish for any finite, non-zero r. To see that this is indeed the case, note that since the quantity inside the curly brackets of Eq. (4.7.15), is positive for all $r > 0$, then so is $(N^2)'$.

Having checked that the backreaction does not cause a qualitative change in $g_{\phi\phi}$, and having found the exact change in N, we may examine the horizon structure of the new solutions. Note that each term in the sum in (4.7.16) is strictly positive, and behaves as $1/r$ at infinity and near the origin, for any M. Hence, the horizon of the $M \neq 0$ solutions is pushed out by quantum fluctuations, as compared with the classical solution of the same ADM mass.

The $M = 0$ solution, which acquires a curvature singularity due to backreaction, also develops a horizon. We regard this result as being indicative of the fact that the $M = 0$ solution is unstable, in the sense that the qualitative features of the solution are changed by quantum fluctuations. Recall that $\langle T_\mu^{\ \nu} \rangle$ in this case appears to be just the Casimir energy of the spacetime as it is associated with a zero temperature Green's function. The appearance of a horizon may be contrasted in an obvious way with 4-dimensional Minkowski spacetime, regarded as the $M = 0$ limit of the Schwarzschild solution. Minkowski spacetime has no Casimir energy associated with it, and is stable in the above sense.

Note that as $M \to 0$, the horizon is located in a region sufficiently close to $r = 0$ that the semi-classical approximation may break down, *i.e.* fluctuations in $\langle T^\mu_\nu \rangle$ will be of the order of $\langle T^\mu_\nu \rangle$. However, if there are n independent scalar fields present, then the ratio of the fluctuations to $\langle T^\mu_\nu \rangle$ becomes negligible in the vicinity of the horizon, as n becomes large. The size of the perturbation on the metric near where the horizon develops may also be estimated. It is approximately an order of magnitude smaller than the curvature of the original solution, a result which is independent of n.

We end with a speculation about the endpoint of evaporation. Notice that although classically there is a clear but puzzling distinction between the black hole solutions of BTZ, with $M \geq 0$, the solutions with conical singularities of Deser and Jackiw [83] corresponding to $-1 < M < 0$, and AdS_3 ($M = -1$), semiclassically the difference between the small M and negative M solutions is not so marked. Our results for $M = 0$ are qualitatively similar to those of Refs. [84], where it is shown that quantum fluctuations on a conical spacetime generate a singularity at the apex of the cone, shielded by an order $\hbar$ horizon. One might speculate from this similarity that evaporation could continue beyond the $M = 0$ solution, perhaps ending at AdS_3.

4.8 Conclusions

In this paper we presented some aspects of quantization on the 2+1 dimensional black hole geometry. We obtained an exact expression for the Green's function in the Hartle-Hawking vacuum and for the expectation value of the energy-momentum tensor, but we found some difficulty in interpreting the particle detector response as Hawking radiation. We feel that further investigation on this question is required. If the black hole evaporates, the results of section 4.7 suggest the possibility that due to quantum fluctuations, the endpoint of evaporation may not look like the classical $M = 0$ solution.

4.9 Appendix A: Scalar field quantization on AdS_3

The derivation of a scalar field propagator on AdS_3 is reviewed. This computation is complicated by the fact that AdS_3 is not globally hyperbolic. In the AdS co-ordinate system defined in Sec. 4.3, spatial infinity is the $\rho = \frac{\pi}{2}$ surface which is seen to be timelike (see Fig. 2). Information can escape or leak in through this surface in a finite co-ordinate time, spoiling the composition law property of the propagator. In order to resolve this problem and define a good quantization scheme on AdS_3, we follow [68] and use the fact that AdS_3 is conformal to half of the Einstein Static Universe (ESU) $R \times S^2$.

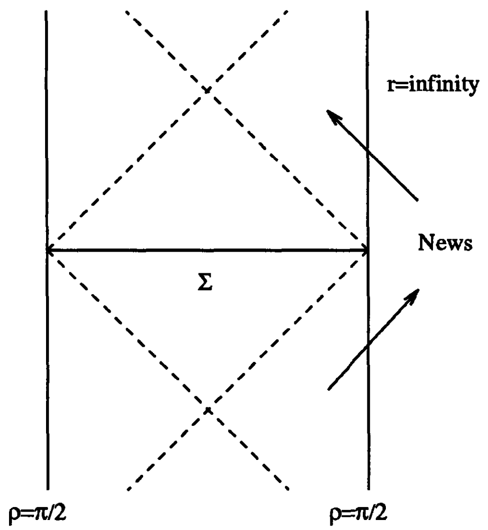

Figure 2: A Penrose diagram of AdS_3. Information can leak in or out through spatial infinity and thus Σ is not a Cauchy surface unless we impose boundary conditions at $r = \infty$.

The metric of ESU is

$$ds^2 = -d\lambda^2 + d\rho^2 + \sin^2 \rho \, d\theta^2$$

where $-\infty < \lambda < \infty$, $0 < \rho \leq \pi$, and $0 < \theta \leq 2\pi$. Positive frequency modes on ESU are solutions of

$$\Box \psi^{\mathrm{E}} - \frac{1}{8} R \psi^{\mathrm{E}} = 0$$

where $R = 2$, and are given by

$$\psi^{\mathrm{E}}_{\ell m} = N_{\ell m} \, e^{-i\omega\lambda} \, Y^{\ell}_{m}(\rho, \theta) \qquad \omega > 0 \tag{4.9.17}$$

where Y_ℓ^m are the spherical harmonics, $\omega = \ell + \frac{1}{2}$, m and ℓ are integers with $\ell \geq 0$, $|m| \leq \ell$, and $N_{\ell m} = \frac{1}{\sqrt{2\ell+1}}$. These modes are orthonormal in the inner product [4]

$$(\psi_1, \psi_2) = -i \int_\Sigma \psi_1 \overleftrightarrow{\partial_\mu} \psi_2^* [-g_\Sigma(x)]^{\frac{1}{2}} d\Sigma^\mu$$

where Σ is a spacelike Cauchy surface. *i.e.* $(\psi_{\ell m}, \psi_{\ell' m'}) = \delta_{\ell\ell'}\delta_{mm'}$, $(\psi_{\ell m}, \psi^*_{\ell' m}) = 0$, and $(\psi^*_{\ell m}, \psi^*_{\ell' m'}) = -\delta_{\ell\ell}\delta_{mm'}$. As usual the field operator is expanded in these modes $\phi = \sum_{\ell,m} \psi_{\ell m} a_{\ell m} + \psi_{\ell m} a^+_{\ell m}$ so that a, a^+ destroy and create particles, and define the vacuum state $|0\rangle_E$.

The two point function is defined as

$$G^+_E(x,x') = {}_E\langle 0|\phi(x)\,\phi(x')|0\rangle_E = \sum \psi^E_{\ell m}(x)\psi^{E*}_{\ell m}(x').$$

Inserting (4.9.17),

$$G^+_{\mathrm{E}}(x,x') = \sum_\ell \frac{1}{2\ell+1} e^{-i(\ell+\frac{1}{2})(\lambda-\lambda')} \sum_m Y^\ell_m(\rho,\theta)\; Y^{*\ell}_m(\rho',\theta').$$

Using $(Y^\ell_m)^* = (-1)^m Y^\ell_{-m}$ and $\sum_{m=-\ell}^{\ell}(-1)^m Y^\ell_m(x) Y^\ell_{-m}(x') = \frac{2\ell+1}{4\pi} P_\ell(\cos\alpha)$ where α is the angle between (ρ,θ) and (ρ',θ'), we get

$$G^+_{\mathrm{E}}(x,x') = \frac{1}{4\pi} e^{-\frac{1}{2}(\lambda-\lambda')} \sum_{\ell=0} e^{-i\ell(\lambda-\lambda')} P_\ell(\cos\alpha).$$

Further, using $\sum_{n=0}^\infty P_n(x)\,z^n = (1-2xz+z^2)^{-\frac{1}{2}}$ for $-1 < x < 1$ and $|z| < 1$ and as usual giving $\Delta\lambda$ a small negative imaginary part for convergence, we get

$$G^+_{\mathrm{E}} = \frac{1}{4\sqrt{2}\pi} \left(\cos(\Delta\lambda - i\epsilon) - \cos\rho\cos\rho' - \sin\rho\sin\rho'\cos\Delta\theta\right)^{-\frac{1}{2}}$$

where the square root is defined with a branch cut along the negative real axis and the argument function is between $(-\pi,\pi)$ [85]. From now we shall call this two point function $G^+_{1,\mathrm{E}}$ and define $G^+_{2,\mathrm{E}}(x,x') = G^+_{1,\mathrm{E}}(\tilde{x},x')$ where $\tilde{x} = (\lambda, \pi-\rho, \theta)$. Then,

$$G^+_{2,\mathrm{E}} = \frac{1}{4\sqrt{2}\pi} \left(\cos(\Delta\lambda - i\epsilon) + \cos\rho\cos\rho' - \sin\rho\sin\rho'\cos\Delta\theta\right)^{-\frac{1}{2}}$$

and $G^+_{2,\mathrm{E}}$ satisfies also the homogeneous equation $(\Box - \frac{1}{8}R)G = 0$. Conformally mapping these solutions to AdS$_3$, where $G^+_{\mathrm{A}} = \sqrt{\cos\rho\cos\rho'}\,G^+_{\mathrm{E}}$ we get

$$G^+_{1,\mathrm{A}}(x,x') = \frac{1}{4\sqrt{2}\pi\ell} \left(\cos(\Delta\lambda - i\epsilon)\sec\rho\sec\rho' - 1 - \tan\rho\tan\rho'\cos\Delta\theta\right)^{-\frac{1}{2}}$$

and

$$G^+_{2,\mathrm{A}}(x,x') = \frac{1}{4\sqrt{2}\pi\ell} \left(\cos(\Delta\lambda - i\epsilon)\sec\rho\sec\rho' + 1 - \tan\rho\tan\rho'\cos\Delta\theta\right)^{-\frac{1}{2}}.$$

It can be seen that $G^+_{1,\mathrm{A}}$ and $G^+_{2,\mathrm{A}}$ are functions of $\sigma(x,x') = \frac{1}{2}[(u-u')^2 + (v-v')^2 + (x-x')^2 + (y-y')^2]$, which is the distance between the spacetime points x, x' in the 4-dimensional embedding space.

In order to deal with the problem of global hyperbolicity, it was shown in [68] that imposing boundary conditions on the ESU modes gives a good quantization scheme on the half of ESU with $\rho \leq \frac{\pi}{2}$, thus inducing a good quantization scheme on AdS$_4$.

It may be checked that this method also works in 2+1 dimensions. The boundary conditions on the ESU modes are either Dirichlet

$$\psi^{\rm E}_{\ell,m}\left(\rho=\frac{\pi}{2}\right)=0 \qquad \text{obeyed by } \psi_{\ell,m} \text{ with } \ell+m= \text{ odd}$$

or Neumann

$$\frac{\partial}{\partial\rho}\psi^{\rm E}_{\ell,m}\left(\rho=\frac{\pi}{2}\right)=0 \qquad \text{obeyed by } \psi_{\ell,m} \text{ with } \ell+m= \text{ even.}$$

It is easily verified that the combination $G^+_{\rm E} = G^+_{1,{\rm E}} \pm G^+_{2,{\rm E}}$ has the right boundary condition where the $+(-)$ signs are for Neumann (Dirichlet) boundary conditions.

Some remarks are in order: if x, x' are restricted such that $-\pi < \lambda(x) - \lambda(x') < \pi$ then

(1) $G^+_{1\,E}$ is real for spacelike points, imaginary for timelike points and singular for x, x' which can be connected by a null geodesic.

(2) $G^+_{2\,E}$ has the same property when $x \to \tilde{x}$, and if $0 \leq \rho(x'), \rho(x) < \frac{\pi}{2}$ then G_2 has singularities when x, x' can be connected by a null geodesic bouncing off $\rho = \frac{\pi}{2}$ boundary.

From this we see that if we take the modes in AdS_3 as

$$\psi^{\rm A}_{\ell,m} = (\cos\rho)^{\frac{1}{2}}\, e^{-i(\ell+\frac{1}{2})\lambda}\, Y^m_\ell(\rho,\theta) \qquad \ell+m= \text{ odd or } \ell+m= \text{ even}$$

then these modes give rise to a well-behaved propagator [68]. The two point function is then

$$G^+_{\rm A} = \sqrt{\cos\rho\cos\rho'}(G^+_{1,{\rm A}} \pm G^+_{2,{\rm A}})$$

where $+(-)$ are for Neumann (Dirichlet). The two point function has singularities whenever x, x' can be connected by a null geodesic directly or by a null geodesic bouncing off infinity (null geodesics remain null geodesics by a conformal transformation). All other properties listed before also stay the same.

Note that it is possible to define a quantization scheme on AdS_3 without using boundary conditions (*i.e.* just using $G^+_{1,A}$), which is referred to as transparent boundary conditions in Ref. [68]. However this requires the use of a two-time Cauchy surface, and its physical interpretation is unclear.

4.10 Appendix B: Calculating the response function

We are interested in an integral of the type

$$J(\omega) = \frac{\ell^2 b}{r_+}\int_{-\infty}^{\infty} e^{-\frac{i\omega}{2\pi T}t}\,(\cosh\alpha_n - \cosh(t-i\epsilon))^{-\frac{1}{2}}\,dt$$

where $T = \frac{r_+}{2\pi\ell(r^2-r_+^2)^{\frac{1}{2}}}$ is the local temperature. $J(\omega) = I_1(\omega) + I_2(\omega) + I_3(\omega)$ where I_1 is the integral from $-\infty$ to $-\alpha_n$, I_2 is from $-\alpha_n$ to α_n, and I_3 is from α_n to ∞. Recall that the square root is defined with the cut along the negative real axis. Then

$$\begin{aligned}
I_1 &= \frac{\ell^2 b}{-ir_+}\int_{-\infty}^{-\alpha_n} e^{-\frac{i\omega t}{2\pi T}}\,(\cosh t - \cosh\alpha_n)^{-\frac{1}{2}} \\
I_3 &= \frac{\ell^2 b}{ir_+}\int_{\alpha_n}^{\infty} e^{-\frac{i\omega t}{2\pi T}}\,(\cosh t - \cosh\alpha_n)^{-\frac{1}{2}} \\
I_2 &= \frac{2\ell^2 b}{r_+}\int_0^{\alpha_n} \cos\frac{\omega t}{2\pi T}\,(\cosh\alpha_n - \cosh t)^{-\frac{1}{2}}dt.
\end{aligned}$$

Using [77]

$$\begin{aligned}
\int_\alpha^\infty \frac{e^{-(\nu+\frac{1}{2})t}}{(\cosh t - \cosh\alpha)^{-\frac{1}{2}}} &= \sqrt{2}\,Q_\nu(\cosh\alpha) \qquad \mathrm{Re}\,\nu > -1 \qquad \alpha > 0 \\
\int_0^\alpha \frac{\cosh(\nu+\frac{1}{2})t}{(\cosh\alpha - \cosh t)^{\frac{1}{2}}} &= \frac{\pi}{\sqrt{2}}P_\nu(\cosh\alpha) \qquad \alpha > 0
\end{aligned}$$

where P_ν and Q_ν are associated Legendre functions of the first and second kind respectively, we get

$$\begin{aligned}
I_3 &= -\frac{i\sqrt{2}\ell^2 b}{r_+}Q_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n) \\
I_2 &= \frac{\sqrt{2}\pi\ell^2 b}{r_+}P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n) \\
I_1 &= \frac{i\sqrt{2}\ell^2 b}{r_+}Q_{\frac{-i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n).
\end{aligned}$$

Now using $Q_\nu(z) - Q_{-\nu-1}(z) = \pi\cot(\nu\pi)P_\nu(z)$ [77]

$$J(\omega) = \frac{2\sqrt{2}\pi\ell^2 b}{r_+}\,P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n)\frac{1}{e^{\omega/T}+1},$$

and $F_{1,2}(\omega)$, defined by (4.6.12) and (4.6.13) in an obvious way, are given by

$$\begin{aligned}
F_1(\omega) &= \frac{1}{2}\,\frac{1}{e^{\omega/T}+1}\,\sum_{n\neq 0} P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n) \\
F_2(\omega) &= \frac{1}{2}\,\frac{1}{e^{\omega/T}+1}\,\sum_{n} P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\beta_n).
\end{aligned}$$

Notice that although the formulae that we used were not correct when $\alpha = 0$, nevertheless the α_0 term came out correctly, since

$$\begin{aligned}
&\int_{-\infty}^{\infty} e^{-i\omega t}(1-\cosh(2\pi T t - i\epsilon))^{-\frac{1}{2}} \\
&= i\int_{-\infty}^{0} e^{-i\omega t}\left|\sqrt{2}\sinh(\pi T t - i\epsilon)\right|^{-1} - i\int_0^\infty e^{-i\omega t}\left|\sqrt{2}\sinh(\pi T t - i\epsilon)\right|^{-1} \\
&= \int_{-\infty}^{\infty} e^{-i\omega t}\left(\sqrt{2}i\sinh(\pi T t) + \epsilon\right)^{-1}.
\end{aligned}$$

This gives [74]

$$F_1^0(\omega) = \frac{T}{2}\int_{-\infty}^{\infty} e^{-i\omega t}\,(2i\sinh(\pi T t) + \epsilon)^{-1} = \frac{1}{2}\,\frac{1}{e^{\omega/T}+1}$$

which is exactly what we got before as $P_\nu(1) = 1$.

Combining the results for F_1 and F_2, we have

$$F(\omega) = \frac{1}{2}\frac{1}{e^{\omega/T}+1}\sum_n \left(P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\alpha_n) \pm P_{\frac{i\omega}{2\pi T}-\frac{1}{2}}(\cosh\beta_n)\right).$$

Chapter 5

Breakdown of the Semi-Classical Approximation at the Black Hole Horizon

On Thursday, when it starts to freeze
And hoar-frost twinkles on the trees,
How very readily one sees
That these are whose–but whose are these?
(The World of Pooh by, A. A. Milne)

Abstract:

The definition of matter states on spacelike hypersurfaces of a 1+1 dimensional black hole spacetime is considered. The effect of small quantum fluctuations of the mass of the black hole due to the quantum nature of the infalling matter is taken into account. It is then shown that the usual approximation of treating the gravitational field as a classical background on which matter is quantized, breaks down near the black hole horizon. Specifically, on any hypersurface that captures both infalling matter near the horizon and Hawking radiation, quantum fluctuations in the background geometry become important, and a semiclassical calculation is inconsistent. An estimate of the size of correlations between the matter and gravity states shows that they are so strong that a fluctuation in the black hole mass of order $e^{-M/M_{Planck}}$ produces a macroscopic change in the matter state.

This chapter is based on [45]

5.1 Introduction

Since the original papers of Hawking [5, 1] arguing that black holes should radiate thermally, and that this leads to an apparent loss of information, it has been hoped that investigations of this apparent paradox would lead to a better understanding of quantum gravity. Over the last few years, there has been renewed interest in this general problem. One reason is the construction of 1+1 dimensional models where evaporating black holes can be easily studied [56]. Another reason is the work by 't Hooft [57, 86, 87] suggesting that the black hole evaporation process may not be semiclassical. This idea is based in part on the fact that although Hawking radiation emerges at low frequencies of order M^{-1} at $\mathcal{I}^+$, it originates in very high frequency vacuum modes at $\mathcal{I}^-$ and even close to the black hole horizon, the latter frequencies being about e^M times the Planck frequency [88] (here M is the mass of the black hole in Planck units). 't Hooft also argues that if the black hole evaporation process is to be described by unitary evolution, then there should exist large commutators between operators describing infalling matter near the horizon and those describing outgoing Hawking radiation [87] despite the fact that they may be spacelike separated.

Recently[†] Susskind *et. al.* have argued that the information contained in infalling matter could be transferred to the Hawking radiation at the black hole horizon, thus avoiding information loss [90]. A common argument against this possibility is that from the perspective of an infalling observer, who probably sees nothing special at the horizon, there is no mechanism that could account for such a transfer of information. In response, Susskind suggests a breakdown of Lorentz symmetry at large boosts, and a principle of *complementarity* which says that one can make observations either far above the horizon or near the horizon, but somehow it should make no sense to talk of both [90, 58].

The 1+1 dimensional black hole problem including the effects of quantum gravity was recently studied in Ref. [91]. It was found that there are very large commutators between operators at the horizon, and operators at $\mathcal{I}^+$ measuring the Hawking radiation, agreeing with the earlier work of 't Hooft [87]. Ref. [91] assumes a reflection boundary condition at a strong coupling boundary. Some natural modifications of this boundary condition have been studied recently in [92]. There have been many other studies of quantum gravity on the black hole problem, some of which are listed in [93].

Let us recall the basic structure of the black hole problem [1]. Collapsing matter forms a black hole, which then evaporates by emission of Hawking radiation [5]. The radiation carries away the energy, leaving 'information' without energy trapped inside the black hole. The Hawking radiation arises from the production of particle pairs, one member of the pair falling into the horizon and the other member escaping to form the Hawking radiation outside the black hole. The quantum state of the quantum particles outside the black hole is thus not a pure state, and one may compute the entanglement entropy between the particles that fall into the black hole and the particles that escape to infinity. It is possible to carry out such a computation

[†]Historically, the greatest champions of this view point have been Page [89] and 't Hooft [57, 86].

explicitly in the simple 1+1 dimensional models referred to above. One finds [94, 95] that this entropy equals the quantity expected on the basis of purely thermodynamic arguments [96].

Such calculations are carried out in the semiclassical approximation, where one assumes that the spacetime is a given 1+1 dimensional manifold, and the matter is given by quantum fields propagating on this manifold. How accurate is this description? We wish to examine the viewpoint raised by 't Hooft and Susskind (referred to above) that quantum gravity is important in some sense at the horizon of the black hole. To this end we start with a theory of quantum gravity plus matter, and see how one obtains the semiclassical approximation where gravity is classical but matter is quantum mechanical. The extraction of a semiclassical spacetime from suitable solutions of the Wheeler–DeWitt equation has been studied in [39, 41]. Essentially, one wishes to obtain an approximation where the variables characterizing gravity are 'fast' (i.e. the action varies rapidly with change of these variables) and the matter variables are 'slow' (i.e. the action varies slowly when they change). This separation hinges on the fact that the gravity action is multiplied by an extra power of the Planck mass squared, compared to the matter variables, and this is a large factor whenever the matter densities are small in comparison to Planck density. We recall that the matter density is indeed low at the horizon of a large black hole (this is just the energy in the Hawking radiation). One might therefore expect the semiclassical approximation to be good at the horizon. It is interesting that this will turn out *not* to be the case, as we shall now show in a 1+1 dimensional model.

It was suggested in [97] that a semiclassical description (*i.e.* where gravity is classical but radiating matter quantum) can break down after sufficient particle production. This suggestion is based on the fact that particle creation creates decoherence [98], but on the other hand an excess of decoherence conflicts with the correlations between position and momentum variables needed for the classical variable [99]. In this paper we investigate this crude proposal and find that there is indeed a sense in which the semiclassical approximation breaks down near a black hole. It turns out that the presence of the horizon is crucial to this phenomenon, so what we observe here is really a property of black holes.

Since in black hole physics one is interested in concepts like entropy, information, and unitarity of states, it is appropriate to use a language where one deals with 'states' or 'wavefunctionals' on spacelike hypersurfaces, instead of considering functional integrals or correlation functions over a coordinate region of spacetime. In this description, the dynamical degrees of freedom are 1-geometries, and it is more fundamental to speak of the state of matter on a 1-geometry than on an entire spacetime. Thus, we will need to study the canonical formulation of 1+1 dimensional dilaton gravity. Recall that in this theory the gravity sector contains both the metric and an additional scalar field, the dilaton, which together define a 1-geometry. The space of all possible 1-geometries is called superspace. We assume that our theory of quantum gravity plus matter is described by some form of Wheeler–DeWitt equation [100], which enforces the Hamiltonian constraint on wavefunctionals in superspace. For dilaton gravity alone, a point of superspace is given by the fields $\{\rho(x), \phi(x)\}$. Here we have assumed the notation that the metric along the 1-dimensional geometry is

$ds^2 = e^{2\rho}dx^2$, and ϕ is the dilaton. One of the constraints on the wavefunctionals is the diffeomorphism invariance in the coordinate x. Using this invariance we may reduce the description of superspace so that different points just consist of intrinsically different 1-geometries. More precisely, choose any value of ϕ, say ϕ_0. Let s denote the proper distance along the 1-geometry measured from the point where $\phi = \phi_0$, with s positive in the direction where ϕ decreases. The function $\phi(s)$ along the 1-geometry describes the intrinsic structure of the 1-geometry, and is invariant under spatial diffeomorphisms (we have assumed here for simplicity that ϕ is a monotonic function along the 1-geometry, and that the value ϕ_0 appears at some point along the 1-geometry). Loosely speaking, we may regard superspace as the space of all such functions $\phi(s)$ (for a spacetime with boundary, this description must be supplemented with an embedding condition at the boundary).

Let us now consider the presence of a massless scalar field $f(x)$. Points of superspace now are described by $\{\phi(s), f(\phi(s))\}$, and wavefunctionals on this space, $\Psi[\phi(s), f(\phi(s))]$, satisfy the Wheeler–DeWitt equation

$$(H_{\text{gravity}} + H_{\text{matter}})\Psi[\phi(s), f(\phi)] = 0 . \tag{5.1.1}$$

We are now faced with the question: How do we obtain the semiclassical limit of quantum gravity, starting from some theory of quantum gravity plus matter? At the present point we have only 1-geometries in the description, and we have to examine how the 1+1 dimensional spacetime emerges in some approximation from $\Psi[\phi(s), f(\phi)]$. Obtaining a 1+1 dimensional spacetime has been called the 'problem of time' in quantum gravity, and considerable work has been done on the semiclassical approximation of gravity as a solution to this problem [41]. We wish to reopen this discussion in the context of black hole physics.

In mathematical terms, we have $\Psi[\phi(s), f(\phi(s))]$ giving the complete description of matter plus gravity. What is the state of matter on a time-slice? If we are given a classical 1+1 spacetime, then a time-slice is given by an intrinsic 1-geometry $\phi(s)$ (plus a boundary condition at infinity). Thus the matter wavefunctional on a time-slice $\phi(s)$ should be given by

$$\Psi_{\phi(s)}[f(\phi(s))] \equiv \Psi[\phi(s), f(\phi(s))] , \tag{5.1.2}$$

The semiclassical approximation then consists of approximating the full solution of the Wheeler–DeWitt equation by the product of a semiclassical functional of the gravitational variables alone, times a matter part which is taken to be a solution

$$\psi^{\mathcal{M}}_{\phi(s)}[f(\phi(s))] \tag{5.1.3}$$

of the functional Schrödinger equation on some mean spacetime $\mathcal{M}$ (here the function $\phi(s)$ is like a generalized time coordinate on $\mathcal{M}$). If any quantum field theory on curved spacetime calculation using (5.1.3) can be used to approximate the result obtained using the exact solution of the Wheeler-DeWitt equation of (5.1.2), then we say that the semiclassical approximation is good. On the other hand, if this approximation fails to work, we conclude that quantum fluctuations in geometry are important to whichever question it is that we wished to answer.

For the black hole problem, it is appropriate to make a separation between the matter regarded as forming the black hole, denoted by $F(\phi(s))$, and all other matter $f(\phi(s))$. It is then more natural to regard $F(\phi(s))$ as part of the gravitational degrees of freedom, and it is certainly regarded as a classical background field in the derivation of Hawking radiation using the semiclassical approximation. In this situation we must be more precise about what we require for the semiclassical approximation to be good. Assume that the black hole is formed by the collapse of some wavepacket of matter F, into a region smaller than the Schwarzschild radius. We note that the energy of this matter wavepacket cannot be exactly M, because an eigenstate of energy would not evolve at all over time in the manner needed to describe the collapsing packet. In fact, since the matter will be localized to within the Schwarzschild radius M, there will be a momentum uncertainty much greater than $1/M$ in Planck units, which leads to an energy uncertainty which must also be much larger than $1/M$. This uncertainty is still quite small, but should nevertheless not be ignored. The different possible energy values in this range $(M, M + \Delta M)$ where $\Delta M \gg 1/M$, will give different semiclassical spacetimes. For the semiclassical approximation to be good for any given computation, it must be independent of which of the slightly different spacetimes is chosen. Conversely, if the difference in any quantity of interest becomes significant when evaluated on different spacetimes in the above mass range, then we cannot use a mean 2-geometry to describe physics, and we should say that the semiclassical approximation is not good[‡]

Casting this problem in the language of the preceeding paragraphs, we must ask whether the wavefunctional of matter from the full quantum solution of the Wheeler–DeWitt equation is well approximated by working on a fixed spacetime $\mathcal{M}$ of mass M and ignoring the uncertainty ΔM in M. Now, suppose that the semiclassical approximation were a good one when describing the state of matter on a given time-slice $\phi(s)$. If we consider the different matter states that are obtained on $\phi(s)$ by taking different values for M, which cannot be clearly distinguished because we are averaging over the fluctuations in geometry, then these states should not be 'too different' if there is to be an unambiguous definition of the state on the time-slice. This is a minimal requirement for a semiclassical calculation to be a good approximation to $\Psi[\phi(s), F(\phi(s)), f(\phi(s))]$.

Let the state of quantized matter obtained by working on $\mathcal{M}$ be $\psi^{\mathcal{M}}_{\phi(s)}[f(\phi(s))]$, where in $\mathcal{M}$ the energy of the infalling matter is M. This is a state in the Schrödinger representation, and thus depends on the time-slice specified by the function $\phi(s)$ (plus boundary condition). At slices corresponding to early times (i.e. near $\mathcal{I}^-$, before the black hole formed) for all spacetimes with mass M in the range $(M, M + \Delta M)$, we fix the matter state to be approximately the same in each spacetime. In terms of a natural inner product relating states on a common 1-geometry in different spacetimes (which we define in this paper), this means that

$$\langle \psi^{\mathcal{M}}_{\phi(s)} | \psi^{\bar{\mathcal{M}}}_{\phi(s)} \rangle \approx 1 \tag{5.1.4}$$

[‡]The role of fluctuations in the mass of the infalling matter was also discussed in [91]. Generally, fluctuations in geometry can also arise from other sources, but we shall ignore these here.

on these early time slices, where $\bar{\mathcal{M}}$ is a spacetime with mass $\bar{M}$ in the above range. On each spacetime the matter state evolves in the Schrödinger picture in different ways, so that the inner product (5.1.4) will not be the same on all slices. For the semiclassical approximation to be good at any given slice, we need that (5.1.4) hold on that slice.

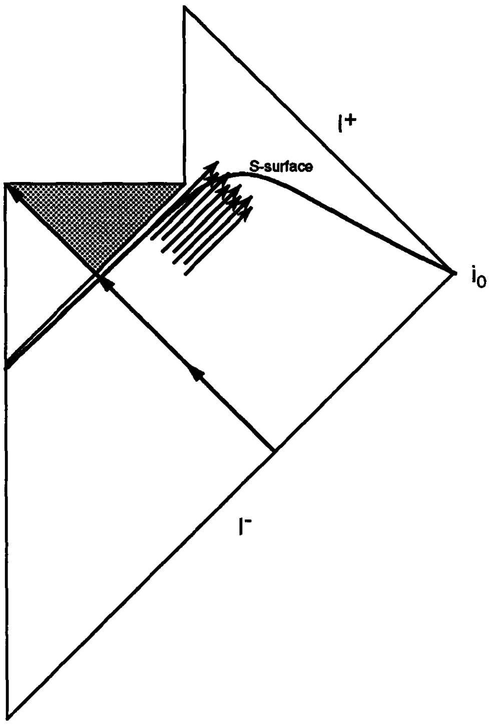

Figure 1: An example of an S-surface, shown in an evaporating black hole spacetime.

Having fixed the matter states on different spacetimes so that they are very similar at early times, we analyse later time slices to check that this property still holds. Any slice is taken to start at some fixed base point near spatial infinity. Consider now a slice that moves up in time near $\mathcal{I}^+$ to capture some fraction of the Hawking radiation. The slice then comes to the vicinity of the horizon, and then moves close to the horizon, so as to reach early advanced times before entering the strong coupling domain (see Fig. 1). The importance of such slices to the black hole paradox has been emphasized by Preskill [103] and Susskind *et. al.* [90] in their arguments relating to information bleaching and to the principle of black hole complementarity. Susskind *et. al.* conjectured that the large Lorentz boost between the two portions of the slice should lead to a problem in the semiclassical description of a black hole. Slices of this type have also been used in the literature as part of a complete spacelike slicing of spacetime, that stays outside the horizon of the black hole [101] and captures the Hawking radiation, and on which semiclassical physics should therefore apply. For these surfaces, which we shall refer to as S-surfaces, we shall show in this paper that it is no longer the case that matter states are approximately the same for different background spacetimes. Indeed, even for $|M - \bar{M}| \sim e^{-M}$ we find that on a 1-geometry $\phi(s)$ of this type,

$$\langle \psi^{\mathcal{M}}_{\phi(s)} | \psi^{\bar{\mathcal{M}}}_{\phi(s)} \rangle \approx 0.$$

As was argued above, the fluctuations in the mass of the hole must be at least of order $\Delta M > 1/M$, so we see that the state of matter on such slices is very ill defined because of the fluctuations in geometry. This shows that at least one natural quantity that we wish to consider in black hole physics, the state of matter on what we have termed an S-surface, is not given well by the semiclassical approximation.

The plan of this paper is the following. In section 2 we review the CGHS model, and give some relevant scales. In section 3 we study the embedding of 1-geometries in different 1+1 dimensional semiclassical spacetimes. In section 4 we compare states of matter on the same 1-geometry, but in different spacetimes. Section 5 is a general discussion of the meaning of these results and of possible connections to other work.

5.2 A review of the CGHS model

There follows a quick review of the CGHS model [56], with reference to the RST model [102] which includes back-reaction and defines some relevant scales in the CGHS solution. Although all calculations in this paper are for a CGHS black hole, the general features of the results that are derived are expected to apply equally well to other black hole models in two and four dimensions.

The Lagrangian for two dimensional string-inspired dilaton gravity is

$$S_G = \frac{1}{2\pi} \int dx dt \, \sqrt{-g} \, e^{-2\phi} \left[R + 4(\nabla\phi)^2 + 4\lambda^2 \right] \tag{5.2.5}$$

where $\phi(x)$ is the dilaton field and λ is a parameter analogous to the Planck scale. Writing

$$ds^2 = -e^{2\rho} dx^+ dx^-$$

where $x^\pm = t \pm x$ are referred to as Kruskal coordinates, (5.2.5) has static black hole solutions

$$e^{-2\rho} = e^{-2\phi} = \frac{M}{\lambda} - \lambda^2 x^+ x^- \tag{5.2.6}$$

and a linear dilaton vacuum (LDV) solution with $M = 0$. More interesting is the solution obtained when (5.2.5) is coupled to conformal matter,

$$S = S_G - \frac{1}{4\pi} \int dx dt \, \sqrt{-g} \, (\nabla f)^2,$$

where f is a massless scalar field. A left moving shock wave in f giving rise to a stress tensor

$$\frac{1}{2} \partial_+ f \partial_+ f = M \delta(x^+ - 1/\lambda)$$

yields a solution

$$e^{-2\rho} = e^{-2\phi} = -\frac{M}{\lambda}(\lambda x^+ - 1)\Theta(x^+ - 1/\lambda) - \lambda^2 x^+ x^- \tag{5.2.7}$$

representing the formation of a black hole of mass M/λ in Planck units (the Penrose diagram for this solution in shown in Fig. 2). For $\lambda x^+ < 1$ (region I), the solution is simply the LDV, whereas the solution for $\lambda x^+ > 1$ (region II),

$$e^{-2\rho} = e^{-2\phi} = \frac{M}{\lambda} - \lambda x^+ \left(\lambda x^- + \frac{M}{\lambda}\right)$$

is a black hole with an event horizon at $\lambda x^- = -M/\lambda$.

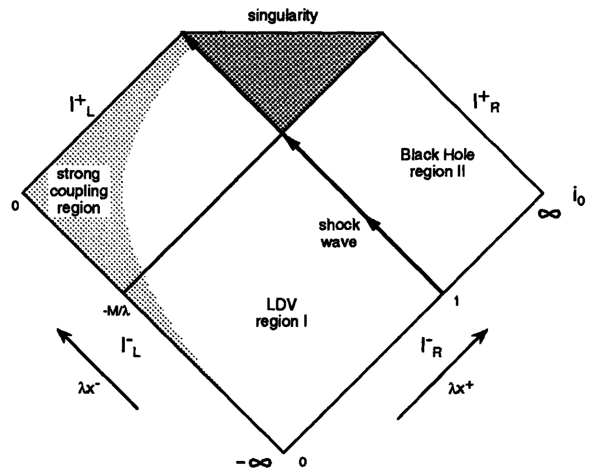

Figure 2: The Penrose diagram of the CGHS solution.

It is possible to define asymptotically flat coordinates in both regions I and II. In region I, we define

$$\lambda x^+ = e^{\lambda y^+}, \qquad \lambda x^- = -\frac{M}{\lambda} e^{-\lambda y^-} \tag{5.2.8}$$

and in region II we introduce the "tortoise" coordinates $\lambda\sigma^\pm$:

$$\lambda x^+ = e^{\lambda\sigma^+}, \qquad \lambda x^- + \frac{M}{\lambda} = -e^{-\lambda\sigma^-} \quad . \tag{5.2.9}$$

The coordinate y^- is used to define right moving modes at $\mathcal{I}_L^-$. To define left moving modes at $\mathcal{I}_R^-$ we can use either y^+ or σ^+. As (5.2.8) and (5.2.9) tell us, both coordinates can be extended to $I \cup II$ so that $y^+ = \sigma^+$. It is easy to see that as $\sigma \to \infty$ or as $y \to \infty$, $\rho \to -\infty$. Notice also that e^ϕ plays the role of the gravitational coupling constant in this theory. It is generally believed that semiclassical theory is reliable in regions where this quantity is small. At infinity $e^\phi \to 0$, and so this is a region of very weak coupling. Even at the horizon, $e^\phi = \sqrt{\lambda/M}$ is small provided that the mass of the black hole is large in Planck units ($M/\lambda \gg 1$). This is assumed to be the case in all calculations so that the weak coupling region extends well inside the black hole horizon.

One virtue of this two dimensional model is that it is straightforward to include the effects of backreaction by adding counterterms to the action S. This was first done by CGHS, but a more tractable model was introduced by RST who found an analytic solution for the metric of an evaporating black hole. However, the RST model still exhibits all the usual paradoxes associated with black hole evaporation (for a review see [103, 6]).

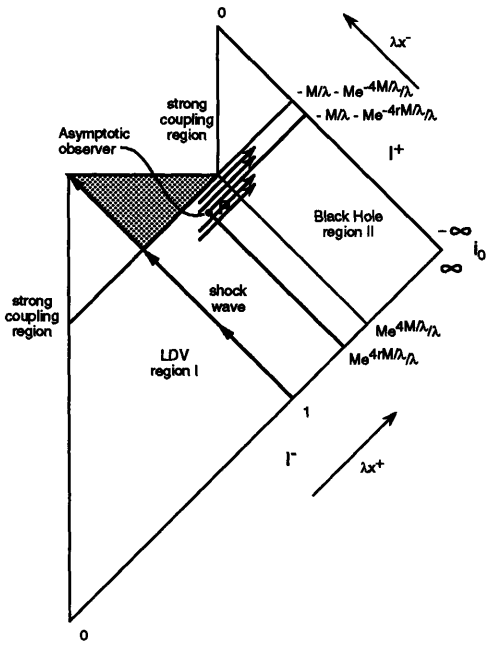

Figure 3: The Penrose diagram of the RST solution with some approximate scales shown.

Although we will carry out our calculations in the simpler CGHS model, the RST solution (whose Penrose diagram is shown in Fig. 3), is a useful guide for identifying certain scales in the evaporation process. These can be usefully carried over to a study of the CGHS solution, and serve to determine the portion of that solution that is unaffected by backreaction: The time scale of evaporation of the hole as measured by an asymptotic observer is $t_E \sim 4M$ in Planck units; the value of x^- at which a proportion r of the total Hawking radiation reaches $\mathcal{I}^+$ is $\lambda x_P^- = -M(1 + e^{-4rM/\lambda})/\lambda$ (by this we mean that the Hawking radiation to the right of this value carries energy rM); the value of x^+, for $x^- = x_P^-$, which corresponds to a point well outside the hole, in the sense that the curvature is weak and the components of the stress tensor are small is $\lambda x_P^+ = Me^{4rM/\lambda}/\lambda$ provided that $\lambda x^+ > e^{2M/\lambda}$. On the basis of these scales, we can define a point P at $(\lambda x_P^+, \lambda x_P^-)$ as defined above, located just outside the black hole, in the asymptotically flat region, and to the left of a proportion r of the Hawking radiation.

5.3 Embedding of 1-geometries

In this section, we shall compare how a certain spacelike hypersurface Σ may be embedded in collapsing black hole spacetimes (5.2.7) of masses M (denoted by $\mathcal{M}$) and $\bar{M} = M + \Delta M$ (denoted by $\bar{\mathcal{M}}$), where ΔM is a fluctuation of at most Planck size.

In 1+1 dimensional dilaton gravity models an invariant definition of a 1-geometry is provided by the value of the dilaton field $\phi(s)$ as a function of the proper distance s along the 1-geometry, measured from some fixed reference point. For spacetimes with boundary, such as the black hole geometries in the CGHS model, this reference point may be replaced by information about how the 1-geometry is embedded at infinity. It is natural to regard asymptotic infinity as a region where hypersurfaces can be nailed down by external observers who are not a part of the quantum system we are considering. We impose the condition that 1-geometries in different spacetimes should be indistinguishable for these asymptotic observers, ensuring that the semiclassical approximation holds for these observers. This condition and the function $\phi(s)$ are enough to define a unique map of Σ from $\mathcal{M}$ to $\bar{\mathcal{M}}$.

It is important to point out at this stage that it is possible that this map is not well defined for some $\bar{\mathcal{M}}$, in the sense that there may exist no spacelike hypersurface in $\bar{\mathcal{M}}$ with the required properties. For the surfaces we consider, this issue does not arise. Further, it can be argued that there is no important effect of this phenomenon on the state of the matter fields, at least as long as one is away from strong curvature regions. (To see this it is helpful to use the explicit quantum gravity wavefunction for dilaton gravity given in [59, 60]). For this reason we shall ignore all spacetimes $\bar{\mathcal{M}}$ where Σ does not fit.

Given an equation for $\Sigma_{\mathcal{M}}$,

$$\lambda x^- = f(\lambda x^+)$$

and expressions for $\rho(x^+, x^-)$ and $\phi(x^+, x^-)$ in $\mathcal{M}$ and $\bar{\rho}(\bar{x}^+, \bar{x}^-)$ and $\bar{\phi}(\bar{x}^+, \bar{x}^-)$ in $\bar{\mathcal{M}}$, we determine the corresponding equation for $\Sigma_{\bar{\mathcal{M}}}$,

$$\lambda \bar{x}^- = \bar{f}(\lambda \bar{x}^+)$$

by requiring that $\phi(s) = \bar{\phi}(\bar{s})$ and similarly $d\phi/ds(s) = d\bar{\phi}/d\bar{s}(\bar{s})$ (it is if these equations have no real solution for a given $\bar{\mathcal{M}}$ that we say that Σ does not fit in $\bar{\mathcal{M}}$). These conditions require one boundary condition which fixes $\Sigma_{\bar{\mathcal{M}}}$ at infinity, and this may be chosen in such a way that the equations for $\Sigma_{\mathcal{M}}$ and $\Sigma_{\bar{\mathcal{M}}}$ are the same in asymptotically flat (tortoise) coordinates sufficiently far from the black hole.

We shall demonstrate that while most surfaces embed in very slightly different ways in spacetimes $\mathcal{M}$ and $\bar{\mathcal{M}}$ with masses differing only at the Planck scale, there is a special class of surfaces for which this is not true (what we mean by embeddings being different will be discussed later). These are the S-surfaces which catch both the Hawking radiation (the Hawking pairs reaching $\mathcal{I}^+$, but not those ending up at the singularity) and the in-falling matter near the horizon (see Fig. 1). It is useful to give an example of such surfaces. A straight line in Kruskal coordinates $x^{\pm}$ going through a point $P \sim (Me^{4rM/\lambda}/\lambda, -M(1 + e^{-4rM/\lambda}))$, is a line of this type, catching a proportion r of the outgoing Hawking radiation, provided the slope of the line is extremely small – of order $e^{-8rM/\lambda}$. The smallness of this parameter will play an important role in our discussion. Although the line is straight in Kruskal coordinates, it will, of course, look bent in the Penrose diagram, ending up at i_0. Far from the horizon, these lines are lines of constant Schwarzschild time $\lambda t = 4rM$, giving an interpretation for minus one half the logarithm of the slope in terms of the time at infinity.

It is worth pointing out that the map from a surface $\Sigma_{\mathcal{M}}$ in $\mathcal{M}$ to the corresponding surface $\Sigma_{\bar{\mathcal{M}}}$ in $\bar{\mathcal{M}}$ defines a map from any point Q on $\Sigma_{\mathcal{M}}$ to a point $\bar{Q}_\Sigma$ on $\Sigma_{\bar{\mathcal{M}}}$ in $\bar{\mathcal{M}}$. Any other choice of surface $\Xi_{\mathcal{M}}$ in $\mathcal{M}$ passing through Q maps Q to a different point $\bar{Q}_\Xi$ in $\bar{\mathcal{M}}$. This uncertainty in the location of a point $\bar{Q}$ in $\bar{\mathcal{M}}$ gives a geometric way of defining the fluctuations in geometry around Q. Generally, we may expect all the images of Q in $\bar{\mathcal{M}}$ to lie within a small region of Planck size. However, we shall see below that this is not the case near a black hole horizon.

5.3.1 Basic Equations

Here we present the basic equations describing the embedding of Σ. In a collapsing black hole manifold $\mathcal{M}$ of mass M (5.2.7), it is convenient to define Σ as

$$\lambda x^- = f(\lambda x^+) - M/\lambda.$$

If we use Kruskal coordinates $\bar{x}^\pm$ to describe Σ in a black hole manifold $\bar{\mathcal{M}}$ of mass $\bar{M}$ as

$$\lambda \bar{x}^- = \bar{f}(\lambda \bar{x}^+) - \bar{M}/\lambda,$$

then in region II of (5.2.7)

$$\frac{M}{\lambda} - \lambda x^+ f(\lambda x^+) = \frac{\bar{M}}{\lambda} - \lambda \bar{x}^+ \bar{f}(\lambda \bar{x}^+) \tag{5.3.10}$$

$$\frac{f + \lambda x^+ f'}{\sqrt{-f'}} = \frac{\bar{f} + \lambda \bar{x}^+ \bar{f}'}{\sqrt{-\bar{f}'}} \tag{5.3.11}$$

where prime denotes a derivative with respect to the argument. The first equation is the requirement of equal $\phi(s)$ and the second of equal $d\phi/ds(s)$.

Once we identify the embedding of Σ in $\bar{\mathcal{M}}$, we can then identify points in both spacetimes by the value of s on Σ. This identification may be described by the function $\bar{x}^+(x^+)$ between coordinates on Σ in each of the spacetimes. To solve the equations (5.3.10) and (5.3.11), for $\bar{x}^+(x^+)$, differentiate (5.3.10) by x^+ and divide by (5.3.11), to get

$$\frac{d\bar{x}^+}{dx^+} = \frac{\sqrt{-f'}}{\sqrt{-\bar{f}'}}.$$

Another combination of these equations gives

$$\sqrt{-\bar{f}'} = \frac{-(f + \lambda x^+ f') \pm \sqrt{(f - \lambda x^+ f')^2 - 4\Delta M f'/\lambda}}{2\lambda \bar{x}^+ \sqrt{-f'}}$$

where $\Delta M = \bar{M} - M$. Combining both equations,

$$\ln(\lambda\bar{x}^+) = 2\int d(\lambda x^+)\frac{f'}{(f + \lambda x^+ f') \mp \sqrt{(f - \lambda x^+ f')^2 - 4\Delta M f'/\lambda}} \tag{5.3.12}$$

which is a general expression for $\bar{x}^+(x^+)$ for any Σ. Similarly, if we label the one geometry by $\lambda x^+ = g(z^-)$ where $z^- = \lambda x^- + M/\lambda$ (using the notation $g = f^{-1}$), we find an analogous expression for $\bar{x}^-(x^-)$:

$$\ln(\lambda\bar{x}^- + \bar{M}/\lambda) = 2\int dz^-\frac{g'}{(g + z^- g') \mp \sqrt{(g - z^- g')^2 - 4\Delta M g'/\lambda}}\ , \tag{5.3.13}$$

In (5.3.12) and (5.3.13), the sign of the square root is determined by requiring that as ΔM tends to zero we get $\bar{x}^\pm = x^\pm$. From these equations one can construct the corresponding one geometry in $\bar{\mathcal{M}}$. In order for the solution to make sense, the expressions inside the square root must be positive. This condition is a manifestation of the fitting problem mentioned above.

5.3.2 A large shift for straight lines

For simplicity, we focus our attention on lines that are straight in the Kruskal coordinates $x^\pm$. Below we present a quick analysis of the embedding of these 1-geometries in neighbouring spacetimes. In the next subsection a more detailed treatment will be given.

Consider the line Σ defined in $\mathcal{M}$ by the equation

$$\lambda x^- = f(\lambda x^+) - \frac{M}{\lambda} = -\alpha^2\lambda x^+ + b.$$

It is easy to see that as a consequence of (5.3.10) and (5.3.11), the function $\bar{f}(\lambda\bar{x}^+)$ describing the deformed line in Kruskal coordinates on $\bar{\mathcal{M}}$ must also be linear. This is a helpful simplification. Let us write the equation for Σ in $\bar{\mathcal{M}}$ as

$$\lambda\bar{x}^- = \bar{f}(\lambda\bar{x}^+) - \frac{\bar{M}}{\lambda} = -\bar{\alpha}^2\lambda\bar{x}^+ + \bar{b}$$

The parameters b and $\bar{b}$ are related by

$$\frac{(\bar{b} + \bar{M}/\lambda)^2}{\bar{\alpha}^2} = \frac{4\Delta M}{\lambda} + \frac{(b + M/\lambda)^2}{\alpha^2} \tag{5.3.14}$$

It is useful to define another quantity δ, so that Σ crosses the shock wave, ($\lambda x^+ = 1$) in $\mathcal{M}$ at $\lambda x^- = -M/\lambda - \delta$ (*i.e.* $\delta = \alpha^2 - b - M/\lambda$). We then find from equation (5.3.12) that

$$2\bar{\alpha}\lambda\bar{x}^+ = 2\alpha\lambda x^+ + \delta/\alpha - \alpha \pm \sqrt{(\alpha^2 - \delta)^2/\alpha^2 + 4\Delta M/\lambda}$$

We still have a free parameter $\bar{\alpha}$. The way to fix it is by imposing the condition that Σ should be the same for an asymptotic observer at infinity, meaning that as expressed in tortoise coordinates σ or $\bar{\sigma}$, Σ should have the same functional form up to unobservable (Planck scale) perturbations. This may be achieved, as we will see later, simply by picking a point on Σ in $\mathcal{M}$, call it x_0^+, and demanding that both lines have the same value of ϕ at the point $x^+ = \bar{x}^+ = x_0^+$. Then

$$\bar{\alpha} = \alpha + \frac{\delta/\alpha - \alpha \pm \sqrt{(\alpha^2 - \delta)^2/\alpha^2 + 4\Delta M/\lambda}}{2x_0^+} \tag{5.3.15}$$

Taking $x_0 \to \infty$ fixes the line at infinity. The result does not depend on whether we take $x_0 \to \infty$ or just take it to be in the asymptotic region $x_0 > Me^{2M/\lambda}/\lambda$.

We can actually derive some quite general conclusions about how the embedding of Σ changes from $\mathcal{M}$ to $\bar{\mathcal{M}}$ from (5.3.14) and (5.3.15). Let us split the possible Σ's into three simple cases, for any value of α and δ (recall that $|\Delta M/\lambda| < 1$):

1. $(\alpha^2 - \delta)^2/\alpha^2 \gg 4\Delta M/\lambda$

   In this case

   $$\bar{\alpha} = \alpha$$

   and

   $$\lambda\bar{x}^+ = \lambda x^+ + \frac{\Delta M}{\lambda(\alpha^2 - \delta)} \quad .$$

2. $(\alpha^2 - \delta)^2/\alpha^2 \ll 4\Delta M/\lambda$

   For $\Delta M/\lambda \geq 0$ (this is taken to avoid fitting problems)

   $$\bar{\alpha} = \alpha$$

   and

   $$\lambda\bar{x}^+ = \lambda x^+ \pm \frac{\sqrt{\Delta M/\lambda}}{\alpha} \quad .$$

3. $(\alpha^2 - \delta)^2/\alpha^2 \sim 4\Delta M/\lambda$

   Again $\Delta M/\lambda \geq 0$, and we find a similar result

   $$\alpha = \bar{\alpha}$$

   and

   $$\lambda\bar{x}^+ \sim \lambda x^+ \pm \frac{\sqrt{\Delta M/\lambda}}{\alpha} \quad .$$

In the last two cases the sign $\pm$ depends on the sign of $\alpha^2 - \delta$.

The above results all show that the slope $\bar{\alpha}$ of the line in $\bar{\mathcal{M}}$ is virtually identical to the slope α in $\mathcal{M}$ (identical in the limit $x_0 \to \infty$). It is also the case that the position of the line in the x^- direction is almost the same in $\mathcal{M}$ and $\bar{\mathcal{M}}$. However, for lines with small values of α and δ, there is a large shift in the location of the line in the x^+ direction in $\bar{\mathcal{M}}$ relative to its position in $\mathcal{M}$. The lines for which this effect occurs are precisely the S-surfaces that we have discussed above. These were defined to have $\alpha^2 \sim Me^{-8rM/\lambda}/\lambda$, and $0 \leq \delta \leq Me^{-4rM/\lambda}/\lambda$, which are both small enough to compensate for the ΔM in the numerator in the expressions above. The large shift, and the fact that it occurs only for a very specific class of lines, precisely the S-surfaces which capture both a reasonable proportion of the Hawking radiation and the infalling matter (see Fig. 1), is the fundamental result behind the arguments presented in this paper. The fact that only a special class of lines exhibit this effect is reassuring, as it means that any effects that are a consequence of this shift can only be present close to the black hole horizon.

5.3.3 Complete hypersurfaces

So far we have not taken the hypersurfaces to be complete, *i.e.*, we have not done the full calculation of continuing them to the LDV and finishing at infinity in the strong coupling regime. We will now perform the full calculation for a certain class of hypersurfaces. They will provide us a convenient example (for calculational purposes) for use in section 4, where we will discuss the implications of the large shift on the time evolution of matter states.

We choose, for convenience, to work with a class of hypersurfaces that all have $d\phi/ds = -\lambda$:

$$\lambda x^- = \begin{cases} -\alpha^2\lambda x^+ - 2\alpha\sqrt{\dfrac{M}{\lambda}} - \dfrac{M}{\lambda} & (\lambda x^+ \geq 1) \\ -\left(\alpha + \sqrt{\dfrac{M}{\lambda}}\right)^2 \lambda x^+ & (\lambda x^+ \leq 1) \end{cases} . \tag{5.3.16}$$

These lines are of type 1 ($(\alpha^2-\delta)^2/\alpha^2 \gg 4\Delta M/\lambda$) discussed in section 3.2. They have one free parameter, the slope α^2. At spacelike infinity, these lines are approximately constant Schwarzschild time lines, $\sigma^0 = -\ln\alpha$, and for different values of α, they provide a foliation of spacetime in a way often discussed in the literature [101] in the context of the black hole puzzle. They always stay outside the event horizon, and they cross the shock wave at a Kruskal distance $\delta = 2\alpha\sqrt{M/\lambda} + \alpha^2$ from the horizon. After crossing the shock wave they continue to the strong coupling region. For an early time Cauchy surface, the parameter α^2 is arbitrarily large ($\alpha^2 \to \infty$ would make the lines approach $\mathcal{I}^-$). As α^2 becomes smaller, the lines move closer to the event horizon. Finally, as $\alpha^2 \to 0$, the upper segment asymptotes to $\mathcal{I}^+$ and to the segment of the event horizon above the shock wave. This is illustrated in Fig. 4. We are mostly interested in the S-surfaces that catch a ratio r of the Hawking radiation emitted by the black hole, which fixes the value of α. For r not too close to 1, the S-surfaces are well within the weak coupling region.

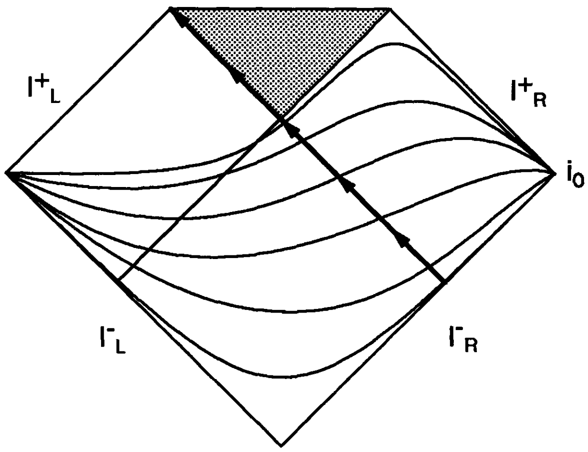

Figure 4: Examples of the complete slices of Sec. 3.3.

We want to find the location of the above lines in a black hole background with a mass $\bar{M} = M + \Delta M$. It is easy to see that in the new background, the lines

$$\lambda \bar{x}^- = \begin{cases} -\bar{\alpha}^2 \lambda \bar{x}^+ - 2\bar{\alpha}\sqrt{\dfrac{\bar{M}}{\lambda}} - \dfrac{\bar{M}}{\lambda} & (\lambda \bar{x}^+ \geq 1) \\ -\left(\bar{\alpha} + \sqrt{\dfrac{\bar{M}}{\lambda}}\right)^2 \lambda \bar{x}^+ & (\lambda \bar{x}^+ \leq 1) \end{cases} . \tag{5.3.17}$$

also satisfy $d\phi/ds \equiv -\lambda$. We only need to identify the new slope $\bar{\alpha}^2$ in terms of the old one, and as before, this is given by the boundary conditions at infinity. Requiring $\lambda \bar{\sigma}_0^+ = \lambda \sigma_0^+$, where σ^+ is the tortoise coordinate defined in (5.2.9), yields

$$\bar{\alpha} = \alpha + \left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\bar{M}}{\lambda}}\right) e^{-\lambda \sigma_0^+} .$$

If we also want to require $\lambda \bar{\sigma}_0^- = \lambda \sigma_0^-$, we need to do the fixing at infinity, which of course sets

$$\bar{\alpha} = \alpha$$

After fixing the $\sigma^\pm$ coordinates at infinity, we may check that $\bar{\sigma}^\pm$ and $\sigma^\pm$ do not differ appreciably as we approach the point P (still considered to be in the asymptotic region) along an S-surface. Taking $\alpha \sim e^{-4rM/\lambda}$ and P to be at $\lambda x_P^+ \sim M e^{4rM/\lambda}/\lambda$, $\lambda x_P^- \sim -M/\lambda - M e^{-4rM/\lambda}/\lambda$ as before, we find that at P

$$\begin{aligned} \lambda \bar{\sigma}_P^+ - \lambda \sigma_P^+ &\approx -\frac{\Delta M}{2M}\left(\frac{\lambda}{M}\right)^{1/2} \\ \lambda \bar{\sigma}_P^- - \lambda \sigma_P^- &\approx -\frac{\Delta M}{2M} , \end{aligned} \tag{5.3.18}$$

which is a small deviation. We conclude that if we had fixed the surface at P instead of infinity, all results would be qualitatively unchanged, as one would expect.

We can now compute the relationship between λx^+ and $\lambda \bar{x}^+$. As we saw in the previous subsection, the points in the original line get "shifted" by a large amount in the new line. It is easy to see that

$$\lambda \bar{x}^+ = \lambda x^+ + \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\bar{M}}{\lambda}}\right) \tag{5.3.19}$$

$$\approx \lambda x^+ - \frac{\Delta M}{2\lambda\alpha}\sqrt{\frac{\lambda}{M}} \,. \tag{5.3.20}$$

For instance, for $\alpha \sim e^{-4rM/\lambda}$, $\Delta M/\lambda \sim \lambda/M$ the shift is of the order of

$$\lambda \bar{x}^+ - \lambda x^+ \sim -\frac{1}{2}\left(\frac{\lambda}{M}\right)^{3/2} e^{4rM/\lambda} \,,$$

which is huge. Even for $\Delta M/\lambda \sim e^{-M/\lambda}$, the shift can be extremely large. As we will see in section 4, instead of the relations $\lambda \bar{x}^+ = \lambda \bar{x}^+(\lambda x^+)$, we will be interested in the induced relations between the asymptotically flat coordinates $\lambda\bar{\sigma}^+$ and $\lambda\sigma^+$, and $\lambda\bar{y}^-$ and λy^-. A huge shift in the Kruskal coordinate close to the shock wave will correspond to a big shift in the coordinate $\lambda\sigma^+$, in which the metric is flat at $\mathcal{I}_R^-$. As a consequence, the relation between $\lambda\bar{\sigma}^+$ and $\lambda\sigma^+$ is nonlinear, as we will discuss in the next section.

Finally, let us mention an immediate consequence of this large shift in the x^+ direction. The map of an S-surface from $\mathcal{M}$ to $\bar{\mathcal{M}}$ induces a map from a point Q close to the horizon to a point $\bar{Q}$ which is shifted a long way up the horizon in terms of Kruskal coordinates. A similar map induced by other surfaces through Q which are not S-surfaces will not shift $\bar{Q}$ by a large amount. We therefore see the presence of large quantum fluctuations near the horizon in the position of $\bar{Q}$ in the sense defined above. These large fluctuations are already a somewhat unexpected result.

5.4 The state of matter on Σ

We have seen in the previous section in some detail the large shift that occurs in the x^+ direction when we map a S-surface Σ from a black hole spacetime $\mathcal{M}$ to one with a mass which differs from $\mathcal{M}$ by an extremely small amount, even compared with the Planck scale. This appears to be a large effect, capable of seriously impairing the definition of a unique quantum matter state on Σ in a semiclassical way. There are, however, many large scales in the black hole problem, and it is premature to draw conclusions from the appearance of this large shift in the Kruskal coordinates, without verifying that there is a corresponding shift in physical (coordinate invariant) quantities. An absolute measure of the shift is given by the asymptotic tortoise coordinate σ^+ at $\mathcal{I}_R^-$. The exponential relationship between x^+ and σ^+ implies that the shift is of Planck size for an x^+ far from the shock wave ($x^+/x_P^+ \sim 1$, where x_P

is again as defined at the end of section 2), and there is no reason to expect this to give rise to a large effect. However, for $x^+/x_P^+ \ll 1$ (close to the horizon), the shift in $\lambda\sigma^+$ is of order M/λ, an extremely large number. This implies that the shift is macroscopic in the sense that, for example, matter falling into the black hole some fixed time after the shock wave will end up at very different points on Σ, depending on whether we work in $\mathcal{M}$ or $\bar{\mathcal{M}}$. Similarly, identical quantum states on $\mathcal{I}_R^-$ should appear very different on Σ in the two cases, meaning that the matter state on Σ is strongly correlated with the fluctuations in geometry.

In this section, we will attempt to make the notion of different quantum states of matter on Σ more precise, allowing us to estimate the scale of entanglement between the matter and spacetime degrees of freedom. In order to do this, it is necessary to have a criterion to quantify the difference between two semiclassical matter states living in different spacetimes $\mathcal{M}$ and $\bar{\mathcal{M}}$, that are identical on $\mathcal{I}^-$ and are then evolved to Σ. The heuristic arguments above show that the expectation values of local operators can be very different for states in $\mathcal{M}$ and $\bar{\mathcal{M}}$ that appear identical on $\mathcal{I}^-$ where there is a fixed coordinate system through which to compare them. Rather than look at expectation values of operators, we construct an inner product

$$\langle \psi_1, \Sigma, \mathcal{M} | \psi_2, \Sigma, \bar{\mathcal{M}} \rangle$$

between Schrödinger picture matter states on the same Σ through which states on $\mathcal{M}$ and $\bar{\mathcal{M}}$ can be compared. The inner product makes use of a decomposition in modes defined using the diffeomorphism invariant proper distance along Σ, through which the states can be compared. Details of this construction can be found in Appendix A.

An important feature of the inner product is that for a Planck scale fluctuation ΔM and for states $|\psi, \mathcal{M}\rangle$ and $|\psi, \bar{\mathcal{M}}\rangle$ that are identical on $\mathcal{I}^-$ it can be checked that

$$\langle \psi, \Sigma, \mathcal{M} | \psi, \Sigma, \bar{\mathcal{M}} \rangle \approx 1 \tag{5.4.21}$$

on any generic surface Σ that does not have a large shift. This is a necessary condition for the consistency of quantum field theory on a mean curved background with a mass in the range $(M, M + \Delta M)$: If states on $\mathcal{M}$ and $\bar{\mathcal{M}}$ are orthogonal on Σ, this is an indication that the approximate Hilbert space structure of the semiclassical approximation is becoming blurred due to an entanglement between the matter and gravity degrees of freedom. Using the inner product, we now show that matter states become approximately orthogonal on S-surfaces for extremely small fluctuations $\Delta M/\lambda \sim e^{-4rM/\lambda}$ in the mass of a black hole, dramatically violating condition (5.4.21).

In general the states that we wish to compare are most easily expressed as Heisenberg picture states on $\mathcal{M}$ and $\bar{\mathcal{M}}$, and the prospect of converting these to Schrödinger picture states, and evolving them to Σ is rather daunting. As explained in Appendix A, there is a short cut to this procedure. For the states we are interested in (those that start as vacua on $\mathcal{I}^-$) the basic information needed for the calculation of the inner product is the relation induced by $\phi(s)$ on Σ between the tortoise coordinates on $\mathcal{M}$ and $\bar{\mathcal{M}}$, namely $\sigma^+ = \sigma^+(\bar{\sigma}^+)$. This allows us to compute the inner product between the Schrödinger picture states by computing the usual Fock space inner product between two different Heisenberg picture states, defined with respect to the modes $e^{-i\omega\sigma^+}$ and $e^{-i\omega\bar{\sigma}^+}$. The latter inner product is given in terms of Bogoliubov coefficients. It should be stressed that this is just a short cut, and that the inner product depends crucially on the surface Σ, which is seen in the form of the function $\sigma^+ = \sigma^+(\bar{\sigma}^+)$.

We will study the overlap

$$\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle \tag{5.4.22}$$

where $|0 \text{ in}, \Sigma, \mathcal{M}\rangle$ is the matter Schrödinger picture state in spacetime $\mathcal{M}$ on the hypersurface Σ which was in the natural left moving sector vacuum state on $\mathcal{I}_R^-$. We shall also use this quantity to estimate the size of $\Delta M = (\bar{M} - M)$ at which the states begin to differ appreciably. To evaluate the inner product (5.4.22), we first need to find the induced Bogoliubov transformation

$$v_\omega = \int_0^\infty d\omega' \; [\; \alpha_{\omega\omega'} \bar{v}_{\omega'} + \beta_{\omega\omega'} \bar{v}^*_{\omega'} \;] \tag{5.4.23}$$

between the in-modes

$$\begin{aligned} \bar{v}_\omega &= \frac{1}{\sqrt{2\omega}}\, e^{-i\omega\bar{\sigma}^+} \\ v_\omega &= \frac{1}{\sqrt{2\omega}}\, e^{-i\omega\sigma^+} \;, \end{aligned} \tag{5.4.24}$$

where σ^+ and $\bar{\sigma}^+$ are related by an induced relation

$$\sigma^+ = \sigma^+(\bar{\sigma}^+) \;. \tag{5.4.25}$$

Let us derive the relation (5.4.25) above, for the example of Section 3.3. As (5.3.20) shows us, the shift $\lambda\bar{x}^+ - \lambda x^+$ can become large and λx^+ above the shock wave maps to $\lambda\bar{x}^+$ further above[§] the shock wave. As λx^+ comes closer to the shock wave and crosses to the other side, the image point $\lambda\bar{x}^+$ can still be located above the shock wave. Only when λx^+ is low enough under the shock wave, does $\lambda\bar{x}^+$ also cross the shock and go below it. Thus, the relation between the coordinates is split into three regions:

$$e^{-\phi} = \begin{cases} \alpha\lambda x^+ + \sqrt{\dfrac{M}{\lambda}} = \alpha\lambda\bar{x}^+ + \sqrt{\dfrac{\bar{M}}{\lambda}} & (\lambda x^+ \geq 1) \\ \left(\alpha + \sqrt{\dfrac{M}{\lambda}}\right)\lambda x^+ = \alpha\lambda\bar{x}^+ + \sqrt{\dfrac{\bar{M}}{\lambda}} & \left(1 \geq \lambda x^+ \geq \dfrac{\alpha + \sqrt{\frac{\bar{M}}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}}\right) \\ \left(\alpha + \sqrt{\dfrac{M}{\lambda}}\right)\lambda x^+ = \left(\alpha + \sqrt{\dfrac{\bar{M}}{\lambda}}\right)\lambda\bar{x}^+ & \left(\dfrac{\alpha + \sqrt{\frac{\bar{M}}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}} \geq \lambda x^+ \geq 0\right) \end{cases} . \tag{5.4.26}$$

[§]For simplicity, we will consider only the case $\Delta M < 0$ in this section. The conclusions will not depend on this assumption.

Rewriting (5.4.26) using the asymptotic coordinates, we then get the relation (5.4.25):

$$\lambda\sigma^+ = \begin{cases} \ln\left[e^{\lambda\bar{\sigma}^+} - \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\overline{M}}{\lambda}}\right)\right] & \left(\lambda\bar{\sigma}^+ \geq \lambda\bar{\sigma}_1^+\right) \\ \ln\left[\left(\frac{\alpha}{\alpha+\sqrt{\frac{M}{\lambda}}}\right)\left(e^{\lambda\bar{\sigma}^+} + \frac{1}{\alpha}\sqrt{\frac{\overline{M}}{\lambda}}\right)\right] & \left(\lambda\bar{\sigma}_1^+ \geq \lambda\sigma^+ \geq 0\right) \\ \lambda\bar{\sigma}^+ + \ln\left[\frac{\alpha+\sqrt{\frac{\overline{M}}{\lambda}}}{\alpha+\sqrt{\frac{M}{\lambda}}}\right] & \left(0 \geq \lambda\bar{\sigma}^+\right) \end{cases} , \tag{5.4.27}$$

where

$$\lambda\bar{\sigma}_1^+ \equiv \ln\left[1 + \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\overline{M}}{\lambda}}\right)\right] . \tag{5.4.28}$$

This coordinate transformation is illustrated in Fig. 5. As can be seen, in the first region (which corresponds to both points being above the shock) the transformation is logarithmic. On the other hand, in the third region when both points are below the shock, the transformation is exactly linear. The form of the transformation for the interpolating region when the other point is above and the other point below the shock should not be taken very seriously, since it depends on the assumptions made on the distribution of the infalling matter. For a shock wave it looks like a sharp jump, but if we smear the distribution to have a width of *e.g.* a Planck length, the jump gets smoothened and the transformation becomes closer to a linear one.

The Bogoliubov coefficients are now found to be

$$\begin{aligned} \alpha_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\, I^+_{\omega\omega'} \\ \beta_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\, I^-_{\omega\omega'} , \end{aligned} \tag{5.4.29}$$

where $I^\pm_{\omega\omega'}$ are the integrals

$$I^\pm_{\omega\omega'} \equiv \int_{-\infty}^{\infty} d\bar{\sigma}^+ \; e^{-i\omega\sigma^+(\bar{\sigma}^+)\pm i\omega'\bar{\sigma}^+} . \tag{5.4.30}$$

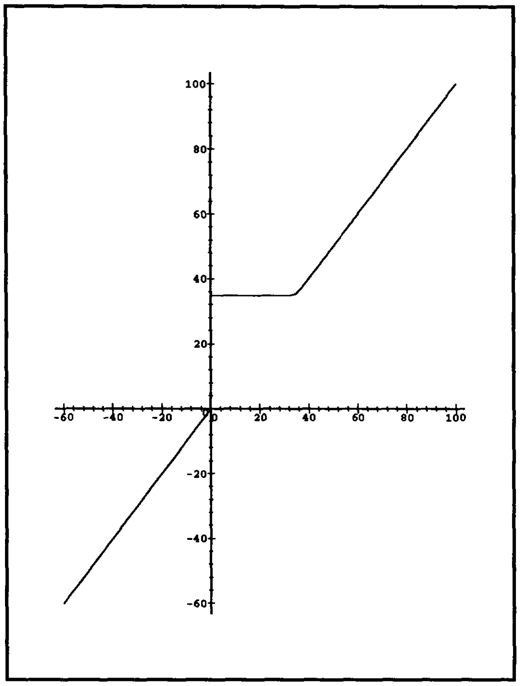

Figure 5: A graph of the function $\lambda\bar{\sigma}^+(\lambda\sigma^+)$. The vertical axis $=\lambda\bar{\sigma}^+$, the horizontal axis $=\lambda\sigma^+$. The part of the graph to the right (left) of the vertical axis corresponds to both points being above (below) the shock wave. (The interpolating part is in the region $\lambda\sigma^+ \in (-0.0013, 0)$ so the plot coincides with a segment of the vertical axis in the figure). The values M/λ =20, $\Delta M/\lambda$ =-1/20 and r =1/2 were used in the plot.

Substituting the relations (5.4.27) above, we get

$$\begin{aligned} I^{\pm}_{\omega\omega'} &= \int_{\bar{\sigma}_1^+}^{\infty} d\bar{\sigma}^+ \exp\left\{-i\frac{\omega}{\lambda}\ln\left[e^{\lambda\bar{\sigma}^+} - \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\overline{M}}{\lambda}}\right)\right] \pm i\omega'\bar{\sigma}^+\right\} \quad (5.4.31) \\ &+ \int_0^{\bar{\sigma}_1^+} d\bar{\sigma}^+ \exp\left\{-i\frac{\omega}{\lambda}\ln\left[\frac{\alpha}{\alpha+\sqrt{\frac{M}{\lambda}}}\left(e^{\lambda\bar{\sigma}^+} + \frac{1}{\alpha}\sqrt{\frac{\overline{M}}{\lambda}}\right)\right] \pm i\omega'\bar{\sigma}^+\right\} \\ &+ \int_{-\infty}^{0} d\bar{\sigma}^+ \exp\left\{-i\omega\bar{\sigma}^+ - i\frac{\omega}{\lambda}\ln\left[\frac{\alpha+\sqrt{\frac{\overline{M}}{\lambda}}}{\alpha+\sqrt{\frac{M}{\lambda}}}\right] \pm i\omega'\bar{\sigma}^+\right\} \\ &\equiv I^{\pm}_{\omega\omega'}(1) + I^{\pm}_{\omega\omega'}(2) + I^{\pm}_{\omega\omega'}(3) \quad . \end{aligned}$$

To identify the first integral, introduce first

$$\lambda\Delta \equiv 1 + \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\overline{M}}{\lambda}}\right) = e^{\lambda\bar{\sigma}_1^+} \ . \quad (5.4.32)$$

and change the integration variable from $\lambda\bar{\sigma}^+$ to λy,

$$\lambda\Delta e^{-\lambda y} \equiv e^{\lambda\bar{\sigma}^+} \ . \tag{5.4.33}$$

We now find that

$$I^{\pm}_{\omega\omega'}(1) = \lambda\Delta^{\pm i\omega'/\lambda} \int_{-\infty}^{0} dy \ \exp\left\{-i\frac{\omega}{\lambda}\ln\left[\lambda\Delta(e^{-\lambda y} - C)\right] \mp i\omega' y\right\} \ , \tag{5.4.34}$$

where

$$C \equiv 1 - \frac{1}{\lambda\Delta} \ . \tag{5.4.35}$$

Now we can recognize the integral to be the same as discussed in [94]. This integral can be identified as an incomplete β-function. However, it is also possible to make the standard approximation of replacing the integrand by its approximate value in the interval $\lambda y \in (-1, 0)$ [5, 94, 104]. Note that this interval corresponds to a region $\lambda\sigma^+ \in (0, \ln[(e-1)\lambda\Delta + 1])$. The latter can be large: for $\lambda\Delta \sim e^{M/\lambda}$ it has size $\sim M/\lambda$. Indeed, comparing with (5.3.20) we notice that $\lambda\Delta - 1$ is equal to the shift $\lambda\bar{x}^+ - \lambda x^+$ above the shock, which could become exponential. So we can use

$$I^{\pm}_{\omega\omega'}(1) \approx \lambda\Delta^{\pm i\omega'/\lambda} \int_{-\infty}^{0} dy \ \exp\left\{-i\frac{\omega}{\lambda}\ln\left[-\lambda\Delta\lambda y\right] \mp i\omega' y\right\} \ . \tag{5.4.36}$$

As discussed in [104], this leads to the approximate relation

$$I^{+}_{\omega\omega'}(1) \approx -e^{\pi\omega/\lambda} \ \left(I^{-}_{\omega\omega'}(1)\right)^* \tag{5.4.37}$$

for the integrals.

The logarithm in (5.4.36) implies that $I^{\pm}_{\omega\omega'}(1)$ contributes significantly in the regime $\omega' - \omega \gg 1$. Since the coordinate transformation (5.4.27) was exactly linear in the third region and we argued that smearing of the incoming matter distribution smoothens the "interpolating part" in the second region, we can argue that $I^{\pm}_{\omega\omega'}(2)$ and $I^{\pm}_{\omega\omega'}(3)$ are negligible in the regime $\omega' - \omega \gg 1$. Therefore, in this limit $I^{\pm}_{\omega\omega'}(1)$ is the significant contribution, and a consequence of (5.4.37) is that the relationship of the Bogoliubov coefficients is (approximately) thermal,

$$\alpha_{\omega\omega'} \approx -e^{\pi\omega/\lambda} \ \beta^*_{\omega\omega'} \ , \tag{5.4.38}$$

with "temperature" $\lambda/2\pi$. Let us emphasize that the "temperature" itself is independent of the magnitude of the fluctuation ΔM. Rather, it is the *validity* of the thermal approximation that is affected: the larger the fluctuation ΔM is, the better approximation (5.4.38) is. Also, the region of $\lambda\sigma^+$ which corresponds to (5.4.38) becomes larger. Consequently, the inner product between $|0 \text{ in}, \Sigma, \bar{\mathcal{M}}\rangle$ and $|0 \text{ in}, \Sigma, \mathcal{M}\rangle$ can become appreciably smaller than 1. We refer to this as the states being "approximately orthogonal", we will elaborate this below.

Let us now calculate the inner product (5.4.22). As was explained before, we have related this inner product to an inner product between two Heisenberg picture states related by the above derived Bogoliubov transformation. For the latter inner product, we can use the general formula given in [107]. We then find (see Appendices):

$$|\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle|^2 = \left(\det(1+\beta\beta^\dagger)\right)^{-\frac{1}{2}} \quad . \tag{5.4.39}$$

We can now make an estimate of the scale of the fluctuations for the onset of the approximate orthogonality. As a rough criterion, let us say that as

$$|\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle|^2 \; < \; \frac{1}{\gamma} \ , \tag{5.4.40}$$

where $\gamma \sim e$, the states become approximately orthogonal. As is shown in Appendix B, the states become approximately orthogonal if

$$|\frac{\Delta M}{\lambda}| = \left|\frac{M}{\lambda} - \frac{\bar{M}}{\lambda}\right| \; > \; \gamma^{48} \sqrt{\frac{M}{\lambda}} \; \alpha \ , \tag{5.4.41}$$

where α is the (square root of the) slope. If the lines do not catch the Hawking radiation, $\alpha > 1$, then the fluctuations are not large enough to give arise to the approximate orthogonality and therefore (5.4.21) is satisfied. On the other hand, if the lines catch the fraction r of the Hawking radiation, $\alpha \approx e^{-4rM/\lambda}$ and the fluctuations can easily exceed the limit. (Recall that $M/\lambda \gg 1$, so α is the significant factor.) Note that the criterion (5.4.41) has been derived for the example hypersurfaces of section 3.3. However, the more general result for any S-surface of the types 1-3 of section 3.2 can be derived equally easily. In general the right hand side of (5.4.41) will depend on both the slope α^2 and the intercept¶ δ. The physics of the result will remain the same as above: for the S-surfaces the approximate orthogonality begins as the fluctuations $\Delta M/\lambda$ satisfy $\ln(\lambda/\Delta M) \sim M/\lambda$.

One might ask what happens to the "in"-vacua at $\mathcal{I}_L^-$ related to the rightmoving modes $e^{-i\omega y^-}, e^{-i\omega \bar{y}^-}$. We can similarly derive the induced coordinate transformations between the coordinates $\lambda\bar{y}^-$ and λy^-. This coordinate relation is virtually linear, and therefore the β Bogoliubov coefficients will be ≈ 0 and the vacua will have overlap ≈ 1. Thus the effect is not manifest in the rightmoving sector.

5.5 Conclusions

Let us review what we have computed in this paper:

It is widely believed that the semiclassical approximation to gravity holds at the horizon of a black hole. We have computed a quantity that is natural in the consideration of the black hole problem, and that does *not* behave semiclassically at the horizon of the black hole. This quantity is the quantum state of matter on a hypersurface which also catches the outgoing Hawking radiation. The crucial ingredient of our approach was that when we try to get the semiclassical approximation from

¶Recall that the hypersurfaces of section 3.3 had $\delta = \alpha^2 + 2\alpha\sqrt{M/\lambda}$ so the rhs of (5.4.41) depends only on α.

the full theory of quantum gravity, the natural quantity to compare between different semiclassical spacetimes is the same 1-geometry, not the hypersurface given by some coordinate relation on the semiclassical spacetimes. By contrast, in most calculations done with quantum gravity being a field theory on a background two dimensional spacetime, one computes n-point Greens functions, where the 'points' are given by chosen coordinate values in some coordinate system. For physics in most spacetimes, the answers would not differ significantly by either method, but in the presence of a black hole the difference is important.

We computed the quantum state on an entire spacelike hypersurface which goes up in time to capture the Hawking radiation, but then comes down steeply to intersect the infalling matter in the weak coupling region near the horizon. We found that quantum fluctuations in the background geometry prevent us from defining an unambiguous state on this S-surface. Matter states defined on an S-surface, evolved from a vacuum state at $\mathcal{I}^-$, are approximately orthogonal for fluctuations in mass of order e^{-M} or greater, a number much smaller than the size of fluctuations expected on general grounds.

One can expand the solution of the Wheeler–DeWitt equation in a different basis, such that for each term in this basis the total mass inside the hole is very sharply defined. If one ignores the Hawking radiation, then one finds that for such sharply defined mass the infalling matter has a large position uncertainty and cannot fall into the hole. Thus one may say that if one wants a good matter state on the S-slice, then the black hole cannot form. Any attempt to isolate a classical description for the metric while examining the quantum state for the matter will be impossible because the 'gravity' and matter modes are highly entangled. It is interesting that if we try to average over the 'gravity states' involved in the range $M \to M + \Delta M$, we generate entanglement entropy between 'gravity' and matter. This entropy is comparable to the entanglement entropy of Hawking pairs.

The computations of sections 3 and 4 show that the states on an S-surface differ appreciably in the region around the horizon. However, to calculate any local quantity close to the horizon, we could equally well have computed the state on spacelike hypersurfaces passing through the horizon without reaching up to $\mathcal{I}^+$. On these surfaces we would find an unambiguous state of matter for black holes with masses differing on the Planck scale. This feature may signal that an effective theory of black hole evaporation might not be diffeomorphism invariant in the usual way. It also indicates that the breakdown in the semiclassical approximation is relevant only if we try to detect *both* the Hawking radiation and the infalling matter. Susskind has pointed to a possible complementarity between the description of matter outside the hole and the description inside. 't Hooft and Schoutens *et. al.* have expressed this in terms of large commutators between operators localized at $\mathcal{I}^+$ and those localized close to the horizon. These notions of complementarity seem to be compatible with our results. It is interesting that we have arrived at them with minimal assumptions about the details of a quantum theory of gravity.

It should be mentioned that although every spacelike hypersurface that captures the Hawking radiation and the infalling matter near the horizon gives rise to the effect we have described, a slice that catches the Hawking radiation, enters the horizon high up, and catches the infalling matter deep inside the horizon can be seen to avoid the large shift. It seems that the quantum state of matter should be well defined on such a slice. The significance of this special case is not yet clear to us, although it is interesting that this type of slice appears to catch not only the Hawking pairs outside the horizon, but also their partners behind the horizon.

Our overall conclusion is that one must consider the entire solution of the quantum gravity problem near a black hole horizon, in particular one must take the solution to the Wheeler–DeWitt equation rather than its semiclassical projection. We believe that the arguments we have presented can be applied equally well to black holes in any number of dimensions.

5.6 Appendix A

In this appendix we shall explain the construction of a natural inner product relating states of a quantum field defined on different spacetimes in which the same hypersurface Σ is embedded. Clearly the Schrödinger picture allows us to compute the value of a state on any hypersurface Σ in the form

$$\Psi[f(x), t_0]$$

where in the chosen coordinate system Σ is the surface $t = t_0$. There is of course a Hamiltonian operator which is coordinate dependent, and which in the chosen coordinate system specifies time evolution on constant t hypersurfaces:

$$H[f(x), \pi_f(x), x, t]\Psi = i\dot{\Psi}$$

We shall assume that the evolution of a state is independent of the coordinate system used, in the sense that a state on a Cauchy surface Σ_1 is taken to define a *unique* state on a later surface Σ_2. This may not always be the case [105], but we will ignore such problems in our reasoning.

In quantum field theory on a fixed background, there is an inner product on the space of states on any hypersurface Σ. However, in order to use this inner product to compare states in different spacetimes, it is necessary to find a natural way of relating two states defined on Σ without reference to coordinates. A natural way of doing this is to use the proper distance along Σ to define a mode decomposition, and to compare the states with respect to this decomposition.

Define

$$\begin{aligned} a(k) &= \int \frac{ds}{\sqrt{4\pi\omega_k}}\, e^{iks} \left(\omega_k f(x(s)) + \frac{\delta}{\delta f(x(s))}\right) \\ a^\dagger(k) &= \int \frac{ds}{\sqrt{4\pi\omega_k}}\, e^{-iks} \left(\omega_k f(x(s)) - \frac{\delta}{\delta f(x(s))}\right) \end{aligned} \tag{5.6.42}$$

so that $[a(k), a^\dagger(k')] = \delta(k-k')$. It is then straightforward to define the familiar inner product on the corresponding Fock space. An easy way to picture the Fock space in this Schrödinger picture language is to transform to a representation $\Psi[a^\dagger(k), t]$, in which $a(k)$ is represented as $\delta/\delta a^\dagger(k)$. The "vacuum" state, annihilated by all the $a(k)$, is just the functional $\Psi = 1$, and excited states arise from multiplication by $a^\dagger(k)$. The inner product is

$$(\Psi_1, \Psi_2) = \int \frac{\prod_k da(k) da^\dagger(k)}{2\pi i} \left(\Psi_1[a^\dagger(k), t]\right)^* \Psi_2[a^\dagger(k), t] \exp\left[-\int dk\, a(k) a^\dagger(k)\right] \tag{5.6.43}$$

where $(a^\dagger(k))^* = a(k)$ [106].

We could carry out the same construction for any coordinate x on Σ, and the inner products would necessarily agree. However, the operators $a(k)$ and $a^\dagger(k)$ constructed using proper distance are special, in that we shall say that two states $\Psi_{\mathcal{M}}[f(x), t_1]$ and $\Psi_{\bar{\mathcal{M}}}[f(x), t_2]$ defined on different spacetimes $\mathcal{M}$ and $\bar{\mathcal{M}}$ but on a common hypersurface located at $t_{\mathcal{M}} = t_1$ or $t_{\bar{\mathcal{M}}} = t_2$, are the same if they are the same Fock states with respect to this decomposition. If they are not identical, their overlap is given by the Fock space inner product (5.6.43), and this is what is meant in Section 4 by

$$\langle \psi_1, \Sigma, \mathcal{M} | \psi_2, \Sigma, \bar{\mathcal{M}} \rangle.$$

Having defined an inner product between two Schrödinger picture states on different spacetimes, we want to extend it to Heisenberg picture states on $\mathcal{M}$ and $\bar{\mathcal{M}}$, since these are the kind of states we are used to working with in quantum field theory in curved spacetime. The inner product we have just defined can be used to relate Heisenberg picture states by transforming each of the states to the Schrödinger picture, evolving them to the common hypersurface Σ, and computing the overlap there. It is useful to have a short-cut to this computation. In order to achieve this, we first relate a Schrödinger picture state $\Psi[a^\dagger(k), t_0]$ to a Heisenberg picture state $\Psi[a^\dagger(k)]$, where now the $a(k)$ are associated with mode functions on $\mathcal{M}$ not Σ: First choose coordinates (x, t) on the spacetime such that the metric is conformally flat, that Σ is a constant time slice, $t = t_0$, and that the conformal factor is unity on Σ. Using these coordinates, we can compute the Hamiltonian, which by virtue of two dimensional conformal invariance is free

$$H = \int dx (\pi_f^2 + (f')^2).$$

We pick a mode basis defined by

$$\begin{aligned} a(k) &= \int \frac{dx}{\sqrt{4\pi\omega_k}}\, e^{ikx} \left(\omega_k f(x) + \frac{\delta}{\delta f(x)} \right) \\ a^\dagger(k) &= \int \frac{dx}{\sqrt{4\pi\omega_k}}\, e^{-ikx} \left(\omega_k f(x) - \frac{\delta}{\delta f(x)} \right) \end{aligned}$$

so that on Σ this is precisely the proper distance mode decomposition. In terms of these modes, the Hamiltonian is simply given by

$$H = \int dk \omega_k a^\dagger(k) a(k)$$

so that transforming the operators $a(k)$ and $a^\dagger(k)$ to the Heisenberg picture simply gives

$$a(k, t) = e^{i\omega_k (t - t_0)} a(k), \qquad a^\dagger(k, t) = e^{-i\omega_k (t - t_0)} a^\dagger(k)$$

and

$$f(x, t) = \int \frac{dk}{\sqrt{4\pi\omega_k}} \left(a(k) e^{-ik\cdot x} + a^\dagger(k) e^{ik\cdot x} \right).$$

Correspondingly the Schrödinger picture state $\Psi[a^\dagger(k), t]$ is identical in form to the Heisenberg picture state: $\Psi[a^\dagger(k)] \equiv \Psi[a^\dagger(k), t]|_{t=t_0}$. We may repeat this procedure on another spacetime $\bar{\mathcal{M}}$, again defining modes on $\bar{\mathcal{M}}$ so that $\Psi_{\bar{\mathcal{M}}}[a^\dagger(k)]$ is identical to the Schrödinger picture state on $\Sigma_{\bar{\mathcal{M}}}$. Then, the inner product (5.6.43) serves as an inner product for Heisenberg picture states $\Psi_{\mathcal{M}}[a^\dagger(k)]$ and $\Psi_{\bar{\mathcal{M}}}[a^\dagger(k)]$ living on spacetimes $\mathcal{M}$ and $\bar{\mathcal{M}}$.

Now in order to compare a Heisenberg picture state $\Psi_{\mathcal{M}}[a^\dagger(k)]$ to any other Heisenberg picture state on $\mathcal{M}$, we may make use of standard Bogoliubov coefficient techniques. Consider another state defined in the Heisenberg picture in terms of mode-coefficients related to modes $v_p(x,t)$ on $\mathcal{M}$. Let us suppose that we associate operators $b(p)$ and $b^\dagger(p)$ with the modes $v_p(x,t)$. Then

$$\begin{aligned} b(p) &= \sum_k \left[\alpha_{kp} a(k) + \beta^*_{kp} a^\dagger(k)\right] \\ b^\dagger(p) &= \sum_k \left[\beta_{kp} a(k) + \alpha^*_{kp} a^\dagger(k)\right] \end{aligned} \tag{5.6.44}$$

where

$$\alpha_{kp} = -i\int dx f_k(x,t)\partial_t v^*_p(x,t), \qquad \beta_{kp} = i\int dx f_k(x,t)\partial_t v_p(x,t).$$

Here $f_k(x,t) = e^{-ik\cdot x}/\sqrt{4\pi\omega_k}$ are the modes defining the $a(k)$.

We may perform a similar calculation on a neighbouring spacetime $\bar{\mathcal{M}}$, also containing Σ, to relate a set of modes $\bar{v}_q(\bar{x},\bar{t})$ to the modes $f_k(\bar{x},\bar{t})$ and similarly to relate the operators $a(k)$ and $a^\dagger(k)$ to the $\bar{b}(p)$ and $\bar{b}^\dagger(p)$ as

$$\begin{aligned} \bar{b}(p) &= \sum_k \left[\bar{\alpha}_{kp} a(k) + \bar{\beta}^*_{kp} a^\dagger(k)\right] \\ \bar{b}^\dagger(p) &= \sum_k \left[\bar{\beta}_{kp} a(k) + \bar{\alpha}^*_{kp} a^\dagger(k)\right] \end{aligned} \tag{5.6.45}$$

Now we can use the inner product (5.6.43) to relate two states $\Psi_{\mathcal{M}}[b^\dagger(p)]$ and $\Psi_{\bar{\mathcal{M}}}[\bar{b}^\dagger(p')]$ directly.

More simply, it follows from (5.6.44) and (5.6.45) that the b's and $\bar{b}$'s are related by

$$\begin{aligned} b(p') &= \sum_p \left[(\bar{v}_p, v_{p'})\bar{b}(p) + (\bar{v}^*_p, v_{p'})\bar{b}^\dagger(p)\right] \\ b^\dagger(p') &= \sum_p \left[-(\bar{v}^*_p, v^*_{p'})\bar{b}^\dagger(p) - (\bar{v}_p, v^*_{p'})\bar{b}(p)\right] \end{aligned} \tag{5.6.46}$$

where

$$(\bar{v}_p, v_{p'}) = -i\int_\Sigma dx \bar{v}_p(x,t)\partial_t v^*_{p'}(x,t) \tag{5.6.47}$$

so that the inner product between states on $\mathcal{M}$ and $\bar{\mathcal{M}}$ may be computed using the standard inner product for states $\Psi[b^\dagger(p)]$ without going through the $a(k)$.

In the examples that we consider, the Bogoliubov coefficients in (5.6.46) need not be evaluated on Σ as in (5.6.47). Suppose for example that we have left moving mode bases $v_p(\sigma^+)$ and $\bar{v}_p(\bar{\sigma}^+)$ defined in terms of tortoise coordinates on $\mathcal{M}$ and $\bar{\mathcal{M}}$ respectively. Then both v_p and $\bar{v}_p$ are functions of x^+ only. We can simply change variables in (5.6.47) from x to σ (the t differentiation becomes an x differentiation which absorbs the factor $dx/d\sigma$), yielding

$$(\bar{v}_p, v_{p'}) = -i\int d\sigma \bar{v}_p(\bar{\sigma}^+(\sigma^+))\, \partial_{\sigma^0}\, v^*_{p'}(\sigma^+) \tag{5.6.48}$$

where $\bar{\sigma}^+$ is given as a function σ^+ through the relations derived by equating points on Σ according to the values of ϕ and $d\phi/ds$, as in Section 3. The integral (5.6.48) looks exactly like the familiar integral for Bogoliubov coefficients, even though it involves mode functions on different manifolds. (5.6.48) may be evaluated on any Cauchy surface in $\mathcal{M}$ (or $\bar{\mathcal{M}}$) since both the mode functions solve the Klein-Gordon equation on $\mathcal{M}$ ($\bar{\mathcal{M}}$).

5.7 Appendix B

We present some details of the calculation of the overlap of the two states on Σ. We now know the Bogoliubov transformation between the modes v_ω and $\bar{v}_\omega$ in the text. Subsequently, the overlap of the two Schrödinger picture states can be found to be

$$\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle = (\det(\alpha))^{-\frac{1}{2}} \quad , \tag{5.7.49}$$

where α is the matrix $(\alpha_{\omega\omega'})$ of Bogoliubov coefficients. The right hand side is the general formula for the overlap of two vacuum states related to modes connected by a Bogoliubov transformation [107]. However, it is more convenient to consider not the overlap but the probability amplitude

$$|\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle|^2 = (\det(\alpha\alpha^\dagger))^{-\frac{1}{2}} \quad , \tag{5.7.50}$$

where the components of the matrix $\alpha\alpha^\dagger$ are

$$(\alpha\alpha^\dagger)_{\omega\omega'} = \int_0^\infty d\omega'' \; \alpha_{\omega\omega''}\alpha^*_{\omega'\omega''} \quad . \tag{5.7.51}$$

The evaluation of the determinant of the matrix $\alpha\alpha^\dagger$ becomes easier if we move into a wavepacket basis. Instead of the modes v_ω we use

$$v_{jn} \equiv a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega \; e^{2\pi i\omega n/a} \; v_\omega \quad . \tag{5.7.52}$$

These wavepackets are centered at $\sigma^+ = 2\pi n/a$, where $n = \ldots, -1, 0, 1, \ldots$, they have spatial width $\sim a^{-1}$ and a frequency $\omega_j \approx ja$, where $j = 0, 1, \ldots$. For more discussion, see [5, 104, 94]. In the new basis, the Bogoliubov coefficients become

$$\alpha_{jn\omega'} = a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega \; e^{2\pi i\omega n/a} \; \alpha_{\omega\omega'} \tag{5.7.53}$$

$$\beta_{jn\omega'} = a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega \; e^{2\pi i\omega n/a} \; \beta_{\omega\omega'}$$

with the normalization

$$\int_0^\infty d\omega'' \; [\alpha_{jn\omega''}\alpha^*_{j'n'\omega''} - \beta_{jn\omega''}\beta^*_{j'n'\omega''}] = \delta_{jj'}\delta_{nn'} \quad . \tag{5.7.54}$$

The thermal relation (5.4.38) becomes

$$\beta^*_{jn\omega'} \approx -e^{-\pi\omega_j/\lambda}\alpha_{jn\omega'} \quad . \tag{5.7.55}$$

Recall that the validity of the thermal approximation corresponded to the region $\lambda\sigma^+ \in (0, \ln[1+(e-1)\lambda\Delta])$ where $\lambda\Delta - 1$ was the shift $\lambda\bar{x}^+ - \lambda x^+$. Let us denote the size of this region as λL. Since the separation of the wavepackets is $\Delta(\lambda\sigma^+) = 2\pi\lambda/a$, we can say that

$$n_{max} = \frac{\lambda L}{\Delta(\lambda\sigma^+)} = \frac{\ln[1+(e-1)\lambda\Delta]}{2\pi\lambda/a} \tag{5.7.56}$$

packets are centered in this region.

Combining (5.7.54) and (5.7.55), we now see that

$$(\alpha\alpha^\dagger)_{jnj'n'} \approx \frac{\delta_{jj'}\delta_{nn'}}{1 - e^{-2\pi\omega_j/\lambda}} \tag{5.7.57}$$

for n, n' "inside" λL. For the other values of n, n' (at least one of them being "outside"),

$$(\alpha\alpha^\dagger)_{jnj'n'} \approx \delta_{jj'}\delta_{nn'} \; . \tag{5.7.58}$$

We are now ready to calculate the overlap (5.7.50). We get

$$\begin{aligned} \ln\left[|\langle 0 \text{ in}, \Sigma, \mathcal{M}|0 \text{ in}, \Sigma, \bar{\mathcal{M}}\rangle|^2\right] &\approx -\frac{1}{2}\left\{ \sum_n^{\text{(inside)}} \sum_j \ln\left[\frac{1}{1-e^{-2\pi\omega_j/\lambda}}\right] \right. \\ &\qquad \left. + \ln\left[\prod_n^{\text{(outside)}} \prod_j 1\right]\right\} \\ &= \frac{1}{2}\, n_{max} \sum_j \ln\left[1 - e^{-2\pi j a/\lambda}\right] \; . \end{aligned} \tag{5.7.59}$$

In order to estimate the last term, we convert the sum to an integral[||]:

$$\begin{aligned} \sum_j \ln\left[1 - e^{-2\pi j a/\lambda}\right] &\to \int_0^\infty dj\; \ln\left[1 - e^{-2\pi j a/\lambda}\right] \\ &= \frac{\lambda}{2\pi a}\left(\frac{-\pi^2}{6}\right) \; . \end{aligned} \tag{5.7.60}$$

Now, combining (5.7.59) and (5.7.60), we finally get a useful formula for the overlap:

$$\begin{aligned} |\langle 0 \text{ in}, \Sigma, \mathcal{M}|0 \text{ in}, \Sigma, \bar{\mathcal{M}}\rangle|^2 &\approx \exp\left\{-\frac{1}{2}\frac{\lambda L}{2\pi\lambda/a}\frac{\lambda}{2\pi a}\frac{\pi^2}{6}\right\} \\ &= e^{-\frac{1}{48}\lambda L} \; . \end{aligned} \tag{5.7.61}$$

Now we can estimate when the overlap is $< \gamma^{-1}$ where γ is a number $\sim e$. The overlap becomes equal to γ^{-1} as

$$\begin{aligned} 48 \ln\gamma &= \lambda L = \ln[1+(e-1)\lambda\Delta] \\ &\approx \ln\left[1 + (e-1)\left(1 - \frac{\Delta M}{2\alpha\lambda}\sqrt{\frac{\lambda}{M}}\right)\right] \; , \end{aligned} \tag{5.7.62}$$

[||]Notice that one might like to exclude frequencies corresponding to wavelengths much larger than the thermal region λL and impose an infra-red cut off at $j_{min}a \sim 1/L$. It turns out that for $\infty > \lambda L > 2\pi$ $(0 < j_{min}a < \lambda/2\pi)$ the effect of imposing this cut off is negligible. Therefore we can just as well take the integral over the full range.

where we used (5.3.20) in the last step. Solving for ΔM, we get

$$\frac{\Delta M}{\lambda} \approx -\frac{2}{e-1}\gamma^{48}\sqrt{\frac{M}{\lambda}}\,\alpha\ . \tag{5.7.63}$$

If $|\Delta M/\lambda|$ is bigger, the states are approximately orthogonal. Notice that since we used (5.3.20) in the end, (5.7.63) is a special result for the hypersurfaces of section 3.3. However, it is straightforward to generalize (5.7.63) to any S-surface of section 3.2 by using the relevant shifts as $\lambda\Delta - 1$ and proceeding as above. In general the right hand side of (5.7.63) will then depend on both α and the intercept δ.

Chapter 6

A Proposal for an Effective Theory for Black Holes

"Tracking what ?" said Piglet, coming closer
"That's just what I ask myself. I ask
myself, what ?"
"What do you think you'll answer"
"I shall have to wait until I catch up
with it" said Winnie-the-Pooh.
(The World of Pooh by, A. A. Milne)

Abstract:

We extend the discussion of chapter 5 of the breakdown of the semiclassical approximation near a black hole horizon to four dimensional black holes. We propose an effective theory for matter interacting with the black hole background taking into account the gravity fluctuation and discuss its properties.

This chapter is based on ongoing work in collaboration with Miguel Ortiz

6.1 Four Dimensional Black Holes

We start with the Schwarzchild black hole metric in Kruskal coordinates.

$$ds^2 = -\frac{32M^3}{r}e^{-r/2M}dUdV + r^2 d\Omega^2$$

Where $U = e^{-u/4M}$, $V = e^{v/4M}$, r is the usual Schwarzchild coordinate and $u = t - r^*$, $v = t + r^*$ and $r^* = r + 2M \ln |r/2M - 1|$ is the tortoise coordinate.

We have

$$-UV = (r/2M - 1)e^{r/2M}$$

We are only interested in the metric near the horizon and we expand around $r = 2M$ hence we write $r = 2M(1 + \bar{\epsilon})$. then one gets $-UV = \bar{\epsilon}$.

Define $x^+ = \sqrt{32M^3}V$, $x^- = \sqrt{32M^3}U$, $\kappa = 4M$ and ρ, A as follows

$$ds^2 = -e^{2\rho}dUdV + Ad\Omega^2$$

Then we get

$$e^{-2\rho} = \sqrt{A} = 2M - \kappa x^+ \kappa x^-$$

We restrict ourself to spherically symmetric spacelike hypersurfaces. This implies that the hypersurface is defined geometrically by its area and its intrinsic geometry in the t, r plane. The area is given by A. It is a scalar under coordinate transformation in the t, r plane. The intrinsic geometry of a spherically symmetric hypersurface is then given by the function $A(s)$ where s is the proper distance in the t, r plane. In two different space times we say the hypersurfaces are the same if they define the same function $A(s)$

Consider two different mass black holes with masses M and $\bar{M}$. We can define a (spherically symmetric) spacelike hypersurface on the first spacetime through a relation between x^+ and x^-, say $\kappa x^- = g(\kappa x^+)$. Given the form of the metric, it is easy to deduce the function $A(s)$ which defines the intrinsic geometry. The condition that $A(s)$ be the same in two different spacetimes defines a simple relation between points labelled by co-ordinates $x^\pm$ on one spacetime and points labelled by coordinates $\bar{x}^\pm$ on another spacetime. Suppose that we define a spacelike hypersurface in the other spacetime by $\bar{\kappa}\bar{x}^- = \bar{g}(\bar{\kappa}\bar{x}^+)$. Then the condition that $A(s) = \bar{A}(\bar{s})$ is equivalent to the two conditions:

$$A(x^+, g(x^+)) = \bar{A}(\bar{x}^+, \bar{g}(\bar{x}^+)) \tag{6.1.1}$$

$$\frac{dA}{ds}(x^+, g(x^+)) = \frac{d\bar{A}}{d\bar{s}}(\bar{x}^+, \bar{g}(\bar{x}^+)) \tag{6.1.2}$$

One then finds that the relationship between the coordinates of the same points in the two different mass black hole is (see also chapter 5.3.1)

$$b \ln(\bar{\kappa}\bar{x}^+) = 2\int dy \frac{g'}{(g + yg') \mp \sqrt{b^2(g + yg')^2 - 4g'(2\Delta M + yg)}} \tag{6.1.3}$$

Where $y = \kappa x^+$ and $b = \kappa/\bar{\kappa}$.

For a S-surface $g = -\alpha^2 \kappa x^+ + \delta$ where $\alpha^2 \ll \delta \ll 1$. We find to a good approximation, a result similar to that found in chapter 5.

$$\bar{\kappa}\bar{x}^+ = \kappa x^+ + \frac{2\Delta M}{\delta}.$$

Two incoming states on different space times of mass M and $\bar{M}$ can be compared on a spacelike hypersurface (Σ) through the proper distance modes on Σ. A vacuum state with respect to modes $\phi(\bar{(x)}^+)$ on the $\bar{M}$ space time is mapped to a vacuum state with respect to modes $\phi(\bar{x}^+(x^+))$ on the M space time. The inner product between the two states can be computed (see appendix A and B in chapter 5). For the natural in-coming vacuum for which the positive frequency modes are of the form e^{-iwv} , one finds that the inner product between the two states on the S-surface considered above is

$$|\langle 0_{in}, M, \Sigma | 0_{in}, \bar{M}, \Sigma \rangle|^2 = \exp\left[-\ln[\frac{2\Delta M}{\sqrt{2M}\delta}]/48\right] \tag{6.1.4}$$

This becomes very small as δ is very small on S-surfaces.

All this is true if the matter field was only coupled to the metric of the t, r plane. However in four dimension this is not the case, so a little care must be taken

As usual we decompose the incoming matter waves as

$$f(x) = \sum_{l,m} r^{-1} Y_{l,m} u_{l,m}(r,t)$$

$u_{l,m}(r,t)$ solves a two dimensional wave equation. The equation is basically that of a free field for $l = 0$ and has a potential term for $l > 0$. If we restrict ourselves to $l = 0$ waves only (as we actually did above) we have proven the orthogonality. In fact that is enough. However one can generalize the construction of chapter 5 to deal with the higher angular momentum modes. We will not show this construction here. The only quantitative difference is that the higher angular momentum modes are not conformally coupled to the two dimensional metric, and hence their propagation is not free. This can be dealt with. The difference between the inner product is most easily seen in the difference of the Bogolubov coefficients between the two states. We find $(a \sim 1)$

$$|\beta_{w,l,w',l'}|^2 = \frac{\Theta(awl - 2M)}{e^{2\pi w/\kappa} - 1}$$

The inner product will be modified and have the form

$$|\langle 0_{in}, l, M, \Sigma | 0_{in}, l, \bar{M}, \Sigma \rangle|^2 = \exp\left[-e^{-l}\ln[\frac{2\Delta M}{\sqrt{2M}\delta}]/48\right] \tag{6.1.5}$$

Thus for high enough l the inner product will be close to 1.

We conclude that for four dimensional Schwarzchild black hole the semiclassical approximation breaks down in the same way as described in chapter 5.

We would like to note that there are two important ingredient coming into the breaking of the semiclassical approximation. One is the large shift in the Kruskal coordinate and the other is the natural chosen incoming state. For space times that have a large shift (but the natural state to consider is for instance the Kruskal vacuum) there will be no breakdown of the semiclassical approximation. It is also possible to show that the semiclassical approximation breaks down for matter propagating on a charged black hole background but not for an extremal black hole; exactly because of the above reason.

6.2 An Effective semiclassical Theory

let us summarize the situation. In a black hole background if we start with the natural state at $\mathcal{I}^-$ we find that due to the small uncertainty of the mass of the black hole the states become almost orthogonal on a certain class of hypersurfaces. If we want to describe the physics in terms of outside observers only (outside the horizon) we have *no* choice but to use those hypersurfaces. We also found that the state on the same hypersurface in two different space times seem to be "thermally related" with a temperature $T = \kappa/2\pi$.

We would like to construct an effective theory (for these observers) that will take into account the quantum gravity fluctuations. In what follows we will explore its properties. One needs to integrate over the un-observed fluctuations of the mass. The assumption of semiclassicality together with the tracing over the unobserved fluctuation around mass M_0 gives a density matrix for the matter field

$$\rho = \frac{1}{\mathcal{N}} \sum_{i=-\mathcal{N}/2}^{\mathcal{N}/2} |f_i(\Sigma)\rangle\langle f_i(\Sigma)|. \tag{6.2.6}$$

Where $|f_i(\Sigma) >$ is the matter state starting from vacuum on $\mathcal{I}^-$ in the space time with mass $M_0 + i\Delta M_0$, and propagated to the surface Σ. Here ΔM_0 is the minimum fluctuation for the states to be different and $\mathcal{N}(\Sigma)$ is the number of orthogonal states in the range of fluctuation of the mass.

We can calculate what are the states $|f_i >$, and what is the mean energy momentum tensor of these states. Now the states on Σ are going to be different only in a region of space time where the β Bogolubov coefficients are non zero. In other words only in the approximate region $\eta < \kappa v < \ln 2\Lambda$ where the size of η, Λ depends on the black hole. For the Schwarzchild black hole $\Lambda = \frac{2|\Delta M|}{\sqrt{2M\delta}}$, for the CGHS model $\Lambda = \frac{|\Delta M|}{\delta}$ and $\eta = \frac{\alpha+\sqrt{\bar{M}/\lambda}}{\alpha+\sqrt{M/\lambda}}$ Notice that $\Lambda > e^M$ for S-surfaces.

Above and below that region the states are basically the same regardless of the small fluctuation of the mass of the black hole. The state on the mean space time is just the vacuum state with energy zero propagated to some hypersurface. This state is the vacuum with respect to some modes $\phi(x)$. On any other space time the state on the hypersurface is the vacuum state with respect to modes $\phi(\bar{x}(x))$. We know on each S-surface the relationship $\bar{v}(v)$. With this information we can find the states $|f_i(\Sigma) >$.

Expanding the matter field operator on a certain hypersurface as

$$\hat{\phi}(v) = \sum_i a_i \phi_i(v) + a_i^\dagger \phi_i^*(v)$$

and

$$\hat{\phi}(\bar{v}(v)) = \sum_i b_i^{\bar{M}} \phi_i(\bar{v}(v)) + b_i^{\dagger \bar{M}} \phi_i^*(\bar{v}(v))$$

One has the relationship

$$a_k = e^{iJ^{\bar{M}}} b_k^{\bar{M}} e^{-iJ^{\bar{M}}}$$

Where $J(a, a^\dagger)$ is a function of the hypersurface. Then

$$|f_i\rangle = e^{-iJ^{\bar{M}}}|0\rangle$$

In a wave packet basis $J = 0$ for wave packets centered in the κx^+ coordinate above 2Λ and below η on Σ. One can see that the interpretation of J is that of an integrated hamiltonian interaction

$$J = \int_0^{t(\Sigma)} H_{int}(t) dt, \tag{6.2.7}$$

where H_{int} is the Hamiltonian interaction between the fluctuating gravitational field and the matter fields.

At this point we would like to remind the reader that all this is true when we consider S-surfaces. There is nothing locally happening as each point could be considered to be on a non S-surface. So these effects have to be interpreted carefully.

6.2.1 Energy

After we trace over the small range of mass fluctuations, from a semiclassical picture the state on $\bar{M}$ becomes a state on M (on some Σ on M). The state on mass M has $T_{\mu\nu} = 0$. One can calculate the energy momentum tensor of all the other induced states.

For notational simplicity define $p(v) = \bar{v}(v)$ (this is understood to be dependent on the hypersurface). Then a standard calculation shows that [4]

$$\langle T_{vv}^{\bar{M}} \rangle \equiv \langle f_i | T_{vv} | f_i \rangle = \frac{1}{12\pi} (p')^{1/2} \partial_v^2 (p')^{-1/2}, \tag{6.2.8}$$

and all other components are zero. Assuming p has at least a continuous third derivative and $p' \longrightarrow 1$ when $v \longrightarrow \pm\infty$ then we find

$$E_{tot}^{\bar{M}} \equiv \int_{-\infty}^{\infty} dv \langle T_{vv}^{\bar{M}} \rangle = \frac{1}{48\pi} \int_{-\infty}^{\infty} (\frac{p''}{p'})^2 > 0.$$

These calculation can be performed in the CGHS model (and in the Schwarzchild case under some assumption about the incoming shock wave). From chapter 5 (and analogues for the Schwarzchild case) we see that in the three regions the function $p(v)$ has the form

$$p(v) = \frac{1}{\kappa} \ln[e^{\kappa v} + B] + C.$$

Further although p is continuous it is not differentiable. This is due to the fact that we have taken the shock wave to be a delta function. Smearing the shock wave a bit we will get a smother function $p(v)$. For this form of p one gets

$$\langle T_{vv}\rangle = \frac{\kappa^2}{48\pi}[1 - \frac{1}{1+Be^{-\kappa v}}].$$

One can calculate B for the CGHS model from chapter 5 ($\kappa = \lambda$, $v = \sigma^+$ $\delta = 2\alpha\sqrt{M/\lambda}$). For $\Delta M < 0$

$$B = \begin{cases} \Lambda & \kappa v > 0 \\ -\tilde{\eta} & \ln(\eta) < \kappa v < 0 \\ 0 & \kappa v < \ln(\eta) \end{cases} \tag{6.2.9}$$

where $\eta = \frac{\alpha+\sqrt{\bar{M}/\lambda}}{\alpha+\sqrt{M/\lambda}}$ and $\tilde{\eta} = \frac{\sqrt{\bar{M}/\lambda}}{\alpha+\sqrt{M/\lambda}}$. For $\Delta M > 0$ we find

$$B = \begin{cases} -\Lambda & \kappa v > \ln(\Lambda+1) \\ \tilde{\Lambda} & 0 < \kappa v < \ln(\Lambda+1) \\ 0 & \kappa v < 0 \end{cases} \tag{6.2.10}$$

where $\tilde{\Lambda} = \frac{\sqrt{\bar{M}/\lambda}}{\alpha}$ and we remind the reader that α is very small. For the Schwarzchild case if we assume that the metric near the shock wave is similar to that of CGHS then the above results apply. The total energy is

$$E_{tot}^{\bar{M}} \approx \frac{\kappa\Lambda}{48\pi}$$

Taking into account the smearing of the shock wave and the positivity of the total energy we can deduce the form of the $T_{vv}^{\bar{M}}(v)$. It is displayed in figures 8 and 9

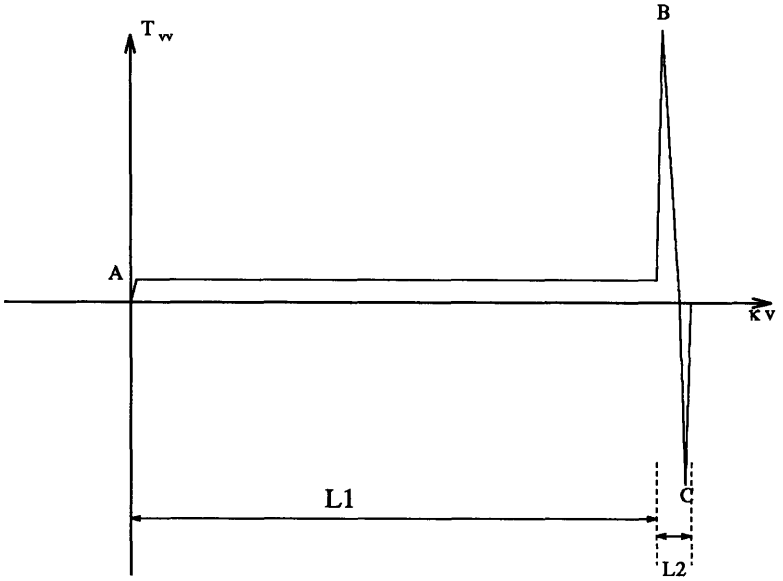

Figure 1: The graph of the $T^{\bar{M}}_{vv}$ is shown for $\Delta M > 0$. $L1 \approx \ln \Lambda$, $L2 \approx 1$. The Hight of peak B is $\frac{\kappa^2}{24\pi}\Lambda^2$, and is twice that of peak C. Point A is at $A = (0, \kappa^2/48\pi)$.

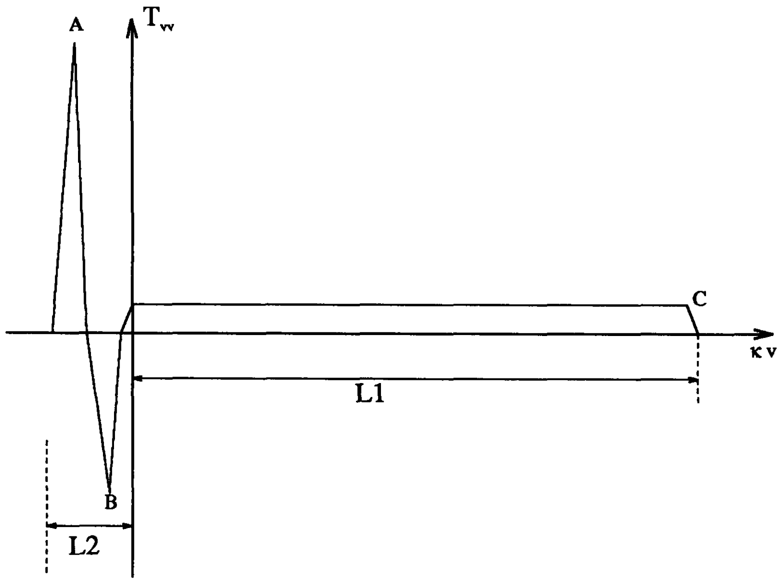

Figure 2: The graph of the $T^{\bar{M}}_{vv}$ is shown for $\Delta M < 0$. $L1 \approx \ln \Lambda$, $L2 \approx |\ln \eta| \approx |\Delta M/2M|$. Hight of peak A is $\frac{\kappa^2}{24\pi}\tilde{\Lambda}^2$, and is twice that of peak B. point C is at $C = (\ln \Lambda, \kappa^2/48\pi)$.

6.2.2 Entropy

What is the physical picture emerging. On the simplest level we see that the semiclassical approximation breaks down and presumably one has to resort to a full theory of gravity. However it is worth pushing the semiclassical approximation a bit further. We see that for S-surfaces intersecting the shock wave a distant δ in κx^- coordinate, there is a large shift of the x^+ coordinate which transcends to a large change of the matter state on the hypersurface for coordinate $e^{-\eta} < x^+ < 2\Lambda$. Thus there is a timelike line which is approximately $r = 2M + 4(\Delta M)_{max}$ (for Schwarzchild) and $r = M + 2(\Delta M)_{max}$ (for CGHS), that is the boundary between the right region (R) where there is no change of the state due to quantum gravity effects and the left region (L) where there is. Further in region L there is a large induced energy momentum tensor. Semiclassical observations can not be made in the R region if it is in the future lightcone of the left region (see figure).

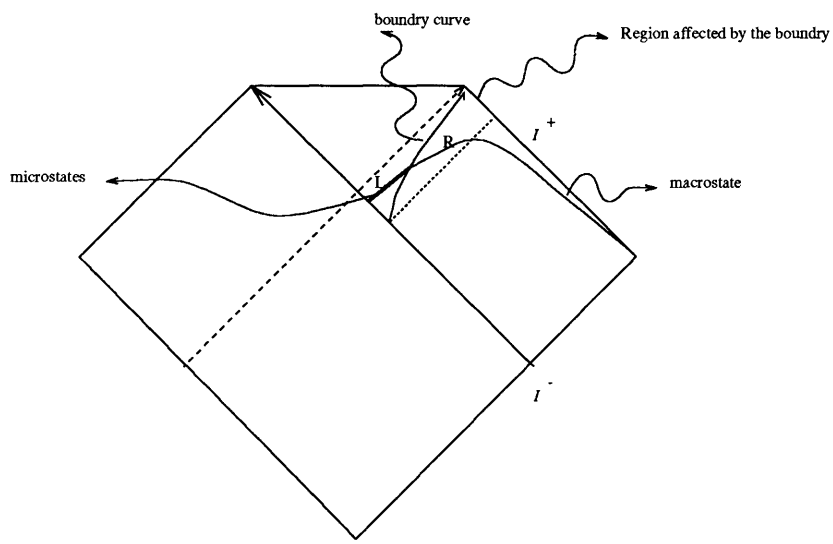

Figure 3: The boundary curve is shown on a CGHS background. For every unique state on the right of the curve there are many possible states on the left.

For a S-surface the state is unique on the right hand side of the boundary curve but on the left there are many corresponding states (microstates) for the same state (macrostate) at $\mathcal{I}^+$. Thus the boundary line can be endowed with energy (maybe like the one computed above), entropy (log of the number of different states) and some dynamics. All observations made at $\mathcal{I}^+$ have to take into account this boundary. The effect of the quantum gravity fluctuations on the matter is then seen to be described in an effective way as if the boundary is some hot membrane. Notice that this is a similar picture to that of the stretched horizon put forward in [90].

Now can the subscribed entropy be the origin of the black hole entropy. It is not clear yet. It is intriguing to notice that if one takes the S-surface that captures an amount of energy from the hawking radiation $E \approx \frac{M - Q\Phi}{2}$, then on that S-surface

$$\ln \mathcal{N} \sim \frac{A}{\hbar G} \sim S_{BH}$$

and this quantity is independent of the number of matter fields present.

Bibliography

[1] S. W. Hawking, *Phys. Rev.* **D14** (1976) 2460.

[2] C. J. Isham, *Canonical quantum gravity and the problem of time*, presented at the 19th International Colloquium of Group Theoretical Methods in Physics, Salamanca, Spain, July 1992 (gr-qc/9210011).

[3] D. Birmingham, M. Blau, M. Rakowski and G. Thompson, *Phys. Reps.* **209** (1991) 129.

[4] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*, Cambridge University Press, Cambridge, UK (1982).

[5] S. W. Hawking, *Comm. Math. Phys.* **43** (1975) 199.

[6] S. Giddings, *Black holes and quantum predictability*, Santa Barbara Report No. UCSBTH-93-16 (hep-th/9306041).

[7] R. Brooks and G. Lifschytz, *Nucl. Phys* **B438** (1995) 211

[8] C. Rovelli and L. Smolin, *Nucl. Phys.* **B331** (1990) 80; C. Rovelli, *Class. Quan. Grav.* **8** (1991) 1613.

[9] L. Smolin, "Recent Developments in Nonperturbative Quantum Gravity", in the Proceedings of the 1991 GIFT International Seminar on Theoretical Physics: *Quantum Gravity and Cosmology*, Saint Feliu de Guixols, Catalonia, Spain.

[10] A. Ashtekar, *Phys. Rev. Lett.* **57** (1986) 2244, *Phys. Rev.* **D36** (1987) 1587; *Lectures on Non-Perturbative Canonical Gravity*, World Scientific, Singapore, 1991.

[11] C. Rovelli and L. Smolin, *Phys. Rev. Lett.* **61** (1988) 1155, *Nucl. Phys.* **B133** (1990) 80; R. Gambini and A. Trias, *Phys. Rev.* **D39** (1989) 3127.

[12] E. Witten, *Nucl. Phys.* **B311** (1988) 46.

[13] S. K. Donaldson and P. B. Kronheimer, *The Geometry of Four-Manifolds*, Oxford University Press, Oxford, 1991.

[14] R. Brooks, *Phys. Lett.* **289B** (1992) 317.

[15] J. Sonnenschein, *Phys. Rev.* **D42** (1990) 2080.

[16] M. F. Atiyah and I. M. Singer, *Proc. Nat'l. Acad. Sci.* (USA) **81** (1984) 2597.

[17] M. F. Atiyah, *Publ. Math. Inst. Hautes Etudes Sci. Paris* **68** (1989) 175.

[18] R. Brooks and G. Lifschytz, work in progress.

[19] A. S. Schwarz, *Lett. Math. Phys.* **2** (1978) 247; *Commun. Math. Phys.* **67** (1979) 1.

[20] A. Ashtekar, V. Husain, C. Rovelli, J. Samuel and L. Smolin, *Class. Quan. Grav.* **6** (1989) L185.

[21] R. Brooks and S. J. Gates, Jr, "Extended Supersymmetry and Super-BF Gauge Theories", MIT-CTP preprint # 2339, July 1994.

[22] B. Broda, *Phys. Lett.* **254B** (1991) 111.

[23] E. Witten, *Commun. Math. Phys.* **117** (1988) 353; "Introduction to Cohomological Field Theories", lectures given at the *Workshop on Topological Methods in Physics*, ICTP, Trieste, Italy, June 1990.

[24] D. B. Ray, *Adv. Math.* **4** (1970) 109; D. B. Ray and I. M. Singer, *Adv. Math.* **7** (1971) 145; W. Müller, *Adv. Math.* **28** (1978) 233; J. Cheeger, *Ann. Math.* **109** (1979) 259

[25] E. Witten, *Nucl. Phys.* **B323** (1989) 113.

[26] G. Thompson, "1992 Trieste Lectures on Topological Gauge Theory and Yang-Mills Theory", in Proceedings, High Energy Physics and Cosmology, ICTP, Trieste, Italy, 1992.

[27] J. B. Hartle and S. W. Hawking, *Phys. Rev.* **D28** (1983) 2960.

[28] D. Rolfson, *Knots and Links*, Publish or Perish, Inc., Houston, 1990.

[29] E. Witten, *Phys. Lett.* **206B** (1988) 601; R. Myers and V. Periwal, *Nucl. Phys.* **B361** (1991) 290; R. Myers, *Nucl. Phys.* **B343** (1990) 705.

[30] B. de Wit, "Formulations of N=2 Supergravity Theories", in *Erice EPS: Unification*, 1980, 327

[31] D. Anselmi and P. Fré, *Nucl. Phys.* **B392** (1993) 401.

[32] E. Witten, *Commun. Math. Phys.* **121** (1989) 351.

[33] G. Horowitz, *Commun. Math. Phys.* **125** (1989) 417.

[34] D. Birmingham, M. Blau and G. Thompson, *Intl. J. Mod. Phys.* **A24** (1990) 4721.

[35] E. Witten, *J. Geom. Phys.* **9** (1992) 303.

[36] M. Blau and G. Thompson, *Nucl. Phys.* **B400** (1993) 345.

[37] G. Lifschytz S. D. Mathur and M. Ortiz *A Note on the Semiclassical Approximation in Quantum Gravity* MIT-CTP-2384 gr-qc/9412040 Submitted to *Phys. Rev.* **D**

[38] R.Arnowitt, S. Deser and C. Misner *Phys. Rev.* **117**, (1960) 1595.

[39] V. Lapchinsky and V. Rubakov, *Acta Phys. Pol.* **B10** (1979) 1041; T. Banks, *Nucl. Phys.* **B249** (1985) 332

[40] J. J. Halliwell *Phys. Rev.* **D36**, (1987) 3626.

[41] J. Hartle, *Progress in quantum cosmology*, in *Proceedings, 12th Conference on General Relativity and Gravitation, Boulder, 1989*, edited by N. Ashby, D. F. Bartlett and W. Wyss (CUP, Cambridge, 1990); A. Vilenkin, *Phys. Rev.* **D39** (1989) 1116; C. Kiefer, *The semiclassical approximation to quantum gravity*, Freiburg University Report No. THEP-93/27,to appear in *Canonical Gravity - from Classical to Quantum*, edited by J. Ehlers and H. Friedrich (Springer, Berlin 1994) (gr-qc/9312015).

[42] S. Habib and R. Laflamme, *Phys. Rev.* **D42** (1990) 4056.

[43] U. Gerlach, *Phys. Rev.* **177** (1969) 1929.

[44] S. D. Mathur, *Black hole entropy and the semiclassical approximation*, MIT report No. CTP-2304, to appear in the proceedings of the International Colloquium on Modern Quantum Field Theory II at TIFR (Bombay), January 1994 (hep-th/9404135).

[45] E. Keski-Vakkuri, G. Lifschytz, S. Mathur and M. E. Ortiz, *Phys. Rev.* **D51** (1995) 1764.

[46] A. Peres *Nuovo Cimento* **XXVI**, (1962) 53.

[47] K. V. Kuchař, *Phys. Rev.* **D39** (1989) 1579.

[48] D. Cangemi and R. Jackiw, *Phys. Lett.* **B337** (1994) 271.

[49] D. Giulini and C. Kiefer, *Consistency of semiclassical gravity*, Freiburg Report THEP-94-20 (gr-qc/9409014).

[50] H. Goldstein *Classical Mechanics*, Addison-Wesley publishing Company, (1980).

[51] A. Komar, *Observables, correspondence and quantized gravitation*, in *Magic Without Magic*, ed. J. Klauder, San Francisco, 1972.

[52] C. Crnkovic and E. Witten, *Covariant description of canonical formalism in geometrical theories*, in *Three Hundred Years of Gravitation*, eds. S. Hawking and W. Israel, Cambridge University Press, Cambridge, UK 1987.

[53] S. Mathur *Real time propagation in the first quantized formalism*, , MIT-CTP-2273, to appear in the proceedings of the Thermal Fields Workshop, Banff, Canada, August 1993 (hep-th/9311025).

[54] C. Kiefer and T. Singh, *Phys. Rev.* **D44**, (1991) 1067.

[55] See for instance C. Misner, K. Thorne and J. A. Wheeler, *Gravitation*, W. H. Freeman, San Francisco (1971).

[56] C. G. Callan, S. B. Giddings, J. A. Harvey and A. Strominger, *Phys. Rev. D* **45** (1992) R1005. For reviews see J. A. Harvey and A. Strominger, *Quantum aspects of black holes*, in the proceedings of the 1992 TASI Summer School in Boulder, Colorado (World Scientific, 1993), and S. B. Giddings, *Toy models for black hole evaporation*, in the proceedings of the International Workshop of Theoretical Physics, 6th Session, June 1992, Erice, Italy, ed. V. Sanchez (World Scientific, 1993).

[57] G. 't Hooft, *Nucl. Phys.* **B335** (1990) 138, and references therein.

[58] L. Susskind, *Phys. Rev.* **D49** (1994) 6606.

[59] D. Louis-Martinez, J. Gegenberg and G. Kunstatter, *Phys. Lett.* **B321**(1994) 193

[60] T. Banks and M. O'Loughlin, *Nucl. Phys.* **B362** (1991) 649; A. Ashtekar and J. Samuel, *Bianchi cosmologies: the role of spatial topology*, Syracuse Report PRINT-91-0253 (1991); K. V. Kuchař, *Geometrodynamics of Schwarzschild black holes*, Utah Preprint UU-REL-94/3/1, (gr-qc/9403003); H. A. Kastrup and T. Thiemann, *Spherically Symmetric Gravity as a Completely Integrable System*, Preprint PITHA 93-35, (gr-qc/9401032).

[61] F. D. Mazzitelli and J. G. Russo, *Phys. Rev.* **D47** (1993) 4490.

[62] T. Regge and C. Teitelboim, *Ann. Phys.* **88** (1974) 286.

[63] G. Lifschytz and M. Ortiz, *Phys. Rev.* **D49** 1929 (1994).

[64] J. B. Hartle and S. W. Hawking *Phys. Rev.* **D13** 2188 (1976)

[65] M. Bañados, C. Teitelboim and J. Zanelli, Phys. Rev. Lett. **69** 1849 (1992); M. Bañados, M. Henneaux, C. Teitelboim and J. Zanelli, *Phys. Rev.* **D48** 1506 (1993).

[66] B. Reznik, *Phys. Rev.* **D45** 2151 (1992).

[67] D. Cangemi, M. Leblanc and R. B. Mann, MIT preprint CTP#2162, gr-qc/9211013; R. B. Mann and S. F. Ross, *Phys. Rev.* **D47** 3319 (1993); G. Clément, *Class. Qu. Grav.* **10** L49 (1993); G. Horowitz and D. Welch, *Phys. Rev. Lett.* **71** 328 (1993); N. Kaloper, Phys. Rev. **D48** 2598 (1993).

[68] S. J. Avis, C. J. Isham and D. Storey, *Phys. Rev.* **D18** 3565 (1978).

[69] S. A Fulling, *J. Phys.* **A10** 917 (1977).

[70] G. W. Gibbons and M. Perry, *Proc. R. Soc. London* **A358** 467 (1978).

[71] R. Kubo, *J. Phys. Soc. Jpn.* **12** 570 (1957); P. C. Martin and J. Schwinger, *Phys. Rev.* **115** 1342 (1959).

[72] A. Achúcarro and M. E. Ortiz, to appear in *Phys. Rev.* **D**; hep-th/9304068.

[73] S. W. Hawking and G. F. R. Ellis *The Large Scale Structure of Space Time*, Cambridge University Press, Cambridge, UK (1973).

[74] S. Takagi, *Prog. Theo. Phys.* **Supp. 88** (1986).

[75] R. C. Tolman, *Relativity, Thermodynamics and Cosmology*, Oxford, UK (1931).

[76] L. V. Ahlfors, *Complex Analysis*, McGraw-Hill, New York (1966).

[77] I. S. Gradshteyn and I. M. Ryzhik, *Tables of integrals, series and products*, Academic Press, San Diego (1965).

[78] S. Weinberg, *Gravitation and Cosmology*, Wiley, New York (1971), chapter 13.4.

[79] W. G. Unruh, *Phys. Rev.* **D14** 870 (1976).

[80] P. Candelas, *Phys. Rev.* **D21** 2185 (1980).

[81] G. W. Gibbons and S. W. Hawking, *Phys. Rev.* **D15** 2738 (1977).

[82] P. Hajicek, *Phys. Rev.* **D15** 2757 (1977).

[83] S. Deser and R. Jackiw, *Ann. Phys.* **153** 405 (1984).

[84] T. Souradeep and V. Sahni, *Phys. Rev.* **D46** 1616 (1993); H. H. Soleng, to appear in *Phys. Scripta,* hep-th/9310007.

[85] E. W. Hobson, *The theory of spherical and ellipsoidal harmonics*, Chelsy, New York (1955).

[86] G. 't Hooft, *Nucl. Phys.* **B256** (1985) 727.

[87] C. R. Stephens, G. 't Hooft and B. F. Whiting, *Class. Qu. Grav.* **11** (1994) 621.

[88] T. Jacobson, *Phys. Rev.* **D48** (1993) 728.

[89] D. N. Page, *Phys. Rev. Lett.* **44** (1980) 301.

[90] L. Susskind, L. Thorlacius and J. Uglum, *Phys. Rev.* **D48** (1993) 3743.

[91] E. Verlinde and H. Verlinde, *A Unitary S-matrix for 2D Black Hole Formation and Evaporation,* Princeton Preprint, PUPT-1380, IASSNS-HEP-93/8, hep-th/9302022 (1993), K. Schoutens, E. Verlinde, and H. Verlinde, *Phys. Rev.* **D48** (1993) 2670.

[92] S. R. Das and S. Mukherji, *Phys. Rev.* **D50** (1994) 930; A. Strominger and L. Thorlacius, *Conformally Invariant Boundary Conditions for Dilaton Gravity*, Univ. of California at Santa Barbara report No. NSF-ITP-94-34, (hep-th/9405084).

[93] A. Bilal and C. Callan, *Nucl. Phys.* **B394** (1993) 73; K. Nakamura, S. Konno, Y. Oshiro and A. Tomimatsu, *Prog. Th. Phys.* **90** (1993) 861; Y. Kazama and Y. Satoh, *Extraction of Black Hole Geometry in Exactly Quantized Two-Dimensional Dilaton Gravity*, Tokyo U. Report UT-Komoba-93-13 (hep-th/9310155); *Space-Time Geometry in Exactly Solvable Quantum Dilaton Gravity*, Tokyo U. Report UT-Komoba-93-13 (hep-th/9310155); *Phys. Rev.* **D48** (1993) 1687; S. P. de Alwis, *Phys. Lett.* **B317** (1993) 46; S. P. de Alwis and D. A. MacIntire, *The Problem of Time and Quantum Black Holes*, Colorado U. Report COLO-HEP-333 (gr-qc/9403052); C. Kiefer, R. Muller and T. P. Singh, *Quantum Gravity and Non-Unitarity in Black Hole Evaporation*, Zurich U. Report ZU-TH-25-93 (gr-qc/9308024); T. Hori, *Prog. Th. Phys.* **90** (1993) 743; T. Hori and M. Kamata, *Wave-Functional of Quantum Black Holes in Two-Dimensions*, Teikyo U. Report TEP-10 (hep-th/9302142).

[94] E. Keski-Vakkuri and S. D. Mathur, *Phys. Rev.* **D50** (1994) 917.

[95] T. M. Fiola, J. Preskill, A. Strominger and S. P. Trivedi, *Black Hole Thermodynamics and Information Loss in Two Dimensions*, Caltech report n.o. CALT-68-1918, (hep-th/9403137); C. Holzhey, F. Larsen and F. Wilczek, *Geometric and Renormalized Entropy in Conformal Field Theory*, Princeton Report No. PUPT-1454, IASSNS-HEP 93/88 (hep-th/9403108).

[96] W. H. Zurek, *Phys. Rev. Lett.* **49** (1982) 1683.

[97] S. D. Mathur, *Black Hole Entropy and the Semiclassical Approximation*, MIT report No. CTP-2304 (hep-th/9404135) (Invited Talk given at the International Colloquium on Modern Quantum Field Theory II at TIFR (Bombay), January 1994).

[98] E. Calzetta and F. D. Mazzitelli, *Phys. Rev.* **D42** (1990) 4066.

[99] J. P. Paz and S. Sinha, *Phys. Rev.* **D44** (1991) 1038.

[100] P. A. M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School Monograph 2, Yeshiva University, NY, 1964; R. Arnowitt, S. Deser, C. Misner, in *Gravitation*, ed. L. Witten, Wiley, NY (1962) p. 227; B. DeWitt, *Phys. Rev.* **160** (1967) 1113; C. Misner, K. Thorne, J. Wheeler, *Gravitation*, W. H. Freeman, San Francisco (1971); A. Hanson, T. Regge, C. Teitelboim, *Constrained Hamiltonian Systems*, Lincei Academy report No. 2 (1976); C. J. Isham, Lectures given at 1983 Les Houches School on Relativity Groups and Topology, June 27 - Aug. 4 1983; Proc. Roy. Soc. **A368** (1979) 33; M. Pilati, *Phys. Rev.* **D26** (1982) 2645.

[101] See for example S. B. Giddings, *Phys. Rev.* **D49** (1994) 4078.

[102] J. G. Russo, L. Susskind and L. Thorlacius, *Phys. Rev. D***46** (1992) 3444; *Phys. Rev.* **D47** (1993) 533.

[103] J. Preskill, *Do black holes destroy information?*, CalTech Report No. CALT-68-1819 (hep-th/9209058);

[104] S. B. Giddings and W. M. Nelson, *Phys. Rev.* **D46** (1992) 2486.

[105] K. V. Kuchař, *Phys. Rev.* **D39** (1989) 1579.

[106] L. D. Faddeev and A. A. Slavnov, *Gauge Fields: An Introduction to Quantum Theory*, Frontiers in Physics # 83, Addison-Wesley, Redwood City (1991).

[107] B. S. DeWitt, *Phys. Rep.* **19C** (1975) 297.