40

Topics in Semiclassical and Quantum Gravity

by

Esko Olavi Keski-Vakkuri

Filosofian Kandidaatti in Theoretical Physics and Mathematics
University of Helsinki
(1989)

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

May 1995

Signature of the Author ______________________

Department of Physics
May, 1995

Certified by ______________________

Professor Samir D. Mathur
Thesis Supervisor

Accepted by ______________________

Professor George F. Koster
Chairman, Graduate Committee

Topics in Semiclassical and Quantum Gravity

by

Esko Olavi Keski-Vakkuri

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

This thesis focuses in understanding various concepts and aspects related to the black hole information puzzle and in developing new ways to test the validity of the assumptions that are behind Hawking's original proposal. We start in chapter 2 with a brief discussion of particle production and entropy generation in scalar quantum field theory in expanding spacetimes with many-particle initial states. In chapter 3, we study the Hawking radiation for the geometry of an evaporating 1+1 dimensional black hole. We compute Bogoliubov coefficients and the stress tensor. We calculate the entropy of entanglement produced in the evaporation process, both for a 1+1 dimensional and 3+1 dimensional black hole. We present a straightforward computation through the density matrix of Hawking radiation. On the other hand, we use a recent result of Srednicki to estimate the entropy. It is found that the one space dimensional result of Srednicki is the pertinent one to use, in both the 1+1 and the 3+1 dimensional cases. In chapters 4 and 5, we investigate the validity of the semiclassical approximation in the black hole evaporation. First, we consider the definition of matter states on spacelike hypersurfaces. We take into account the quantum fluctuations in the black hole background spacetime and study their effect on the time evolution of matter states. We show that on any hypersurface that captures both infalling matter near the horizon and Hawking radiation, quantum fluctuations in the background become important. This suggests that we cannot describe the matter state by a semiclassical evolution up to this stage. We estimate that the correlations between the matter and gravity are so strong that a fluctuation of order $\exp(-M/M_{\text{Planck}})$ in the mass of the black hole produces a macroscopic change in the matter state. In chapter 5 we discuss how the existence of classical turning points can affect the validity of the semiclassical approximation. We show how turning points can appear in the evolution of a two dimensional black hole. We argue that turning points can create more complicated phase correlations than what can be seen in the leading order semiclassical approximation without back reaction. We demonstrate this in the context of simple quantum mechanical models.

However, we show that the effect is not present in a simple minisuperspace model of quantum matter in a closed universe.

Thesis supervisor: Dr. Samir D. Mathur
Title: Assistant Professor

Acknowledgements

One of the great pleasures in graduating is to have an opportunity to thank some of those who have helped me in the past about five years:

- First, I want to thank Samir Mathur, my advisor for the past three years. His cheerful attitude and love for physics was very infectious and ever so encouraging. I learned to respect very much his strong physics intuition and his way of striving to get to the root of a problem. He is always interested in discussing physics and delighted to find new connections between diverse ideas.

- I am also grateful to Gilad Lifschytz and Miguel Ortiz for our fruitful collaboration. This work is Chapter 4 of this thesis.

- Xiao-Gang Wen facilitated a diversion to the physics of fractional quantum Hall effect and co-authored my first research paper at MIT. I value very highly the additional opportunity to work with a superb physicist like him. I would also like to thank him for agreeing to be on my thesis committee on short notice and for all the other valuable help that he has provided.

- Roman Jackiw was the other member of my thesis committee. I thank him for that, and for the things that I learned from him both indirectly through his papers and directly from his graduate course.

- Barton Zwiebach supervised a reading project in my second semester, taught two excellent classes and helped me in other ways. Alan Guth, Ken Olum, and Sean Carroll listened to a mini seminar series and asked many insightful questions. Alan helped me in my postdoctoral job search.

- I thank my fellow CTP students for the many discussions and friendship. Gilad Lifschytz shared an interest in black holes and other things and I learned a great deal from him. My officemate Ranganathan Krishan and I spent many hours discussing physics and everything else; I thank him for his company. I thank Tom Wynter for his warm friendship which has lasted over his move to Paris. Dongsu Bak, Oren Bergman, Bhuvnesh Jain, many thanks for your pleasant company. I would really like to thank many more people separately but it would make a too long list.

- I also thank the CTP postdocs that I have met, Miguel, Sean, Mike Crescimanno, and others for discussions in and out of physics.

- The Center for Theoretical Physics and the Physics Department supported financially my studies, I am very grateful for this.

- I thank my mentors in Helsinki: Antti Kupiainen who introduced me to modern physics and Antti Niemi who was my advisor and led me to research. I express my deep gratitude to them. I also want to thank my former coworkers, teachers and fellow students in Helsinki.
- I also want to thank warmly Suomen Akatemia, Jenny ja Antti Wihurin rahasto and Magnus Ehrnroothin Säätiö for their partial support.
- I thank my personal friends Arto Anttila and Arto Tolvanen for their sharp wit and deeper perspective which brought much joy.
- Finally and most importantly, I thank my family: Irma, Simo and Pekka Keski-Vakkuri for all their love and support which has been a stable source of strength and encouragement throughout my life; and my fiancée Marja Bister for her love and radiant smile which made all bleaker moments disappear quickly. I dedicate this thesis to them.

If anyone wonders why,
after so many other histo-
ries have been written, I also
should have had the idea of
writing one, let him begin
by reading through all those
others, then turn to mine,
and after that he may wonder,
if he will.
FLAVIUS ARRIANOS
A.D. (95-180)

Contents

Chapter 1

Introduction and Overview

1.1 The Black Hole Information Problem

Every physics freshman has heard about black holes and has pondered upon the fate of an astronaut or a cosmonaut travelling into a black hole. The puzzling aspects of black holes are however far from being limited to the level of classical physics - even more tantalizing problems are encountered when one tries to combine black holes with quantum physics. In mid-seventies it was proposed by Stephen Hawking [1] that black holes formed in a gravitational collapse are not stable objects, but they begin to radiate thermally and lose their energy. Thus, first of all, black holes are not so black as they were thought to be. Further, he made the remarkable suggestion [2] that the black hole will evaporate completely and this process will not follow the rules of quantum mechanics. If this picture is correct, it means that the true laws of Nature should be based on some far deeper conceptual basis than quantum mechanics, on something new that we would not have even vague ideas about.

Hawking's theory can be described as follows. One can imagine that the matter which forms the black hole is in a pure quantum state $|\Psi\rangle$ (a superposition of s-waves, say). Initially the matter is very diffuse so that the spacetime is approximately flat. Later, as the black hole forms, an apparent horizon separating the interior of the black hole from the external world will also form. Other than that, the horizon is not really a very special place, for instance a part of this room could be inside a black hole horizon right now. Thus there is no reason why the total quantum state should not contain correlations which connect the both sides of the horizon. In other words, the quantum state is of the form

$$|\Psi\rangle = \sum_{iJ} c_{iJ}\ |\psi^i, \text{inside}\rangle \otimes |\chi^J, \text{outside}\rangle$$

where $|\psi^i, \text{inside}\rangle$ ($|\chi^J, \text{outside}\rangle$) represents states which are inside (outside) the horizon. Hawking showed that the inside (outside) states are particles with negative (positive) energy; the negative energy particles will be trapped into the black hole and decrease its energy and the positive energy particles will radiate out to infinity where the radiation turns out to be thermal with a temperature

$$T_H = \frac{\hbar c^3}{8\pi k_B G M} \, .$$

This temperature is called the Hawking temperature and the radiation is called Hawking radiation. In this formula, G is Newton's constant, k_B is Boltzmann's constant and M is the (instantaneous) mass of the black hole. For astrophysical black holes, the Hawking temperature is very small. For example, a solar mass black hole would have a Hawking temperature of the order $T_H \sim 10^{-7}K$.

At this point the total state $|\Psi\rangle$ is still a pure state. However, a crucial point of Hawking's argument is that there should be no good reason to expect the black hole to stop shrinking and eventually disappearing in a final explosive stage. After the black hole is gone, the states $|\psi^i, \text{inside}\rangle$ are gone with it and then also the correlations which linked them to the particles of the radiation. As a result, the final quantum state will be a *mixed* state described with a density matrix

$$\rho = \sum_{IJ} a_{IJ} \, |\chi^I, \text{outside}\rangle\langle\chi^J, \text{outside}| \, .$$

This would mean that a pure state has evolved into a mixed state, which is in contradiction with the unitary time evolution rule of quantum mechanics. The correlations lost along with the black hole represent fundamental information of the system which has disappeared. This is why Hawking's problem is often called the 'black hole information problem'.

1.2 Overview of the Thesis

This thesis work focuses in understanding various concepts and aspects related to the black hole information puzzle and in developing some new ways to test the validity of the assumptions that are behind Hawking's original proposal, and the assumptions behind the alternate viewpoint presented by t'Hooft [3].

Coarse grained entropy in an expanding universe

I start in **chapter 2** with a simple model of scalar field theory in an expanding universe and first review briefly the phenomena of spontaneous and stimulated particle production. This is a central feature of quantum field theory in a curved spacetime. In a flat Minkowski spacetime, when we define a vacuum, all inertial observers will agree with the definition. However, in a curved spacetime there is no unique choice for a class of observers with respect to whom to define what is meant by a vacuum. Therefore, we could have one natural definition of a vacuum in one region of spacetime, but a different one in some other region of the spacetime. The former vacuum state would then look like an exited state in the latter region. Hawking radiation is an example of this more general phenomenon. I also discuss the notion of entropy in an expanding universe. In an expanding universe the time evolution of matter states is naturally unitary. Therefore, strictly speaking there is no entropy generation either. More precisely, no *fine grained* entropy is generated. Usually in thermodynamics the notion of entropy means a *coarse grained* entropy. For example, the time evolution of a classical gas follows the classical equations of motion and according to the Liouville's theorem the phase space volume is conserved in the process. However, if the system is complex enough, in practise it is impossible to keep track of all the degrees of freedom and one is forced to adopt a coarse grained picture of the system. Then the volume of the phase space may appear to grow. This then leads to an apparent increase in the entropy of the system. Similar points of view may be taken in quantum field theory and there are various coarse graining scenarios with different physics motivations. I investigated a particular scenario proposed by Brandenberger, Mukhanov and Prokopec [4]. Their work considered a vacuum initial state. I found out that if their procedure is applied to more general many-particle (mixed) initial state, the amount of generated entropy depends non-trivially on the initial average number of particles per mode. In the case of bosons, the number of produced particles in an expanding universe becomes larger the more particles there was to begin with, but the generated entropy behaves in an opposite way and becomes less.

Fine grained entropy of the Hawking radiation

In **chapter 3**, we discuss the *fine grained* entropy S_{HR} of the Hawking radiation emerging from an evaporating black hole. Our main interest is to calculate S_{HR} using field theory techniques.

Recently there has appeared new string theory motivated models of two dimensional gravity. It was found that the new models allow black hole like solutions to the field equations and, coupled to quantum matter, they also Hawking radiate. These models became then very popular since they made it possible to investigate many features of black hole evaporation in a greatly simplified setting. Especially, it was possible to find a model which allowed an exact analytical treatment of a black hole which evaporates completely in a finite time. This model was introduced by Russo, Susskind and Thorlacius [4]. We study Hawking radiation in this model.

Our first route to calculate the entropy of the radiation is a straightforward one. We will first find the density matrix $\rho_{\mathbf{rad}}$ which describes the radiation. Then we will compute the entropy $S_{\mathrm{HR}} = -\mathrm{Tr}(\rho_{\mathbf{rad}} \ln \rho_{\mathbf{rad}})$.

In order to find the density matrix of the radiation, we need to investigate the Bogoliubov transformation which relates the observations of inertial observers in the far past and far future regions in the background spacetime. The Bogoliubov transformation encodes the structure of the Hawking radiation, as Hawking discovered in his original work [1]. In the case of two dimensional dilaton gravity black holes, the Bogoliubov transformation has been studied by Giddings and Nelson [6]. Their calculation ignored the backreaction of the Hawking radiation to the black hole geometry. In other words, they considered an eternal black hole which radiates infinitely. We repeate their analysis in the case where the backreaction has been included and the black hole evaporates in a finite time. *A priori* the Bogoliubov transformation could be quite different from the eternal black hole case. However, the differences turned out not to be very significant. After we have obtained the Bogoliubov transformation, we can deduce the form of the density matrix $\rho_{\mathbf{rad}}$ of the Hawking radiation and evaluate the entropy $S_{\mathbf{rad}}$.

Our second method of calculating the entropy is motivated by the results of Bombelli et. al. [7] and Srednicki [9] on the entanglement of two subregions of space in a vacuum state, in various spacetime dimensions. They considered a flat space free field theory in a vacuum state $|0\rangle$, and formed a reduced density matrix

$$\rho = \mathrm{Tr}_{\mathbf{inside}}\ |0\rangle\langle 0|$$

by taking a trace over the degrees of freedom inside a spherical region of space. In two and three space dimensions, the entropy $S = -\mathrm{Tr}_{\mathbf{outside}}(\rho \ln \rho)$ was found to depend on the radius R of the sphere as follows:

$$S \sim (\frac{R}{\varepsilon})^{d-1} ,$$

where d is dimensionality of the space. The three dimensional result is interesting since it depends quadratically on the radius of the sphere, like the Bekenstein-Hawking entropy for black holes. To get a finite result, it was necessary to introduce an ultraviolet cutoff ε, which could be interpreted *e.g.* as the radial thickness of the sphere. This entropy is attributed to the correlations between the inside and outside of the spherical region. Therefore the answer is independent of the order of the traces over 'inside' and 'outside' degrees of freedom. The leading contribution comes from short distance correlations, therefore the answer is ultraviolet divergent. In order to make contact with the Bekenstein-Hawking entropy, one would need to justify a specific choice for the UV cutoff.

In the one space dimensional case, the role of a 'sphere' is played by a segment of length R. In this case the entropy depends on the ratio of the length and the UV cutoff through a logarithm. Therefore the overall coefficient in front of the logarithm is meaningful. In the higher dimensional cases the overall coefficient depends on the specific regularization procedure.

We do not present an analytical derivation of the results of [7] and [9]. (They were originally obtained by using numerical considerations). Instead, we will present a heuristic derivation. Our main interest is to show how the one dimensional result of Srednicki can be used to calculate the entropy S_{HR} of the Hawking radiation. For an initial vacuum state, we compute the entanglement entropy of the subregion of past null infinity which contains the starting points of all rays which experience a redshift and give rise to the Hawking radiation. A natural cutoff scale is given by a relation to the characteristic wavelength of Hawking radiation. The resulting entanglement entropy is equal to the total fine grained entropy of the radiation. Perhaps more surprisingly, we show that the entropy of radiation from a *three* dimensional black hole can also be derived using the *one* dimensional Srednicki result. This is due to the fact that most of the Hawking radiation is in s-waves.

These two different routes give microscopic field theoretic computations of the entropy S_{HR}. It had been derived earlier by Zurek, using purely thermodynamic arguments [13]. The result shows that the entropy of the radiation is bigger than the Bekenstein-Hawking entropy of a black hole, a fact which is not always appreciated. The result should not be so surprising, since the black hole evaporates into vacuum and the process is not adiabatic. In fact, as we also discuss, the entropy of the radiation can be made arbitrarily large by feeding repeatedly matter into the hole, at the same time the Bekenstein-Hawking entropy never increases its initial value.

The results of Bombelli *et. al.* and Srednicki have been proved analytically by Holzhey [10], Callan and Wilczek [11] (who coined the name *geometric entropy* for this approach), Kabat and Strassler [12], Susskind and Uglum [13] and Fiola *et. al.* [14]. Fiola *et. al.* also considered the entropy of two dimensional evaporating black holes and found results similar to ours. It should also be mentioned that in a subsequent work Holzhey *et. al.* [15] investigated further the cutoff dependency of geometric entropy and introduced an elegant renormalization scheme to obtain a finite cutoff independent entropy. Our approach is similar in spirit to this and can be recast in their formalism.

The validity of the semiclassical approximation

In the remainder of the thesis, we move to a central issue in the black hole information problem. Hawking's theory is based on the assumption that the black hole evaporation process can be adequately studied in the framework of semiclassical gravity. Since for most of the time of evaporation the black hole remains a macroscopic object and the curvature of the gravitational field outside its horizon is negligible over planckian distances, there would seem to be no reason to expect quantum gravitational effects to play any role in the process. However, this view point has already for some time been challenged by 't Hooft [3], Page [8], Susskind [17] and others [18]; they have pointed out that superplanckian energy scales are important in the black hole problem because of the exponential redshift between asymptotic observers and the region close to the horizon which gives rise to the Hawking radiation. They have argued that the semiclassical approximation gives an insufficient picture of the black hole evolution.

One possible way to investigate the validity of the semiclassical approximation for is to study how it emerges from the Wheeler–de Witt equation of quantum gravity. In **chapter 4** we study if the semiclassical picture is consistent with quantum fluctuations in the background metric. We do this in the simplified context of two dimensional dilaton gravity. We consider the evolution of matter states on spacelike hypersurfaces (one-geometries) which provide a foliation of a 1+1 dimensional spacetime. From semiclassical physics one would expect the time evolution of matter states to be insensitive to Planck scale fluctuations in the background spacetime.

In order to describe both outgoing radiation quanta and infalling matter quanta, one needs to follow the evolution of matter up to hypersurfaces which traverse both trough the outgoing radiation and infalling matter, we shall call these surfaces S-surfaces. We find that in an evolution up these surfaces the matter states are very sensitive to the background fluctuations. We compute a natural inner product on a one-geometry between matter states which started out as vacuum states in backgrounds within a Planck scale fluctuation. As the states evolve up to S-surfaces, they become almost orthogonal[1]. This does not appear to be consistent with the spirit

[1]This is orthogonality is non-trivial, unlike the orthogonality between vacuua of fields of different mass or between a vacuum and a state with one low energy exitation.

of the semiclassical description[2]. Rather, this effect seems to be in the spirit of 't Hoofts's arguments [3], where he claims that large fluctuations appear in appropriate operator quantities, thus perhaps leading to quantum gravity effects in the black hole evaporation.

Does this mean that the semiclassical approximation is really insufficient to capture the physics of black hole evaporation? It is important to investigate if the effect which we found is spurious. Let us first recapitulate what is done in chapter 4. We are interested in matter states. Matter states are defined on spacelike hypersurfaces. The evolution of a state from one hypersurface to the next one gives a notion of time evolution. There are infinitely many hypersurfaces and correspondingly we could think of many different ways to do the time evolution of matter states. We would like to check if the quantum fluctuations in the background spacetime can affect the time evolution. In chapter 4 we find that this indeed can happen, even if we consider a time evolution in a regime where would expect the semiclassical description to be sufficient. This brings us to ask if the sensitivity to the background fluctuations is inevitable or incurable. We investigate this issue further in **chapter 5**.

First of all, it can be shown that in a black hole spacetime there are only two basic different categories of hypersurfaces [20]. One can consider a time evolution of a matter state which uses hypersurfaces which all belong into the same category. The time evolution which we consider in section 4.3.3 of chapter 4 is an example of such time evolution, and we find that in this case the matter state becomes increasingly sensitive to the background fluctuations as it evolves forward in time. It turns out that the two categories of hypersurfaces are related with each other by a time reversal symmetry [20]. Therefore, one can consider a backward time evolution of a matter state with hypersurfaces all in the other category (than above), and now one finds that the matter state becomes more sensitive to the background fluctuations as it evolves towards earlier times [20]. This motivates us to check what would happen in a time evolution which first uses hypersurfaces in the first category and then crosses over to the other category. Could it be that the matter state is sensitive to the background fluctuations only somewhere in the middle, but not any more at later times? Further, could the sensitivity in the middle leave any kind of an 'imprint' to the state which could still be detected at late times?

In chapter 5, we show that crossing over from one category of hypersurfaces to the other category creates a turning point in the evolution. Thus, we need to first

[2]For additional discussion, see [19].

investigate in general what kind of effects turning points can create in the semiclassical approximation. As we will discuss, the semiclassical approximation is not valid in the vicinity of a turning point. This is because the WKB approximation which is a part of the semiclassical approximation breaks down near a turning point. We will give estimates of the size of this region and also discuss when the region is big enough to be relevant at all. We also remind the reader that even if the WKB approximation is not valid at the turning point, it can again be used after the system has evolved sufficiently far from the turning point. In this case, one needs to scrutinize the potential tunneling issues and join the two WKB solutions at the turning point in an appropriate way. In the same manner, even if the semiclassical approximation breaks down at the turning point, we need to study if it becomes again applicable after the turning point.

We will first discuss simple quantum mechanical models of a light particle coupled to a heavy particle. The light particle is the analogue of quantum matter and the heavy particle is the analogue of gravitational degrees of freedom. In this context, using an exact quantum mechanical solution, we will demonstrate that a turning point (in the motion of the heavy particle) can indeed leave a permanent imprint into the total wavefunction which survives until the end of the time evolution. This imprint would be missed in the leading order semiclassical calculation. We then move on to investigate a simple minisuperspace model of quantum matter propagating in a closed radiation dominated Robertson-Walker cosmology. In this case, there is a turning point at the point when the universe has reached its maximum size and begins to recollapse. We study the imprint of the turning point into the state, and find that soon after the turning point the imprint will become insignificant. Thus, in this case the time evolution of states is sufficiently described with a leading order semiclassical approximation.

Finally, we study dilaton gravity black holes and show how the turning points appear in crossing from one category of hypersurfaces to another. We would like to investigate if the turning point can cause lasting effects to the states, or if these effects decay away as they do in cosmology. This work is not yet done. So far we have studied the tunneling issues at the turning point, we will discuss some preliminary results.

Bibliography

[1] S. W. Hawking, *Comm. Math. Phys.* **43** (1975) 199.

[2] S. W. Hawking, *Phys. Rev.* **D14** (1976) 2460.

[3] G. t'Hooft, *Nucl. Phys. B256* (1985) 727; *Nucl. Phys.* (1990) 138, and references therein; C. R. Stephens, G. 't Hooft and B. F. Whiting, *Class. Qu. Grav.* **11** (1994) 621.

[4] R. Brandenberger, V. Mukhanov and T. Prokopec, *Phys. Rev.* **D48** (1993) 2443; Phys. Rev. Lett. **69** (1992) 3606; T. Prokopec, *Class. Quant. Grav.* **10** (1993) 2295.

[5] J. G. Russo, L. Susskind and L. Thorlacius, *Phys. Rev.* **D46** (1992) 3444; *Phys. Rev.* **D47** (1993) 533.

[6] S. B. Giddings and W. M. Nelson, *Phys. Rev.* **D46** (1992) 2486.

[7] L. Bombelli, R. K. Koul, J. Lee and R. D. Sorkin, *Phys. Rev.* **D34** (1986) 373.

[8] M. Srednicki, *Phys. Rev. Lett.* **71** (1993) 666.

[9] W. H. Zurek, *Phys. Rev. Lett.* **49** (1982) 1683.

[10] C. Holzhey, Ph.D. Thesis, Princeton University.

[11] C. Callan and F. Wilczek, *Phys. Lett.* **B333** (1994) 55.

[12] D. Kabat and M. Strassler, *Phys. Lett.* **B329** (1994) 46.

[13] L. Susskind and J. Uglum, *Phys. Rev.* **D50** (1994) 2700.

[14] T. Fiola, J. Preskill, A. Strominger and S. Trivedi, *Phys. Rev.* **D50** 3987.

[15] C. Holzhey, F. Larsen and F. Wilczek, *Nucl. Phys.* **B424** (1994) 443.

[16] D. N. Page, *Phys. Rev. Lett.* **44** (1980) 301.

[17] L. Susskind, L. Thorlacius and J. Uglum, *Phys. Rev.* **D48** (1993) 3743; L. Susskind, *Phys. Rev.* **D49** (1994) 6606.

[18] see eg. T. Jacobson, *Phys. Rev.* **D48** (1993) 728, 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; S. R. Das and S. Mukherji, *Phys. Rev.* **D50** (1994) 930.

[19] G. Lifschytz, S. D. Mathur, and M. Ortiz, *A Note on the Semi-Classical Approximation in Quantum Gravity,* MIT Report No. MIT-CTP-2384 (gr-qc/9412040).

[20] E. Keski-Vakkuri and S. D. Mathur, unpublished

Chapter 2

Coarse-grained Entropy and Stimulated Emission in Curved Space

One of the interesting features of quantized fields in a curved spacetime [1] is that the concept of particles becomes very observer-dependent. For instance, in an expanding Universe spontaneous particle creation can occur. One defines generally a vacuum state such that all inertial observers in the past region agree that the spacetime looks empty of particles. As a result of the expansion of the Universe, the above vacuum state looks full of particles using modes natural to inertial observers in the far future region. Stated differently, a no-particle initial state can evolve to a many-particle state. However, since one starts with a pure state, one ends with a pure state. Thus there must be subtle correlations between the particles in the final state. In particular, there is no entropy production in this process even if lots of particles are produced. But, it may be that some of these subtle correlations are very difficult to detect and/or that they may be quite sensitive to interactions between the produced particles. One may then consider such information about the system to be "less relevant" and either discard it altogether or apply some kind of a "statistical averaging" procedure to it. This way one can try to associate a "coarse-grained" entropy to the final state of the system, hopefully in as natural way as possible. There has been a lot of work in this direction by Hu, Kandrup and collaborators [2].

Recently, novel such approaches have been proposed. Brandenberger, Mukhanov and Prokopec (BMP) discussed in [3, 4] among other issues a coarse-graining procedure based on averaging over the so called squeeze angles which appear in the S-matrix of particle production. On the other hand, Gasperini and Giovannini [5], together

This chapter is based on work which has appeared in *Physical Review* **D49** (1994) 2122.

with Veneziano (GGV) [6] related entropy generation to an increased dispersion of a superfluctuant operator. Both groups were especially interested in the entropy generation related to the production of gravitational waves and density fluctuations in inflationary universe models.

In this chapter, we study the coarse-graining procedure based on averaging over the squeeze angles, which we shall call the BMP approach. We investigate the entropy generation starting not from an initial vacuum state with zero entropy, but allowing the system to be initially in some generic many-particle (mixed) state with non-zero entropy. If one starts with many bosons it is known [7] that the particle production will be amplified as a result of boson statistics, as one would expect. So, in general one can ask whether the entropy generation (in the coarse-grained sense) would also be amplified or not. Indeed, as a consistency check it is necessary to investigate if definitions of coarse-grained entropy will lead to a growing entropy even if initial state is allowed to be an arbitrary many-particle state with initial entropy. In [5, 6] the GGV entropy was shown to be growing at least in certain classes of initial states. Interestingly, it was found that their entropy generation did not depend at all on the number of particles or entropy of the initial state. Here we will attempt to investigate the BMP entropy in similar situations. At least in the case of an initial density matrix where particles appear as pairs of opposite momenta, and initial entropy depends on the average occupation number per mode, we can show that the BMP entropy grows, though the entropy generation is *attenuated.* The BMP entropy *does* depend on the initial number of particles in a non-trivial way. In the end we comment briefly on the case of an initial thermal density matrix.

A scalar field in a D-dimensional curved spacetime is described by an action

$$S = \int d^D x \sqrt{-g}\, \frac{1}{2}[g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - (m^2 + \xi R(x))\phi^2], \tag{2.1}$$

where $R(x)$ is the Ricci scalar curvature of the metric and ξ is a coupling constant. Assume that the metric depends explicitly on time and that it is asymptotically flat in the far past and far future : $g_{\mu\nu}(\vec{x}, t) \to C_\pm \eta_{\mu\nu}$, as $t \to \pm\infty$. In this case there are two natural ways to quantize the field ϕ in the Heisenberg picture [1]. One can either use modes which look like plane waves in the far past region, or modes which look like plane waves in the far future region, respectively. One then associates two sets of annihilation/creation operators to these modes, the 'in' and 'out' operators. These in turn define two vacua, one for the 'in' annihilation operators and one for the 'out' operators.

The 'in' and 'out' modes can be related via a Bogoliubov transformation, which can be given in terms of annihilation/creation operators as

$$a_{\vec{k}}^{in} = \alpha_{\vec{k}\vec{p}}^{*} a_{\vec{p}}^{out} - \beta_{\vec{k}\vec{p}}^{*} a_{\vec{p}}^{\dagger out} . \tag{2.2}$$

This transformation is generated by a S-matrix

$$a_{\vec{k}}^{in} = \mathbf{S} a_{\vec{k}}^{out} \mathbf{S}^{-1} , \tag{2.3}$$

which has the explicit form [8]

$$\begin{aligned} \mathbf{S} &= \frac{1}{\sqrt{\det(\alpha_{\vec{k}\vec{p}})}} : \exp\{\frac{1}{2}[\alpha^{-1}\beta^{*}]_{\vec{k}\vec{k}'} a_{\vec{k}}^{\dagger out} a_{\vec{k}'}^{\dagger out} \\ &\qquad + [\alpha^{-1} - 1]_{\vec{k}\vec{k}'} a_{\vec{k}}^{\dagger out} a_{\vec{k}'}^{out} - \frac{1}{2}[\alpha^{-1}\beta]_{\vec{k}\vec{k}'} a_{\vec{k}}^{out} a_{\vec{k}'}^{out}\} : \ . \end{aligned} \tag{2.4}$$

The factor $\frac{1}{\sqrt{\det(\alpha_{\vec{k}\vec{p}})}}$ is the in-out vacuum amplitude. We use the convention of [9] where the coefficients α have taken to be real. The S-matrix is known to generate a unitary transformation between the 'in' and 'out' representations if the gravitational field has a compact support [10]. For Robertson-Walker type universes the in-out vacuum amplitude is zero and the 'in' and 'out' representations are thus unitarily inequivalent.

The S-matrix relates the in- and out-vacuum states in the following way

$$\begin{aligned} \mid 0, in\rangle &= \mathbf{S} \mid 0, out\rangle \\ &= \frac{1}{\sqrt{\det(\alpha)}} \exp\{\frac{1}{2}[\alpha^{-1}\beta]_{\vec{k}\vec{k}'} a_{\vec{k}}^{\dagger out} a_{\vec{k}'}^{\dagger out}\} \mid 0, out\rangle . \end{aligned} \tag{2.5}$$

This is the statement that an inertial observer in the far future region sees the in-vacuum state as full of out-particles. Similarly, the density matrix of the system expanded using in-modes ($\equiv \rho_i$) can be related to an expression using out-modes ($\equiv \rho_f$) as follows

$$\begin{aligned} \rho_i &= \prod_{\vec{k}} \sum_{n_{\vec{k}}=0}^{\infty} f(n_{\vec{k}}) \mid n_{\vec{k}}, in\rangle\langle in, n_{\vec{k}} \mid \\ &= \prod_{\vec{k}} \sum_{n_{\vec{k}}=0}^{\infty} f(n_{\vec{k}}) \mathbf{S} \mid n_{\vec{k}}, out\rangle\langle out, n_{\vec{k}} \mid \mathbf{S}^{-1} \ \equiv \ \mathbf{S} \rho_f \mathbf{S}^{-1} , \end{aligned} \tag{2.6}$$

using (2.3) and (2.5). Suppose now that the system is initially in an arbitrary many-particle state. In this state the average occupation number per mode (using in-modes) is given by

$$\bar{n}_{\vec{k}}^{i} = \frac{1}{Tr\rho_i} Tr(\rho_i a_{\vec{k}}^{\dagger in} a_{\vec{k}}^{in}) . \tag{2.7}$$

In the far future region an inertial observer sees the average occupation number per mode using out-modes as

$$\bar{n}^f_{\vec{k}} = \frac{1}{Tr\rho_f} Tr(\mathbf{S}\rho_f\mathbf{S}^{-1} a^{\dagger out}_{\vec{k}} a^{out}_{\vec{k}}) \tag{2.8}$$

Using the cyclicity of the trace and the properties of the Bogoliubov transformation one can derive the relation between $\bar{n}^f_{\vec{k}}$ and $\bar{n}^i_{\vec{k}}$ to be

$$\bar{n}^f_{\vec{k}} = \mid \alpha_{\vec{k}\vec{p}} \mid^2 \bar{n}^i_{\vec{p}} + \mid \beta_{\vec{k}\vec{p}} \mid^2 (1 + \bar{n}^i_{\vec{p}}) \ . \tag{2.9}$$

This is the formula for "stimulated emission" [7]. It tells us that even if the spontaneous creation of particles is weak, $\mid \beta_{\vec{k}\vec{p}} \mid^2 \ll 1$, the particle production $\bar{n}^f_{\vec{k}} - \bar{n}^i_{\vec{k}}$ can become arbitrarily large, if the initial average occupation number per mode is arbitrarily large. This amplification of particle production is a result of the boson statistics of the particles. For fermions the particle production would be attenuated [11].

Let us now discuss for simplicity metrics of the form $ds^2 = dt^2 - a^2(t)d\vec{x}^2$, where $a(t)$ is a scale parameter of the universe. We again just require that $a(t) \to a_\pm$ asymptotically as $t \to \pm\infty$. As a result of the invariance under spatial translations, the Bogoliubov coefficents can be written as

$$\begin{aligned} \alpha_{\vec{k}\vec{p}} &= \alpha_{\vec{k}}\delta_{\vec{k},\vec{p}} \equiv \cosh r_{\vec{k}}\delta_{\vec{k},\vec{p}} \\ \beta_{\vec{k}\vec{p}} &= \beta_{\vec{k}}\delta_{\vec{k},-\vec{p}} \equiv e^{i\theta_{\vec{k}}} \sinh r_{\vec{k}}\delta_{\vec{k},-\vec{p}} \ . \end{aligned} \tag{2.10}$$

The parameters $r_{\vec{k}}$, $\theta_{\vec{k}}$ are called squeeze parameter and squeeze angle in the Quantum Optics literature, and the S-matrix is called a two-mode squeeze operator [12]. If one starts with a vacuum state, the final state (2.5) is called a squeezed vacuum. In the initial vacuum case, if one expands the corresponding $\mathbf{S}\rho_f\mathbf{S}^{-1}$ in the 'out' basis of energy eigenstates, one finds [4] that the off-diagonal components of $\mathbf{S}\rho_f\mathbf{S}^{-1}$ have an oscillatory dependence of the angles θ_k. In the BMP coarse-grained entropy approach it is assumed that these angles represent irrelevant information about the system (*e.g.,* in the sense that they would be very difficult to measure) and they are therefore averaged over. After the averaging only the diagonal elements of $\mathbf{S}\rho_f\mathbf{S}^{-1}$ survive and one then defines a coarse-grained entropy with the resulting reduced density matrix ρ_{red} with the usual formula $S = -k_B Tr(\rho_{red} \log \rho_{red})$. The result is [3, 4]

$$\begin{aligned} S^{f,0} \equiv \sum_{\vec{k}} s^{f,0}_{\vec{k}} &\equiv k_B \sum_k (\cosh^2 r_{\vec{k}} \log \cosh^2 r_{\vec{k}} - \sinh^2 r_{\vec{k}} \log \sinh^2 r_{\vec{k}}) \\ &= k_B \sum_{\vec{k}} [(\bar{n}^{f,0}_{\vec{k}} + 1)\log(\bar{n}^{f,0}_{\vec{k}} + 1) - (\bar{n}^{f,0}_{\vec{k}})\log(\bar{n}^{f,0}_{\vec{k}})] \ , \end{aligned} \tag{2.11}$$

where the notation $\bar{n}_{\vec{k}}^{f,0}$ means the LHS of (2.9) in the case of an initial vacuum state. Now we turn to consider initial density matrices ρ_i which can describe generic many-particle states with non-zero entropy. Let us assume that the initial density matrix has the form

$$\rho_i = \prod_{\vec{k},(k_z>0)} \sum_{n=0}^{\infty} f_{\vec{k}}(n) \mid n_{\vec{k}}, n_{-\vec{k}}, in\rangle\langle in, n_{\vec{k}}, n_{-\vec{k}} \mid \ , \tag{2.12}$$

where the coefficients $f_{\vec{k}}(n)$ are of the form $f_{\vec{k}}(n) = (\bar{n}_{\vec{k}}^i)^n/(\bar{n}_{\vec{k}}^i + 1)^{n+1}$. That is, we start with a many-particle state where particles appear in pairs of opposite momenta, with an initial average occupation number spectrum $\bar{n}_{\vec{k}}^i = \bar{n}^i_{-\vec{k}}$ and with (ordinary) entropy given by $-k_B Tr(\rho_i \log \rho_i)$. Writing the initial entropy in more explicit form, it is

$$S^i \equiv \sum_{\vec{k},(k_z>0)} s_{\vec{k}}^i = k_B \sum_{\vec{k},(k_z>0)} [(\bar{n}_{\vec{k}}^i + 1)\log(\bar{n}_{\vec{k}}^i + 1) - (\bar{n}_{\vec{k}}^i)\log(\bar{n}_{\vec{k}}^i)] \ . \tag{2.13}$$

When we expand the resulting final density matrix $\mathbf{S}\rho_f\mathbf{S}^{-1}$ (2.7) in 'out' energy eigenstates, we find that also in this case the off-diagonal elements have an oscillatory dependence on the angles $\theta_{\vec{k}}$. Therefore, following the BMP approach and averaging over the angles, only the diagonal elements will survive. Thus the reduced density matrix of the final state has the form

$$\rho_{red} = \prod_{\vec{k},(k_z>0)} \sum_{n=0}^{\infty} \bar{f}_{\vec{k}}(n) \mid n_{\vec{k}}, n_{-\vec{k}}, out\rangle\langle out, n_{\vec{k}}, n_{-\vec{k}} \mid \ , \tag{2.14}$$

where $\bar{f}_{\vec{k}}(n) = \langle n_{\vec{k}}, n_{-\vec{k}}, out \mid \mathbf{S}\rho_f\mathbf{S}^{-1} \mid out, n_{\vec{k}}, n_{-\vec{k}}\rangle$. After some effort, one can show that the coefficients have the form

$$\bar{f}_{\vec{k}}(n) = \frac{(\bar{n}_{\vec{k}}^f)^n}{(\bar{n}_{\vec{k}}^f + 1)^{n+1}} \ , \tag{2.15}$$

where $\bar{n}_{\vec{k}}^f$ is the LHS of (2.9) with the $\bar{n}_{\vec{k}}^i$ of (2.12). Thus, the final coarse-grained entropy is

$$S^f \equiv \sum_{\vec{k},(k_z>0)} s_{\vec{k}}^f = k_B \sum_{\vec{k},(k_z>0)} [(\bar{n}_{\vec{k}}^f + 1)\log(\bar{n}_{\vec{k}}^f + 1) - (\bar{n}_{\vec{k}}^f)\log(\bar{n}_{\vec{k}}^f)] \ . \tag{2.16}$$

The entropy depends only on the occupation number spectrum of particles in the final state. This result is in agreement with a similar formula given in [3] by a more heuristic argument to define entropy of a statistical system with a definite spectrum which is valid both in and far out of thermodynamical equilibrium. Let us now compare the entropy generation per mode in the initial vacuum and initial many-particle cases. Denote

$$\Delta_0 s_{\vec{k}} \equiv s_{\vec{k}}^{f,0} - 0 \tag{2.17}$$

$$\Delta s_{\vec{k}} \equiv s_{\vec{k}}^{f} - s_{\vec{k}}^{i} , \tag{2.18}$$

where (2.17) applies to the former case and (2.18) to the latter case. As a first consistency check, we find that $\Delta s_{\vec{k}} \geq 0$, so the coarse-graining led to a growing entropy in our many-particle case. However, as we compare (2.18) and (2.17) we find that

$$\Delta s_{\vec{k}} \leq \Delta_0 s_{\vec{k}} ; \tag{2.19}$$

i.e., the entropy generation is *attenuated* if one starts with many particles present in the mode $\vec{k}$. The equality holds iff $\bar{n}_{\vec{k}}^{i} = 0$. This result is easiest to see in the following way. Consider the difference $\Delta s_{\vec{k}} - \Delta_0 s_{\vec{k}}$. Substitute (2.16) and (2.13) to (2.18), and (2.11) to (2.17). Then substitute $\bar{n}_{\vec{k}}^{f}$ as a function of $\bar{n}_{\vec{k}}^{i}$ and $\mid \beta_{\vec{k}} \mid^2 \equiv \sinh^2 r_{\vec{k}}$ by using (2.9) and (2.10). The difference $\Delta s_{\vec{k}} - \Delta_0 s_{\vec{k}}$ depends then symmetrically on $\bar{n}_{\vec{k}}^{i}$ and $\mid \beta_{\vec{k}} \mid^2$. By drawing a 3d plot one can see that it is always non-positive and it decreases monotonically as either variable increases. Further, as both variables approach infinity,

$$\Delta s_{\vec{k}} - \Delta_0 s_{\vec{k}} \ \rightarrow \ \log 2 - 1 \approx -0.31 \tag{2.20}$$

asymptotically. This finite value is the maximum difference between the generated entropies per mode. Thus, unlike the GGV entropy, the BMP entropy generation is *not* independent of the number of particles in the initial state, but has some 'memory' about the initial occupation numbers. Since entropy is a measure of loss of information, it would appear that more information about the initial state of the system is conveyed to the final coarse-grained state when stimulated emission dominates the spontaneous particle production (since the entropy generation is attenuated).

The next case to be investigated would be an initial thermal density matrix

$$\rho_i = \prod_{\vec{k}} Z_{\vec{k}}^{-1} \exp(-\beta \omega_{\vec{k}}^{in} a_{\vec{k}}^{\dagger in} a_{\vec{k}}^{in}) \ .$$

This situation is somewhat trickier to deal with, for the following reason. Initially, the particles of opposite momenta are uncorrelated. However, in the expansion of the universe the particles are produced in pairs of opposite momenta. This induces correlations between the opposite momenta in the final density matrix. It would have the form

$$\mathbf{S} \rho_f \mathbf{S}^{-1} = \prod_{\vec{k},(k_z>0)} \sum_{n,m,n',m'=0}^{\infty} f_{\vec{k}}(n,m;n',m') \mid n_{\vec{k}}, m_{-\vec{k}}, out\rangle\langle out, n'_{\vec{k}}, m'_{-\vec{k}} \mid \ , \tag{2.21}$$

where

$$f_{\vec{k}}(n,m;n',m') = \tag{2.22}$$
$$\langle n_{\vec{k}}, m_{-\vec{k}}, out \mid \frac{1}{Z_{\vec{k}}Z_{-\vec{k}}} e^{-\beta\omega^{in}_{\vec{k}}[(\alpha_{\vec{k}}a^{\dagger out}_{\vec{k}} - \beta_{\vec{k}}a^{out}_{-\vec{k}})(\alpha_{\vec{k}}a^{out}_{\vec{k}} - \beta^*_{\vec{k}}a^{\dagger out}_{-\vec{k}}) + (\vec{k} \mapsto -\vec{k})]} \mid out, n'_{\vec{k}}, m'_{-\vec{k}}\rangle \ .$$

Again, one can see that the $n \neq n'$ or $m \neq m'$ components have an oscillatory dependence of the squeeze angles and they vanish in the coarse graining. However, the diagonal coefficients (those of the reduced density matrix) will have a form $f_{\vec{k}}(n,m)$ where the dependence on n and m does not factorize. Hence the opposite momenta have acquired correlations through the particle production and the reduced density matrix is not of the same type as the initial one. As advocated in [6], one would like to ignore the correlations between different modes. Further, one should not do this by replacing the two-mode squeeze operator by a one-mode squeeze operator, since the particles are then not created in the correct way as pairs of opposite momenta. We would like to propose that the correlations between opposite momenta could be ignored by proceeding to define $\bar{f}_{\vec{k}}(n) = \sum_m f_{\vec{k}}(n,m)$ and $\bar{f}_{-\vec{k}}(m) = \sum_n f_{\vec{k}}(n,m)$. Then we would define the final reduced density matrix to be

$$\rho_{red} = \prod_{\vec{k},(k_z>0)} \sum_{n,m=0}^{\infty} \bar{f}_{\vec{k}}(n)\bar{f}_{-\vec{k}}(m) \mid n_{\vec{k}}, m_{-\vec{k}}, out\rangle\langle out, n_{\vec{k}}, m_{-\vec{k}} \mid \ , \tag{2.23}$$

which is of the same type as the initial density matrix. Now the final entropy would be given by

$$S^f = -k_B \sum_{\vec{k}} \sum_n \bar{f}_{\vec{k}}(n) \log \bar{f}_{\vec{k}}(n) \ . \tag{2.24}$$

Unfortunately, at the present we do not have explicit formulas for the coefficients $\bar{f}_{\vec{k}}(n)$ or the final entropy. It would be very interesting to see if the resulting expressions could depend on the final average occupation number in the same fashion as in the earlier case. We hope to be able to return to this question in the future.

Finally, let us clarify that even if we found a different result as in the GGV approach, that the entropy generation depends on the number of particles in the initial state, we are not arguing that it would mean that the BMP approach is 'better' than the GGV approach. As stated in [2], it is good to have different definitions of entropy, corresponding to loss of different information about the system. Both BMP and GGV approaches have the virtue of giving the correct average occupation number for particles in the final state. Otherwise the GGV approach appears to discard information about the system a bit more generously, since it leads to a greater growth of entropy.

Acknowledgements

It is a pleasure to thank prof. S.D. Mathur for guidance and numerous discussions. I am also grateful to profs. R. Brandenberger, A.J. Niemi and S. Stenholm for helpful discussions.

Bibliography

[1] N.D. Birrell and P.C.W. Davies, *Quantum Fields in Curved Space*, Cambridge Univ. Press (1982)

[2] B. Hu and D. Pavon, *Phys. Lett.* **B180** (1986) 329; H. Kandrup, *J. Math. Phys.* **28** (1987) 1398; *Phys. Rev.* **D37** (1988) 3505; B. Hu and H. Kandrup, *Phys. Rev.* **D35** (1987) 1776; S. Habib and H. Kandrup, *Ann. Phys.* **191** (1989) 335.

[3] R. Brandenberger, V. Mukhanov and T. Prokopec, *Phys. Rev.* **D48** (1993) 2443; *Phys. Rev. Lett.* **69** (1992) 3606.

[4] T. Prokopec, *Class. Quant. Grav.* **10** (1993) 2295.

[5] M. Gasperini and M. Giovannini, *Phys. Lett.* **B301** (1993) 334; *Class. Quant. Grav.* **10** (1993) L133.

[6] M. Gasperini, M. Giovannini and G. Veneziano, *Phys. Rev.* **D48** (1993) R439.

[7] L. Parker, *Phys. Rev.* **183** (1969) 1057; R.M. Wald, *Phys. Rev.* **D13** (1976) 3176; J. Audretsch, in *Quantum Mechanics in Curved Space-Time*, ed. J. Audretsch and V. de Sabbata, Plenum Press (1990).

[8] The explicit formula for the S-matrix can be found (in the case of a RW-universe) *e.g.* in J. Ambjørn, R.J. Hughes and N.K. Nielsen, *Ann. Phys.* **150** (1983) 92.

[9] B. DeWitt, *Phys. Rep.* **19** (1975) 297.

[10] R.M. Wald, *Ann. Phys.* **118** (1979) 490.

[11] L. Parker, *Phys. Rev.* **D3** (1971) 346.

[12] B. Schumaker, *Phys. Rep.* **135** (1986) 317. The relations between entropy, correlations and squeezing have also been studied in Quantum Optics, see *e.g.* S.M. Barnett and S.J.D. Phoenix, *Phys. Rev.* **A40** (1989) 2404; S. Abe, *Phys. Lett.* **A166** (1992) 163.

Chapter 3

Evaporating Black Holes and Entropy

3.1 Introduction

Several interesting phenomena are related to the discovery of Hawking radiation [1]. It is intriguing that black holes seem to obey laws of thermodynamics [2]. The information contained in the matter which made up the black hole is lost into the singularity. Hawking radiation appears in the evaporation of the hole, but the outgoing modes are not in a pure state; instead they are mixed with modes of the field that fall into the singularity. The precise significance of black hole thermodynamics, and its relation to the ordinary ideas of thermodynamics and information theory, are matters of debate.

Recently the discovery of 1+1 dimensional models for black holes [3, 4] has led to a more accurate understanding of the semiclassical features of black hole geometry and Hawking radiation. In particular the model of [4] (the RST model) can be exactly solved to yield the semiclassical geometry of a black hole formed by a shock wave of infalling matter, and evaporating by massless scalars to an endpoint (the 'thunderpop'). It may even be possible to obtain a complete quantum gravity plus matter description of the black hole evaporation process in 1+1 dimensions [5].

In this chapter we study some features of the semiclassical geometry and Hawking radiation in semiclassical models. For the RST model of the evaporating hole we compute the Bogoliubov coefficients. We perform a point splitting calculation to compute the stress tensor at $\mathcal{I}^+$. We also compute the stress tensor in the evaporating region using the anomalous trace of the matter stress tensor. The RST model is solved

This chapter is based on work in collaboration with Samir D. Mathur which has appeared in *Physical Review* **D50** (1994) 917.

also for the case where the hole is formed by one shock wave, evaporation occurs for a time, and then a second shock wave increases the mass of the hole again. (This geometry is used for clarifying some aspects of the entropy produced by the hole, as discussed below.)

We then study the 'entropy of entanglement' of the Hawking radiation, by two methods. We can compute the density matrix obtained by tracing the field modes inside the horizon. This was the approach taken by Hawking and also carried out in [6] for the 1+1 dimensional case, and it yields a density matrix that is close to thermal after the initial stage of the collapse and formation of the hole. We consider the case where the hole has a finite lifetime (due to the evaporation) and thereby estimate the entropy of the entire radiation produced. We then compare this result to that obtained by using a calculation first studied by Bombelli *et. al.* [7] and more recently by Srednicki [8]. Consider a scalar field in flat Minkowski space in the vacuum state, and 'trace out' the degrees of freedom inside a ball of radius R. The entropy of the reduced density matrix is the 'entropy of entanglement' between the region inside the ball and its complement. The entropy S depends on R and also on the ultraviolet cutoff, which gives the 'sharpness of separation' between the regions. In the one-space-dimension case, the infrared cutoff also appears. We find that both for the 1+1 dimensional black hole and for the 3+1 dimensional hole the one space dimension result of Srednicki is the pertinent one to use, and the leading dependence of the entropy on the black hole mass is reproduced.

The plan of this chapter is as follows. In section 3.2 the RST model is reviewed. In section 3.3 the Bogoliubov coefficients for a scalar field in the evaporating black hole background are computed. Section 3.4 contains the calculation of stress-tensor. Section 3.5 studies the two shock wave solution. We discuss the entropy of the Hawking radiation for 1+1 dimensional black holes in section 3.6, and for 3+1 dimensional black holes in section 3.7. Finally, a discussion is presented in section 3.8.

3.2 The RST model

The model of Russo, Susskind and Thorlacius (RST) [4] is a modified version of the model of two dimensional dilaton gravity coupled to quantum matter introduced in [3]. A key idea of RST was to introduce an additional counterterm which restores a global symmetry originally present in the classical dilaton gravity + matter action. This allowed them to solve the theory analytically in the large N limit. The properties of the RST model have been extensively discussed in the literature [4, 9, 10] so we will just mention the facts we will need for later use.

The semiclassical effective action of RST can be written as follows

$$\begin{aligned} S &= \frac{1}{2\pi}\int d^2x\sqrt{-g}\{(e^{-2\phi}-\frac{N}{24}\phi)R+4e^{-2\phi}[(\nabla\phi)^2+\lambda^2]-\frac{1}{2}\sum_{i=1}^{N}(\nabla f_i)^2\} \quad (3.1) \\ &- \frac{N}{96\pi}\int d^2x\sqrt{-g(x)}\int d^2x'\sqrt{-g(x')}R(x)G(x,x')R(x')\ , \end{aligned}$$

where R is the 1+1 dimensional scalar curvature, ϕ is the dilaton field and f_i, $i = 1\ldots, N$ are N massless conformal scalar fields. $G(x,x')$ is the Green's function for the d'Alembertian in curved space. The constant λ plays the role of Planck mass.

The analysis of the semiclassical equations of motion that follow from the action (3.2) can be simplified by the following two steps. First, one can write the metric in the conformal gauge, given by

$$g_{\pm\mp} = -\frac{1}{2}e^{2\rho}\ ,\ g_{\pm\pm} = 0\ ,$$

where $x^{\pm} = x^0 \pm x^1$. Secondly, one can make a field redefinition and introduce the fields

$$\begin{aligned} \Omega &= \kappa^{-1}e^{-2\phi}+\frac{\phi}{2}+\frac{1}{4}\ln\frac{\kappa}{4} \\ \chi &= \kappa^{-1}e^{-2\phi}+\rho-\frac{\phi}{2}-\frac{1}{4}\ln(4\kappa)\ , \end{aligned} \quad (3.2)$$

where $\kappa \equiv \frac{N}{12}$. The coordinates $x^{\pm}$ can be fixed so that $\Omega = \chi$, so the dilaton field ϕ differs from the conformal factor of the metric ρ by a constant. The matter fields are assumed to reflect from the strong coupling boundary $\Omega = \Omega_{cr} = \frac{1}{4}$ [4]. The semiclassical equations can now be reduced to

$$\begin{aligned} \partial_+\partial_- f_i &= 0 \\ \partial_+\partial_-\Omega &= -\lambda^2 \\ \kappa\partial_\pm^2\Omega &= -\pi T^f_{\pm\pm}+\kappa t_\pm(x^\pm)\ . \end{aligned} \quad (3.3)$$

Here $T^f_{\pm\pm}$ are the components for outgoing and ingoing matter energy of the stress-tensor, which is defined as follows. Since the classical matter action is written as

$$S_f = -\frac{1}{4\pi}\int d^2x\sqrt{-g}\sum_{i=1}^{N}(\nabla f_i)^2\ , \quad (3.4)$$

the stress tensor is

$$T^f_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta S_f}{\delta g^{\mu\nu}}\ .$$

The components representing outgoing and ingoing matter are normalized as

$$T^f_{\pm\pm} = \frac{1}{2\pi}\sum_i \partial_\pm f_i \partial_\pm f_i \ ,$$

in the conformal gauge.

The functions $t_\pm(x^\pm)$ are fixed by boundary conditions. We assume that the incoming matter energy flux at $\mathcal{I}^-$ vanishes sufficiently rapidly at early and late times. Then, $t_+(x^+) = 1/4(x^+)^2$. The solution for the field Ω can now be found to be

$$\Omega = -\lambda^2 x^+(x^- + \frac{\pi}{\kappa\lambda}P_+(x^+)) + \frac{\pi}{\kappa\lambda}M(x^+) - \frac{1}{4}\ln(-4\lambda^2 x^+ x^-) \ , \tag{3.5}$$

where

$$\begin{aligned} M(x^+) &= \lambda \int_0^{x^+} du \ u \ T^f_{++}(u) \\ P_+(x^+) &= \int_0^{x^+} du \ T^f_{++}(u) \ . \end{aligned} \tag{3.6}$$

Consider now an incoming matter shock wave that carries energy M. The stress tensor is then given by

$$T^f_{++}(x^+) = \frac{M}{\lambda x_0^+}\delta(x^+ - x_0^+) \ , \tag{3.7}$$

which is the only non-vanishing component. We substitute this in the equations (3.6) above. Following [9], we set $\lambda x_0^+ = 1$.

After solving the RST equations one finds the following results. The Penrose diagram for the evaporating black hole spacetime is given in Fig. 1.

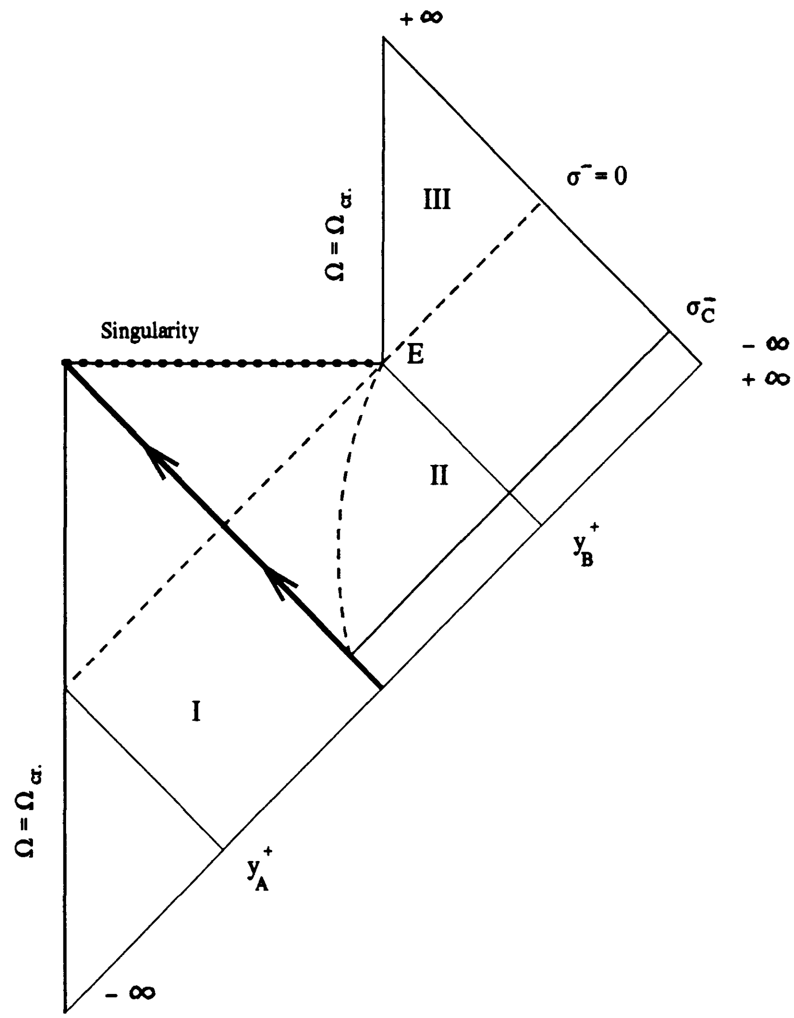

Figure 1: The black hole geometry formed by an incoming shock wave (thick line with arrows). Evaporation occurs in region II; regions I and II are linear dilaton vacuua.

The spacetime is seen to be divided into three regions. The 'lowest' region I, before the incoming shockwave at x_0^+, is the linear dilaton vacuum, bounded from the left by the timelike strong coupling boundary $\Omega = \Omega_{cr}$. The incoming shockwave forms the black hole by causing the boundary $\Omega = \Omega_{cr}$ to become spacelike. It can be shown that the scalar curvature R diverges at the spacelike portion of the boundary. The apparent horizon also forms after the incoming shockwave. Following [4], the apparent horizon is defined by the condition $\partial_+\phi = 0$. After the black hole forms, it starts to evaporate, and the apparent horizon shrinks until it meets the singularity at the endpoint of evaporation. At the endpoint of evaporation a short (delta function) burst of negative energy is seen to emerge from the black hole. This is called the 'thunderpop' [4], and it travels along the null line $x^- = x_s^-$ to future null infinity. (x_s^- is a light-cone coordinate of the endpoint to be specified later.) Thus the region II bounded by the thunderpop and the incoming shockwave is the curved region of the evaporating black hole. After the thunderpop, the spacetime becomes again flat and the boundary $\Omega = \Omega_{cr}$ becomes timelike. The corresponding region III is a linear dilaton vacuum.

In the linear dilaton vacuum region I the metric is

$$ds^2 = (\lambda^2 x^+ x^-)^{-1} dx^+ dx^- \ .$$

We can write it as $ds^2 = -dy^+ dy^-$ using coordinates

$$\begin{aligned} y^- &= -\frac{1}{\lambda}\ln(-\lambda x^-) \\ y^+ &= \frac{1}{\lambda}\ln(\lambda x^+) - y_0^+ \ . \end{aligned} \tag{3.8}$$

The shift y_0^+ is introduced to set the origin of the coordinate y^+ to the point A where the reflected ray of the event horizon (see Fig. 1.) meets $\mathcal{I}^-$.

The event horizon, the singularity and the apparent horizon meet at point E, the end point of evaporation. In our conventions, its coordinates are

$$(x_s^-,\ x_s^+) = (-\frac{\pi M}{\kappa\lambda^2}(1 - e^{-\frac{4\pi M}{\kappa\lambda}})^{-1},\ \frac{\kappa}{4\pi M}(e^{\frac{4\pi M}{\kappa\lambda}} - 1)) \ . \tag{3.9}$$

We can now specify what the shift y_0^+ is. In region I the reflecting boundary $\Omega = \Omega_{cr} = \frac{1}{4}$ is the line

$$y^+ = y^- + \frac{1}{\lambda}\ln\frac{1}{4} - y_0^+ \ .$$

Reflecting the line $x^- = x_s^-$ (off the boundary $\Omega = \Omega_{cr}$) back to $\mathcal{I}^-$, we find that the point A has $y_A^+ = -\frac{1}{\lambda}\ln(-\lambda x_s^-) + \frac{1}{\lambda}\ln\frac{1}{4} - y_0^+$. Setting $y_A^+ = 0$ yields $y_0^+ = -\frac{1}{\lambda}\ln(-4\lambda x_s^-)$.

The region II is the curved region of the evaporating black hole. It can be joined continuously but not smoothly with region III along the line $x^- = x_s^-$, with region III being again a linear dilaton vacuum, but with the coordinate x^- shifted. From the joining conditions the metric in III can be found to be

$$ds^2 = (\lambda^2 x^+ (x^- + \frac{\pi M}{\lambda^2\kappa}))^{-1} dx^+ dx^- \ .$$

If one makes the coordinate transformation

$$\begin{aligned} \sigma^+ &= \frac{1}{\lambda}\ln(\lambda x^+) \\ \sigma^- &= -\frac{1}{\lambda}\ln(\frac{\lambda x^- + \frac{\pi M}{\lambda\kappa}}{\lambda x_s^- + \frac{\pi M}{\lambda\kappa}}) \ , \end{aligned} \tag{3.10}$$

this metric becomes $ds^2 = -d\sigma^+ d\sigma^-$. The coordinate σ^- has been normalized so that the thunderpop is at $\sigma^- = 0$. The reflecting boundary $\Omega = \Omega_{cr}$ in region III is at $\sigma^+ = \sigma^- + \frac{1}{\lambda}\ln(\lambda x_s^+)$.

The metric in II becomes asymptotically flat near $\mathcal{I}^+$. We can extend the coordinates $\sigma^\pm$ from region III into region II in the neighbourhood of $\mathcal{I}^+$. Then the metric in region II also has the asymptotic form $ds^2 \to -d\sigma^+ d\sigma^-$ near $\mathcal{I}^+$. Thus, $\sigma^\pm$ are the physical coordinates near $\mathcal{I}^+$.

Finally, we identify some points of interest in the Penrose diagram. The point A where the reflection of the null line $x^- = x_s^-$ meets $\mathcal{I}^-$ we already set to be at $y_A^+ = 0$. The point B is the projection of the end point E along a null ray to past null infinity. We find it to be located at $y_B^+ = \frac{4\pi M}{\kappa\lambda^2}$. The point C is defined by projecting the point where the apparent horizon meets the incoming shockwave along a null ray to future null infinity. It is at

$$\sigma_C^- = -\frac{1}{\lambda}\ln\left(\frac{\kappa\lambda}{4\pi M}\left(\exp\left(\frac{4\pi M}{\lambda\kappa}\right) - 1\right)\right) .$$

Since $M >> \kappa\lambda$, to a good accuracy $\sigma_C^- = -\frac{4\pi M}{\lambda^2\kappa}$. The absolute value of σ_C^- is thus the total (retarded) time of evaporation of the black hole.

3.3 Bogoliubov transformations

In this section we calculate the Bogoliubov coefficients for the relation between the natural mode functions in the 'in' region close to $\mathcal{I}^-$ and the 'out' region close to $\mathcal{I}^+$. In [1] the Bogoliubov coefficients were estimated for modes travelling close to the horizon that forms in a spherical collapse of a star. In [6] the Bogoliubov coefficients were computed for the 1+1 dimensional eternal black hole geometry, *i.e.* without taking into account the backreaction on the metric due to the Hawking evaporation. Our notations follow those in [6] where we refer to for all introductory steps.

We choose the following form for modes at $\mathcal{I}^-$, $\mathcal{I}^+$ respectively

$$\begin{aligned} u_\omega &= \frac{1}{\sqrt{2\omega}} e^{-i\omega y^+} \qquad (in) \\ v_\omega &= \frac{1}{\sqrt{2\omega}} e^{-i\omega\sigma^-} \qquad (out) \, , \end{aligned} \tag{3.11}$$

where the normalization factor is in agreement with the form (3.4) of the matter action.

We define the Bogoliubov coefficients $\alpha_{\omega\omega'}$ and $\beta_{\omega\omega'}$ for the relation between the 'in' and 'out' modes as follows

$$v_\omega = \int_0^\infty d\omega' \, [\alpha_{\omega\omega'} u_{\omega'} + \beta_{\omega\omega'} u^*_{\omega'}] \, . \tag{3.12}$$

To compute the Bogoliubov coefficients, we pull the 'out' mode back to $\mathcal{I}^-$. First we divide the mode into two parts at the point $\sigma^- = 0$. The 'upper' piece $v_\omega = \frac{1}{\sqrt{2\omega}} e^{-i\omega\sigma^-} \theta(\sigma^-)$ reflects from the timelike boundary in region III without experiencing any blueshift. At $\mathcal{I}^-$ it becomes

$$v_\omega = \frac{1}{\sqrt{2\omega}} {}^{-i\omega(y^+ - y^+_B)} \theta(y^+ - y^+_B) \, ,$$

where y^+_B was defined in the previous section.

However, the 'lower' piece $v_\omega = \frac{1}{\sqrt{2\omega}} e^{-i\omega\sigma^-} \theta(-\sigma^-)$ gets distorted. It first experiences a blueshift when pulled back to region I. This is done by replacing the coordinate σ^- with the coordinate y^- using the relation

$$\sigma^- = -\frac{1}{\lambda} \ln[\frac{-e^{-\lambda y^-} + \pi M/\lambda\kappa}{\lambda x^-_s + \pi M/\lambda\kappa}] \, .$$

Then we reflect the mode from the timelike boundary in region I back to $\mathcal{I}^-$, by replacing y^- with

$$y^- = y^+ + y^+_0 - \frac{1}{\lambda} \ln \frac{1}{4} = y^+ - \frac{1}{\lambda} \ln(-\lambda x^-_s) \, .$$

The final relation between the coordinates σ^- and y^+ can then be written as

$$\begin{aligned} \sigma^- &= -\frac{1}{\lambda} \ln[\frac{\lambda x^-_s e^{-\lambda y^+} + \pi M/\lambda\kappa}{\lambda x^-_s + \pi M/\lambda\kappa}] \\ &= -\frac{1}{\lambda} \ln[\lambda\Delta(e^{-\lambda y^+} - C)] \, , \end{aligned} \tag{3.13}$$

where

$$\begin{aligned} \lambda\Delta &\equiv \frac{\lambda x^-_s}{\lambda x^-_s + \pi M/\lambda\kappa} = e^{4\pi M/\lambda\kappa} \\ C &\equiv -\frac{\pi M/\lambda\kappa}{\lambda x^-_s} = 1 - e^{-4\pi M/\lambda\kappa} \, . \end{aligned} \tag{3.14}$$

As an important aside, notice that (3.13) implies a relation between a small distance $d\sigma^-$ at $\mathcal{I}^+$ centered at σ^- and the corresponding small distance dy^+ at $\mathcal{I}^-$ which results from mapping the former distance along null rays which reflect from the boundary back to the past null infinity. This relation can be found to be

$$dy^+ = \{1 + (e^{4\pi M/\kappa\lambda} - 1)e^{\lambda\sigma^-}\}^{-1} d\sigma^- \, . \tag{3.15}$$

The result (3.15) tells us that a distance $d\sigma^-$ at $\mathcal{I}^+$ (before the endpoint of evaporation) maps to a much smaller distance dy^+ at $\mathcal{I}^-$, with the 'squeeze factor' becoming exponentially larger as the black hole evaporates. This observation will turn out to be crucial in applying the Srednicki calculation for black holes, as will be discussed in the section 6.

Returning back to the behaviour of the modes, (3.13) tells us that the 'lower' part of v_ω becomes

$$v_\omega = \frac{1}{\sqrt{2\omega}} \exp\{\frac{i\omega}{\lambda} \ln[\lambda\Delta(e^{-\lambda y^+} - C)]\}\theta(-y^+) \tag{3.16}$$

when pulled back to $\mathcal{I}^-$.

It is now straightforward to proceed to find the Bogoliubov coefficients. The result is

$$\begin{aligned}
\alpha_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\{\int_{-\infty}^{0} dy^+ \exp\{\frac{i\omega}{\lambda}\ln[\lambda\Delta(e^{-\lambda y^+} - C)] + i\omega' y^+\} \qquad (3.17)\\
&\quad + e^{i\omega' y_B^+} \int_0^\infty dz^+ \exp\{-i(\omega - \omega')z^+\}\}\\
\beta_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\{\int_{-\infty}^{0} dy^+ \exp\{\frac{i\omega}{\lambda}\ln[\lambda\Delta(e^{-\lambda y^+} - C)] - i\omega' y^+\}\\
&\quad + e^{-i\omega' y_B^+} \int_0^\infty dz^+ \exp\{-i(\omega + \omega')z^+\}\} \ .
\end{aligned}$$

These results resemble the ones of Giddings and Nelson (GN), with the following three differences: 1) there are additional terms resulting from the 'upper' part of the mode v_ω, *i.e.*, the part after the endpoint of evaporation, 2) $\lambda\Delta$ is different, and 3) we have $C = 1 - e^{-4\pi M/\kappa\lambda}$ while GN had $C = 1$. (Of course, $C \approx 1$ since $M/\kappa\lambda >> 1$. However, the small difference turns out to be significant if one tries to investigate the behaviour of the modes very near the endpoint.)

The integrals in the above expressions can be evaluated further. Substituting first $x = e^{\lambda y^+}$ and then $t = Cx$, we get

$$\begin{aligned}
\alpha_{\omega\omega'} &= \frac{1}{2\pi}\,(\frac{\omega'}{\omega})^{1/2}\ \{\ (\lambda\Delta)^{i\omega/\lambda}\ C^{i(\omega-\omega')/\lambda} \int_0^C dt\ (1-t)^{i\omega/\lambda}\ t^{-1-i(\omega-\omega')/\lambda} \qquad (3.18)\\
&\quad + e^{i\omega' y_B^+}\ \frac{-i}{\omega - \omega' - i\Delta M}\} \ .
\end{aligned}$$

The remaining integral can be identified as an Incomplete Beta function. The coefficient $\beta_{\omega\omega'}$ is computed similarly. The final expressions are

$$\begin{aligned}
\alpha_{\omega\omega'} &= \frac{1}{2\pi\lambda}\,(\frac{\omega'}{\omega})^{1/2}\ \{(\lambda\Delta)^{i\omega/\lambda}\ C^{i(\omega+\omega')/\lambda}\ B_C(-\frac{i\omega}{\lambda} + \frac{i\omega'}{\lambda} + \Delta M, i + \frac{i\omega}{\lambda}) \qquad (3.19)\\
&\quad - i\lambda\ e^{i\omega' y_B^+}\ (\omega - \omega' - i\Delta M')^{-1}\}
\end{aligned}$$

$$\beta_{\omega\omega'} = \frac{1}{2\pi\lambda}\,(\frac{\omega'}{\omega})^{1/2}\,\{(\lambda\Delta)^{i\omega/\lambda}\,C^{i(\omega-\omega')/\lambda}\,B_C(-\frac{i\omega}{\lambda}-\frac{i\omega'}{\lambda}+\Delta M, i+\frac{i\omega}{\lambda})$$
$$-\,i\lambda\,e^{-i\omega' y_B^+}\,(\omega+\omega'-i\Delta M')^{-1}\}\ .$$

Note that in the semiclassical approximation the thunderpop is a delta function at $\mathcal{I}^+$, thus there is a part of the field modes that is not captured by the modes $e^{-i\omega\sigma^-}$ at $\mathcal{I}^+$ for any finite range of ω. Thus the Bogoliubov coefficients computed below need to be supplemented with an infinite frequency component to completely describe the field at $\mathcal{I}^+$.

Let us now turn to the issue of examining the nature of the radiation by studying the approximate behaviour of the Bogoliubov coefficients. We expect outgoing thermal radiation at constant temperature $\frac{\lambda}{2\pi}$ to be seen in the region $\sigma^- \in (-\frac{4\pi M}{\kappa\lambda^2}, 0)$ of $\mathcal{I}^+$, except perhaps at the beginning and very end of the evaporation process. The coordinate y^+ corresponding to this region is very small, so we can approximate

$$\ln[\lambda\Delta(e^{-\lambda y^+} - C)] \approx \ln[1 - e^{4\pi M/\kappa\lambda}\lambda y^+]\ . \tag{3.20}$$

If we are not too close to the endpoint, the term 1 in (3.20) is negligible. For the essence of Hawking radiation, we have the frequency $\omega \sim \lambda$ at $\mathcal{I}^+$ and very high frequencies $\omega' \sim \lambda e^{4\pi(M+y^+)/\kappa\lambda}$ at $\mathcal{I}^-$. For such ω, ω' we can ignore the second integrals in the expression (3.17) for $\alpha_{\omega\omega'}, \beta_{\omega\omega'}$. We therefore get

$$\alpha_{\omega\omega'} \approx \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\{\int_{-\infty}^{0} dy^+ \exp\{\frac{i\omega}{\lambda}\ln[-e^{4\pi M/\lambda\kappa}\lambda y^+] + i\omega' y^+\}\} \tag{3.21}$$
$$\beta_{\omega\omega'} \approx \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\{\int_{-\infty}^{0} dy^+ \exp\{\frac{i\omega}{\lambda}\ln[-e^{4\pi M/\lambda\kappa}\lambda y^+] - i\omega' y^+\}\}\ .$$

This means that we get the thermal relation (see [6])

$$\alpha_{\omega\omega'} \approx -e^{\pi\omega/\lambda}\beta_{\omega\omega'} \tag{3.22}$$

and we see that the outgoing radiation is thermal with constant temperature $T_H = \frac{\lambda}{2\pi}$ in the region $\sigma^- \in (-\frac{4\pi M}{\kappa\lambda^2}, 0)$ at $\mathcal{I}^+$.

3.4 The stress tensor

Since we have worked out the relations between the 'in' and 'out' modes, we can easily calculate the VEV of the stress-energy at $\mathcal{I}^+$, *i.e.* $\langle 0, in \mid T_{\mu\nu}(\sigma^-) \mid 0, in\rangle^{ren}$ by using the point-splitting method. The non-vanishing component of $\langle T_{\mu\nu}\rangle^{ren}$ is the

$--$ component. Since we are computing the 'in' VEV at $\mathcal{I}^+$, we need first the form of the 'in' mode at $\mathcal{I}^+$. For $\sigma^- \leq 0$ it is

$$u_\omega = \frac{1}{\sqrt{2\omega}} \exp\{\frac{i\omega}{\lambda} \ln[\frac{1}{\lambda\Delta} e^{-\lambda\sigma^-} + C]\}\theta(-\sigma^-) \ . \tag{3.23}$$

Again, for $\sigma^- > 0$ (after the endpoint) there is no redshift and it is then trivial to see that the stress-tensor vanishes in this region. We concentrate only in the region before the endpoint. Using the point-splitting method, we first calculate

$$\langle T_{--}(\sigma^-)\rangle = \lim_{\sigma_1^-, \sigma_2^- \to \sigma^-} \frac{1}{2\pi} \frac{\partial}{\partial\sigma_1^-} \frac{\partial}{\partial\sigma_2^-} \int_0^\infty d\omega \ u_\omega(\sigma_1^-) u_\omega^*(\sigma_2^-) \ , \tag{3.24}$$

where the LHS means the VEV with respect to the 'in'-vacuum. (We use this notation in the following.) For $\sigma^- < 0$ the step functions are just constant and can be ignored. Taking the partial derivatives and making a series expansion in $(\sigma_1^- - \sigma_2^-)$ yields then a term which diverges quadratically in the limit, and a finite term. The divergent term must be subtracted (it is just the usual vacuum divergence) and the finite term gives the renormalized value for the stress-energy:

$$\langle T_{--}(\sigma^-)\rangle^{ren} = \frac{\lambda^2}{48\pi}[1 - \frac{1}{(1 + (e^{4\pi M/\kappa\lambda} - 1)e^{\lambda\sigma^-})^2}] \ . \tag{3.25}$$

This result is for one conformal scalar field, for N fields it must be multiplied by N. Note that in the region $\sigma^- \in (-\frac{4\pi M}{\kappa\lambda^2}, 0)$ we can approximate

$$\langle T_{--}(\sigma^-)\rangle^{ren} \approx \frac{\lambda^2}{48\pi} = \frac{\pi}{12}(T_H)^2 \ , \tag{3.26}$$

which is the correct value for outgoing thermal radiation at temperature $T_H = \frac{\lambda}{2\pi}$.

(Comment: if one would use the *approximate* behaviour of the mode very near the end point, one would find that $\langle T_{--}\rangle^{ren} \to 0$ at roughly Planck distance from the end point. One should not, however, trust such a local treatment when dealing with modes.)

The above formula gave $\langle T_{--}\rangle$ only at $\mathcal{I}^+$. However, in 1+1 dimensions it is easy to calculate $\langle T_{\mu\nu}\rangle^{ren}$ *everywhere* by using the trace anomaly, integrating the covariant conservation equation and applying the boundary conditions (see [11] for details). Using this route, we end up with the following results for the stress tensor everywhere in region II:

$$\begin{aligned} \langle T_{x^-x^-}\rangle^{ren} &= \frac{N}{48\pi} \frac{1}{(x^-)^2} (\frac{e^{-2\rho} + 1/4}{e^{-2\rho} - 1/4}) \{ (\frac{\lambda^2 x^+ x^- + 1/4}{e^{-2\rho} - 1/4})^2 - 1 \} \\ \langle T_{x^+x^+}\rangle^{ren} &= \frac{N}{48\pi} \frac{1}{(x^+)^2} (\frac{e^{-2\rho} + 1/4}{e^{-2\rho} - 1/4}) \{ (\frac{\lambda^2 x^+ (x^- + \pi M/\kappa\lambda^2) + 1/4}{e^{-2\rho} - 1/4})^2 - 1 \} \end{aligned} \tag{3.27}$$

$$\langle T_{x^+x^-}\rangle^{ren} = -\frac{N\lambda^2}{24\pi}(\frac{1}{e^{-2\rho}-1/4})$$
$$\{1+\frac{e^{-2\rho}}{(e^{-2\rho}-1/4)^2}\frac{[\lambda^2x^+(x^-+\frac{\pi M}{\kappa\lambda^2})+1/4][\lambda^2x^+x^-+1/4]}{\lambda^2x^+x^-}\}\ ,$$

where ρ is given implicitely via the equation

$$e^{-2\rho}+\frac{\rho}{2} = -\lambda^2x^+(x^-+\frac{\pi M}{\kappa\lambda^2})-\frac{1}{4}\ln[-\lambda^2x^+x^-]+\frac{\pi M}{\kappa\lambda}\ . \tag{3.28}$$

In the above the stress tensor is given in the 'Kruskal' coordinates $x^\pm$ since one is generally interested in its behaviour in both sides of the apparent horizon.

3.5 Two incoming shock waves

In this section we present the solution of the RST equations in the case of two incoming shock waves. We use this geometry later in section 6 where we discuss the entropy of the black hole. We want to first form a black hole of mass M_0, which then starts to evaporate. Then the second shock wave at a later time carries additional energy to the black hole. For instance, it could restore the black hole back to its original mass. The question then is: How much is the entanglement entropy after the evaporation process? If the second shock wave carries just enough energy to restore the black hole back to its original mass, but not more, is this entanglement entropy related to Bekenstein entropy? We discuss this question in section 6, here we will just derive the results that we will use.

Recalling how we defined the incoming shock wave, it is clear that for two incoming shock waves we should replace (3.7) with

$$T^f_{++}(x^+) = \frac{M_0}{\lambda x_0^+}\delta(x^+-x_0^+)+\frac{M_1}{\lambda x_1^+}\delta(x^+-x_1^+)\ . \tag{3.29}$$

The first term on the RHS is the first shock wave at x_0^+ (we set again $\lambda x_0^+ = 1$) carrying energy M_0 and the second term is a later shock wave at x_1^+ carrying energy M_1. We do not specify x_1^+, M_1 here, but x_1^+ should be chosen so that the black hole formed by the first shock wave has not yet evaporated away when the second shock wave reaches it. On could think of M_1 being the energy needed to restore the black hole back to its original size, but for now that is not essential.

The spacetime curve for the apparent horizon can be found by solving the equation $\partial_+\Omega = 0$ (this follows from setting $\partial_+\phi = 0$ for the dilaton field; which is the RST definition of the apparent horizon). The solution is

$$x^+ = -\frac{1}{4\lambda}\frac{1}{(x^-+\frac{\pi M_0}{\kappa\lambda^2}+\frac{\pi M_1}{\kappa\lambda^3 x_1^+}\theta(x^+-x_1^+))}\ , \tag{3.30}$$

for $x^+ > x_0^+$. The boundary of the spacetime is the critical line $\Omega = \Omega_{cr} = \frac{1}{4}$, which defines another curve $x^+ = x^+(x^-)$. The (final) endpoint of evaporation is where the above two curves meet. We find it to be located at

$$
\begin{aligned}
x_s^- &= -\frac{\pi}{\kappa\lambda^2}(M_0 + \frac{M_1}{\lambda x_1^+})(1 - e^{-4\pi(M_0+M_1)/\kappa\lambda})^{-1} \\
x_s^+ &= \frac{\kappa}{4\pi(M_0 + \frac{M_1}{\lambda x_1^+})}(e^{4\pi(M_0+M_1)/\kappa\lambda} - 1) \ .
\end{aligned} \tag{3.31}
$$

We can now join the region II with the linear dilaton vacuum in region III with an appropriate shift of the coordinate x^-. For the metric in region III, we find

$$
ds^2 = \frac{dx^+ dx^-}{\lambda^2 x^+(x^- + \frac{\pi}{\kappa\lambda^2}(M_0 + M_1/\lambda x_1^+))} \ .
$$

This in turn tells us how to define the coordinates $\sigma^\pm$ which are the 'physical' coordinates near $\mathcal{I}^+$. We define

$$
\begin{aligned}
\sigma^+ &= \frac{1}{\lambda}\ln(\lambda x^+) \\
\sigma^- &= -\frac{1}{\lambda}\ln[\frac{\lambda x^- + \frac{\pi}{\kappa\lambda}(M_0 + M_1/\lambda x_1^+)}{\lambda x_s^- + \frac{\pi}{\kappa\lambda}(M_0 + M_1/\lambda x_1^+)}] \ ,
\end{aligned} \tag{3.32}
$$

On the other hand, we still have the 'physical' coordinates in region I

$$
\begin{aligned}
y^+ &= \frac{1}{\lambda}\ln(\lambda x^+) - y_0^+ \\
y^- &= -\frac{1}{\lambda}\ln(-\lambda x^-) \ ,
\end{aligned} \tag{3.33}
$$

where $y_0^+ = -\frac{1}{\lambda}\ln(-4\lambda x_s^-)$, with x_s^- given now by (3.31). The reflecting boundary in region I is the line $y^+ = y^- + \frac{1}{\lambda}\ln\frac{1}{4} - y_0^+$.

It is straightforward to see that the relation between the coordinates σ^-, y^+ now becomes

$$
\sigma^- = -\frac{1}{\lambda}\ln[\lambda\Delta'(e^{-\lambda y^+} - C')] \ , \tag{3.34}
$$

where

$$
\begin{aligned}
\lambda\Delta' &= e^{4\pi(M_0+M_1)/\lambda\kappa} \\
C' &= 1 - e^{-4\pi(M_0+M_1)/\lambda\kappa} \ .
\end{aligned} \tag{3.35}
$$

As we noticed in section 3, this implies a relation between a distance $d\sigma^-$ at $\mathcal{I}^+$, centered at σ^- and the corresponding small distance dy^+ at $\mathcal{I}^-$. In the two shock wave case, the relation is

$$
dy^+ = \{1 + (e^{4\pi(M_0+M_1)/\kappa\lambda} - 1)e^{\lambda\sigma^-}\}^{-1} d\sigma^- \ . \tag{3.36}
$$

Again, a distance $d\sigma^-$in $\mathcal{I}^+$ (before the endpoint of evaporation) maps to a much smaller distance dy^+ in $\mathcal{I}^-$ and this 'squeeze factor' becomes exponentially large as the black hole evaporates. In the present case the 'squeeze factor' reaches the value $\sim e^{4\pi(M_0+M_1)/\kappa\lambda}$ which exceeds the value $\sim e^{4\pi M_0/\kappa\lambda}$ that would be obtained in the absence of the second shock wave. The result (3.36) will be used in the next section.

3.6 Entropy for 1+1 dimensional black holes

As argued by Hawking, the process of pair creation by the gravitational field of the black hole creates a state which is 'mixed' between the regions exterior and interior to the horizon. If we compute the reduced density matrix that describes the field in the exterior region, then the entropy computed for this density matrix gives the 'entropy of entanglement' [12] between the interior and exterior regions of the hole. If we do not take into account the backreaction from the created vradiation then an infinite number of particle pairs are produced and the entropy of entanglement will be found to be infinite also. But in the simple RST model we can estimate the produced entropy for the semiclassical geometry that includes backreaction. We shall do this in two ways:

1. We directly compute
$$S = -Tr\{\rho \ln \rho\} \tag{3.37}$$
where ρ is the reduced density matrix describing the field in the exterior region. Here we use the fact that ρ is to a good approximation a thermal density matrix.

2. We use the result of Srednicki described in the introduction. Thus a complete spacelike hypersurface is considered, and the part corresponding to the interior of the horizon is assumed to be, effectively, the traced over region considered in this approach.

1. The essential idea is to define reasonably localised regions on $\mathcal{I}^+$ such that the density matrix can be described as approximately thermal in those regions. Hawking presented this analysis for the 3+1 dimensional black hole [1]; it was worked out explicitly for the 1+1 case in [6] (without backreaction). We follow the notations of [6] in the following. Define a complete set of orthonormal wavepackets on $\mathcal{I}^+$:
$$v_{jn} = \varepsilon^{-1/2} \int_{j\varepsilon}^{(j+1)\varepsilon} d\omega \; e^{2\pi i \omega n/\varepsilon} \; v_\omega \, , \tag{3.38}$$
where $j = 0, 1, 2, \ldots$ and n are integers. The wavepacket v_{jn} is peaked about $\mathcal{I}^+$ coordinate $\sigma^- = 2\pi n/\varepsilon$, has a spatial width $\sim \varepsilon^{-1}$ and a frequency $\omega_j \approx j\varepsilon$. In this

basis the reduced density matrix obtained by tracing out the field inside the black hole is

$$\rho \;=\; N^{-1} \sum_{\{n_{jn}\}} e^{-\frac{2\pi}{\lambda}\sum_{jn} n_{jn}\omega_j} |\{n_{jn}\}\rangle\langle\{n_{jn}\}| \; . \tag{3.39}$$

This has the form of a thermal density matrix. We can write

$$\rho \;=\; \prod_{jn} \rho_{jn} \tag{3.40}$$

with

$$\rho_{jn} \;=\; N_{jn}^{-1} \sum_{n_{jn}=0}^{\infty} r(n_{jn}) |n_{jn}\rangle\langle n_{jn}| \tag{3.41}$$

$$r(n_{jn}) \;=\; e^{-\frac{2\pi}{\lambda} n_{jn}\omega_j} \; . \tag{3.42}$$

We get

$$S \;=\; \sum_{jn} S_{jn} \tag{3.43}$$

where

$$S_{jn} \;=\; -Tr_{(jn)}\{\rho_{jn} \ln \rho_{jn}\} \;=\; -\sum_{n_{jn}=0}^{\infty} r(n_{jn}) \ln r(n_{jn}) \; . \tag{3.44}$$

A brief computation gives

$$S_{jn} \;=\; \beta\omega_j (e^{\beta\omega_j} - 1)^{-1} \;-\; \ln(1 - e^{-\beta\omega_j}) \; . \tag{3.45}$$

Thus

$$\begin{aligned} S_n \;=\; \sum_{j=0}^{\infty} S_{jn} \;&=\; \sum_{j=0}^{\infty} \{\beta\omega_j (e^{\beta\omega_j} - 1)^{-1} \;-\; \ln(1 - e^{-\beta\omega_j})\} \\ &\to \int_0^{\infty} dj \{\beta j\varepsilon (e^{\beta j\varepsilon} - 1)^{-1} \;-\; \ln(1 - e^{-\beta j\varepsilon})\} \;=\; \frac{\pi^2}{3\beta\varepsilon} \; . \end{aligned} \tag{3.46}$$

The entropy from N evaporating fields is

$$S = N \sum_n S_n \; . \tag{3.47}$$

The separation between wavepackets (3.38) is $\Delta\sigma^- = 2\pi/\varepsilon$. Thus in time T the number n ranges from $n = 1$ to $n = \varepsilon T/2\pi$. From (3.46, 3.47) we get

$$S = \frac{N\pi T}{6\beta} \; . \tag{3.48}$$

The total evaporation time is $T = \frac{48\pi M}{N\lambda^2}$. This gives the estimate of the total entropy created in the Hawking radiation:

$$S_{\text{total}} \;=\; \frac{4\pi M}{\lambda} \; . \tag{3.49}$$

Note that this is *twice* the Bekenstein entropy, which for the 1+1 dimensional hole is $S_{\mathrm{Bek}} = \frac{2\pi M}{\lambda}$. The result (3.49) is the entropy of a thermal distribution of bosons at temperature β^{-1} and with energy $E = M$, which in one dimension is given by $S = 2\beta E$. For a discussion of how the entropy of radiation at $\mathcal{I}^+$ relates to the Bekenstein entropy, see [13].

2. We now investigate the application of Srednicki's result to the black hole. We split the discussion into 3 parts:

- (i) We recall the result of [8], and discuss the issue of infrared divergence.
- (ii) We discuss how this result may be applied to the black hole, after making some plausible arguments for the required modifications.
- (iii) We carry out the calculations for the entropy.

(i) The computation of Srednicki for the one space dimension case may be described as follows. We consider a free scalar field on a 1-dimensional lattice, with lattice spacing a. Let this field be in the vacuum state. We select a region of length R of this lattice and trace over the field degrees of freedom outside this region. This gives a reduced density matrix ρ, from which we compute $S = -Tr\{\rho \ln \rho\}$ which is the entropy of entanglement of the selected region with the remainder of the lattice. It is immaterial whether we trace over the interior or exterior of the selected region; since the field was in a pure state overall the entanglement entropy is the same in both cases. This entropy is given by [8]

$$S = \kappa_1 \ln(R/a) + \kappa_2 \ln(\mu R) , \tag{3.50}$$

for one scalar field. (For N species the result must be multiplied by N.) Here μ is an infrared cutoff.

The infrared term is very sensitive to boundary conditions. As an example consider taking a periodic lattice, and let the scalar field be periodic as well. If the field is massless, then the zero mode of the field varies over an infinite range, in the vacuum state. If we trace over the degrees of freedom in a subregion of the lattice, then the mean value of the field inside is correlated to the mean value outside, but this mean value can take on an infinite range of values. The entanglement entropy will thus be infinite.

If we take antiperiodic boundary conditions for the scalar field then the zero mode does not exist. With this choice (3.50) becomes [14]

$$S = \frac{1}{3}\ln(R/a) + \frac{1}{6}\ln(I/R) \tag{3.51}$$

where I is the infrared cutoff coming from the finite size of the lattice. Another way to kill the zero mode of the scalar field is to have a vanishing boundary condition for the field at say $x = 0$. Let the traced over region extend from $x = x_1$ to $x = x_2$. Now modes with wavelength much greater than x_2 effectively vanish over the interval (x_1, x_2), so they do not serve to entangle this region with the remainder of the line $x > 0$. Thus the entanglement entropy will be finite without the need for an explicit infrared regulator. A third way of dealing with the zero mode is to simply assume that the field has a small mass μ. Then no other infrared regulation should be needed and the results should not be sensitive to choice of boundary conditions.

(ii) For the black hole, we start by considering a complete spacelike hypersurface through the evaporating geometry, described as follows. Starting at spatial infinity ($\sigma^- = -\infty$), we move near $\mathcal{I}^+$ to a point with $\sigma^- = \sigma_1^- < 0$. (The black hole vanishes at $\sigma^- = 0$.) Then we smoothly bend this hypersurface so that it enters the horizon and reaches the timelike segment of the line $\Omega = \Omega_{cr}$. (This timelike segment occurs before this critical line becomes the singularity.) Thus the hypersurface is kept spacelike all through, and avoids the singularity by passing below it. An observer collecting radiation far from the black hole will see the part of this hypersurface that lies along $\mathcal{I}^+$, and we would wish to trace over the part that cannot be observed from outside the black hole. This would give the reduced density matrix, and thereby the entropy of entanglement between the field inside and outside the hole. Let us be more precise about what we take as the 'observed' part. Suppose that the observer at $\mathcal{I}^+$ carries an instrument with her which she uses to collect the outgoing radiation. At first, near $\sigma^- = -\infty$ nothing comes out from the black hole. The observer has to wait for quite a while before the black hole starts to radiate. At some point the radiation starts and becomes approximately thermal. We have discussed earlier that this happens roughly at $\sigma_0^- = -\frac{4\pi M}{\kappa\lambda^2}$. Correspondingly, around this point the observer turns on her instrument. The observer then collects radiation up to some retarded time σ_1^-, when she turns her instrument off again. Thus the part of the hypersurface between σ_0^- and σ_1^- corresponds to observations, and the rest of it will be traced over. Notice that the observer can neither start nor stop collecting radiation at an instant, but there will be a short time scale $d\sigma^-$ which she needs to turn on or shut off her instrument in a proper way. The time interval needed should be sufficiently rapid to give a good accuracy for specifying the turn-on point σ_0^- and the shut-off point σ_1^-, on the other hand it should not be so small that it creates disturbances in the matter field that generate radiation comparable to the Hawking radiation collected. It is reasonable to assume that the time intervals $d\sigma^-(\sigma_0^-)$, $d\sigma^-(\sigma_1^-)$ needed should be given by the typical wavelength in the outgoing thermal radiation. Thus we can assume that $d\sigma^- \sim \beta_H \sim \frac{1}{\lambda}$ (our result for the entropy of the hole will turn out not to depend on this choice).

We split the contribution of different modes to the entropy, as follows: a) the modes of wavelength $\omega^{-1} >> M/\lambda^2$ (these modes are effectively constant over the observed region (σ_0^-, σ_1^-), so they may be taken as a contribution to the 'zero mode'); b) the leftmoving modes with $\omega^{-1} \lesssim M/\lambda^2$; c) the rightmoving modes with $\omega^{-1} \lesssim M/\lambda^2$.

Modes of type a) will give a divergence in the entanglement entropy even in flat Minkowski space (without black hole). This happens when the mass μ of the field is taken to zero, or (if the field is massless) as the observed part of the hypersurface is taken further and further away from the line $\Omega = \Omega_{cr}$ where the field vanishes. Since we are interested in entropy of entanglement of the rightmoving Hawking radiation (which occurs over a distance M/λ^2), we subtract this (diverging) contribution arising from the large wavelength modes $\omega^{-1} >> M/\lambda^2$. We also ignore the leftmoving modes b), as they do not contribute to the Hawking radiation. The rightmoving modes c) are of interest to us, but after particle pairs have been created, the state of the field is not the vacuum state for the geometry on the spacelike hypersurface under consideration. Srednicki's result, on the other hand, applies to a vacuum state for the field. The essential idea is to follow the rightmoving (*i.e.* outgoing) modes of the field back from $\mathcal{I}^+$, through the collapsing matter to the line $\Omega = \Omega_{cr}$ where they reflect to left moving modes that can be followed back to $\mathcal{I}^-$. Here these left movers are in the vacuum state, so that we may apply a Srednicki type approach to estimate the entanglement entropy in these modes.

As we follow the radiation modes back to $\mathcal{I}^-$ in the manner indicated above, we observe the following. The region $\sigma^- \in (\sigma_0^-, \sigma_1^-)$ in $\mathcal{I}^+$ corresponds to a region $y^+ \in (y^+(\sigma_0^-), y^+(\sigma_1^-))$ in $\mathcal{I}^-$, where the relation between $\sigma_i^-, i = 0, 1$ and $y^+(\sigma_i^-), i = 0, 1$ is given by the formula (3.13) in section 3.3. Thus, the latter region is the region of starting points for ingoing rays which will experience redshift and give arise to the collected radiation in the 'observed' region of $\mathcal{I}^+$. Also, this region is separated from the rest of $\mathcal{I}^-$ by 'cuts' of size $dy^+(\sigma_0^-)$ and $dy^+(\sigma_1^-)$, the size of which follows from the size of the corresponding cuts near $\mathcal{I}^+$ by the relation (3.15) given in section 3.3. Now we can apply the result of [8]. We disentangle the finite region $y^+ \in (y^+(\sigma_0^-), y^+(\sigma_1^-))$ from the rest of $\mathcal{I}^-$ by cuts of size $a_0 = dy^+(\sigma_0^-), a_1 = dy^+(\sigma_1^-)$. Ignoring 'zero modes' and the leftmovers, we expect to create an entropy for each scalar field (see discussion in section 3.8 below)

$$S = \kappa_3 \ln(\frac{R}{a_0}) + \kappa_3 \ln(\frac{R}{a_1}) \ , \tag{3.52}$$

where $R = y^+(\sigma_1^-) - y^+(\sigma_0^-)$ and $\kappa_3 = \frac{1}{4}\kappa_1$. (A factor of $\frac{1}{2}$ because we are considering only rightmovers and another factor of $\frac{1}{2}$ because the contribution to S is separated over the two 'cuts'.)

(iii) Let us rewrite (3.52) as

$$2\kappa_3 \ln(R\lambda) + \kappa_3 \ln(\frac{1}{a_0\lambda}) + \kappa_3 \ln(\frac{1}{a_1\lambda}) \ . \tag{3.53}$$

We substitute $a_i = dy^+(\sigma_i), i = 0, 1$ given by the relation (3.15):

$$\begin{aligned} a_1 &= \{1 + (e^{4\pi M/\kappa\lambda} - 1)e^{\lambda\sigma_1^-}\}^{-1} d\sigma^- \\ &\approx e^{-4\pi M/\kappa\lambda} e^{-\lambda\sigma_1^-} \frac{1}{\lambda} \ , \\ a_0 &\sim \frac{1}{\lambda} \ . \end{aligned} \tag{3.54}$$

The former approximate formula is valid for $\sigma_1^- \in (-\frac{4\pi M}{\kappa\lambda^2}, 0)$ and the latter just results from the fact that the redshift is negligible for the rays at earlier times.

After substitution we get (ignoring the infrared cut-off term)

$$\begin{aligned} S &= \kappa_3 \ln(e^{4\pi M/\kappa\lambda} e^{\lambda\sigma_1^-}) + \kappa_3 \ln(\frac{1}{\lambda}\lambda) + 2\kappa_3 \ln(R\lambda) \\ &= \kappa_3(\frac{4\pi M}{\kappa\lambda} + \sigma_1^-) + 2\kappa_3 \ln(R\lambda) \ . \end{aligned} \tag{3.55}$$

The second term is negligible with respect to the first term, since it can be calculated that $R \sim \frac{1}{\lambda}$. The first term is the significant term. As we take σ_1^- closer and closer to the endpoint $\sigma^- = 0$ (the observer collects more outgoing radiation), the first term approaches

$$\kappa_3 \frac{4\pi M}{\kappa\lambda} \ . \tag{3.56}$$

The above result for entropy must be multiplied by N, the number of scalar fields. Let us deduce the value of κ_3 by comparing (3.56) with the entropy of entanglement estimated directly from the density matrix of the outgoing radiation (eq. (3.49)). This gives

$$S \approx \frac{4\pi M}{\lambda} \ , \ \text{if} \ \kappa_3 = \frac{1}{12} \ . \tag{3.57}$$

If our assumptions regarding the separation of the 'zero mode', of right and left movers are correct (*i.e.*, if $\kappa_1 = 4\kappa_3$, κ_1 as given in (3.50)), then (3.57) agrees with the calculation of Srednicki which gave $\kappa_1 = \frac{1}{3}$. We discuss this issue further in section 8.

Note that if we consider the evaporation right up to the endpoint, then the cutoff scale (over which the radiation measurement is switched off) must go to zero, and the entropy becomes infinite. But since we are using the semiclassical geometry, it is not justified to go below distance λ^{-1} (or $(N\lambda)^{-1}$, for large N) in our analysis. With this restriction, the entropy from the cutoff scale of the endpoint is ignorable compared to the entropy of the hole, for holes that evaporate over classical time scales $4\pi M/\kappa\lambda^2 >> \lambda^{-1}$.

The significance of the Bekenstein entropy for a black hole is a matter of debate. One hypothesis is that the horizon behaves as a membrane with $e^{S_{\mathrm{Bek}}}$ states, so that there is an upper limit to the entanglement entropy of the matter outside the hole with the hole itself [15]. Thus if a sufficiently large amount of matter were thrown into the hole then a part of the information would have to leak back out through subtle correlations in the Hawking radiation [16].

In a semiclassical treatment of the gravitational field, on the other hand, there seems to be no limit to the amount of information that can disappear into the black hole. Thus the entanglement entropy can also grow without bound when matter is repeatedly thrown into the hole and the black hole mass allowed to decrease back to M by evaporation.[1] It is possible to verify in the simple 1+1 dimensional evaporating RST solution that the entanglement entropy can indeed exceed the Bekenstein value by an arbitrary amount. We illustrate this by taking a simple example: the RST model with two incoming shock waves, discussed in the previous section.

We again apply the Srednicki approach to estimate entropy. Consider an observer at $\mathcal{I}^+$ collecting radiation, who switches on the measuring device at σ_0^- and switches it off at σ_1^- with the corresponding switch-on-off intervals $d\sigma^-(\sigma_i^-)$ as before. The only difference in the two shockwave case is that now the relation between the $d\sigma^-$'s at $\mathcal{I}^+$ and the corresponding dy^+'s at $\mathcal{I}^-$ is different. The relation was given by the formula (3.36) in section 4. Now we need to substitute

$$\begin{aligned} a_1 &= \{1 + (e^{4\pi(M_0+M_1)/\kappa\lambda} - 1)e^{\lambda\sigma_1^-}\}^{-1} d\sigma^- \\ &\approx e^{-4\pi(M_0+M_1)/\kappa\lambda} e^{-\lambda\sigma_1^-}\frac{1}{\lambda}\,, \\ a_0 &\sim \frac{1}{\lambda} \end{aligned} \tag{3.58}$$

into the equation (3.53) above. Also the distance R will be different, but it is still of the order of $\frac{1}{\lambda}$ and the R-dependent term can thus be ignored. All this gives for the

[1] We are grateful to S. Trivedi for pointing this out to us.

entropy (for N fields)

$$\begin{aligned} S &\approx N\kappa_3 \ln(e^{4\pi(M_0+M_1)/\kappa\lambda}e^{\lambda\sigma_1^-}) + N\kappa_3\ln(\frac{1}{\lambda}\lambda) + \ldots \\ &\approx N\kappa_3(\frac{4\pi(M_0+M_1)}{\kappa\lambda} + \sigma_1^-) + \ldots\ , \end{aligned} \tag{3.59}$$

where $+\ldots$ represents the ignored contributions from the subleading R-dependent term and the infrared cutoff term. As $\sigma_1^- \to 0$, the entropy becomes

$$S \approx N\kappa_3 \frac{4\pi(M_0+M_1)}{\kappa\lambda}\ . \tag{3.60}$$

Substituting $\kappa_3 = 1/12$ gives then

$$S \approx \frac{4\pi(M_0+M_1)}{\lambda}\ . \tag{3.61}$$

Note that the Bekenstein entropy, on the other hand, need not exceed $S_{Bek} = \frac{2\pi M_0}{\lambda}$ at any time if the mass of the hole never exceeds M_0. The final entropy of entanglement (3.61) could be made arbitrarily large (as one could see by considering *e.g.* a n-shock wave case).

3.7 Entropy for 3+1 dimensional black holes

Let us now turn to the 3+1 dimensional black hole. We will consider only the Schwarzschild case. The Bekenstein entropy for this hole is

$$S_{\rm bek} = A/4 = 4\pi(2M)^2/4 = 4\pi M^2 \tag{3.62}$$

The entropy collected in the form of radiation by an observer at $\mathcal{I}^+$ is also proportional to M^2 [13, 12] though it requires use of the 'transmission coefficients' $\Gamma(\omega)$ for its computation.

One is tempted to compare such entropies to the result obtained by carrying out the flat space calculation of Bombelli *et. al.* or Srednicki for the case of three space dimensions. In the latter calculation one considers a massless scalar field, say, in the vacuum state in three space dimensions. One traces over the field modes inside an imaginary sphere of radius R, and computes the entropy of the corresponding reduced density matrix. In this calculation it is convenient to decompose the scalar field into angular modes $Y_{l,m}(\theta,\phi)$, so that we obtain a 1-dimensional problem in the radial coordinate r for each such angular mode. The different angular modes decouple from each other, so we have to just add the entropies resulting from the computation for each mode. The radial co-ordinate is taken as a 1-dimensional lattice with lattice spacing a. The entropy is found by numerical computation to be [8]

$$S \approx .30(R/a)^2 \tag{3.63}$$

Thus we seem to reproduce the $\sim R^2$ dependence expected of the black hole entropy. But if we accept that (3.63) applies to the black hole then we are faced with the question: What should be the value of the cutoff a?

As we now argue, the result (3.63) is not the one relevant to the 3+1 dimensional black hole. In fact, the 1-dimensional result (3.50) is again the relevant one to use. To see this, consider decomposing the scalar field in the black hole geometry into modes with angular dependence $Y_{l,m}(\theta,\phi)$. Since we have assumed spherical symmetry, these modes decouple from each other. Thus we can consider studying the evolution of different fields (labelled by $\{l,m\}$) on the 1+1 dimensional geometry obtained by the spherically symmetric reduction of the 3+1 dimensional geometry. Proceeding in this way, one would need to find the 'entanglement entropy' of fields in one space dimension, which (for free fields) is given by the the result (3.50).

At this point one sees an important difference between the flat space 3-dimensional problem and the 3+1 dimensional black hole. This difference occurs in the number of angular modes $Y_{l,m}$ that contribute significantly to the entropy. Consider first the flat space problem. If we had taken a lattice with lattice spacing a all over 3-space, then on the boundary sphere of radius R we could consider angular modes with $l \lesssim R/a$. If we first reduce the Hamiltonian to a sum over modes and then put the r co-ordinate on a lattice, then again it is found that for $l >> R/a$ the degrees of freedom on the two sides of the $r = R$ boundary effectively decouple [8]. This happens because the radial wave equation is dominated at large l by a 'mass term' arising from the angular Laplacian, and such a term does not couple neighbouring sites of the r co-ordinate lattice. Since $-l \leq m \leq l$, the number of angular modes contributing to the entropy is $O\{(R/a)^2\}$, which explains the leading power dependence of the entropy (3.63) on the cutoff a. (Treating the 1-dimensional problem for each angular mode should give a $ln(R/a)$ dependence multiplying the dependence $\sim (R/a)^2$, but the above argument is too crude to distinguish the presence or absence of logarithmic terms.)

Thus we see that the difference between the a dependence of (3.50) and (3.63) can be understood in terms of the large number ($O\{(R/a)^2\}$) of angular modes contributing to the entropy in the 3-dimensional flat space problem. But in the 3+1 dimensional black hole, we know that most of the radiation comes out in only a few angular modes! In fact for a reasonable first estimate of the Hawking radiation one can require that only the s-wave modes ($l = 0$) emerge from the hole. We now compute the entropy of the 3+1 hole with such an approximation.

While we cannot solve the geometry of the evaporating 3+1 dimensional hole as accurately as for the RST model, for the purposes of our calculation we can consider the Schwarzschild metric with a time dependent mass M. The surface gravity of the black hole is $\kappa = (4M)^{-1}$. The temperature is $T = \kappa/2\pi = (8\pi M)^{-1}$. The luminosity in the s-wave mode is

$$L_s = \frac{dE}{dt} = \frac{1}{2\pi}\int_0^\infty \frac{d\omega\omega}{e^{\omega/T}-1} = \frac{\pi T^2}{12} \tag{3.64}$$

From (3.64) we compute $M(y)$, the mass of the hole at the $\mathcal{I}^+$ Schwartzschild coordinate y. (We take $y = 0$ at the endpoint of evaporation, so y is negative in the part of $\mathcal{I}^+$ where the Hawking radiation is being received.) We have

$$M(y) = (\frac{-y}{256\pi})^{1/3} . \tag{3.65}$$

As mentioned above, we approximate the evaporating geometry by just letting M depend on time. Letting v be the Minkowski co-ordinate at $\mathcal{I}^-$. This approximation then gives

$$dy = -4M(y)d(\ln(v_0 - v)) \tag{3.66}$$

when the co-ordinate v is close to the value v_0 which reflects at $r = 0$ to move along the event horizon. Integrating (3.66) gives

$$\ln(v_0 - v_f) - \ln(v_0 - v_i) = 96\pi(M_f^2 - M_i^2) \tag{3.67}$$

We have $M_i = M$, and we let M_f be of the order of Planck mass. Further, $\ln(v_i - v_0)$ can be ignored compared to $\ln(v_0 - v_f)$. Following the discussion of entropy in the 1-dimensional case, we write (for one evaporating field, one 'cut' and rightmovers only):

$$S \approx -\frac{1}{12}\ln(a) = -\frac{1}{12}\ln(\delta v) \tag{3.68}$$

Here δv is found from (3.66) by setting $dy = \Lambda$ (for some chosen scale Λ over which the observation of radiation is switched off):

$$\delta v = \frac{\Lambda}{4M_f}(v_0 - v_f) \approx \frac{\Lambda}{4M_f}e^{-96\pi M^2} \tag{3.69}$$

Substituting in (3.68) we get

$$S \approx 8\pi M^2 = 2S_{\text{Bek}} \tag{3.70}$$

We have ignored terms logarithmic in M; these corrections are smaller than the contribution of the $l \neq 0$ modes which we have also neglected.

We can now compare the result (3.70) for the entanglement entropy with the thermodynamic entropy collected at $\mathcal{I}^+$. Following [13], we compare the change in the thermodynamic entropy received at $\mathcal{I}^+$ to the change in the Bekenstein entropy of the hole:

$$R = dS/dS_{\text{Bek}} = \frac{\int_0^\infty dx x^2 \sigma(x)\{xe^x/(e^x-1) - \ln(e^x-1)\}}{\int_0^\infty dx x^3 \sigma(x)/(e^x-1)} \tag{3.71}$$

Here

$$\sigma(\omega) = \sum_{l,m} \Gamma_{l,m}(\omega)/[27(\omega M)^2] \tag{3.72}$$

is the absorptivity per unit area of the black hole. In the above used approximation to Hawking radiation we used only the $l = 0$ mode, and let the 'transmission coefficient' Γ be unity for all ω. This gives $\sigma(\omega) = 1/[27(\omega M)^2]$. Substituting this in (3.71) gives $R = 2$, in accordance with (3.70). For more realistic models, taking into a account the transmission coefficients $\Gamma_{l,m}(\omega)$, one obtains $R \sim 1.3 - 1.6$ [17]. To reproduce the effect of nontrivial $\Gamma(\omega)$ we would need to extend the Srednicki calculation to fields with position dependent 'mass term'; we do not investigate this further here.

We can also compare the above derivations to the direct computation of the entropy of the density matrix obtained in the evaporation process; *i.e.*, carry out the calculation analogous to (3.39) to (3.49). We again have $S_n = \pi^3 T/3\varepsilon$, but now $T = (8\pi M_n)^{-1} \equiv T_n$. Following the evaporation process we find $M_n = (n/128\varepsilon)^{1/3}$. This gives, as expected,

$$S = \sum_{n=1}^{n_{\max}} \frac{\pi^3 T_n}{3\varepsilon} \approx 8\pi M^2 = 2S_{\text{Bek}} . \tag{3.73}$$

In our above application of Srednicki's result, we found that the one space dimensional formulation was more relevant, rather than the three space dimensional one. On the other hand if an observer stands near the horizon of the black hole, she sees thermal radiation with power in a large number of angular momentum modes. Then it is possible that by carrying out the above calculations with a different choice of hypersurface (*e.g.* with the 'outside' part of the hypersurface corresponding to a static frame near the horizon) one would find relevant the equation (3.63).

3.8 Discussion

In this chapter we have carried out the computation of Bogoliubov coefficients, stress tensor and the entanglement entropy for an evaporating black hole. Before one had explicit models of evaporating geometries, such quantities had been calculated only in the absence of backreaction. With backreaction, it is possible to obtain for example the stress tensor in the region between the event horizon and the apparent horizon.

With regard to the issue of entropy, for the static black hole geometry (*i.e.*, ignoring the back reaction and evaporation) a calculation similar in spirit to the calculations of Bombelli *et. al.* and Srednicki was performed by Frolov and Novikov [18]. These authors computed the entanglement entropy of the field modes outside the horizon with the field modes inside the horizon. The quantum fluctuations of the horizon suggest a short-distance cutoff at the horizon. Using this cutoff, one obtains an entropy which is of the order of the Bekenstein entropy. Our calculation, on the other hand, considers the geometry of an evaporating hole. Our cutoff is at $\mathcal{I}^+$: It has the interpretation of the 'time of switching' of the instrument measuring the Hawking radiation and is thus related to the temperature of this radiation (see section 3.6). We find hte entropy of the Hawking radiation to be that computed by Zurek [13], instead of the Bekenstein value.

Concerning the application of Srednicki's result to the black hole entanglement entropy, we have made two assumptions. First, we have assumed that the infrared divergence comes from modes with wavelength very large compared to the system size; after these modes are removed, the entropy can be split into a contribution from rightmoving modes and a contribution from leftmoving modes. Second, we assumed that when we cut the region R out of a line, the entropy of system, say, can be split into two parts, one coming from each of the two cuts at the two ends of R. What we do now is offer some heuristic arguments to justify these assumptions.

First we wish to understand the appearance of the logarithmic dependence in the entanglement entropy. Consider a segment of the real line, $0 \leq x \leq I_1 + I_2$, divided into two regions near $x = I_1$ by a 'cut' of length $a << I_1, I_2$. Further, assume $I_1 < I_2$. The scalar field we take to vanish at $x = 0$, $x = I_1 + I_2$. What is the entanglement entropy of I_1 with I_2 (with cutoff scale a)?

Suppose $I_1 = I_2$. The field modes have wave numbers

$$k = \frac{n\pi}{I_1 + I_2}\ ,\ n = 1, 2, \ldots\ . \tag{3.74}$$

The mode $k = \pi/(I_1 + I_2)$ is 'shared' between the two sides I_1,I_2, and we assume that it gives an entropy s. Now consider modes with $1 \lesssim n \lesssim (I_1 + I_2)/a$. For any scale ω^{-1} of the wavelength we can make localized wavepackets just as was done in section 6 (eq. (3.38)). The number of such wavepackets is $\sim n$. Only the wavepacket that overlaps both sides of the cut at $x = I_1$ contributes to the entanglement entropy; and we again take this entropy to be s. (This is an assumption.) Equivalently, we find that each mode (3.74) contributes $\frac{s}{n}$ to the entropy, which becomes

$$S \approx s \sum_{n=1}^{I_1/a} \frac{1}{n} \approx s \ln \frac{I_1}{a} \ . \tag{3.75}$$

Now suppose instead that $I_1 << I_2$. For $k^{-1} >> I_1$, the mode essentially vanishes over I_1, and so cannot entangle this region with I_2. For wavelengths $k^{-1} \lesssim I_1$ we make localized wavepackets in the region $0 < x < \alpha x$, $\alpha \gtrsim 1$, just as in the discussion above, and thus find again eq. (3.75), where we note that I_1 is the smaller segment.

Now we address a more complicated case. We have the segments

- $I_1 : \ 0 < x < I_1$
- $R : \ I_1 < x < I_1 + R$
- $I_2 : \ I_1 + R < x < I_1 + I_2 + R$.

Let $R << I_1 << I_2$. I_1 and R are separated by a cut of size $a_1 << R$, and R and I_2 are separated by $a_2 << R$. We want the entanglement entropy of R with the remainder (*i.e.*, $I_1 \cup I_2$). Again, we do not need to consider modes with wavelength $k^{-1} >> I_1$, since these effectively vanish over R, I_1. The modes with $R < k^{-1} < I_1$ give a contribution $S \approx s \ln \frac{I_1}{R}$. The modes with $k^{-1} \lesssim R$ can be made into wavepackets that get 'partitioned' at two places; $x = I_1, x = I_1 + R$. These give the contributions $S \approx s \ln \frac{R}{a_1}$, $S \approx s \ln \frac{R}{a_2}$. Overall, we then get

$$S \approx s \ln \frac{R}{a_1} + s \ln \frac{R}{a_2} + s \ln \frac{I_1}{R} \ . \tag{3.76}$$

If $a_1 = a_2 \equiv a$, $I_1 = I_2 \equiv I$, we can write

$$S \approx 2s \ln \frac{R}{a} + s \ln \frac{I}{R} \ , \tag{3.77}$$

which resembles (3.50) with $\kappa_2 = -\frac{1}{2}\kappa_1$. The value of s we can fix by the direct calculation of the entropy in the 1+1 dimensional black hole evaporation preceeding eq. (3.49). For the modes with $k^{-1} \lesssim R$, one can clearly make a breakup between right and left movers. Thus we conclude $s = \frac{1}{6}$, after doubling up the obtained answer for the rightmovers alone.

The above gives a heuristic understanding of the entropy of entanglement, which should be useful in applying the result with a variety of boundary conditions.

In conclusion, we note that the 'exponential expansion' of coordinates near the horizon gives rise to thermal radiation, as was shown by Hawking. By using the result of Srednicki, the same coordinate transformation gives the entanglement entropy produced by the thermal radiation. Thus we seem to be one step closer to understanding the nature of black hole thermodynamics.

Acknowledgements

We thank M. Crescimanno, A. Dabholkar, C. Itzykson and G. Lifschytz for many discussions. We are grateful to S. Trivedi for discussions and for informing us of the reference [13].

This work was supported in part by funds provided by the U.S. Department of Energy (D.O.E.) under contract ♯DE-AC02-76ER03069 and cooperative agreement DE-FC02-94ER40818.

Bibliography

[1] S.W. Hawking, *Comm. Math. Phys.* **43** (1975) 199; *Phys. Rev.* **D14** (1976) 2460

[2] J.D. Bekenstein, *Phys. Rev.* **D7** (1973) 2333; *Phys. Rev.* **D9** (1973) 3292

[3] C.G. Callan, S.B. Giddings, J.A. Harvey and A. Strominger, *Phys. Rev.* **D45** (1992) R1005; For recent reviews see J.A. Harvey and A. Strominger, in the Proceedings of the TASI Summer School, June 3-28, 1992, Boulder, Colorado (World Scientific, 1993); S.B. Giddings, in the Proceedings of the International Workshop on Theoretical Physics, 6th Session, June 21-28, Erice, Italy (World Scientific, 1993)

[4] J.G. Russo, L. Susskind and L. Thorlacius, *Phys. Rev.* **D46** (1992) 3444; *Phys. Rev.* **D47** (1993) 533

[5] E. Verlinde and H. Verlinde, *Nucl. Phys.* **B406** (1993) 43; K. Schoutens, E. Verlinde and H. Verlinde, *Phys. Rev.* **D48** (1993) 2690; A. Bilal and C. G. Callan, *Nucl. Phys.* **B394** (1993) 73; S. P. de Alwis, *Phys. Lett.* **B289** (1992) 278; *Phys. Lett.* **B300** (1993) 330.

[6] S.B. Giddings and W. Nelson, *Phys. Rev.* **D46** (1992) 2486

[7] L. Bombelli, R.K. Koul, J. Lee and R.D. Sorkin, *Phys. Rev.* **D34** (1986) 373

[8] M. Srednicki, *Phys. Rev. Lett.* **71** (1993) 666

[9] A. Strominger, *Phys. Rev.* **D48** (1993) 5769

[10] A. Strominger and L. Thorlacius, *Phys. Rev. Lett.* **72** (1994) 1584

[11] J.A. Harvey and A. Strominger, in ref. [3]

[12] D. Page, *Black Hole Information*, University of Alberta report No. Alberta-Thy-23-93, hep-th/9305040; L. Susskind, L. Thorlacius and J. Uglum, *Phys. Rev.* **D48** (1993) 3743

[13] W.H. Zurek, *Phys. Rev. Lett.* **49** (1982) 1683

[14] M. Srednicki (private communication)

[15] L. Susskind, L. Thorlacius and J. Uglum, in ref. [12]

[16] see also J.D. Bekenstein, *Phys. Rev. Lett.* **70** (1993) 3680

[17] D. Page, *Phys. Rev. Lett.* **50** (1993) 1013

[18] V. Frolov and I. Novikov, *Phys. Rev.* **D48** (1993) 4545.

Chapter 4

Breakdown of the Semiclassical Approximation at the Black Hole Horizon

4.1 Introduction

Since the original papers of Hawking [1, 2] 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 [3]. Another reason is the work by 't Hooft [4, 5, 6] 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 [7] (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 [6] despite the fact that they may be spacelike separated.

Recently[1] 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 [9]. A common argument against this possibility is that

This chapter is based on work in collaboration with Gilad Lifschytz, Miguel Ortiz and Samir D. Mathur which has appeared in *Physical Review* **D51** (1995) 1764.

[1] Historically, the greatest champions of this view point have been Page [8] and 't Hooft [4, 5].

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 [9].

The 1+1 dimensional black hole problem including the effects of quantum gravity was recently studied in Ref. [10]. 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 [6]. Ref. [10] assumes a reflection boundary condition at a strong coupling boundary. Some natural modifications of this boundary condition have been studied recently in [11]. De Alwis *et. al.* [12] have investigated the semiclassical approximation in canonical quantum gravity, and the relation of the black hole problem to the issue of time. There have been many other studies of quantum gravity on the black hole problem, some of which are listed in [13].

Let us recall the basic structure of the black hole problem [2]. Collapsing matter forms a black hole, which then evaporates by emission of Hawking radiation [1]. 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 explicitly in the simple 1+1 dimensional models referred to above. One finds [14, 15] that this entropy equals the quantity expected on the basis of purely thermodynamic arguments [16].

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 [17]. 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 [18] 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 [19], but on the other hand an excess of decoherence conflicts with the correlations between position and momentum variables needed for the classical variable [20].

In this chapter we investigate this crude proposal and find that there is indeed a sense in which the semiclassical approximation breaks down in the 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 [21], 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

$$(H_{\text{gravity}} + H_{\text{matter}})\Psi[\phi(s), f(\phi)] = 0 . \tag{4.1}$$

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 [17]. 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

$$\Psi_{\phi(s)}[f(\phi(s))] \equiv \Psi[\phi(s), f(\phi(s))] , \tag{4.2}$$

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

$$\psi^{\mathcal{M}}_{\phi(s)}[f(\phi(s))] \tag{4.3}$$

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 (4.3) can be used to approximate the result obtained using the exact solution of the Wheeler-DeWitt equation of (4.2), 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[2]

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

[2]The role of fluctuations in the mass of the infalling matter was also discussed in [10]. Generally, fluctuations in geometry can also arise from other sources, but we shall ignore these here.

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 chapter), this means that

$$\langle \psi^{\mathcal{M}}_{\phi(s)} | \psi^{\bar{\mathcal{M}}}_{\phi(s)} \rangle \approx 1 \tag{4.4}$$

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 (4.4) will not be the same on all slices. For the semiclassical approximation to be good at any given slice, we need that (4.4) hold on that slice.

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).

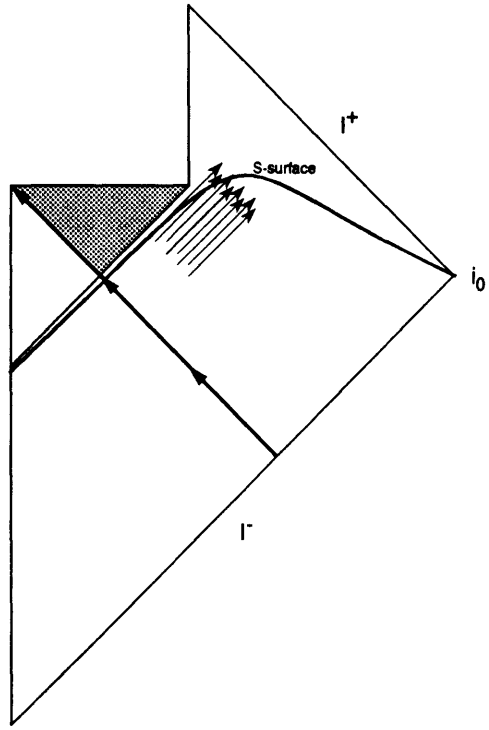

Figure 1: An example of an S-surface, shown in an evaporating black hole spacetime.

The importance of such slices to the black hole paradox has been emphasized by Preskill [24] and Susskind *et. al.* [9] 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 [22] 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 chapter 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 chapter is the following. In section 4.2 we review the CGHS model, and give some relevant scales. In section 4.3 we study the embedding of 1-geometries in different 1+1 dimensional semiclassical spacetimes. In section 4.4 we compare states of matter on the same 1-geometry, but in different spacetimes. Section 4.5 is a general discussion of the meaning of these results and of possible connections to other work.

4.2 A review of the CGHS model

There follows a quick review of the Callan-Giddings-Harvey-Strominger (CGHS) model [3], with reference to the Russo-Susskind-Thorlacius (RST) model [23] which i ncludes back-reaction and defines some relevant scales in the CGHS solution. Although all calculations in this chapter 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

$$S_G = \frac{1}{2\pi}\int dx dt\, \sqrt{-g}\, e^{-2\phi}\left[R + 4(\nabla\phi)^2 + 4\lambda^2\right] \tag{4.5}$$

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, (4.5) has static black hole solutions

$$e^{-2\rho} = e^{-2\phi} = \frac{M}{\lambda} - \lambda^2 x^+ x^- \tag{4.6}$$

and a linear dilaton vacuum (LDV) solution with $M = 0$. More interesting is the solution obtained when (4.5) is coupled to conformal matter,

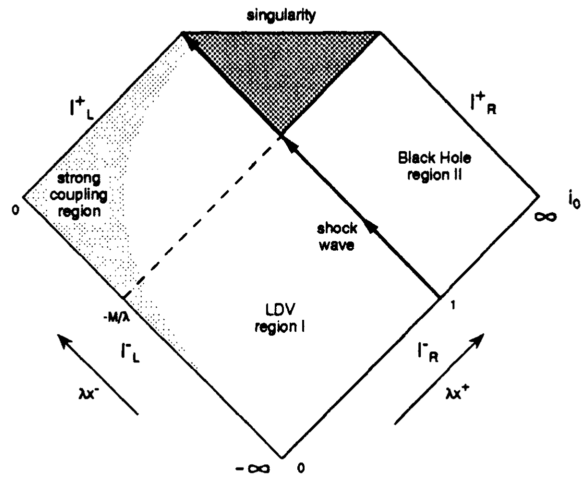

Figure 2: The Penrose diagram of the CGHS solution.

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

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

$$e^{-2\rho} = e^{-2\phi} = -\frac{M}{\lambda}(\lambda x^+ - 1)\Theta(x^+ - 1/\lambda) - \lambda^2 x^+ x^- \tag{4.7}$$

representing the formation of a black hole of mass M/λ in dimensionless 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 $\lambda x^- = -M/\lambda$.

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

$$\lambda x^+ = e^{\lambda y^+}, \qquad \lambda x^- = -\frac{M}{\lambda}e^{-\lambda y^-} \tag{4.8}$$

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

$$\lambda x^+ = e^{\lambda\sigma^+}, \qquad \lambda x^- + \frac{M}{\lambda} = -e^{-\lambda\sigma^-} \ . \tag{4.9}$$

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 (4.8) and (4.9) 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 [24, 25]).

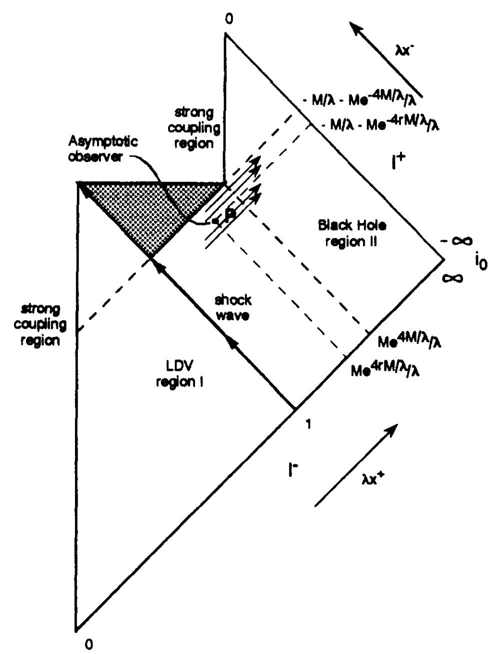

Figure 3: The Penrose diagram of the RST solution with some approximate scales shown.

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.

4.3 Embedding of 1-geometries

In this section, we shall compare how a certain spacelike hypersurface Σ may be embedded in collapsing black hole spacetimes (4.7) 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 [26]). 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.

4.3.1 Basic Equations

Here we present the basic equations describing the embedding of Σ. In a collapsing black hole manifold $\mathcal{M}$ of mass M (4.7), 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 (4.7)

$$\frac{M}{\lambda} - \lambda x^+ f(\lambda x^+) = \frac{\bar{M}}{\lambda} - \lambda \bar{x}^+ \bar{f}(\lambda \bar{x}^+) \tag{4.10}$$

$$\frac{f + \lambda x^+ f'}{\sqrt{-f'}} = \frac{\bar{f} + \lambda \bar{x}^+ \bar{f}'}{\sqrt{-\bar{f}'}} \tag{4.11}$$

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 (4.10) and (4.11), for $\bar{x}^+(x^+)$, differentiate (4.10) by x^+ and divide by (4.11), 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,

$$\ln(\lambda \bar{x}^+) = 2 \int d(\lambda x^+) \frac{f'}{(f + \lambda x^+ f') \mp \sqrt{(f - \lambda x^+ f')^2 - 4\Delta M f'/\lambda}} \tag{4.12}$$

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 $\bar{x}^-(x^-)$:

$$\ln(\lambda \bar{x}^- + \bar{M}/\lambda) = 2 \int dz^- \frac{g'}{(g + z^- g') \mp \sqrt{(g - z^- g')^2 - 4\Delta M g'/\lambda}} \, , \tag{4.13}$$

In (4.12) and (4.13), 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.

4.3.2 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 (4.10) and (4.11), 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

$$\frac{(\bar{b} + \bar{M}/\lambda)^2}{\bar{\alpha}^2} = \frac{4\Delta M}{\lambda} + \frac{(b + M/\lambda)^2}{\alpha^2} \tag{4.14}$$

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 (4.12) 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

$$\bar{\alpha} = \alpha + \frac{\delta/\alpha - \alpha \pm \sqrt{(\alpha^2 - \delta)^2/\alpha^2 + 4\Delta M/\lambda}}{2x_0^+} \tag{4.15}$$

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 (4.14) and (4.15). 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 chapter. 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.

4.3.3 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 $d\phi/ds = -\lambda$:

$$\lambda x^- = \begin{cases} -\alpha^2\lambda x^+ - 2\alpha\sqrt{\dfrac{M}{\lambda}} - \dfrac{M}{\lambda} & (\lambda x^+ \geq 1) \\ -\left(\alpha + \sqrt{\dfrac{M}{\lambda}}\right)^2 \lambda x^+ & (\lambda x^+ \leq 1) \end{cases} . \tag{4.16}$$

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 [22] 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.

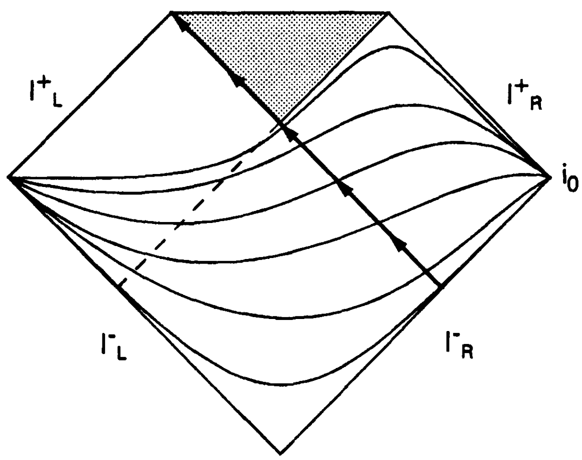

Figure 4: Examples of the complete slices of Sec. 3.3.

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

$$\lambda \bar{x}^- = \begin{cases} -\bar{\alpha}^2 \lambda \bar{x}^+ - 2\bar{\alpha}\sqrt{\dfrac{\bar{M}}{\lambda}} - \dfrac{\bar{M}}{\lambda} & (\lambda \bar{x}^+ \geq 1) \\ -\left(\bar{\alpha} + \sqrt{\dfrac{\bar{M}}{\lambda}}\right)^2 \lambda \bar{x}^+ & (\lambda \bar{x}^+ \leq 1) \end{cases} . \tag{4.17}$$

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 (4.9), 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

$$\begin{aligned} \lambda \bar{\sigma}_P^+ - \lambda \sigma_P^+ &\approx -\frac{\Delta M}{2M}\left(\frac{\lambda}{M}\right)^{1/2} \\ \lambda \bar{\sigma}_P^- - \lambda \sigma_P^- &\approx -\frac{\Delta M}{2M} , \end{aligned} \tag{4.18}$$

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

$$\lambda \bar{x}^+ = \lambda x^+ + \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\bar{M}}{\lambda}}\right) \tag{4.19}$$

$$\approx \lambda x^+ - \frac{\Delta M}{2\lambda\alpha}\sqrt{\frac{\lambda}{M}}\,. \tag{4.20}$$

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.

4.4 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 4.2), 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.

The heuristic arguments above suggest 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). 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. 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

$$\langle \psi, \Sigma, \mathcal{M} | \psi, \Sigma, \bar{\mathcal{M}} \rangle \approx 1 \tag{4.21}$$

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 $\Delta M/\lambda \sim e^{-4\tau M/\lambda}$ in the mass of a black hole, dramatically violating condition (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

$$\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle \tag{4.22}$$

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 (4.22), we first need to find the induced Bogoliubov transformation

$$v_\omega = \int_0^\infty d\omega' \left[\alpha_{\omega\omega'} \bar{v}_{\omega'} + \beta_{\omega\omega'} \bar{v}^*_{\omega'} \right] \tag{4.23}$$

between the in-modes

$$\begin{aligned} \bar{v}_\omega &= \frac{1}{\sqrt{2\omega}} e^{-i\omega\bar{\sigma}^+} \\ v_\omega &= \frac{1}{\sqrt{2\omega}} e^{-i\omega\sigma^+} , \end{aligned} \tag{4.24}$$

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

$$\sigma^+ = \sigma^+(\bar{\sigma}^+) . \tag{4.25}$$

Let us derive the relation (4.25) above, for the example of Section 3.3. As (4.20) 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[3] 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:

$$e^{-\phi} = \begin{cases} \alpha\lambda x^+ + \sqrt{\dfrac{M}{\lambda}} = \alpha\lambda\bar{x}^+ + \sqrt{\dfrac{\bar{M}}{\lambda}} & (\lambda x^+ \geq 1) \\ \left(\alpha + \sqrt{\dfrac{M}{\lambda}}\right)\lambda x^+ = \alpha\lambda\bar{x}^+ + \sqrt{\dfrac{\bar{M}}{\lambda}} & \left(1 \geq \lambda x^+ \geq \dfrac{\alpha + \sqrt{\frac{\bar{M}}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}}\right) \\ \left(\alpha + \sqrt{\dfrac{M}{\lambda}}\right)\lambda x^+ = \left(\alpha + \sqrt{\dfrac{\bar{M}}{\lambda}}\right)\lambda\bar{x}^+ & \left(\dfrac{\alpha + \sqrt{\frac{\bar{M}}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}} \geq \lambda x^+ \geq 0\right) \end{cases} . \quad (4.26)$$

Rewriting (4.26) using the asymptotic coordinates, we then get the relation (4.25):

$$\lambda\sigma^+ = \begin{cases} \ln\left[e^{\lambda\bar{\sigma}^+} - \dfrac{1}{\alpha}\left(\sqrt{\dfrac{M}{\lambda}} - \sqrt{\dfrac{\bar{M}}{\lambda}}\right)\right] & \left(\lambda\bar{\sigma}^+ \geq \lambda\bar{\sigma}_1^+\right) \\ \ln\left[\left(\dfrac{\alpha}{\alpha + \sqrt{\frac{M}{\lambda}}}\right)\left(e^{\lambda\bar{\sigma}^+} + \dfrac{1}{\alpha}\sqrt{\dfrac{\bar{M}}{\lambda}}\right)\right] & \left(\lambda\bar{\sigma}_1^+ \geq \lambda\sigma^+ \geq 0\right) \\ \lambda\bar{\sigma}^+ + \ln\left[\dfrac{\alpha + \sqrt{\frac{\bar{M}}{\lambda}}}{\alpha + \sqrt{\frac{M}{\lambda}}}\right] & \left(0 \geq \lambda\bar{\sigma}^+\right) \end{cases} , \quad (4.27)$$

where

$$\lambda\bar{\sigma}_1^+ \equiv \ln\left[1 + \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\bar{M}}{\lambda}}\right)\right] . \quad (4.28)$$

This coordinate transformation is illustrated in Fig. 5.

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

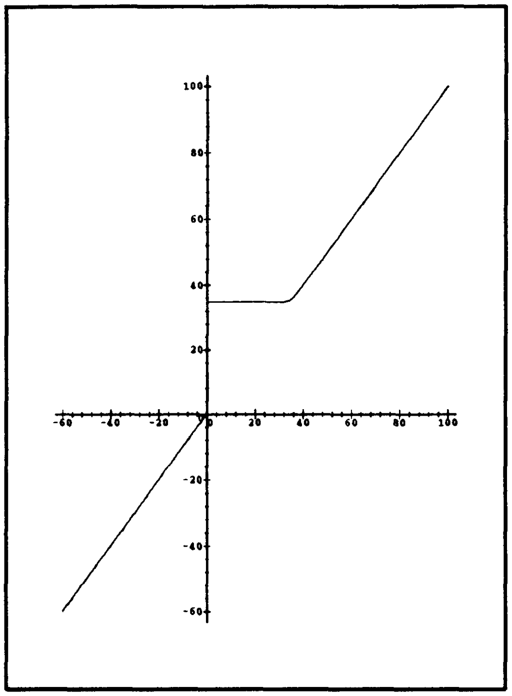

Figure 5: A graph of the function $\lambda\bar{\sigma}^+(\lambda\sigma^+)$. The vertical axis = $\lambda\bar{\sigma}^+$, the horizontal axis = $\lambda\sigma^+$. The part of the graph to the right (left) of the vertical axis corresponds to both points being above (below) the shock wave. (The interpolating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

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

$$\begin{aligned} \alpha_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\, I^+_{\omega\omega'} \\ \beta_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\, I^-_{\omega\omega'} \ , \end{aligned} \tag{4.29}$$

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

$$I^{\pm}_{\omega\omega'} \equiv \int_{-\infty}^{\infty} d\bar{\sigma}^+ \; e^{-i\omega\sigma^+(\bar{\sigma}^+) \pm i\omega'\bar{\sigma}^+} \quad . \tag{4.30}$$

Substituting the relations (4.27) above, we get

$$\begin{aligned} I^{\pm}_{\omega\omega'} &= \int_{\bar{\sigma}_1^+}^{\infty} d\bar{\sigma}^+ \; \exp\left\{-i\frac{\omega}{\lambda}\ln\left[e^{\lambda\bar{\sigma}^+} - \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\overline{M}}{\lambda}}\right)\right] \pm i\omega'\bar{\sigma}^+\right\} \qquad (4.31) \\ &+ \int_0^{\bar{\sigma}_1^+} d\bar{\sigma}^+ \; \exp\left\{-i\frac{\omega}{\lambda}\ln\left[\frac{\alpha}{\alpha+\sqrt{\frac{M}{\lambda}}}\left(e^{\lambda\bar{\sigma}^+} + \frac{1}{\alpha}\sqrt{\frac{\overline{M}}{\lambda}}\right)\right] \pm i\omega'\bar{\sigma}^+\right\} \\ &+ \int_{-\infty}^{0} d\bar{\sigma}^+ \; \exp\left\{-i\omega\bar{\sigma}^+ - i\frac{\omega}{\lambda}\ln\left[\frac{\alpha+\sqrt{\frac{\overline{M}}{\lambda}}}{\alpha+\sqrt{\frac{M}{\lambda}}}\right] \pm i\omega'\bar{\sigma}^+\right\} \\ &\equiv I^{\pm}_{\omega\omega'}(1) + I^{\pm}_{\omega\omega'}(2) + I^{\pm}_{\omega\omega'}(3) \quad . \end{aligned}$$

To identify the first integral, introduce first

$$\lambda\Delta \equiv 1 + \frac{1}{\alpha}\left(\sqrt{\frac{M}{\lambda}} - \sqrt{\frac{\overline{M}}{\lambda}}\right) = e^{\lambda\bar{\sigma}_1^+} \, . \tag{4.32}$$

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

$$\lambda\Delta e^{-\lambda y} \equiv e^{\lambda\bar{\sigma}^+} \, . \tag{4.33}$$

We now find that

$$I^{\pm}_{\omega\omega'}(1) = \lambda\Delta^{\pm i\omega'/\lambda}\int_{-\infty}^{0} dy \; \exp\left\{-i\frac{\omega}{\lambda}\ln\left[\lambda\Delta(e^{-\lambda y} - C)\right] \mp i\omega' y\right\} \quad , \tag{4.34}$$

where

$$C \equiv 1 - \frac{1}{\lambda\Delta} \, . \tag{4.35}$$

Now we can recognize the integral to be the same as discussed in [14]. 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 $\lambda y \in (-1, 0)$ [1, 14, 27]. 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 (4.20) 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

$$I^{\pm}_{\omega\omega'}(1) \approx \lambda\Delta^{\pm i\omega'/\lambda}\int_{-\infty}^{0} dy \; \exp\left\{-i\frac{\omega}{\lambda}\ln\left[-\lambda\Delta\lambda y\right] \mp i\omega' y\right\} \quad . \tag{4.36}$$

As discussed in [27], this leads to the approximate relation

$$I^{+}_{\omega\omega'}(1) \approx -e^{\pi\omega/\lambda}\, \left(I^{-}_{\omega\omega'}(1)\right)^{*} \tag{4.37}$$

for the integrals.

The logarithm in (4.36) implies that $I^{\pm}_{\omega\omega'}(1)$ contributes significantly in the regime $\omega' - \omega \gg 1$. Since the coordinate transformation (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 (4.37) is that the relationship of the Bogoliubov coefficients is (approximately) thermal,

$$\alpha_{\omega\omega'} \approx -e^{\pi\omega/\lambda}\, \beta^{*}_{\omega\omega'}\ , \tag{4.38}$$

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 (4.38) is. Also, the region of $\lambda\sigma^{+}$ which corresponds to (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 (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 [30]. We then find (see Appendices):

$$|\langle 0\text{ in}, \Sigma, \mathcal{M}|0\text{ in}, \Sigma, \bar{\mathcal{M}}\rangle|^2 = \left(\det(1+\beta\beta^{\dagger})\right)^{-\frac{1}{2}}\ . \tag{4.39}$$

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

$$|\langle 0\text{ in}, \Sigma, \mathcal{M}|0\text{ in}, \Sigma, \bar{\mathcal{M}}\rangle|^2 \ < \ \frac{1}{\gamma}\ , \tag{4.40}$$

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

$$\left|\frac{\Delta M}{\lambda}\right| = \left|\frac{M}{\lambda} - \frac{\bar{M}}{\lambda}\right| \ > \ \gamma^{48}\sqrt{\frac{M}{\lambda}}\,\alpha\ , \tag{4.41}$$

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 (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 (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 (4.41) will depend on both the slope α^2 and the intercept[4] 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.

4.5 Conclusions

Let us review what we have computed in this chapter:

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 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.

[4]Recall that the hypersurfaces of section 3.3 had $\delta = \alpha^2 + 2\alpha\sqrt{M/\lambda}$ so the rhs of (4.41) depends only on α.

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 nearthe 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 4.3 and 4.4 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 Verlinde *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.

Acknowledgements

We would like to thank R. Brooks, M. Crescimanno, J. Cohn, A. Dabholkar, S. Das, S. P. de Alwis, J. Goldstone, A. Guth, J. J. Halliwell, B. L. Hu, R. Jackiw, P. Kelly, K. Kuchař, G. Kunstatter, J. Samuel, A. Sen, C. Stephens, L. Susskind, S. Trivedi, H. Verlinde, A. Vilenkin, B. Zwiebach and especially J. G. Demers for discussions.

4.6 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 [28], 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

$$\begin{aligned} a(k) &= \int \frac{ds}{\sqrt{4\pi\omega_k}} e^{iks} \left(\omega_k f(x(s)) + \frac{\delta}{\delta f(x(s))}\right) \\ a^\dagger(k) &= \int \frac{ds}{\sqrt{4\pi\omega_k}} e^{-iks} \left(\omega_k f(x(s)) - \frac{\delta}{\delta f(x(s))}\right) \end{aligned} \tag{4.42}$$

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

$$(\Psi_1, \Psi_2) = \int \frac{\prod_k da(k) da^\dagger(k)}{2\pi i} \left(\Psi_1[a^\dagger(k), t]\right)^* \Psi_2[a^\dagger(k), t] \exp\left[-\int dk\; a(k) a^\dagger(k)\right] \tag{4.43}$$

where $(a^\dagger(k))^* = a(k)$ [29].

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 (4.43), 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 (4.43) 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

$$\begin{aligned} b(p) &= \sum_k \left[\alpha_{kp} a(k) + \beta^*_{kp} a^\dagger(k)\right] \\ b^\dagger(p) &= \sum_k \left[\beta_{kp} a(k) + \alpha^*_{kp} a^\dagger(k)\right] \end{aligned} \tag{4.44}$$

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

$$\begin{aligned} \bar{b}(p) &= \sum_k \left[\bar{\alpha}_{kp} a(k) + \bar{\beta}^*_{kp} a^\dagger(k)\right] \\ \bar{b}^\dagger(p) &= \sum_k \left[\bar{\beta}_{kp} a(k) + \bar{\alpha}^*_{kp} a^\dagger(k)\right] \end{aligned} \tag{4.45}$$

Now we can employ the inner product (4.43) 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 (4.44) and (4.45) that the b's and $\bar{b}$'s are related by

$$\begin{aligned} b(p') &= \sum_p \left[(\bar{v}_p, v_{p'})\bar{b}(p) + (\bar{v}^*_p, v_{p'})\bar{b}^\dagger(p)\right] \\ b^\dagger(p') &= \sum_p \left[-(\bar{v}^*_p, v^*_{p'})\bar{b}^\dagger(p) - (\bar{v}_p, v^*_{p'})\bar{b}(p)\right] \end{aligned} \tag{4.46}$$

where

$$(\bar{v}_p, v_{p'}) = -i \int_\Sigma dx \bar{v}_p(x,t) \partial_t v^*_{p'}(x,t) \tag{4.47}$$

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 (4.46) need not be evaluated on Σ as in (4.47). 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 (4.47) from x to σ (the t differentiation becomes an x differentiation which absorbs the factor $dx/d\sigma$), yielding

$$(\bar{v}_p, v_{p'}) = -i \int d\sigma \bar{v}_p(\bar{\sigma}^+(\sigma^+))\, \partial_{\sigma^0}\, v^*_{p'}(\sigma^+) \tag{4.48}$$

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 (4.48) looks exactly like the familiar integral for Bogoliubov coefficients, even though it involves mode functions on different manifolds. (4.48) 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 $\mathcal{M}$ ($\bar{\mathcal{M}}$).

4.7 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

$$\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle = (\det(\alpha))^{-\frac{1}{2}} \quad , \tag{4.49}$$

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 [30]. However, it is more convenient to consider not the overlap but the probability amplitude

$$|\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle|^2 = (\det(\alpha\alpha^\dagger))^{-\frac{1}{2}} \quad , \tag{4.50}$$

where the components of the matrix $\alpha\alpha^\dagger$ are

$$(\alpha\alpha^\dagger)_{\omega\omega'} = \int_0^\infty d\omega'' \; \alpha_{\omega\omega''}\alpha^*_{\omega'\omega''} \quad . \tag{4.51}$$

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

$$v_{jn} \equiv a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega \; e^{2\pi i\omega n/a} \; v_\omega \quad . \tag{4.52}$$

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 [1, 27, 14]. In the new basis, the Bogoliubov coefficients become

$$\begin{aligned} \alpha_{jn\omega'} &= a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega \; e^{2\pi i\omega n/a} \; \alpha_{\omega\omega'} \\ \beta_{jn\omega'} &= a^{-\frac{1}{2}} \int_{ja}^{(j+1)a} d\omega \; e^{2\pi i\omega n/a} \; \beta_{\omega\omega'} \end{aligned} \tag{4.53}$$

with the normalization

$$\int_0^\infty d\omega'' \left[\alpha_{jn\omega''}\alpha^*_{j'n'\omega''} - \beta_{jn\omega''}\beta^*_{j'n'\omega''}\right] = \delta_{jj'}\delta_{nn'} \quad . \tag{4.54}$$

The thermal relation (4.38) becomes

$$\beta^*_{jn\omega'} \approx -e^{-\pi\omega_j/\lambda}\alpha_{jn\omega'} \quad . \tag{4.55}$$

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

$$n_{max} = \frac{\lambda L}{\Delta(\lambda\sigma^+)} = \frac{\ln[1+(e-1)\lambda\Delta]}{2\pi\lambda/a} \tag{4.56}$$

packets are centered in this region.

Combining (4.54) and (4.55), we now see that

$$(\alpha\alpha^\dagger)_{jnj'n'} \approx \frac{\delta_{jj'}\delta_{nn'}}{1 - e^{-2\pi\omega_j/\lambda}} \tag{4.57}$$

for n, n' "inside" λL. For the other values of n, n' (at least one of them being "outside"),

$$(\alpha\alpha^\dagger)_{jnj'n'} \approx \delta_{jj'}\delta_{nn'} \; . \tag{4.58}$$

We are now ready to calculate the overlap (4.50). We get

$$\begin{aligned} \ln\left[|\langle 0 \text{ in}, \Sigma, \mathcal{M}|0 \text{ in}, \Sigma, \bar{\mathcal{M}}\rangle|^2\right] &\approx -\frac{1}{2}\left\{\sum_n^{\text{(inside)}}\sum_j \ln\left[\frac{1}{1 - e^{-2\pi\omega_j/\lambda}}\right]\right. \\ &\quad \left. + \ln\left[\prod_n^{\text{(outside)}}\prod_j 1\right]\right\} \\ &= \frac{1}{2}\, n_{max}\sum_j \ln\left[1 - e^{-2\pi ja/\lambda}\right] \; . \end{aligned} \tag{4.59}$$

In order to estimate the last term, we convert the sum to an integral[5]:

$$\begin{aligned} \sum_j \ln\left[1 - e^{-2\pi ja/\lambda}\right] &\rightarrow \int_0^\infty dj \; \ln\left[1 - e^{-2\pi ja/\lambda}\right] \\ &= \frac{\lambda}{2\pi a}\left(\frac{-\pi^2}{6}\right) \; . \end{aligned} \tag{4.60}$$

[5]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.

Now, combining (4.59) and (4.60), we finally get va useful formula for the overlap:

$$
\begin{aligned}
|\langle 0 \text{ in}, \Sigma, \mathcal{M} | 0 \text{ in}, \Sigma, \bar{\mathcal{M}} \rangle|^2 &\approx \exp\left\{ -\frac{1}{2} \frac{\lambda L}{2\pi\lambda/a} \frac{\lambda}{2\pi a} \frac{\pi^2}{6} \right\} \\
&= e^{-\frac{1}{48}\lambda L} \quad .
\end{aligned}
\tag{4.61}
$$

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

$$
\begin{aligned}
48 \ln \gamma &= \lambda L = \ln\left[1 + (e-1)\lambda\Delta\right] \\
&\approx \ln\left[1 + (e-1)\left(1 - \frac{\Delta M}{2\alpha\lambda}\sqrt{\frac{\lambda}{M}}\right)\right] \quad ,
\end{aligned}
\tag{4.62}
$$

where we used (4.20) in the last step. Solving for ΔM, we get

$$
\frac{\Delta M}{\lambda} \approx -\frac{2}{e-1}\gamma^{48}\sqrt{\frac{M}{\lambda}}\,\alpha \ . \tag{4.63}
$$

If $|\Delta M/\lambda|$ is bigger, the states are approximately orthogonal. Notice that since we used (4.20) in the end, (4.63) is a special result for the hypersurfaces of section 3.3. However, it is straightforward to generalize (4.63) 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 (4.63) will then depend on both α and the intercept δ.

Bibliography

[1] S. W. Hawking, *Comm. Math. Phys.* **43** (1975) 199.

[2] S. W. Hawking, *Phys. Rev.* **D14** (1976) 2460.

[3] 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).

[4] G. 't Hooft, *Nucl. Phys.* **B335** (1990) 138, and references therein.

[5] G. 't Hooft, *Nucl. Phys.* **B256** (1985) 727.

[6] C. R. Stephens, G. 't Hooft and B. F. Whiting, *Class. Qu. Grav.* **11** (1994) 621.

[7] T. Jacobson, *Phys. Rev.* **D48** (1993) 728.

[8] D. N. Page, *Phys. Rev. Lett.* **44** (1980) 301.

[9] L. Susskind, L. Thorlacius and J. Uglum, *Phys. Rev.* **D48** (1993) 3743; L. Susskind, *Phys. Rev.* **D49** (1994) 6606.

[10] 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.

[11] S. R. Das and S. Mukherji, *Phys. Rev.* **D50** (1994) 930; A. Strominger and L. Thorlacius, *Phys. Rev.* **D50** (1994) 5177.

[12] S. P. de Alwis, *Phys. Lett.* **B317** (1993) 46; S. P. de Alwis and D. A. MacIntire, *Phys. Rev.* **D50** (1994) 5164.

[13] S. P. de Alwis, *Phys. Lett.* **B289** (1992) 278; *Phys. Rev.* **D46** (1992) 5429; 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, *Phys. Rev.* **D50** (1994) 2368; *Phys. Rev.* **D50** (1994) 3889; *Phys. Rev.* **D48** (1993) 1687; C. Kiefer, R. Muller and T. P. Singh, *Mod. Phys. Lett.* **A9** (1994) 2661; 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).

[14] E. Keski-Vakkuri and S. D. Mathur, *Phys. Rev.* **D50** (1994) 917.

[15] T. M. Fiola, J. Preskill, A. Strominger and S. P. Trivedi, *Phys. Rev.* **D50** 3987; C. Holzhey, F. Larsen and F. Wilczek, *Nucl. Phys.* **B424** (1994) 443.

[16] W. H. Zurek, *Phys. Rev. Lett.* **49** (1982) 1683.

[17] V. Lapchinsky and V. Rubakov, *Acta Phys. Pol.* **B10** (1979) 1041; T. Banks, *Nucl. Phys.* **B249** (1985) 332; 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); J. J. Halliwell, *Phys. Rev.* **D36** (1989) 3621; 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).

[18] 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).

[19] E. Calzetta and F. D. Mazzitelli, *Phys. Rev.* **D42** (1990) 4066.

[20] J. P. Paz and S. Sinha, *Phys. Rev.* **D44** (1991) 1038.

[21] 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.

[22] See for example S. B. Giddings, *Phys. Rev.* **D49** (1994) 4078.

[23] J. G. Russo, L. Susskind and L. Thorlacius, *Phys. Rev. D***46** (1992) 3444; *Phys. Rev.* **D47** (1993) 533.

[24] J. Preskill, *Do black holes destroy information?*, CalTech Report No. CALT-68-1819 (hep-th/9209058). Presented at International Symposium on Black Holes, Membranes, Wormholes and Superstrings, Woodlands, TX, 16-18 Jan. 1992.

[25] S. Giddings, *Black holes and quantum predictability*, Santa Barbara Report No. UCSBTH-93-16 (hep-th/9306041). Presented at 7th Nishinomiya-Yukawa Memorial Symposium: Quantum Gravity, Nishinomiya, Japan 19-20 Nov. 1992, and at 1992 YITP Workshop on Quantum Gravity, Uji, Japan Nov. 24-27 1992.

[26] T. Banks and M. O'Loughlin, *Nucl. Phys.* **B362** (1991) 649; D. Louis-Martinez, J. Gegenberg and G. Kunstatter, *Phys. Lett.* **B321** (1994) 193; K. V. Kuchař *Phys. Rev.* **D50** (1994) 3961; H. A. Kastrup and T. Thiemann, *Nucl. Phys.* **B425** (1994) 665.

[27] S. B. Giddings and W. M. Nelson, *Phys. Rev.* **D46** (1992) 2486.

[28] K. V. Kuchař, *Phys. Rev.* **D39** (1989) 1579.

[29] L. D. Faddeev and A. A. Slavnov, *Gauge Fields: An Introduction to Quantum Theory*, Frontiers in Physics # 83, Addison-Wesley, Redwood City (1991).

[30] B. S. DeWitt, *Phys. Rep.* **19C** (1975) 297.

Chapter 5

Turning Points in the Semiclassical Approximation

5.1 Introduction

In this chapter we discuss various ways how the existence of classical turning points can affect the validity of the semiclassical approximation. We argue that turning points can create more complicated phase correlations than what can be seen in the leading order semiclassical approximation without backreaction. We discuss this in the context of simple quantum mechanical models and in the case of a minisuperspace model of quantum matter in a closed universe. We also show how turning points appear naturally in the evolution of a two dimensional black hole. We comment briefly on the various consequences which may result.

5.2 Simple Quantum Mechanical Examples

In this section we review the semiclassical approximation. To keep the discussion as simple as possible, we consider simple quantum mechanical examples of a light particle coupled to a heavy particle. We ask what happens if the heavy particle encounters a turning point. We give a criterion for the validity of the semiclassical approximation and show that it breaks down in the vicinity of the turning point. We make an estimate of the size of the region where this happens. Then we ask if the region can be large enough in order of the breakdown of the approximation to be relevant. We present another criterion for this. Next, we compare an exact solution of a quantum mechanical model with a semiclassical solution. We show that the presence of a turning point may leave a lasting 'imprint' to the phase correlations

Based on work in progress with Samir D. Mathur

of the wavefunction. This 'imprint' cannot be seen in the leading order semiclassical approximation.

5.2.1 A Heavy and a Light Particle

We start with simple quantum mechanical examples. Consider a light particle coupled to a very massive one. The heavy particle m_1 moves in a potential $U(x_1)$ and the light particle m_2 is coupled to the heavy particle through a potential $u(x_1, x_2)$. The total Hamiltonian of the system is thus

$$H \equiv H_1 + H_{12} = \frac{p_1^2}{2m_1} + U(x_1) + \frac{p_2^2}{2m_2} + u(x_1, x_2) \ . \tag{5.1}$$

Let us promote the Hamiltonian as a quantum operator $\hat{H}$ and study its zero energy eigenequation

$$\hat{H} \ \Psi(x_1, x_2) = 0. \tag{5.2}$$

This equation is often used as a simple toy analogue of the Wheeler-de Witt equation, viewing the heavy particle analogous to the gravitational degrees of freedom and the light particle analogous to the matter degrees of freedom. The heavy particle - gravity analogue can be made better if the potential $U(x_1)$ is assumed to be of the form $U(x_1) = m_1 V(x_1)$. In particular, the model can then be used as a simple preparatory example in discussing how the semiclassical approximation (= quantum field theory in curved spacetime) to quantum gravity is recovered from the Wheeler-de Witt equation. The standard discussion in the literature [1] goes as follows. An approximate solution to (5.2) can be found by writing an ansatz

$$\Psi(x_1, x_2) = e^{\frac{i}{\hbar}[m_1 S_0 + m_1^0 S_1 + m_1^{-1} S_2 + \cdots]}$$

and solving the equation order by order in m_1. At order $\mathcal{O}(m_1^0)$, the approximate solution is given by the wavefunction

$$\Psi(x_1, x_2) = \psi_{\rm WKB}(x_1) \ \chi(x_1, x_2) \ , \tag{5.3}$$

where $\psi_{\rm WKB}$ is a WKB wavefunction for the heavy particle

$$\psi_{\rm WKB}(x_1) = \frac{1}{(-2V(x_1))^{(1/4)}} \ e^{\pm \frac{i}{\hbar} m_1 S_0(x_1)} \ . \tag{5.4}$$

The exponent S_0 satisfies the Hamilton-Jacobi equation for the zero energy classical motion of the heavy particle

$$\frac{1}{2}(\frac{dS_0}{dx_1})^2 + V(x_1) = 0 \tag{5.5}$$

so

$$S_0(x_1) = \int dx_1 \sqrt{-2V(x_1)} \ . \tag{5.6}$$

In the Hamilton-Jacobi equation, dS_0/dx_1 is the velocity of the heavy particle, so we can find its classical zero energy trajectory $x_1(t)$ by integrating first

$$t(x_1) = \int \frac{dx_1}{\sqrt{-2V(x_1)}} \tag{5.7}$$

and then turning the result inside out to obtain $x_1(t)$. Now we can think that the classical motion of the heavy particle defines a 'clock' measuring time t. This is often called a 'WKB time'. Since the heavy particle potential was assumed to be of the form $U = m_1 V$, the time is independent of the mass of the heavy particle, making the notion of the clock more elegant. The clock is useful in the following way. The function $\chi(x_1, x_2)$ satisfies an equation

$$i\hbar \frac{dS_0}{dx_1} \frac{\partial}{\partial x_1} \chi(x_1, x_2) = \hat{H}_{12} \chi(x_1, x_2) \tag{5.8}$$

which we can rewrite into a more suggestive form using the fact that

$$\frac{dS_0}{dx_1} \frac{\partial}{\partial x_1} = \frac{\partial}{\partial t} \ .$$

The equation (5.8) can now be recoqnized as a time dependent Schrödinger equation for the light particle,

$$i\hbar \frac{\partial}{\partial t} \chi = \hat{H}_{12}(x_1(t), x_2) \ \chi \ . \tag{5.9}$$

[Note that the role that $x_1(t)$ plays in this equation resembles the situation in adiabatic approximation in quantum mechanics, exepct that the motion of the heavy particle is not assumed to be 'slow'.] Now let us summarize what was done. We started by looking at a zero energy time independent Schrödinger equation (5.2). We found an approximate solution, which tells us that the heavy particle motion is essentially classical whereas the light particle behaves in a quantum mechanical way. Although the starting equation does not give any time evolution for the total state, the heavy and the light particle are mutually correlated in such a way that the light particle can be thought to time evolve according to a Schrödinger equation with the time measured by the position of the heavy particle. Finally, we point out that since exponent of the WKB part of the total wavefunction is proportional to m_1, the heavy particle part of the wavefunction is much more rapidly oscillating as the χ part which describes the light particle.

We would like to comment that there seems to sometimes be some confusion about the proper name of this approximation and its relation to the Born-Oppenheimer approximation. There are three differences to the latter:

1. The wavefunction χ was not assumed to be an energy eigenstate of $\hat{H}_{12}$.
2. For the heavy particle, an additional WKB approximation is used.
3. The light particle satisfies a time dependent Schrödinger equation with respect to a time defined by the classical motion of the heavy particle.

The name 'semiclassical approximation' is therefore appropriate. To be more precise, we call this the leading order semiclassical approximation to make a distinction with a treatment where higher order ($\mathcal{O}(m_1^{-1})$) corrections would be included.

We will now turn our attention to the question what happens to the semiclassical approximation if the heavy particle potential U can have classical turning points. Discussions in the literature about this issue seem to be mostly concerned about tunneling issues. However, we will try to avoid tunneling effects and demonstrate that there are other additional issues involved.

The WKB approximation in one dimensional quantum mechanics problems is known to be applicable as long as the de Broglie wave length λ_{dB} of the particle varies only very slightly over distances which are of the order of the wave length itself. Since the de Broglie wave length is inversely proportional to the momentum of the particle, and by definition the momentum approaches zero at a classical turning point, in the vicinity of the turning point λ_{dB} grows very quickly and the WKB approximation is no longer applicable. Therefore, in the present case it should also be obvious that the semiclassical approximation must lose its validity when the heavy particle approaches a turning point. Our first task is incorporate this feature into the formalism and then make some estimates on the size of the region where the semiclassical approximation is invalid.

To avoid tunneling effects, we assume that the heavy particle has only one classical turning point, at $x_1 = a$. It is important to keep in mind that the position of a turning point depends on the available energy of the particle. In the semiclassical approximation the heavy particle followed a zero energy classical trajectory, therefore we will assume that $x_1 = a$ is a turning point for zero energy. It is natural to specify the integration limits in (5.6) by writing

$$S_0(x_1) = \int_{x_1}^{a} dx_1 \sqrt{-2V(x_1)} \tag{5.10}$$

so the heavy particle phase factor vanishes at the turning point.

Now we investigate how rapidly the heavy particle and light particle phase factors oscillate near the turning point. Notice that the time dependece in (5.9) arises through the classical motion $x_1 = x_1(t)$ of the heavy particle. Near the turning point, the motion of the heavy particle will slow down. For simplicity, let us assume that the light particle wavefunction χ is an energy eigenfunction of $\hat{H}_{12}$ with an energy eigenvalue $E(x_1(t))$. The parameter $x_1(t)$ is slowly evolving. If we also assume that the energy levels are sufficiently separated, the time evolution of the light particle wavefunction is given by the adiabatic approximation. Thus,

$$\chi(x_1(t), x_2) = \exp\{-\frac{i}{\hbar}\int_0^t E(x_1(t'))dt'\}\ \chi(a, x_2)\ , \tag{5.11}$$

where the time is taken to be zero at the turning point: $x_1(0) = a$. Now, we can change the integration variable from t to x_1. Using $\dot{x}_1 = dS_0/dx_1$ with (5.5) we get

$$\chi(x_1, x_2) = \exp\{-\frac{i}{\hbar}\int_a^{x_1} dx_1' \frac{E(x_1')}{\sqrt{-2V(x_1')}}\}\ \chi(a, x_2)\ .$$

We can now compare the heavy particle phase factor

$$\varphi_h = \frac{1}{\hbar}\ m_1 \int_{x_1}^a dx_1' \sqrt{-2V(x_1')} \tag{5.12}$$

with the light particle phase factor

$$\varphi_l = \frac{1}{\hbar}\int_{x_1}^a dx_1' \frac{E(x_1')}{\sqrt{-2V(x_1)}}\ . \tag{5.13}$$

An essential feature of the semiclassical approximation is that the heavy particle sector was supposed to oscillate more rapidly than the light particle sector, or $\partial\varphi_h/\partial x_1 \gg \partial\varphi_l/\partial x_1$. This is equivalent to

$$E(x_1) \ll -2m_1 V(x_1) = m_1(\frac{dx_1}{dt})^2\ . \tag{5.14}$$

In other words, the semiclassical approximation applies to the region where the kinetic energy of the heavy particle is much larger than the energy of the light particle. This will be violated when the heavy particle slows down as it approaches the turning point. We will now make a simple estimate of when the semiclassical region is not applicable. We approximate the heavy particle potential near the turning point

$$U(x_1) \equiv m_1\ V(x_1) \approx m_1\ V'(a)\ (a - x_1)\ ,$$

and assume that $E(x_1) \approx E(a)$ so that we can pull it outside the integral in (5.13).

Then, the semiclassical approximation is not good in a region of size

$$|\Delta x_1| \sim \frac{E}{|V'(a)|}$$

near the turning point. But this is irrelevant if the region is so small that the heavy particle wavefunction does not have time to go through more than several full oscillations. Using this additional condition $\varphi_h(a - \Delta x_1) > 2\pi$ we find that if

$$E \gtrsim (\frac{(\hbar\ U'(a))^2}{m_1})^{\frac{1}{3}} , \tag{5.15}$$

the breakdown of the semiclassical approximation can be thought to be significant near the classical turning point.

A good strategy to gain more insight to the turning point effects is to consider a simple example where everything can be calculated exactly. For this purpose, let us choose specific potentials $U(x_1), u(x_1, x_2)$ and consider the model

$$H = \frac{p_1^2}{2m_1} + m_1 r x_1 + \frac{p_2^2}{2m_2} + \frac{\kappa}{2}(x_1 - x_2)^2 . \tag{5.16}$$

This system describes a heavy particle moving on an incline (with r denoting the slope of the incline) and a light particle moving on a line, connected to the heavy particle by a spring with a spring constant κ. Let us now find an exact solution of the time independent Schrödinger equation

$$\hat{H}\ \Psi(x_1, x_2) = -E_T\ \Psi(x_1, x_2) ,$$

where an arbitrary energy E_T reflects the freedom in the choice of a zero energy. Let us factorize the equation by changing the coordinates x_1, x_2 to the center of mass coordinate X and the relative separation y. This gives the two equations

$$\{-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial y^2} + \frac{\kappa}{2}(y - \frac{r\mu}{\kappa})^2\}\ G(y) = E\ G(y) \tag{5.17}$$

$$\{-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial X^2} + m_1 r X\}\ F(X) = -(E + E_T)\ F(X) .$$

Here M is the total mass and μ is the reduced mass. Since $m_1 \gg m_2$, we can assume that $M \approx m_1$ and $\mu \approx m_2$. We will speak somewhat loosely and call $X \approx x_1$ as the position of the heavy particle and $y \approx x_2$ as the position of the light particle. The first equation is the simple harmonic oscillator Schrödinger equation with the familiar discrete eigenenergies $E = E_n$ and eigenfunctions $G(y) = G_n(y)$. Consider now the second equation. We introduce the wavenumber $k(X)$ and write the second equation as

$$F''(X) + k^2(X)\ F(X) = 0 .$$

Now, the heavy particle has a classical turning point at

$$X = -\frac{E_n + E_T}{m_1 r} \equiv a(E_n) \ . \tag{5.18}$$

We would like to emphasize that *the location of the turning point depends on the energy of the light particle.* This is different from what was seen in the semiclassical approximation, there the location of the turning point did not depend on what the light particle was doing. We will show that this fact may give rise to additional effects which can be missed in the semiclassical approximation. We adjust the zero energy in the model (choose E_T) so that the turning point is deep in the incline. Therefore the heavy particle can be assumed to be confined in the classically allowed region $X < a$. The exact solution of the heavy particle equation (it can be found in [2]) is

$$F(X) = \frac{1}{\sqrt{k(X)}}\, v_\lambda(s(X)) \ ,$$

where

$$s(X) = \int_X^a dX'\ k(X') = \frac{2}{3}\sqrt{\frac{2rm_1M}{\hbar^2}} e^{i\pi}(a - X)^{\frac{3}{2}} \ . \tag{5.19}$$

We do not bother to write the form of v_λ explicitly[1]. It is enough to know that there are two independent solutions, corresponding to $\lambda = 1/6,\ 5/6$ and that in the region far away from the turning point v_λ reduces to a superposition of two WKB solutions:

$$v_\lambda \sim A\ e^{i\int_X^a dX' k(X')} + B\ e^{-i\int_X^a dX' k(X')}$$

but near the turning point it behaves like

$$v_\lambda \sim k^{3\lambda} \ .$$

Now we take into account the joining condition at the turning point. Since the heavy particle is deep in the incline, the wavefunction F can only be exponentially decaying in the region $X > a$. Then we know that sufficiently far in the classically allowed region $X < a$ the function F takes the form

$$F(X) = \frac{1}{\sqrt{k(X)}} \cos(-s(X) - \frac{\pi}{4}) \ .$$

This means that as the heavy particle propagates from some far away point to the turning point and back, the function F undergoes a phase change

$$\begin{aligned} \Delta\varphi(E_n) &= 2\ s_n(X) + \frac{\pi}{2} \\ &= \frac{4}{3\hbar}\sqrt{2m_1rM}(a(E_n) - X)^{\frac{3}{2}} + \frac{\pi}{2} \ . \end{aligned} \tag{5.20}$$

[1]see[2] pp. 134-137.

We have used a notation $\Delta\varphi(E_n)$ to emphasize that the accumulated phase depends on energy of the light particle through the location of the turning point (5.18). As (5.18) shows us, this dependence can become quite significant: if the slope of the heavy particle potential is very gentle (r very small), even small fluctuations in the light particle energy can lead to a great change location in the turning point and thus also in the accumulated phase $\Delta\varphi$. We get

$$\frac{\partial\Delta\varphi}{\partial E_n} = -\frac{2}{\hbar}\sqrt{\frac{2M}{m_1 r}}[a(E_n) - X]^{\frac{1}{2}} \; . \tag{5.21}$$

Is this behaviour something that we can reproduce in the semiclassical approximation? It turns out that in the present case we can, but not always. Let us discuss the present case first. We reanalyze the model using the semiclassical approximation and compare the phase behaviour to the exact solution. We rewrite the Schrödinger equation first as

$$\{\hat{H}_X + \hat{H}_{Xy}\}\ \Psi = \{-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial X^2} + MrX + E_T - \frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial y^2} + \frac{\kappa}{2}(y - \frac{r\mu}{\kappa})^2 - m_2 rX\}\ \Psi = 0 \; . \tag{5.22}$$

[This split makes the potential in $\hat{H}_X$ to be proportional to the same mass M which appears in the kinetic term.] An approximate solution is then

$$\Psi_{sc}(X,y) \approx \frac{1}{(-2V(X))^{\frac{1}{4}}}\ e^{\pm\frac{i}{\hbar}MS_0(X)}\ \chi(X,y) \; , \tag{5.23}$$

where

$$\begin{cases} S_0(X) = \frac{2}{3}\sqrt{2r}[a(0) - X]^{\frac{3}{2}} \\ a(0) \equiv -E_T/(Mr) \\ V(X) = r(X - a(0)) \end{cases} .$$

The function χ satisfies

$$i\hbar\frac{\partial S_0}{\partial X}\frac{\partial}{\partial X}\ \chi = \hat{H}_{Xy}\ \chi \; . \tag{5.24}$$

We define the WKB time using the classical trajectory of the center of mass:

$$t_{\mathrm{WKB}}(X) = \int_X^a \frac{dX'}{\partial S_0/\partial X} = \pm\sqrt{\frac{2}{r}}\ [a(0) - X]^{\frac{1}{2}} \; , \tag{5.25}$$

note that in the leading order semiclassical approximation the location of the turning point a is independent of the energy of the light particle. The overall sign choice will be chosen to give a negative (positive) t_{WKB} for a motion towards (away from) the turning point.

Next, we will assume that χ is an energy eigenstate with simple harmoning oscillator eigenenergy E_n, so that

$$\hat{H}_{Xy}\ \chi = (E_n - m_2 rX)\ \chi \; .$$

Using (5.24), (5.25) we can solve χ to time evolve as

$$\chi_n = e^{-\frac{i}{\hbar}E_n t_{\rm WKB}+\frac{i}{\hbar}m_2 r\int^{t_{WKB}} X(t')dt'}\, G_n(y) \, .$$

[Notice that a time evolution where X propagates to the turning point and back to the same point again is periodic. Therefore in general models there might be an additional Berry's phase in the exponent. However, in this case there is not.]

Now the total approximate wavefunction is

$$\Psi \sim e^{+\frac{i}{\hbar}\frac{2}{3}M\sqrt{2r}(a(0)-X)^{\frac{3}{2}}-\frac{i}{\hbar}E_n t_{\rm WKB}+\frac{i}{\hbar}m_2 r\int^{t_{WKB}} X(t')dt'}\, G_n \, . \tag{5.26}$$

We can see that the semiclassical result reproduces the phase behaviour of the exact solution in the WKB region. The first term is the same as (5.19) (recall that $M \approx m_1$). The third term is an overall phase factor that we will not care about. The second term can be identified with (5.21): far away from the turning point $\sqrt{a(0)-X} \approx \sqrt{a(E_n)-X}$, so (5.21) reduces to

$$\frac{\partial \Delta\varphi}{\partial E_n} \approx -\Delta t_{\rm WKB} \, .$$

Further, all higher derivatives $\partial^m \Delta\varphi/\partial E_n^m$ decay to zero. Thus, far away from the turning point there are no lasting effects from bouncing (other than the constant phase shift by $\pi/2$) and the semiclassical approximation is again sufficient.

This is *not* true in general. For a generic potential $U(X) = m_1 V(X)$, one can derive that

$$\frac{\partial \Delta\varphi}{\partial E_n} = -\frac{M}{\hbar^2}\int_X^{a(E_n)} \frac{dX'}{k(X', a(E_n))}$$

and compare this with the WKB time

$$t_{\rm WKB} = \frac{M}{\hbar}\int_X^{a(0)} \frac{dX'}{k(X', a(0))} \, .$$

Sufficiently far away from the turning point

$$k(X, a(E_n)) \approx k(X, a(0))$$

so the first derivative of the phase always reduces to the semiclassical form

$$\frac{\partial \Delta\varphi}{\partial E_n} \to \frac{1}{\hbar}\,\Delta t_{\rm WKB} \, . \tag{5.27}$$

However, if we look at the second derivative, it does not always decay to zero but can instead approach a *nonzero* value. This means that there is a leftover phase factor which is of second order in the light particle energy. Therefore different energy levels may time evolve in a more complicated way than can be seen using the semiclassical approximation without properly taking into account energy dependent fluctuations in the position of the turning point. In general the situation is then as follows. Looking at an exact quantum mechanical solution, a bounce from a turning point creates initially a complicated relative phase dependence for states corresponding to different energy levels of the light particle. After a sufficiently long time the relative phase shifts clear out and the first order energy dependence reduces to the usual one which could be recovered from a semiclassical treatment. However, the bounce from the turning point may create additional higher order dependence so the semiclassical approximation is insufficient to capture the phase correlations of the system.

It is not hard to find example cases where the higher order corrections can be very large. Consider for example the inverted harmonic oscillator potential $V(X) = -bX^2$. We will consider the region $x < 0$ and add an infinite wall at $X = -\varepsilon$ (with ε infinitesimal) to again avoid tunneling effects and an to keep the time of flight to the turning point finite. The light particle energy dependent turning point is at $a(E_n) = -\sqrt{E_n/(bM)}$ and the wavenumber is

$$k(X, a(E_n)) = \frac{M}{\hbar} \sqrt{2b\ (X^2 - a^2(E_n))} \ .$$

In the semiclassical approximation, the wavenumber is

$$k(X, a(0)) = \frac{M}{\hbar} \sqrt{2b\ (X^2 - a^2(0))}$$

with $a(0) = -\varepsilon$. Now we get

$$\frac{\partial \Delta\varphi}{\partial E_n} = \frac{1}{\hbar\sqrt{2b}} \cosh^{-1}\left(\sqrt{\frac{\varepsilon^2 bM}{E_n}} \cosh(\sqrt{2b}\ t_{\mathrm{WKB}})\right) \ .$$

This is defined after a time $\Delta t_{\mathrm{WKB}} > t_*$ where

$$t_* \sim \frac{1}{\sqrt{2b}} \cosh^{-1}\left(\sqrt{\frac{E_n}{\varepsilon^2 bM}}\right) \ .$$

Then the first derivative quickly reduces to (5.27). However, for the second derivative we get

$$\frac{\partial^2 \Delta\varphi}{\partial E_n^2} = \frac{1}{2\hbar\sqrt{2b}E_n} \frac{X}{\sqrt{X^2 - \frac{E_n}{bM}}}$$

then, far away from the turning point it tends to a finite value

$$\frac{\partial^2 \Delta\varphi}{\partial E_n^2} \to \frac{1}{2\hbar\sqrt{2b}E_n} \ .$$

If the potential well slopes very gently (b small), this residual energy dependence in the phase factor can be quite significant, and leads to additional interference effects between different energy eigenstates. This is the 'imprint' in the wavefunction from the turning point which we mentioned in the introduction. Notice that the last formula can be rewritten as

$$\frac{\partial^2 \Delta\varphi}{\partial E_n^2} \sim \frac{\sqrt{-V''(a(E_n))}}{\hbar M (V'(a(E_n)))^2} \tag{5.28}$$

omitting irrelevant numerical factors. Then we can see that for a general potential, if the turning point is near a local maximum, there will be a significant residual energy dependece which is given by the formula (5.28). This is the main result of this section.

5.3 Example in Minisuperspace Quantum Cosmology

Now that we have seen that turning points may spoil the the simple phase correlations of the semiclassical approximation, we would like to investigate an example in quantum gravity. We discuss a simplified quantum cosmological model of quantum matter evolving in a closed universe and check whether the turning point effect found in the previous section would spoil the semiclassical approximation in this case. It would be quite surprising if this should happen in the case of a macroscopic universe. At least in our simplistic model example this will not however be the case.

vvvWe start with

$$S = S_{\mathbf{grav.+cl.matter}} = \int d^4x \sqrt{-g} \left(\frac{R}{16\pi G} - \rho_m\right) ,$$

where ρ_m is a function related to the energy density of the pertinent classical matter and we will specify it later. The first term in the action is the Einstein-Hilbert action of classical gravity and G is the Newton's constant.

Using a spherically symmetric ansatz for the metric,

$$\begin{aligned} ds^2 &= \sigma^2 [N_0^2(t) dt^2 - a_0^2(t) d\Omega_3^2] \\ &\equiv N^2(t) dt^2 - a^2(t) d\Omega_3^2 , \end{aligned} \tag{5.29}$$

where N, a have dimensions of length, t is dimensionless and $\sigma^2 = 2G/3\pi$, the action reduces to

$$\begin{aligned} S &= \int dt L , \\ L &= \frac{N}{2\sigma^2} \{a(1 - \frac{\dot{a}^2}{N^2}) - 4\pi^2 \rho_m a^3\} . \end{aligned} \tag{5.30}$$

We will now choose ρ_m to satisfy

$$4\pi^2 \rho_m a^4 \equiv C^2 = \text{constant}$$

which means that ρ_m is the energy density of a classical radiation fluid filling the universe. The constant C has the dimension of length. Now the Lagrangian becomes

$$L = \frac{N}{2\sigma^2} \{a(1 - \frac{\dot{a}^2}{N^2}) - \frac{C^2}{a}\}$$

and, varying with respect to N, we get a classical equation of motion

$$1 + \frac{\dot{a}^2}{N^2} - \frac{C^2}{a^2} = 0 \; .$$

The model is time reparametrization invariant, which means that we need to make a gauge choice and choose a particular N. We choose

$$N = \sigma \equiv \frac{1}{M} \; ,$$

the latter notation is useful since M is essentially the Planck mass, $M \sim m_{\text{Planck}}$. Likewise, σ is the Planck length $\sigma \sim l_{\text{Planck}}$. The classical equation of motion corresponds to the dynamics of a closed radiation filled Robertson-Walker cosmology, with the solution

$$a = \sqrt{C^2 - (\frac{t}{M})^2} \; .$$

We chose the origin of time so that the expansion of the universe starts at $t = -MC$, it reaches the maximum size $a_{\text{max}} = C$ at $t = 0$, after which it recollapses to zero size at $t = +MC$. Since $MC \approx a_{\text{max}}/l_{\text{Planck}}$, a macroscopic size universe has a macroscopic lifetime.

If the model is quantized, we obtain the Wheeler–de Witt equation

$$\frac{1}{2}\{-\frac{1}{aM^2}(\frac{\partial^2}{\partial a^2} + \frac{\gamma}{a}\frac{\partial}{\partial a}) + M^2 a(1 - \frac{C^2}{a^2})\} \; \Psi(a) = 0 \; ,$$

where $0 \leq \gamma \leq 1$ is a free parameter reflecting an operator ordering ambiguity in the kinetic term. However, the γ-dependent term is known to be irrelevant in the WKB regime, so we will set $\gamma = 0$ henceforth. We rewrite the WdW equation then as

$$\hat{\mathcal{H}}_{\text{gr+c.m.}} \; \Psi(a) = \frac{1}{2}\{\frac{1}{aM^2}\hat{p}^2 + \frac{M^2}{a}U(a)\} \; \Psi(a) = 0 \; ,$$

where $\hat{p} = -\partial/\partial a$ and $U(a) = C^2 - a^2$.

Let us now include additional matter in the model. We assume that the energy of the additional matter is much lower than that of the classical radiation fluid which was already included in $\hat{\mathcal{H}}_{\mathrm{gr+c.m.}}$. We will then treat the additional matter quantum mechanically and consider it to propagate in the background of the expanding/recollapsing universe. The WdW equation is now

$$\{\hat{\mathcal{H}}_{\mathrm{gr.+c.m.}} + \hat{\mathcal{H}}_{\mathrm{qu.matter}}\}\ \Psi = 0\ .$$

We will not be more specific about what $\hat{\mathcal{H}}_{\mathrm{qu.matter}}$ is. For our simple model calculation it is sufficient to choose Ψ to be an eigenstate of $\hat{\mathcal{H}}_{\mathrm{qu.matter}}$,

$$\hat{\mathcal{H}}_{\mathrm{qu.matter}}\ \Psi = EM^2\ \Psi\ ,$$

writing the eigenenergy in a way that will be convenient later. Here E has the same dimension as length. Our assumption of the quantum matter being much less energetic than the radiation fluid means that $E << C$.

As in the quantum mechanics examples, we are interested in the phase behaviour of Ψ. The potential energy part of $\hat{\mathcal{H}}_{gr.+c.m.}$ has a turning point at the maximum size of the universe. After adding quantum matter, the location of the turning point (*i.e.*, the maximum size of the universe) depends on E. The new classical turning point is at

$$a_{\mathrm{max}}(E) = \sqrt{C^2 + E^2} - E$$

and the classically allowed region is $a \leq a_{\mathrm{max}}(E)$. We know that in the tunneling region the solution can only be exponentially decaying since the tunneling region is infinite. Thus, from the WKB joining conditions we know that far away in the classically allowed region the wavefunction will reduce to a superposition

$$\Psi \sim e^{iS+i\frac{\pi}{4}} + e^{-iS-i\frac{\pi}{4}}$$

where

$$S = \frac{1}{\hbar}\int_a^{a_{\mathrm{max}}(E)} da' p(a')$$

with the momentum

$$p(a) = M^2\ \sqrt{C^2 - 2aE - a^2}\ .$$

Thus

$$S = \frac{M^2}{2\hbar}\ (C^2+E^2)\ \{\frac{\pi}{2} - \sin^{-1}(\frac{a+E}{\sqrt{C^2+E^2}}) - (\frac{a+E}{\sqrt{C^2+E^2}})\sqrt{1-(\frac{a+E}{\sqrt{C^2+E^2}})^2}\}\ .$$

We will now consider the time behaviour of the first two derivatives vvof S with respect to EM^2. The scale factor a of the universe is related to the WKB time by

$$\frac{a^2}{C^2} = 1 - (\frac{t}{CM})^2 \equiv 1 - \tau^2\ .$$

The first derivative is

$$\frac{\partial S}{\partial(EM^2)} = C \ \{\varepsilon[\frac{\pi}{2} - \sin^{-1}(\frac{\varepsilon + \sqrt{1-\tau^2}}{\sqrt{1+\varepsilon^2}})] - \sqrt{\tau^2 - 2\varepsilon\sqrt{1-\tau^2}}\} \ ,$$

where $\varepsilon = E/C$ is very small. As we can see in Figure 1., the time dependence of the first derivative becomes very quickly linear after an expected troublesome period (where the semiclassical approximation is not valid) near the turning point.

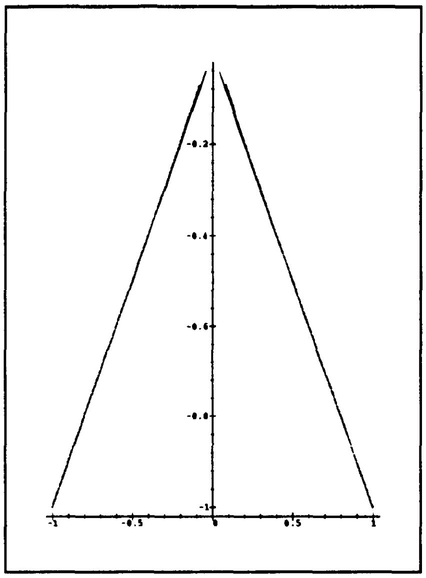

Figure 1: Time dependence of the first derivative of S, plotted with two parametrizations $\varepsilon = 0.00001$ and $\varepsilon = 0.001$.

The second derivative is

$$\frac{\partial^2 S}{\partial(EM^2)^2} = \frac{1}{M^2}\{\frac{(1+2\varepsilon^2)\sqrt{1-\tau^2}}{(1+\varepsilon^2)\sqrt{\tau^2 - 2\varepsilon\sqrt{1-\tau^2}}} + \frac{\pi}{2} - \sin^{-1}(\frac{\varepsilon + \sqrt{1-\tau^2}}{\sqrt{1+\varepsilon^2}})\} \ .$$

As Figure 2. shows, around the turning point the second derivative decays into a constant value. However, the constant value is not equal to zero.

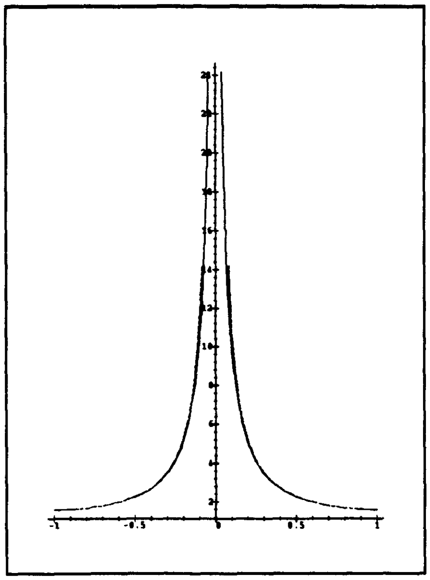

Figure 2: Time dependence of the second derivative of S, plotted with two parametrizations $\varepsilon = 0.00001$ and $\varepsilon = 0.001$.

The key fact is that the nonzero end value is of the order $1/M^2(\approx 1/m^2_{\rm Planck})$ and is thus negligibly small as long as the energy of the quantum matter is much less than than Planck energy. [This also means that the decay in the figure is much faster than it seems, since the units vertical axis are in $1/M^2$.] Thus the semiclassical approach, which does not take into account the quantum matter energy dependence in the location of the turning point, is accurate enough in our cosmological example. It the quantum matter energy would be $\sim M^2$, so that the residual second derivative would be significant, it would mean that we need to take into account its backreaction to the spacetime geometry itself. Then it would be as significant as the radiation fluid which we had included into the gravitational action, and our model would cease to make any sense. In short, it seems that the residual second order energy dependence in the phase factor which we have discovered is related to the issue whether the backreaction of the quantum subsystem to the 'classical' subsystem is significant or not.

5.4 Turning Points and Black Hole Evolution in Dilaton Gravity

Models of 1+1 dimensional gravity have been around for quite a while [3]. More recently, models with black hole like solutions have become popular. We are especially interested in the black hole spacetimes which are vacuum solutions in the CGHS model. This model has recently been studied in the framework of Dirac quantization and solutions to the Wheeler–De Witt equation have been found [4, 5]. [The equivalence of the solutions of [4] and [5] is shown in [6].] Consider all hypersurfaces in a black hole spacetime. The Wheeler–de Witt equation describes evolution from one hypersurface to the next one. In such an evolution, turning points may appear as we shall show in this section. We will examine the behaviour of the (pure gravity) wavefunction in the vicinity of a turning point. We need to identify the classically allowed region where the wavefunction is oscillatory. We will also investigate the possibility of having exponentially decaying (or growing) behaviour in classically disallowed regions. In other words, we need to study if tunneling issues can arise.

Here we will be using the results presented in [4, 7]. [For additional discussion, see [8].] We start with a brief review. Consider a class of two dimensional dilaton gravity models of the form

$$S = \int d^2x\sqrt{-g}[\frac{1}{2}g^{\alpha\beta}\partial_\alpha\chi\partial_\beta\chi - V(\chi) + D(\chi)R] \ . \tag{5.31}$$

Here χ is the dilaton scalar field and $(g_{\alpha\beta})$ is a two dimensional metric. The action can be rewritten in a form where the kinetic term disappears. First rescale the metric

$$(g_{\alpha\beta}) = \Omega^{-2}(\bar{g}_{\alpha\beta})$$

with a rescaling factor Ω satisfying the equation

$$-1 + 4\frac{dD}{d\chi}\frac{d\ln\Omega}{d\chi} = 0 \ ,$$

then perform a field redefinition

$$\chi \to \bar{\phi} = D(\chi)$$

and introduce a new potential

$$\bar{V}(\bar{\phi}) = \frac{V(\chi(\bar{\phi}))}{\Omega^2(\chi(\bar{\phi}))} \ .$$

As a result, the action (5.31) becomes

$$S = \int d^2x\sqrt{-\bar{g}}[\bar{\phi}\bar{R} - \bar{V}(\bar{\phi})] \ . \tag{5.32}$$

We need to relate (5.31),(5.32) to the CGHS dilaton gravity action

$$S = \int d^2x\sqrt{-g}\ e^{-2\phi}\ [R + 4(\nabla\phi)^2 + 4\lambda^2] \ . \tag{5.33}$$

The actions (5.31),(5.32),(5.33) can be related with the following rules:

$$\begin{aligned} D(\chi) &= \frac{1}{8}\chi^2 = \bar{\phi} = e^{-2\phi} \\ \Omega^2(\chi) &= \chi^2 \\ \bar{V}(\bar{\phi}) &= -\frac{\lambda^2}{2} . \end{aligned}$$

The metrics $g, \bar{g}$ can be parametrized as follows:

$$\begin{aligned} g_{\alpha\beta}dx^\alpha dx^\beta &= \Omega^{-2}\bar{g}_{\alpha\beta}dx^\alpha dx^\beta \\ &= \Omega^{-2}e^{2\bar{\rho}}\left[-N(dx^0)^2 + (dx^1 + N_\perp dx^0)^2\right] , \end{aligned} \tag{5.34}$$

where $N, N_\perp$ are the lapse and shift functions and $x^0 = (x^+ + x^-)/2, x^1 = (x^+ - x^-)/2$ are the time and the space components of the Kruskal coordinates. In the Kruskal gauge ($N = 1, N_\perp = 0$) there are then additional relations

$$\phi = \rho = \bar{\rho} - \ln\chi = \bar{\rho} + \phi - \ln\sqrt{8} .$$

Using the parametrization (5.35) in (5.32) one can rewrite the action in a form from which it is easy to read out the canonical momenta

$$\begin{aligned} \Pi_{\bar{\phi}} &= \frac{2}{N}(N_\perp\bar{\rho}' + N_\perp' - \dot{\bar{\rho}}) \\ \Pi_{\bar{\rho}} &= \frac{2}{N}(-\dot{\bar{\phi}} + N_\perp\bar{\phi}') \\ \Pi_N &= \Pi_{N_\perp} = 0 . \end{aligned}$$

The Hamiltonian is the sum of the time reparametrization and spatial diffeormorphism constraints

$$H = \int dx^1 \left[N\mathcal{H} + N_\perp\mathcal{H}_\perp\right]$$

where

$$\begin{aligned} \mathcal{H} &= 2\bar{\phi}'' - 2\bar{\phi}'\bar{\rho}' - \frac{1}{2}\Pi_{\bar{\rho}}\Pi_{\bar{\phi}} + e^{2\bar{\rho}}\bar{V}(\bar{\phi}) \\ \mathcal{H}_\perp &= \bar{\rho}'\Pi_{\bar{\rho}} + \bar{\phi}\Pi_{\bar{\phi}} - \Pi_\rho' . \end{aligned}$$

The time reparametrization (or, the Hamiltonian) constraint $\mathcal{H} = 0$ can be written as a Hamilton-Jacobi equation

$$\frac{1}{2}\left[\frac{\partial S}{\partial\bar{\phi}}\frac{\partial S}{\partial\bar{\rho}} - g[\bar{\phi}, \bar{\rho}]\right] = 0$$

where

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

It was found [4] that this is solved by a functional

$$S[\bar\phi,\bar\rho,\mathcal{C}] = \int dx^1 \{Q_{\mathcal{C}} + \bar\phi' \ln[\frac{2\bar\phi' - Q_{\mathcal{C}}}{2\bar\phi' + Q_{\mathcal{C}}}]\} \tag{5.35}$$

which depends on an integration constant $\mathcal{C}$ through the functional

$$Q_{\mathcal{C}} = 2\sqrt{(\bar\phi')^2 + (\mathcal{C} + j(\bar\phi))e^{2\bar\rho}} \tag{5.36}$$

where $j(\bar\phi)$ is given by $dj(\bar\phi)/d\bar\phi = \bar V(\bar\phi)$. For the CGHS model,

$$j(\bar\phi) = -\frac{\lambda^2}{2}\bar\phi\ .$$

When the theory is quantized using the Dirac quantization approach in the Schrödinger representation, the Hamiltonian constraint becomes the WdW equation. As usual, the operator ordering is ambiguous. In [4] a specific operator ordering was used to write the WdW equation as

$$\hat{\mathcal{H}}\ \Psi(\bar\phi,\bar\rho) = \frac{1}{2}(g - Q_{\mathcal{C}}\ \hat\Pi_{\bar\phi}\ Q_{\mathcal{C}}^{-1}\ \hat\Pi_{\bar\rho})\ \Psi(\bar\phi,\bar\rho) = 0\ . \tag{5.37}$$

This equation is solved by a wavefunctional $\Psi = \exp(\pm iS)$ where S is the solution (5.35) of the H-J equation. It is interesting that the exact solution is similar to the WKB solution. However, this seems to be tied with the specific choice of operator ordering. As we shall show later, $Q_{\mathcal{C}}$ tends to zero at a turning point, so (5.37) becomes singular. It seems possible that a wavefunctional which would be a solution to the WdW equation with a nonsingular operator ordering at a turning point could differ from the WKB solution at the turning point. Then the situation would be similar to the case of the models which we discussed earlier.

The integration constant $\mathcal{C}$ can be expressed as a functional of the canonical variables $\bar\rho, \bar\phi, \Pi_{\bar\rho}, \Pi_{\bar\phi}$, using the relations

$$\Pi_{\bar\phi} = \frac{\delta S}{\delta\bar\phi} = \frac{g}{Q_{\mathcal{C}}} \tag{5.38}$$

$$\Pi_{\bar\rho} = \frac{\delta S}{\delta\bar\rho} = Q_{\mathcal{C}}\ . \tag{5.39}$$

The result is

$$\mathcal{C} = e^{-2\bar\rho}\ (\frac{1}{4}\Pi_{\bar\rho}^2 - (\bar\phi')^2) - j(\bar\phi)\ . \tag{5.40}$$

We will now relate $\mathcal{C}$ to the mass M of a black hole in this special case. In the CGHS model (with the respective field definitions) the black hole solution is given by[2]

$$e^{2\rho} = e^{2\phi} = M - x^+x^-\ . \tag{5.41}$$

[2]We will use $\lambda = 1$ henceforth.

We can express $Q_{\mathcal{C}}$ and $\Pi_{\bar{\rho}}$ in the CGHS fields:

$$Q_{\mathcal{C}} = 4e^{-2\phi}\sqrt{(\phi')^2 + (2\mathcal{C}e^{2\phi} - 1)e^{2\rho}} \tag{5.42}$$

and

$$\Pi_{\bar{\rho}} = -2\dot{\bar{\phi}} = 4e^{-2\phi}\dot{\phi} \ ,$$

we are using the Kruskal gauge $N = 1, N_{\perp} = 0$. From (5.41) we find $\phi', \dot{\phi}$. Substituting these to $Q_{\mathcal{C}}, \Pi_{\bar{\rho}}$ and using (5.39) we can finally identify

$$\mathcal{C} = \frac{M}{2} \ . \tag{5.43}$$

Now we are all set to discuss the existence of turning points. Let us consider all different ways of foliating the black hole spacetime with hypersurfaces. It can be shown [9] that the foliations fall into two basic categories. In the case of an eternal black hole, the two categories have a simple description:

1. hypersurfaces which cross the event horizon of the black hole, we call surfaces in this category *ingoing hypersurfaces*

2. hypersurfaces which avoid the black hole and cross the event horizon of the white hole, these will be called *external hypersurfaces*

A boundary case is given by the surfaces which do neither. In a fluctuating black hole spacetime these surfaces are pathological in a certain sense [9] so we will not promote this case as a separate (third) category. As an example of this division into categories, consider a class of hypersurfaces Σ_γ which are straight lines in the Kruskal coordinates

$$\Sigma_\gamma \ : \quad x^- = -\alpha^2 x^+ + \gamma \ .$$

For any fixed slope α^2, it is evident that $\gamma > 0$ ($\gamma < 0$) correponds to external (ingoing) hypersurfaces. The boundary case $\gamma = 0$ surfaces correspond to constant Schwarzschild time slices

$$\sigma^0 = \frac{1}{2}(\sigma^+ + \sigma^-) = -\ln\alpha \ .$$

It is apparent that there is a one-to-one correspondence between the ingoing and external hypersurfaces. This is a result of the time reversal symmetry in the eternal black hole spacetime. Time reversal reverses the roles of black and white holes and therefore maps the external to ingoing surfaces and vice versa. We have previously discussed the role of hypersurfaces (or, 1-geometries) in the black hole context and described how they are used to give a notion of time in studying the evolution of quantum matter in a fluctuating black hole spacetime background. We found that if we use only external surfaces to foliate the spacetime, the quantum matter states will become very sensitive to subplanckian fluctuations in the spacetime and their time evolution is thus ill-behaved. On the other hand, we could proceed in the opposite way and consider the backward time evolution of a final matter state using a foliation with ingoing hypersurfaces. In this case the matter state becomes more sensitive to the background fluctuations as it evolves towards earlier times [9], so this leads to an equally bad time evolution for the matter state. A natural question is then to ask if these problems can be overcome if one considers a time evolution which starts with surfaces in one category (external) and crosses over to the other category (ingoing) somewhere in the middle. Since in the previous cases the matter state is well behaved in the asymptotic regions, it might be that all sensitivity is confined to a time interval in the middle, with no sensitivity remaining in final result of the time evolution. This would mean that the semiclassical approximation after all sufficiently describes the black hole evolution, up to the last stages of its lifetime.

We will now discuss the case of crossing from one category to the other and show that this is equal to going over a turning point. More precisely, the boundary of the two categories forms a turning point in the evolution. Let us first rewrite the surfaces Σ_γ as

$$x^0 = \frac{1-\alpha^2}{1+\alpha^2}x^1 + \frac{\gamma}{1+\alpha^2} \, .$$

Then

$$e^{-2\phi} = M + \frac{4\alpha^2(x^1)^2 - 2\gamma(1-\alpha^2)x^1 - \gamma^2}{(1+\alpha^2)^2}$$

and we find that

$$\phi'(\phi) = \frac{d\phi}{dx^1}(\phi) = \pm\frac{2e^{2\phi}}{1+\alpha^2}\,\sqrt{\gamma^2 + \alpha^2(e^{-2\phi} - M)} \tag{5.44}$$

for a Σ_γ hypersurface. [The overall sign choice is arbitrary, we will choose it to be negative.] It is evident that this quantity has a global maximum at the boundary case $\gamma = 0$. We can now calculate how the momenta (5.38,5.39) depend on the surfaces Σ_γ. We already found what Q_C looks like in the CGHS notation, and similarly

$$g = 8e^{-2\phi}\,\left[(\phi')^2 - \phi'' + \phi'\rho' - e^{2\rho}\right] \, .$$

When we foliate the spacetime with Σ_γ, we should make the gauge choice

$$e^{2\rho} = \frac{4\alpha^2}{(1+\alpha^2)^2}\,e^{2\phi}$$

(or, $N = \frac{2\alpha}{1+\alpha^2}, N_\perp = 0$) instead of $\rho = \phi$ (which corresponds to a foliation with surfaces x^0=const. and the Kruskal gauge choice $N = 1, N_\perp = 0$). Substituting this relation to Q_C, g along with (5.44), yields

$$g \equiv 0 \quad , \quad Q_C = \frac{8|\gamma|}{1+\alpha^2} \tag{5.45}$$

so

$$\begin{aligned} \Pi_{\bar\rho} &= \frac{8|\gamma|}{1+\alpha^2} \\ \Pi_{\bar\phi} &= 0. \end{aligned} \tag{5.46}$$

The latter equation is consistent with $\Pi_{\bar\phi} = -2(\dot\rho - \dot\phi) = 0$. The first equation shows that as γ changes from negative to positive (crossing the boundary of categories), the momentum $\Pi_{\bar\rho}$ decreases to zero at the boundary and then increases again. Thus the boundary of the two categories is a turning point in dilaton gravity.

Let us now examine how the wavefunction behaves near the turning point. We rewrite S in terms of ϕ, ρ:

$$S = \int dx^1 \; \{Q_C - 2e^{-2\phi}\phi' \ln[\frac{-4e^{-2\phi} - Q_C}{-4e^{-2\phi} + Q_C}]\}$$

and change the integration variable from x^1 to ϕ. We will first consider only the outer parts the hypersurfaces and keep the upper limit of integration at some finite value ϕ_0. Then

$$S = \int_{-\infty}^{\phi_0} d\phi \; \{\frac{Q_C}{\phi'} - 2e^{-2\phi} \ln[\frac{-4e^{-2\phi} - Q_C}{-4e^{-2\phi} + Q_C}]\} \; . \tag{5.47}$$

Now we substitute ϕ' from (5.44) and Q_C from (5.45) and change again the integration variable to

$$\eta = \sqrt{\gamma^2 + \alpha^2(e^{-2\phi} - M)} \; .$$

Now we get

$$S = \frac{2}{\alpha^2} \int_{\infty}^{\eta_0} d\eta \; \{2|\gamma| + \eta \ln[\frac{\eta - |\gamma|}{\eta + |\gamma|}]\} \; .$$

The contribution from the infinite end of integration will be zero. We end up with

$$S = \frac{1}{\alpha^2} \; \{2|\gamma|\eta_0 + (\eta_0^2 - \gamma^2) \ln[\frac{\eta_0 - |\gamma|}{\eta_0 + |\gamma|}]\} \; . \tag{5.48}$$

Now, suppose that we had restricted the hypersurfaces to the external region outside the event horizons, where $e^{-2\phi} \geq M$. Then we should take the integration limit $\eta_0 \to |\gamma|$, which yields

$$S = 2\frac{\gamma^2}{\alpha^2} \; .$$

In this case then the wavefunctional remains purely oscillatory throughout. However, as was pointed out already in [7], the wavefunctional can become exponentially decaying (or growing) if the first term Q_C in (5.47) becomes imaginary, or in the second term the logarithm becomes imaginary. Since we now know that there is a turning point in the evolution, the crucial issue is to understand what would be the joining conditions for the wave functional at the turning point. In the case of minisuperspace closed cosmology, we knew that the wavefunction could only be exponentially decaying in the tunneling region $a > a_{\mathbf{max}}$. Further, that model reduced essentially to ordinary one dimensional quantum mechanical problem where tunneling and joining conditions are well understood. However, the present situation is infinite dimensional, the degrees of freedom are hypersurfaces (one-geometries) and thus we are dealing with a function space. Moreover, we do not yet know what would the 'potential' look like in this function space and we do not know what the joining conditions would look like in the present setting. Still, it is possible to make a few interesting observations.

1. One possibility is that at the turning point the hypersurfaces tunnel into the direction of one-geometries which can only fit in spacetimes of more massive black holes. More precisely, considers the direction of hypersurfaces with

$$\phi'(\phi) = \frac{d\phi}{dx^1}(\phi) = \pm\frac{2e^{2\phi}}{1+\alpha^2}\,\sqrt{\gamma^2+\alpha^2(e^{-2\phi}-\bar{M})}$$

where $\bar{M} \neq M$. For these surfaces, we find that

$$Q_C = \frac{8}{1+\alpha^2}\,\sqrt{\gamma^2+\alpha^2(M-\bar{M})}\ .$$

Thus, around the turning point $\gamma = 0$ the wavefunction is oscillatory towards the direction of hypersurfaces in a black hole spacetime with lower mass ($\bar{M} < M$). On the other hand, as $\gamma = 0$ and $\bar{M} > M$, the wavefunction could become exponentially decaying (growing). However, the exponent of the integral is actually

$$iS = \pm\frac{8\alpha\sqrt{\bar{M}-M}}{1+\alpha^2}\,\int_{-\infty}^{\phi_0}\frac{d\phi}{\phi'} + \text{(oscillatory terms)}$$

where the integral is infinite.

2. Another more intriguing possibility is that the surfaces start to penetrate into the region behind the event horizon. This causes the logarithmic term in S become imaginary. This possibility may allow a finite tunneling direction, investigations of this are underway.

Bibliography

[1] For recent reviews on the topic, see for example 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); C. Isham, *Canonical Quantum Gravity and the Problem of Time*, Lectures presented at the NATO Advanced Summer Institute "Recent Problems in Mathematical Physics", Salamanca, June 15-27, 1992; K. Kuchař, *Time and Interpretations of Quantum Gravity*, in *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*, eds. G. Kunstatter, D. Vincent and J. Williams (World Scientific, Singapore, 1992).

[2] E. Merzbacher, *Quantum Mechanics*, Second edition, John Wiley & Sons (1970), Chapter 7.

[3] C. Teitelboim, *Phys. Lett.* **B126** (1983) 41 and "The Hamiltonian Structure of Two-Dimensional Space-Time and its Relation with the Conformal Anomaly", in *Quantum Theory of Gravity*, S. Christensen, ed. (Adam Hilger, Bristol, 1984); R. Jackiw, "Liouville Field Theory: A Two-Dimensional Model for Gravity?", in *Quantum Theory of Gravity*, S. Christensen, ed. (Adam Hilger, Bristol, 1984); R. Jackiw, *Nucl. Phys.* **B252** (1985) 343.

[4] D. Louis-Martinez, J. Gegenberg and G. Kunstatter, *Phys. Lett.* **B321** (1993) 193.

[5] D. Cangemi and R. Jackiw, *Phys. Rev.* **D50** (1994) 3913; *Phys. Lett.* **B337** (1994) 271.

[6] E. Benedict, *Phys. Lett.* B340 (1994) 43.

[7] J. Gegenberg and G. Kunstatter, *Phys. Rev.* **D47** (1993) R41292.

[8] G. Lifschytz, S. D. Mathur, and M. Ortiz, *A Note on the Semi-Classical Approximation in Quantum Gravity*, MIT Report No. MIT-CTP-2384 (gr-qc/9412040).

[9] E. Keski-Vakkuri and S. D. Mathur, unpublished.