Merging Black Hole and Cosmological Horizons

by

Alain Prat

B.Sc., The University of British Columbia, 2002

M.Sc., The University of British Columbia, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2015

Abstract

This thesis investigates the merging of horizons which occurs when a black hole crosses a cosmological horizon. We study the simplest spacetime which has both a black hole and cosmological horizon, namely Schwarzschild-deSitter (SdS) spacetime. First we develop a new coordinate system for SdS spacetime, which allows us to properly illustrate and analyze the merging of horizons. We then use a combination of numerical and analytical methods to study the structure of the merging horizons, including the null generators which make up the horizon, as well as the presence of caustic points on the horizon. We find an analytical formula for the location in spacetime where the black hole and cosmological horizon first touch. Next we study the area of the horizons. Using numerical methods, we find several intriguing results regarding the behavior of horizon area on time, and in the limit of small black hole mass. The first result is that the time at which the black hole first touches the cosmological horizon is also the time at which the rate of horizon area increase is maximal. The second and third results concern the horizon area in the limit of small black hole mass. The second result is that in this limit, all of the increase in horizon area occurs prior to horizon merger. The third and final result is that in the limit of small black hole mass, the increase in horizon area can be thought of as being due in equal parts to two effects: to the joining of new generators not previously on the horizon, and the expansion of generators on the horizon for all times. The first and third results just mentioned are both corroborated using analytical techniques. Finally, we conclude by discussing how the study of merging horizons in this thesis is a valuable first step to undertaking a similar study of the horizons which occur in merging black hole binaries.

Preface

This dissertation is original and independent work by the author, A. Prat. At the time of this writing, the results of this thesis have not been published.

Table of Contents

List of Figures

Acknowledgements

I want to thank my supervisor, family, colleagues and friends for creating the kind of environment that made the research in this thesis possible and enjoyable. First and foremost, my supervisor's positive attitude and affable nature helped me persevere through the more difficult moments of my doctoral work. My parents' support and encouragement was also of great help. Finally, my colleagues and friends in the department helped create a positive research environment that made the research all the more enjoyable. This work was partly supported by an NSERC PGS D scholarship.

Dedication

This thesis is dedicated to my late uncle Frédéric Prat, whose passion for science was inspirational.

Chapter 1

Introduction

1.1 Overview

This thesis investigates the merger of horizons that occurs when a black hole crosses an observer's cosmological horizon. Our main motivation is to use a simple spacetime with merging horizons as a mathematical laboratory where we can investigate various questions about the mathematics of the merging horizons. This thesis is broadly organized as follows: in section 1.2 we introduce the concept of a cosmological horizon, and discuss the motivation for considering the merger of a black hole with a cosmological horizon. In chapter 2 we review some mathematical properties of the spacetime (Schwarzschild deSitter spacetime; SdS spacetime from now on) which will be considered in this thesis and give a more precise fomulation of the problem statement considered here. In chapter 3 we introduce a new coordinate system for SdS spacetime, developed specifically for the purposes of this thesis. In chapters 4 and 5, we present the main results of this thesis: an analysis of the structure and area of the horizons that result from a merger of a black hole with a cosmological horizon. These results are obtained using both numerical and analytical methods. In chapter 6, we use the main results from the previous chapters and discuss how they could be related to the merging of horizons that occurs in binary black holes. As well, we summarize the results and discuss avenues for future research.

1.2 Merging black hole and cosmological horizons

The cosmological horizon The measured acceleration of the universe, as inferred from supernova data [24, 26], indicates that either our theory of gravity on large scales is in serious need of repair, or that the universe contains a mysterious fluid, often called dark energy, which is driving this accelerated expansion. The revision of cosmological models to include dark energy, typically in the form of vacuum energy or a cosmological constant, has led to striking revisions of our understanding of the universe. For example, constraints on cosmological parameters strongly suggest continued expansion of the universe [27], even if the energy density of the universe is greater than the critical density required to make it flat. Another striking feature of a universe with accelerated expansion is the presence of a *cosmological horizon*[1], which in a perfectly homogeneous universe is a sphere[2] surrounding any observer in spacetime such that beyond this sphere, no information can ever reach the observer.

The cosmological horizon is not to be confused with the *particle horizon*. The particle horizon separates objects that are sufficiently close (as measured using comoving distance) that the light they emitted in the past has had time to reach us, so that we may currently observe them. The particle horizon limits what we can currently observe; beyond the current particle horizon, distant objects have not yet come into view, so to speak. The cosmological horizon, on the other hand, limits which part of the universe will ever be accessible to a hypothetical eternal observer. Currently there are distant objects (galaxies and so on) so far that the light they emit will never reach us, no matter how long we would be willing to wait (see pages 128-129 of [15] for a careful discussion of particle and cosmological horizons for deSitter spacetime).

This inability to receive light from objects beyond the cosmological horizon is often described in terms of distant objects travelling faster than light. However, this "faster than light" description is

[1] For FRW universes accelerated expansion implies a cosmological horizon; for more general cosmological models one can find exceptions (See [21].)

[2] In inhomogeneous cosmological models the cosmological horizon can of course be nonspherical. However the topology of the cosmological horizon in these cases is nonetheless spherical.

incorrect for a number of reasons. First, the proper distance to distant objects is an inherently ambiguous concept since it depends on the choice of the time coordinate. Second, even in cases where one can define an unambiguous proper distance, such as in homogeneous FRW cosmological models, only in pure deSitter spacetime does the faster than light limit coincide with the cosmological horizon (see [5] for a discussion of this and other misconceptions).

The cosmological horizon is better understood in terms of stretching of space between an observer and distant objects. In cosmological models where the expansion of the universe is accelerating, the distance between an observer and a distant object may grow so quickly that even light will not traverse this great distance, even after an eternity. Another equally valid way of describing the cosmological horizon is in terms of infinite redshift. Currently there are objects whose light, when it finally reaches us in the far future, will have been stretched by the expansion of the universe into light of unfathomably long wavelengths. As one moves further away from these objects, one finds a point where objects would suffer an infinite redshift, and the light emitted by them would take an infinite amount of time to reach us. These objects lie at the boundary of our current cosmological horizon (see page 129 of [15] for a discussion of this infinite redshift in the context of deSitter spacetime).

Although one can only speak of objects being inside or outside the cosmological horizon at any one time, it is possible for an object to cross the cosmological horizon. For example, in a universe with a cosmological constant and a single object of negligible mass, we would essentially have deSitter spacetime. Thus the cosmological horizon would be at the same proper distance from the observer at all times, but the proper distance between the object and the observer would be increasing due to the expansion of the universe. Consequently, there would be a critical time where the object would cross the observer's cosmological horizon (page 129, [15]). On can similarly deduce that in a universe with a cosmological constant, an observer and a small black hole would move away from each other such that the black hole would effectively cross the cosmological horizon, and its horizon would merge with the cosmological horizon. The study of this merging of horizons and its consequences is the subject of this thesis.

The inability to access information beyond the cosmological horizon is reminiscent of the inability to access information inside a black hole (see page 300 of [32] for the definition of a black hole). In a universe with a cosmological horizon, the observable universe is separated from a possibly much larger universe by two types of horizons: the black hole horizons and the cosmological horizon. An important distinction between these two types of horizons is that the cosmological horizon is observer dependent, whereas the black hole horizons are thought of as observer independent. All observers outside a black hole should agree on the location of the black hole horizon, regardless of their location and trajectory in spacetime. On the other hand, each observer in a universe with accelerated expansion has a cosmological horizon centered on their location. We will come back to this distinction in section 2.4, where we will explain how to incorporate both types of horizons into a single definition of horizon. This will be crucial for exploring the merger of the two horizons.

Motivation Here we seek to address to address a number of questions about the horizon which results when a black hole crosses and merges with a cosmological horizon. This could be, for example, a supermassive black hole in the center of a distant galaxy, merging with what is currently our cosmological horizon. There are several motivations for considering such a merger. Our primary motivation for considering this merger is the similarity between it and the merger of horizons that occurs during a head on collision in a binary black hole merger (see [23] for a thorough numerical exploration of a head on binary black hole merger).

By studying the merger of black hole and cosmological horizons, we create a mathematical laboratory where we can investigate the structure and area of the horizons during the merger. The results obtained in this study then lead to natural questions about the merger of binary black hole horizons. An alternative approach to studying the structure and area of merging horizons is to use perturbation theory to construct the spacetime of an extreme mass-ratio binary black hole system (as in [14]), or to use a Rindler horizon approximation for the larger black hole horizon, as in [13].

The analysis of the structure and the area of the cosmological and black hole horizons are in chapters 4 and 5, respectively. Our analysis of the structure of the horizons will focus on the location and

structure of the caustic, as well as the structure of the horizon at late times. Our analysis of the horizon area can be seperated into coordinate dependent and independent results. The coordinate independent results include an analysis of the relative importance of different horizon generators to the final horizon area, as well as a characterization of the time of maximal area increase. The coordinate dependent results will include a quantitative analysis of the horizon area in the extreme mass ratio limit. Although these latter results are coordinate dependent, in the sense that the choice of time coordinate can affect the area, we will argue that we can make general statements about the qualitative behavior of the horizon area in the extreme mass ratio limit which should hold regardless of the coordinate system.

Although our main motivation is the application of our results to the horizons of extreme mass ratio binary black holes, there are other applications of our results as well. In the context of cosmology, examining the combined effect of both black hole and cosmological horizons allows us to give a more precise answer to the question: which part of the universe is in principle observable? Although both black holes and the cosmological horizon independently shroud parts of the universe from view and influence, a full understanding requires combining these two types of horizons.

A full understanding of the observable part of the universe is also related to the paradoxes (or perceived paradoxes) created by the presence of horizons. For example, many of the questions related to information loss in the context of black hole horizons have recently been extended to cosmological horizons as well (see for example, [10]). By considering a case where black hole and cosmological horizons meet, we are examining the horizons for a spacetime where these two sets of questions overlap.

Still in the context of cosmology, calculations of the total entropy of the observable universe rely on estimating the area of the cosmological horizon, as well as the area of the black hole horizons inside the cosmological horizon (see [7] for an example of such a calculation). In addition, calculations of the rate of change of entropy of the observable universe requires knowledge of the rate of change of horizon area as black holes merge with the cosmological horizon (see [31] for an example of calculating the rate of change of entropy for FRW cosmologies). Both of these types of calculations

have so far ignored the corrections due to distortions in the shapes of black hole and cosmological horizons as they approach one another. Here we provide the first step towards including such corrections.

Another motivation for the work in this thesis is the possibility of using a simple analytical spacetime with merging horizons as a useful analytical testbed for the numerical event horizon solvers used in binary black hole simulations. To our knowledge, the only other analytically known spacetime which has been shown to contain event horizon mergers is the exact solution known as the Kastor-Traschen solution [19]. This solution has been used as a test bed for a numerical event horizon solver in [4].

Our work is also related to the broader question of the topological transition in event horizons. Ever since the discovery of black string instability and pinch off in five dimensional spacetime [12], there has been interest in topological transitions of event horizons in higher dimensional spacetimes. For example, Emparan and Hassad [9] have studied the self-similar geometry at the intersection of a black hole and cosmological horizon merger for spacetimes with dimension greater than six. It should be noted that this merger is fundamentally different from the one studied in this thesis, since it occurs due to a changing parameter in the spacetime, as opposed to being due to the movement of an observer.

We address our basic questions regarding the merger of horizons by considering the simplest cosmological spacetime which has both a black hole and a cosmological horizon: Schwarzschild deSitter (SdS) spacetime. This choice of spacetime is motivated not only by the fact that the metric is known analytically, but also by the fact that it is a good approximation to the late time spacetime structure of our own universe, according to the Λ-CDM model.

Chapter 2

Preliminaries: Schwarzschild deSitter Spacetime

2.1 Introduction

In 1916, Karl Schwarzschild published the first analytical solution [29] to Einstein's newly developed field equations of gravitation [8]. This solution for the spacetime metric is the now famous Schwarzschild metric, which represents the vacuum spacetime geometry outside any spherically symmetric body. If the entire spacetime is considered to be devoid of matter, so that the energy-momentum tensor vanishes, and one uses so-called Schwarzschild coordinates to represent the metric, one obtains the well known line element

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2(\sin^2\theta d\phi^2 + d\theta^2)\,, \tag{2.1}$$

where $0 < r < \infty$. In the above line element, there is a coordinate singularity at $r = 2M$ and a curvature singularity at $r = 0$. In the decades that followed, it was gradually understood that there is an event horizon at $r = 2M$, and that Schwarzschild spacetime corresponds to the simplest black hole geometry one can imagine: a static non-rotating black hole in a vacuum spacetime. It was also understood that the curvature singularity at $r = 0$ represent a breakdown of the classical description of spacetime geometry.

Another historically important exact solution to Einstein's equation is the so-called deSitter spacetime, first discovered by Willem deSitter [6]. Mathematically, the deSitter spacetime is the maximally symmetric vacuum solution of Einstein's equation with a cosmological constant. One commonly used coordinate system is the so-called static coordinates of deSitter spacetime, in which the line element takes the form

$$ds^2 = (1-Hr)\,dt^2 - (1-Hr)^{-1}\,dr^2 - r^2(\sin^2\theta d\phi^2 + d\theta^2)\,, \tag{2.2}$$

where $H = \sqrt{\Lambda/3}$ is the Hubble constant and $0 < r < \infty$. Provided that one ignores the coordinate singularity at $r = 1/H$ in the above, the static coordinates can be said to cover the whole of the deSitter manifold. The physical interpretation of deSitter spacetime is obtained by considering homogeneous spacelike hypersurface slicings of the deSitter manifold. This results in an FRW cosmological model for a universe devoid of matter, but with a cosmological constant driving the expansion of space. One can obtain different cosmologies depending on the choice of spatial slices. For example, if one uses the so-called "closed" slicing of deSitter spacetime, one can interpret the full deSitter manifold as a so-called "big bounce" universe. This is an FRW cosmology with spatially homogeneous constant time slices of spherical geometry, and with a scale factor that first shrinks to a minimum value and then grows indefinitely. It is from this behavior of the scale factor that the big bounce universe gets its curious name (the name is in keeping with the names "big bang" and "big crunch"). Another cosmological model is the so-called deSitter universe or steady-state universe. This is also an FRW cosmology, where the spatial geometry of the constant time slices is flat and Euclidean and the scale factor increases exponentially in time. In the so-called planar coordinates of the deSitter universe, the line element takes the form

$$ds^2 = dt^2 - e^{2Ht}\left(dx^2 + dy^2 + dz^2\right)\,, \tag{2.3}$$

where t is the cosmic time and (x,y,z) are the Euclidean coordinates on the flat slices. It is this latter cosmological model that will be of interest to us in this thesis. As discussed below, we will consider a part of Schwarzschild deSitter spacetime which can be interpreted as a black hole embedded in a deSitter universe.

Schwarzschild deSitter (SdS) spacetime is a generalization of both Schwarzschild and deSitter spacetime, and encompasses both as special cases. It can be thought of as the generalization of Schwarzschild spacetime obtained when one allows for a positive cosmological constant in Einstein's equations. Thus it can be interpreted as the spacetime geometry of the simplest black hole in the presence of the cosmological constant. Like Schwarzschild spacetime, it is a static and spherically symmetric spacetime. A well-known but nevertheless remarkable fact about Schwarzschild spacetime that it is the unique vacuum spherically symmetric spacetime, as encapsulated by Birkhoff's theorem [3]. It is possible to generalize this theorem to SdS spacetime and include the cosmological constant (see [28] for a proof of the theorem). The static and spherically symmetric nature of SdS spacetime is captured by its four Killing vectors, one of which is associated with time and three of which are associated with spherical symmetry. The two parameters characterizing SdS spacetime are the mass M of the black hole and the value Λ of the cosmological constant, with SdS spacetime reducing to Schwarzschild spacetime for $M = 0$ and deSitter spacetime for $\Lambda = 0$. As will be shown in section 2.7, these two parameters can be combined into a single dimensionless parameter $\varepsilon = M\sqrt{\Lambda/3}$ in such a way that, up to constant conformal rescaling, the spacetime only depends on the single parameter ε. Thus the essential character of the spacetime geometry only depends on a single parameter, and without loss of generality we only need to consider the effect of changing the parameter ε on the spacetime structure.

SdS spacetime is rarely used in an astrophysical context due to the fact that the corrections to the Schwarzschild or Kerr metric due to the cosmological constant are negligible for most observationally relevant astrophysical phenomena (however, see [30] for an example of an application). Here SdS spacetime is used as the simplest example of a spacetime with both a black hole and cosmological horizon. Although SdS spacetime is rarely used in an astrophysical context, it is interesting to point out that according to the Λ-CDM model of cosmology, SdS spacetime is relevant to the late time spacetime structure of our universe. The cosmological constant is becoming the dominant component of energy density in the universe, and localized matter structures are slowly approaching increasingly isolated supermassive black holes. In the far future, an observer in outer space would presumably find spacetime structure to be approximately that of a single isolated black hole in a universe with a cosmological constant; that is, Schwarzschild deSitter spacetime (or more precisely,

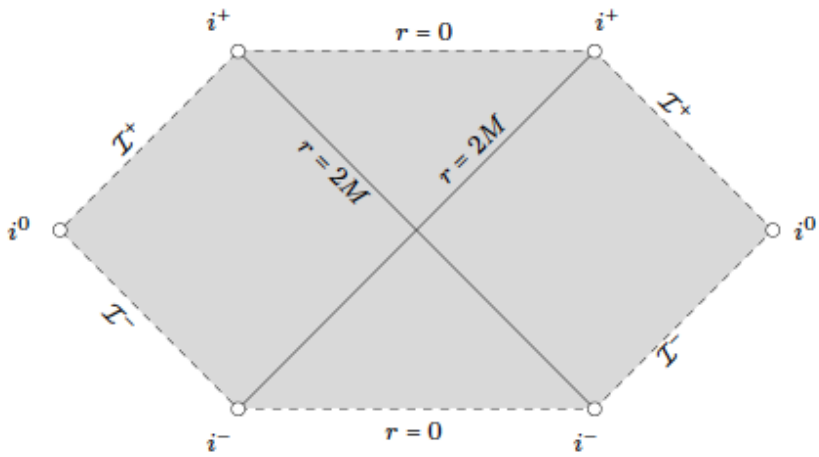


Figure 2.1: Penrose diagram of Scharzschild spacetime. The spacetime can be thought of as consisting of two regions: a black hole region and a white hole region. The line $r = 2M$ with negative slope separates these two regions.

Kerr deSitter spacetime if one takes into account rotation of the black hole).

The causal structure of the full SdS spacetime manifold has many subtleties, some of which can be readily understood by considering the Penrose diagram of the spacetime (see [20] for a more detailed discussion of the Penrose diagram of SdS spacetime). In everything that follows, we restrict ourselves to values of Λ and M such that $9\Lambda M^2 < 1$. In this case the Penrose diagram is as shown in figure 2.2. Some of the features of this Penrose diagram can be understood by first recalling the Penrose diagram of Schwarzschild spacetime, shown in figure 2.1. Like Schwarzschild spacetime, SdS spacetime has both black hole and white hole regions, which are time reverses of each other. In contrast to Schwarzschild spacetime, however, SdS spacetime can be viewed as having an infinite series of alternating black hole and white hole regions, as shown in figure 2.2. These can be thought of as distinct black and white holes, each existing in causally separated parallel universes. More commonly however, one constructs SdS spacetime using the topological identification process shown in figure 2.3 (the resulting spacetime is what is normally called SdS spacetime; the case with an infinite number of black holes is rarely discussed). When the identification process in figure 2.3 is used, the spacelike hypersurfaces of constant time can be taken to have spherical topology, as in the big bounce universe of deSitter spacetime discussed previously. The topological identification also means that there is only one black hole and one white hole region. On the other hand, when SdS

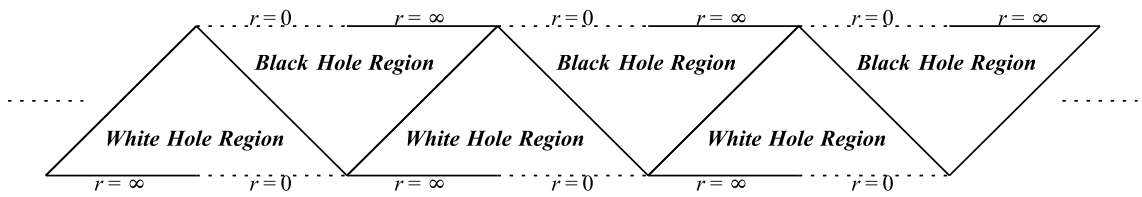


Figure 2.2: The Penrose diagram of the full spacetime manifold for SdS spacetime. There are an infinite number of black hole and white hole regions continuing indefinitely in both directions. The black hole regions can be thought of as black holes embedded in a steady state universe, and the white hole regions are simply the time reverse of the black hole regions.

spacetime is viewed as having an infinite number of black holes, one usually thinks of the spacelike slices of constant time as non-compact and with infinite spatial extent. These considerations of the global topology of SdS will not be important for our purposes since we will be limiting ourselves to a part of the spacetime containing a single black hole (such as one of the black hole regions shown in 2.2).

We can further understand the causal structure of SdS spacetime by recalling the Penrose diagram of deSitter spacetime, shown in figure 2.4. As with SdS spacetime, there is an identification procedure which results in a spacetime where the spacelike hypersurfaces can be chosen to have spherical topology. Furthermore, with an appropriate choice of slicing the spacelike hypersurfaces are perfect three spheres, and we obtain the "big bounce" cosmological model discussed previously. Alternatively, one can consider spacelike slices covering only the upper half of deSitter spacetime, also shown in figure 2.4. In this case, we have an FRW cosmology with flat spacelike hypersurfaces and a scale factor growing exponentially in time. The resulting cosmological model is sometimes called the deSitter universe, as discussed above, although it now often also goes under the name of the steady state universe. The reason for this latter name is that the deSitter universe obeys the so-called "perfect cosmological principle", which extends the usual "cosmological principle" by requiring that not only all vantage points in space be equivalent, but all vantage points in spacetime be equivalent.

In addition to the use of a Penrose diagram, one can also represent deSitter spacetime as a hyperboloid embedded in three dimensional Minkowski spacetime, as shown in figure 2.5. The hy-

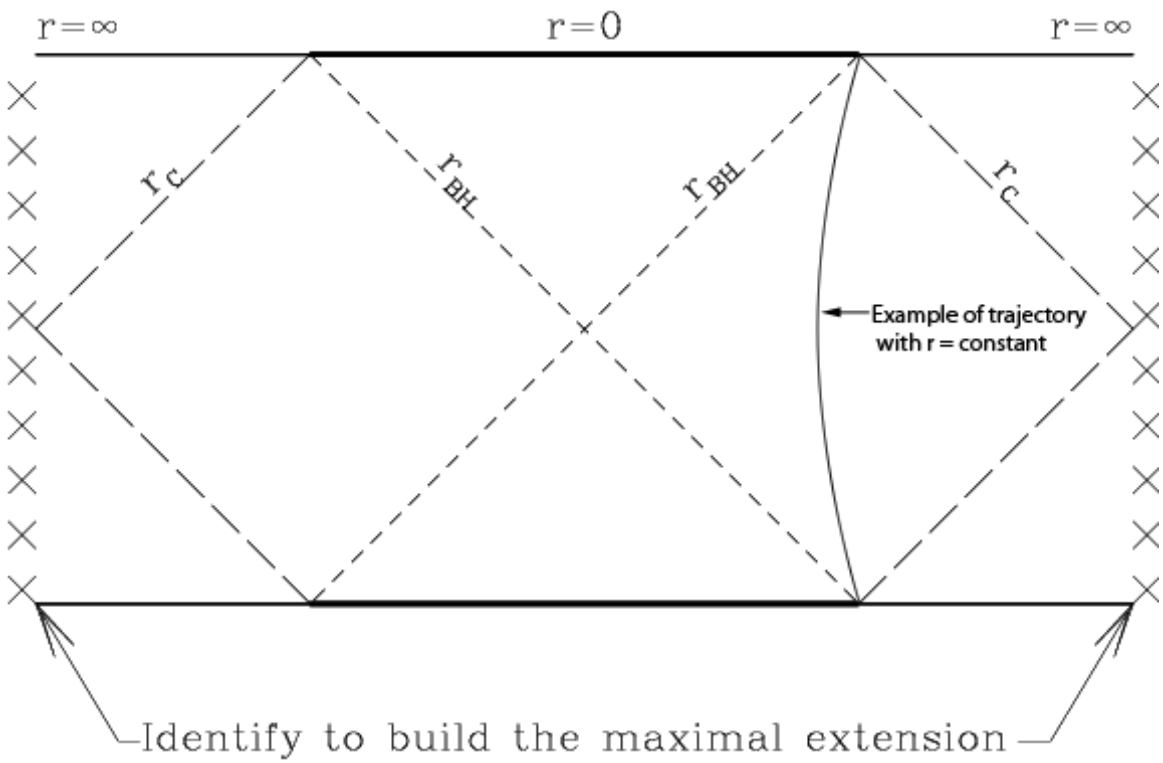


Figure 2.3: Penrose diagram of SdS spacetime. The identification procedure involves gluing the two lines of crosses above. The resulting spacetime has spacelike slices of constant time which have spherical topology.

perboloid results from suppressing two of the spatial dimensions of deSitter spacetime. It nicely illustrates the slices used in both the big bounce and steady state cosmology of deSitter spacetime (see [18] for similar diagrams constructed for SdS spacetime).

As with deSitter spacetime, one can consider only half of SdS spacetime (this is the region labelled "Schwarzschild coordinates region" in figure 2.7). In the limit $M = 0$ this part of SdS spacetime reduces to the steady state universe previously discussed, and in the limit $\Lambda = 0$ it reduces to the black hole half of Schwarzschild spacetime. Putting these two limits together, we interpret the region of SdS spacetime in figure 2.7 as a black hole embedded in a steady-state universe. It is well known that in the steady-state universe there is a cosmological horizon surrounding each freely floating observer, and that the proper distance between such observers increases exponentially with time. Based on this we expect the black hole embedded in the steady-state universe to eventually merge with the cosmological horizon of a freely falling observer drifting away from the black hole.

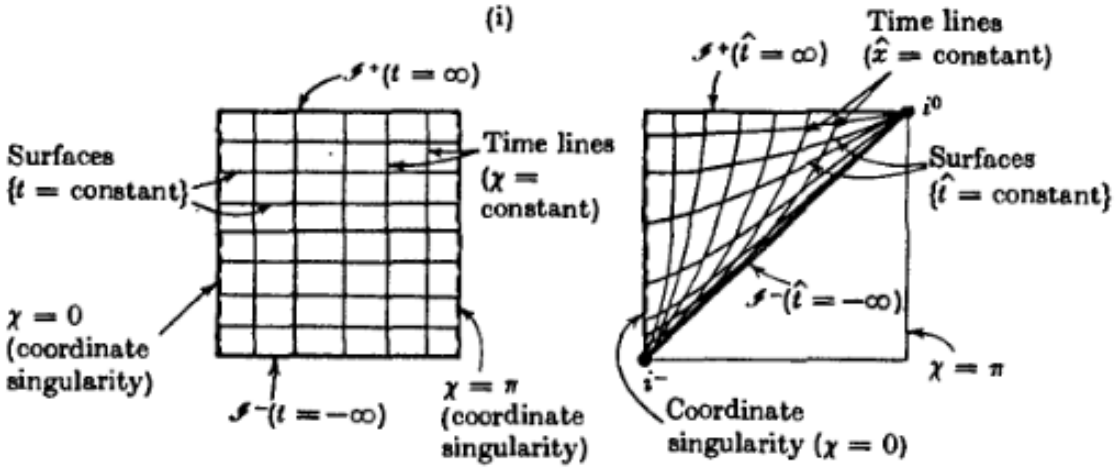


Figure 2.4: Penrose diagram for deSitter spacetime. The diagram on the left shows the full deSitter manifold, where it is understood that the left and right vertical edges are to be glued through an identification process. The spacelike hypersurfaces of constant time (the horizontal lines) are three spheres. In the diagram on the right, only the upper half of the deSitter manifold is considered. Here the spacelike hypersurfaces are three dimensional flat Euclidean space. In terms of the cosmology, the situtation on the left is the big bounce cosmology and on the right we have the steady state cosmology. Figure from © [15], page 127, by permission from publisher.

The coordinates used to cover this part of the spacetime will be developed in chapter 3. The spacelike hypersurfaces of these coordinates reduce to the flat spacelike hypersurfaces of deSitter spacetime in the limit that $M = 0$ (as shown in figures 2.4 and 2.5). In some sense, they can be thought of as a generalization of what has traditionally been called the planar coordinates of deSitter spacetime. In the limit $M = 0$, these coordinates are closely related to the well known Lemaitre coordinates [22] (this is discussed in section 3.3.2). These coordinates will have several advantages over traditional coordinate systems such as the Schwarzschild coordinates (section 2.2). For example, they will be free of coordinate singularities and will have geodesic timelike coordinate curves.

In section 2.2 we review the Schwarzschild coordinates for SdS spacetime and discuss some additional features of the spacetime which will be relevant in later sections. In section 2.3 we find the equations for the Schwarzschild coordinates of radial timelike geodesics. These equations will be important for the coordinate system developed in chapter 3. In section 2.4 we discuss the horizons of observers stationary next to the black hole. Traditionally, it is these stationary Killing horizons

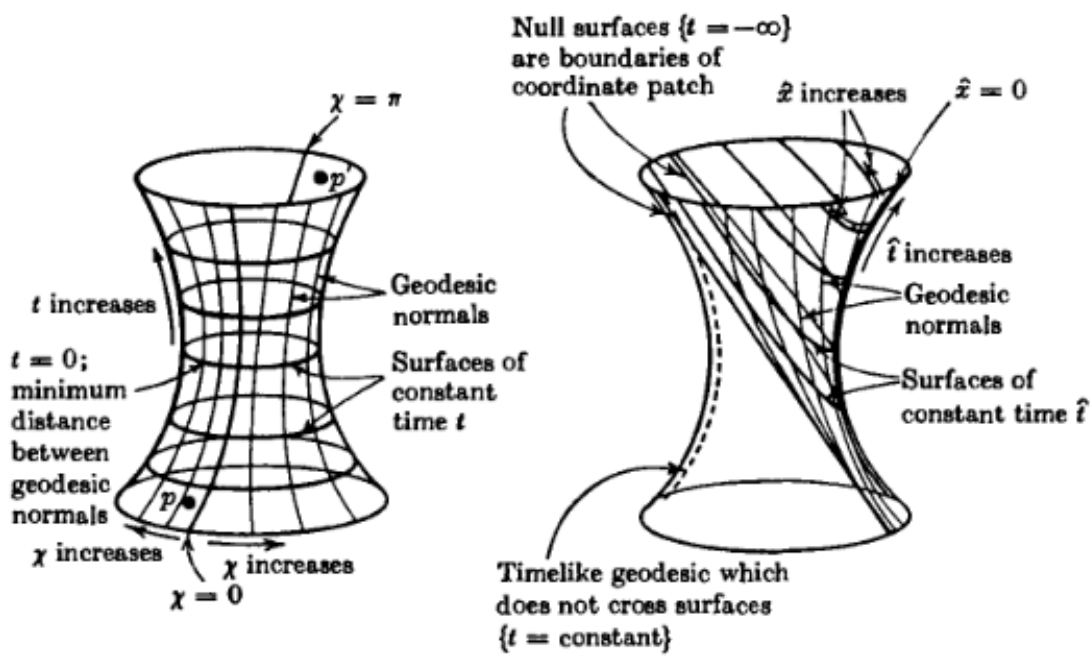


Figure 2.5: The hyperboloid on the left is the full deSitter manifold. The slices of constant time are circles. These shrink to a minimum size and then grow again, giving rise to the so-called big bounce cosmology. The diagram on the left shows a set of constant time slices which only cover half of the hyperboloid. These are the constant time slices of the steady state universe. They are spatially flat and grow exponentially. Figure from © [15], page 125, by permission from publisher.

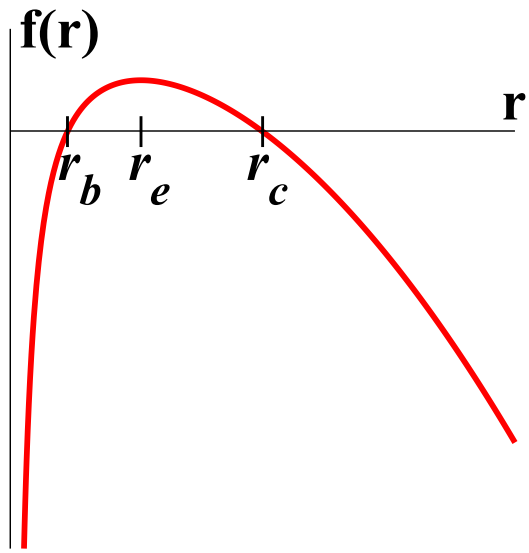


Figure 2.6: Function $f(r) = 1 - 2M/r - H^2r^2$.

which are called the horizons of SdS spacetime. Understanding these stationary horizons is an important starting point before undertaking the study of merging horizons in SdS spacetime. In section 2.5 we give a more precise formulation of the problem statement considered in this thesis, although the full formulation of the problem will not come until chapter 4. In section 2.7 we define a dimensionless parameter ε which will play an important role in calculations throughout this thesis. Finally, in section 2.6 we discuss how the $r \to \infty$ limit of SdS spacetime can be well approximated by deSitter spacetime (in a sense that we will make more precise). This fact about SdS spacetime will be important when we set up the equations for the null geodesic generators of the merging horizons in chapter 4.

2.2 Schwarzschild coordinates

SdS spacetime is traditionally presented in Schwarzschild coordinates, where the line element is

$$ds^2 = f(r)dt^2 - f(r)^{-1}dr^2 - r^2(\sin^2\theta d\phi^2 + d\theta^2)\,, \tag{2.4}$$

$$f(r) = 1 - \frac{2M}{r} - \frac{\Lambda}{3}r^2\,. \tag{2.5}$$

The Schwarzschild coordinates of SdS spacetime encompass as special cases both the Schwarzschild coordinates of Schwarzschild spacetime (eq. 2.1) and the static coordinates of deSitter spacetime (eq. 2.2). Throughout this thesis, we will use units such that $G = c = 1$ and take the metric signature to be $(+,-,-,-)$. In the above, the 2-surfaces $(t,r) =$ constant are spheres of area $4\pi r^2$, with θ and ϕ the polar and azimuthal angles on these spheres, respectively. The spheres $r = r_b$ and $r = r_c$ can be interpreted as the location of black hole and cosmological horizons, respectively, for a set of static observers. This will be discussed in more detail in section 2.4. The Schwarzschild coordinates can be used to cover either the black hole or white hole part of SdS spacetime (i.e. any of the triangular regions in figure 2.2). Figure 2.7b illustrates how the Schwarzschild coordinates would cover the black hole region of SdS spacetime. This region of spacetime can be thought of as consisting of three parts: the interior of a black hole for $r < r_b$, an external expanding universe for $r > r_c$, and an intermediate region $r_b < r < r_c$. The regions $r < r_b$, $r_b < r < r_c$ and $r > r_c$ are shown on the Penrose diagram in figure 2.7a. The coordinate t is timelike in the region $r_b < r < r_c$, and spacelike in the regions $r < r_b$ and $r > r_c$. In the intermediate region $r_b < r < r_c$ observers can in principle remain stationary with $(r(\tau), \phi(\tau), \theta(\tau)) =$ constant, neither caught up in the expansion of the universe nor irrevocably swallowed by the black hole. Such observers would in general have a proper acceleration, except at a critical equilibrium radius $r_e \in (r_b, r_c)$, where it is possible to have a stationary observer outside the black hole whose worldline is a timelike geodesic. This can be seen by looking at the effective potential for timelike radial geodesics, as shown in figure 2.8 (the effective potential is found in section 2.3). This unstable equilibrium can be thought of as resulting from the balance of the expansion pulling the observer to larger r and the gravitational pull of the black hole pulling the observer to smaller r. The timelike trajectory $r(\tau) = r_e$ is shown in figure 2.9. The equilibrium radius r_e will play an important role in the coordinate system we will develop in

chapter 3. Its value can be calculated as

$$r_e = \frac{1}{H}(MH)^{\frac{1}{3}}\,, \tag{2.6}$$

where $H = \sqrt{\Lambda/3}$. H can be interpreted as the Hubble constant in the case of pure deSitter spacetime. Another value which we be needed in many calculations further on in this thesis is $f(r_e)$. From (2.4) and (2.6), we have

$$f(r_e) = 1 - 3(MH)^{2/3}\,. \tag{2.7}$$

2.3 Radial timelike geodesics

Let $\gamma(\tau) = (t(\tau), r(\tau), 0, \pi/2)$ be the Schwarzschild coordinates of an arbitary timelike radial geodesic. SdS spacetime is static and has the Killing vector

$$\xi^{(t)} = (1,0,0,0)\,.$$

This leads to the following conservation law for $\gamma(\tau)$:

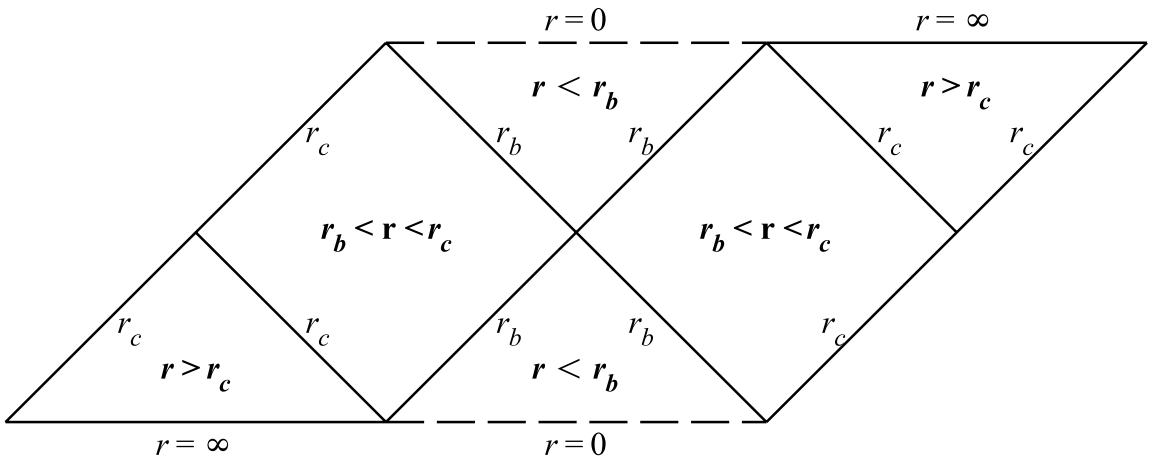
$$\langle \frac{d\gamma}{d\tau}, \xi^{(t)} \rangle = E = \text{constant}\,.$$

That is:

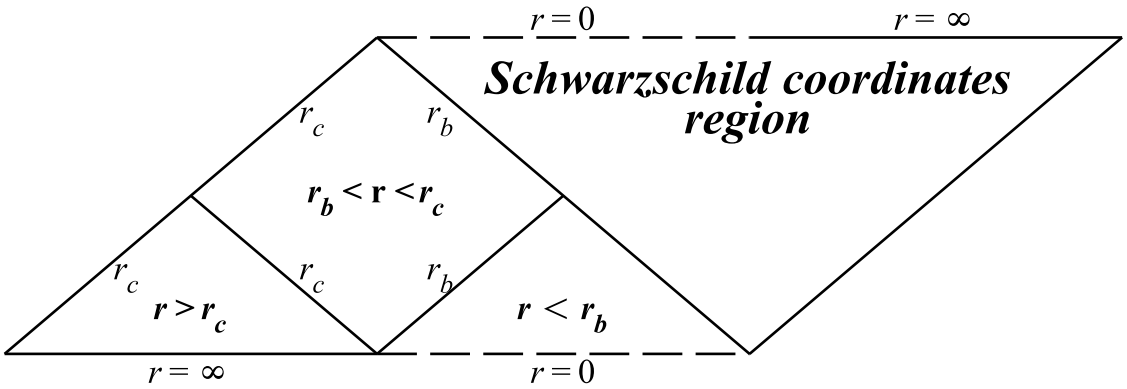
$$\frac{dt}{d\tau} = \frac{E}{f(r)}\,, \tag{2.8}$$

where (2.4) has been used. Another equation for $\gamma(\tau)$ comes from the requirement that τ be proper time:

$$\langle \frac{d\gamma}{d\tau}, \frac{d\gamma}{d\tau} \rangle = 1\,.$$



(a)



(b)

Figure 2.7: a) Penrose diagram of SdS spacetime with the regions $r < r_a$, $r_a < r < r_b$ and $r > r_c$ indicated. b) The black hole portion of SdS spacetime covered by Schwarzschild coordinates. This is the region of the spacetime that will be considered in this thesis.

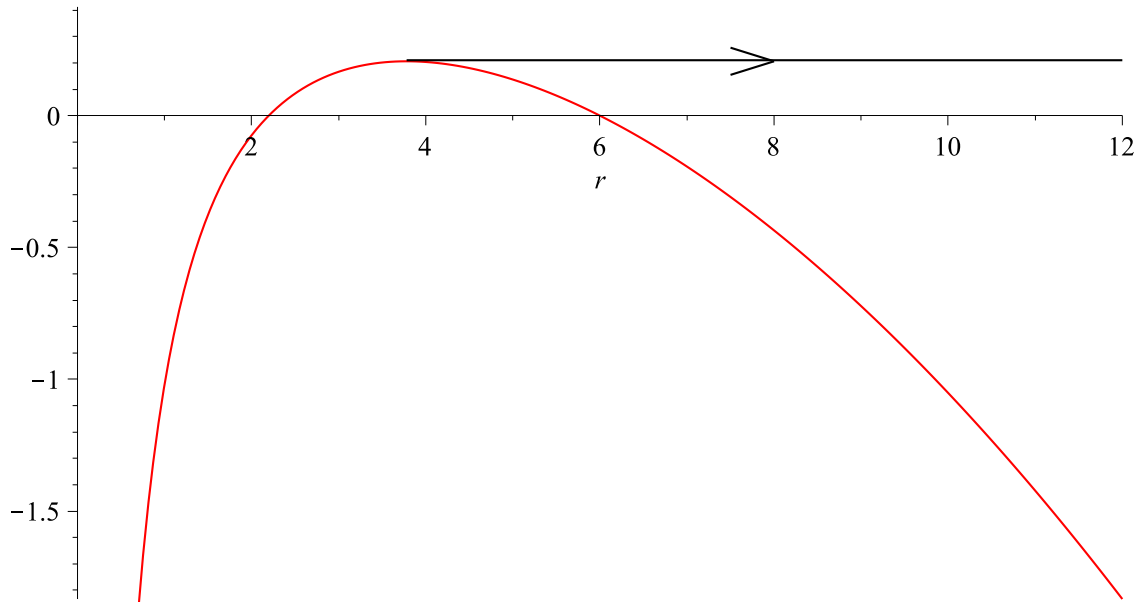


Figure 2.8: Effective potential diagram for radial timelike geodesic trajectories in SdS spacetime. There is an unstable equilibrium at $r = r_e$. The black line represent the motion of a geodesic observer moving away from the equilibrium location.

This gives

$$f(r)\left(\frac{dt}{d\tau}\right)^2 - f(r)^{-1}\left(\frac{dr}{d\tau}\right)^2 = 1\,. \tag{2.9}$$

Substituting (2.8) into the above gives the equation for $r(\tau)$:

$$E^2 - \left(\frac{dr}{d\tau}\right)^2 = f(r)\,.$$

The above equation can be recast in effective potential form as

$$\left(\frac{dr}{d\tau}\right)^2 + V_{\text{eff}}(r) = E^2\,,$$

where

$$V_{\text{eff}}(r) = f(r)\,.$$

A plot of the effective potential $V_{\text{eff}}(r)$ is shown in figure 2.8. We see that there is an unstable equilibrium point at $r = r_e$. If we impose the requirement

$$r(\tau = -\infty) = r_e \tag{2.10}$$

on a trajectory, this leads to the following value for E:

$$E = f(r_e)^{1/2}\,.$$

Substituting the above into (2.8) and (2.9) leads to the following equations for $t(\tau)$ and $r(\tau)$:

$$\frac{dt}{d\tau} = \frac{f(r_e)^{1/2}}{f(r)}, \tag{2.11}$$

$$\frac{dr}{d\tau} = \pm(f(r_e) - f(r))^{1/2}\,, \tag{2.12}$$

where the $+$ sign is for trajectories drifting towards larger values of r and the $-$ sign is for trajectories drifting towards $r = 0$. There is also a trajectory with $r(\tau) = r_e$ for all τ. The family of trajectories satisfying the above equations will play an important role in the coordinate system developed in chapter 3. As well, the horizons considered in this thesis will be those an observer drifting away from the equilbrium (as shown in figures 2.8 and 2.9), whose coordinates are described by a solution of the above equation (these horizons will be discussed more in section 2.5).

2.4 Horizons of stationary observers

Consider the Penrose diagram of SdS spacetime, as shown in figure 2.7a. Notice first from this diagram that unlike Schwarzschild spacetime, there is no clear distinction between future null infinity and future timelike infinity, and instead the two combine into what could be called future null/timelike infinity (see section 5.2 of [15] for a discussion of future null/timelike infinity for deSitter spacetime). Because of this, it is not possible to use the usual definition of a black hole event horizon as the boundary of the causal past of future null infinity (see section 12.1 of [32] for

a detailed definition of a black hole event horizon). Instead, we need an alternative definition of event horizon. One possibility is to define the black hole horizon as the boundary of the causal past of future null/timelike infinity. However, since we are ultimately interested in describing the part of the universe that is accessible to a specific observer, we will use the concept of a *causal horizon*, which is based on the causal past of a timelike trajectory. The *causal horizon* of an observer (i.e. timelike trajectory) is defined as the boundary of the causal past of the observer's trajectory. Formally, if $\gamma(\tau)$ is the observer's trajectory, and $\gamma(\mathbb{R})$ is the image of the real line under the mapping that is $\gamma(\tau)$, then the causal horizon of $\gamma(\tau)$ is $\mathrm{Bd}(J^-(\gamma(\mathbb{R})))$. The terminology "causal horizon" may not be familiar to some readers, and has not yet gained widespread and popular usage. For example, some authors prefer to use the words "cosmological horizon" in the cosmological context, and "Rindler horizon" in the context of accelerated observers, even though these are both observer dependent horizons which can be subsumed under the broader concept of causal horizons. Here we follow the terminology in [17] by using the words "causal horizon".

The causal horizon can naturally incorporate both the black hole and cosmological horizons. For example, for an observer stationary outside a Schwarzschild black hole, the causal horizon of this observer would coincide with what we normally think of as the black hole horizon; the spherical surface at the Schwarzschild radius. More generally, the formal definition of a black hole event horizon (as in section 12.1 of [32]) coincides with the causal horizon of any observer with a timelike wordline reaching future timelike infinity. In a cosmological context, it has long been known that for a geodesic observer in a deSitter universe, the causal horizon is a spherical surface centered on this observer, and is at a cosmological proper distance equal to the Hubble radius. Historically, this cosmological horizon of the deSitter universe was one of the first examples where it was necessary to generalize the concept of the event horizon by considering observer dependent causal horizons. For example, the concept of the causual horizon is used in a seminal article on the thermodynamics of cosmological event horizons [11], although it is simply called an "event horizon" in that article.

For simplicity, from now on we will often use the word *horizon*, where it is implicitly understood that we are refering to the *causal horizon,* unless otherwise specified. We will also abuse terminology slightly and use the word *horizon* to mean either the full horizon, viewed as a 3d null

hypersurface, or the 2d surface which is formed by taking the intersection of this null hypersurface with a 3d spacelike hypersurface of constant time. For example, when we use the word horizon in the context of the deSitter universe, we could be referring to the full horizon, which is a null hypersurface obtained by taking the boundary of the causal past of a geodesic observer's trajectory. However, we could also be referring to the intersection of this null hypersurface with a spacelike hypersurface $t =$ constant, where t is the time coordinates from the planar coordinates leading to the line element in (2.3). This intersection results in a spherical surface surrounding the observer, and is normally what one thinks of when referring to the horizon.

To illustrate the presence of both cosmological and black hole horizons in SdS spacetime, consider first the trajectory of an observer with $r(\tau) =$ constant and $r_b < r(\tau) < r_c$, where r_b and r_c are the two roots of $f(r) = 0$ (we assume that $0 < 9\Lambda M^2 < 1$, so that $f(r)$ has precisely two real roots). The example trajectory of $r(\tau) = r_e$ is shown in figure 2.9, and other trajectories with $r(\tau) =$ constant would have this same shape. From the Penrose diagram in figure 2.9, we can deduce that the causal horizon of this observer consists of the spheres $r = r_b$ and $r = r_c$ (also see figure 2.10a). The outer sphere $r = r_c$ is the cosmological horizon and the inner sphere r_b is the black hole horizon. Now consider any observer satisfying $r_b < r(\tau) < r_c$ for all $\tau \in \mathbb{R}$. From the Penrose diagram, we can once again conclude that $r = r_c$ and $r = r_b$ are the cosmological and black hole horizons, respectively. Thus the concentric spheres $r = r_b$ and $r = r_c$ are the horizons for a large set of observers who neither fall into the black hole by crossing $r = r_b$, nor get irrevocably caught up in the accelerated expansion of the space by crossing $r = r_c$. In addition to being the horizons for a large family of observers, the horizons $r = r_b$ and $r = r_c$ are both Killing horizons associated with the Killing vector of the spacetime. For these reasons, it is natural to call $r = r_b$ and $r = r_c$ the horizons of SdS spacetime, and to our knowledge, all previous work on the horizons of SdS spacetime have been dealing with these horizons exclusively.

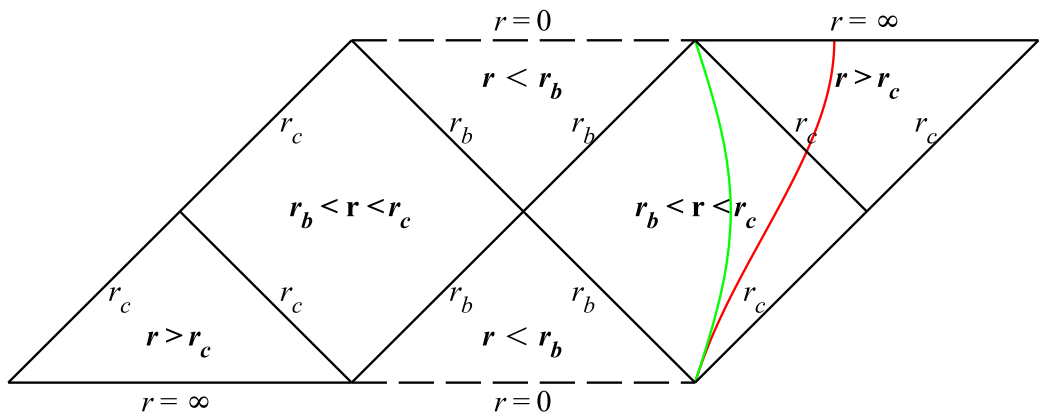


Figure 2.9: Penrose diagram with two observer trajectories shown. The green trajectory corresponds to an observer with $r(\tau) = r_e$ for all τ. The red trajectory corresponds to an observer drifting away from the unstable equilibrium $r = r_e$ (see figure 2.8).

2.5 Problem statement

Here we will study the horizons for an observer caught up in the expansion of space and drifting away from the black hole along a geodesic. This trajectory is shown on the Penrose diagram in figure 2.9. This is a natural trajectory to consider since it resembles what we are currently experiencing in our universe. According to the Λ-CDM model of cosmology, we are currently in the Λ dominated phase of the universe's expansion, and black holes are constantly drifting away from us and crossing our cosmological horizon.

The horizons for such an observer consist of a black hole merging with a cosmological horizon. Since this merging of horizons is fundamentally 3+1 dimensional, it cannot be deduced from the 1+1 dimensional Penrose diagram alone (see figure 2.10 for an illustration of the past causal horizon of both stationary and drifting observers on the Penrose diagram). In order to deduce this merging of horizons and study it, we must calculate the individual trajectories of light rays eminating from the observer's trajectory at late times. Schwarzschild coordinates are inadequate for such a calculation, and so we must develop another coordinate system, as will be discussed in chapter 3. The calculations of the lights rays which make up the horizon and the analysis of this horizon will be in

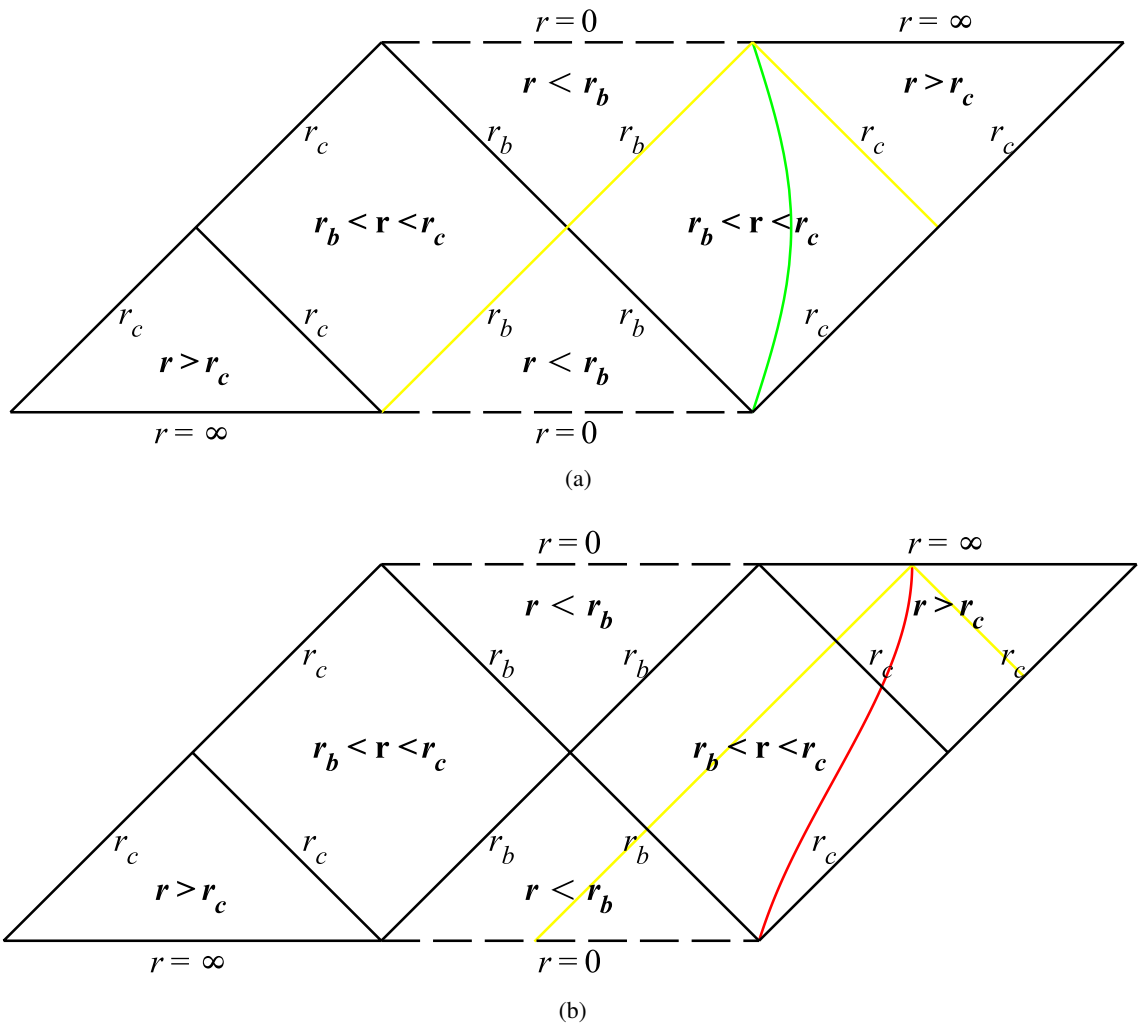


Figure 2.10: a) Penrose diagram with causal horizon (yellow lines) of an observer stationary next to the black hole. b) Penrose diagram with causal horizon (yellow lines) of an observer drifting away from the black hole.

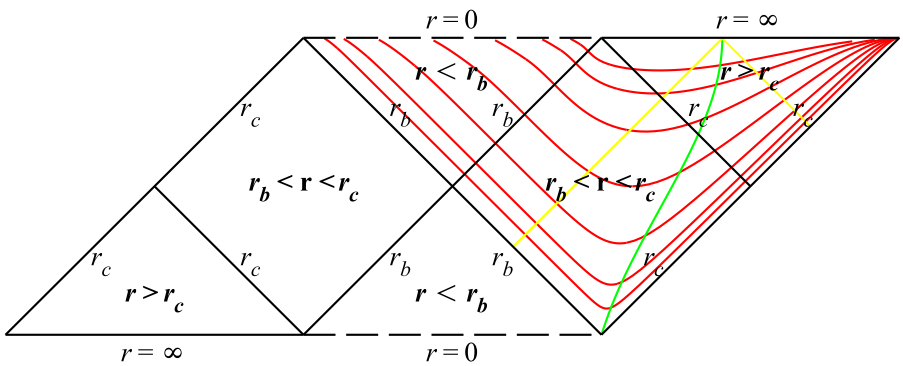


Figure 2.11: Penrose diagram showing the drifting observer (green curve), the causal horizon of this observer (yellow lines) and the spacelike hypersurfaces of constant time (red curves). The red curves are the spacelike hypersurfaces for the coordinate system developed in chapter 3. The upper and lower red curves correspond to very late times and very early times, respectively. Notice how the intersection of the lower red curves and the yellow lines is arbitrarily close to $r = r_b$ (left side) or $r = r_c$ (right side) at early times. This is what allows us to conclude that at early times, the horizons of the drifting observer consist of the two concentric spheres $r = r_b$ and $r = r_c$.

chapter 4.

Although the merging of horizons discussed in the previous paragraph cannot be deduced explicitly by looking at the Penrose diagram, its existence can be inferred by using our knowledge of the observer's horizon at both early times and late times. In chapter 3 we will introduce a coordinate system whose spacelike hypersurfaces are shown in the Penrose diagram in 2.11. From this diagram, we see that at early times, the spacelike hypersurfaces of constant time intersect the horizon arbitrarily close to the spheres r_b and r_c. Thus the early time horizons of the drifting observer are the same as the horizons of the stationary observer; they are concentric spheres, with the inner sphere being the black hole horizon and the outer sphere being the cosmological horizon.

We can also deduce the shape of the horizon for our drifting observer at late time. At late times, the spacetime in the neighborhood of the drifting observer can be well approximated by the deSitter universe, and the trajectory of the drifting observer approaches the trajectory of an observer expanding with the Hubble flow in the deSitter universe. Both of these facts will be discussed in more detail in section 2.6 below. Thus the horizon surrounding the drifting observer will approach that of

an observer caught up in the Hubble flow of deSitter spacetime. That is, the horizon will approach a closed surface surrounding the observer (the surface will be spherical for the appropriate choice of time slicing).

In conclusion, if at early times the horizons are two concentric spherical surfaces and at late times the horizon is a single closed surface, then we can deduce that at some point in time there is a transition between the two, and this transition must necessarily involve the merger of the black hole and cosmological horizons. It is this transition which we study in this thesis.

2.6 $r \to \infty$ limit of SdS spacetime

In section 2.5, we claimed that in the limit that $r \to \infty$, SdS spacetime in some sense approaches deSitter spacetime, and that this could be used to conclude that the late time behavior of a drifting observer's horizon must be the same as that of an observer caught up in the Hubble flow of the deSitter universe (also known as the steady state universe). This conclusion about the late time behavior of the observer's horizon is not only important for the heuristic arguments used in section 2.5, but also critical for the calculations to be performed in chapter 4, where we will solve the equations for the null geodesic generators which make up the horizon. These equations will require initial conditions, or more precisely, final conditions, which will be obtained using knowledge of the late time behavior of the horizon. We will use the fact that at late times the drifting observer is in a part of SdS spacetime well approximated by deSitter spacetime, so that the horizon at late times should resemble the horizon of an observer drifting in the deSitter universe. That is, at late times the horizon should be a closed surface surrounding the observer. Notice that for $r > r_c$, the coordinate r in SdS spacetime is timelike, so that the limit $r \to \infty$ is the late time limit.

The claim that SdS spacetime approaches deSitter spacetime in the limit $r \to \infty$ is an inherently ambiguous one, since it requires a way of comparing two spacetimes. The definition we will use is that SdS spacetime approaches deSitter spacetime as $r \to \infty$, if there exists a coordinate system

of SdS spacetime where the lapse, shift and 3-metric of SdS spacetime all approach the lapse, shift and 3-metric in a coordinate system of deSitter spacetime. Usually comparing two spacetimes by comparing their metric components in specific coordinate systems is hopelessly difficult, since one cannot disentangle the difference in metric components due to coordinate changes from those due to genuine changes in spacetime geometry. Fortunately, the Schwarzschild coordinates of SdS spacetime are linked in a simple and natural way to the static coordinates of deSitter spacetime, so that according to our definition, we can claim that at late times SdS spacetime does indeed approach deSitter spacetime (see (2.4) and (2.2) for the line elements of SdS and deSitter spacetime in Schwarzschild and static coordinates, respectively).

Given the late time behavior of SdS spacetime in the sense defined above, it follows that the late time behavior of the null geodesics and timelike geodesics in SdS spacetime must approach the late time behavior of null and timelike geodesics in deSitter spacetime. Again, there is an inherent ambiguity involving the definition of closeness used when comparing null geodesics in one spacetime to null geodesics in another spacetime. This can be resolved using a specific coordinate system as was done when comparing spacetimes above.

To summarize, we use the $r \to \infty$ behavior of SdS spacetime to infer that the late time behavior of the horizon must be a closed surface surrounding the drifting observer. This knowledge will in turn be used in chapter 4 to set up the "initial" conditions (when flowing backwards in time) for the geodesic equations of the null generators which make up the horizon.

2.7 The dimensionless parameter $\varepsilon = HM$

Although the metric components of SdS spacetime depend on both the black hole mass M and the cosmological constant Λ, the essential geometry of the spacetime can be characterized by a single

dimensionless parameter ε, defined as follows:

$$\varepsilon = HM\,, \tag{2.13}$$

where $H = \sqrt{\Lambda/3}$ is the Hubble constant. Once ε is specified, changing M or Λ simply amounts to a constant rescaling of the metric. This can be seen by introducing the dimensionless time $\bar{t}$ and radial coordinate $\bar{r}$ as follows:

$$\begin{aligned} \bar{t} &= Ht, \\ \bar{r} &= Hr\,. \end{aligned}$$

The line element (2.4) now becomes

$$ds^2 = \frac{1}{H^2}\left[\left(1-\frac{2\varepsilon}{\bar{r}}-\bar{r}^2\right)d\bar{t}^2 - \left(1-\frac{2\varepsilon}{\bar{r}}-\bar{r}^2\right)^{-1}d\bar{r}^2 - \bar{r}^2(\sin^2\theta d\phi^2 + d\theta^2)\right] \tag{2.14}$$

and we see that apart from a constant overall rescaling by the Hubble length $1/H$, the line element only depends on the dimensionless parameter ε. Thus all values of M and Λ which yield the same value for $\varepsilon = HM$ lead to conformally equivalent spacetimes. Due to the conformal equivalence of these spacetimes, the essential character of the geometry only depends on the parameter ε. For this reason, throughout this thesis we will often discuss the dependence of the metric on the parameters M and Λ by referring to the single parameter ε.

The parameter ε is related to the relative size of the spherical black hole and cosmological horizons of stationary observers (as measured using the circumference of a great circle, for example). For $\varepsilon \ll 1$ the black hole is much smaller than the cosmological horizon and as $\varepsilon \to \varepsilon_c^-$, where $\varepsilon_c = 1/3\sqrt{3}$, the black hole size approaches that of the cosmological horizon size. For $\varepsilon > \varepsilon_c$ there is neither a black hole or cosmological horizon, and instead we have a naked singularity [20]. In this thesis we will restrict ourselves to values of M and Λ such that $0 < \varepsilon < \varepsilon_c$, in which case the structure of the spacetime is as shown in the Penrose diagram in figure 2.9, and the function $f(r)$ will have precisely two positive roots, r_b and r_c. The spheres $r = r_b$ and $r = r_c$ are the locations of the black hole and cosmological horizons for an observer stationary outside the black hole (stationary in the

sense of the timelike Kiling vector of (2.4)). The roots r_b and r_c, as well as the ratio r_b/r_c, can be computed for $\varepsilon \ll 1$ by using pertubation methods to find the solution to $f(r) = 0$. This gives

$$\begin{aligned} r_b &= 2M(1+\mathscr{O}(\varepsilon^2)), \\ r_c &= \frac{1}{H}\left(1-\varepsilon+\mathscr{O}(\varepsilon^2)\right), \\ \frac{r_b}{r_c} &= \varepsilon+\mathscr{O}(\varepsilon^2). \end{aligned} \tag{2.15}$$

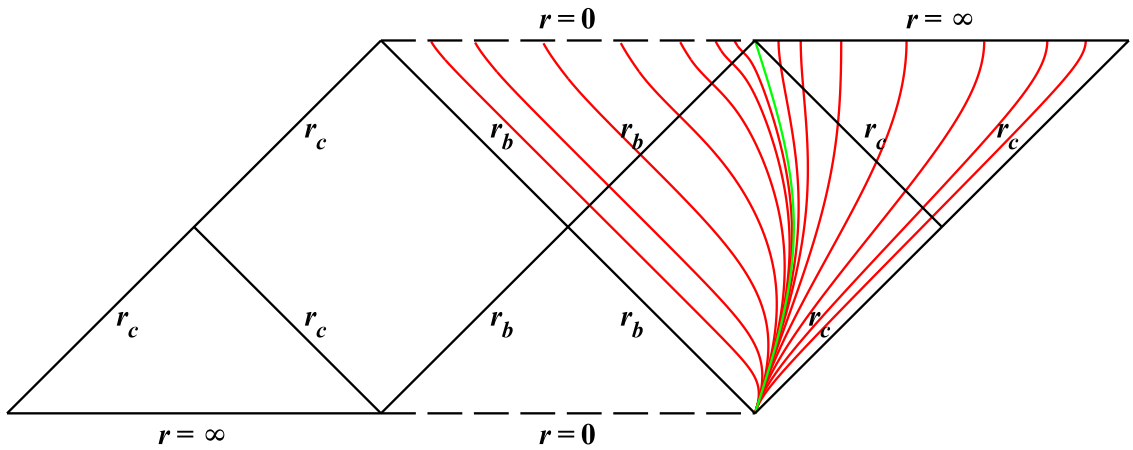
From the above we see that for $\varepsilon \ll 1$ the black hole radius approaches the Schwarzschild radius $r_b = 2M$ and the cosmological horizon approaches the Hubble radius $r_c = \frac{1}{H}$. As well, we see that the relative size of r_b and r_c is characterized by the parameter ε, as mentioned previously. Throughout this thesis we will take the point of view that the case where $\varepsilon = 0$ corresponds to pure deSitter spacetime (i.e. $M = 0$), with the introduction of $\varepsilon \ll 1$ equivalent to introducing a small black hole into the spacetime. In this way considering $\varepsilon \ll 1$ amounts to considering a perturbation of deSitter spacetime caused by a small black hole. This perturbation of deSitter spacetime can be seen in the line element (2.14), where $\varepsilon = 0$ gives the line element for deSitter spacetime and $\varepsilon \ll 1$ introduces a small perturbation to the $g_{\bar{t}\bar{t}}$ and $g_{\bar{r}\bar{r}}$ metric components. The limit $\varepsilon \ll 1$ will be exploited in some of the analytical calculations in this thesis. For example, it will be used to approximate the caustic structure of merging horizons in chapter 4 and the horizon area in chapter 5. It is a very natural limit to consider, given that for a typical supermassive black hole and the currently accepted value of Λ, we would have $M \approx 10^{12}\,\mathrm{m}$ and $\Lambda \approx 10^{-52}\,\mathrm{m}^{-2}$, giving $\varepsilon \approx 10^{-14}$.

Chapter 3

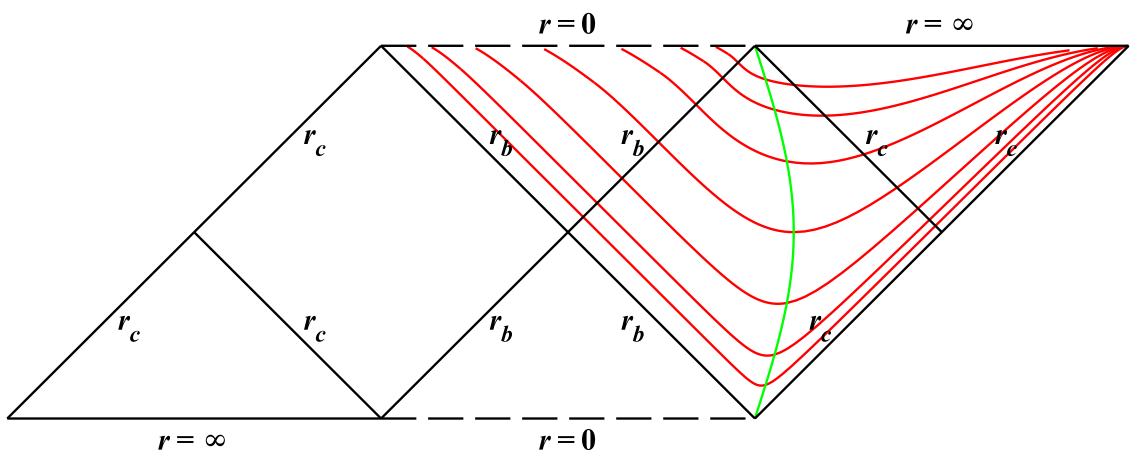
Lemaitre-Planar Coordinates

We wish to develop a coordinate system that acomplishes several purposes. First, we would like these new coordinates to cover the same region of spacetime as covered by Schwarzschild coordinates (see figure 2.7b), so that we may interpret the spacetime region under consideration as a black hole in the Λ dominated phase of an expanding universe. Second, in contrast to Schwarzschild coordinates, we would like these coordinates to be free of coordinate singularities. Also in contrast with Schwarzschild coordinates, if (T, R, θ, ϕ) are the new coordinates, we would like the metric signature to be $(+,-,-,-)$ on the whole region of spacetime covered by the coordinates, so that T and R can always be interpreted as timelike and spacelike variables, respectively. Lastly, we would like the $T =$ constant hypersurfaces to be interpreted as the hypersurfaces of constant time of a homogeneous and isotropic universe containing a black hole. This interpretation will come from the fact that the coordinate T will be the proper time for a family of geodesic observers, along with the fact that these hypersurfaces reduce to the usual hypersurfaces of an FRW cosmology in the limit $M \to 0$. The requirement that the spacelike hypersurfaces intersect the black hole can be verified by, for example, comfirming the existence of an apparent horizon on these hypersurfaces.

The coordinates which will achieve all of these purposes will be developed in this chapter, and are a special case of the general class of coordinate systems known as Gaussian normal (GN) coordinates. These coordinates are built using a family of freely falling observers (i.e. timelike geodesics) and hypersurfaces orthogonal to these observers. Given an arbitary spacetime, it is not possible in general to set up global GN coordinates over the whole of that spacetime, and one must content



(a)



(b)

Figure 3.1: a) The timelike coordinate curves of LP coordinates, shown on the Penrose diagram of SdS spacetime. b) The spacelike coordinate curves of LP coordinates.

one's self to a local use of GN coordinates. However, according to the Frobenius theorem (see section 2.3.3 of [25], or section B.3 of [32]), if we have a family of timelike trajectories (i.e. timelike congruence) with vanishing vorticity tensor, then there exists a time coordinate such that the hypersurfaces $t =$ constant are everywhere orthogonal to this congruence. One set of spacetimes where we have such a congruence, and it is possible to set up global GN coordinates, are the familiar spacetimes of FRW cosmology. For example, GN coordinates for the steady state universe lead to the familiar line element

$$ds^2 = d\tau^2 - e^{2H\tau}\left(d\rho^2 + \rho^2(d\theta^2 + \sin^2\theta d\phi^2)\right) \tag{3.1}$$

which has flat hypersurfaces and an exponential scale factor. The $(\tau, \rho, \theta, \phi)$ coordinates above are often called planar coordinates. The GN coordinates which we will develop for SdS spacetime will reduce to the planar coordinates of the steady state universe in the limit $M = 0$, and can be thought of as a generalization of these coordinates to include the case where there is a single Schwarzschild black hole. They are also closely related to the well known Lemaitre coordinates [22] in the case where the cosmological constant is zero (this will be discussed in section 3.3.2). For this reason, we will call them Lemaitre-planar (LP) coordinates. They will be based on a family of timelike geodesics eminating from the equilibrium point $r = r_e$ (see figure 3.3a for the effective potential diagram and figure 2.9 for the timelike trajectory $r(\tau) = r_e$ on the Penrose diagram). The family of such trajectories is shown on the Penrose diagram in figure 3.1a. The spacelike hypersurfaces of LP coordinates are shown in figure 3.1b.

As an interesting historical note, it would seem that the earliest construction of GN coordinates for Schwarzschild deSitter spacetime go all the way back to Lemaitre [22]. These coordinates (let us simply call them Lemaitre coordinates) reduce to the well known Lemailtre coordinates for Schwarzschild spacetime in the limit that $\Lambda \to 0$. They also have the attractive feature of having spatially flat hypersurfaces. However, the part of spacetime covered by these coordinates is not the part of the spacetime we are interested in considering (see figure 3.2), and so these coordinates will not be useful for our purposes. The difference between Lemaitre coordinates and LP coordinates can be understood by comparing the family of radial timelike geodesics used as timelike coordinate

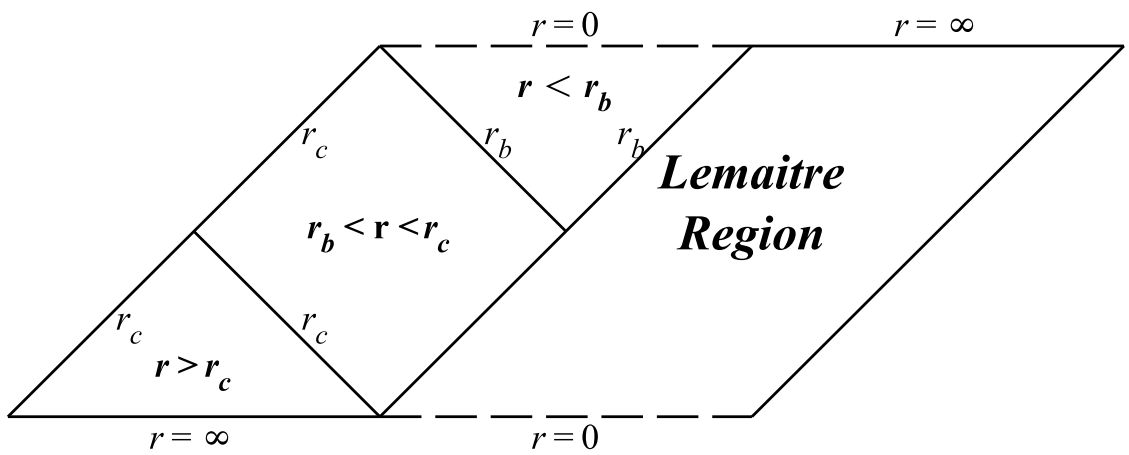
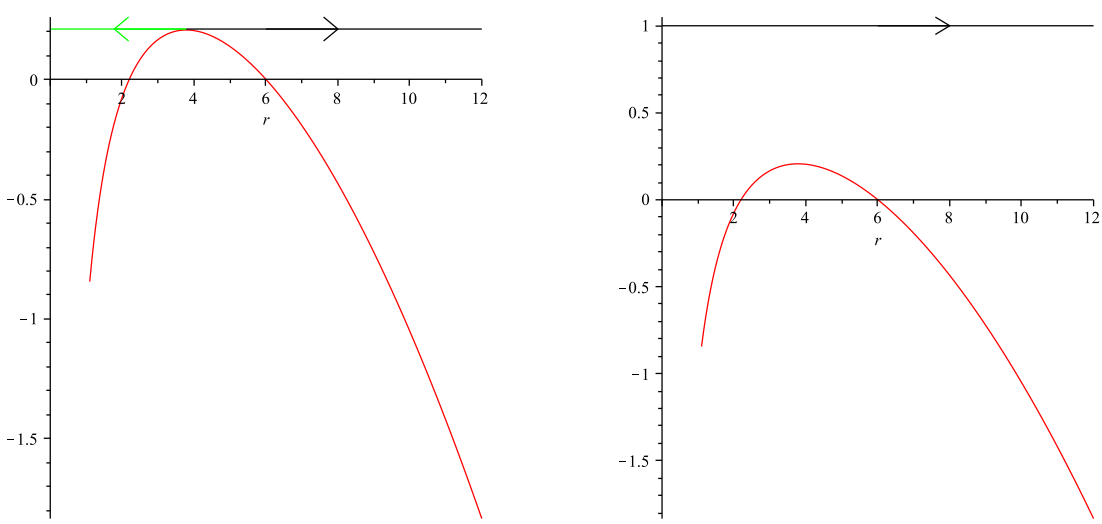


Figure 3.2: Penrose diagram showing the region of SdS spacetime covered by Lemaitre coordinates.



(a) (b)

Figure 3.3: Effective potential diagrams for radial timelike geodesic trajectories. a) The black line and green lines indicate two families of radial timelike geodesics trajectories: those going towards $r = \infty$ and those going to $r = 0$. The "energy at infinity" is chosen so that these trajectories all eminate from the equilibrium $r = r_e$. These two families are the trajectories used as timelike coordinate curves for the LP coordinate system. b) The black line indicates a family of outgoing radial timelike geodesics. The "energy at infinity" for these trajectories is $E = 1$; it turns out this is the only value which produces flat spacelike hypersurfaces. This family of trajectories are used as the timelike coordinate curves of Lemaitre's coordinate system for SdS spacetime [22].

curves for these two coordinate systems, as shown on the effective potential diagram in figure 3.3. The geodesics in Lemaitre coordinates are all radially outgoing, whereas the geodesics in LP coordinates either fall into the singularity or drift out to $r = \infty$, always starting from $r = r_e$. In Lemaitre coordinates, the family of timelike coordinate curves are all essentially the same: the radial coordinate as a function of proper time, $r(\tau)$, is the same for all of them, and they only differ in their relationship between coordinate time and proper time. In LP coordinates, on the other hand, we have two families of trajectories whose radial coordinate as a function of proper time are all the same; those moving radially inward and those moving radially outward.

We construct LP coordinates in section 3.1. This is done by starting from Schwarzschild coordinates and integrating along the coordinate curves of LP coordinates. This allows us to find the explicit mapping from Schwarzschild to LP coordinates. This mapping is used in section 3.2 to find the metric components of SdS spacetime in LP coordinates. In section 3.3 we discuss some additional features of LP coordinates which will be useful in later sections.

3.1 Construction of LP coordinates

In terms of observers, the timelike coordinate curves of LP coordinates can be thought of as a set of observers drifting radially away from the equilibrium $r = r_e$, as shown on the effective potential diagram in figure 3.3a and on the Penrose diagram in figure 3.1a. One set of observers drift into the black hole while the other ones drift away from the black hole, caught up in the expansion of space. The time coordinate is defined as the time measured by clocks carried by each of these observers. We place a requirement that the clocks of two sufficiently nearby observers should, to lowest approximation in the distance between them, agree on the time of an event taking place on the midpoint of a rod connecting them. Mathematically this amounts to the requirement of having spacelike hypersurfaces $T =$ constant orthogonal to the timelike coordinate curves. The spatial coordinates consist of the angles (θ, ϕ), which remain constant for any of our observers, and a radial variable R, which is also constant for any observer, and whose value is the Schwarzschild

radius of the observer's position when that observer's clock value reads some chosen value $T = T^*$.

We now turn to the mathematical formulation of LP coordinates. Our construction of LP coordinates is based on four requirements. The first two requirements specify the coordinate curves of the coordinate system, and the second two requirements specify the coordinates which are to be used along these curves. Suppressing the (θ, ϕ) variables for the moment, let (T, R) be the LP coordinates. We require that:

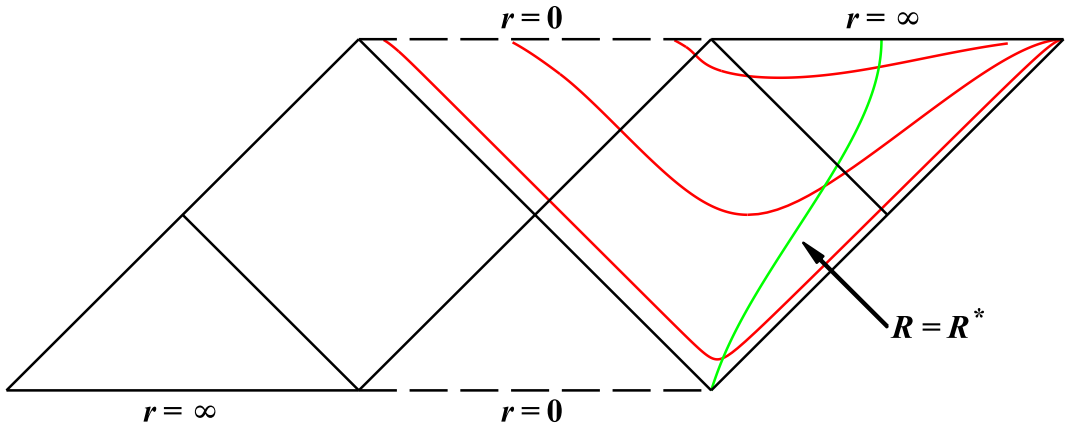
(i) The timelike coordinate curves $R =$ constant are radial timelike geodesics satisfying $r(\tau = -\infty) = r_e$, where $r(\tau)$ is the Schwarzschild radial coordinate of the trajectory, parametrized by proper time, and r_e is the equiibrium radius (see figure 3.3a for the effective potential diagram and figure 3.1a for a plot of the timelike coordinate curves on the Penrose diagram).

(ii) The spacelike coordinate curves $T =$ constant are orthogonal to the timelike coordinate curves (see figure 3.1b for a plot of the spacelike coordinate curves on the Penrose diagram).

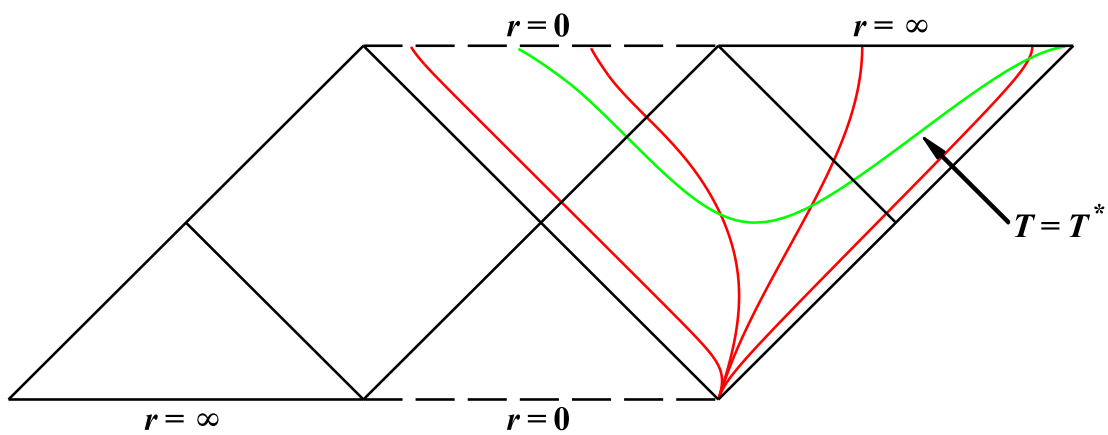
(iii) The timelike coordinate T is chosen such that any spacelike coordinate curve $T =$ constant intersects a chosen timelike coordinate curve $R = R^* > r_e$ at a point with coordinates $(T, R) = (\tau, R^*)$, where τ is the proper time along the curve $R = R^*$. This intersection of curves is shown in figure 3.4a. As will be shown in section 3.2, R^* can be thought of as essentially arbitrary, in the sense that its value does not affect the metric components in the LP coordinate system.

(iv) The spacelike coordinate R is chosen such that any curve $R =$ constant intersects a chosen spacelike coordinate curve $T = T^*$ with $R = r$, where r is the radial Schwarzschild coordinate of the intersection point. This intersection of curves is shown in figure 3.4b. As will be shown in section 3.2, the effect of changing T^* on the metric components in the LP coordinate system amounts to simply shifting the time variable T, and thus T^* is essentially arbitrary.

Translated into precise mathematical language, conditions (i) and (ii) specify the directions of the



(a)



(b)

Figure 3.4: a) The green curve is the arbitrarily chosen timelike coordinate curve $R = R^*$ and the red curves are the spacelike coordinate curves $T =$ constant. The value of the time coordinate T on any one of these spacelike curves is assigned to be $T = \tau$, where τ is the value of the proper time along the curve $R = R^*$ at the point of intersection between the two curves. b) A similar procedure is used to assign a value to R for the timelike coordinate curves $R =$ constant (shown in red). The coordinate R is defined such that $R = r$, where r is the value of the Schwarzschild coordinate at the intersection point of a timelike coordinate curve with the chosen spacelike curve $T = T^*$ (shown in green).

timelike and spacelike coordinate tangent vectors $\frac{\partial}{\partial T}$ and $\frac{\partial}{\partial R}$, respectively. Expressed in the basis of coordinate tangent vectors in Schwarzschild coordinates, $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial r}$, these directions are

$$\frac{\partial}{\partial T} \propto \frac{f(r_e)^{\frac{1}{2}}}{f(r)}\frac{\partial}{\partial t} + \mathrm{sgn}(r-r_e)\left[f(r_e)-f(r)\right]^{1/2}\frac{\partial}{\partial r}, \tag{3.2}$$

$$\frac{\partial}{\partial R} \propto \mathrm{sgn}(r-r_e)\left[f(r_e)-f(r)\right]^{\frac{1}{2}}\frac{\partial}{\partial t} + f(r_e)^{1/2}f(r)\frac{\partial}{\partial r}. \tag{3.3}$$

The first relation above comes from the geodesic equations for timelike radial geodesics (equations (2.11)-(2.12)). The second relation is obtained by simply requiring that $\langle \frac{\partial}{\partial T}, \frac{\partial}{\partial R}\rangle = 0$. Conditions (iii) and (iv) can also be translated into precise mathematical language. Letting $r(T,R)$ be the function mapping (T,R) coordinates to the Schwarzschild coordinate r, we have:

(iii) $\Leftrightarrow r(T,R^*) = r_0(\tau = T; R^*)$, where $r_0(\tau; R^*)$ is a timelike geodesic satisfying $r_0(\tau = -\infty) = r_e$ and $r_0(\tau = T^*) = R^*$,

(iv) $\Leftrightarrow r(T^*,R) = R$.

Consider any of the timelike coordinate curves $R =$ constant. Suppressing the coordinate R for a moment since it is held constant, let $r(T)$ and $t(T)$ be the Schwarzschild coordinates along this coordinate curve. From condition (i) above, we have that $r(T)$ is a monotonic function of T (except for the one curve where $r(T) = r_e$ for all T). Therefore we can define the inverse $T(r)$ and the function $t(r) = t(T(r))$. By dividing the components of the tangent vector in equation (3.2) we obtain that $t(r)$ satisfies the following:

$$\left.\frac{dt}{dr}\right|_{R=\text{constant}} = \frac{f(r_e)^{1/2}}{f(r)\left(f(r_e)-f(r)\right)^{1/2}}\mathrm{sgn}(r-r_e). \tag{3.4}$$

Next we apply the same procedure to a spacelike coordinate curve $T =$ constant. We suppress the T coordinate and consider the functions $r(R)$ and $t(R)$ along the curve. For $T = T^*$, we have $r(R) = R$ and the function $r(R)$ is monotonic. It then follows by continuity that for values of T sufficiently near T^*, $r(R)$ is once again monotonic. Therefore, for values of T near T^*, we can define the

function $t(r) = t(R(r))$. We divide the components of the tangent vector in equation (3.3) to obtain

$$\left.\frac{dt}{dr}\right|_{T=\text{constant}} = \frac{(f(r_e) - f(r))^{1/2}}{f(r_e)^{1/2} f(r)} \operatorname{sgn}(r - r_e)\,. \tag{3.5}$$

We will obtain the explicit mapping from LP coordinates (T,R) to Schwarzschild coordinates (t,r) by integrating equations (3.4) and (3.5). Note that the right hand side of (3.4) is ill-defined where $f(r) = 0$ and where $f(r) = f(r_e)$ (i.e. at $r = r_b$, $r = r_c$ and $r = r_e$) and the right hand side of (3.5) is ill-defined for $f(r) = 0$. This difficulty can be overcome by temporarily restricting the coordinate transformation from LP to Schwarzschild coordinates to any region small enough that it does not intersect $r = r_b$, $r = r_e$ or $r = r_c$. We have thus placed two temporary restrictions on the coordinate mapping from LP to Schwarzschild coordinates: the first is that T be sufficiently near T^* and the other is that the region considered not intersect $r = r_b$, $r = r_e$ or $r = r_c$. Once the explicit mapping from LP coordinates to Schwarzschild coordinates in this restricted region is obtained, these temporary restriction will be lifted and the mapping will be extended to the full Schwarzschild region (i.e. the region shown in figure 3.1). This will be done by simply defining the LP coordinates using the formula obtained for the functions $t(T,R)$ and $r(T,R)$. The justification for this procedure will ultimately come from the fact that the resulting metric components in LP coordinates will be well defined over the region of spacetime of interest (the metric components will be derived in section 3.2).

Let us now derive the functions $r(T,R)$ and $t(T,R)$. To do this, consider two paths joining an arbitrary point (T,R) to the point (T^*,R^*), where it is understood that (T,R) and R^* are chosen so that the restrictions mentioned in the previous paragraph are met. The first path connects the point (T,R) to (T^*,R^*) by first going along $R =$ constant until the point (T^*,R), and then going along the path $T = T^*$ until the point (T^*,R^*) is reached. The second path connects the point (T,R) to (T^*,R^*) by first going along $T =$ constant until the point (T,R^*), and then going along the path $R = R^*$ until the point (T^*,R^*) is reached. Notice that both of these paths consist entirely of LP coordinate curves.

We now integrate dt/dr along these two paths. Moving along the first path, we integrate (3.4) from

(T,R) to (T^*,R) along the coordinate curve $R=$ constant, and then integrate (3.5) from (T^*,R) to (T^*,R^*) along the coordinate curve $T=T^*$. This gives the following equation:

$$\int_{r(T,R)}^{r(T^*,R)} \left.\frac{dt}{dr}\right|_{R=\text{constant}} dr + \int_{r(T^*,R)}^{r(T^*,R^*)} \left.\frac{dt}{dr}\right|_{T=\text{constant}} dr = t(T^*,R^*) - t(T,R)\,. \tag{3.6}$$

Moving along the second path, we integrate (3.5) from (T,R) to (T,R^*) along the coordinate curve $T=$ constant, and then integrate (3.4) from (T,R^*) to (T^*,R^*) along the coordinate curve $R=$ constant. This gives the following equation:

$$\int_{r(T,R)}^{r(T,R^*)} \left.\frac{dt}{dr}\right|_{T=\text{constant}} dr + \int_{r(T,R^*)}^{r(T^*,R^*)} \left.\frac{dt}{dr}\right|_{R=\text{constant}} dr = t(T^*,R^*) - t(T,R)\,. \tag{3.7}$$

We can find the function $r(T,R)$ by first combining (3.6) and (3.7) into the following equation:

$$\int_{r(T,R)}^{r(T^*,R)} \left.\frac{dt}{dr}\right|_{R=\text{constant}} dr + \int_{r(T^*,R)}^{r(T^*,R^*)} \left.\frac{dt}{dr}\right|_{T=\text{constant}} dr = \int_{r(T,R)}^{r(T,R^*)} \left.\frac{dt}{dr}\right|_{T=\text{constant}} dr + \int_{r(T,R^*)}^{r(T^*,R^*)} \left.\frac{dt}{dr}\right|_{R=\text{constant}} dr\,.$$

Manipulating the limits of integration in the above and rearranging terms, we obtain the following:

$$\int_{r(T^*,R)}^{r(T,R)} \left(\left.\frac{dt}{dr}\right|_{R=\text{constant}} - \left.\frac{dt}{dr}\right|_{T=\text{constant}} \right) dr = \int_{r(T^*,R^*)}^{r(T,R^*)} \left(\left.\frac{dt}{dr}\right|_{R=\text{constant}} - \left.\frac{dt}{dr}\right|_{T=\text{constant}} \right) dr\,. \tag{3.8}$$

Using (3.4) and (3.5), we have

$$\left.\frac{dt}{dr}\right|_{R=\text{constant}} - \left.\frac{dt}{dr}\right|_{T=\text{constant}} = \frac{1}{f(r_e)^{1/2}} \frac{\operatorname{sgn}(r-r_e)}{(f(r_e)-f(r))^{1/2}}\,.$$

From condition (iii) above, and the geodesic equation (2.12), we have that $r(T,R)$ satisfies the following:

$$\frac{\partial r}{\partial T} = \operatorname{sgn}(r-r_e)(f(r_e)-f(r))^{1/2} \tag{3.9}$$

Combining the previous two equations, we have

$$\left.\frac{dt}{dr}\right|_{R=\text{constant}} - \left.\frac{dt}{dr}\right|_{T=\text{constant}} = \frac{1}{f(r_e)^{1/2}}\left(\frac{\partial r}{\partial T}\right)^{-1}. \tag{3.10}$$

Substituting the above into (3.8), we obtain

$$\int_{r(T^*,R)}^{r(T,R)} \left(\frac{\partial r}{\partial T}\right)^{-1} dr = \int_{r(T^*,R^*)}^{r(T,R^*)} \left(\frac{\partial r}{\partial T}\right)^{-1} dr\,. \tag{3.11}$$

Interpreted physically, the above is the statement that the timelike curves $R =$ constant correspond to a family of observers whose clocks are synchronized. It is for this reason that Gaussian normal coordinates are sometimes also called synchronous coordinates. Let us evaluate the integral on the right hand side above explictly. We do so by making a change of variables. Fixing $R = R^*$, define $r(T) = r(T,R^*)$. $r(T)$ is a monotonic function and can be inverted to give the function $T(r)$. Using $T(r)$ as the new variable of integration, the integral becomes

$$\int_{r(T^*,R^*)}^{r(T,R^*)} \left(\frac{\partial r}{\partial T}\right)^{-1} dr = \int_{T^*}^{T} d\tau = T - T^*\,. \tag{3.12}$$

Using the above and (3.9) in (3.11), we obtain the following implicit equation for the function $r(T,R)$:

$$\boxed{\int_{R}^{r(T,R)} \frac{d\rho}{(f(r_e)-f(\rho))^{1/2}} = \text{sgn}(R-r_e)(T-T^*)\,,} \tag{3.13}$$

where we have used $r(T^*,R) = R$ (condition (iv) above). We have also removed $\text{sgn}(r-r_e)$ from the integral and replaced it with $\text{sgn}(R-r_e)$. It is possible to do this because our definition of the timelike coordinate curves (condition (i) at the beginning of this section) ensures that none of them cross $r = r_e$, and furthermore our definition of the coordinate R (condition (iv) above) ensures that $\text{sgn}(r-r_e) = \text{sgn}(R-r_e)$. The integral on the left hand side above can be found analytically. Once this is done it is possible to write an explicit formula for $r(T,R)$ in terms of elementary functions and their inverses. This will only be useful for the purpose of extending the coordinate mapping to all values of the Schwarzschild coordinates, and so we relegate it to section 3.3.3.

To obtain the function $t(T,R)$, we manipulate the limits of integration and rearrange terms in (3.7) to get

$$t(T,R) = t(T^*,R^*) + \int_{r(T^*,R^*)}^{r(T,R^*)} \left(\left.\frac{dt}{dr}\right|_{R=\text{constant}} - \left.\frac{dt}{dr}\right|_{T=\text{constant}} \right) dr + \int_{r(T^*,R^*)}^{r(T,R)} \left.\frac{dt}{dr}\right|_{T=\text{constant}} dr\,. \quad (3.14)$$

The first integral on the right hand side above is the same as the right hand side of (3.8), and has in fact already been found. Combining (3.10) and (3.12), we have

$$\int_{r(T^*,R^*)}^{r(T,R^*)} \left(\left.\frac{dt}{dr}\right|_{R=\text{constant}} - \left.\frac{dt}{dr}\right|_{T=\text{constant}} \right) dr = \frac{1}{f(r_e)^{1/2}}(T-T^*)\,.$$

Substituting the above and (3.5) into (3.14), we obtain the following formula for the function $t(T,R)$:

$$\boxed{t(T,R) = t(T^*,R^*) + \frac{1}{f(r_e)^{1/2}}(T-T^*) + \int_{r(T^*,R^*)}^{r(T,R)} \frac{(f(r_e)-f(\rho))^{1/2}}{f(r_e)^{1/2} f(\rho)} \operatorname{sgn}(\rho - r_e) d\rho\,.} \quad (3.15)$$

The integral on the right hand side can be found analytically. Again, this explicit formula is not necessary except for the purpose of extending the mapping, and so we relegate it to section 3.3.3. Note that the function $t(T,R)$ above is only known once the function $r(T,R)$ has been found from (3.13). The equations (3.13) and (3.15), taken together, provide a change of coordinates from Schwarzschild coordinates (t,r) to the LP coordinates (T,R). In so doing they allow us to find the metric components of SdS spacetime in LP coordinates, as will be done in the next section. Additional details regarding LP coordinates will be dealt with in section 3.3. For example, in section 3.3.3, we deal with the issue of ensuring that (3.13) and (3.15) provide a mapping over the whole region of SdS spacetime covered by Schwarzschild coordinates. That is, we ensure that the mapping from LP coordinates to Schwarzschild coordinates can indeed be extended beyond the restricted region for which it was originally derived.

3.2 Metric components in LP coordinates

We are interested in obtaining the metric components of SdS spacetime in LP coordinates, and so we must first obtain the four components of the Jacobian matrix of partial derivatives associated with the coordinate transformation from Schwarzschild to LP coordinates. This is achieved by taking partial derivatives of (3.13) and (3.15). Taking the partial derivative of (3.13) with respect to T or R, we get the first two Jacobian components:

$$\frac{\partial r}{\partial T} = \operatorname{sgn}(r-r_e)\,(f(r_e)-f(r))^{\frac{1}{2}}\,, \tag{3.16}$$

$$\frac{\partial r}{\partial R} = \left[\frac{f(r_e)-f(r)}{f(r_e)-f(R)}\right]^{1/2}\,. \tag{3.17}$$

Next, taking the partial derivative of (3.15) and using the above where necessary, we get the other two components of the Jacobian:

$$\frac{\partial t}{\partial T} = \frac{f(r_e)^{\frac{1}{2}}}{f(r)}, \tag{3.18}$$

$$\frac{\partial t}{\partial R} = \frac{[f(r_e)-f(r)]\operatorname{sgn}(R-r_e)}{f(r_e)^{1/2}f(r)\,[f(r_e)-f(R)]^{1/2}}\,. \tag{3.19}$$

In the above, it is understood that $r = r(T,R)$ is a function of T and R. Using these four Jacobian components, we can apply the tensor transformation law to the Schwarzschild metric components (2.4) to obtain the line element in LP coordinates. This yields

$$ds^2 = dT^2 - \frac{f(r_e)-f(r)}{f(r_e)\,[f(r_e)-f(R)]}dR^2 - r^2(d\theta^2+\sin^2\theta d\phi^2)\,, \tag{3.20}$$

where $r = r(T,R)$ is given implicitly by (3.13). The constant R^*, which was used as part of the definition of the coordinates (condition (iii) in section 3.1), does not appear explicitly in the above line element, nor does it appear in the equation (3.13) defining $r(T,R)$. Thus we see that the value of R^* has no impact on the metric and is essentially arbitrary. The constant T^* (defined in condition (iv) of section 3.1) affects the above line element through its presence in (3.13). However, as will be discussed in section 3.3.5, there is a freedom in shifting the variable T by a constant amount. For

this reason, T^* can be thought of as arbitrary and we will take $T^* = 0$ from now on. Notice that the metric component g_{RR} above is undefined at $R = r_e$. However, we can define g_{RR} at $R = r_e$ in such a way that it is continuous, as will be discussed in section 3.3.4.

3.3 Additional features of LP coordinates

In what follows, we describe some additional features of LP coordinates which will be useful in later sections, or which shed further light on the coordinate system.

3.3.1 The case $M = 0$: deSitter spacetime in planar coordinates

In the case that $M = 0$, (3.20) reduces to the deSitter line element in planar coordinates, as in (3.1). This can be seen by first noting that when $M = 0$, we have:

$$\begin{aligned} f(r) &= 1 - H^2 r^2, \\ f(r_e) &= 1\,, \end{aligned}$$

where r_e is given by (2.6) and $f(r)$ is given by (2.5). Using the above, the equation (3.13) defining $r(T,R)$ can be readily integrated to give

$$r(T,R) = R\,e^{HT}\,,$$

where recall that $T^* = 0$. Substituting the above into (3.20) and setting $M = 0$, we get

$$ds^2 = dT^2 - e^{2HT}\left(dR^2 + R^2(d\theta^2 + \sin^2\theta d\phi^2)\right)\,, \tag{3.21}$$

which is precisely the same as (3.1). The above is the line element of an FRW cosmology with flat spatial hypersurfaces and an exponentially growing scale factor. It has traditionally been asso-

ciated with the cosmology of a "steady state" universe obeying the perfect cosmological principle. The above line element also approximates the geometry of a universe going through a period of inflation. The coordinates (T, R, θ, ϕ) are often called planar coordinates, though in the context of inflation the name inflationary coordinates is used as well. Mathematically, the coordinates cover the upper half of the hyperboloid corresponding to deSitter spacetime, as illustrated in figure 2.5. Physically, the coordinate curves $R =$ constant correspond to a family of freely floating observers in a homogeneous and isotropic universe, and the coordinate T corresponds to proper time measured by these observers. The coordinate R is sometimes called the comoving distance and the coordinate T is sometimes called the cosmic time.

3.3.2 The case $H = 0$: Schwarzschild spacetime in Lemaitre coordinates

In the case that $H = 0$, we have

$$\begin{aligned} f(r) &= 1 - \frac{2M}{r}, \\ f(r_e) &= 1\,, \end{aligned}$$

where r_e is given by (2.6) and $f(r)$ is given by (2.5). The line element (3.20) reduces to

$$ds^2 = dT^2 - \left(\frac{R}{r}\right) dR^2 - r^2 d\Omega^2\,. \tag{3.22}$$

Let us introduce the new variable ρ as

$$\rho = \frac{2}{3}\frac{R^{3/2}}{(2M)^{1/2}}\,. \tag{3.23}$$

(3.22) now becomes

$$ds^2 = dT^2 - \frac{2M}{r} d\rho^2 - r^2 d\Omega^2\,.$$

The above is the line element of Schwarzschild spacetime, with (T, ρ, θ, ϕ) the well known Lemaitre coordinates [22]. In the case that $H = 0$, the LP coordinates are therefore closely related to the

Lemaitre coordinates, the only difference being the redefinition of the radial variable through (3.23). The function $r(T,\rho)$ in the line element above can be found by first obtaining $r(T,R)$ from (3.13), and then replacing R by ρ using (3.23). This gives

$$r(T,R) = \left(\frac{3}{2}\right)^{2/3} (2M)^{1/3}(\rho - T)^{2/3}\,.$$

The region of spacetime covered by these coordinates is the black hole half of Schwarzschild spacetime (the upper half in figure 2.1). The spacelike hypersurfaces $T =$ constant in this coordinate system are intrinscially flat. To see this, first set $T =$ constant in (3.22) to get

$$ds^2 = -\left(\frac{R}{r}\right) dR^2 - r^2 d\Omega^2\,. \tag{3.24}$$

Now consider the Jacobian component (3.17) with $H = 0$:

$$\frac{\partial r}{\partial R} = \left(\frac{R}{r}\right)^{1/2}\,.$$

Using the above to transform the line element (3.24) to (r,θ,ϕ) coordinates, we obtain the line element of three dimensional Euclidean space in spherical coordinates. Physically, the coordinate curves $R =$ constant correspond to a family of observers falling into the black hole from infinity, and all with the same "energy at infinity". They only differ in the reading on their clocks when they pass through some chosen value of the Schwarzschild radius r. For example, an observer with coordinate $R = R_0$ would reach the singularity $r = 0$ at time $T = \frac{2}{3}\sqrt{\frac{R_0^3}{2M}}$. The coordinate T measures proper time as measured by any one of these observers.

3.3.3 Explicit coordinate transformation from Schwarzschild to LP coordinates

In deriving the coordinate transformation from LP to Schwarzschild coordinates, as given by equations (3.13) and (3.15), we placed certain restrictions on the values of the coordinates (T,R). This was to avoid certain technical difficulties, such as the presence of a divergent integrand in (3.15). The easiest way to overcome these restrictions is to perform the integrals in (3.13) and (3.15) ana-

lytically, and then use the resulting explicit form of the coordinate transformation as the definition of LP coordinates. Thus the true definition of the LP coordinates is provided in this section, in the sense that the definition given by equations (3.13) and (3.15) does not apply to the entire spacetime region of interest. Since the explicit coordinate transformation given in this section is obtained by finding the antiderivatives of the integrands in (3.13) and (3.15), it will lead to the same Jacobian, as given by (3.16)-(3.19), and thus to the same line element (3.20).

We start by giving the mapping from LP to Schwarzschild coordinates. By performing the integral in (3.13), we obtain the following formula for the Schwarzschild coordinate r as a function of the LP coordinates (T,R):

$$\boxed{r(T,R) = \begin{cases} r_e F_1^{-1}\left(F_1\left(\frac{R}{r_e}\right)e^{HT}\right), & R \leq r_e \\ r_e F_2^{-1}\left(F_2\left(\frac{R}{r_e}\right)e^{HT}\right), & R \geq r_e \end{cases}} \tag{3.25}$$

where

$$\begin{aligned} F_1(\rho) &= \left(\rho+1+\sqrt{\rho(\rho+2)}\right)\left[\frac{1-\rho}{\sqrt{3\rho(\rho+2)}+2\rho+1}\right]^{\frac{1}{\sqrt{3}}}, \\ F_2(\rho) &= \left(\rho+1+\sqrt{\rho(\rho+2)}\right)\left[\frac{\rho-1}{\sqrt{3\rho(\rho+2)}+2\rho+1}\right]^{\frac{1}{\sqrt{3}}}. \end{aligned} \tag{3.26}$$

Recall that we have taken $T^* = 0$ in the above. The functions $F_1(\rho)$ and $F_2(\rho)$ are strictly monotonic, and therefore invertible, for $\rho \leq 1$ and $\rho \geq 1$, respectively. The ranges of these functions are $F_1 \geq 0$ and $F_2 \geq 0$, so that the domains of the inverses $F_1^{-1}(y)$ and $F_2^{-1}(y)$ are well defined for all $y \geq 0$.

The formula for the Schwarzschild time coordinate $t(T,R)$ is a little more involved. First we rewrite (3.15) as

$$t(T,R) = t^* + \frac{1}{f(r_e)^{1/2}}\left(T + \int_{r^*}^{r} \frac{\mathrm{sgn}(\rho - r_e)\left(f(r_e)-f(\rho)\right)^{1/2}}{f(\rho)} d\rho\right), \tag{3.27}$$

where $t^* = t(T^*,R^*)$ and $r^* = r(T^*,R^*)$ and we have once again taken $T^* = 0$. Recall that R^* does not appear anywhere in the line element (3.20), so that its choice is essentially arbitrary. In addition

to the freedom in choosing R^*, there is a freedom in shifting the variable T by an arbitrary constant amount (this is discussed in section 3.3.5 below). Because of these two freedoms, the choice for the values of t^* and r^* are arbitrary, and we take $t^* = 0$ and $r^* = (r_e + r_c)/2$ without loss of generality. By evaluating the integral in the equation above, we obtain the following formula for $t(T,R)$:

$$\boxed{t(T,R) = \frac{1}{f(r_e)^{1/2}} \left(T + A\left(r(T,R)\right) - A\left(r^*\right)\right),} \tag{3.28}$$

where $r(T,R)$ is given above, and $A(\rho)$ is the antiderivative of the integrand in (3.27):

$$\begin{aligned} A(\rho) \quad = \quad & \frac{\sqrt{r_b^2 + 2r_b r_e}\,(r_e - r_b)}{H\,(r_a - r_b)\,(r_b - r_c)} \operatorname{arctanh}\left(\frac{r_b r_e + r_e\rho + r_b\rho}{\sqrt{r_b^2 + 2r_b r_e}\sqrt{\rho^2 + 2\rho r_e}}\right) \\ & - \frac{\sqrt{r_c^2 + 2r_c r_e}\,(r_e - r_c)}{H\,(r_a - r_c)\,(r_b - r_c)} \operatorname{arctanh}\left(\frac{r_c r_e + r_e\rho + r_c\rho}{\sqrt{r_c^2 + 2r_c r_e}\sqrt{\rho^2 + 2\rho r_e}}\right) \\ & - \ln\left(1 + \frac{\rho}{r_e} + \frac{1}{r_e}\sqrt{\rho^2 + 2r_e\rho}\right) \\ & - \frac{\sqrt{r_a^2 + 2r_a r_e}\,(r_e - r_a)}{H\,(r_a - r_b)\,(r_a - r_c)} \ln\left(2\frac{r_a r_e + r_e\rho + r_a\rho + \sqrt{r_a^2 + 2r_a r_e}\sqrt{\rho^2 + 2\rho r_e}}{r_e\,(\rho - r_a)}\right). \end{aligned} \tag{3.29}$$

In the above, r_a, r_b and r_c are the roots of $f(r) = 0$, with $r_a < 0 < r_b < r_e < r_c$. Notice that the mapping from LP to Schwarzschild coordinates, as given by (3.25)-(3.29), is asymptotically discontinuous at $r = r_b$ and $r = r_c$. That is, as one approaches values of (T,R) such that $r(T,R) = r_b$ or $r(T,R) = r_c$, one finds that $t \to \infty$ or $t \to -\infty$. These discontinuities are a consequence of the fact that the Schwarzschild coordinates are not defined along $r = r_b$ and $r = r_c$. Therefore, strictly speaking, the equations (3.25)-(3.29) provide three mappings from LP coordinates to Schwarzschild coordinates: one for $0 < r < r_b$, one for $r_b < r < r_c$, and one for $r > r_c$. For any value of the Schwarzschild variables (t,r) in any one of these three regions, we can solve (3.25) and (3.28) for a unique value of the LP coordinates (T,R), with $R > 0$. This leads to the inverse mapping, which takes Schwarzschild coordinates into LP coordinates, and is given by

$$R(t,r) = \begin{cases} r_e F_1^{-1}\left(F_1\left(\frac{r}{r_e}\right) e^{-HT(t,r)}\right), & r \le r_e \\ r_e F_2^{-1}\left(F_2\left(\frac{r}{r_e}\right) e^{-HT(t,r)}\right), & r \ge r_e \end{cases} \tag{3.30}$$

$$T(t,r) = f(r_e)^{1/2}t + A(r^*) - A(r)\,. \tag{3.31}$$

(3.30)-(3.31) provide the inverse mapping over any one of the three regions previously discussed. We have thus constructed a coordinate mapping from LP to Schwarzschild coordinates which covers the region of SdS spacetime of interest, namely the region covered by Schwarzschild coordinates, as shown in figure 2.7b (where it is understood that points with $r = r_b$ and $r = r_c$ should not be included in the region covered by Schwarzschild coordinates). The fact that the coordinate transformation from LP to Schwarzschild coordinates is not defined for $r = r_b$ and $r = r_c$ does not present any real difficulty, since in the end what matters is that the metric components in LP coordinates, as given by equation (3.20), are well defined for all values of r. The issue of the continuity of the metric components is dealt with in section 3.3.4.

3.3.4 Continuity of metric components

Consider the g_{RR} component of the line element (3.20):

$$g_{RR} = -\frac{f(r_e) - f(r)}{f(r_e)\left[f(r_e) - f(R)\right]}\,, \tag{3.32}$$

where $r = r(T,R)$, so that the above is a function of both T and R. Notice that the above is undefined at $R = r_e$. However, by the very definition of our coordinate system (see conditions (i) and (iv) in section 3.1), we have $r(T,R_e) = r_e$, so that the numerator vanishes as well as $R \to r_e$. We will now show that the limit of g_{RR} as $R \to r_e$ is well-defined, so that by defining g_{RR} at $R = r_e$ as being equal to this limit, we will have a continuous metric. Using (3.17), we can rewrite the above as

$$g_{RR} = -\frac{1}{f(r_e)}\left(\frac{\partial r}{\partial R}\right)^2\,. \tag{3.33}$$

From the above, we see that we can find the limit of g_{RR} as $R \to r_e$ by calculating $\partial r/\partial R$ at $R = r_e$. To do this, we find the lowest order behavior of $r(T,R)$ near R. First consider $F_1(\rho)$ and $F_2(\rho)$ in

(3.26). For ρ near 1, these have the following asymptotic forms:

$$\begin{aligned} F_1(\rho) &\sim \left(2+\sqrt{3}\right)\left(\frac{1-\rho}{6}\right)^{\frac{1}{\sqrt{3}}} , \\ F_2(\rho) &\sim \left(2+\sqrt{3}\right)\left(\frac{\rho-1}{6}\right)^{\frac{1}{\sqrt{3}}} . \end{aligned}$$

The inverses $F_1^{-1}(y)$ and $F_2^{-1}(y)$ therefore have the following asymptotic forms:

$$\begin{aligned} F_1^{-1}(y) &\sim -6\left(\frac{y}{2+\sqrt{3}}\right)^{\sqrt{3}}+1 , \\ F_2^{-1}(y) &\sim 6\left(\frac{y}{2+\sqrt{3}}\right)^{\sqrt{3}}+1 . \end{aligned}$$

Substituting the two preceeding equations into 3.25, we obtain

$$r(T,R) = r_e F_i^{-1}\left(F_i\left(\frac{R}{r_e}\right)e^{HT}\right) \sim r_e + (R-r_e)e^{\sqrt{3}HT} ,$$

where $i = 1,2$. Notice that the asymptotic form of $r(T,R)$ near $R = r_e$ is the same for both $R \leq r_e$ and $R \geq r_e$. Also, as one would expect, $r(T,R)$ is linear in R near $R = r_e$. From the above we can compute that

$$\lim_{R\to r_e} \frac{\partial r}{\partial R} = e^{\sqrt{3}HT} .$$

Substituting the above in (3.33), we get

$$\lim_{R\to r_e} g_{RR} = -\frac{1}{f(r_e)}e^{2\sqrt{3}HT} ,$$

so that the limit of g_{RR} as $R \to r_e$ is well defined. Let us now rewrite the line element for LP coordinates in (3.20) as

$$ds^2 = dT^2 - g_{RR}\,dR^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) ,$$

with

$$g_{RR} = \begin{cases} -\frac{f(r_e)-f(r)}{f(r_e)[f(r_e)-f(R)]} & \text{for } R \neq r_e, \\ -\frac{1}{f(r_e)}e^{2\sqrt{3}HT} & \text{for } R = r_e . \end{cases}$$

Using the above definition, we have a line element which is defined and continuous for all values of $R > 0$ and $-\infty < T < \infty$. As was shown in section 3.3.3, these values of LP coordinates cover the whole region of SdS spacetime relevant to this thesis (i.e. the region covered by the coordinate curves shown in figure 3.1). Thus we have a well-defined and continuous line element over the region of spacetime considered in this thesis. This will be essential when computing the cosmological and black hole horizons in chapter 4.

3.3.5 Killing symmetries

It was claimed in section 3.3.3 that there is a fundamental ambiguity in the definition of the LP coordinate T, and that one can add an arbitrary constant to this variable and still obtain essentially the same coordinate system. In this section, we show how this ambiguity in the definition of T is related to one of the Killing symmetries of SdS spacetime, and how moving along the flow associated with this Killing field maps the LP coordinate curves onto themselves, and therefore does not change the coordinate system in any fundamental way.

The absence of the Schwarzschild coordinate t in the line element (2.4) allows us to deduce the existence of the Killing vector $\xi^{(t)} = \partial_t$ for SdS spacetime. The integral curves of this Killing field are those for which the other three Schwarzschild coordinates (r, θ, ϕ) are held constant. To understand the role of this Killing symmetry in LP coordinates, we first introduce a new radial coordinate ρ as

$$\rho(R) = \frac{1}{H} \ln\left(F_2\left(\frac{R}{r_e}\right)\right) , \tag{3.34}$$

where F_2 is defined in (3.26). In the above, and in everything that follows, we are restricting our attention to $R > r_e$. Notice that this corresponds to $-\infty < \rho < \infty$. The role of the Killing symmetry in SdS spacetime is the same for both $R < r_e$ and $R > r_e$, so that considering only $R > r_e$ is sufficient. Furthermore, the Killing symmetry does not map points with $R < r_e$ to points with $R > r_e$ or vice-versa, so that these two regions can be considered separately without any difficulty. From the above,

we have

$$d\rho = (f(r_e) - f(R))^{-1/2}\, dR\ .$$

Using the above, we can transform the line element (3.20) from LP coordinates into (T, ρ, θ, ϕ) coordinates to get

$$ds^2 = dT^2 - (f(r_e) - f(r))\, d\rho^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2)\ , \tag{3.35}$$

where $r = r(T, \rho)$. By combining (3.34) with (3.25), we find that $r(T, \rho)$ is given by

$$r(T, \rho) = r_e F_2^{-1}\left(e^{H(\rho+T)}\right)\ .$$

Since the line element (3.35) only depends on ρ and T through $r(T, \rho)$, it follows that any transformation of ρ and T that does not change r will be a symmetry of the spacetime. By examining the expression for $r(T, \rho)$ above, we have that for any constant T_0, the transformation

$$\begin{aligned} T &\rightarrow T + T_0, \\ \rho &\rightarrow \rho - T_0 \end{aligned}$$

will leave r unchanged and will be a symmetry of the spacetime. The above symmetry transformation is precisely the same as the one associated with the Killing field $\xi^{(t)}$ in Schwarzschild coordinates, as can be deduced by simply realizing that both have the same integral curves, namely those obtained by setting (r, θ, ϕ) to constant values. Using (3.34), we can take the above transformation and find the equivalent transformation in LP coordinates. This gives

$$T \rightarrow T + T_0, \tag{3.36}$$

$$R \rightarrow r_e F_2^{-1}\left(F_2\left(\frac{R}{r_e}\right)e^{-HT_0}\right)\ . \tag{3.37}$$

Since the right hand sides of (3.36) and (3.37) depend only on T and R, respectively, the above transformation takes the LP coordinate curves $T =$ constant and $R =$ constant and maps them onto themselves. Thus the LP coordinate system that we have created is in fact a family of coordinate systems, all of which have the same line element as given by (3.20), and which are all related by the

transformation given above. Which of these coordinate systems we choose is arbitary. This choice of coordinates is the freedom in shifting the coordinate T to we alluded to in section 3.3.3. Note that this freedom is still present even after one has chosen the values for the constants T^* and R^* introduced in section 3.1.

3.3.6 Spacelike hypersurfaces: intrisinsic geometry

Let us consider the spacelike hypersurfaces $T =$ constant. The line element on these hypersurfaces is found by setting $dT = 0$ in (3.20). This gives

$$ds^2 = -\frac{f(r_e) - f(r)}{f(r_e)\left[f(r_e) - f(R)\right]} dR^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{3.38}$$

Now consider using the coordinates (r, θ, ϕ) on the spacelike hypersurface $T =$ constant, where r is the Schwarzschild radial variable first introduced in (2.4). When used as a coordinate on the hypersurface $T =$ constant, r is a spacelike variable (this would not be the case if one attempted to combine the variable r with the time coordinate T from LP coordinates into a "hybrid" coordinate system (T, r, ϕ, θ); in such a case the variable r would be timelike at some locations in spacetime, much as it is in Schwarzschild coordinates). From (3.17), we have

$$dR = \left[\frac{f(r_e) - f(r)}{f(r_e) - f(R)}\right]^{-1/2} dr \ .$$

Substituting the above into (3.20), we get

$$ds^2 = -\frac{1}{f(r_e)} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{3.39}$$

We recognize the above as the line element for a three dimensional cone with opening angle of $\beta = \arcsin\left(f(r_e)^{1/2}\right)$. Unlike the familiar two dimensional cone, this three dimensional cone is not locally intrinsically flat in the sense of having a vanishing Riemann curvature tensor. In fact the

Riemann tensor can be calculated to be

$$R_{\theta\phi\theta\phi} = 3r^2(\sin^2\theta)\varepsilon^{2/3} \ ,$$

where it is understood that there are other non-zero components of the Riemann tensor, related to the above by an appropriate shuffling of indices. The associated Ricci scalar is

$$\mathscr{R} = -\frac{6}{r^2}\varepsilon^{2/3} \ ,$$

where recall that $\varepsilon = HM$. The above reveals a curvature singularity at $r = 0$. Recognizing that the spacelike hypersurfaces are not flat will be important in section 4.4.3, where we find that the late time shape of the horizons is not perfectly spherical, in spite of the fact that the spacetime at late times is well approximated by deSitter spacetime. This non-sphericity of the horizon is an artifact of choosing spacelike hypersurfaces which are not intrinsically flat, as will be explained.

3.3.7 Spacelike hypersurfaces: polar coordinates

In this section we describe the coordinates which will be used on the spacelike hypersurfaces $T =$ constant when analyzing and illustrating the horizon in chapter 4. The horizon we will find can be thought of as a series of 2-surfaces, each living on one of the hypersurfaces $T =$ constant. We will use (r, ϕ, θ) as coordinates on these hypersurfaces, leading to the line element (3.39), and the 2-surfaces will be visualized as living in 3d Euclidean space, with (r, ϕ, θ) playing the role of spherical coordinates. This way of representing the horizon introduces metrical distortions since the hypersurfaces $T =$ constant have the geometry of a cone and are not truly Euclidean. However, visualizing the horizon as a series of 2-surfaces in 3d Euclidean space has the advantage that it easily allows one to illustrate the basic change in horizon topology that occurs during merger. As well, it will be useful for identifying and illustrating key features of the horizons, such as the presence of caustic points. In our graphical representation we will also suppress the θ coordinate by setting $\theta = \pi/2$, so that the horizon will be illustrated as a series of curves in the Euclidean plane, with

(r, ϕ) as polar coordinates. The spherical symmetry of SdS spacetime allows us to recover any null geodesic from one constrained to $\theta = \pi/2$ by a simple rotation, and the full 2-surface can be visualized by rotating the curves (such as those in figure 4.4) about the x-axis. Another advantage of using (r, ϕ, θ) instead of (R, ϕ, θ) coordinates is that the line element (3.39) is considerably simpler than (3.38), and this will greatly simplify calculations of horizon area in chapter 5.

Chapter 4

Structure of Merging Black Hole and Cosmological Horizons

4.1 Introduction

The observer dependent *causal horizon* we would like to find is defined as the boundary of the causal past of the trajectory of the observer moving radially away from the black hole. The trajectory of the observer is shown on the effective potential diagram in figure 3.3a and the causal horizon is illustrated schematically in figure 2.10b. Our objective is to approximate this causal horizon using a family of null geodesics called the *null generators*. Before performing such calculations, we outline a justification for our procedure, as well as a strategy for computing the null generators.

4.1.1 Null generators of the horizon

We first introduce some notation. Let $\gamma(\tau)$ be the timelike geodesic trajectory of our observer and let $\gamma(\mathbb{R})$ be the image of the real line under the mapping $\gamma(\tau)$. $\gamma(\mathbb{R})$ is the set of spacetime events that make up the world line of our observer. Let $J^-(\gamma(\mathbb{R}))$ and $I^-(\gamma(\mathbb{R}))$ be the causal and chronological pasts of $\gamma(\mathbb{R})$, respectively, and let $\mathscr{H}$ be the horizon, so that $\mathscr{H} \equiv \mathrm{Bd}(J^-(\gamma(\mathbb{R})))$.

Next we give a definition of the null generators, and justify their use in describing the horizon. Let us define the null generators of $\mathscr{H}$ to be those null geodesics which lie entirely within $\mathscr{H}$. To understand why null generators give a complete description of the horizon, we invoke theorem 8.1.6 of [32], and apply it to the closed set $\gamma(\mathbb{R})$. Notice that $\gamma(\mathbb{R})$ is a closed set since $\mathrm{Bd}(\gamma(\mathbb{R})) = \gamma(\mathbb{R})$. According to the theorem, every point in $S \equiv \mathrm{Bd}(I^-(\gamma(\mathbb{R}))) - \gamma(\mathbb{R})$ lies on at least one null geodesic which is contained entirely in S. Since the set S is precisely the horizon $\mathscr{H}$, this guarantees that the null generators give a complete description of $\mathscr{H}$. The fact that $S = \mathscr{H}$ follows from $\mathrm{Bd}(J^-(\gamma(\mathbb{R}))) = \mathrm{Bd}(I^-(\gamma(\mathbb{R})))$ and $\gamma(\mathbb{R}) \cap \mathrm{Bd}(I^-(\gamma(\mathbb{R}))) = \emptyset$, where the former is discussed at the bottom of page 191 of [32], and the latter follows from the fact that $\gamma(\mathbb{R}) \subset I^-(\gamma(\mathbb{R}))$, along with the fact that $I^-(\gamma(\mathbb{R}))$ is an open set (page 190 of [32]).

The third step is to develop a strategy for computing the null generators. This strategy will depend crucially on the late time behavior of the generators. The basic idea will be to use knowledge of the late time behavior of the generators as "initial conditions" for the null geodesic equations.

Before developing our strategy, it is useful to get a better understanding of the future behavior of the generators. We once again use theorem 8.1.6 of [32], this time using the second part of the theorem and applying it to the null geodesics associated with the closed set $\gamma(\mathbb{R})$. As discussed in the previous paragraph, these null geodesics are the null generators of the horizon. The theorem states that these generators are either future inextendible or have a future endpoint on $\gamma(\mathbb{R})$. Let us show that none of these generators have a future endpoint on $\gamma(\mathbb{R})$, and so by the theorem just cited, they are in fact all future inextendible. The proof is by contradiction. Let $\mathscr{C}(\lambda)$ be a generator of the horizon, and suppose it has a future endpoint $p \in \gamma(\mathbb{R})$. Since $\gamma(\mathbb{R}) \subset I^-(\gamma(\mathbb{R}))$, we have $p \in I^-(\gamma(\mathbb{R}))$, and therefore, $p \in \mathrm{Int}(J^-(\gamma(\mathbb{R})))$ (where we have used $I^-(S) = \mathrm{Int}(J^-(S))$; page 191 of [32]). Since $\mathrm{Int}(J^-(\gamma(\mathbb{R})))$ is an open set (page 190 of [32]) and $p \in \mathrm{Int}(J^-(\gamma(\mathbb{R})))$, there is a neighborhood $\mathscr{O}$ of the point p such that $\mathscr{O} \subset \mathrm{Int}(J^-(\gamma(\mathbb{R})))$. By the definition of future endpoint (page 193 of [32]), there must be a value λ_0 such that $\mathscr{C}(\lambda) \in \mathscr{O}$ for all $\lambda > \lambda_0$. We have thus shown that $\mathscr{C}(\lambda) \in \mathrm{Int}(J^-(\gamma(\mathbb{R})))$ for all $\lambda > \lambda_0$, which contradicts the fact that we must always have $\mathscr{C}(\lambda) \in \mathscr{H} = \mathrm{Bd}(J^-(\gamma(\mathbb{R})))$. That is, it contradicts the fact that a generator, by definition, is contained entirely on the horizon $\mathscr{H}$, as discussed in the paragraph above. We have thus established

that the generators of the horizon are future inextendible. By definition, this means that they do not a have a future endpoint (note that this is a stronger statement than claiming that they do not have a future endpoint in $\gamma(\mathbb{R})$; hence the power of the theorem). Formally, the absence of a future endpoint for a generator $\mathscr{C}$ means that for every point $p \in \mathscr{C}$, there exists a neighborhood $\mathscr{O}$ of p such that given any $\lambda_0 \in \mathbb{R}$, it is the case that $\gamma(\lambda) \notin \mathscr{O}$ for some $\lambda > \lambda_0$. Less formally, the absence of a future endpoint means that the generators either run into a singularity or continue indefinitely. By contrast, a future endpoint would mean that a null geodesic was stopped abruptly, and could in some sense be extended by continuing where it left off (these issues are discussed in detail on page 193 of [32])). The horizon generators which we are interested in here are inextendible due to the fact that they continue indefinitely towards future timelike/null infinity, as illustrated by the 45 degree yellow lines in the Penrose diagram (figure 2.10b).

In order to implement our strategy for computing the generators, we need a more precise undertanding of the future behavior of the generators. Having established that the family of null generators are future inextendible, we can use the Penrose diagram (figure 2.10b) to conclude that they will satisfy $r(\lambda = \infty) = \infty$, where λ is an affine parameter and $r(\lambda)$ is the Schwarzschild coordinate of such a curve. If we use the time coordinate T from the LP coordinates developed in chapter 3 as the parameter along the null generators, the condition $r(\lambda = \infty) = \infty$ becomes $r(T = \infty) = \infty$ instead (this can be seen by looking at the spacelike coordinate curves of the LP coordinate system, shown in figure 3.1b). We can gain further insight into the generators which make up the horizon by using our knowledge of SdS spacetime as $r \to \infty$. As discussed in section 2.6, SdS spacetime approaches deSitter spacetime in the limit that $r \to \infty$. We know the shape of the causal horizon for any geodesic observer in deSitter spacetime consists of a closed surface surrounding that observer. Note that by the shape of the causal horizon, here we mean the intersection of a spacelike hypersurface with the causal horizon $\mathscr{H}$. Let us denote the spacelike hypersurfaces $T =$ constant by Σ_T and the intersection with $\mathscr{H}$ by $S_T = \Sigma_T \cap \mathscr{H}$. Based on our knowledge of the shape of the causal horizon in deSitter spacetime, we expect S_T to approach a closed surface surrounding our observer as $T \to \infty$. Such a surface is shown in the last frame of figure 4.4 (it is a curve since one dimension is suppressed). It is simply the cosmological horizon surrounding the observer once the black hole and observer have drifted sufficiently far apart. The null generators which make up this

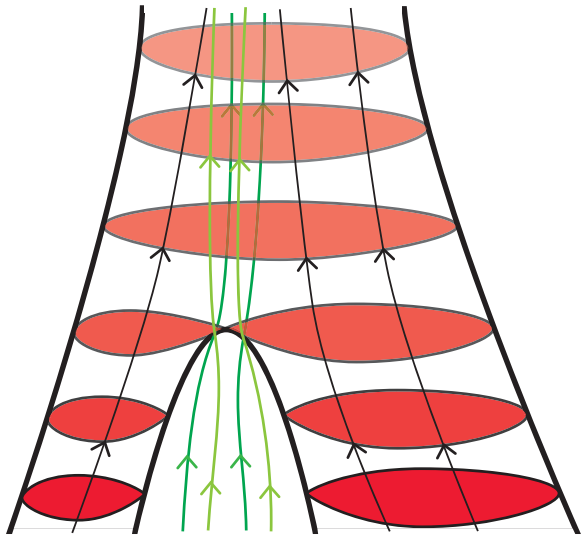


Figure 4.1: The "trouser" shaped horizon of two merging black holes. The green curves are horizon generators which enter through caustic points on the "inseam" of the trouser. Notice that these generators cross at the caustic points.

closed surface will satisfy $r(T=\infty)=\infty$ and $\phi(T=\infty)=0$. Furthermore, since the null generators which make up the horizon are all future inextendible, they will all reach this final closed surface as $T\to\infty$.

We now have a strategy for finding the null generators which make up the horizon: set up the null geodesic equations, and look for a family of solutions which form a closed surface as $T\to\infty$ and satisfy $r(T=\infty)=\infty$ and $\phi(T=\infty)=0$. Since the null geodesic equations cannot be solved analytically except in certain special circumstances, we use a combination of analytical and numerical methods. In section 4.4.3, analytical methods are used to find an approximate series solution for the desired family of geodesics in the limit that $T\to\infty$. This approximation is then used to set initial conditions for the numerical solution of the equations. The numerical method used is discussed in 4.3.3 and the numerical results are presented in 4.4.1.

Let us conclude this section by making some additional remarks about null generators and the structure of the causal horizon. As discussed, $\mathscr{H}$ is a null hypersurface generated by a family of future inextendible null geodesics. It thus enjoys the same properties as a traditional black hole

event horizon. In particular, we have the following (page 203 of [25]): (i) new generators can enter the horizon through special points on the horizon called caustic points, (ii) once a generator enters the horizon, it can never leave, and (iii) through each point on $\mathscr{H}$, there is either a unique generator going through that point, or it is a caustic point where new generators enter. A well known example of caustics in the context of black holes are those on the "inseam" of the "trouser" shaped horizon associated with the head-on merger of two non-rotating black holes, as shown in figure 4.1. Notice how generators cross at the caustic points. In the merger of black hole and cosmological horizons studied in this thesis, there is also a set of caustic points. These are illustrated in the diagram shown in figure 4.5. There is some resemblance with figure 4.1, the main difference being that at early times we have one horizon (the black hole) inside a larger horizon (the comosmological horizon). One of the focal points of this chapter will be the analysis of the mathematical structure of the caustic points associated with the merging of the black hole and cosmological horizons. Another focal point in the analysis of the horizons will be a precise determination of the merger point where the black hole and cosmological horizon first touch. The analysis of the merger point will be done in section 4.5.

4.1.2 The family of radial geodesic observers

So far we have been describing the causal horizon $\mathscr{H}$ for any observer drifting away from the black hole along a radial geodesic satisfying $r(\tau = -\infty) = r_e$ and $r(\tau) > r_e$. The observer's trajectory is shown on the effective potential diagram in figure 3.3 and on the Penrose diagram in figure 2.10b. This trajectory is part of a family of observers all drifting away from the black hole along radial goedesics (this family is plotted on the Penrose diagram in figure 4.2a). Let us prove that the family of causal horizons associated this family of observers are all equivalent, in the sense that any one can be obtained from any other by applying a symmetry transformation of the spacetime. Notice first that this family of geodesics precisely coincides with the timelike coordinate curves in the LP coordinate system which satisfy $r > r_e$ (these are the curves to the right of the green curve in figure 3.1a). As was shown in section 3.3.5, the flow associated with the Killing vector field $\xi^{(t)}$ maps this family of curves onto itself. It follows that this family of geodesics, and therefore also their

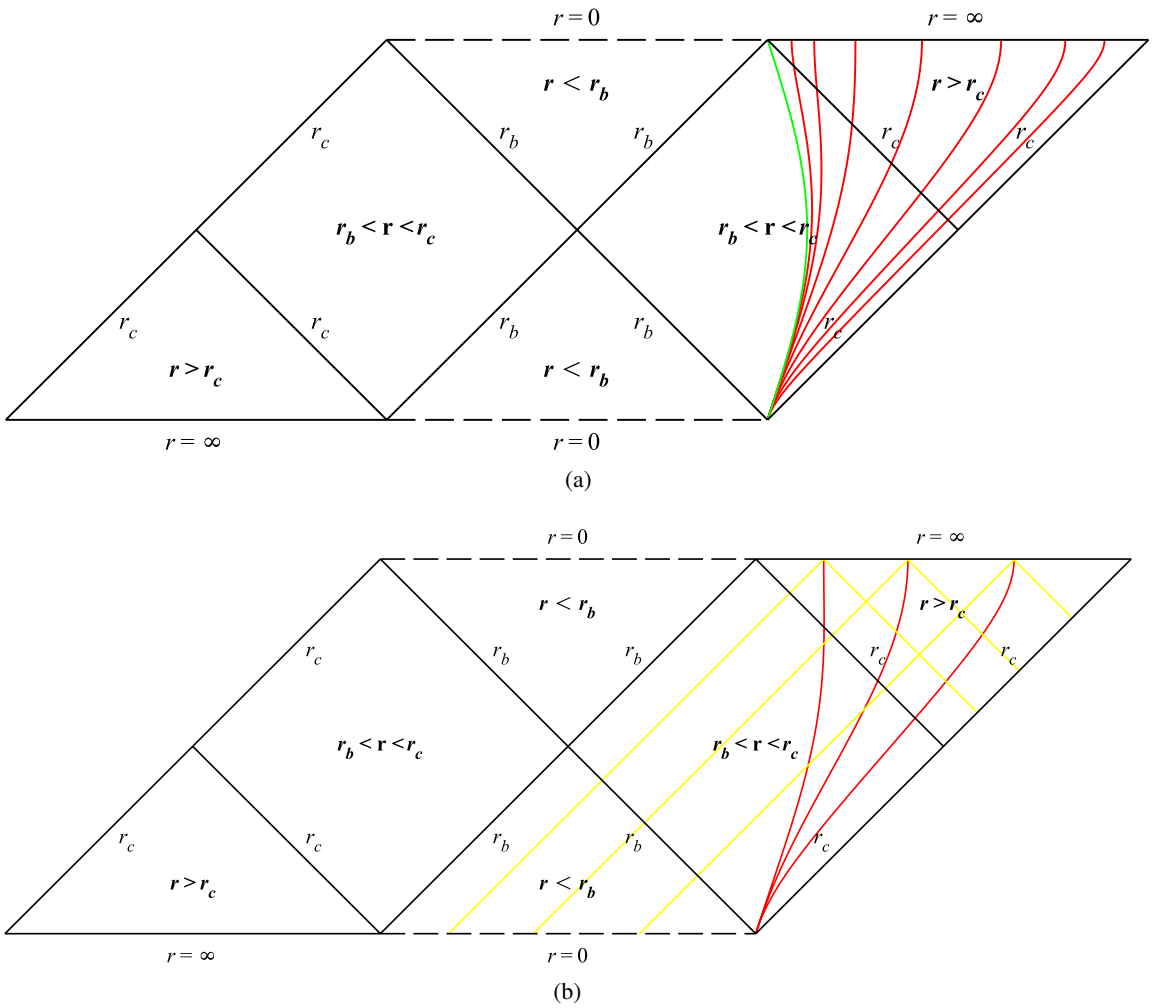


Figure 4.2: a) The family of radial geodesic observers drifting away from $r = r_e$. Notice that this set of trajectories was used in the construction of the timelike coordinate curves of the LP coordinate system (figure 3.1a). b) Illustration of the causal horizons (yellow) for three radial geodesic observers. The observer trajectories, as well as the causal horizons, are related by the symmetry transformation associated with the Killing vector $\xi^{(t)}$.

causal horizons, are related to each other by the flow associated with $\xi^{(t)}$, and thus are related by a symmetry of the spacetime. This family of horizons is illustrated in figure 4.2b.

4.1.3 Non-radial and non-geodesic observers

More generally, the observers moving along radial geodesic trajectories are part of an even larger family of observers moving along trajectories which satisfy the following two requirements:

(i) $r_b < r(\tau = -\infty) < r_c$,

(ii) the trajectory approaches one of the outward radial geodesics as $\tau \to \infty$ (as measured using the proper distance along the hypersurfaces $T =$ constant, for example).

Physically, conditions (i)-(ii) require the observer to start close to the black hole, then move along a possibly non-radial and/or accelerated trajectory, and finally end up caught up in the Hubble flow of spacetime. We will now sketch a proof that the causal horizons for the observers in the family of trajectories satisfying condition (i)-(ii) are all equivalent, in the sense of being related by a Killing symmetry. As was explained at the beginning of this section, the causal horizon $\mathscr{H}$ of any radial geodesic observer is calculated by propagating null geodesics backwards in time, starting from a set of light rays forming a closed surface surrounding the observer. Observers statisfying condition (ii) above have the same late time behavior as radial geodesic observers, and so their causal horizons at late times are also expected to have the same behavior: a closed surface surrounding the observer, corresponding to a deSitter horizon for a homogeneous expanding universe. This deSitter horizon is the cosmological horizon in the steady state universe. Since the late time behavior of these causal horizons is the same, and this late time behavior uniquely determines the causal horizon for all prior times, the causal horizon of the observers satisfying (i)-(ii) above are expected to match those of the radial geodesic observers. In the preceding, we have attempted to motivate the idea that it is the late time behavior of the observers' trajectories which determines the causal horizon, so that the

horizons calculated in this thesis apply to a broader class of observers than simply those which are moving on radial geodesic trajectories. We have not given a rigourous proof however.

An interesting further generalization of the above considerations would be to ask the following question: are the casual horizons of all observers reaching $r = \infty$ the same, in the sense of being related by a Killing symmetry transformation? This would include trajectories which are non-geodesic and non-radial for all times. If the casual horizon for such observers are indeed all the same, one could give a complete categorization of the possible casual horizons of observers in SdS spacetime (when considering the part of the spacetime covered by LP coordinates; see figure 3.1). Observers with $r_b < r(\tau) < r_c$ for all τ would have the familiar spherical Killing horizons $r = r_b$ and $r = r_c$ (as discussed in section 2.4), and observers with $r(\tau = \infty) = \infty$ would have the causal horizons discussed in this thesis. Note that all observers which do not fall into the black hole and reach the singularity at $r = 0$ must in fact be in one the two categories just mentioned. This can be deduced by looking at the Penrose diagram. The question of the causal horizon of observers moving along non-radial and non-geodesic trajectories for all times will not be explored in this thesis, however. Instead it will be addressed in a forthcoming publication.

4.1.4 Chapter organization

This chapter is organized as follows. In section 4.2, we set up the null geodesic equations which will be used to calculate the causal horizon $\mathscr{H}$ described above. This is accomplished by first exploiting the symmetries of SdS spacetime to find the null geodesic equations in Schwarzschild coordinates (section 4.2.1). These equations contain conserved quantitites which can be used to naturally incorporate the initial direction of propagation of a backwards light ray. The equations in Schwarzschild coordinates are then used to find the equations for the functions $r(T, \alpha)$ and $\phi(T, \alpha)$, where T is the time variable from LP coordinates, r is the radial Schwarzschild coordinate, ϕ is the azimuthal angle common to both of these coordinate systems (the angle θ will be set to $\pi/2$ and thus will be irrelevant), and α is an angle parametrizing the initial direction of propagation of a

light ray (section 4.2.2). As was discussed in section 3.3.7, when considering a single hypersurface $T =$ constant, the illustration and analysis of the horizon structure is particularly simple if instead of using LP coordinates (R, ϕ, θ), one instead uses (r, ϕ, θ) as coordinates on these three dimensional hypersurfaces.

In section 4.3, we outline the method used for solving the null geodesic equations. First, the equations for $r(T)$ and $\phi(T)$ are modified by compactifying both T and r (section 4.3.1). This allows for a more efficient numerical integration of the equations, and gives a method for dealing with the tricky issue of setting up initial conditions for light rays whose starting point is effectively future null/timelike infinity. The issue of initial conditions (or more precisely, asymptotic requirements as $T \to \infty$) is dealt with in section 4.3.2. The numerical method used is discussed in section 4.3.3.

In section 4.4, we analyze and illustrate the overall shape of the horizons. First, the results of the numerical integration are used to display the qualitative structure of the horizons by plotting their shape for various times (section 4.4.1). Analytical methods are then used to study in detail various aspects of the global horizon structure. We analyze the motion of lightlike geodesics using effective potentials, and categorize their possible behavior based on their initial conditions (section 4.4.2). This analysis corroborates our numerical results, and identifies key quantities whose calculation is useful for later analytical results. Next we consider the structure of the horizons at late times (section 4.4.3). This allows us to set up an explicit formula which can be used to set the initial conditions when numerically integrating the null geodesic equations (note that these initial conditions will in fact already have been used for numerical purposes in section 4.4.1, even though section 4.4.3 comes after section 4.4.1).

Finally, we focus on the analysis of the location of the merger point when the two horizons first merge (section 4.5). We give an explicit formula for the location of the merger point, in the limit that $\varepsilon \to 0$ (section 4.5).

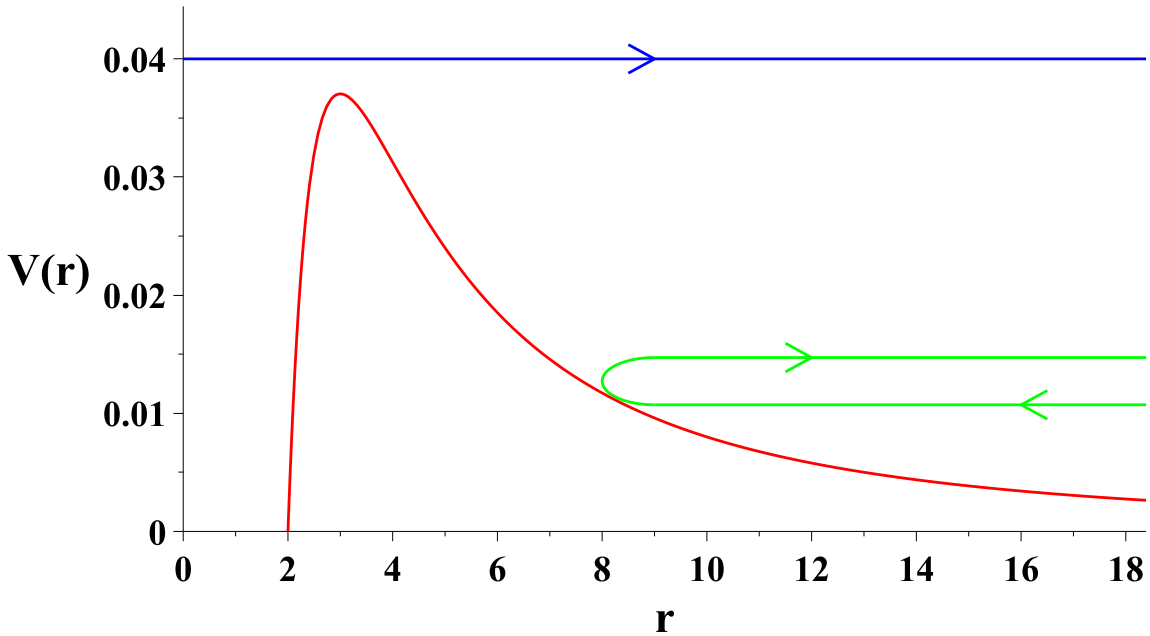


Figure 4.3: The effective potential diagram for null geodesics. The red curve is a plot of $V(r) \equiv \frac{V_{eff}(r)}{\sin^2\alpha} = \frac{1}{r^2}\left(1-\frac{2M}{r}\right)$. The green and blue curves are schematic illustrations of trajectories with and without turning points, respectively. The existence of turning points depends on the "initial conditions", as characterized by the value of the parameter α. The precise conditions are $\sin^2\alpha \leq 27H^2M^2$ (no turning points) and $\sin^2\alpha > 27H^2M^2$ (turning point). The parameter values used to produce the plot are $M = 1$, $27H^2M^2 = \frac{1}{4}$, with $\sin^2\alpha \approx 3/4$ and $\sin^2\alpha \approx 2/9$ for the green and blue trajectories, respectively.

4.2 Null geodesic equations

In the following section, we exploit the symmetries of SdS spacetime to find the equations for $t(\lambda)$, $r(\lambda)$ and $\phi(\lambda)$, where (t, r, ϕ) are the usual Schwarzschild coordinates and λ is an affine parameter along a backwards null geodesic. In section 4.2.2, we use the equations for $r(\lambda)$ and $t(\lambda)$ to find the equation for $T(\lambda)$, where T is the time variable in LP coordinates. This equation is then combined with the equations for $r(\lambda)$ and $\phi(\lambda)$ to yield the equations for $r(T)$ and $\phi(T)$.

4.2.1 Schwarzschild coordinates

As discussed in section 2.1, SdS spacetime possesses four Killing vectors. The vectors $\xi^{(\phi)}$, $\xi^{(1)}$, $\xi^{(2)}$ are associated with spherical symmetry and the vector $\xi^{(t)}$ is associated with the stationarity of the spacetime. The conservation law associated with $\xi^{(\phi)}$ can be used to deduce that there is a family of geodesics whose trajectories lie entirely on the hypersurface $\theta = \pi/2$. The symmetries associated with $\xi^{(1)}$ and $\xi^{(2)}$ are then used to conclude that all other geodesics can be obtained from this family by applying the transformations associated with the flow corresponding to $\xi^{(1)}$ and $\xi^{(2)}$. The upshot of this is that only the geodesics with $\theta = \pi/2$ need to be found, and without loss of generality we can set $\theta = \pi/2$ immediately in the Schwarzschild line element. Next, we use the conservation laws associated with the two Killing vectors

$$\begin{aligned} \xi^{(t)} &= (1,0,0,0) = \frac{\partial}{\partial t}, \\ \xi^{(\phi)} &= (0,0,1,0) = \frac{\partial}{\partial \phi} \ . \end{aligned}$$

Letting $(t(\lambda), r(\lambda), \phi(\lambda))$ be the coordinates of our null geodesic trajectory, the conservation laws associated with these Killing vectors are

$$\begin{aligned} \frac{dt}{d\lambda} f(r) &= e = \text{constant}, \\ \frac{d\phi}{d\lambda} r^2 &= l = \text{constant} \ . \end{aligned}$$

The requirement that the trajectory be null leads to

$$f(r)\left(\frac{dt}{d\lambda}\right)^2 - f(r)^{-1}\left(\frac{dr}{d\lambda}\right)^2 - r^2\left(\frac{d\phi}{d\lambda}\right)^2 = 0 \ .$$

Combining the above three equations, we have the equations for the Schwarzschild coordinates of a null geodesic trajectory:

$$\begin{aligned}
\frac{dt}{d\lambda} &= \frac{e}{f(r)}, \\
\frac{d\phi}{d\lambda} &= \frac{l}{r^2}, \\
\left(\frac{dr}{d\lambda}\right)^2 + \left(\frac{l}{r}\right)^2 f(r) &= e^2.
\end{aligned}$$

It will turn out to be convenient to eliminate the parameters l and e in favor of the single parameter $\alpha \in [0, 2\pi)$, defined through the relation

$$\tan\alpha = H\left(\frac{l}{e}\right) , \tag{4.1}$$

where $H = \sqrt{\frac{\Lambda}{3}}$ as before. The motivation for introducing the parameter α is that in the case $M = 0$ it can be interpreted as an angle parametrizing the intersection of the horizon with $\theta = \pi/2$ (here we use the term "horizon" to mean the 2-surface living on a spacelike hypersurface $T =$ constant, as opposed to the full three dimensional hypersurface). In the case $M \neq 0$, the interpretation of α is more subtle, but it nevertheless holds true that the values $\alpha \in [0, 2\pi)$ parametrize the curve formed by the horizon's intersection with $\theta = \pi/2$ (see figure 4.4 for examples of such curves). We will return to the interpretation of the parameter α in section 4.4.3.

Rescaling the affine parameter λ so that $\lambda \to \left(l^2 + \frac{e^2}{H^2}\right)\lambda$ and replacing l and e in favor of α, the geodesic equations become

$$\frac{dt}{d\lambda} = \frac{-H\cos\alpha}{f(r)}, \tag{4.2}$$

$$\frac{d\phi}{d\lambda} = \frac{-\sin\alpha}{r^2}, \tag{4.3}$$

$$\left(\frac{dr}{d\lambda}\right)^2 + V_{eff}(r) = H^2, \tag{4.4}$$

where

$$V_{eff}(r) = \frac{\sin^2\alpha}{r^2}\left(1 - \frac{2M}{r}\right) . \tag{4.5}$$

The equation for $r(\lambda)$ has been written in effective potential form. From the effective potential diagram (figure 4.3) we see that for some values of α, light rays reaching $r = \infty$ will have a turning point. We have two cases:

(I) If $\sin^2\alpha \leq 27H^2M^2$, then there are no turning points and $\frac{dr}{d\lambda} > 0$ for all λ.

(II) If $\sin^2\alpha > 27H^2M^2$, then there is a turning point at some $\lambda = \lambda^*$ and we have $\frac{dr}{d\lambda} < 0$ for $\lambda < \lambda^*$ and $\frac{dr}{d\lambda} > 0$ for $\lambda > \lambda^*$.

We can use knowledge of the two cases above to tranform equation (4.4) into a first order equation, keeping in the mind the presence of the turning point at $\lambda = \lambda^*$ in the second case. This gives

$$\sin^2\alpha \leq 27H^2M^2 \quad \Longrightarrow \quad \frac{dr}{d\lambda} = \sqrt{H^2 - V_{eff}(r)}\ , \tag{4.6}$$

$$\sin^2\alpha > 27H^2M^2 \quad \Longrightarrow \quad \frac{dr}{d\lambda} = \begin{cases} -\sqrt{H^2 - V_{eff}(r)} & \text{if}\,\lambda < \lambda^* \\ \sqrt{H^2 - V_{eff}(r)} & \text{if}\,\lambda > \lambda^* \end{cases} \tag{4.7}$$

Alternatively, equation (4.4) can be transformed into a set of coupled first order equations:

$$\frac{dr}{d\lambda} = w, \tag{4.8}$$

$$\frac{dw}{d\lambda} = -\frac{1}{2}V'_{eff}(r)\ . \tag{4.9}$$

The equations (4.2)-(4.3), combined with either (4.6)-(4.7) or (4.8)-(4.9) and appropriate initial conditions, form a complete set of equations that can be used to calculate the Schwarzschild coordinates of a null geodesic in SdS spacetime. In this thesis, we will use either of these sets of equations, depending on the circumstances. First, however, they have to be transformed into equations which give the Schwarzschild coordinates (r, ϕ) of a null geodesic as a function of LP coordinate time T. This will be done in the next section.

4.2.2 LP coordinates and spherical coordinates

We are interested in illustrating and analyzing the structure of the horizons. Although this can be done in the LP coordinates developed in chapter 3, it will turn out that it is easier to instead use the following two step process. First, we restrict our attention to a single spacelike hypersurface $T =$ constant. Second, on this hypersurface we use the coordinates (r, ϕ, θ), where r is the radial Schwarzschild coordinate. Notice that when restricting attention to a single hypersurface $T =$ constant, the coordinate r is always a spacelike coordinate, even though it is a timelike coordinate for $r < r_b$ and $r > r_c$ when viewed as part of the Schwarzschild coordinates (t, r, ϕ, θ). Using the coordinates (r, ϕ, θ) on the spacelike hypersurfaces $T =$ constant has two advantages. First, the line element instrinsic to these hypersurfaces takes the particularly simple form (3.39) in these coordinates. This greatly simplifies the analysis of the horizons. Second, as discussed in section 3.3.7, we can interpret the coordinates (r, ϕ, θ) as spherical coordinates on the hypersurfaces and use this interpretation to illustrate the shape of the horizons. Notice that by "horizons" here we mean the surfaces which are formed by the intersection of a spacelike hypersurface $T =$ constant with the full null hypersurface which is the boundary of the causal past of the observer. In other words, horizons are the snapshots in time of the full spacetime horizon.

As explained in section 3.3.7, the horizons can be visualized as being formed by the surface obtained when rotating the curves such as those in figure 4.4. Therefore it suffices to use (r, ϕ) as polar coordinates in the Euclidean plane to analyze and illustrate the horizons. Since we will be using the coordinates (r, ϕ) to analyze and illustrate the horizons on the hypersurfaces $T =$ constant, we want to find the equations for $r(T)$ and $\phi(T)$. Notice the slight abuse of notation since we have used notation of the form $r(\cdot)$ for both $r(\lambda)$ and $r(T)$, even though these are different functions. $r(T)$ and $r(\lambda)$ give the Schwarzschild coordinates of a null geodesic as a function of LP coordinate time T or affine parameter λ, respectively. If the 1-to-1 function $T(\lambda)$ for a null geodesic is known, either one of these functions can be used to find the other (it is shown below that $T(\lambda)$ is strictly increasing and therefore 1-to-1). The same abuse of notation applies to $\phi(\lambda)$ and $\phi(T)$. In cases where they may be ambiguity as to which of these functions is involved in an equation, it will be

explicitly mentioned.

The equations for $r(T)$ and $\phi(T)$ are obtained by combining either the first order equations (4.6)-(4.7), or the coupled first order equations (4.8)-(4.9) for $r(\lambda)$, and the equation (4.3) for $\phi(\lambda)$ with the equation for $T(\lambda)$. The procedure is slighly different in the two cases, and so we consider them separately. In either case, the equation for $T(\lambda)$ is obtained by using the transformation law

$$\frac{dT}{d\lambda} = \frac{\partial T}{\partial r}\frac{dr}{d\lambda} + \frac{\partial T}{\partial t}\frac{dt}{d\lambda} \ . \tag{4.10}$$

First order equations

Substituting (4.2) and (4.6)-(4.7) for $dt/d\lambda$ and $dr/d\lambda$ into (4.10), and using (3.18)-(3.19) to obtain the Jacobian coefficients $\partial T/\partial r$ and $\partial T/\partial t$, we get

$$\frac{dT}{d\lambda} = \frac{-H\, f(r_e)^{\frac{1}{2}} \cos\alpha - \mathrm{sgn}\left(\frac{dr}{d\lambda}\right)\left(f(r_e)-f(r)\right)^{1/2} \mathrm{sgn}(r-r_e)\sqrt{H^2 - V_{eff}(r)}}{f(r)} \ . \tag{4.11}$$

As will be shown in section 4.2.2 below, we always have $\frac{dT}{d\lambda} > 0$, so that $T(\lambda)$ is a strictly increasing function. Because of this, for every value of T we have a unique value of λ, and the first order equations for $r(T)$ and $\phi(T)$ are obtained by dividing the first order equations for $r(\lambda)$ and $\phi(\lambda)$ in (4.6)-(4.7) and (4.3) by $\frac{dT}{d\lambda}$ above. This yields the following equations for $r(T)$ and $\phi(T)$:

$$\frac{dr}{dT} = \frac{-f(r)\sqrt{H^2 - V_{eff}(r)}}{\mathrm{sgn}\left(\frac{dr}{dT}\right) H\, f(r_e)^{\frac{1}{2}} \cos\alpha + \left(f(r_e)-f(r)\right)^{\frac{1}{2}} \mathrm{sgn}(r-r_e)\sqrt{H^2 - V_{eff}(r)}} \ , \tag{4.12}$$

$$\frac{d\phi}{dT} = \left(\frac{\sin\alpha}{r^2}\right) \frac{f(r)}{H\, f(r_e)^{\frac{1}{2}} \cos\alpha + \mathrm{sgn}\left(\frac{dr}{dT}\right)\left(f(r_e)-f(r)\right)^{\frac{1}{2}} \mathrm{sgn}(r-r_e)\sqrt{H^2 - V_{eff}(r)}} \ . \tag{4.13}$$

It will be useful to rewrite the above equations more succinctly as

$$\frac{dr}{dT} = G_{\pm}(r), \tag{4.14}$$

$$\frac{d\phi}{dT} = K_{\pm}(r) \ . \tag{4.15}$$

with the plus and minus signs corresponding to the cases $\frac{dr}{dT} > 0$ and $\frac{dr}{dT} < 0$, respectively. The above equations are most useful when performing analytical computations of the horizons. For example, they will be used in analyzing the late time behavior of the horizons in section 4.4.3, or when finding the merger point location in section 4.5. When solving the null geodesic equations numerically, these equations break down near the turning points, and so instead we will use the coupled first order equations in (4.8)-(4.9). This system of equations are converted into equations for $r(T)$ and $\phi(T)$ instead of $r(\lambda)$ and $\phi(\lambda)$ in the next section.

Coupled first order equations

We once again substitute (4.2) for $dt/d\lambda$ and use (3.18)-(3.19) to obtain the Jacobian coefficients $\partial T/\partial r$ and $\partial T/\partial t$ in the equation for $dT/d\lambda$, given by (4.10) above. The only difference is that now we use $dr/d\lambda = w$, as given by the first of the two coupled equations (4.8)-(4.9) for $r(\lambda)$, and obtain

$$\frac{dT}{d\lambda} = \frac{-H f(r_e)^{\frac{1}{2}} \cos\alpha - w\left(f(r_e) - f(r)\right)^{1/2} \operatorname{sgn}(r - r_e)}{f(r)} \ .$$

As before, the equations for $r(T)$ and $\phi(T)$ are obtained by simply dividing the equations for $r(\lambda)$ and $\phi(\lambda)$ by the above expression for $dT/d\lambda$. This gives

$$\frac{dr}{dT} = \frac{-wf(r)}{H f(r_e)^{\frac{1}{2}} \cos\alpha + w\left(f(r_e) - f(r)\right)^{1/2} \operatorname{sgn}(r - r_e)}, \tag{4.16}$$

$$\frac{dw}{dT} = \frac{\frac{1}{2}V'_{eff}(r) f(r)}{H f(r_e)^{\frac{1}{2}} \cos\alpha + w\left(f(r_e) - f(r)\right)^{1/2} \operatorname{sgn}(r - r_e)}, \tag{4.17}$$

$$\frac{d\phi}{dT} = \left(\frac{\sin\alpha}{r^2}\right) \frac{f(r)}{H f(r_e)^{\frac{1}{2}} \cos\alpha + w\left(f(r_e) - f(r)\right)^{\frac{1}{2}} \operatorname{sgn}(r - r_e)} \ . \tag{4.18}$$

The above set of three coupled equations will be solved numerically to find the coordinates of the null geodesics, and these coordinates will be used to illustrate the shape of the horizons in section 4.4. The details of the numerical integration process will be discussed in section 4.3.3.

Proof that $\frac{dT}{d\lambda} > 0$

Let $N(r)$ be the numerator of the right hand side of (4.10):

$$N(r) = -H f(r_e)^{\frac{1}{2}} \cos\alpha - \mathrm{sgn}\left(\frac{dr}{d\lambda}\right)(f(r_e) - f(r))^{1/2}\,\mathrm{sgn}(r - r_e)\sqrt{H^2 - V_{eff}(r)}\ . \tag{4.19}$$

We will show that

$$r > r_c \Rightarrow N(r) < 0\ , \tag{4.20}$$

$$r_b < r < r_c \Rightarrow N(r) > 0\ . \tag{4.21}$$

Combining the above with (see figure 2.6 for a plot of $f(r)$)

$$r > r_c \Rightarrow f(r) < 0\ , \tag{4.22}$$

$$r_b < r < r_c \Rightarrow f(r) > 0\ , \tag{4.23}$$

in (4.6)-(4.7) and (4.3), it will follow that the right hand side of (4.10) is always positive, provided that $r > r_b$. As will be discussed in section 4.4.2, the null geodesics we are interested in satisfy $r(T) > r_b$ for all T. We can thus be assured that the right hand side of (4.10) is always positive, as claimed, provided that (4.20)-(4.21) hold, and we are restricting attention to the null geodesics considered in this thesis.

Let us first prove (4.20). Throughout this part of the proof, it is understood that we are taking $r > r_c$. First note that

$$\begin{aligned} \mathrm{sgn}\left(\frac{dr}{d\lambda}\right) &= 1, \\ \mathrm{sgn}(r - r_e) &= 1\ . \end{aligned}$$

The first equation follows from the fact that there are no turning points with $r > r_c$, which will be proven in section 4.4.2 below, while the second of these equations follows from the fact that $r_c > r_e$,

which can be seen in figure 2.6. Using the above, (4.19) becomes

$$N(r) = -H\, f(r_e)^{\frac{1}{2}} \cos\alpha - (f(r_e) - f(r))^{1/2} \sqrt{H^2 - V_{eff}(r)}\ . \tag{4.24}$$

In addition, (4.22) implies that the following inequalities hold:

$$\begin{aligned} (f(r_e) - f(r))^{1/2} &> f(r_e)^{\frac{1}{2}}, && (4.25)\\ \sqrt{H^2\cos^2\alpha - \frac{1}{r^2} f(r)\sin^2\alpha} &> H\mid\cos\alpha\mid\ . && (4.26) \end{aligned}$$

Now we use the definition of the effective potential in equation (4.5) to convert (4.26) into the following inequality:

$$\sqrt{H^2 - V_{eff}(r)} > H\mid\cos\alpha\mid\ . \tag{4.27}$$

Applying (4.25) and (4.27) to (4.24), we obtain

$$N(r) < -H\, f(r_e)^{\frac{1}{2}} (\cos\alpha + \mid\cos\alpha\mid)\ .$$

The right hand side above vanishes if $\cos\alpha \leq 0$ and is negative if $\cos > 0$, so that we have established that $N(r) < 0$ if $r > r_c$.

Next we prove (4.21). Throughout the proof, it is understood that we are taking $r_b < r < r_c$. As will be shown in section 4.4.2, only null geodesics with $\cos\alpha < 0$ ever cross $r = r_c$ in order to satisfy $r_b < r(T) < r_c$ for some value of T. We can therefore assume that $\cos\alpha < 0$.

First, we use (4.23) to obtain the inequalities

$$\begin{aligned} \sqrt{H^2\cos^2\alpha - \frac{1}{r^2} f(r)\sin^2\alpha} &< H\mid\cos\alpha\mid,\\ (f(r_e) - f(r))^{1/2} &< f(r_e)^{\frac{1}{2}}\ . && (4.28) \end{aligned}$$

Using the definition of the effective potential in equation (4.5) and the fact that $\cos\alpha < 0$, the first

inequality above becomes

$$\sqrt{H^2 - V_{eff}(r)} < -H\cos\alpha \ . \tag{4.29}$$

Multiplying the inequalities (4.28) and (4.29), we obtain

$$-H\, f(r_e)^{\frac{1}{2}} \cos\alpha - (f(r_e) - f(r))^{1/2} \sqrt{H^2 - V_{eff}(r)} > 0 \ .$$

The next step is to consider $\mathrm{sgn}\left(\frac{dr}{d\lambda}\right)\mathrm{sgn}(r - r_e)$. Suppose $\mathrm{sgn}\left(\frac{dr}{d\lambda}\right)\mathrm{sgn}(r - r_e) = -1$. Then since $\cos\alpha < 0$, both terms in $N(r)$ are positive and $N(r) > 0$ follows immediately. If $\mathrm{sgn}\left(\frac{dr}{d\lambda}\right)\mathrm{sgn}(r - r_e) = -1$, then $N(r) > 0$ follows from the above inequality. Thus we have shown that $N(r) > 0$ when $r_b < r < r_c$. Since we have now shown that both (4.20) and (4.21) hold for the null geodesics in this thesis, this completes the proof that $dT/d\lambda > 0$.

4.3 Null geodesic calculations

In what follows, we outline the procedure used to solve the null geodesic equations (4.12)-(4.13) numerically.

4.3.1 Compactification of the time variable

For the purposes of solving the null geodesic equations numerically, it is useful to compactify the time variable T. We will denote the compactified variable $\hat{T}$. Compactification serves two purposes. First, it is easier to implement the asymptotic requirements as $T \to \infty$ once the variable T is compactified. Second, a wise choice of compactification of the time variable effectively introduces a variable step size in time. This is useful since the horizons undergo the largest change in shape and area in a $\mathscr{O}(\frac{1}{H})$ time interval about the merger point, where $H = \sqrt{\Lambda/3}$ is the Hubble parameter.

With this in mind, we introduce the compactified time variable

$$\hat{T} = \tanh\left(\frac{HT}{2}\right) . \tag{4.30}$$

The infinte interval $T \in (-\infty, \infty)$ has been compactified to the finite interval $\hat{T} \in (-1, 1)$.

Consider either the equations (4.14)-(4.15) or the coupled equations (4.16)-(4.18) for $r(T)$ and $\phi(T)$. Let us convert these into equations for $r(\hat{T})$ and $\phi(\hat{T})$. Transforming (4.14)-(4.15) using the change of variables (4.30), we get the following equations for $r(\hat{T})$ and $\phi(\hat{T})$:

$$\frac{dr}{d\hat{T}} = \frac{2}{H(1-\hat{T}^2)} G_{\pm}(r), \tag{4.31}$$

$$\frac{d\phi}{d\hat{T}} = \frac{2}{H(1-\hat{T}^2)} K_{\pm}(r) . \tag{4.32}$$

Similarly, applying the same change of variables to the coupled equations (4.16)-(4.18), we obtain the following set of coupled equations for $r(\hat{T})$, $w(\hat{T})$ and $\phi(\hat{T})$:

$$\frac{dr}{d\hat{T}} = \left(\frac{2}{H(1-\hat{T}^2)}\right) \frac{-wf(r)}{Hf(r_e)^{\frac{1}{2}}\cos\alpha + w\left(f(r_e)-f(r)\right)^{1/2}\operatorname{sgn}(r-r_e)}, \tag{4.33}$$

$$\frac{dw}{d\hat{T}} = \left(\frac{2}{H(1-\hat{T}^2)}\right) \frac{\frac{1}{2}V'_{eff}(r)f(r)}{Hf(r_e)^{\frac{1}{2}}\cos\alpha + w\left(f(r_e)-f(r)\right)^{1/2}\operatorname{sgn}(r-r_e)}, \tag{4.34}$$

$$\frac{d\phi}{d\hat{T}} = \left(\frac{2}{H(1-\hat{T}^2)}\right)\left(\frac{\sin\alpha}{r^2}\right) \frac{f(r)}{Hf(r_e)^{\frac{1}{2}}\cos\alpha + w\left(f(r_e)-f(r)\right)^{\frac{1}{2}}\operatorname{sgn}(r-r_e)} . \tag{4.35}$$

Next we consider the initial conditions (or more precisely, limiting conditions as $\hat{T} \to 1^-$) that will accompany these equations.

4.3.2 Asymptotic requirements

We impose the following asymptotic requirements on $r(T)$ and $\phi(T)$ as $T \to \infty$:

$$\begin{aligned} r(T) &\to \infty, \\ \phi(T) &\to 0\,. \end{aligned} \tag{4.36}$$

The first requirement is a consequence of the fact that the null rays are the boundary of the causal past of an observer reaching future null/timelike infinity at $r = \infty$, as shown in figure 2.10b. The second requirement follows from the fact that the light rays remain a finite distance from our observer as $T \to \infty$, and the fact that the observer moves along a trajectory satisfying $\phi(\tau) = 0$ for all τ. These requirements were discussed in more detail in section 4.1, where they were shown to result from the expected late time behavior of the null geodesic generators of the causal horizon.

Expressed in terms of the compactified time variable (4.30), these conditions become

$$\begin{aligned} \hat{r}(\hat{T} = 1^-) &= \infty, \\ \phi(\hat{T} = 1^-) &= 0\,. \end{aligned} \tag{4.37}$$

The above can in some sense be thought of as providing initial conditions for the equations (4.31)-(4.32). The main difference between the above and proper initial conditions is that the above are actually limiting conditions on $r(\hat{T})$ and $\phi(\hat{T})$ as $\hat{T} \to 1^-$.

Although the conditions (4.37) are necessary conditions for any light ray trajectory which is part of the horizon, they are insufficient for setting up initial conditions when finding the trajectories numerically, since the right hand side of the equations (4.31)-(4.32) is ill defined at $\hat{T} = 1$. Furthermore, neither of these conditions capture the expected behavior of the horizon as $T \to \infty$, which is a simple closed curve parametrized by $\alpha \in [0, 2\pi)$, as shown in the last frame of figure 4.4, and discussed in section 4.1. Notice that the horizon in figure 4.4 is a simple closed curve instead of a closed surface since we have set $\theta = \pi/2$ and supressed one of the spatial dimensions.

To set up initial conditions for the equations (4.31)-(4.32), and in order to understand the behavior of the solutions of (4.12)-(4.13) as $T \to \infty$, we will use a series expansion to obtain an approximate solution for $T \to \infty$. How this series expansion is used to set initial conditions for numerical purposes is explained in the next section. The derivation of the series solution will be deferred until section 4.4.3, where we will perform a detailed analysis of the late time shape of the horizons.

4.3.3 Numerical methods

Initial conditions

The initial conditions to (4.31)-(4.32) are set by using approximations to $r(\hat{T})$ and $\phi(\hat{T})$ near $\hat{T} = 1$. When presenting these approximations, it is useful to first define $\xi > 0$ as

$$\xi = 1 - \hat{T} \ .$$

In section 4.4.3, we will obtain an approximate series expansion for $r(\xi)$ and $\phi(\xi)$ in the limit that $\xi \ll 1$. These expansions are of the form

$$r(\xi) = \frac{1}{\xi}\left(1 + r_1\xi + r_2\xi^2 + r_3\xi^3 + r_4\xi^4 + \mathscr{O}\left(\xi^5\right)\right) \ , \tag{4.38}$$

$$\phi(\xi) = \phi_1\xi + \phi_2\xi^2 + \phi_3\xi^3 + \phi_4\xi^4 + \mathscr{O}\left(\xi^5\right) \ . \tag{4.39}$$

where r_0, r_1, r_2 and r_3 are given by (4.76) and ϕ_1, ϕ_2, ϕ_3 and ϕ_4 are given by (4.80)-(4.83). The approximations to $r(\hat{T})$ and $\phi(\hat{T})$ near $\hat{T} = 1$ are then found by making the substitution $\xi = 1 - \hat{T}$ in the above.

Equations (4.31)-(4.32) are solved by integrating backwards from some initial value $\hat{T}_0 = 1 - \xi_0$ for some small but finite value of ξ_0, with the approximations to $r(\hat{T})$ and $\phi(\hat{T})$ above providing the initial conditions. The value of ξ_0 should be chosen small enough that the error introduced by

approximating the initial conditions is negligible, but not so small that numerical round-off errors due to the finite number of digits cause a loss of significance either in evaluating the right hand sides of (4.31)-(4.32), or in evaluating the expression for $r(\hat{T})$ and $\phi(\hat{T})$ that are used to set the initial conditions.

Let T_0 be the value of T corresponding to $\hat{T}_0$, as given by (4.30). A useful benchmark for determining a minimum size for the value of ξ_0 is that we should have $T_0 - T_{merger} \gg \frac{1}{H}$, where T_{merger} is the time at which horizon merger occurs. This benchmark is based on the expectation that the largest change in horizon area and shape takes place over an $\mathscr{O}(\frac{1}{H})$ time interval about T_{merger}, where $H = \sqrt{\Lambda/3}$ is Hubble parameter. Inverting (4.30) and using $\hat{T} = 1 - \xi$, we have

$$\begin{aligned} T &= \frac{1}{H}\ln\left(\frac{1+\hat{T}}{1-\hat{T}}\right) \\ &= \frac{1}{H}\ln\left(\frac{2-\xi}{\xi}\right) \\ &= \frac{1}{H}\left[-\ln\xi + \ln 2 + \mathscr{O}(\xi)\right] \\ &\approx \frac{2.3}{H}\left[-\log_{10}\xi + 0.3 + \mathscr{O}(\xi)\right] \ . \end{aligned}$$

For example, using the value $\xi_0 = 10^{-10}$ in the above gives $T_0 \approx 23\frac{1}{H} \gg \frac{1}{H}$. Throughout this thesis, we use the value $\xi_0 = 10^{-4}$, unless otherwise indicated.

Integration method

In order to find the coordinates (r, ϕ) as a function of LP coordinate time T for a null geodesic, one approach is to numerically integrate the null geodesic equations for $r(\hat{T})$ and $\phi(\hat{T})$, either in the uncoupled form (4.31)-(4.32) or the coupled form (4.33)-(4.35). Numerically integrating the uncoupled equations (4.31)-(4.32) is problematic at the turning points, since it would require an infinite number of time steps as one gets arbitrarily close to the turning point. To avoid this difficulty, we will use the the coupled equations (4.33)-(4.35) for trajectories with turning points, and the uncoupled equations (4.31)-(4.32) for null geodesic trajectories without turning points. The

advantage of the uncoupled equations is that there are two equations instead of three, allowing for a more efficient numerical integration. The uncoupled equations, either in the compactified form (4.31)-(4.32) or the raw form (4.14)-(4.15), will also be used in analytical calculations involving the null geodesics. For example, in section 4.4.3 they will be used to find the null geodesics at late times, and in section 4.5 they will be used to find the location of the merger point where the horizons first touch.

Equations (4.31)-(4.32) or (4.33)-(4.35) are solved subject to the initial conditions described in section 4.3.3 using MAPLE 14's numeric dsolve procedure, with the integration preceding backwards from the initial time $\hat{T}_0$. The MAPLE environment variable "Digits" is set to 15, instead of using the default value of 10, so that floating point calculations are done with 15 digits. This is equivalent to using double precision floating point calculations, and has the advantage of allowing the value of ξ_0 from section 4.3.3 to be, for example, $\xi_0 = 10^{-10}$ without creating a loss of significance due to rounding of floating point numbers. Note that the system of equations (4.31)-(4.32) or (4.33)-(4.35) are both considered "stiff" systems, in the sense that excessively small integration time steps are required to deal with the rapid variation in $r(\hat{T})$ near the initial time $\hat{T} = 1$. For this reason, we use the option stiff='true' when calling MAPLE's dsolve procedure.

For null geodesic trajectories with $\cos\alpha < 0$, both the numerator and denominator of the right hand side of either (4.31) or (4.17) vanish for $r = r_c$, and we encounter a potential difficulty during the numerical integration. How this potential difficulty is handled is discussed in detail in section 4.3.3 below.

The equations are solved for a set of values for the parameter α appearing in the equations. The definition of α is given by equation 4.1, and the interpretation of α is that the values $\alpha \in [0, 2\pi)$ parametrize the generators which make up the closed curve in the last frame of figure 4.4. This closed curve is a representation of the late time horizon, where one spatial dimension has been suppressed so that the horizon is a closed curve instead of a closed surface. If one considered the full two dimensional closed surface, it would be parametrized by two angles α and β. However, the rotational symmetry of the spacetime allows us to ignore the angle β. A more detailed interpretation

of the parameter α can be found in section 4.4.3.

The specific values of α used in a calculation depends on the problem at hand. For example, when considering the area of the horizon, we will want to use values of α covering the entire interval $[0, 2\pi)$. Notice that by reflection symmetry, we can always limit ourselves to values $\alpha \in [0, \pi)$ without any loss of generality.

The sphere $r = r_c$

For null geodesic trajectories with $\cos\alpha \leq 0$, the functions $G_+(r)$ and $K_+(r)$ in (4.31)-(4.32) are undefined when $r = r_c$, since both numerator and denominator of the right hand side of (4.12)-(4.13) vanish at this point. The same difficulty occurs in the right hand sides of the coupled equations (4.33)-(4.35). Fundamentally, this problematic behavior stems from the fact that one of the null geodesic equations in Schwarzschild coordinates (equation (4.2)) contains a singularity, and that although changing to LP coordinates eliminates this infamous coordinate singularity, it does so only by introducing a vanishing numerator to balance the vanishing denominator in the equations.

For trajectories with $\cos\alpha \leq 0$ and which cross $r = r_c$, this creates a problem when integrating the equations (4.31)-(4.32) or (4.33)-(4.35) numerically. Depending on the numerical algorithm and initial conditions used, we have found that it is possible in some cases to integrate the equations (4.31)-(4.32) through $r = r_c$ to obtain (by brute force) a numerical solution satisfying $r(T_c) = r_c$, where T_c is defined by $r(T_c) = r_c$. However, the accuracy of such a solution is questionable, since it is possible that the numerical algorithm would attempt to compute $G_+(r)$ or $K_+(r)$ with r sufficiently close to r_c that a loss of significance would occur. This would be the result of attempting to compute the ratio of two floating point numbers sufficiently close to zero. The same potential loss of accuracy arises when dealing with the coupled equations (4.33)-(4.35). The strategy for dealing with the two cases is slightly different, and so we discuss them separately.

Uncoupled equations In the case of the uncoupled equations (4.31)-(4.32), our strategy for circumventing this difficulty will be to approximate $G_+(r)$ and $K_+(r)$ in the vicinity of r_c by using the functions $g_+(r)$ and $k_+(r)$, defined by

$$g_+(r) = \begin{cases} \lim_{r\to r_c} G_+(r) & \text{for} \quad r_c - \Delta < r < r_c + \Delta\,, \\ G_+(r) & \text{for} \quad r \leq r_c - \Delta \quad \text{and} \quad r \geq r_c + \Delta\,, \end{cases} \tag{4.40}$$

$$k_+(r) = \begin{cases} \lim_{r\to r_c} K_+(r) & \text{for} \quad r_c - \Delta < r < r_c + \Delta\,, \\ K_+(r) & \text{for} \quad r \leq r_c - \Delta \quad \text{and} \quad r_c \geq r + \Delta\,. \end{cases} \tag{4.41}$$

The basic approximation used is that $G_+(r)$ and $K_+(r)$ are assumed to be constant for $r \in (r_c - \Delta, r_c + \Delta)$, and equal to their limiting values as $r \to r_c$. Notice that our approximation is equivalent to using a first order Taylor approximation to $r(T)$ for $r \in (r_c - \Delta, r_c + \Delta)$, with the expansion taking place about r_c and the derivative $r'(T)$ being approximated by $\lim_{r\to r_c} G_+(r)$. Equivalently, we are essentially taking the numerical method across $(r_c - \Delta, r_c + \Delta)$ to be one time step using the forward Euler method.

In the above, Δ is chosen to be small enough to limit the error introduced by this assumption, but not so small that rounding errors in the numerator and denominator in $G_+(r)$ cause a loss of significance (the very problem we are trying to avoid).

The limit in (4.40) can be found using L'Hopital's rule. First recall from (4.12) and (4.14) the definition of $G_+(r)$:

$$G_+(r) = \frac{-f(r)\sqrt{H^2 - V_{eff}(r)}}{H\, f(r_e)^{\frac{1}{2}} \cos\alpha + (f(r_e) - f(r))^{\frac{1}{2}} \operatorname{sgn}(r - r_e)\sqrt{H^2 - V_{eff}(r)}}\,.$$

Provided that $\cos\alpha < 0$, we have

$$\begin{aligned}
&\lim_{r\to r_c} G_+(r) \\
&= -\left(\lim_{r\to r_c}\sqrt{H^2 - V_{eff}(r)}\right)\left(\lim_{r\to r_c}\frac{f(r)}{H\,f(r_e)^{\frac{1}{2}}\cos\alpha + (f(r_e)-f(r))^{\frac{1}{2}}\sqrt{H^2-V_{eff}(r)}}\right) && (4.42)\\
&= \lim_{r\to r_c}\frac{H|\cos\alpha|f'(r)}{\frac{1}{2}\left(f(r_e)-f(r)\right)^{-\frac{1}{2}}f'(r)\sqrt{H^2-V_{eff}(r)} + (f(r_e)-f(r))^{\frac{1}{2}}\,\frac{1}{2}\left(H^2-V_{eff}(r)\right)^{-\frac{1}{2}}V'_{eff}(r)} && (4.43)\\
&= H\cos\alpha\frac{f'(r_c)}{\frac{1}{2}f'(r_e)^{-\frac{1}{2}}f'(r_c)H\cos\alpha + f(r_e)^{\frac{1}{2}}\,\frac{1}{2}\,\frac{1}{H\cos\alpha}V'_{eff}(r_c)}\ . && (4.44)
\end{aligned}$$

In going from (4.42) to (4.43) above, we have used $V_{eff}(r_c) = H\sin\alpha$ for the first limit and L'Hopital's rule for the second limit. In going from (4.43) to (4.44) we have used $\cos\alpha < 0$ and $f(r_c) = 0$ as well.

We can simplify (4.44) by noticing that

$$V_{eff}(r) = \frac{\sin^2\alpha}{r^2}\left(f(r) + H^2r^2\right)\ , \tag{4.45}$$

which follows from (4.5) and (2.5). Differentiating the above and using $f(r_c) = 0$, we obtain

$$V'_{eff}(r_c) = \frac{f'(r_c)}{r_c^2}\sin^2\alpha\ .$$

Substituting the above in (4.44) and simplifying, we get

$$\lim_{r\to r_c} G_+(r) = \frac{2H^2r_c^2 f(r_e)^{1/2}}{H^2r_c^2 + f(r_e)^{1/2}\tan^2\alpha}\ . \tag{4.46}$$

$f(r_e)$ is given in terms of M and H in (2.7), and r_c is the largest positive root of $f(r)$. The above allows us to calculate $\lim_{r\to r_c} G_+(r)$ for $\cos\alpha < 0$, given numerical values for M and H. For $\cos\alpha = 0$, we simply have $\lim_{r\to r_c} G_+(r) = G_+(r_c) = 0$. Notice that as $\cos\alpha \to 0$, the above expression goes to zero, so that $\lim_{r\to r_c} G_+(r)$ is continous as a function of α.

Before calculating $\lim_{r\to r_c} K_+(r)$, recall the definition of $K_+(r)$ from (4.13) and (4.15):

$$K_+(r) = \left(\frac{\sin\alpha}{r^2}\right) \frac{f(r)}{H f(r_e)^{\frac{1}{2}} \cos\alpha + \operatorname{sgn}\left(\frac{dr}{dT}\right)(f(r_e)-f(r))^{\frac{1}{2}} \operatorname{sgn}(r-r_e)\sqrt{H^2 - V_{eff}(r)}} \ . \quad (4.47)$$

In the case that $\cos\alpha < 0$, $\lim_{r\to r_c} K_+(r)$ can be calculated by noticing from the above that we have

$$\lim_{r\to r_c} K_+(r) = \frac{\sin\alpha}{r_c^2\sqrt{H^2 - V_{eff}(r_c)}} \lim_{r\to r_c} G_+(r) \ .$$

Using $V_{eff}(r_c) = H\sin\alpha$, $\cos\alpha < 0$ and (4.46), the above simplifies to

$$\lim_{r\to r_c} K_+(r) = \frac{-2H^2 f(r_e)^{1/2}\tan\alpha}{H^2 r_c^2 + f(r_e)^{1/2}\tan^2\alpha} \ . \quad (4.48)$$

In the case that $\cos\alpha = 0$, (4.47) simplifies to

$$K_+(r) = \left(\frac{\operatorname{sgn}(\sin\alpha)}{r}\right) \frac{f(r)}{(f(r_e)-f(r))^{\frac{1}{2}} \operatorname{sgn}(r-r_e)\sqrt{f(r)}} \ ,$$

where we have used $|\sin\alpha| = 1$ and (4.45). Although the above is undefined at $r = r_c$, a trivial simplification reveals that

$$\lim_{r\to r_c} K_+(r) = 0 \ .$$

Let us summarize our strategy for dealing with the possible loss of accuracy when numerically integrating the uncoupled equations (4.31)-(4.32) across $r = r_c$. For null geodesics with $\cos\alpha \le 0$ and $r'(T) > 0$ at $r = r_c$, the uncoupled equations (4.31)-(4.32) are replaced with the following equations instead:

$$\frac{dr}{d\hat{T}} = \frac{2}{H(1-\hat{T}^2)} g_+(r) \ , \quad (4.49)$$

$$\frac{d\phi}{d\hat{T}} = \frac{2}{H(1-\hat{T}^2)} k_+(r) \ , \quad (4.50)$$

where the functions $g_+(r)$ and $k_+(r)$ are given by (4.40)-(4.41), and $\lim_{r\to r_c} G_+(r)$ and $\lim_{r\to r_c} K_+(r)$ are given by (4.46) and (4.48) for $\cos\alpha < 0$, and by $\lim_{r\to r_c} G_+(r) = \lim_{r\to r_c} K_+(r) = 0$ for $\cos\alpha =$

0.

The coupled equations The difficulty that arises when numerically integrating the uncoupled equations (4.31)-(4.32) also occurs in the coupled equations (4.33)-(4.34). To see why, consider null geodesics with $r'(T) < 0$ when $r(T) = r_c$. By combining equations (4.6)-(4.7) with (4.8), we get that, as long as $r'(T) > 0$, we must have

$$w = \sqrt{H^2 - V_{eff}(r)}\ .$$

Now substituting the above into (4.34)-(4.35), these equations become

$$\begin{aligned}
\frac{dr}{d\hat{T}} &= \left(\frac{2}{H(1-\hat{T}^2)}\right) G_+(r), \\
\frac{dw}{d\hat{T}} &= -\left(\frac{1}{H(1-\hat{T}^2)}\right) \frac{V'_{eff}(r)}{\sqrt{H^2 - V_{eff}(r)}} G_+(r), \\
\frac{d\phi}{d\hat{T}} &= -\left(\frac{2}{H(1-\hat{T}^2)}\right) \left(\frac{\sin\alpha}{r^2\sqrt{H^2 - V_{eff}(r)}}\right) G_+(r)\ .
\end{aligned}$$

We see that the same difficulty in evaluating $G_+(r)$ that arose in the uncoupled equations also arises in the above. The strategy for dealing with the coupled equations (4.33)-(4.34) is to replace them with a modified version of the above equation when $w < 0$ and $r \in (r_c - \Delta, r_c + \Delta)$, where the modification that we make is to replace $G_+(r)$ with $\lim_{r\to r_c} G_+(r)$ and set $r = r_c$ in the remaining terms. This is basically the same strategy that was used in (4.49)-(4.50) above, where we replace the right hand sides of the equations with a constant for $r \in (r_c - \Delta, r_c + \Delta)$. The main differences in the case of the coupled equations are that we are dealing with three equations, and the right hand sides of these equations are functions of both r and w.

4.4 Horizon shapes

4.4.1 Numerical results

We begin our illustration and analysis of the shape and structure of the horizon by using numerical results to plot the coordinates $(r\cos\phi, r\sin\phi)$ at several instants of time T, as shown in figure 4.4. The functions $r(T)$ and $\phi(T)$ are found by solving the equations (4.31)-(4.32) or (4.33)-(4.35) numerically using the procedure outlined in section 4.3. As discussed in sections 3.3.7 and 4.2.2, the horizon can be thought of as a series of 2-surfaces, each embedded in one of the hypersurfaces $T =$ constant, where T is the time coordinate from LP coordinates. These 2-surfaces can be visualized by rotating the curves in 4.4 about the x-axis, where $x = r\cos\phi$. Representing the horizon as a 2-surface in Euclidean space, or a curve in the Euclidean plane, has the advantage that it easily allows us to identify key features of the horizon shape and structure. Some of these key features will be discussed further in this section and section 4.4.2. The disadvantage of this representation is that it introduces metrical distortions since the true geometry of the hypersurfaces $T =$ constant is that of a 3-cone, which is a curved manifold (see section 3.3.6 for a discussion of the geometry of the spacelike hypersurfaces).

Consider the progression of frames in figure 4.4. Since integration runs backwards in time, consider first the last frame. In this frame, the horizon differs only slightly from the "final shape". As was explained in section 4.1, the horizon (viewed as a 2-surface) is expected to converge to a "final shape" which is a closed surface surrounding the observer. In section 4.4.3 below, we will derive an explicit formula for the horizon at late times using a series expansion. This will confirm our expectation of the existence of such a final shape, and also allow us to precisely characterize the slight deviations from the final shape that occur at late times before the horizon finally settles down. Note that this final shape is not a circle (or not a sphere, if one considers the surface formed by rotating the curve). Instead it is prolate in the vertical direction, as can be deduced by looking at the discrepancy between the horizontal and vertical range of the figure. We will return to a precise characterization of this shape in section 4.4.3. The non-sphericity of the late time horizon shape may

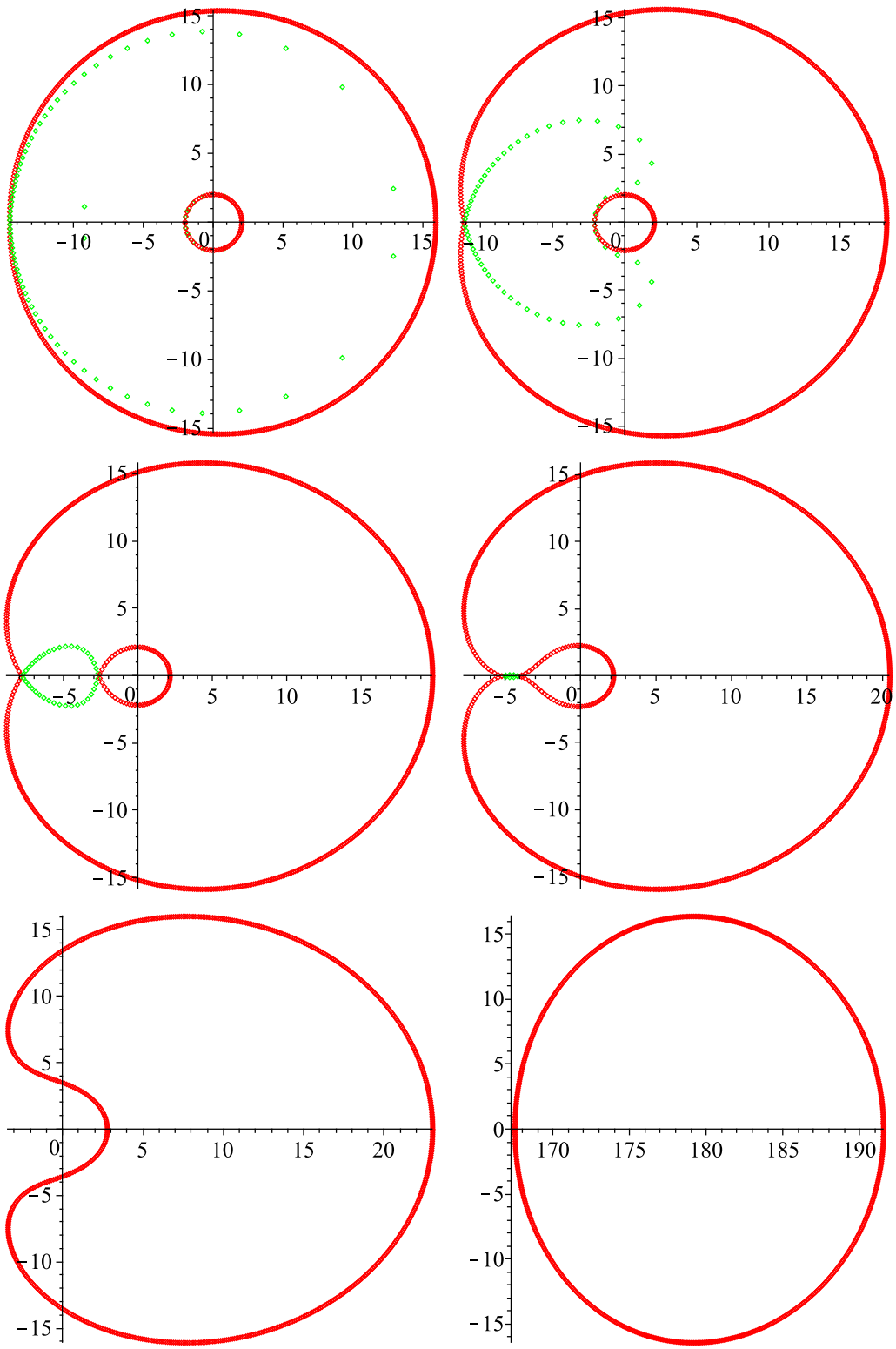


Figure 4.4: Six frames showing the progression from two distinct horizons at early times to a single cosmological horizon at late times. Parameter values: $M = 1$ and $\Lambda = 1/90$. Times for the six frames: $\hat{T} = 0$, $\hat{T} \approx 0.6$, $\hat{T} \approx 0.71$, $\hat{T} \approx 0.74$, $\hat{T} \approx 0.81$, $\hat{T} \approx 0.99$

be surprising, given that the spacetime geometry at late times approaches that of deSitter spacetime, as discussed in section 2.6, and given that the horizon surrounding an observer in deSitter spacetime is spherical. The reason for the discrepancy is that the spacelike hypersurfaces $T =$ constant of our LP coordinate system do not coincide precisely with the usual planar constant time surfaces of the steady-state universe of deSitter spacetime. So although the full three dimensional horizon, when viewed as a null hypersurface embedded in spacetime, approaches at late times the same null hypesurface as a horizon in deSitter spacetime, it has been sliced slightly differently in the LP coordinate system developed in this thesis, as compared to the usual planar deSitter steady-state universe slicing. In the limit that $M = 0$, the horizon shape in LP coordinates is spherical, as expected from the fact that LP coordinates reduce to the planar coordinates of deSitter spacetime in this limit, as was discussed in section 3.3.1.

In the fifth frame, we begin to see a significant distortion in the horizon shape due to the influence of the black hole. In order to easily illustrate the change in shape that occurs due to the black hole, we have used an unrealistically large value for the parameter characterizing the size of the black hole relative to the cosmological horizon. A realistic value based on the Λ-CDM model would be on the order of $\varepsilon = 10^{-14}$ (see section 2.7). Also in the fifth frame, we see that the horizons have moved significantly closer to $r = 0$. This approach towards $r = 0$ occurs roughly as $r \propto e^{HT}$ for large values of T (r is decreasing since we are thinking of progressively smaller values of T). This is related to the fact that moving forward in time, an observer caught up in the expansion of spacetime drifts away with approximately $r \propto e^{HT}$ at late times, with the cosmological horizon roughly centered on the observer's position.

In the fourth frame, we are just before (i.e. at an earlier time) the critical point where the horizons merge. Here we can see that we are dealing with two closely spaced but nevertheless disconnected horizons. With a clear separation into two horizons, we can define the black hole horizon as the inner horizon, and the cosmological horizon as the outer horizon. On the other hand, in the fifth and sixth frames, the horizon would be considered to be solely a cosmological horizon, although possibly with significant distortions due to the black hole. At some time between the fourth and fifth frames, there is a critical moment in time where the horizons first touch. We call this time the

merger of the horizons, and the precise location of this event in both space and time will be called the merger point. Determining the coordinates of the merger point will be the focus of section 4.5. As explained in section 4.1, the merger point is one of several caustic points where new generators can enter the horizon. The other caustic points are the points on the horizon which intersect the line $\phi = \pi$ (or the line $\phi = \pi$, $\theta = \pi/2$ if one considers the horizons as 2-surfaces and not curves). Although it is difficult to see in figure 4.4, the horizon in the first four frames is not smooth along these points. These nonsmooth points on the horizon converge until they finally join together at the merger point. Prior to their merger, the green points in between the horizons are the locations of generators which are not yet part of the horizons, but will later enter the caustic points and become horizon generators. These are similar to the green curves entering the horizon through the caustic points along the "inseam" of the "trousers" diagram shown in figure 4.1. In all six frames, the red points are the locations of the horizon generators.

In the third frame, we see that the horizons are further apart, and there are more (green) generators which have yet to join the horizon. In the second frame, there are even more green generators, and we see that these are moving towards the line $\phi = 0$ (when going backwards in time). Going backwards in time indefinitely, these green generators will rotate indefinitely, with three possible behaviors: either spiralling inwards into the black hole at $r = r_b$, spiralling outwards into the cosmological horizon $r = r_c$, or spiralling into the light sphere at $r = 3M$. None of these generators ever join the horizons, although all except those approaching $r = 3M$ will get arbitrarily close to the horizons as we go backwards in time. The generator (or family of generators if we consider the horizon as a 2-surface and not a curve) that approaches $r = 3M$ is the limiting case that separates the generators which approach either horizon, and this critical separator only occurs for one critical value of the parameter $\alpha \in [0, \pi)$, or two values of the parameter if $\alpha \in [0, 2\pi)$ (recall that $\alpha \in [0, 2\pi)$ parametrizes the generators of the sixth frame). These different qualitative behaviors for both the green and red generators will be analyzed in more detail in section 4.4.2, culminating in a classification of different generators according to different qualitative behavior. For example, one important distinction is between those generators which at the very earliest times started on the horizons, and those which at the very earliest times start not on the horizon, only to later join the horizon. These are similar to the red and green points, respectively, of the first frame, except that

one must imagine the limit as $T \to -\infty$. That is, one must imagine a very early frame. This distinction between these two types of generators will be especially relevant in section 5.3.2, where we will calculate the horizon area increase in time due to these two types of generators, and compare them to the total horizon area increase.

In the first frame, the horizon is very nearly that of a stationary observer: two nested spheres at $r = r_b$ and $r = r_c$. This agrees with our expectations, since the observer drifting away from the black hole starts at the equilibrium point $r = r_e$, and thus is close to the black hole for an arbitarily long time. During this time the observer can receive signals from anywhere in the region $r_b < r < r_c$, much like an observer with $r_b < r(\tau) < r_c$ for all τ (see section 2.4 for a further discussion of the horizons of such observers).

An alternative way of depicting the merging of the black hole and cosmological horizons is to use a three dimensional plot, as we have done in figure 4.5. The vertical direction is time and the horizontal directions are spatial, and as with the 2d plots, one spatial dimension has been suppressed. In this plot we can see the transition from two separate horizons into one horizon at later times, as well as the non-smoothness of the horizons along the caustic points prior to merger. As with the "trousers" diagram of binary black hole merger (figure 4.1), these caustic points occur along an "inseam", although the diagram does not resemble a pair of pants (if it was a pair of pants, one pant leg would be inside the other). Note that unlike our 2d plot, on our 3d figure we have only plotted the horizons, and not plotted the generators which eventually join through the caustic points along the inseam.

The analysis of the light rays which make up the horizon is carried out in more detail in the next section. This analysis will serve two purposes. First, we will focus on certain key generators which are crucial for further analysis that will be performed in the following sections. Second, a qualitative analysis of the behavior of the generators will allow us to gain further insight into the structure of the horizon, and confirm some of the numerical results presented above.

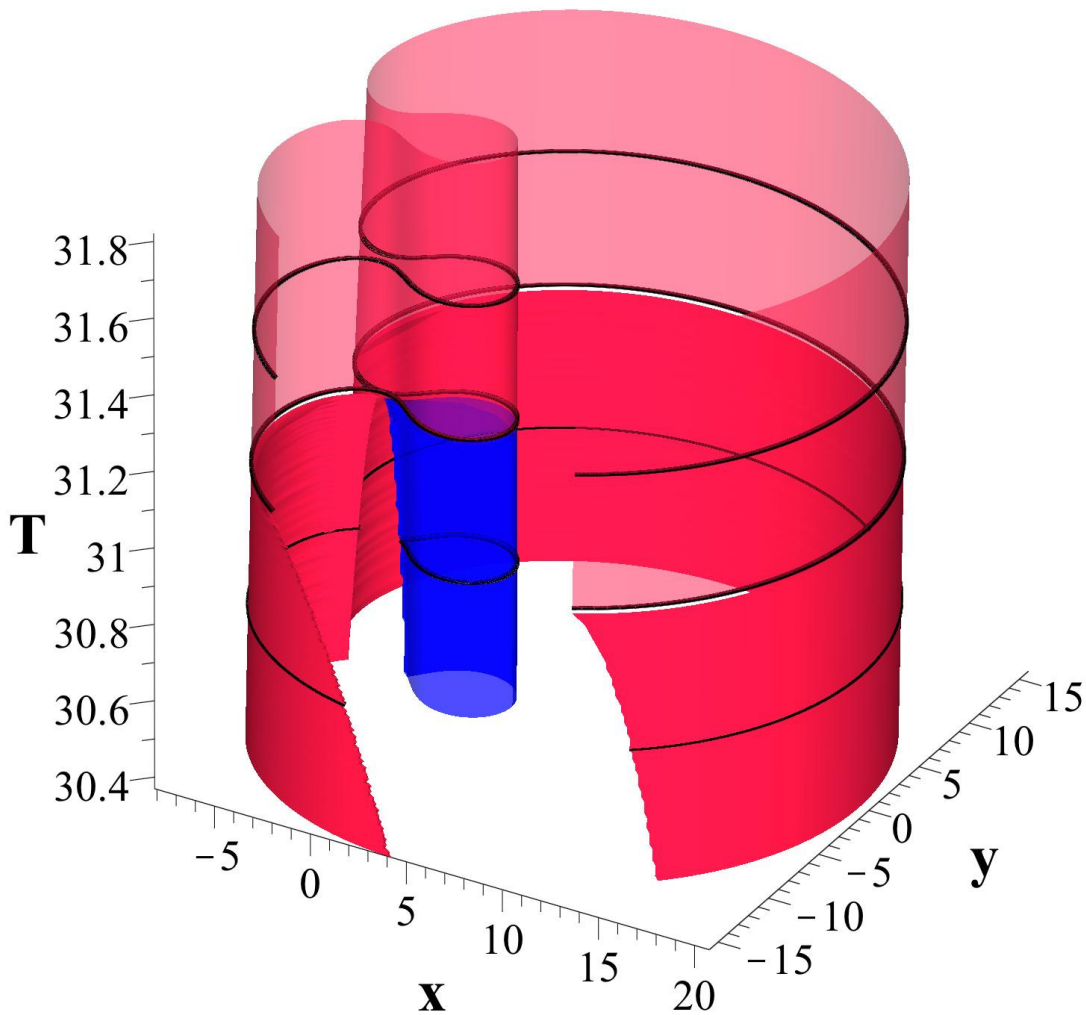


Figure 4.5: 3d plot of the merging of black hole and cosmological horizons. The vertical axis is time in LP coordinates and the horizontal axes are the coordinates $x = r\cos\phi$ and $y = r\sin\phi$. The blue and crimson surfaces represent the black hole and cosmological horizons respectively. We clearly see that the crimson and blue surfaces are not smooth prior to merger.

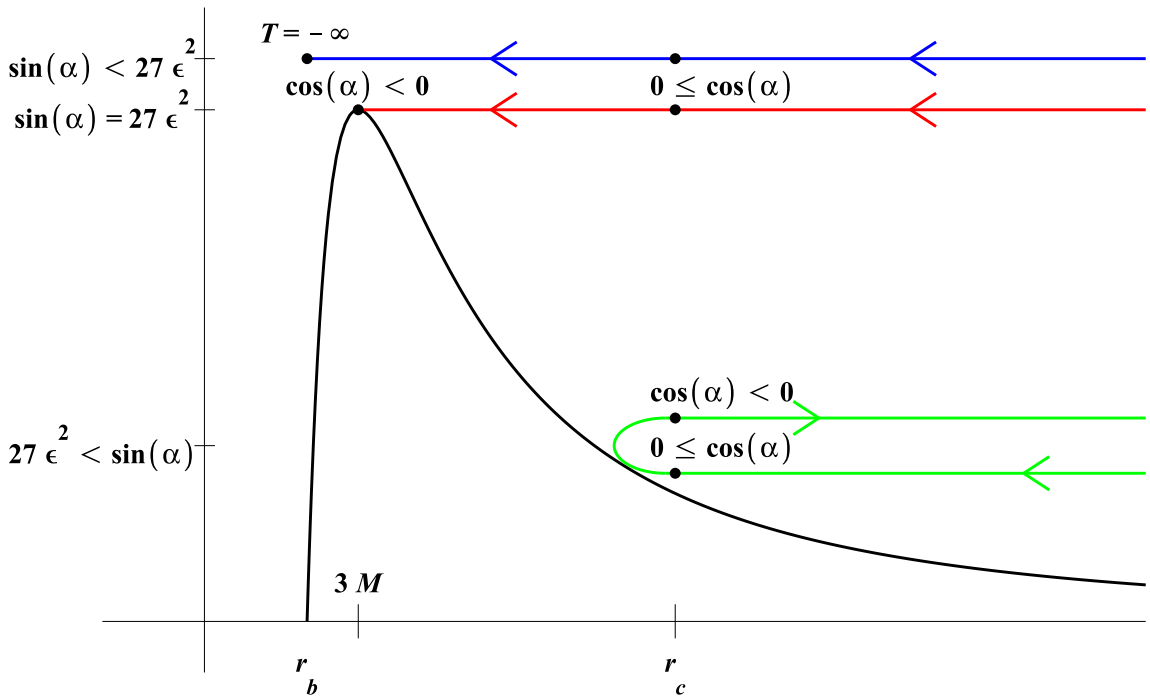


Figure 4.6: Schematic illustration of the possible behaviors for $r(T)$ as $T \to -\infty$. The solid black circles are points where $T = -\infty$. The direction of the arrows are thought of as indicating the direction of decreasing T, or equivalently, of decreasing λ. The solid black curve is the effective potential $\frac{V(r)}{\sin^2\alpha} = \frac{1}{r^2}(1 - \frac{2M}{r})$, and the three coloured curves indicate the different cases for the trajectories, based on the value of $\sin\alpha$. Consider first the case where $\sin\alpha > 27\varepsilon^2$, as shown by the green curve. If $\cos\alpha \geq 0$, then we have $T = -\infty$ at r_c, without a turning point in the behavior of $r(T)$. On the other hand, if $\cos(\alpha) < 0$, then we have $T = -\infty$ at r_c also, but this time with a turning point in the behavior of $r(T)$. For the cases where $\sin^2\alpha = 27\varepsilon^2$ and $\sin^2\alpha < 27\varepsilon^2$, if $\cos\alpha \geq 0$ then $T = -\infty$ at r_c. If $\cos\alpha < 0$, then we have $r(T = -\infty) = 3M$ and $r(T = -\infty) = r_b$ for the cases $\sin^2\alpha = 27\varepsilon^2$ and $\sin^2\alpha < 27\varepsilon^2$, respectively.

4.4.2 Classification of null generators

As discussed in section 4.1, the horizon can be thought of as a null surface generated by a family of light rays, known as the null generators. We can gain further insight into the overall structure of the horizon by classifying these generators into different categories, based on the behavior of the generators at early times. This classification will be done by using the parameter α, which parametrizes the generators such that each generator has precisely one value of α. The interpretation of α was briefly discussed in section 4.2.1, and will be revisited in more detail in section 4.4.3. Even

though the coordinates of the generators depend on the parameter α as well as the time T, so that we could write $r = r(T, \alpha)$, we will suppress the dependence on α and simply write $r = r(T)$.

The first classification of generators is based on the limiting behavior of the radial Schwarzschild coordinate $r(T)$ as $T \to -\infty$. As can be seen in the first frame of figure 4.4, generators appear to start from either the sphere $r = r_b$ or the sphere $r = r_c$. We will confirm this by combining the effective potential diagram (figure 4.3) with an understanding of the behavior of $T(\lambda)$, as given by (4.10). Note that in everything that follows we are considering null geodesics satisfying $r(\lambda = \infty) = \infty$. These can be visualized as moving to the right with increasing λ for large values of r in figure 4.3.

Since we are interested in the behavior of $r(T)$ as $T \to -\infty$, let us start by considering equation (4.10) and looking for finite values λ_0 such that $T(\lambda_0^+) = -\infty$. To identify such values, notice that

$$\left.\frac{dT}{d\lambda}\right|_{\lambda_0^+} = \infty \Longleftrightarrow T(\lambda_0^+) = -\infty \ ,$$

which follows from the fact that $T'(\lambda) > 0$, as was shown in section 4.2.2. Thus it suffices to find values of λ_0 where the right hand side of (4.11) diverges. This requires that the numerator is nonvanishing and the denominator vanishes. The vanishing of the denominator leads to

$$f(r(\lambda_0)) = 0 \quad \Longrightarrow \quad r(\lambda_0) = r_b \quad \text{or} \quad r(\lambda_0) = r_c \ , \tag{4.51}$$

which can be seen from figure 2.6. The condition that the numerator be nonvanishing leads to the following requirement:

$$\operatorname{sgn}\left((r - r_e) \left.\frac{dr}{d\lambda}\right|_{\lambda_0} \cos\alpha \right) \geq 0 \ . \tag{4.52}$$

Putting together the previous two requirements, we conclude that

$$\begin{aligned} r(\lambda_0) \quad &= \quad r_b \quad \text{and} \quad \operatorname{sgn}\left(\left.\frac{dr}{d\lambda}\right|_{\lambda_0} \cos\alpha \right) \leq 0 \\ \Longrightarrow \quad & \quad T(\lambda_0^+) = -\infty \\ \Longrightarrow \quad & \quad r(T = -\infty) = r_b \ , \end{aligned} \tag{4.53}$$

and

$$\begin{aligned} r(\lambda_0) &= r_c \quad \text{and} \quad \operatorname{sgn}\left(\left.\frac{dr}{d\lambda}\right|_{\lambda_0} \cos\alpha \right) \geq 0 \\ \Longrightarrow \quad & T(\lambda_0^+) = -\infty \\ \Longrightarrow \quad & r(T=-\infty) = r_c \ , \end{aligned} \tag{4.54}$$

where we have used the fact that $\operatorname{sgn}(r_b - r_e) = -1$ and $\operatorname{sgn}(r_c - r_e) = 1$, respectively, as can be seen from figure 2.6. It will turn out that there are only two generators (i.e. two values of α) that do not satisfy either of the above conditions, so that $r(T=-\infty) = r_b$ or $r(T=-\infty) = r_c$ for all except two null generators. To classify the values of α according to either $r(T=-\infty) = r_c$ or $r(T=-\infty) = r_b$, let λ_1 and λ_2 be defined such that $r(\lambda_1) \in \{r_b, r_c\}$ and $r(\lambda_2) \in \{r_b, r_c\}$, with $\lambda_1 < \lambda_2$. From the effective potential diagram 4.3, we see that for any trajectory with $r(\lambda = \infty) = \infty$, there are three possibilities for $r(\lambda_1)$ and $r(\lambda_2)$. These are

$$\sin^2\alpha < 27H^2M^2 \quad \Longrightarrow \quad r(\lambda_1) = r_b, \quad r(\lambda_2) = r_c \ , \tag{4.55}$$

$$\sin^2\alpha > 27H^2M^2 \quad \Longrightarrow \quad r(\lambda_1) = r_c, \quad r(\lambda_2) = r_c \ , \tag{4.56}$$

$$\sin^2\alpha = 27H^2M^2 \quad \Longrightarrow \quad \lambda_1 \text{ does not exist}, \quad r(\lambda_2) = r_c \ . \tag{4.57}$$

Furthermore, the signs of $r'(\lambda_1)$ and $r'(\lambda_2)$ are as follows:

$$\sin^2\alpha < 27H^2M^2 \quad \Longrightarrow \quad r'(\lambda_1) > 0 \quad \text{and} \quad r'(\lambda_2) > 0 \ , \tag{4.58}$$

$$\sin^2\alpha > 27H^2M^2 \quad \Longrightarrow \quad r'(\lambda_1) < 0 \quad \text{and} \quad r'(\lambda_2) > 0 \ , \tag{4.59}$$

$$\sin^2\alpha = 27H^2M^2 \quad \Longrightarrow \quad r'(\lambda_2) > 0 \ . \tag{4.60}$$

Combining (4.55), (4.58) and (4.53), we conclude that

$$\begin{aligned} \cos\alpha &< 0 \quad \text{and} \quad \sin^2\alpha < 27H^2M^2 \\ \Longrightarrow \quad & r(\lambda_1) = r_b \quad \text{and} \quad T(\lambda_1^+) = -\infty \\ \Longrightarrow \quad & r(T=-\infty) = r_b \ . \end{aligned} \tag{4.61}$$

Similarly, combining (4.56), (4.59) and (4.54), we have

$$\begin{aligned} \cos\alpha &< 0 \quad \text{and} \quad \sin^2\alpha > 27H^2M^2 \\ \Longrightarrow \quad & r(\lambda_1) = r_c \quad \text{and} \quad T(\lambda_1^+) = -\infty \\ \Longrightarrow \quad & r(T=-\infty) = r_c \ . \end{aligned} \tag{4.62}$$

Lastly, combining (4.55)-(4.57), (4.58)-(4.58) and (4.54), we get

$$\cos\alpha \geq 0 \quad \Longrightarrow \quad r(\lambda_2) = r_c \quad \text{and} \quad T(\lambda_2^+) = -\infty \quad \Longrightarrow \quad r(T=-\infty) = r_c \ . \tag{4.63}$$

The three cases (4.61), (4.62) and (4.63) give the limiting behavior of $r(T)$ as $T \to \infty$ for all values of α except those for which

$$\cos\alpha < 0 \quad \text{and} \quad \sin^2\alpha = 27H^2M^2 \ .$$

From figure 4.6 we see that the above corresponds to the critical case where the light ray trajectory approaches the unstable equilibrium at $r = 3M$ and satisfies

$$r(\lambda = -\infty) = 3M \ .$$

Since $T'(-\infty) > 0$, as shown in section 4.2.2, we have $T(\lambda = -\infty) = -\infty$ and the above implies that

$$\begin{aligned} \cos\alpha &< 0 \quad \text{and} \quad \sin^2\alpha = 27H^2M^2 \\ \Longrightarrow \quad & r(\lambda=-\infty) = 3M \quad \text{and} \quad T(\lambda=-\infty) = -\infty \\ \Longrightarrow \quad & r(T=-\infty) = 3M \ . \end{aligned} \tag{4.64}$$

Finally, we can put together (4.61), (4.62), (4.63) and (4.64), so that we have the behavior of $r(T)$

as $T \to -\infty$ for all values of α:

$$\begin{aligned}
\cos\alpha &< 0 \quad \text{and} \quad \sin^2\alpha < 27H^2M^2 \quad \Longrightarrow \quad r(T=-\infty) = r_b \ , \\
\cos\alpha &< 0 \quad \text{and} \quad \sin^2\alpha = 27H^2M^2 \quad \Longrightarrow \quad r(T=-\infty) = 3M \ , \\
\cos\alpha &\geq 0 \quad \text{or} \quad \sin^2\alpha > 27H^2M^2 \quad \Longrightarrow \quad r(T=-\infty) = r_c \ .
\end{aligned}$$

The above results are summarized in figure 4.6. They confirm the basic separation of generators into those which start either near $r = r_b$ or $r = r_c$, as was discovered numerically and can be seen in the first frame of figure 4.4. Furthermore, we have given precise values of the parameter α where this separation occurs. Notice that the above result is not merely stating that the horizon at early times consists of the spheres $r = r_b$ and $r = r_c$, as could easily be deduced from the Penrose diagram in figure 3.1b. Instead, our result is that all generators except those with $\sin^2\alpha \neq 27H^2M^2$ and $\cos\alpha < 0$, including those which do not start on the horizon, approach $r = r_b$ or $r = r_c$ as $T \to -\infty$.

A further subdivision separates null generators into the following two categories: those which start on either the black hole or cosmological horizon, as opposed to those which only join these horizons later. This distinction is best understood by looking at figure 4.7, where the color of the generator indicates its origin. Red and blue generators start on the cosmological and black hole horizons, respectively, whereas green generators do not start on either horizon, but instead join the horizons by entering through the caustic points along $\phi = \pi$. Based on the last frame in this figure, we see that we can introduce $\alpha_1 \in [0,\pi)$ and $\alpha_2 \in [0,\pi)$ such that generators satisfying $\alpha_1 \leq \alpha \leq \alpha_2$ or $\alpha_1 \leq 2\pi - \alpha \leq \alpha_2$ will be those which did not begin on either horizon (i.e. green generators).

Suppose we assume that the horizons have the basic shape as shown in figure 4.4, with caustic points on the horizon occuring if and only if $\phi = \pi$. This claim can be proven analytically through a detailed analysis of the properties of the function $\phi(T;\alpha)$, but we will not pursue this here. With this basic assumption it follows from the fact that generators can only enter through caustic points, as discussed in section 4.1, that we have

$$\alpha_1 \leq \alpha \leq \alpha_2 \quad \text{or} \quad \alpha_1 \leq 2\pi - \alpha \leq \alpha_2 \Longleftrightarrow |\phi(T=-\infty)| \geq \pi \ . \tag{4.65}$$

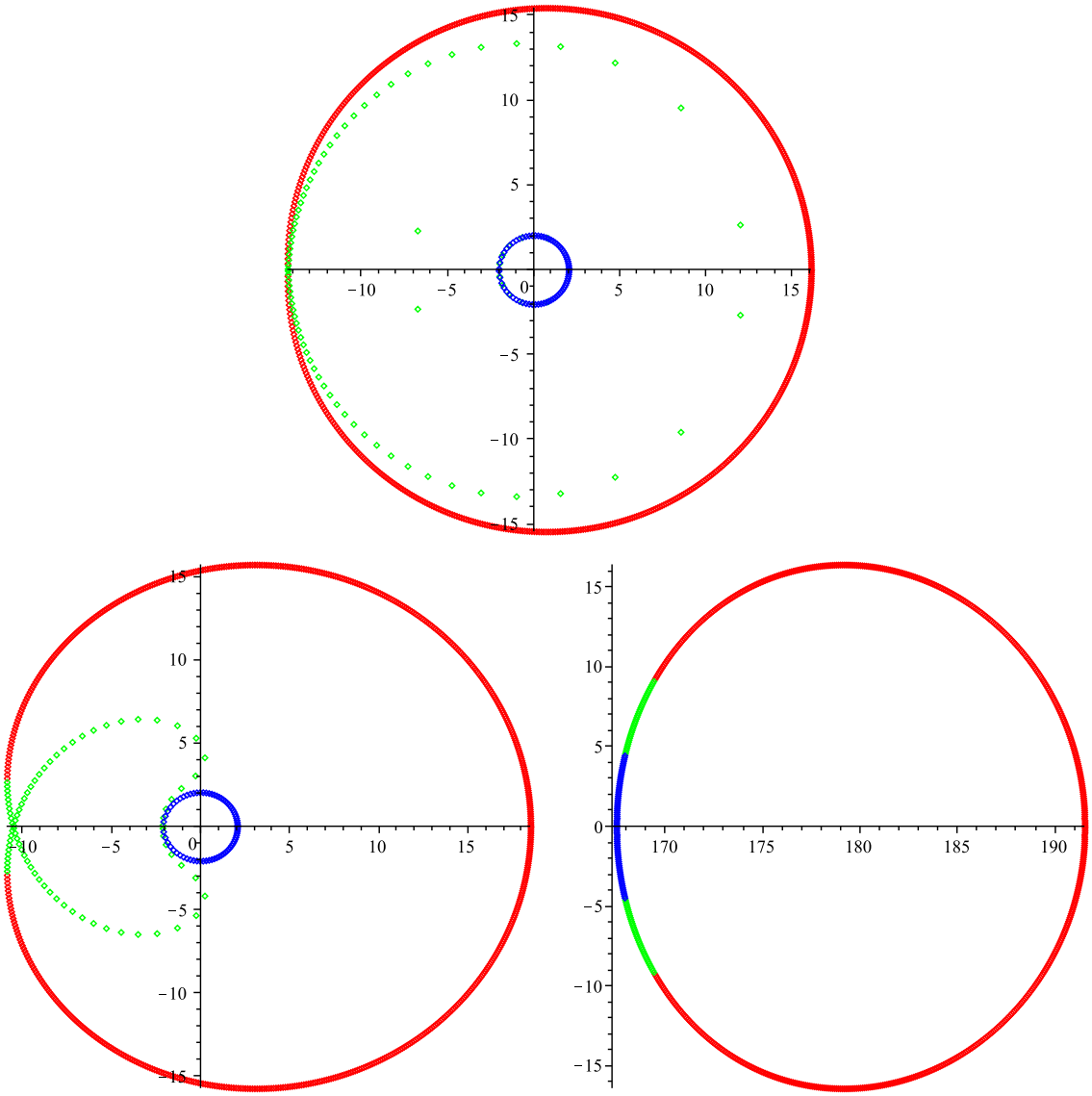


Figure 4.7: Three frames showing the merging of horizons, with a color scheme to indicate the origin of each null generator. Blue and red are used to indicate null generators which begin on the black hole and cosmological horizons, respectively. Green is used to indicate generators which begin on neither horizon. This color scheme is unlike the one in figure 4.4, where we used the colors to indicate if a null generator was currently part of either the cosmological horizon, the black hole horizon, or neither.

This above condition will be important in section 5.3.2, where a calculation of α_1 and α_2 for $\varepsilon \ll 1$ will be used to find the relative area increase due to the three types of generators, in the limit that $\varepsilon \to 0$.

4.4.3 Horizon at late times

For sufficiently large values of LP coordinate time T, we can see from the effective potential diagram (figure 4.6) that $\frac{dr}{dT} > 0$, so that $r(\hat{T})$ obeys the following equation:

$$\frac{dr}{d\hat{T}} = \frac{2}{H}\frac{1}{1-\hat{T}^2}G_+(r) \ , \tag{4.66}$$

where $G_+(r)$ is the right hand side of (4.12) for the case $\frac{dr}{dT} > 0$. Let $\xi = 1 - \hat{T}$ replace $\hat{T}$ as the independent variable. The equation (4.66) above becomes the following equation for $r(\xi)$:

$$\frac{dr}{d\xi} = \frac{-1}{H\xi\,(1-\xi/2)}G_+(r) \ . \tag{4.67}$$

We will look for an approximate asymptotic solution of the form

$$r(\xi) = \frac{1}{\xi}\sum_{i=0}^{n} r_i \xi^i + \mathscr{O}(\xi^{n+1}) \ , \tag{4.68}$$

where the coefficients r_i are allowed to depend on M, H and α. Notice that $\xi \to 0^+$ corresponds to $T \to \infty$, and that (4.68) automatically incorporates the asymptotic requirement (4.36). In section 4.4.3, we briefly discuss the movitation behind choosing an expansion of the form (4.68). We then find the expansions for $r(\xi)$ and $\phi(\xi)$, which will be (or already have been) used for several purposes. In section 4.4.1, they were used as initial conditions when solving the null geodesic equations from section 4.3.1 numerically. In section 4.4.3 below, we will use $r(\xi)$ and $\phi(\xi)$ to describe and analyze the shape of the horizon as $T \to \infty$. In section 4.4.3, we discuss the validity of the above expansion.

The coefficient r_0

Although we have allowed r_0 to depend on α for the sake of generality, we can see that if the shape of the horizon at late times is to be a closed surface at a finite distance from the observer (as discussed in section 4.1), then r_0 must be independent of α. The reasoning is as follows. Suppose $r_0(\alpha_1) \neq r_0(\alpha_2)$, so that r_0 does depend on α. Then if $r(\xi;\alpha)$ is the solution to equation (4.67), with the expansion given by (4.68), we have that $\|(r(\xi;\alpha_1),\phi(\xi;\alpha_1)) - (r(\xi;\alpha_2),\phi(\xi;\alpha_2))\| \to \infty$ as $\varepsilon \to 0$, where the norm $\|\cdot\|$ is, for example, the geodesic distance along the hypersurface $T =$ constant, with the metric (3.39) providing the measure of distance, and we have set $\theta = \pi/2$ for simplicity. In other words, if r_0 depends on α, the distance between two geodesics with different values of r_0 would become infinite as $T \to \infty$ (i.e. $\xi \to 0^+$), and this would clearly contradict our requirement and expectation that the null geodesics form a closed surface a finite distance from the observer as $T \to \infty$. This expectation was discussed in 4.1, and is motivated by the fact that as $T \to \infty$, the observer is in a region of spacetime well approximated by deSitter spacetime.

As can be deduced from (4.72) below, the coefficient r_0 in the expansion (4.68) is essentially arbitrary, in the sense that changing the value of r_0 simply amounts to a shift in the time variable T. This is unlike the coefficients r_1, r_2 and r_3 in the expansion, which will be determined by substituting the expansion (4.68) into the equation (4.67). Since changing the value of r_0 corresponds to a simple shift of the time variable T, we will use the following convenient value:

$$r_0 = \frac{2}{H} \ .$$

Motivation for the expansion

To understand the expansion (4.68), let us first rewrite it as

$$r(\xi) = \frac{r_0}{\xi} + r_1 + \sum_{k=2}^{n} r_k \xi^k + \mathscr{O}(\xi^{n+1}) \ . \tag{4.69}$$

Consider the first term on the right hand side, which characterizes the lowest order asymptotic behavior of $r(\xi)$:

$$r(\xi) \sim \frac{r_0}{\xi} \ . \tag{4.70}$$

To understand this lowest order behavior, we invert the relation (4.30) and use $\xi = 1 - \hat{T}$ to get

$$e^{HT} = \frac{2}{\xi} - 1 \ . \tag{4.71}$$

Using the above in (4.69), this gives

$$r(T) = \frac{r_0}{2} e^{HT} + \mathscr{O}(1) \ . \tag{4.72}$$

As discussed in section 4.4.3, r_0 is independent of α. Provided that we have $\phi \sim \xi$, which can be deduced from (4.77) below, it follows from the above that the family of light rays with $\alpha \in [0, 2\pi)$ are all a finite distance away from each other. Here the measurement of distance is made using the proper length of spacelike geodesics on the hypersurfaces $T =$ constant. Now compare the above with the r coordinate as a function of proper time for our drifting observer, which can be deduced from the $r \to \infty$ limit of equation (2.12):

$$r_{obs}(T) = A e^{HT} + \mathscr{O}(1) \ . \tag{4.73}$$

The above has precisely the same form as (4.72), so that the light rays with $r(T)$ given by (4.72) can be interpreted as not only a finite distance from one another, but a finite distance from the observer as well. As with the constant r_0, changing the constant A in the above amounts to shifting the time variable T by a constant amount. Provided that we choose $A = r_0/2$, the observer's trajectory will be precisely centered on the family of light rays with coordinates given by (4.72). Changing the constant r_0 amounts to shifting between the different observer trajectory curves in figure 4.2a, and changing the constant A amounts to shifting between the different light cones shown in figure 4.2b. By choosing $A = r_0/2$, we are requiring that the observer's trajectory (one of the curves in figure 4.2a) is correctly matched with one of the light cones from figure 4.2b.

The fact that (4.72) and (4.73) have the same exponential dependence on T can be seen as a conse-

quence of the way we constructed LP coordinates. By choosing timelike geodesic trajectories as the timelike coordinate curves, and by constructing spacelike coordinate curves which are orthogonal to the timelike curves, we have ensured that the coordinates of the null generators at late times follow the coordinates of the observer. This is similar to the way that the cosmological horizon is a finite distance from an observer in deSitter spacetime, where the distance is measured along the flat planar slices of the deSitter universe.

Another way of understanding the asymptotic behavior (4.70) is to consider the equation (4.12) for large values of r, so that it becomes

$$\frac{dr}{dT} = Hr + \mathscr{O}(1) \ .$$

The above has the approximate asymptotic solution (4.72), which then becomes (4.70) after using (4.71).

Next consider the second term on the right hand side of (4.69). This term is $\mathscr{O}(1)$, and we can write

$$r(\xi) - \frac{r_0}{\xi} \sim r_1 \ .$$

Notice that the right hand side above is the lowest order term which depends on α, and therefore will be the largest term in the series which contributes to the shape of the horizon as $T \to \infty$. The fact that this term is also independent of ξ suggests that it will play a role in determining a final shape of the horizon. In section 4.4.3 below, we show that there is indeed a final shape which is independent of T, and we will use the $\mathscr{O}(1)$ and $\mathscr{O}(\xi)$ terms of $r(\xi)$ and $\phi(\xi)$, respectively, to derive an anlalytical formula for this shape. This final shape can also be seen in the last frame of figure 4.4.

Finally, consider the terms which are $\mathscr{O}(\xi)$ and smaller in (4.69). These terms will play a role in determining the slight deformations that occur in the final shape for large values of T (i.e. for $\xi \ll 1$). An example of such a deformation can be seen in the second to last frame of figure 4.4.

We have motivated the form of the series (4.69), and we will confirm its accuracy by comparing it to numerical solutions in section 4.4.3. However, as will be discussed in section 4.4.3, we have good reason to believe that the series (4.69) is not an exact solution but instead is only an asymptotic series.

The expansion coefficients r_1, r_2 and r_3

To find the unkown coefficients r_1, r_2 and r_3, we first expand the function $G_+(r)$:

$$\begin{aligned} G_+(r) &= Hr - f(r_e)^{1/2}\cos\alpha + \left(\frac{f(r_e)\cos 2\alpha - 1}{2H}\right)\frac{1}{r} \\ &\quad + \left(\frac{\left(f(r_e)-\frac{1}{2}\right)f(r_e)^{1/2}\cos\alpha\sin^2\alpha + MH}{H^2}\right)\frac{1}{r^2} + \mathscr{O}\left(\frac{1}{r^3}\right) . \end{aligned} \tag{4.74}$$

From (4.68) we have

$$\begin{aligned} \frac{1}{r} &= \frac{1}{r_0}\xi - \frac{r_1}{r_0^2}\xi^2 + \mathscr{O}(\xi^3), \\ \frac{1}{r^2} &= \frac{\xi^2}{r_0^2} + \mathscr{O}(\xi^3) . \end{aligned} \tag{4.75}$$

Substituting the above and (4.68) into (4.74) and collecting terms of like powers, we obtain

$$\begin{aligned} G_+(\xi) &= Hr_0\frac{1}{\xi} + Hr_1 - f(r_e)^{1/2}\cos\alpha + \left(\frac{f(r_e)\cos 2\alpha - 1}{2Hr_0} + Hr_2\right)\xi \\ &\quad + \left(Hr_3 - \frac{f(r_e)\cos 2\alpha - 1}{2H}\left(\frac{r_1}{r_0^2}\right) + \frac{\left(f(r_e)-\frac{1}{2}\right)f(r_e)^{1/2}\cos\alpha\sin^2\alpha + MH}{H^2 r_0^2}\right)\xi^2 + \mathscr{O}(\xi^3) . \end{aligned}$$

Substituting the above and the expansion

$$\frac{1}{1-\xi/2} = 1 + \frac{\xi}{2} + \frac{\xi^2}{4} + \frac{\xi^3}{8} + \mathscr{O}(\xi^4)$$

into (4.67) and collecting like powers again, we obtain

$$\begin{aligned}\frac{dr}{d\xi} = -r_0\frac{1}{\xi^2} - \left(\frac{r_0}{2} + r_1 - \frac{f(r_e)^{1/2}\cos\alpha}{H}\right)\frac{1}{\xi} \\ - \left(\frac{r_0}{4} + \frac{r_1}{2} - \frac{f(r_e)^{1/2}\cos\alpha}{2H} + \frac{f(r_e)\cos 2\alpha - 1}{2H^2 r_0} + r_2\right) \\ - \left(\frac{r_0}{8} + \frac{r_1}{4} - \frac{f(r_e)^{1/2}\cos\alpha}{4H} + \frac{f(r_e)\cos 2\alpha - 1}{4H^2 r_0^2}(r_0 - 2r_1) + \frac{r_2}{2} + r_3 \right. \\ \left. + \frac{\left(f(r_e) - \frac{1}{2}\right) f(r_e)^{1/2}\cos\alpha\sin^2\alpha + MH}{H^3 r_0^2}\right)\xi + \mathscr{O}(\xi^2)\ .\end{aligned}$$

From (4.68), we also know that

$$\frac{dr}{d\xi} = -\frac{r_0}{\xi^2} + r_2 + 2r_3\xi + \mathscr{O}(\xi^2)\ .$$

Equating like powers of ξ in our previous two expressions, we can systematically solve for r_1, r_2 and r_3 to get

$$\begin{aligned} r_1 &= \frac{f(r_e)^{1/2}\cos\alpha - 1}{H}, \\ r_2 &= \frac{1 - f(r_e)\cos 2\alpha}{8H}, \\ r_3 &= \frac{1}{24H}\left[\left(f(r_e) - \frac{1}{4}\right) f(r_e)^{1/2}\cos 3\alpha - \frac{3}{2}f(r_e)\cos 2\alpha - \frac{3}{4}f(r_e)^{1/2}\cos\alpha + \frac{3}{2}\right] - \frac{M}{12}, \end{aligned} \tag{4.76}$$

where we have used $r_0 = 2/H$. This process can be continued ad infinitum to obtain all higher order coefficients in the series. More practically, one can use a computer algebra program such as MAPLE to find the coefficients. For example, one can apply the "series" method of the dsolve routine of MAPLE 14 to equation (4.67) to determine the higher order coefficients in the expansion.

Expansion for $\phi(\xi)$

We can use the expansion for $r(\xi)$ to obtain an expansion for $\phi(\xi)$, valid for $\xi \ll 1$. For this purpose it is useful to first use r as the independent variable instead of ξ. It is possible to do this for $\xi \ll 1$ since $r(\xi)$ is a monotonic function. The equation for $\phi(r)$ can be found by dividing the equation for $\phi(T)$ by the equation for $r(T)$ in (4.12). This gives

$$\frac{d\phi}{dr} = \frac{-\sin\alpha}{r^2}\frac{1}{(H^2 - V_{eff}(r))^{1/2}} \ .$$

Expanding the right hand side for large r gives

$$\frac{d\phi}{dr} = \frac{-\sin\alpha}{H}\left(\frac{1}{r^2} + \frac{\sin^2\alpha}{2H^2}\frac{1}{r^4} - \frac{M\sin^2\alpha}{H^2}\frac{1}{r^5}\right) + \mathscr{O}\left(\frac{1}{r^6}\right) \ .$$

Now integrate with respect to r to get

$$\phi(r) = \frac{\sin\alpha}{H}\left(\frac{1}{r} + \frac{1}{3}\frac{\sin^2\alpha}{2H^2}\frac{1}{r^3} - \frac{M\sin^2\alpha}{4H^2}\frac{1}{r^4}\right) + \mathscr{O}\left(\frac{1}{r^5}\right) \ . \tag{4.77}$$

(4.68) implies the following expansions:

$$\begin{aligned}
\frac{1}{r} &= \frac{\xi}{r_0}\left(1 - \frac{r_1}{r_0}\xi + \frac{r_1^2 - r_2 r_0}{r_0^2}\xi^2 + \frac{2r_1 r_2 r_0 - r_3 r_0^2 - r_1^3}{r_0^3}\xi^3 + \mathscr{O}(\xi^4)\right), \\
\frac{1}{r^3} &= \frac{\xi^3}{r_0^3}\left(1 - \frac{3r_1}{r_0}\xi + \mathscr{O}(\xi^2)\right), \\
\frac{1}{r^4} &= \frac{\xi^4}{r_0^4} + \mathscr{O}(\xi^5) \ .
\end{aligned}$$

Substituting the above into (4.77) and collecting like powers of ξ, we obtain the following series for $\phi(\xi)$:

$$\begin{aligned}\phi(\xi) &= \frac{\sin\alpha}{H}\frac{\xi}{r_0}\left[1-\frac{r_1}{r_0}\xi+\frac{1}{r_0^2}\left(r_1^2-r_2r_0+\frac{\sin^2\alpha}{6H^2}\right)\xi^2\right. \\ &\quad \left.+\frac{1}{r_0^3}\left(2r_1r_2r_0-r_3r_0^2-r_1^3-r_1\frac{\sin^2\alpha}{2H^2}-\frac{M\sin^2\alpha}{4H^2}\right)\xi^3+\mathscr{O}(\xi^4)\right] ,\end{aligned} \tag{4.78}$$

where r_1, r_2 and r_3 are given by (4.76). Writing the above as

$$\phi(\xi)=\phi_1\xi+\phi_2\xi^2+\phi_3\xi^3+\phi_4\xi^4+\mathscr{O}\left(\xi^5\right) , \tag{4.79}$$

we find the following explicit formulas for the coefficients of the expansion:

$$\phi_1 = \frac{\sin\alpha}{2}, \tag{4.80}$$

$$\phi_2 = -\frac{\sin\alpha}{4}\left(f(r_e)^{1/2}\cos\alpha-1\right), \tag{4.81}$$

$$\phi_3 = \frac{\sin\alpha}{24}\left(\frac{1}{4}(9f(r_e)-1)\cos 2\alpha-6f(r_e)^{1/2}\cos\alpha+\frac{3}{2}f(r_e)+\frac{5}{2}\right), \tag{4.82}$$

$$\begin{aligned}\phi_4 &= \frac{\sin\alpha}{16}\left[\frac{2}{3}\left(\frac{1}{4}-f(r_e)\right)f(r_e)^{1/2}\cos 3\alpha+\frac{1}{4}\left(9f(r_e)+\frac{MH}{2}-1\right)\cos 2\alpha\right. \\ &\quad \left.-\left(f(r_e)+\frac{5}{2}\right)f(r_e)^{1/2}\cos\alpha+\frac{3}{2}f(r_e)+\frac{5}{24}MH+\frac{1}{2}\right] .\end{aligned} \tag{4.83}$$

where we have used $r_0 = 2/H$.

Validity of the expansion

Our choice for the form of the series (4.68) was not motivated by any rigorous mathematical theorem or technique. However, it is nevertheless accurate, as can be seen in figure 4.8. To understand our series solution further, it is useful to make a change of variables and define $\bar{r}(\xi)$ as

$$\bar{r}(\xi)=\xi r(\xi)-r_0 . \tag{4.84}$$

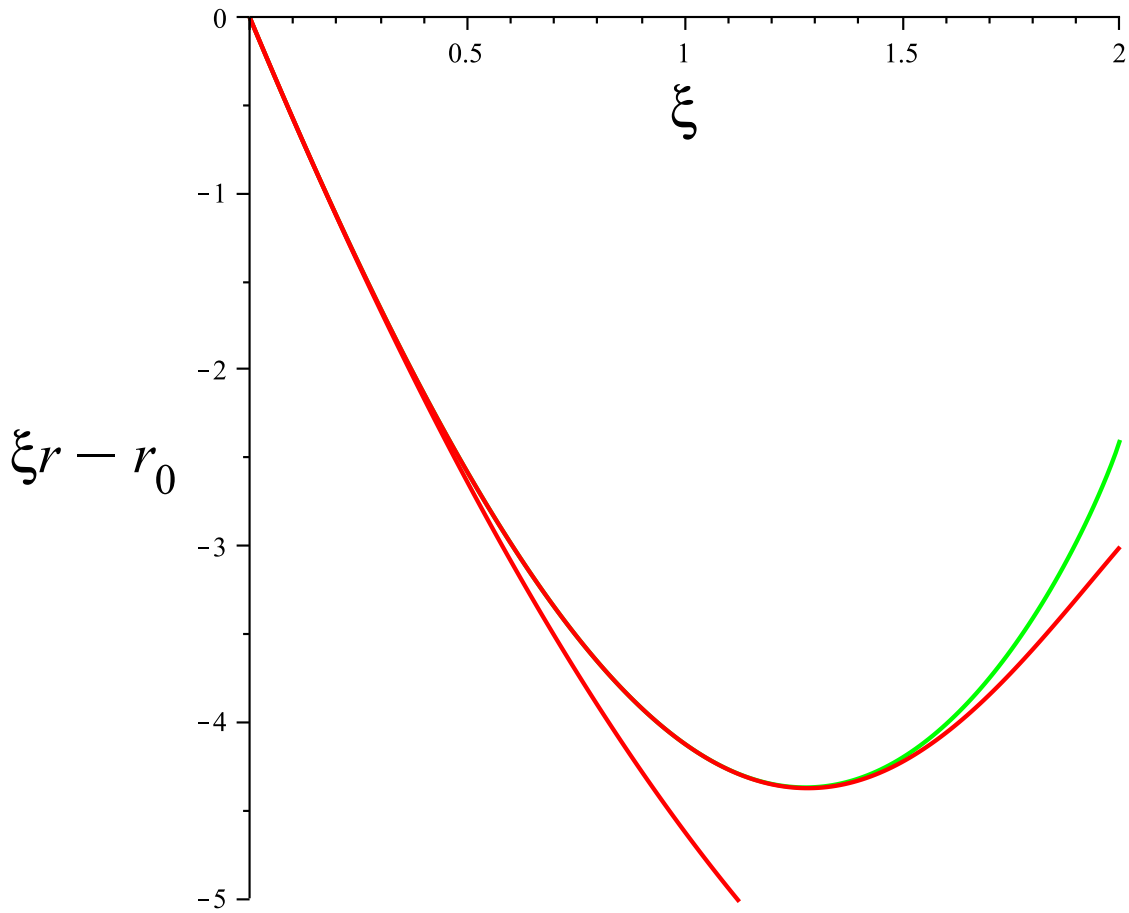


Figure 4.8: The green curve is the numerical solution to (4.85), and the two red curves are the series solution to (4.85) at third order and eleventh order. Notice the remarkable accuracy of the eleventh order series solution, especially for $0 < \xi < 1$. Parameter values used are: $M = 1$, $H = (6\sqrt{3})^{-1}$, $\alpha = \pi/4$. The initial condition leading to the numerical solution is set by using $\bar{r}(\xi_0) = \bar{r}_{(2)}(\xi_0)$, where $\xi_0 = 10^{-10}$ and $\bar{r}_{(2)}(\xi)$ is the second order series solution.

The equation (4.67) for $r(\xi)$ can then be transformed into the following equation for $\bar{r}(\xi)$:

$$\frac{d\bar{r}}{d\xi} = \frac{1}{\xi}\left(\frac{-2\xi}{H(2-\xi)}G_+\left(\frac{r_0+\bar{r}}{\xi}\right)+r_0+\bar{r}\right) . \tag{4.85}$$

In figure 4.8, we plot the numerical solution to the above equation, which for our purposes can be considered the exact solution, along with the series solution to the above at two different orders. The series solution is obtained using MAPLE's dsolve procedure with the series method, and leads to the same answer as would be obtained by substituting (4.68)-(4.76) into (4.84). We see the increased accuracy with higher order, as well as the remarkable accuracy of the higher order solution, especially for values $0 < \xi < 1$ (notice that $\xi \to 2$ corresponds to $T \to -\infty$). This agreement with the numerical solution increases our confidence in the series expansion, especially for smaller values of ξ.

The difficulty in applying traditional mathematical techniques and theorems to equation (4.85) lies in the fact that it is ill-defined at $\xi = 0$. Although the right hand side of (4.85) is undefined at $\xi = 0$, it approaches a finite limit as $\xi \to 0$ provided that

$$\bar{r}(\xi) \sim \xi .$$

Based on this we can construct the approximate solution

$$\bar{r}(\xi) = r_1\xi + o(\xi) ,$$

where $o(\cdot)$ is little-o notation, and r_1 is given by (4.76). Continuing this process, we can construct the following series:

$$\bar{r}(\xi) = \sum_{i=1}^{n} r_i\xi^i + o(\xi^n) , \tag{4.86}$$

where the coefficients r_i are determined iteratively using the procedure outlined in section (4.4.3). The above is an approximation to the solution of (4.85) in the sense of being an asymptotic series. However, we cannot be assured that it approaches an exact solution of (4.85) as $n \to \infty$. In fact it is unlikely to be an exact solution. For example, if it is an exact solution then we ought to be able to use it to find the asymptotic behavior of the horizon area at late times. With this in mind, let us

substitute the power series (4.68) into the formula for the area element dA in (5.4), which will be derived in chapter 5. Doing so, and collecting terms of like powers in ξ, we find

$$dA_p = \frac{|\sin\alpha|}{H^2} + o\left(\xi^4\right) \ , \tag{4.87}$$

where we have attached the subscript p to dA to remind ourselves that we are calculating dA in the context of approximating $r(\xi)$ by a power series. Integrating the above to obtain the cosmological horizon area, as in equation (5.7), we get

$$A_p(\xi) = 4\pi\left(\frac{1}{H}\right)^2 + o\left(\xi^4\right) \ , \tag{4.88}$$

where the subscript p attached to $A(\xi)$ has the same meaning as before. By choosing the specific values $M = 1$ and $H = (6\sqrt{3})^{-1}$ and using MAPLE to find the series for $r(\xi)$ and $A(\xi)$, we can extend this result and calculate that the fifth order term vanishes as well. Beyond fifth order the MAPLE calculations become prohibitively time consuming. If this vanishing of terms extends to all orders, then using a power series for $r(\xi)$ leads to the result that the horizon area is constant in time, which is clearly false. This provides a hint that the power series (4.68) is not an exact solution, and that either it must be a divergent series, or there must be additional terms in the solution to (4.85) which are not positive integer powers of ξ. The idea that the exact solution to (4.85) cannot be expressed as a series in positive integer powers of ξ is actually a natural expectation, given if there was such an exact solution, then the area at late times would be of the form

$$A(\xi) = 4\pi\left(\frac{1}{H}\right)^2 - a\xi^n + o(\xi^n) \ ,$$

for some $a > 0$ and natural number n. Then using (4.71) to return to the original time variable T, the above becomes

$$A(T) = 4\pi\left(\frac{1}{H}\right)^2 - a\,2^n\,e^{-nHT} + o(\xi^n) \ .$$

However, there is no reason to think that as $T \to \infty$ the area $A(T)$ of horizons should have the asymptotic behavior

$$A(T) - A(\infty) \sim e^{-nHT}$$

for some natural number n, and we once again have a hint that a power series cannot be an exact solution.

Although the preceding discussion suggests that (4.86) does not converge to the exact solution of (4.85) as $n \to \infty$, we have found that numerical evidence strongly suggests that the series in (4.86) does converge as $n \to \infty$. In figure 4.9, we plot the value of the series solution at two different values of ξ, for different values of n in (4.86). Suppose we are correct and the series does converge, but does not converge to the exact solution. Let $\bar{r}_e(\xi)$ be the exact solution to (4.85) and $\bar{r}_p(\xi)$ be the asymptotic power series solution, and define the difference function $d(\xi)$ as

$$d(\xi) = \bar{r}_e(\xi) - \bar{r}_p(\xi)\,.$$

Now suppose we describe the change in shape of the horizon using a transverse deformation tensor $\tilde{B}_{ab}$ and decompose it into a trace part and two trace free parts in the usual manner [25]:

$$\tilde{B}_{ab} = \theta\,\delta_{ab} + \sigma_{ab} + \omega_{ab}\ .$$

Given that the horizon consists of a family of null generators which form a null geodesic congruence which is hypersurface orthogonal, it follows that the rotation ω_{ab} for the congruence vanishes [25], so that we are only left with the expansion term containing θ and the shear term σ_{ab} in the above. From the result (4.87), it appears that the asymptotic power series solution $\bar{r}_p(\xi)$ only describes a change in shape of the horizon, without affecting its area, and thus the difference $d(\xi)$ must contain the dependence on ξ which probably accounts for the change in area of the horizons, as captured by the expansion parameter θ. On the other hand, the power series $\bar{r}_p(\xi)$ describes a dependence on ξ which decribes a shearing of the horizon. That is, it describes a change in shape which does not affect the area and which does not cause any rotation.

The final issue which must be addressed is the question of the uniqueness of the solution to (4.85). Because the right hand side is not defined at $\xi = 0$, we are not aware of any theorems that can be applied to ensure the uniqueness of a solution to the equation (4.85) with the "initial condition"

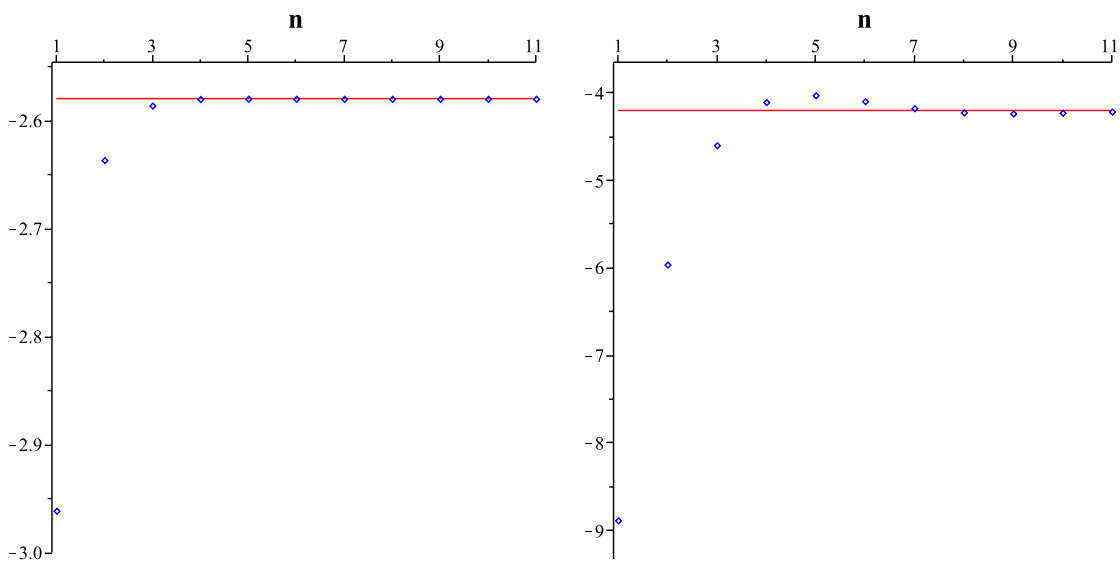


Figure 4.9: Plot of the series solution (4.86) (blue points) vs order n for $\xi = 0.5$ (left) and $\xi = 1.5$ (right). The horizontal red line is the numerically determined solution to (4.85), evaluated at $\xi = 0.5$ (left) or $\xi = 1.5$ (right). The blues points appear to converge to a limiting value, suggesting that the series in (4.86) does converge to a final value as $n \to \infty$. Although it appears that the series solution (blue points) converges to the numerical solution (red line), we believe that there is a residual difference between the two that cannot be eliminated, even in the limit as $n \to \infty$. This difference is too small to be seen on the plots, however. Parameter values used: $M = 1$, $H = (3\sqrt{6})^{-1}$, $\alpha = \pi/4$. Initial conditions for the numerical solution as set using the $n = 2$ series solution at $\xi = 10^{-10}$.

$\bar{r}(0) = 0$ (more precisely, because it is not possible to define a solution for $\xi = 0$, we can only define a limiting condition on $\bar{r}(\varepsilon)$, which should be written $\bar{r}(0^+) = 0$). However, one can define the initial condition $\bar{r}(\xi_0) = \bar{r}_0$ for some $\xi_0 > 0$. The Picard-Lindelof theorem can then be used to ensure a unique solution for some interval about ξ_0. To ensure a unique solution satisfying $\bar{r}(0^+) = 0$, on must show that there is a unique value of $\bar{r}_0$ such that $\bar{r}(0^+) = 0$. Doing this is beyond the scope of this thesis. However, we have plotted the phase portrait associated with (4.85) and have confirmed numerically that there is a single trajectory satisfying $\bar{r}(0^+) = 0$.

The shape of the horizon at late times

As can be seen in figure 4.4, the horizon at late times settles down to a final shape. As discussed in section 4.4.1, this shape is not spherical (i.e. not circular when one dimension is suppressed). This may seem suprising, but is simply an artifact of the choice of time slicing. Although the horizon at late times is in a part of SdS spacetime that is well approximated by deSitter spacetime, the slices $T =$ constant in this part of the spacetime do not coincide with the usual constant time slices of deSitter spacetime in the so-called planar coordinates. The unusual choice of time slices obscures the fact that the horizon, when viewed as a null hypersurface, does indeed approach the horizon of deSitter spacetime.

In this section, we combine the expansions for $r(\xi)$ and $\phi(\xi)$ (equations (4.68) and (4.78)) to describe this final shape of the horizon at late times analytically.

In order to describe this final shape, it is useful to introduce the variables ρ and ψ as follows:

$$\begin{aligned}\rho &= \left((r\cos\phi - r_{avg})^2 + (r\sin\phi)^2\right)^{1/2} , \\ \psi &= \arctan\left(\frac{r\sin\phi}{r\cos\phi - r_{avg}}\right) .\end{aligned}$$

where $r_{avg} = (r(\xi;\alpha=0)+r(\xi;\alpha=\pi))/2$. ρ and ψ can be thought of as polar coordinates centered at the point with coordinates $r = r_{avg}$ and $\phi = 0$, where r_{avg} is the midpoint of the two parts of the horizon which intersect the line $\phi = 0$. Substituting the expansions (4.68) and (4.78) for $r(\xi)$ and $\phi(\xi)$ into the above, we can obtain the expansions for $\rho(\xi)$ and $\psi(\xi)$. Doing so, and keeping only the lowest order term in each expansion, we obtain

$$\begin{aligned}\rho &= \frac{1}{H}\sqrt{f(r_e)\cos^2\alpha + \sin^2\alpha} + \mathscr{O}(\xi), \\ \psi &= \arctan\left(\frac{1}{\sqrt{f(r_e)}}\tan\alpha\right) + \mathscr{O}(\xi) .\end{aligned}$$

Truncating at the lowest order and eliminating α, we can obtain $\rho(\psi)$ in the limit that $\xi \to 0$:

$$\rho(\psi) = \frac{1}{H}\sqrt{1 + \frac{f(r_e) - 1}{1 + f(r_e)\tan^2\psi}}\,. \tag{4.89}$$

Plotting the above gives the same shape as the last frame in figure 4.4, with ρ a maximum at $\psi = \pm\pi/2$. Notice that in the limit that the black hole mass vanishes, i.e., for $M = 0$, we have $f(r_e) = 1$ and therefore $\rho(\psi) = 1/H$, so that the horizon shape is circular (or spherical if the θ variable is restored). This is precisely what we would expect, since the case $M = 0$ corresponds to deSitter spacetime, and the cosmological horizon shape in deSitter spacetime is spherical. Notice also that for $M = 0$ the spacelike hypersurfaces $T =$ constant of the LP coordinates used in this thesis are the same as the constant time slices of the conventional planar coordinates of deSitter spacetime (as was discussed in section 3.3.1).

4.5 Merger point

In this section we find the location in spacetime of the point where the black hole and cosmological horizons first touch (as in the fourth frame of figure 4.4). More precisely, we find the value of the Schwarzschild coordinate r for this event in spacetime, in the limit that $\varepsilon \to 0$ (i.e. the limit of an infinitesimally small black hole). Due to the Killing symmetry discussed in section 3.3.5, it is not possible to assign an unambiguous value for the time coordinate T of this merger point. Although we introduce the merger time T_p below, only its derivative is meaningful, since it is always possible to add a constant to T_p.

Although we are finding the value the Schwarzschild coordinate r, and thus dealing with a particular coordinate system, the result is essentially coordinate independent. This is due to the fact that the coordinate r has a simple geometric interpretation which can be implemented in any other coordinate system.

4.5.1 Defining equations

Let (T_p, r_p) be the LP coordinates of the merger point and let α_p be the value of the parameter α for the unique generator that goes through the merger point (we use the subscript p instead of m to indicate the merger point). Note that this generator is really only unique when suppressing the θ coordinate by setting $\theta = \pi/2$; otherwise it unique up to rotations about the "x-axis" shown in figure 4.4.

There are three equations relating the three unknowns T_p, r_p and α_p. These equations are

$$\phi(T_p;\alpha_p) = \pi \,, \tag{4.90}$$

$$r(T_p;\alpha_p) = r_p \,, \tag{4.91}$$

$$\left.\frac{\partial T}{\partial \alpha}\right|_{\phi=\pi,\alpha=\alpha_p} = 0 \,, \tag{4.92}$$

where $\phi(T;\alpha)$ and $r(T;\alpha)$ are the solutions (4.13) and (4.12), respectively. The function $T(\phi;\alpha)$ in the third equation above is simply the inverse of $\phi(T;\alpha)$, with α as a passive variable not playing any role in the inversion process. The inverse $T(\phi;\alpha)$ is guaranteed to exist by virtue of the fact that $d\phi/dT > 0$ for $\sin\alpha \neq 0$, which will be the case for values of α in a sufficiently small neighborhood about α_p.

(4.90) comes from the fact that we expect the merger point to be located along the "x-axis" (see figure 4.4), and so the ϕ coordinate of the merger point is $\phi = \pi$. This expectation is based on the numerical results presented in section 4.4.1, and is not a logical necessity. (4.91) simply follows from the definition of T_p, r_p and α_p. To understand (4.92), consider the function $T_\pi(\alpha) = T(\phi = \pi;\alpha)$. The equation $T_\pi(\alpha) = T_0$ has no roots for $T_0 > T_p$, two roots for $T_0 < T_p$ and precisely one root for $T_0 = T_p$. Furthermore, by the definitions of α_p and T_p, we clearly have $T_\pi(\alpha_p) = T_p$. Hence it follows that $T_\pi(\alpha)$ has a local maximum at α_p, and we have

$$\left.\frac{dT_\pi}{d\alpha}\right|_{\alpha=\alpha_p} = 0 \,.$$

The above equation is precisely (4.92), written using the notation $T_\pi(\alpha) = T(\phi = \pi; \alpha)$.

Below we derive two coupled equations relating r_p and α_p. The first of these equations combines (4.90) and (4.91) into one equation by eliminating T_p. The second of these equations will follow from (4.92).

4.5.2 First equation

In order to derive the equation resulting from (4.90)-(4.91), it will be useful to first consider a family of spacetime points (T_π, r_π) parametrized by α, and defined by

$$\phi(T_\pi; \alpha) = \pi \ , \tag{4.93}$$

$$r(T_\pi; \alpha) = r_\pi \ . \tag{4.94}$$

Once the above are coupled to (4.92), then we will set $\alpha = \alpha_p$ and the above will become the desired equations (4.90)-(4.91). The reason for not setting $\alpha = \alpha_p$ right away is that it will be necessary to leave α arbibtrary when computing $\frac{dT_\pi}{d\alpha}$ in the next section.

For the moment, the only restrictrion we place on the values of α is that they belong to a small neighborhood about α_p. This will ensure that (4.93)-(4.94) has a solution, and that the generators under consideration have a turning point, in the sense that $\frac{dr}{dT} > 0$ for $T > T^*$ and $\frac{dr}{dT} < 0$ for $T < T^*$, where T^* is the time of the turning point. We only consider generators which have a turning point since it turns out this is a necessary condition for (4.92) to have a solution.

In order to utilize the condition (4.93), we first define the functions $\phi_-(r)$ and $\phi_+(r)$ as

$$\phi_\pm(r) \equiv \phi(T_\pm(r)) \ ,$$

where $\phi(T)$ is the solution to (4.13) and $T_-(r)$ and $T_+(r)$ are the inverses of $r(T)$ for $T < T^*$ and

$T > T^*$, respectively. These inverses are guaranteed to exist since $\frac{dr}{dT} > 0$ for $T > T^*$ and $\frac{dr}{dT} < 0$ for $T < T^*$.

The equations obeyed by $\phi_\pm(r)$ are obtained by dividing equation (4.13) by (4.12) and taking $\frac{dr}{dT} > 0$ or $\frac{dr}{dT} < 0$ for $\phi_+(r)$ and $\phi_-(r)$, respectively. This gives

$$\frac{d\phi_\pm}{dr} = \mp \frac{\sin\alpha}{r^2\left(H^2 - V_{\text{eff}}(r)\right)^{1/2}} \ . \tag{4.95}$$

We integrate $\phi_-(r)$ from r_π to the turning point r^* and $\phi_+(r)$ from r^* to $r = \infty$. This gives the following equation relating α, r_π and r^*:

$$\int_\pi^0 d\phi = \int_{r_\pi}^{r^*} \frac{\sin\alpha}{r^2\left(H^2 - V_{\text{eff}}(r)\right)^{1/2}} dr - \int_{r_*}^{\infty} \frac{\sin\alpha}{r^2\left(H^2 - V_{\text{eff}}(r)\right)^{1/2}} dr \ , \tag{4.96}$$

where the lower and upper limits of integration on the left hand side follow from (4.93)-(4.94) and $\phi(r = \infty) = 0$, respectively. Manipulating the limits of integration, we can rewrite the above more succinctly as

$$\pi = \left(2\int_{r^*}^{\infty} - \int_{r_\pi}^{\infty}\right) \frac{\sin\alpha}{r^2\left(H^2 - V_{\text{eff}}(r;\alpha)\right)^{1/2}} dr \ , \tag{4.97}$$

where we have written $V_{eff}(r) = V_{eff}(r;\alpha)$ to emphasize the dependence on α. Since r^* is the turning point for the trajectory with parameter α, we have from (4.4) the following equation relating r^* and α:

$$V_{eff}(r^*;\alpha) = H^2 \ . \tag{4.98}$$

The above can be used to eliminate α from (4.97), so that it becomes an equation simply relating r_π and r^*. In this way r^* replaces the role of α in (4.97).

Next, we define u_π and u_* as

$$u_\pi = \frac{M}{r_\pi}, \tag{4.99}$$

$$u_* = \frac{M}{r^*} \ . \tag{4.100}$$

Written in terms of u_*, the turning point condition (4.98) becomes

$$u_*^2(1-2u_*) = \frac{\varepsilon^2}{\sin^2\alpha} \ . \tag{4.101}$$

The two integrals on the right hand side of (4.97) can be viewed as functions of u_π and u_*. These integrals are best recast using the new variables of integration

$$\begin{aligned} v &= \frac{r^*}{r} \ , \\ u &= \frac{M}{r} \ . \end{aligned}$$

Using these new variables of integration, and replacing the dependence on α, r_π and r^* in favor of a dependence on u_* and u_π instead, (4.97) becomes

$$2I(u_*) - I_1(u_*, u_\pi) = \pi \ , \tag{4.102}$$

where

$$I(u_*) = \int_0^1 \frac{dv}{\left(1-2u_*-v^2(1-2u_*v)\right)^{1/2}} \ , \tag{4.103}$$

$$I_1(u_*, u_\pi) = \int_0^{u_\pi} \frac{du}{\left(u_*^2(1-2u_*)-u^2(1-2u)\right)^{1/2}} \ . \tag{4.104}$$

The equation (4.102), along with the definition of the integrals defined above, gives us an equation relating u_* and u_π. Equivalently, this equation can be tought of as an equation relating r_π and α, with these related to u_π and u_* by (4.99) and (4.101), respectively. Similarly, by setting $\alpha = \alpha_p$ and $r_\pi = r_p$, we obtain an equation relating α_p and r_p. In the next section we use (4.92) to derive another equation relating α_p and r_p.

4.5.3 Second equation

The second equation relating r_p and α_p will be derived using the following steps: First, we find $T'_\pi(\alpha)$. Second, we simplify $T'_\pi(\alpha)$ by taking the limit $T_f \to \infty$, where T_f will be introduced below. Third, we simplify $T'_\pi(\alpha)$ by taking the limit $\bar{u} \to u_*$, where $\bar{u}$ will also be introduced below. In the fourth and final step, we find the equation $T'_\pi(\alpha) = 0$.

First step: Calculation of $T'_\pi(\alpha)$

The second equation relating r_p and α_p will be derived from the condition (4.92), and thus requires us to first find an expression for the function $T(\phi = \pi; \alpha) \equiv T_\pi(\alpha)$. We start by integrating the differential dT from the time $T_\pi(\alpha)$ to some arbitrary final time T_f, with the restrictions that $T_f > T^*$ and that T_f is independent of α. This gives

$$\begin{aligned} \int_{T_\pi}^{T_f} dT &= \int_{T_\pi}^{T^*} dT + \int_{T^*}^{T_f} dT \\ \Longrightarrow \qquad T_f - T_\pi(\alpha) &= \int_{r_\pi}^{r^*} \frac{dT_-}{dr} dr + \int_{r^*}^{r_f} \frac{dT_+}{dr} dr \,. \end{aligned} \tag{4.105}$$

where $r_f = r(T_f; \alpha)$, and $T_-(r; \alpha)$ and $T_+(r; \alpha)$ are the inverses of $r(T; \alpha)$ for $T < T^*$ and $T > T^*$, respectively, with α held fixed throughout the inversion process. In going from the first to the second equation above, we have used (4.94) in changing the limit of integration from T_π to r_π.

Note that in the above we have assumed that the generator under consideration has a turning point, as in the previous section. As was mentioned, this assumption is based on the fact that it is a necessary condition for the equations (4.90)-(4.92) to have a solution.

It should also be noted that since $\frac{dr}{dT} \to 0$ as $T \to T^*$, the integrands in the integrals on the right hand

side above are divergent. However, the integrals are convergent, in keeping with the fact that it takes a finite time for a generator to reach the turning point.

Since $r(T;\alpha)$ is given by (4.14), with $\frac{dr}{dT} < 0$ for $T < T^*$ and $\frac{dr}{dT} > 0$ for $T > T^*$, $\frac{dT_-}{dr}$ and $\frac{dT_+}{dr}$ can be obtained by simply inverting the right hand side of (4.14):

$$\begin{aligned} \frac{dT_-}{dr} &= (G_-(r))^{-1} \ , \\ \frac{dT_+}{dr} &= (G_+(r))^{-1} \ . \end{aligned}$$

(4.105) now becomes

$$T_f - T_\pi(\alpha) = \int_{r^*}^{r_f} \frac{dr}{G_+(r;\alpha)} - \int_{r^*}^{r_\pi} \frac{dr}{G_-(r;\alpha)} \ , \tag{4.106}$$

where we have written $G_\pm(r) = G_\pm(r;\alpha)$ to emphasize the dependence on α, and we have flipped the limits of integration in one of the integrals. As when deriving the first equation in the previous section, it is useful to introduce a new variable of integration:

$$u = \frac{M}{r} \ . \tag{4.107}$$

(4.106) now reads:

$$T_f - T_\pi(\alpha) = M \left(\int_{u_f}^{u_*} \frac{du}{u^2 G_+(\frac{M}{u};\alpha)} - \int_{u_\pi}^{u_*} \frac{du}{u^2 G_-(\frac{M}{u};\alpha)} \right) \ , \tag{4.108}$$

where u_π, u_* and u_f are given by (4.99), (4.100) and $u_f = M/r_f$, respectively. For technical reasons which will be discussed below, it will be useful to split the two integrals above as follows:

$$\begin{aligned} T_f - T_\pi(\alpha) = M \Bigg(&- \int_0^{u_f} \frac{du}{u^2 G_+(\frac{M}{u};\alpha)} + \int_0^{\bar{u}} \frac{du}{u^2 G_+(\frac{M}{u};\alpha)} + \int_{\bar{u}}^{u_*} \frac{du}{u^2 G_+(\frac{M}{u};\alpha)} \\ &+ \int_0^{u_\pi} \frac{du}{u^2 G_-(\frac{M}{u};\alpha)} - \int_0^{\bar{u}} \frac{du}{u^2 G_-(\frac{M}{u};\alpha)} - \int_{\bar{u}}^{u_*} \frac{du}{u^2 G_-(\frac{M}{u};\alpha)} \Bigg) \ , \end{aligned} \tag{4.109}$$

where $\bar{u}$ is a constant independent of α and satifies

$$u_\pi < \bar{u} < u_* \ . \tag{4.110}$$

Such a constant $\bar{u}$ is guaranteed to exist, provided one restricts values of α to a sufficiently small interval about α_p.

Grouping the integrals with the same limits of integration in (4.109), we get

$$T_f - T_\pi(\alpha) = M \left(\int_{\bar{u}}^{u_*} \frac{1}{u^2} \left(\frac{1}{G_+(\frac{M}{u};\alpha)} - \frac{1}{G_-(\frac{M}{u};\alpha)} \right) du + \int_0^{\bar{u}} \frac{1}{u^2} \left(\frac{1}{G_+(\frac{M}{u};\alpha)} - \frac{1}{G_-(\frac{M}{u};\alpha)} \right) du \right.$$
$$\left. + \int_0^{u_\pi} \frac{du}{u^2 G_-(\frac{M}{u};\alpha)} - \int_0^{u_f} \frac{du}{u^2 G_+(\frac{M}{u};\alpha)} \right) \ . \tag{4.111}$$

When differentiating the above with respect to α, one encounters potential singularities at the two problematic points u_* and u_c, where $u_c \equiv r_c/M$. The first integral contains u_* in its domain of integration and the second and third integrals contain u_c in their domains of integration. The fourth integral contain neither. This follows from the inequalities (4.110) and $u_f < u_c < u_\pi$. This latter inequality can be deduced from $r_\pi < r_c < r_f$, where $r_f > r_c$ assumes that T_f has been chosen to be sufficiently large, and where $r_\pi < r_c$ is most easily deduced by looking at the effective potential diagram (figure 4.6), while keeping in mind that $T_\pi < T^*$ and $r(T = -\infty) = r_c$.

By splitting the domains of integration as we have done, we can deal with these singularities separately. The first singularity occurs since differentiating the first integral in (4.111) with respect to α forces us to evaluate the integrand at $u = u_*$, where it is undefined. We can obviate this difficulty by making a change of variables to a new integration variable v defined as follows:

$$v = \frac{u}{u_*} \ . \tag{4.112}$$

Using the change of variables (4.112), (4.111) becomes

$$T_f - T_\pi(\alpha) = M\left(\int_{\bar{u}/u_*}^{1} \frac{1}{u_* v^2}\left(\frac{1}{G_+(\frac{M}{u_* v};\alpha)} - \frac{1}{G_-(\frac{M}{u_* v};\alpha)}\right) dv \right.$$
$$\left. + \int_0^{\bar{u}} \frac{1}{u^2}\left(\frac{1}{G_+(\frac{M}{u};\alpha)} - \frac{1}{G_-(\frac{M}{u};\alpha)}\right) du + \int_0^{u_\pi} \frac{du}{u^2 G_-(\frac{M}{u};\alpha)} - \int_0^{u_f} \frac{du}{u^2 G_+(\frac{M}{u};\alpha)}\right) . \quad (4.113)$$

It turns out that if one makes this same change of variables in the second and third integrals above, then when differentiating the integrand with respect to α, one obtains an integrand with a $(u_* v - u_c)^{-1}$ divergence at $v = u_c/u_*$. However, if one retains the integration variable u, this difficulty does not arise. This is the technical reason for splitting the integrals that we alluded to earlier. The fact that the choice of integration variables can affect the ability to calculate $T'_\pi(\alpha)$ may seem surprising. It is a consequence of the fact that if one does not choose integration variables carefully, then $T'_\pi(\alpha)$ is only defined in the sense of having two diverging terms cancelling. This is reminiscent of an improper integral whose value is only defined when one considers the Cauchy principal value of this integral.

An alternative method to the one outlined in (4.109)-(4.113) would be to replace the upper limits of integration in the two integrals in (4.108) with a family of functions $\hat{u}(\alpha;\delta)$, such that $\hat{u}(\alpha;0) = u_*(\alpha)$ and $\hat{u}'(\alpha;0) = u'_*(\alpha)$, and then take the limit $\delta \to 0$ only after the differentiation of the integrals with respect to α. In general such a method is difficult to implement numerically, since the "principal value" one obtains when taking $\delta \to 0$ depends on the cancellation of two diverging terms, one of which is an integral that cannot be evaluated analytically. This leads to the subtraction of two very large numbers and a loss of significance in the numerical calculations. However, by sheer luck it turns out that one can manipulate the expressions involved and analytically cancel the two diverging quantities, so that the limit $\delta \to 0$ can be implemented numerically.

As an equivalent method to the one described in the previous paragraph, one can take the limit $\bar{u} \to u_*$ once the expression for $T'_\pi(\alpha)$ is obtained. This is what will be done below.

The final steps before differentiating (4.113) is to find an explicit expression for $G_{\pm}(\frac{M}{u},\alpha)$ in terms of the integration variable u, and to use (4.101) in order to eliminate the parameter α in favor of the turning point u_*, as was done in the previous section. Recall from (4.14) that $G_{\pm}(r;\alpha)$ is given by

$$G_{\pm}(r;\alpha) = \frac{-f(r)\sqrt{H^2 - V_{eff}(r)}}{\pm H\, f(r_e)^{\frac{1}{2}} \cos\alpha + (f(r_e) - f(r))^{\frac{1}{2}} \operatorname{sgn}(r - r_e)\sqrt{H^2 - V_{eff}(r)}} \ , \tag{4.114}$$

where $V_{eff}(r;\alpha)$ is given by (4.5). Using (4.101) and (4.107), we can replace α and r in $V_{eff}(r;\alpha)$ in favor of u_* and u to give

$$V_{eff}\left(\frac{M}{u};\alpha(u_*)\right) = \frac{\varepsilon^2 u^2}{u_*^2(1-2u_*)M^2}\,(1-2u) \ , \tag{4.115}$$

where we have written $\alpha = \alpha(u^*)$ to emphasize the dependence on u_*. Also using (4.107), we have

$$\cos\alpha = -\sqrt{1 - \frac{\varepsilon^2}{u_*^2(1-2u_*)}} \ , \tag{4.116}$$

where we have used $\cos\alpha < 0$, which is a necessary condition for generators with a turning point (discussed in more detail in section 4.4.2). Let us define $F_{\pm}(u,u_*)$ as follows:

$$F_{\pm}(u,u_*) = \frac{1}{G_{\pm}\left(\frac{M}{u},\alpha(u_*)\right)} \ . \tag{4.117}$$

Now we substitute (4.107), (4.115) and (4.116) into (4.114), and substitute the resulting expression for $G_{\pm}(\frac{M}{u};\alpha(u_*))$ into the above to obtain

$$F_{\pm}(u,u_*) = \frac{1}{1-2u-\frac{\varepsilon^2}{u^2}}\left(\pm\frac{(1-3\varepsilon^{\frac{2}{3}})^{\frac{1}{2}}\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}}{(u_*^2(1-2u_*)-u^2(1-2u))^{\frac{1}{2}}} - \left(2u+\frac{\varepsilon^2}{u^2}-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}}\operatorname{sgn}(u_e-u)\right) \ , \tag{4.118}$$

where we have used the result (2.6) for r_e, the definitions (2.5) and (2.13) for $f(r)$ and ε, respectively, and we have also defined $u_e = r_e/M$. Written in terms of $F_{\pm}(u,u_*)$, (4.113) becomes

$$T_f - T_\pi(\alpha) = M\left(\int_{\bar{u}/u_*}^{1}\frac{H(u_*v,u_*)}{u_*v^2}dv + \int_0^{\bar{u}}\frac{H(u,u_*)}{u^2}du + \int_0^{u_\pi}\frac{F_-(u,u_*)}{u^2}du - \int_0^{u_f}\frac{F_+(u,u_*)}{u^2}du\right) \ , \tag{4.119}$$

where it is understood that $u_* = u_*(\alpha)$ and $u_\pi = u_\pi(\alpha)$, and we have defined $H(u,u_*)$ as

$$H(u,u_*) = F_+(u,u_*) - F_-(u,u_*) \ . \tag{4.120}$$

Using (4.118), we have the following explicit formula for $H(u,u_*)$:

$$H(u,u_*) = \frac{2(1-3\varepsilon^{\frac{2}{3}})^{\frac{1}{2}}\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}}{\left(1-2u-\frac{\varepsilon^2}{u^2}\right)\left(u_*^2(1-2u_*)-u^2(1-2u)\right)^{\frac{1}{2}}} \ . \tag{4.121}$$

We are now in a position to differentiate (4.119) and obtain $T_\pi'(\alpha)$. Recalling that T_f and $\bar{u}$ have been chosen to be independent of α, we have

$$\begin{aligned}\frac{dT_\pi}{d\alpha} = -M\Bigg(&\int_{\bar{u}/u_*}^{1}\frac{1}{v^2}\frac{d}{du_*}\left(\frac{H(u_*v,u_*)}{u_*}\right)dv + \frac{H(\bar{u},u_*)}{u_*\bar{u}} + \int_0^{\bar{u}}\frac{1}{u^2}\frac{\partial H}{\partial u_*}du \\ &+\frac{F_-(u_\pi,u_*)}{u_\pi^2}\frac{du_\pi}{du_*} + \int_0^{u_\pi}\frac{1}{u^2}\frac{\partial F_-}{\partial u_*}du\Bigg)\frac{du_*}{d\alpha} - \frac{F_+(u_f,u_*)}{u_f^2}\frac{\partial u_f}{\partial\alpha} - \frac{du_*}{d\alpha}\int_0^{u_f}\frac{1}{u^2}\frac{\partial F_+}{\partial u_*}du \ ,\end{aligned} \tag{4.122}$$

where we have utilized the chain rule in order to differentiate the term in brackets with respect to u^*. Let us define the expressions E_1, E_2 and E_3 as follows:

$$\begin{aligned} E_1(\bar{u}) &= \int_{\bar{u}/u_*}^{1}\frac{1}{v^2}\frac{d}{du_*}\left(\frac{H(u_*v,u_*)}{u_*}\right)dv \ , \\ E_2(\bar{u}) &= \frac{H(\bar{u},u_*)}{u_*\bar{u}} + \int_0^{\bar{u}}\frac{1}{u^2}\frac{\partial H}{\partial u_*}du \ , \\ E_3(u_f) &= \frac{F_+(u_f,u_*)}{u_f^2}\frac{\partial u_f}{\partial\alpha} + \frac{du_*}{d\alpha}\int_0^{u_f}\frac{1}{u^2}\frac{\partial F_+}{\partial u_*}du \ . \end{aligned} \tag{4.123}$$

In the next step we show that E_3 vanishes in the limit $T_f \to \infty$. We then consider simplifications of $E_1(\bar{u})$ and $E_2(\bar{u})$ as $\bar{u} \to u_*^-$.

Second step: Simplification of $T'_\pi(\alpha)$ as $T_f \to \infty$

Since the constant T_f was arbitrary, we can take the limit $T_f \to \infty$, which leads to $u_f \to 0$. Let us show that taking this limit eliminates the last two terms in (4.122). Consider the last term first. From (4.118), we can compute:

$$\frac{\partial F_\pm}{\partial u_*} = \mp \frac{\left(1-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}} u_*(1-3u_*)u^2}{(u_*^2(1-2u_*)-\varepsilon^2)^{\frac{1}{2}}\left(u_*^2(1-2u_*)-u^2(1-2u)\right)^{\frac{3}{2}}} \ . \tag{4.124}$$

The limiting behavior of the above as $u \to 0$ is

$$\frac{\partial F_\pm}{\partial u_*} \sim u^2 \ .$$

Hence it follows that

$$\lim_{u_f \to 0} \int_0^{u_f} \frac{1}{u^2} \frac{\partial F_+}{\partial u_*} du = 0 \ , \tag{4.125}$$

so that the last term in (4.122) vanishes in the limit that $T_f \to \infty$. Note that multiplication of the above by $du_*/d\alpha$ does not affect this conclusion, since differentiating (4.101) and isolating $du_*/d\alpha$, we have

$$\frac{du_*}{d\alpha} = \frac{\varepsilon^2 \cos\alpha}{u_*(3u_*-1)\sin^3\alpha} \ ,$$

from which it follows that

$$0 < \frac{du_*}{d\alpha} < \infty \tag{4.126}$$

since $0 < u_* < 1/3$ and $\pi/2 < \alpha_p < \pi$. Next consider the second last term in (4.122). The behavior of $\partial u_f/\partial\alpha$ as $T_f \to \infty$ can be deduced using the approximation to $u(T;\alpha)$ as $T \to \infty$ that was derived in section 4.4.3. Using (4.68) and (4.76), in the limit that $T \to \infty$, we have

$$\frac{\partial r}{\partial \alpha} = \mathscr{O}(1) \ .$$

Using the above and the definition of u in (4.107), we get

$$\frac{\partial u}{\partial \alpha} \sim u^2 \tag{4.127}$$

for $u \to 0$. From (4.118), we have

$$F(u, u_*) \sim u \ .$$

Applying the above and (4.127) to $u_f = u(T_f; \alpha)$, we get

$$\frac{F_+(u_f, u_*)}{u_f^2} \frac{\partial u_f}{\partial \alpha} \sim u_f \ ,$$

so that the second last term in (4.122) vanishes as $T_f \to \infty$. That is

$$\lim_{u_f \to 0} \frac{F_+(u_f, u_*)}{u_f^2} \frac{\partial u_f}{\partial \alpha} = 0 \ . \tag{4.128}$$

Third step: Simplification of $T_\pi'(\alpha)$ as $\bar{u} \to u_*$

First we show that the first term in (4.122) vanishes in the limit $\bar{u} \to u_*$. With some work, one can compute the integrand to be

$$\begin{aligned}\left[\frac{1}{v^2}\frac{d}{du_*}\left(\frac{H(u_* v, u_*)}{u_*}\right)\right]_{v=u/u_*} &= \frac{2\left(1-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}}}{u_*\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}(u_*-u)^{\frac{1}{2}}} \\ &\times \left(\frac{u_*^2u^2-\varepsilon^2(u_*^2+u_*u+u^2)}{(u_*+u-2(u_*^2+u_*u+u^2))^{\frac{3}{2}}(u^2-2u^3-\varepsilon^2)}\right. \\ &\left. + \frac{u_*^2(1-2u_*)(4u_*-1)u^2-\varepsilon^2\left(u_*^2(1-2u_*)+2(3u-1)u^2\right)}{(u_*+u-2(u_*^2+u_*u+u^2))^{\frac{1}{2}}(u^2-2u^3-\varepsilon^2)^2}\right) \ .\end{aligned}$$

As one can verify, both the numerator and denominator of the term in brackets above are nonvanishing as $\bar{u} \to u_*$, so that

$$\left[\frac{1}{u_* v^2}\frac{d}{du_*}\left(\frac{H(u_* v, u_*)}{u_*^2}\right)\right]_{v=u/u_*} \sim \frac{1}{(u_*-u)^{\frac{1}{2}}} \ .$$

Thus we have

$$\lim_{\bar{u}\to u_*} \frac{1}{u_*} \int_{\bar{u}/u_*}^{1} \frac{1}{v^2} \frac{d}{du_*} \left(\frac{H(u_* v, u_*)}{u_*^2} \right) dv = 0 \ . \tag{4.129}$$

Next we show that although the second and third terms in (4.119) both diverge as $\bar{u} \to u_*$, they can be combined into a single term such that the limit $\bar{u} \to u_*$ does not contain any diverging terms. The expression we are interested was defined as E_2 in (4.123). Using (4.120) and (4.124), the explicit expression for $E_2(\bar{u})$ is

$$\begin{aligned} E_2(\bar{u}) = 2\left(1-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}} \Bigg(& \frac{\bar{u}\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}}{u_*\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)\left(u_*^2(1-2u_*)-\bar{u}^2(1-2\bar{u})\right)^{\frac{1}{2}}} \\ & - \frac{u_*(1-3u_*)}{\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}} \int_0^{\bar{u}} \frac{du}{\left(u_*^2(1-2u_*)-u^2(1-2u)\right)^{\frac{3}{2}}} \Bigg) \ . \end{aligned} \tag{4.130}$$

As we can see, both terms above diverge as $\bar{u} \to u_*^-$. However, it is possible to manipulate the expressions such that the divergences can be eliminated, even if one cannot evaluate the integral analytically. This is accomplished by writing the integral above as a sum of a convergent integral and a diverging integral that can be evaluated analytically. This is done as follows:

$$\begin{aligned} \int_0^{\bar{u}} \frac{du}{\left(u_*^2(1-2u_*)-u^2(1-2u)\right)^{\frac{3}{2}}} &= \int_0^{\bar{u}} \frac{du}{\left(u_*+u-2(u_*^2+u_*u+u^2)\right)^{\frac{3}{2}}(u_*-u)^{\frac{3}{2}}} \\ &\quad - \int_0^{\bar{u}} \frac{du}{\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}(u_*-u)^{\frac{3}{2}}} + \int_0^{\bar{u}} \frac{du}{\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}(u_*-u)^{\frac{3}{2}}} \\ &= \int_0^{\bar{u}} \frac{\left(u_*+u-2(u_*^2+u_*u+u^2)\right)^{-\frac{3}{2}} - \left(2u_*(1-3u_*)\right)^{-\frac{3}{2}}}{(u_*-u)^{\frac{3}{2}}} du + \int_0^{\bar{u}} \frac{du}{\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}(u_*-u)^{\frac{3}{2}}} \ . \end{aligned} \tag{4.131}$$

The first term on the right hand side is a convergent integral, and the second integral is diverging but can be evaluated analytically. Consider the first integral. Define $A(u,u_*)$ and A_* as follows:

$$A(u) \equiv u_* + u - 2(u_*^2 + u_* u + u^2), \tag{4.132}$$

$$A_* \equiv A(u_*) = 2u_*(1-3u_*) \ . \tag{4.133}$$

We can factor the numerator of the first integral in (4.131):

$$\left(u_* + u - 2(u_*^2 + u_* u + u^2)\right)^{-\frac{3}{2}} - \left(2u_*(1-3u_*)\right)^{-\frac{3}{2}} = A^{-\frac{3}{2}} - A_*^{-\frac{3}{2}}$$
$$= \frac{(A_* - A)\left(A^{-2} + A^{-1}A_*^{-1} + A_*^{-2}\right)}{AA_*\left(A^{-\frac{3}{2}} + A_*^{-\frac{3}{2}}\right)} . \quad (4.134)$$

The term $A_* - A$ above is a quadratic in u, and can be also factored:

$$\begin{aligned} A_* - A &= 2u^2 + (2u_* - 1)u + u_* - 4u_*^2 \\ &= (u - u_*)(2u - 1 + 4u_*) . \end{aligned}$$

Using the above in (4.134), the first integral in (4.131) becomes

$$\int_0^{\bar{u}} \frac{\left(u_* + u - 2(u_*^2 + u_* u + u^2)\right)^{-\frac{3}{2}} - \left(2u_*(1-3u_*)\right)^{-\frac{3}{2}}}{(u_* - u)^{\frac{3}{2}}} du$$
$$= \int_0^{\bar{u}} \frac{(2u - 1 + 4u_*)\left(A^{-2} + A^{-1}A_*^{-1} + A_*^{-2}\right)}{AA_*\left(A^{-\frac{3}{2}} + A_*^{-\frac{3}{2}}\right)(u_* - u)^{\frac{1}{2}}} du . \quad (4.135)$$

Since $r^* > 3M$ (see figure 4.6), we have $0 < u_* < 1/3$ and thus $A(u) > 0$ for $0 < u \leq u_*$. This ensures that the integrand above has the following behavior as $u \to u_*^-$:

$$\frac{(2u - 1 + 4u_*)\left(A^{-2} + A^{-1}A_*^{-1} + A_*^{-2}\right)}{AA_*\left(A^{-\frac{3}{2}} + A_*^{-\frac{3}{2}}\right)(u_* - u)^{\frac{1}{2}}} \sim \frac{1}{(u_* - u)^{\frac{1}{2}}} ,$$

so that (4.135) is a convergent integral as $\bar{u} \to u_*^-$. The second integral in (4.131) is divergent as $\bar{u} \to u_*^-$, but can be evaluated analytically:

$$\int_0^{\bar{u}} \frac{du}{\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}(u_* - u)^{\frac{3}{2}}} = \frac{2}{\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}}\left(\frac{1}{(u_* - \bar{u})^{\frac{1}{2}}} - \frac{1}{u_*^{\frac{1}{2}}}\right) .$$

Using the above and (4.135), we can now write (4.131) as the sum of three terms, the first two of which are convergent as $\bar{u} \to u_*$:

$$\int_0^{\bar{u}} \frac{du}{\left(u_*^2(1-2u_*) - u^2(1-2u)\right)^{\frac{3}{2}}} = \int_0^{\bar{u}} \frac{(2u-1+4u_*)\left(A^{-2}+A^{-1}A_*^{-1}+A_*^{-2}\right)}{AA_*\left(A^{-\frac{3}{2}}+A_*^{-\frac{3}{2}}\right)(u_*-u)^{\frac{1}{2}}} du$$
$$- \frac{2}{u_*^{\frac{1}{2}}\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}} + \frac{2}{\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}(u_*-\bar{u})^{\frac{1}{2}}} \,.$$

Substituting the above into (4.130), we get the following expression:

$$E_2(\bar{u}) = 2\left(1-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}} \left(\frac{\bar{u}\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}}{u_*\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)\left(u_*^2(1-2u_*)-\bar{u}^2(1-2\bar{u})\right)^{\frac{1}{2}}} \right.$$
$$\left. - \frac{2u_*(1-3u_*)}{\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}(u_*-\bar{u})^{\frac{1}{2}}} \right) - 2\left(1-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}} \frac{u_*(1-3u_*)}{\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}}$$
$$\times \left(\int_0^{\bar{u}} \frac{(2u-1+4u_*)\left(A^{-2}+A^{-1}A_*^{-1}+A_*^{-2}\right)}{AA_*\left(A^{-\frac{3}{2}}+A_*^{-\frac{3}{2}}\right)(u_*-u)^{\frac{1}{2}}} du - \frac{2}{u_*^{\frac{1}{2}}\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}} \right) \,. \quad (4.136)$$

The first term in large brackets is a subtraction of two divergent terms as $\bar{u} \to u_*^-$. Let us now show that this term in fact vanishes as $\bar{u} \to u_*^-$. First we rewrite it as

$$\frac{\bar{u}\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}}{u_*\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)\left(u_*^2(1-2u_*)-\bar{u}^2(1-2\bar{u})\right)^{\frac{1}{2}}} - \frac{2u_*(1-3u_*)}{\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}(u_*-\bar{u})^{\frac{1}{2}}}$$
$$= \frac{\left(u_*^2(1-2u_*)-\varepsilon^2\right)^{\frac{1}{2}}}{u_*} \left(\frac{\bar{u}}{\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)A(\bar{u})^{\frac{1}{2}}} - \frac{u_*}{\left(u_*^2(1-2u_*)-\varepsilon^2\right)A_*^{\frac{1}{2}}} \right) \frac{1}{(u_*-\bar{u})^{\frac{1}{2}}} \,. \quad (4.137)$$

We can see we are clearly dealing with a $0/0$ limit, and so we apply L'Hopital's rule:

$$\lim_{\bar{u}\to u_*^-} \left(\frac{\bar{u}}{\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)A(\bar{u})^{\frac{1}{2}}} - \frac{u_*}{\left(u_*^2(1-2u_*)-\varepsilon^2\right)A_*^{\frac{1}{2}}} \right) \frac{1}{(u_*-\bar{u})^{\frac{1}{2}}}$$
$$= \lim_{\bar{u}\to u_*^-} -2(u_*-\bar{u})^{\frac{1}{2}} \frac{d}{d\bar{u}} \left(\frac{\bar{u}}{\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)A(\bar{u})^{\frac{1}{2}}} \right) \,. \quad (4.138)$$

We can compute:

$$\frac{d}{d\bar{u}}\left(\frac{\bar{u}}{(\bar{u}^2(1-2\bar{u})-\varepsilon^2)A(\bar{u})^{\frac{1}{2}}}\right)$$
$$= \frac{\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)A(\bar{u})-\bar{u}\left((2\bar{u}-6\bar{u}^2)A(\bar{u})+\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)\frac{1}{2}A'(\bar{u})\right)}{\left(\bar{u}^2(1-2\bar{u})-\varepsilon^2\right)^2 A(\bar{u})^{\frac{3}{2}}} .$$

From (4.132)-(4.133), we see that $A(\bar{u})$ and $A'(\bar{u})$ are bounded as $\bar{u}\to u_*^-$. Furthermore, the denominator is non-vanishing as $\bar{u}\to u_*^-$ since (4.101) implies that $u_*^2(1-2u_*)>\varepsilon^2$ and $A(u)>0$ for $0<u\leq u_*$ as discussed previously. Therefore we have

$$\left|\lim_{\bar{u}\to u_*^-}\frac{d}{d\bar{u}}\left(\frac{\bar{u}}{(\bar{u}^2(1-2\bar{u})-\varepsilon^2)A(\bar{u})^{\frac{1}{2}}}\right)\right|<\infty ,$$

and so

$$\lim_{\bar{u}\to u_*^-}-2(u_*-\bar{u})^{\frac{1}{2}}\frac{d}{d\bar{u}}\left(\frac{\bar{u}}{(\bar{u}^2(1-2\bar{u})-\varepsilon^2)A(\bar{u})^{\frac{1}{2}}}\right)=0$$

as expected. Combining the above with (4.138), we have that the expression in (4.137) vanishes as $\bar{u}\to u_*^-$, so that (4.136) finally becomes

$$\lim_{\bar{u}\to u_*^-}E_2(\bar{u})=2\left(1-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}}\frac{u_*(1-3u_*)}{(u_*^2(1-2u_*)-\varepsilon^2)^{\frac{1}{2}}}$$
$$\times\left(\frac{2}{u_*^{\frac{1}{2}}\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}}-\int_0^{u_*}\frac{(2u-1+4u_*)\left(A^{-2}+A^{-1}A_*^{-1}+A_*^{-2}\right)}{AA_*\left(A^{-\frac{3}{2}}+A_*^{-\frac{3}{2}}\right)(u_*-u)^{\frac{1}{2}}}du\right) . \quad (4.139)$$

Fourth step: The equation $T'(\alpha)=0$

Taking the limits $T_f\to\infty$ (i.e. $u_f\to 0$) and $\bar{u}\to u_*^-$ and using (4.125), (4.128) and (4.129) in (4.122), we get

$$\frac{dT_\pi}{d\alpha}=-M\left(\lim_{\bar{u}\to u_*^-}E_2(\bar{u})+\frac{F_-(u_\pi,u_*)}{u_\pi^2}\frac{du_\pi}{du_*}+\int_0^{u_\pi}\frac{1}{u^2}\frac{\partial F_-}{\partial u_*}du\right)\frac{du_*}{d\alpha} ,$$

where $\lim_{\bar{u}\to u_*} E_2$ is given by (4.139). As shown in the steps leading up to (4.126), we have $0 < u'_*(\alpha) < \infty$, so that the equation $T'_\pi(\alpha) = 0$ implies that the term in brackets above vanishes:

$$\lim_{\bar{u}\to u_*^-} E_2(\bar{u}) + \frac{F_-(u_\pi, u_*)}{u_\pi^2}\frac{du_\pi}{du_*} + \int_0^{u_\pi} \frac{1}{u^2}\frac{\partial F_-}{\partial u_*} du = 0 \ . \tag{4.140}$$

The only expression that remains to be computed in the above equation is $u'_\pi(u_*)$. It is obtained by differentiating (4.102) with respect to u_*. This gives

$$\begin{aligned} & 2\frac{dI}{du_*} - \frac{\partial I_1}{\partial u_\pi}\frac{du_\pi}{du_*} - \frac{\partial I_1}{\partial u_*} = 0 \\ \Rightarrow \quad & \frac{du_\pi}{du_*} = \frac{2\frac{dI}{du_*} - \frac{\partial I_1}{\partial u_*}}{\frac{\partial I_1}{\partial u_\pi}} \ . \end{aligned} \tag{4.141}$$

From (4.103)-(4.104), we can compute the following:

$$\begin{aligned} \frac{dI}{du_*} &= \int_0^1 \frac{1-v^3}{(1-2u_*-v^2(1-2u_*v))^{\frac{3}{2}}} dv \ , \\ \frac{\partial I_1}{\partial u_*} &= u_*(3u_*-1)\int_0^{u_\pi} \frac{du}{(u_*^2(1-2u_*)-u^2(1-2u))^{\frac{3}{2}}} \ , \end{aligned} \tag{4.142}$$

$$\frac{\partial I_1}{\partial u_\pi} = \frac{1}{(u_*^2(1-2u_*)-u_\pi^2(1-2u_\pi))^{\frac{1}{2}}} \ . \tag{4.143}$$

Notice how the integration variable $v = u/u_*$ allows us to avoid dealing with the problem of evaluating the integrand of $I(u_*)$ at $u = u_*$ when computing $I'(u_*)$. The first integral above can be simplified by factoring both numerator and denominator. Doing this and switching back to the integration variable $u = u_* v$, we get

$$\frac{dI}{du_*} = \int_0^{u_*} \frac{u_*^2 + u_* u + u^2}{A(u)^{\frac{3}{2}}(u_*-u)^{\frac{1}{2}}} du \ , \tag{4.144}$$

where $A(u)$ was defined in (4.133). We can now put everything together into a single equation relating u_π and u_*. Substituting (4.142), (4.143) and (4.144) into (4.141), and then substituting

(4.141) and (4.139) into (4.140), we arrive at the following equation:

$$\int_0^{u_\pi} \frac{1}{u^2}\frac{\partial F_-}{\partial u_*}du + 2\left(1-3\varepsilon^{\frac{2}{3}}\right)^{\frac{1}{2}} \frac{u_*(1-3u_*)}{(u_*^2(1-2u_*)-\varepsilon^2)^{\frac{1}{2}}}$$
$$\times\left(\frac{2}{u_*^{\frac{1}{2}}\left(2u_*(1-3u_*)\right)^{\frac{3}{2}}} - \int_0^{u_*} \frac{(2u-1+4u_*)\left(A^{-2}+A^{-1}A_*^{-1}+A_*^{-2}\right)}{AA_*\left(A^{-\frac{3}{2}}+A_*^{-\frac{3}{2}}\right)(u_*-u)^{\frac{1}{2}}}du\right)$$
$$+\frac{F_-(u_\pi,u_*;\varepsilon)}{u_\pi^2}\left(u_*^2(1-2u_*)-u_\pi^2(1-2u_\pi)\right)^{\frac{1}{2}}$$
$$\times\left(2\int_0^{u_*}\frac{u_*^2+u_*u+u^2}{A(u)^{\frac{3}{2}}(u_*-u)^{\frac{1}{2}}}du - u_*(3u_*-1)\int_0^{u_\pi}\frac{du}{(u_*^2(1-2u_*)-u^2(1-2u))^{\frac{3}{2}}}\right)=0\,, \quad (4.145)$$

where $A(u)$, A_*, $F_-(u,u_*;\varepsilon)$ and $\partial F_-/\partial u_*$ are given by (4.132), (4.133), (4.118) and (4.124), respectively. The above is the second equation relating u_π and u_*. When coupled to (4.102), these two equations can be solved simultaneously to give numerical values for u_p and u_*, provided one chooses a specific value for the parameter ε. Once u_p is known, the spacetime coordinates (T_p, r_p) of the merger point are given by (4.108) and (4.107) respectively. Due to the Killing symmetry of the spacetime, the value of T_p is arbitrary since it depends on choosing a reference time (for example, one could choose $T=0$ to be the time when the area of the horizons is at the midpoint between the initial and final areas). For this reason, we will only be interested in calculating the value of r_p. In the next section we solve (4.102) and (4.145) for u_p in the limit that $\varepsilon \to 0$. This is the limit in which the Schwarzschild radius of the black hole is infinitesimally smaller than the Hubble radius associated with the cosmological constant.

4.5.4 Merger point in the small black hole mass limit

As discussed in section (2.7), for any astrophysically realistic black holes, the value of the parameter ε is exceedingly small. For this reason, taking the limit $\varepsilon \to 0$ gives an excellent approximation for the location and structure of the horizon merger point in an astrophysically realistic situation. Note that the limit $\varepsilon \to 0$ is nontrivial, in the sense that u_p converges to a finite value in this limit.

Converting this to a value of r_p using $r_p = u_p M$, we have that in the limit $\varepsilon \to 0$, the merger point is located at a spacetime point with coordinates such that $r \propto M$, where r is the radial coordinate of Schwarzschild coordinates. We seek to find the numerical value of the proportionality constant in the relation $r_p \propto M$. More precisely, we expect an asymptotic series for $u_p(\varepsilon)$ of the form

$$u_p(\varepsilon) = u_p(0) + \mathscr{O}(\varepsilon) \ .$$

The corresponding asymptotic series for $r_p(\varepsilon)$ has the form

$$r_p(\varepsilon) = M \left[\frac{1}{u_p(0)} + \mathscr{O}(\varepsilon) \right] \ . \tag{4.146}$$

Note that since $\varepsilon = HM$, the limit $\varepsilon \to 0$ can be thought of as either the result of taking $M \to 0$ or $H \to 0$. Here we think of taking $H \to 0$, so that the Hubble radius becomes infinitely large, while the black hole mass remain constant. In this way the above expansion yields a nontrivial result for $r_p(\varepsilon)$.

We solve equations (4.102) and (4.145) numerically for u_* and u_p using MAPLE 14, in the limit that $\varepsilon \to 0$. We employ the following strategy. First we plot the functions $u_\pi(u_*)$ which result from solving either equation (4.102) or (4.145) independently. The graphs of $u_\pi(u_*)$ for these two equations are constructed by choosing a set of values for u_*, and for each such value, solving the resulting one variable equation for u_π using MAPLE's fsolve procedure. Looking for a point where these curves intersect, we can determine graphically that we must have $0.2165 < u_* < 0.2175$ and $0.193 < u_p < 0.195$. Using these bounds, we then solve the coupled equations using the fsolve procedure. Obtaining bounds for u_* and u_p is an essential first step since they are used as optional inputs to constrain the search for a solution to the coupled equations. We have found that without them, the fsolve procedure will not find a solution to the coupled equations. The numerical result of solving the coupled equations is

$$\begin{aligned} u_* &= 0.2171541500..., \\ u_p &= 0.1945829820.... \ . \end{aligned} \tag{4.147}$$

Substituting u_p above in (4.146) and taking $\varepsilon = 0$, we obtain

$$\boxed{r_p \approx 5.25M} \,. \tag{4.148}$$

The above is the final result culminating from all the steps in this section (i.e. section 4.5). Using (4.147) for the value of u_* in (4.101), we can also obtain a final result for α_p, which is the value of the parameter α for the unique generator that goes through the merger point. Thinking of $\alpha_p = \alpha_p(\varepsilon)$, we expect from (4.101) that we have the following expansion:

$$\alpha_p(\varepsilon) = \varepsilon \left[\frac{1}{u_*(0)\,(1-2u_*(0))^{\frac{1}{2}}} + \mathscr{O}(\varepsilon) \right] \,.$$

Substituting the numerical value of u_* from (4.147) into the above, we obtain

$$\frac{\alpha_p}{\varepsilon} \approx 6.12 \,, \tag{4.149}$$

where higher order terms in ε have been neglected. Note that unlike (4.146), in the limit $\varepsilon \to 0$ one obtains the trivial result $\alpha_p = 0$. However, one can easily define a parameter which does not vanish in the limit $\varepsilon \to 0$. For example, we define the parameter b as

$$b \equiv \frac{1}{H} \sin\alpha \,.$$

Substituting (4.149) into the above, we obtain

$$b_p \approx 6.12M \,.$$

This result is similar to (4.148), in that b_p does not vanish as $\varepsilon \to 0$. The interpretation of the parameter b is that it is analogous to a kind of impact parameter. It terms of the late time behavior of the null generators, it is the y-coordinate of such a generator as the horizon settles down to its final shape (the last frame of figure 4.4).

Although (4.148) specifies the spacetime coordinates of the merger point using Schwarzschild coordinates, it is clear that the result is coordinate independent, in the sense that the location of the

spacetime event where the horizon merger occurs is independent of choice of coordinates. Furthermore, although we have specified the location of the merger point using a specific coordinate system, the result can be transformed to another coordinate by utilitzing the geometric interpretation of the Schwarzschild coordinate r, i.e., by recognizing that the spheres associated with the spherical symmetry of the spacetime have area $4\pi r^2$.

Chapter 5

Area of Merging Black Hole and Cosmological Horizons

One of the key quantities that one can calculate regarding a black hole event horizon is the horizon area. The usefulness of horizon area stems from both astrophysical considerations and more theoretical questions. On the astrophysical side, the relationship between horizon area and mass can prove useful, for example, when calculating the final mass of the resulting black hole in numerical simulations of black hole mergers [1]. On the theoretical side, the pioneering work of Bekenstein [2] and Hawking, along with the extensive work that followed, has firmly linked black hole area as a measure of black hole entropy. This link between horizon area and entropy has also been considered in the context of cosmological horizons [11].

In this chapter we investigate questions related to the total surface area of the cosmological and black hole horizons. In section 5.1, we outline the method used to calculate the area of the horizons, including a description of the numerical methods used. In section 5.2, we use numerical methods to investigate the dependence of horizon area on time and on the small parameter $\varepsilon = HM$. These numerical calculations lead to three main results. The first concerns the area of the horizons in the limit of $\varepsilon \to 0$. The second result relates the time of maximal area increase to the time of merger. The third and final numerical result concerns the relative contribution of different horizon generators towards the overall area increase. Notice that since the area depends on the particular choice of time slicing, these results depend on the choice of time coordinate, although obviously

they are independent of the spacelike coordinates one chooses to use on the constant time slices. Although our results depend on the choice of time slicing, we will argue that the qualitative aspects of these results ought to hold in any coordinate system.

In section 5.3, we use analytical methods to further investigate two of the numerical results obtained in section 5.2. The purpose of these analytical results is both to confirm and corroborate the numerical results, as well as to extend them to beyond a specific choice of time coordinate. The first analytical result is a partial proof that the time of maximal horizon area increase precisely coincides with the time of merger where the horizons first touch. This result is coordinate independent in the sense that it will be shown to hold in a family of coordinate systems. The second analytical result is a proof that in the limit that $\varepsilon \to 0$, the area increase of the cosmological horizon due to the merger with a black hole can be attributed as being due in equal parts to two causes: the expansion of generators already on the horizon, and the joining of new generators with the horizon. This result is coordinate independent in the sense that there is a canonical time slicing that one can use to measure the area, and it is with this coordinate system that we calculate area.

5.1 Horizon area calculations

5.1.1 Horizon area formula

Recall that in the so-called LP coordinates used in this thesis, the induced metric on a spacelike hypersurface $T =$ constant is given by the following, (in spherical coordinates; see equation (3.39)):

$$ds^2 = \frac{1}{f(r_e)} dr^2 + r^2 \left(d\theta^2 + \sin^2\theta \, d\phi^2 \right) \ , \tag{5.1}$$

where we have eliminated the minus signs so that the metric has signature $(+,+,+)$. The above is part of a broader class of spacelike metrics:

$$ds^2 = F(r,\tau)dr^2 + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right) \ . \tag{5.2}$$

Given any coordinate system for SdS spacetime where two of the spacelike variables coincide with the θ and ϕ variables of Schwarzschild coordinates, the line element on a spacelike hypersurface $\tau =$ constant can always be written in the above form. Thus the above represents the spacelike hypersurface geometry for a broad class of coordinate systems for SdS spacetime. The motivation for considering the above broader class of line elements, as opposed to (5.1), is that our proof in section 5.3.1 will make use of the area formula for the broad class of line elements in (5.2).

To derive the formula for the area, let S_τ be the 2d surface formed by the intersection of the spacelike hypersurface $\tau =$ constant and the horizon (viewed as a null hypersurface in spacetime). Taking the further intersection of S_τ with the surface $\theta = \pi/2$, we obtain a curve such as the ones shown in figure 4.4. We can parametrize any one of these curves using an angle $\alpha \in [0, 2\pi)$, and specify the spatial coordinates of the curve as $(r(\alpha), \phi(\alpha), \pi/2)$. Given any one of these curves, spherical symmetry allows one to recover S_τ by simply rotating the curve about the axis $y = 0$, where $y = r\sin\theta$. Defining $\beta \in [0, \pi)$ as the angle corresponding to this rotation, we can use α and β as intrinsic coordinates on S_τ.

To obtain the area element dA on S_τ, we must first find the induced metric on the surface S_τ. It is clear that owing to the rotational symmetry, this metric will be independent of the coordinate β. Hence it suffices to find the induced metric along any curve $\beta =$ constant, since the induced metric everywhere else will be the same. Let us consider the induced metric on the curve $\beta = 0$ (i.e. the curve in the plane $\theta = \pi/2$). On this curve, the coordinate tangent vectors of the (α, β) coordinate system are

$$\begin{aligned} \frac{\partial}{\partial\alpha} &= \frac{dr}{d\alpha}\frac{\partial}{\partial r} + \frac{d\phi}{d\alpha}\frac{\partial}{\partial\phi} \ , \\ \frac{\partial}{\partial\beta} &= \sin\phi\frac{\partial}{\partial\theta} \ , \end{aligned} \tag{5.3}$$

where it is understood that $r = r(\alpha)$ and $\phi = \phi(\alpha)$ are functions of α only and describe the curve formed by the intersection of S_τ and $\theta = \pi/2$, as mentioned previously. Letting $\sigma_{\alpha\alpha}$, $\sigma_{\alpha\beta}$ and $\sigma_{\beta\beta}$ be the components of the induced metric on S_T, we have that

$$\begin{aligned}
\sigma_{\alpha\alpha} &= \langle \frac{\partial}{\partial\alpha}, \frac{\partial}{\partial\alpha} \rangle = \left(\frac{dr}{d\alpha}\right)^2 \langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \rangle + \left(\frac{d\phi}{d\alpha}\right)^2 \langle \frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi} \rangle = \left(\frac{dr}{d\alpha}\right)^2 g_{rr} + \left(\frac{d\phi}{d\alpha}\right)^2 g_{\phi\phi} \\
&= \left(\frac{dr}{d\alpha}\right)^2 F(r,\tau) + \left(\frac{d\phi}{d\alpha}\right)^2 r^2 \ , \\
\sigma_{\alpha\beta} &= \langle \frac{\partial}{\partial\alpha}, \frac{\partial}{\partial\beta} \rangle = 0 \ , \\
\sigma_{\beta\beta} &= \langle \frac{\partial}{\partial\beta}, \frac{\partial}{\partial\beta} \rangle = \sin^2\phi \langle \frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta} \rangle = r^2 \sin^2\phi \ ,
\end{aligned}$$

where we have used the coordinate tangent vectors from (5.3) and the metric components from (5.1), with $\theta = \pi/2$. The metric determinant associated with the above induced metric is

$$\sigma = \sigma_{\alpha\alpha}\sigma_{\beta\beta} - \sigma_{\alpha\beta}^2 = r^2 \sin^2\phi \left(\left(\frac{dr}{d\alpha}\right)^2 F(r,\tau) + \left(\frac{d\phi}{d\alpha}\right)^2 r^2 \right) \ ,$$

and hence the area element dA is

$$dA = \sqrt{\sigma}\, d\alpha d\beta = r \mid \sin\phi \mid \left(F(r,\tau) \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right)^{1/2} d\alpha d\beta \ . \tag{5.4}$$

We must integrate dA over that portion of S_τ which is the horizon, which either has two disconnected pieces (the black hole and cosmological horizons) or simply one piece (the cosmological horizon). In the former case we have

$$A_{\text{bh and ch}} = 2\pi \int_0^{\alpha_c} r \sin\phi \left(F(r,\tau) \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right)^{1/2} d\alpha \tag{5.5}$$

$$+ 2\pi \int_{\alpha_b}^{\pi} r \sin\phi \left(F(r,\tau) \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right)^{1/2} d\alpha \ , \tag{5.6}$$

where α_c and α_b are the critical values of α where new null generators can enter the horizon (i.e.

caustic points). When there is only a cosmological horizon, we have

$$A_{ch} = 2\pi \int_0^{\pi} r \sin\phi \left(F(r,\tau) \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right)^{1/2} d\alpha \ . \tag{5.7}$$

Notice that in the above we have integrated over $\alpha \in [0, 2\pi)$ and $\beta \in [0, \pi)$. The integral over β yields π and the integral over α was simplified to an integral over $\alpha \in [0, \pi)$ by introducing a factor of 2. Because we are integrating over $\alpha \in [0, \pi)$, we have $\sin\phi > 0$ and therefore have substituted $|\sin\phi| = \sin\phi$ in the expression for dA above.

5.1.2 Numerical approximations

For the numerical calculations in section 5.2, it suffices to restrict our attention to the case where the line element is given by (5.1). The integral (5.7) is then approximated using Simpson's method. Let $I(\alpha)$ be the integrand in the above integrals. That is:

$$I(\alpha) = 2\pi r \mid \sin\phi \mid \left(\frac{1}{f(r_e)} \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right)^{1/2} . \tag{5.8}$$

Suppose $r(\alpha)$ and $\phi(\alpha)$ are known at a sequence of evenly spaced grid points $(\alpha_0, \ldots, \alpha_n)$, with $\alpha_0 = 0$, $\alpha_n = \pi$ and n and even number. Then by Simpson's method we have

$$A_{ch} = \int_0^{\pi} I(\alpha) d\alpha \approx \frac{\Delta\alpha}{3} \sum_{k=0}^{\frac{n}{2}-1} \left(I(\alpha_{2k}) + 4I(\alpha_{2k+1}) + I(\alpha_{2k+2}) \right) \ , \tag{5.9}$$

where $\Delta\alpha$ is the spacing between neighboring values of α_i (i.e. $\Delta\alpha = \alpha_{i+1} - \alpha_i$). Computing the above requires knowledge of $r'(\alpha_i)$ and $\phi'(\alpha_i)$, both of which can be approximated using central differences:

$$\begin{aligned} \left.\frac{dr}{d\alpha}\right|_{\alpha_i} &\approx \frac{r(\alpha_{i+1}) - r(\alpha_{i-1})}{2\Delta\alpha} \ , \\ \left.\frac{d\phi}{d\alpha}\right|_{\alpha_i} &\approx \frac{\phi(\alpha_{i+1}) - \phi(\alpha_{i-1})}{2\Delta\alpha} \ . \end{aligned} \tag{5.10}$$

The endpoints $\alpha_0 = 0$ and $\alpha_n = \pi$ can be dealt with by exploiting $r(-\alpha) = r(\alpha)$ and $\phi(-\alpha) = -\phi(\alpha)$ so that the central differences become

$$\begin{aligned} \left.\frac{dr}{d\alpha}\right|_{\alpha_0} &= \left.\frac{dr}{d\alpha}\right|_{\alpha_n} = 0 \,, \\ \left.\frac{d\phi}{d\alpha}\right|_{\alpha_0} &\approx \frac{\phi(\alpha_1)}{\Delta\alpha} \,, \\ \left.\frac{d\phi}{d\alpha}\right|_{\alpha_n} &\approx \frac{\pi - \phi(\alpha_{n-1})}{\Delta\alpha} \,. \end{aligned} \tag{5.11}$$

Using (5.10)-(5.11) in (5.8)-(5.9), we can estimate (5.7) using only knowledge of $r(\alpha)$ and $\phi(\alpha)$ at the sequence of points $(\alpha_0, \ldots, \alpha_n)$. The method used for computing $r(\alpha)$ and $\phi(\alpha)$ was described in section 4.3.

To estimate the integrals in (5.5), let us first define α_{i_1} and α_{i_2} such that

$$\alpha_{i_1-1} < \alpha_c < \alpha_{i_1} < \ldots < \alpha_{i_2} < \alpha_b < \alpha_{i_2+1} \,.$$

We then separate each integral into two pieces:

$$\begin{aligned} \int_0^{\alpha_c} I(\alpha)d\alpha &= \int_0^{\alpha_{i_1-p}} + \int_{\alpha_{i_1-p}}^{\alpha_c} I(\alpha)d\alpha \,, \\ \int_{\alpha_b}^{\pi} I(\alpha)d\alpha &= \int_{\alpha_b}^{\alpha_{i_2+l}} + \int_{\alpha_{i_2+l}}^{\pi} I(\alpha)d\alpha \,, \end{aligned} \tag{5.12}$$

where $p = 1$ if i_1 is odd and $p = 2$ if i_1 is even. Similarly, $l = 1$ if i_2 is odd and $l = 2$ if i_2 is even. The reason for separating the above integrals into two pieces is that one of the pieces is most easily dealt with using Simpson's method, whereas the other piece can be most easily approximated using the trapezoid method. Using Simpson's method, we have

$$\begin{aligned} \int_0^{\alpha_{i_1-p}} I(\alpha)d\alpha &\approx \frac{\Delta\alpha}{3} \sum_{k=0}^{\frac{i_1-p}{2}-1} \left(I(\alpha_{2k}) + 4I(\alpha_{2k+1}) + I(\alpha_{2k+2})\right) \,, \\ \int_{\alpha_{i_2+l}}^{\pi} I(\alpha)d\alpha &\approx \frac{\Delta\alpha}{3} \sum_{k=\frac{i_2+l}{2}-1}^{\frac{n}{2}-1} \left(I(\alpha_{2k}) + 4I(\alpha_{2k+1}) + I(\alpha_{2k+2})\right) \,. \end{aligned} \tag{5.13}$$

Notice that $p \in \{1,2\}$ and $l \in \{1,2\}$ are always chosen so that $i_1 - p$ and $i_2 + l$ are even. We thus have a sum over an odd number of integers in the above, as required by Simpson's method. Using the trapezoid method, we approximate the remaining integrals in (5.12). This gives

$$\begin{aligned} \int_{\alpha_{i_1-p}}^{\alpha_c} I(\alpha)d\alpha &\approx (p-1)\frac{I(\alpha_{i_1-2})+I(\alpha_{i_1-1})}{2}\Delta\alpha + \frac{I(\alpha_{i_1-1})+I(\alpha_c)}{2}(\alpha_c - \alpha_{i_1-1}) \ , \quad (5.14) \\ \int_{\alpha_{i_2+l}}^{\pi} I(\alpha)d\alpha &\approx (l-1)\frac{I(\alpha_{i_2+1})+I(\alpha_{i_2+2})}{2}\Delta\alpha + \frac{I(\alpha_b)+I(\alpha_{i_2+1})}{2}(\alpha_{i_2+1} - \alpha_b) \ . \end{aligned}$$

If $p = 1$ in the above, the first term vanishes and we have are applying the trapezoid method using the two grid points $\{\alpha_{i_1-1}, \alpha_c\}$. By contrast, if $p = 2$ then we have two terms in the above and we are using the three grid points $\{\alpha_{i_1-2}, \alpha_{i_1-1}, \alpha_c\}$ in the trapezoid method. A similar distinction applies to the $l = 1$ and $l = 2$ cases. The above formulae require knowledge of α_b and α_c, which are not known in general. However, using the fact that

$$\phi(\alpha_b) = \phi(\alpha_c) = \pi$$

and knowledge of $\phi(\alpha)$ at the grid points α_{i_1-1} and α_c (or α_{i_2} and α_b), we can use a linear interpolant between these grid points and approximate α_c and α_b as

$$\begin{aligned} \alpha_c &= \alpha_{i_1-1} + \frac{\pi - \phi(\alpha_{i_1-1})}{\phi(\alpha_{i_1}) - \phi(\alpha_{i_1-1})}\Delta\alpha \ , \quad (5.15) \\ \alpha_b &= \alpha_{i_2} + \frac{\phi(\alpha_{i_2}) - \pi}{\phi(\alpha_{i_2}) - \phi(i_{i_2+1})}\Delta\alpha \ . \end{aligned}$$

Using (5.15) in (5.14) and approximating the integrals in (5.12) using (5.13)-(5.14), we arrive at an approximation of the integrals in (5.5), which requires only knowledge of $r(\alpha)$ and $\phi(\alpha)$ at the uniform grid of point $(\alpha_0, \ldots, \alpha_n)$. The method for computing $r(\alpha)$ and $\phi(\alpha)$ was outlined in section 4.3.

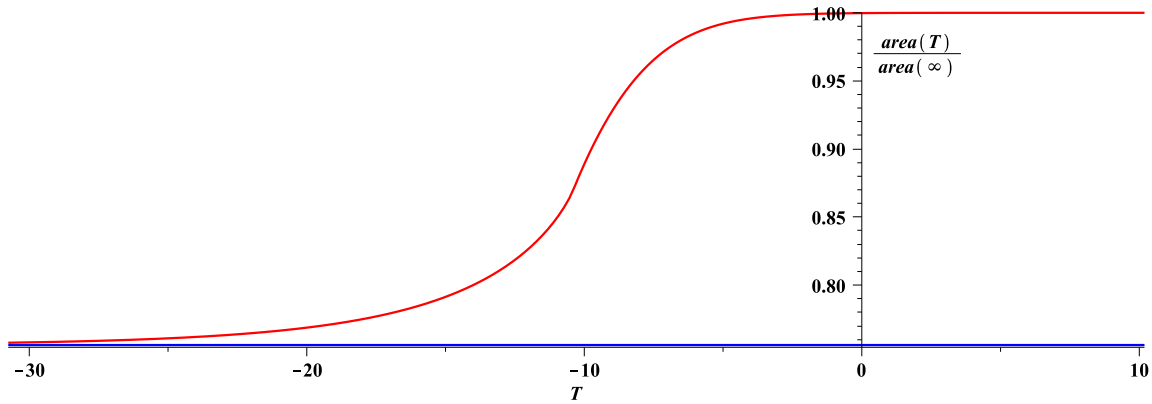


Figure 5.1: Total area of the horizons vs LP coordinate time T. The area has been normalized to the limiting value of the area as $T \to \infty$. The horizontal blue line is the limiting value as $T \to -\infty$. Parameter values used are: $M = 1$, $\Lambda = 1/18$, $n = 400$ and $\xi_0 = 10^{-4}$.

5.2 Numerical results

We begin by discussing the general features of the area vs time graph in section 5.2.1 below. We then focus on three aspects of the horizon area, as revealed by numerical computations. In section 5.2.2, we analyze the location of the inflection point in the area vs time graph, and present numerical evidence that it coincides precisely with the time of merger. In section 5.2.3, we analyze the horizon area in the limit $\varepsilon \to 0$, where recall that $\varepsilon = HM$. Specifically, we compute the horizon area at the time of merger for $\varepsilon \ll 1$. What we find is that for $\varepsilon \ll 1$, the area at the time of merger differs from the final area by an amount which is vanishingly small, with an $\mathscr{O}(\varepsilon)$ dependence on ε (this will be true when the area is measured in dimensionless units). Also, we analyze the influence of different horizon generators for $\varepsilon \ll 1$, and find that in this limit, the horizon area increase can be attributed as being due in equal parts to new generators joining the horizons, and existing generators expanding on the horizon.

5.2.1 Horizon area vs time graph: general features

Using the procedure outlined in section 4.3.3, we can compute the coordinates of the null generators of the horizon. We then use the procedure described in section 5.1.2 to compute the total area of the horizons numerically. The resulting graph is shown in figure 5.1.

The early time behavior of the area is as expected. That is, as $T \to -\infty$, the area approches the value

$$A(-\infty) = 4\pi \left(r_b^2 + r_c^2\right) \tag{5.16}$$

which is simply the sum of the areas of the spheres with radii $r = r_b$ and $r = r_c$. This agrees with the fact that at early times, the horizons are simply the concentric spheres at $r = r_b$ and $r = r_c$, as was discussed in section 2.5, and deduced in section 4.4.2.

As $T \to \infty$, the area approaches the value

$$A(\infty) = 4\pi \left(\frac{1}{H}\right)^2 . \tag{5.17}$$

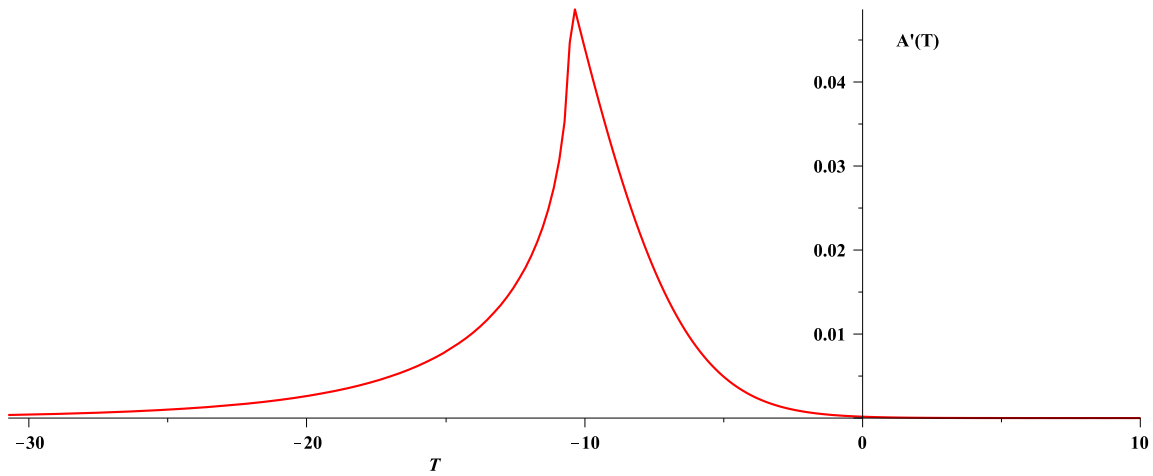
The above is what one would expect for the area of a spherical cosmological horizon in deSitter spacetime with a Hubble radius $1/H$. This agrees with the fact that in the limit $r \to \infty$, SdS spacetime is well approximated by deSitter spacetime, as was discussed in section 2.6. The above value is obtained numerically when integrating foward in time, after setting the initial conditions at some late time T_0, as described in section 4.3.3. It can also be obtained analytically by calculating the area of the final horizon shape found in section 4.4.3. As discussed in section 4.4.3, the LP coordinates used in this thesis lead to a non-spherical shape for the horizon. However, the area is still what would be obtained if the slicing was chosen such that the horizon is a sphere with radius $1/H$. This is in keeping with the fact that for a stationary horizon, the area of the horizon is independent of the choice of slicing.

The overall shape of the graph is also as expected: it is monotonically increasing, with an inflection

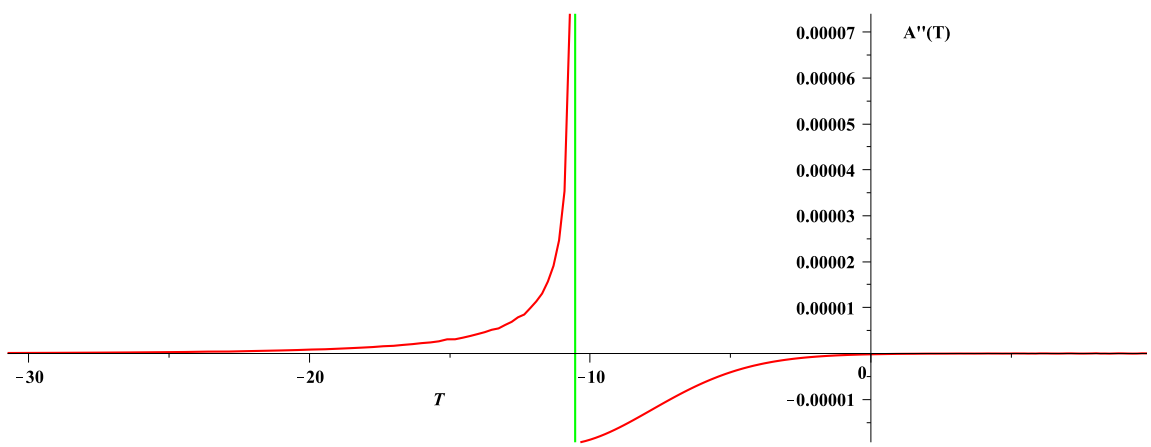
point at the moment where the rate of area increase is greatest. The monotonicity of the area graph is hardly surprising, in light of the famous area theorem of Hawking. However, it should be noted that the area theorem, as originally proven by Hawking [16], only applies to proper event horizons, and not the causal horizons like the one considered in this thesis. Nevertheless, it is not difficult to see how the original proof could be extended to causal horizons, given that the original proof rests on the focusing theorem, along with the observation that horizon generators cannot leave an event horizon through a future caustic point. The focusing theorem clearly applies to the congruence of null geodesics which are the null generators of a causal horizon. Furthermore, the presence of a future caustic point on a causal horizon would seem irreconcilable with the definition of a causal horizon as the boundary of the causal past of an observer's trajectory. Although the generalization of the area theorem to causal horizons seems straight forward enough, to our knowledge, a proof has not been published at this time. The existence of an unpublished proof has been alluded to in [17] however.

5.2.2 Inflection point in the horizon area vs time graph

The location and structure of the inflection point in the graph turn out to be quite intriguing. First, we have found numerically that the time at which the inflection point occurs coincides with the time of merger of the horizons. This can be seen in figure 5.2b, where the green vertical line has been drawn at the time of horizon merger. We see that $A''(T)$ changes sign precisely at the merger time, in keeping with the observation that the merger time coincides with the time where the inflection point occurs. The graphs of $A''(T)$ also reveals that we do not have $A''(T) = 0$ at the inflection point, unlike what we would expect if $A(T)$ was a twice differentiable function. Instead $A''(T)$ diverges at the inflection point. This immediately suggest a hypothesis regarding the reason for the inflection and merger times coinciding: perhaps there is something about the merger of the horizons which causes $A''(T)$ to diverge as the merger time is approached. For example, we might hypothesize that the presence of the caustic prior to merger causes $A''(T)$ to diverge right before merger time.



(a)



(b)

Figure 5.2: Examples plots of $A'(T)$ and $A''(T)$ for merging black hole and cosmological horizons. Both plots reveal a discontinuity in $A''(T)$ at the critical merger time T_p. This is most apparent in the plot of $A''(T)$, where we see a divergence just prior to the merger time (indicated by a vertical green line). Furthermore, this plot reveals that $A''(T)$ goes from positive to negative at the merger point. Parameter values used: $M = 1$, $\Lambda = 1/18$, $n = 400$, $\xi_0 = 10^{-4}$.

We thus have two intriguing hypotheses about the horizon area as a function of time. The first result is that $A''(T)$ diverges right before the merger point, and the second is that the merger time coincides precisely with the inflection point in the horizon area vs time graph. The first of these hypotheses will be shown to hold in a broad class of coordinate systems in section 5.3.1 below.

5.2.3 Horizon area in the limit $\varepsilon \to 0$

Area increase prior to merger

The parameter $\varepsilon = MH$ characterizes the size of the black hole relative to the size of the cosmological horizon. As discussed briefly at the end of section 2.7, for a typical supermasive black hole in a universe with a cosmological constant given by the current value according to the Λ-CDM model, we have $\varepsilon \approx 10^{-14}$, which is an exceedingly small number. For this reason, it is interesting and relevant to consider questions regarding the horizon area for small values of ε. When considering questions related to the total area increase, it is useful to introduce the dimensionless area $\hat{A}(T)$ as

$$\hat{A}(T) = \frac{A(T) - A(-\infty)}{A(\infty) - A(-\infty)} \,, \tag{5.18}$$

where $A(T)$ is the area at LP coordinate time T, and $A(\pm\infty)$ are given by (5.16) and (5.17). Notice that $\hat{A}(-\infty) = 0$ and $\hat{A}(\infty) = 1$. The above can also be written as

$$\hat{A}(T) = \frac{\Delta A(T)}{\Delta A} \,, \tag{5.19}$$

where $\Delta A(T) = A(T) - A(-\infty)$ is the area increase after time T, and $\Delta A = A(\infty) - A(-\infty)$ is the total area increase. Another useful quantity to introduce is a normalized value of ε. Recall from section 2.7 that SdS spacetime only has a black hole and cosmological horizon provided that $\varepsilon < \varepsilon_c = (3\sqrt{3})^{-1}$. With this in mind, we define

$$\hat{\varepsilon} = \frac{\varepsilon}{\varepsilon_c} = 3\sqrt{3}\varepsilon \,.$$

Therefore, the condition $\varepsilon < \varepsilon_c$ becomes simply $\hat{\varepsilon} < \hat{\varepsilon}_c = 1$. The advantage of this rescaling of the parameter ε is that it is easy to gauge the "smallness" of $\hat{\varepsilon}$, since it can be easily compared to the critical value $\hat{\varepsilon}_c = 1$.

Consider first the plot of area vs time for $\hat{\varepsilon} \approx 0.1$, as shown in figure 5.3a. One of the remarkable features of this graph is that the inflection point in the graph occurs at a time where the area is very nearly equal to its final value. This leads us to hypothesize that $\hat{A}(T_c) \to 1$ in the limit that $\hat{\varepsilon} \to 0$, where T_c is the time where the inflection point occurs. In figure 5.3, we plot $\hat{A}(T_c)$ vs $\hat{\varepsilon}$, and can clearly see strong numerical evidence supporting this hypothesis. We have created the same graph using the merger time instead of the inflection time T_c, and have found that the graph is identical. This is in keeping with the result from section 5.2.2, where we found numerically that the merger time and inflection point time coincided. The implication of the result in figure 5.3 is that in the limit that $\hat{\varepsilon} \to 0$, all of the area increase occurs prior to merger. That is, when an infinitesimally small black hole merges with the cosmological horizon, all of the area increase takes place before the black hole and cosmological horizon first touch. Quantitatively, the graph suggests that we have

$$\hat{A}(T_c) = 1 + \mathscr{O}(\hat{\varepsilon}) \ . \tag{5.20}$$

More precisely, by fitting a straight line to the graph in figure 5.3, we have the approximate formula

$$\hat{A}(T_c) \approx 1 - 1.8\hat{\varepsilon} \ . \tag{5.21}$$

Obviously the coefficient of $\hat{\varepsilon}$ in the above formula depends on how the horizon area was calculated, which ultimately depends on our choice of time coordinate. However, the fact that $\hat{A}(T_c) \to 1$ as $\hat{\varepsilon} \to 0$, as well as the linear dependence on $\hat{\varepsilon}$ of $\hat{A}(T_c)$ for $\hat{\varepsilon} \ll 1$, both ought to be results that hold for any choice of time coordinate. For any given choice of time coordinate, one can find a formula akin to the one above. The resulting formula then provides a very good approximation to the area at the time of merger for small values of $\hat{\varepsilon}$. Notice that based on figure 5.3, we see that the above formula should provide a reasonable approximation of $\hat{A}(T_c)$ for values of $\hat{\varepsilon}$ as large as $\hat{\varepsilon} \approx 0.1$. Also notice that the above formula gives the area for any values of the parameters M and H such that $HM \ll 1$, since given values for the parameters M and H, one can compute a value for $\hat{\varepsilon} = 3\sqrt{3}HM$ and then

combine the above with (5.18), (5.16) and (5.17) to find the change in area.

Rate of area increase at the time of merger

Another natural question to ask regarding the limit $\hat{\varepsilon} \to 0$ is whether or not $A'(T_c) \to 0$ in this limit. That is, whether or not the maximum rate of change of the area with respect to time vanishes in the limit $\hat{\varepsilon} \to 0$. To properly answer this question, it is useful to first nondimensionalize $A'(T)$. To do so, we nondimensionlize the numerator of dA/dT by diving it by $\Delta A = A(\infty) - A(-\infty)$, and nondimensionalize the denominator of dA/dT by dividing it by $1/H$. The motivation for these rescalings is that the numerator of dA/dT should be divided by the total change in area ΔA, and the denominator of dA/dT should be divided by the approximate time scale associated with the area increase, which is $1/H$. Letting $\bar{A}'$ be this nondimensional rate of change of area, we have

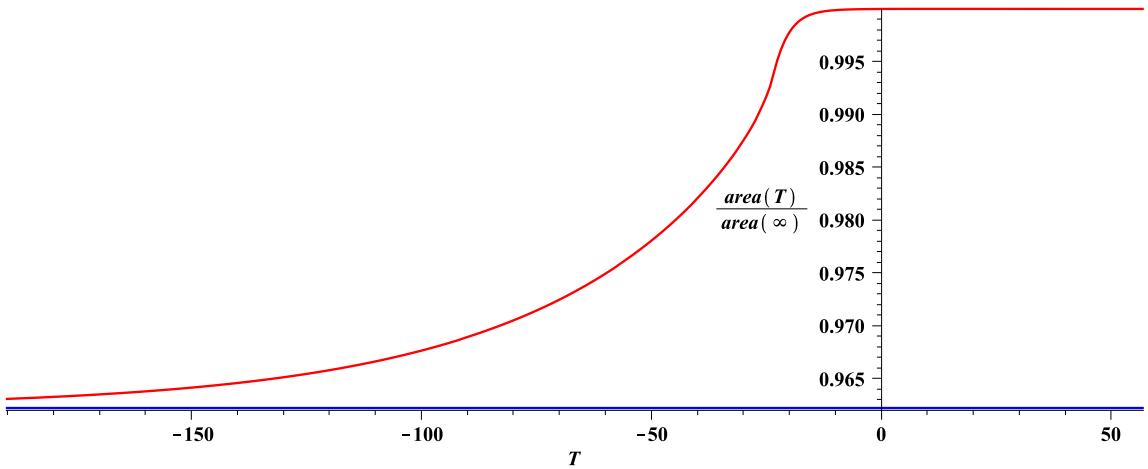
$$\bar{A}'(T) = \frac{A'(T)}{H\Delta A} \ . \tag{5.22}$$

A plot of $\bar{A}'(T_c)$ vs $\hat{\varepsilon}$ is shown in figure 5.4. Despite the considerable scatter of the points due to numerical error, there is a general trend showing that we have

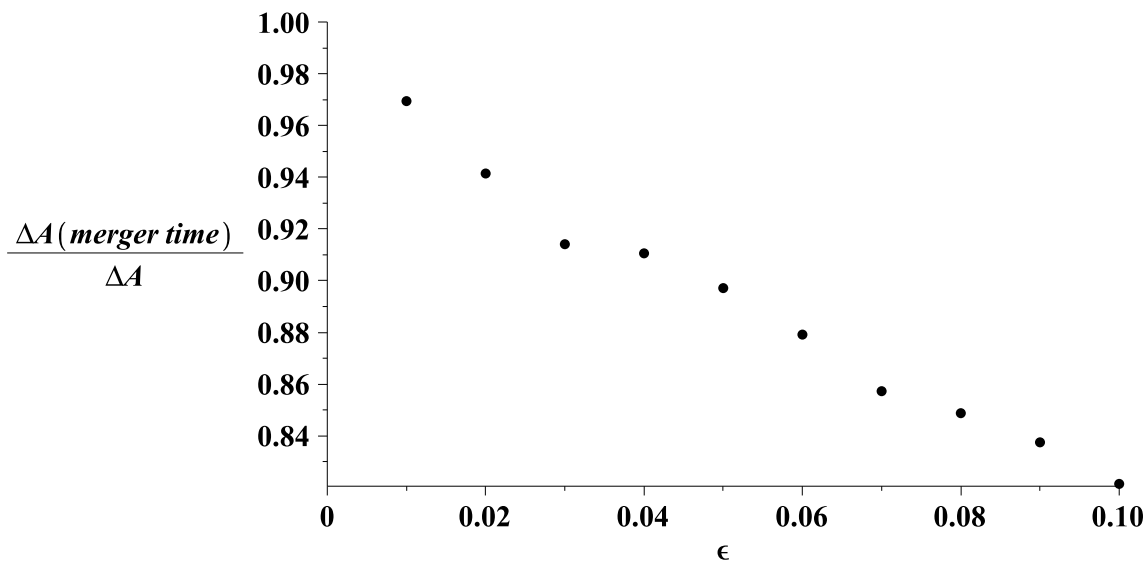
$$\bar{A}'(T_c) \to \text{constant} \quad \text{as} \quad \hat{\varepsilon} \to 0 \ . \tag{5.23}$$

If correct, the above suggests that as $\hat{\varepsilon} \to 0$, the $\bar{A}'(T)$ vs T graph (which would have the same basic shape as the $A'(T)$ graph in figure 5.2a) has a peak value at T_c which is roughly independent of $\hat{\varepsilon}$ for $\hat{\varepsilon} \ll 1$. Alternatively, we can recognize that

$$\begin{aligned} \Delta A &= 4\pi \left(\frac{1}{H^2} - r_c^2 - r_b^2 \right) \\ &= 4\pi \frac{1}{H^2} \left(1 - (Hr_c)^2 - (Hr_b)^2 \right) \\ &= 4\pi \frac{1}{H^2} \left(2\varepsilon + \mathscr{O}(\varepsilon^2) \right) \ , \end{aligned} \tag{5.24}$$



(a)



(b)

Figure 5.3: a) Example of horizon area vs time for a small value of ε. Notice that the time of maximum area increase, which is also the merger time, occurs at a time when the area has nearly reached its final value. b) A plot of the nondimensional area $\hat{A}$, as given by equation (5.18), at the merger time, for small values of the parameter $\hat{\varepsilon}$. Notice that $\hat{A}$ approaches 1 as $\hat{\varepsilon} \to 0$. This means that in the limit of small black hole mass, all of the area increase occurs before merger.

where we have used (2.15) in the last step. Using the above and the definition (5.22), the result (5.23) becomes:

$$A'(T_c) \sim \frac{1}{H}\hat{\varepsilon} \quad \text{as} \quad \hat{\varepsilon} \to 0 \, .$$

If we imagine taking the limit $\hat{\varepsilon} \to 0$ by setting $H =$ constant and taking $M \to 0$, then the above becomes:

$$A'(T_c) \sim M \quad \text{as} \quad M \to 0 \, . \tag{5.25}$$

In other words, if an infinitesimally small black hole merges with a cosmological horizon, the rate of change of area increase at the merger point will scale like the mass M of the black hole. By fitting a straight line to the points in figure 5.4, we can find the approximation

$$\bar{A}'(T_c) \approx 2.4 \quad \text{for} \quad \hat{\varepsilon} \ll 1 \, . \tag{5.26}$$

Using (5.22) and (5.24), viewing the above as a result for $M \ll 1/H$, we have

$$A'(T_c) \approx 60M \quad \text{for} \quad M \ll \frac{1}{H} \, . \tag{5.27}$$

The above result gives an approximate formula for the rate of change of area at the time of merger for small black hole masses. As in the approximate formula (5.21), the particular coefficient above depends on the choice of time coordinate, which in this case is the time T associated with the LP coordinates developed for this thesis. However, for any other coordinate system one can find the appropriate coefficient in a formula similar to the above. Furthermore, the linear dependence on M obtained in the above formula is a result which we would expect to hold in any coordinate system.

Average rate of area increase

A third related question that one can ask regarding the area vs time graph in the limit that $\hat{\varepsilon} \to 0$ is how the average of $A'(T)$ behaves in this limit, where the average is to be computed over some predetermined time interval. As before, it is useful to use the normalized area from (5.19) and the

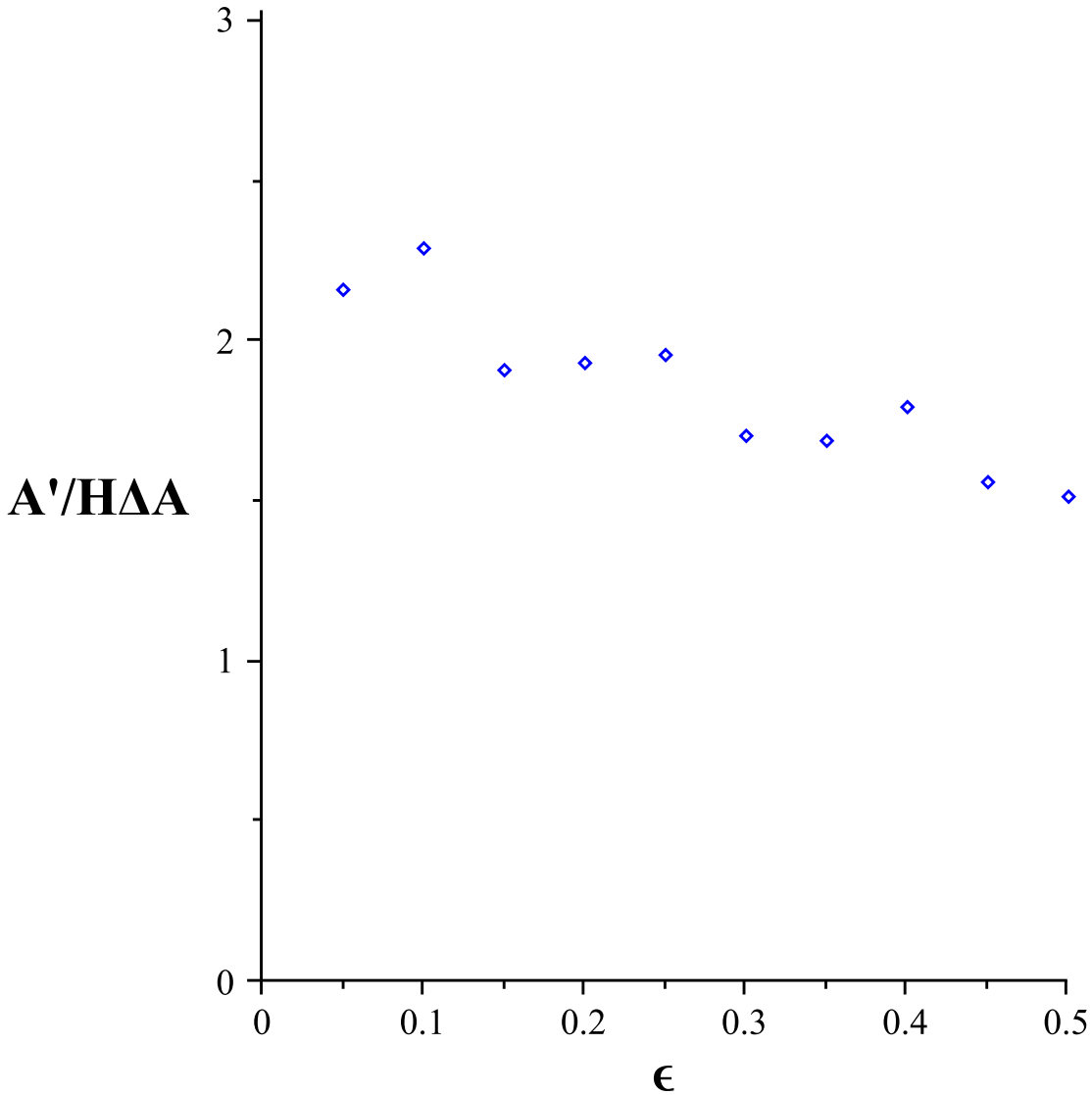


Figure 5.4: Plot of the maximum nondimensionalized rate of area increase, as given by equation (5.22), for several small values of the parameter $\hat{\varepsilon}$. We can see that this approaches a constant as $\hat{\varepsilon} \to 0$. The upshot of this is that the maximum rate of area increase $A'(T_c)$ scales like the mass of the black hole for $M \ll 1/H$, as encapsulated by equation (5.25).

nondimensional derivative introduced in (5.22). Let us define the time T_{half} as

$$\hat{A}(T_{half}) = \frac{1}{2} \ . \tag{5.28}$$

That is, T_{half} is the time at which half of the total change in area has taken place. Now we consider the average of $\bar{A}'(T)$ over the time interval from T_{half} to T_c. Letting this average be $\bar{A}'_{avg}$, we have

$$\begin{aligned} \bar{A}'_{avg} &= \frac{A(T_c) - A(T_{half})}{H\Delta A\left(T_c - T_{half}\right)} \\ &= \frac{\hat{A}(T_c) - \hat{A}(T_{half})}{H\left(T_c - T_{half}\right)} \\ &= \frac{1}{H\left(T_c - T_{half}\right)}\left(\frac{1}{2} + \mathscr{O}(\varepsilon)\right) \ , \end{aligned} \tag{5.29}$$

where we have used (5.19) in the first step, (5.22) in the second step, and (5.20) and (5.28) in the last step. In figure 5.5, we plot $H(T_c - T_{half})$ vs $\hat{\varepsilon}$ for several small values of $\hat{\varepsilon}$. There is considerable scatter in the points due to numerical error, but we can nevertheless conclude that

$$H(T_c - T_{half}) \to \text{constant} \quad \text{as} \qquad \hat{\varepsilon} \to 0 \ .$$

The same way that (5.26) lead to (5.27), the above and (5.29) allow us to conclude that

$$A'_{avg} \sim M \quad \text{for} \quad M \ll \frac{1}{H} \ .$$

That is, the derivative $A'(T)$, averaged over the time from T_{half} to T_c scales like the mass of the black hole, in the limit of small black hole mass. Thus the same scaling which applied to the maximum derivative, as in (5.25), applies to the average derivative as well.

In summary, we have looked at three aspects of the horizon area vs time graph for $\hat{\varepsilon} \ll 1$. The first is the fraction of area increase which occurs before merger, which we found to be unity in the limit that $\hat{\varepsilon} \to 0$. That is, in the limit of small black hole mass (or equivalently, small cosmological constant), all of the area increase occurs prior to merger. The second question we considered was the rate of

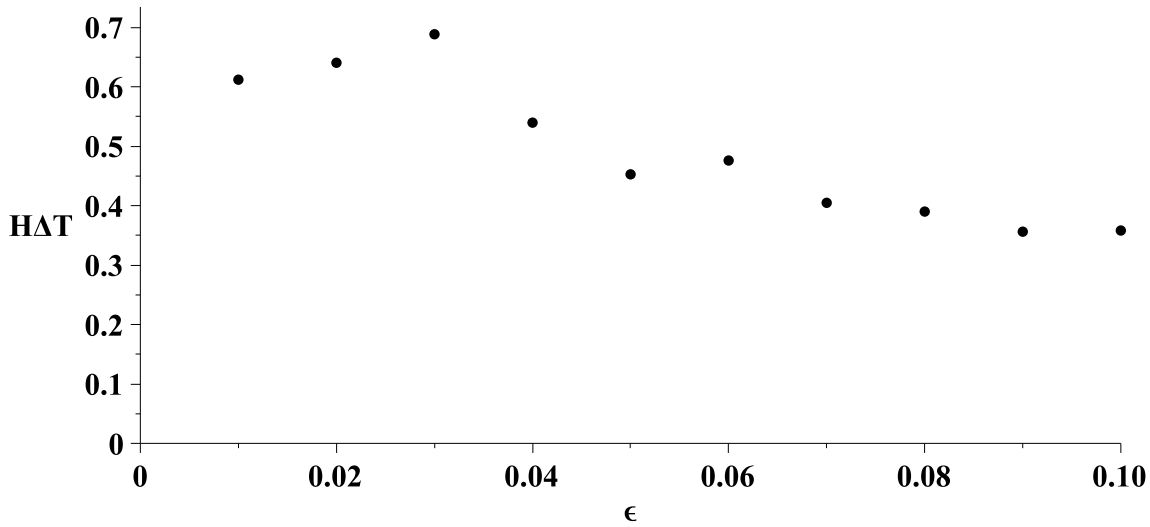


Figure 5.5: Plot of the nondimensional time interval $H\Delta T = H(T_{merger} - T_{half})$ for several small values of the paramter $\hat{\varepsilon}$, where T_{merger} is the merger time and T_{half} is the time where half of the area increase has occured. As we can see, $H\Delta T$ appears to approach a constant as $\hat{\varepsilon} \to 0$. This implies that for small black hole masses, the average rate of change of area over the time interval from T_{half} to T_{merger} goes like the mass of the black hole. This scaling is the same as was obtained for the maximum rate of area increase, as shown in figure 5.4.

change with resepct to time of horizon area, as calculated at the time of merger (as hypothesized in section 5.2.2, and assumed throughout this section, this is also the time at which the rate of change is maximal). This maximum rate of change was found to scale like the mass of the black hole for small ε. Finally, we considered the average rate of change of the area vs time, over a suitable interval of time, and found that it also followed the same scaling, being proportional to mass for $\hat{\varepsilon} \ll 1$.

Horizon area and null generators

The final question we investigate regarding the horizon area for $\hat{\varepsilon} \ll 1$ relates not to the dependence of area on time, but rather to the contribution of different null generators to the horizon area increase. As was discussed briefly at the end of section 4.4.2, null generators can be separated into two basic categories: those which are part of the horizon for all times, and those which only join the horizon at some time by entering through the caustic points. Let us call these the *existing generators* and

the *new generators*, respectively. This distinction between two types of generators is most easily seen in figure 4.7, where the ew generators are drawn in green. Using this distinction between these two types of generators, we can think of the horizon area as being due to two effects: the joining of new generators previously not on the horizon, and the expansion of generators which are on the horizon for all times (notice that those generators which join the horizon can also expand once on the horizon; this type of area increase still falls under the category of being due to "joining of new generators").

Recall that different horizon generators are parametrized using the parameter $\alpha \in [0, 2\pi)$. At the end of section 4.4.2 we defined α_1 and α_2 such that generators with $\alpha_1 \leq \alpha \leq \alpha_2$ or $\alpha_1 \leq 2\pi - \alpha \leq \alpha_2$ are the new generators. The meaning of these angles is most easily understood by looking at the green part of the last frame of figure 4.7. In section 4.4.2, we also explained that the value of $\phi(T = -\infty)$ can be used to calculate the values of α_1 and α_2. Specifically, we have

$$\alpha_1 \leq \alpha \leq \alpha_2 \quad \text{or} \quad \alpha_1 \leq 2\pi - \alpha \leq \alpha_2 \Longleftrightarrow |\phi(T = -\infty)| \geq \pi \ .$$

Based on the above, we can approximate α_1 and α_2 numerically using the value of $\phi(T_i)$, where T_i is a very early time, as compared with T_{merger} (more precisely, we have $T_i - T_{merger} \ll 1/H$). Let us define $A_{ng}(T)$ and $A_{eg}(T)$ to be the contribution to the horizon area from the new generators and existing generators, respectively, at some time T. That is, $A_{ng}(T)$ is the contribution to the area from generators satisfying $\alpha_1 \leq \alpha \leq \alpha_2$ or $\alpha_1 \leq 2\pi - \alpha \leq \alpha_2$, and $A_{eg}(T)$ is the contribution to the area from the remaining generators. By definition, we have

$$\begin{aligned} A_{ng}(-\infty) &= 0 \ , \\ A_{eg}(-\infty) &= 4\pi \left(r_b^2 + r_c^2\right) \ . \end{aligned}$$

The area of $A_{eg}(-\infty)$ is of course simply the area of the spherical black hole and cosmological horizons, with Schwarzschild radii r_b and r_c, respectively. Let us also define $\Delta A_{ng}(T)$ and $\Delta A_{eg}(T)$ to be the area increase due to new generators and existing generators, respectively, after time T.

That is, we have

$$\begin{aligned}\Delta A_{ng}(T) &= A_{ng}(T) - A_{ng}(-\infty)\ ,\\ \Delta A_{eg}(T) &= A_{eg}(T) - A_{eg}(-\infty)\ .\end{aligned}$$

As previously discussed, we can approximate α_1 and α_2 by using the condition $|\phi(T_i)| \geq \pi$. This is done by integrating the null geodesic equations backwards until time T_i, using the procedure described in section 4.3. Once we have approximate values of α_1 and α_2, we can use these to find approximate values for $A_{ng}(T)$ and $A_{eg}(T)$, where the horizon area is calculated using the procedure described in section 5.1, and the horizon generators are calculated from the null geodesic equations as described in section 4.3. Letting T_f be the final time of integration (actually, the "initial time" if we think of integration as proceeding backwards), we can compute $\Delta A_{ng}(T_f)$ and $\Delta A_{eg}(T_f)$. For $T_f - T_{merger} \gg 1/H$, $\Delta A_{ng}(T_f)$ and $\Delta A_{eg}(T_f)$ are then thought of as approximations to $\Delta A_{ng}(\infty)$ and $\Delta A_{eg}(\infty)$. Computing the ratio of $\Delta A_{eg}(T_f)$ to $\Delta A_{ng}(T_f)$ for T_f sufficiently large and for different values of $\hat{\varepsilon}$, we obtain the plot shown in figure 5.6. According to this plot, it appears that the ratio plotted approaches a constant as $\hat{\varepsilon} \to 0$. As we will show analytically in section 5.3.2, the ratio of $\Delta A_{eg}(\infty)$ to $\Delta A_{ng}(\infty)$ approaches unity as $\hat{\varepsilon} \to 0$. This means that in the limit of infinitesimally small black hole mass, new generators and existing generators contribute an equal amount to the horizon area increase. This is a surprising mathematical coincidence. Questions about generalizing this result, or the reason for its origin, will be discussed further in section 6.2. Notice that in figure 5.6, we see in the limit $\varepsilon \to 0$, the ratio of $\Delta A_{eg}(T_f)$ to $\Delta A_{ng}(T_f)$ appears to approach a constant with value approximately equal to 0.8, and not unity. This is due to numerical errors in the way that $\Delta A_{eg}(T_f)$ and $\Delta A_{ng}(T_f)$ are calculated, and is an issue we are currently addressing. Analytical methods in section 5.3.2 will confirm that this ratio is indeed unity.

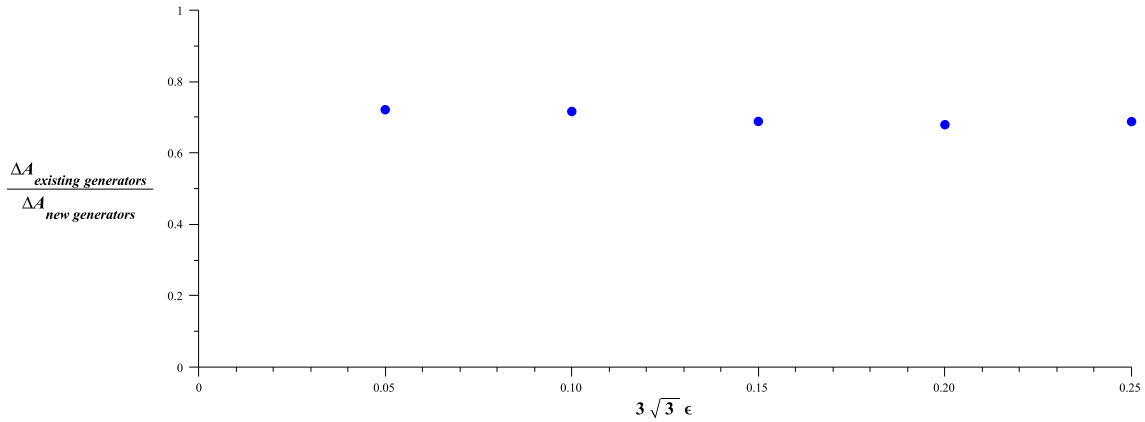


Figure 5.6: Ratio of two types of area increase: that due to the expansion of existing generators, and that due to the joining of new generators not previously on the horizon. This ratio was calculated numerically for several small values of $\hat{\varepsilon}$. As will be shown analytically in section 5.3.2, this ratio is equal to unity. In the plot above, the ratio appears to approach a value close to 0.8 instead of 1. This is due to numerical errors.

5.3 Analytical results

5.3.1 Time of maximum area increase

In section 5.2.2, we presented numerical evidence showing that the time of maximal area increase (i.e. the inflection point in figure 5.1) coincides with the time at which the merger of horizons occurs. In this section we extend this result using analytical methods. The plots in figure 5.2 provide a hint as to how to proceed. As we can see in 5.2b, the second derivative of $A(T)$ diverges prior to the inflection point. This suggests that there may be something about the merger time which causes a divergence in $A''(T)$. In this section, we show that $A''(\tau_p^-) = \infty$ for a broad class of coordinate systems, where $A(\tau)$ is the total horizon area at coordinate time τ and τ_p is the time of merger. This establishes that

$$\lim_{\tau \to \tau_p^-} \frac{d^2 A}{d\tau^2} > 0 \,.$$

The second result which would be needed to complete the proof is $A''(\tau_p^+) < 0$. However, it is not possible to show that $A''(\tau_p^+) < 0$ for an arbitrary coordinate system since the sign of $A''(\tau)$ after merger ultimately depends on the choice of time coordinate. Since showing that $A'(\tau)$ has a

maximum at the time of merger requires us to show that both $A''(\tau_p^-) > 0$ and $A''(\tau_p^+) < 0$ for a given choice choice of time coordinate, our proof is somewhat incomplete. Despite the incompleteness of our proof, we can motivate that $A''(\tau_p^-) > 0$ and $A''(\tau_p^+) < 0$ should occur for a family of coordinate systems. The reasoning is as follows. If we think of taking a coordinate system which is in some sense close to LP coordinates, then by continuity the fact that $A''(\tau_p^+) < 0$ for LP coordinates (as can be seen from figure 5.2b) implies that $A''(\tau_p^+) < 0$ must hold for any other sufficiently "near" coordinate system. Since LP coordinates belong to the class of coordinate systems considered below, this other nearby coordinate system can be constrained be within this class, from which it follows that $A''(\tau_p^-) > 0$ for this other coordinate system as well. This establishes that $A''(\tau_p^-) > 0$ and $A''(\tau_p^+) < 0$ hold for coordinate systems which are sufficiently "near" to LP coordinates, and which satisfy the constraints of the coordinate systems considered below.

Second derivative prior to merger

Consider a family of coordinate systems such that the line element on the constant time hypersurfaces is given by (5.2). From (5.5), we know that the area before merger is given by the following expression:

$$\begin{aligned} A(\tau) &= 2\pi \int_0^{\alpha_c} r \mid \sin\phi \mid \left(F(r,\tau) \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right)^{1/2} d\alpha \\ &\quad + 2\pi \int_{\alpha_b}^{\pi} r \mid \sin\phi \mid \left(F(r,\tau) \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right)^{1/2} d\alpha \ , \end{aligned} \tag{5.30}$$

where $\alpha \in [0, 2\pi)$ parametrizes the curve which is the intersection of the surface S_τ and $\theta = \pi/2$ (see section). In other words, α parametrizes the curves such as those shown in figure 4.4. Instead of using α to parametrize these curves, it is useful instead to use the arclength along the curve as a parameter. The relationship between arclength l and α is

$$dl^2 = \left(F(r,\tau) \left(\frac{dr}{d\alpha}\right)^2 + r^2 \left(\frac{d\phi}{d\alpha}\right)^2 \right) d\alpha^2 \ .$$

Since arclength is a strictly increasing function of α we can use it as an integration variable, so that the integrals in (5.30) become

$$A(\tau) = 2\pi \int_0^{l_c} r \sin\phi \, dl + 2\pi \int_{l_b}^{L} r \sin\phi \, dl \ ,$$

where l_c and l_b are the arc length values corresponding to α_c and α_b, respectively, and L is the total arc length of the curve. Note that $l_c = l_c(\tau)$, $l_b = l_b(\tau)$ and $L = L(\tau)$ are all functions of τ, and that there is also a dependence on τ through $r = r(\tau, l)$ and $\phi = \phi(\tau, l)$. Differentiating the area formula above with respect to τ, we get

$$\begin{aligned}\frac{1}{2\pi}A'(\tau) &= r\left(\tau, l_c(\tau)\right) \sin\left(\phi\left(\tau, l_c(\tau)\right)\right) l_c'(\tau) + \int_0^{l_c} \frac{\partial}{\partial \tau}\left(r \sin\phi\right) dl \\ &\quad + r\left(\tau, L(\tau)\right) \sin\left(\phi\left(\tau, L(\tau)\right)\right) L'(\tau) - r\left(\tau, l_b(\tau)\right) \sin\left(\phi\left(\tau, l_b(\tau)\right)\right) l_b'(\tau) + \int_{l_b}^{L} \frac{\partial}{\partial \tau}\left(r \sin\phi\right) dl \ .\end{aligned}$$

Differentiating with respect to τ again and simplifying slightly, we obtain the following expression for $A''(\tau)$:

$$\begin{aligned}\frac{1}{2\pi}A''(\tau) &= \left[\left.\frac{\partial r}{\partial l}\right|_{l=l_c} l_c'(\tau) + 2\left.\frac{\partial r}{\partial \tau}\right|_{l=l_c}\right] \sin\left(\phi\left(\tau, l_c(\tau)\right)\right) l_c'(\tau) \\ &+ \left[\left.\frac{\partial \phi}{\partial l}\right|_{l=l_c} l_c'(\tau) + 2\left.\frac{\partial \phi}{\partial \tau}\right|_{l=l_c}\right] r\left(\tau, l_c(\tau)\right) \cos\left(\phi\left(\tau, l_c(\tau)\right)\right) l_c'(\tau) + r\left(\tau, l_c(\tau)\right) \sin\left(\phi\left(\tau, l_c(\tau)\right)\right) l_c''(\tau) \\ &\quad + \int_0^{l_c} \frac{\partial^2}{\partial \tau^2}\left(r \sin\phi\right) dl + \left[\left.\frac{\partial r}{\partial l}\right|_{l=L} L'(\tau) + 2\left.\frac{\partial r}{\partial \tau}\right|_{l=L}\right] \sin\left(\phi\left(\tau, L(\tau)\right)\right) L'(\tau) \\ &+ \left[\left.\frac{\partial \phi}{\partial l}\right|_{l=L} L'(\tau) + 2\left.\frac{\partial \phi}{\partial \tau}\right|_{l=L}\right] r\left(\tau, L(\tau)\right) \cos\left(\phi\left(\tau, L(\tau)\right)\right) L'(\tau) + r\left(\tau, L(\tau)\right) \sin\left(\phi\left(\tau, L(\tau)\right)\right) L''(\tau) \\ &\quad - \left[\left.\frac{\partial r}{\partial l}\right|_{l=l_b} l_b'(\tau) + 2\left.\frac{\partial r}{\partial \tau}\right|_{l=l_b}\right] \sin\left(\phi\left(\tau, l_b(\tau)\right)\right) l_b'(\tau) \\ &\quad - \left[\left.\frac{\partial \phi}{\partial l}\right|_{l=l_b} l_b'(\tau) + 2\left.\frac{\partial \phi}{\partial \tau}\right|_{l=l_b}\right] r\left(\tau, l_b(\tau)\right) \cos\left(\phi\left(\tau, l_b(\tau)\right)\right) l_b'(\tau) \\ &\quad - r\left(\tau, l_b(\tau)\right) \sin\left(\phi\left(\tau, l_b(\tau)\right)\right) l_b''(\tau) + \int_{l_b}^{L} \frac{\partial^2}{\partial \tau^2}\left(r \sin\phi\right) dl \ . \qquad (5.31)\end{aligned}$$

We want to find $A''(\tau)$ in the limit $\tau \to \tau_p^-$. To do so, first consider the functions $l_c(\tau)$ and $l_b(\tau)$. Recall that at any time $\tau < \tau_p$ prior to merger, $l_c(\tau)$ and $l_b(\tau)$ are the values of the parameter l

for which $(r(\tau,l_c),\phi(\tau,l_c))$ and $(r(\tau,l_b),\phi(\tau,l_b))$ are the locations of the caustic points where new generators enter the horizon (these are the nonsmooth points on the horizon in figure 4.4). By spherical symmetry, we can assume without loss of generality that these caustic points occur along $\phi=\pi$, so that $l_c(\tau)$ and $l_b(\tau)$ are given implicitly by the following equation:

$$\phi(\tau,l)=\pi\ . \tag{5.32}$$

That is, we have

$$\begin{aligned}\phi(\tau,l_c(\tau)) &= \pi\ ,\\ \phi(\tau,l_b(\tau)) &= \pi\ ,\end{aligned}$$

for all values $\tau<\tau_p$. Similarly, $L(\tau)$ is defined as the value of $l>0$ for which $\phi(\tau,l)=0$. This is half of the total arc length of any of the curves shown in figure 4.4, since the top half of these curves starts at $\phi=0$ for $l=0$ and ends at $\phi=0$ for $l=L$. Using $\phi(\tau,L)=0$ and the above, we can immediately conclude that the terms involving $\sin(\phi(\tau,l_c(\tau)))$, $\sin(\phi(\tau,l_b(\tau)))$ or $\sin(\phi(\tau,L(\tau)))$ in (5.31) vanish for all values $\tau<\tau_p$, and we are left with

$$\begin{aligned}\frac{1}{2\pi}A''(\tau) &= \left[\left.\frac{\partial\phi}{\partial l}\right|_{l=l_c} l_c'(\tau)+2\left.\frac{\partial\phi}{\partial\tau}\right|_{l=l_c}\right] r(\tau,l_c(\tau))\cos\left(\phi\left(\tau,l_c(\tau)\right)\right) l_c'(\tau)\\ &\quad-\left[\left.\frac{\partial\phi}{\partial l}\right|_{l=l_b} l_b'(\tau)+2\left.\frac{\partial\phi}{\partial\tau}\right|_{l=l_b}\right] r(\tau,l_b(\tau))\cos\left(\phi\left(\tau,l_b(\tau)\right)\right) l_b'(\tau)\\ &\quad+\left[\left.\frac{\partial\phi}{\partial l}\right|_{l=L} L'(\tau)+2\left.\frac{\partial\phi}{\partial\tau}\right|_{l=L}\right] r(\tau,L(\tau))\cos\left(\phi\left(\tau,L(\tau)\right)\right) L'(\tau)\\ &\quad+\int_0^{l_c}\frac{\partial^2}{\partial\tau^2}\left(r\sin\phi\right)dl+\int_{l_b}^{L}\frac{\partial^2}{\partial\tau^2}\left(r\sin\phi\right)dl\ .\end{aligned} \tag{5.33}$$

Next we consider the asymptotic behavior of $l_c(\tau)$ and $l_b(\tau)$ for $\tau<\tau_p$ and τ near τ_p. For any value of $\tau<\tau_p$, (5.32) yields two solutions for l, which are $l_c(\tau)$ and $l_b(\tau)$. For $\tau>\tau_p$, horizon merger has occured and (5.32) no longer has any roots (as in the fifth frame in figure 4.4). Letting $l_p=l_c(\tau_p)=l_b(\tau_p)$, this implies that $\phi(\tau,l)$ has the following Taylor expansion about (τ_p,l_p):

$$\phi(\tau,l)=\pi+\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}(\tau-\tau_p)+\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}(l-l_p)^2+\mathscr{O}\left((\tau-\tau_p)^2+(l-l_p)^3\right)\ , \tag{5.34}$$

where $\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p} < 0$ and $\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p} < 0$. From the above and (5.32) we can deduce that $l_c(\tau)$ and $l_b(\tau)$ have the following asymptotic behavior:

$$l_c(\tau) = l_p - \left[\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\right)\left(\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{-1}\right]^{\frac{1}{2}}(\tau_p-\tau)^{\frac{1}{2}} + \mathscr{O}\left((\tau-\tau_p)^{\frac{3}{2}}\right) , \tag{5.35}$$

$$l_b(\tau) = l_p + \left[\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\right)\left(\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{-1}\right]^{\frac{1}{2}}(\tau_p-\tau)^{\frac{1}{2}} + \mathscr{O}\left((\tau-\tau_p)^{\frac{3}{2}}\right) . \tag{5.36}$$

Differentiating the above, we get the asymptotic behavior of the first derivatives of $l_c(\tau)$ and $l_b(\tau)$:

$$l_c'(\tau) = \frac{1}{2}\left[\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\right)\left(\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{-1}\right]^{\frac{1}{2}}(\tau_p-\tau)^{-\frac{1}{2}} + \mathscr{O}\left((\tau-\tau_p)^{\frac{1}{2}}\right) ,$$

$$l_b'(\tau) = -\frac{1}{2}\left[\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\right)\left(\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{-1}\right]^{\frac{1}{2}}(\tau_p-\tau)^{-\frac{1}{2}} + \mathscr{O}\left((\tau-\tau_p)^{\frac{1}{2}}\right) .$$

From (5.34) and (5.35)-(5.36), we can also deduce the following:

$$\left.\frac{\partial\phi}{\partial l}\right|_{l=l_c} = -2\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{\frac{1}{2}}(\tau_p-\tau)^{\frac{1}{2}} + \mathscr{O}(\tau-\tau_p) ,$$

$$\left.\frac{\partial\phi}{\partial l}\right|_{l=l_b} = 2\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{\frac{1}{2}}(\tau_p-\tau)^{\frac{1}{2}} + \mathscr{O}(\tau-\tau_p) ,$$

$$\left.\frac{\partial\phi}{\partial\tau}\right|_{l=l_c} = \left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p} + \mathscr{O}(\tau-\tau_p) ,$$

$$\left.\frac{\partial\phi}{\partial\tau}\right|_{l=l_b} = \left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p} + \mathscr{O}(\tau-\tau_p) .$$

Using the results above, we can now find the asymptotic behavior of the first two terms in (5.33):

$$\left[\left.\frac{\partial\phi}{\partial l}\right|_{l=l_c} l_c'(\tau)+2\left.\frac{\partial\phi}{\partial\tau}\right|_{l=l_c}\right] r\left(\tau,l_c(\tau)\right)\cos\left(\phi\left(\tau,l_c(\tau)\right)\right) l_c'(\tau) = \tag{5.37}$$

$$\left[\frac{3}{2}\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\right)^2\left(-\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{-\frac{1}{2}} r(\tau_p,l_p)\right](\tau_p-\tau)^{-\frac{1}{2}}+\mathscr{O}(1) \tag{5.38}$$

$$\left[\left.\frac{\partial\phi}{\partial l}\right|_{l=l_b} l_b'(\tau)+2\left.\frac{\partial\phi}{\partial\tau}\right|_{l=l_b}\right] r\left(\tau,l_b(\tau)\right)\cos\left(\phi\left(\tau,l_b(\tau)\right)\right) l_b'(\tau) = \tag{5.39}$$

$$-\left[\frac{3}{2}\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,l_p}\right)^2\left(-\left.\frac{\partial^2\phi}{\partial l^2}\right|_{\tau_p,l_p}\right)^{-\frac{1}{2}} r(\tau_p,l_p)\right](\tau_p-\tau)^{-\frac{1}{2}}+\mathscr{O}(1) \tag{5.40}$$

Now consider the third term in (5.33). Since $L(\tau)$ is defined by $\phi(\tau,L(\tau))$, the implicit function theorem implies that we have:

$$L(\tau)=L_p+\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,L_p}\left(\left.\frac{\partial\phi}{\partial l}\right|_{\tau_p,L_p}\right)^{-1}(\tau-\tau_p)+\mathscr{O}\left((\tau-\tau_p)^2\right)$$

where $L_p=L(\tau_p)$. Using the above, we find that the third term in (5.33) has the following asymptotic behavior:

$$\left[\left.\frac{\partial\phi}{\partial l}\right|_{l=L} L'(\tau)+2\left.\frac{\partial\phi}{\partial\tau}\right|_{l=L}\right] r\left(\tau,L(\tau)\right)\cos\left(\phi\left(\tau,L(\tau)\right)\right) L'(\tau)$$
$$=3\left(\left.\frac{\partial\phi}{\partial\tau}\right|_{\tau_p,L_p}\right)^2\left(\left.\frac{\partial\phi}{\partial l}\right|_{\tau_p,L_p}\right)^{-1}+\mathscr{O}(\tau-\tau_p)=\mathscr{O}(1) \tag{5.41}$$

Lastly, consider the two integrals in (5.33). Since the first and second partial derivatives of $r(\tau,l)$ and $\phi(\tau,l)$ with respect to τ are $\mathscr{O}(1)$, we have:

$$\int_0^{l_c}\frac{\partial^2}{\partial\tau^2}\left(r\sin\phi\right)dl+\int_{l_b}^{L}\frac{\partial^2}{\partial\tau^2}\left(r\sin\phi\right)dl=\mathscr{O}(1) \tag{5.42}$$

Using (5.37)-(5.39), (5.41) and (5.42), we can deduce that our expression for $A''(\tau)$ in (5.33) has the following asymptotic behavior:

$$\frac{1}{2\pi}A''(\tau) = \left[3\left(\frac{\partial\phi}{\partial\tau}\bigg|_{\tau_p,l_p}\right)^2\left(-\frac{\partial^2\phi}{\partial l^2}\bigg|_{\tau_p,l_p}\right)^{-\frac{1}{2}} r(\tau_p,l_p)\right](\tau_p-\tau)^{-\frac{1}{2}} + \mathscr{O}(1)$$

From the above it follows that:

$$\lim_{\tau\to\tau_p^-}\frac{1}{2\pi}A''(\tau) = \infty$$

5.3.2 Horizon area increase and horizon generators

In this section we derive approximate formulas for $\Delta A_{ng}(\infty)$ and $\Delta A_{eg}(\infty)$, valid for $\varepsilon \ll 1$, where $\Delta A_{ng}(T)$ and $\Delta A_{eg}(T)$ were defined in section 5.2.3. We then use these formulas to prove the main result of this section, which is that:

$$\frac{\Delta A_{ng}(\infty)}{\Delta A_{eg}(\infty)} \to 1 \quad \text{as} \quad \varepsilon \to 0 \tag{5.43}$$

That is, we prove that in the limit $\varepsilon \to 0$, the total area increase due to the joining of new generators is equal to the total area increase due to the expansion of generators. This result was discussed in section 5.2.3, where we provided numerical evidence hinting at its validity.

We prove (5.43) in three steps. In section (5.3.2), we find approximate formulas for α_1 and α_2, valid for $\varepsilon \ll 1$, where α_1 and α_2 were defined in section 5.2.3. In sections 5.3.2 and 5.3.2, we find the area increase due to new generators and expanding generators, respectively, for $\varepsilon \ll 1$. That is, we find approximate formulas for $\Delta A_{ng}(\infty)$ and $\Delta A_{eg}(\infty)$, valid for $\varepsilon \ll 1$. Finally, we combine the results from sections 5.3.2 and 5.3.2 to find the ratio of $\Delta A_{ng}(\infty)$ to $\Delta A_{eg}(\infty)$, in the limit that $\varepsilon \to 0$.

Calculation of α_1 and α_2 for $\varepsilon \ll 1$

Recall from section 5.2.3 that we defined $\alpha_!$ and α_2 such that generators with $\alpha_1 \leq \alpha \leq \alpha_2$ or $\alpha_1 \leq 2\pi - \alpha \leq \alpha_2$ are the new generators which join the horizon at some time. We wish to find formulas for $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$, viewed as functions of ε, in the limit that $\varepsilon \ll 1$. This is the limit in which the Schwarzschild radius of the black hole is much smaller than the Hubble radius associated with the cosmological constant. We will do this in three steps. First we find the defining equations for $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$, then we find the formula for $\alpha_1(\varepsilon)$, and finally we find the formula for $\alpha_2(\varepsilon)$. The formula for $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$ will be used in sections 5.3.2 and 5.3.2 to find the area increase due to expanding generators and new generators.

Defining equations We start by defining $\phi_0(\alpha;\varepsilon)$ as:

$$\phi_0(\alpha;\varepsilon) = \lim_{T\to-\infty} \phi(T,\alpha;\varepsilon)$$

provided such a limit exists. Next notice that

$$\phi_0(\alpha_1;\varepsilon) = \pi, \tag{5.44}$$

$$\phi_0(\alpha_2;\varepsilon) = \pi \tag{5.45}$$

which follows from the fact that null generators must enter the horizon through the caustic point, together with the fact that the caustic is located at $\phi = \pi$. Also note that $\phi(\alpha_1;\varepsilon)$ and $\phi(\alpha_2;\varepsilon)$ are guaranteed to exist since as $T \to -\infty$, the null generators with α_1 and α_2 asymptotically approach the null generators of the horizon for a stationary observer, which are know to have $\phi =$ constant. (5.44)-(5.45) implicitly define the functions $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$. To make further progress, we must integrate the null geodesics equations (4.3) and (4.6)-(4.7) in order to find integral expressions for $\phi_0(\alpha_1;\varepsilon)$ and $\phi_0(\alpha_2;\varepsilon)$. From (4.3) and (4.6)-(4.7), we have that null geodesics, when thought of as

described by functions $\phi(r)$, must satisfy one of the two equations:

$$\text{If}\,\alpha \in [0,\alpha^*):\ \frac{d\phi}{dr} = \begin{cases} \frac{-\sin\alpha}{r^2\sqrt{H^2-V_{eff}(r)}} & \text{for}\, r > r^* \\ \frac{\sin\alpha}{r^2\sqrt{H^2-V_{eff}(r)}} & \text{for}\, r < r^* \end{cases} \tag{5.46}$$

$$\text{If}\,\alpha \in [\alpha^*,\pi):\ \frac{d\phi}{dr} = \frac{-\sin\alpha}{r^2\sqrt{H^2-V_{eff}(r)}} \tag{5.47}$$

In the first case above, there is a turning point of the motion, with r^* the Schwarzschild coordinate of the turning point. α^* and r^* are given by (as discussed in section 4.4.2):

$$\begin{aligned} \sin^2\alpha^* &= 27M^2H^2,\ \alpha^* \in \left(\frac{\pi}{2},\pi\right) \\ V_{eff}(r^*) &= H^2,\ r^* > 3M \end{aligned} \tag{5.48}$$

We hypothesize that the null generator with $\alpha = \alpha_1$ has a turning point of the motion, whereas the null generator with $\alpha = \alpha_2$ does not have a turning point. This hypothesis will then be vindicated by our calculations of α_1 and α_2 under this assumption. Using our hypothesis, we have that $\phi_0(\alpha_1;\varepsilon)$ and $\phi_0(\alpha_2;\varepsilon)$ will be found by integrating (5.46) and (5.47), respectively. Before integrating (5.46)-(5.47), we need appropriate limits of integration. These can be obtained by considering $r(T)$ and $\phi(T)$ as $T \to \pm\infty$. Based on our results from section 4.4.2, the null generator with $\alpha = \alpha_1$ has the following limiting behavior:

$$\begin{aligned} r(T=-\infty) &= r_c \\ r(T=\infty) &= \infty \\ \phi(T=-\infty) &= \phi_0(\alpha_1;\varepsilon) \\ \phi(T=\infty) &= 0 \end{aligned} \tag{5.49}$$

Similarly, the null generator with $\alpha = \alpha_2$ obeys:

$$\begin{aligned} r(T=-\infty) &= r_b \\ r(T=\infty) &= \infty \\ \phi(T=-\infty) &= \phi_0(\alpha_2;\varepsilon) \\ \phi(T=\infty) &= 0 \end{aligned} \tag{5.50}$$

Using (5.49)-(5.50) as the limits of integration, we can integrate (5.46)-(5.47) to obtain integral expressions for $\phi_0(\alpha_1;\varepsilon)$ and $\phi(\alpha_2;\varepsilon)$. These are:

$$\begin{aligned} \phi_0(\alpha_1;\varepsilon) &= \int_{r^*}^{\infty} \frac{\sin\alpha_1}{r^2\sqrt{H^2-V_{eff}(r)}}dr + \int_{r^*}^{r_c} \frac{\sin\alpha_1}{r^2\sqrt{H^2-V_{eff}(r)}}dr \\ \phi_0(\alpha_2;\varepsilon) &= \int_{r_b}^{\infty} \frac{\sin\alpha_2}{r^2\sqrt{H^2-V_{eff}(r)}}dr \end{aligned}$$

Combining the above with (5.44)-(5.45), we obtain:

$$\begin{aligned} \pi &= \int_{r^*}^{\infty} \frac{\sin\alpha_1}{r^2\sqrt{H^2-V_{eff}(r)}}dr + \int_{r^*}^{r_c} \frac{\sin\alpha_1}{r^2\sqrt{H^2-V_{eff}(r)}}dr \\ \pi &= \int_{r_b}^{\infty} \frac{\sin\alpha_2}{r^2\sqrt{H^2-V_{eff}(r)}}dr \end{aligned}$$

The above implicitly define the functions $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$. The depedence of α_1 and α_2 on ε can be more easily seen by first making the following change of variables in the above integrals:

$$u = \frac{M}{r}$$

The integrals become:

$$\pi = \frac{1}{M}\int_0^{u^*} \frac{\sin\alpha_1}{\sqrt{H^2 - V_{eff}\left(\frac{M}{u}\right)}}du + \frac{1}{M}\int_{u_c}^{u^*} \frac{\sin\alpha_1}{\sqrt{H^2 - V_{eff}\left(\frac{M}{u}\right)}}du \tag{5.51}$$

$$\pi = \frac{1}{M}\int_0^{u_b} \frac{\sin\alpha_2}{\sqrt{H^2 - V_{eff}\left(\frac{M}{u}\right)}}du \tag{5.52}$$

where $u^* = M/r^*$, $u_c = M/r_c$ and $u_b = M/r_b$. Substituting the explicit expression for $V_{eff}(r)$:

$$V_{eff}(r) = V_{eff}\left(\frac{M}{u}\right) = \frac{\sin\alpha}{M^2}u^2(1-2u) \tag{5.53}$$

into (5.51)-(5.52), and manipulating the limits of integration in (5.51), we obtain

$$\pi = 2\int_0^{u^*(\alpha_1,\varepsilon)} \frac{1}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_1}\right)^2 - u^2(1-2u)}}du - \int_0^{u_c(\varepsilon)} \frac{1}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_1}\right)^2 - u^2(1-2u)}}du \tag{5.54}$$

$$\pi = \int_0^{u_b(\varepsilon)} \frac{1}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_2}\right)^2 - u^2(1-2u)}}du \tag{5.55}$$

where we have written $u^* = u^*(\alpha_1,\varepsilon)$, $u_c = u_c(\varepsilon)$ and $u_b = u_b(\varepsilon)$ to emphasize the depedendence on α_1 or ε. In the above we clearly see that we are dealing with implicit relations for $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$. Our task is now to extract the limiting behavior of $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$ for $\varepsilon \ll 1$. This will be done by expanding the above integrals for $\varepsilon \ll 1$. First define:

$$I(\alpha_1,\varepsilon) = \int_0^{u^*(\alpha_1,\varepsilon)} \frac{1}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_1}\right)^2 - u^2(1-2u)}}du, \tag{5.56}$$

$$I_1(\alpha_1,\varepsilon) = \int_0^{u_c(\varepsilon)} \frac{1}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_1}\right)^2 - u^2(1-2u)}}du,$$

$$I_2(\alpha_2,\varepsilon) = \int_0^{u_b(\varepsilon)} \frac{1}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_2}\right)^2 - u^2(1-2u)}}du$$

so that (5.54)-(5.55) become:

$$2I - I_1 = \pi, \tag{5.57}$$

$$I_2 = \pi \tag{5.58}$$

Calculation of α_1 Consider (5.57) first. For the integrals I and I_1, it will be useful to make the substitutions $v = u/u^*$ and $w = u/u_c$, respectively, so that the integrals become:

$$I(\alpha_1, \varepsilon) = \int_0^1 \frac{u^*}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_1}\right)^2 - (u^* v)^2(1 - 2u^* v)}} dv \tag{5.59}$$

$$I_1(\alpha_1, \varepsilon) = \int_0^1 \frac{u_c}{\sqrt{\left(\frac{\varepsilon}{\sin\alpha_1}\right)^2 - (u_c w)^2(1 - 2u_c w)}} dw \tag{5.60}$$

Recall that by definition, r^* and r_c satisfy:

$$V_{eff}(r^*) = H^2,$$

$$f(r_c) = 0$$

so that u^* and u_c satisfy

$$V_{eff}\left(\frac{M}{u^*}\right) = H^2,$$

$$f\left(\frac{M}{u_c}\right) = 0$$

Using $V_{eff}(r) = \frac{\sin^2\alpha}{r^2}(1 - \frac{2M}{r})$ and$f(r) - 1 - \frac{2M}{r} - H^2 r^2$, the above two equations become

$$(u^*)^2(1 - 2u^*) = \left(\frac{\varepsilon}{\sin\alpha_1}\right)^2, \tag{5.61}$$

$$u_c^2(1 - 2u_c) = \varepsilon^2 \Rightarrow u_c = \varepsilon + \mathscr{O}(\varepsilon^2) \tag{5.62}$$

Using (5.61) in (5.59)-(5.60) to eliminate α_1 and ε in favor of u^*, the integrals become:

$$I(u^*) = \int_0^1 \frac{1}{\sqrt{1-2u^*-v^2(1-2u^*v)}}dv \tag{5.63}$$

$$I_1(u^*,u_c) = \frac{u_c}{u^*}\int_0^1 \frac{1}{\sqrt{1-2u^*-\left(\frac{u_c}{u^*}\right)^2 w^2(1-2u_c w)}}dw \tag{5.64}$$

where we have written $I = I(u^*)$ and $I_1 = I_1(u^*,u_c)$ to emphasize that these can be viewed as functions of u^* and u_c instead of α_1 and ε. For the purpose of expanding the above integrals for $\varepsilon \ll 1$, it is easier to treat $u^* \ll 1$ as the small parameter, and furthermore view $u_c = u_c(u^*)$ as a function of u^*, whose behavior is defined implicitly by (5.57). That is:

$$2I(u^*) - I_1(u^*,u_c) = \pi \tag{5.65}$$

Once $u_c(u^*)$ for $u^* \ll 1$ is determined using the above, the behavior of $\alpha_1(\varepsilon)$ for $\varepsilon \ll 1$ can then in turn be obtained from (5.61)-(5.62). We start by expanding $u_c(u^*)$ in a Taylor series:

$$u_c(u^*) = u_c(0) + u_c'(0)u^* + \frac{1}{2}u_c''(0)(u^*)^2 + \mathscr{O}\left((u^*)^3\right)$$

From (5.61)-(5.62), we see that $u_c(0) = 0$, so that the above simplifies to

$$u_c(u^*) = u_c'(0)u^* + \frac{1}{2}u_c''(0)(u^*)^2 + \mathscr{O}\left((u^*)^3\right)$$

Let us expand the integral in (5.64) to lowest order in u^*. We have:

$$\begin{aligned}
1-2u^*-\left(\frac{u_c}{u^*}\right)^2 w^2(1-2u_c w) &= 1-u_c'(0)^2w^2+\mathscr{O}(u^*) \\
\Rightarrow \sqrt{1-2u^*-\left(\frac{u_c}{u^*}\right)^2 w^2(1-2u_c w)} &= 1-\frac{1}{2}u_c'(0)^2w^2+\mathscr{O}(u^*) \\
\Rightarrow \left(1-2u^*-\left(\frac{u_c}{u^*}\right)^2 w^2(1-2u_c w)\right)^{-\frac{1}{2}} &= 1+\frac{1}{2}u_c'(0)^2w^2+\mathscr{O}(u^*) \\
\Rightarrow \int_0^1 \frac{1}{\sqrt{1-2u^*-\left(\frac{u_c}{u^*}\right)^2 w^2(1-2u_c w)}}dw &= 1+\frac{1}{6}u_c'(0)+\mathscr{O}(u^*)
\end{aligned}$$

Therefore, to lowest order in u^*, (5.64) is

$$\begin{aligned} I_1(u^*,u_c) &= \left(\frac{u_c'(0)u^* + \mathscr{O}\left((u^*)^2\right)}{u^*}\right)\left(1+\frac{1}{6}u_c'(0)+\mathscr{O}(u^*)\right) \\ &= u_c'(0)\left(1+\frac{1}{6}u_c'(0)\right)+\mathscr{O}(u^*) \end{aligned}$$

To lowest order in u^*, (5.63) is

$$I(u^*) = I(0) + \mathscr{O}(u^*)$$

Using the above expansions for I and I_1 in (5.65), this equation becomes, to lowest order

$$2I(0) + u_c'(0)\left(1+\frac{1}{6}u_c'(0)\right)+\mathscr{O}(u^*) = \pi \tag{5.66}$$

Computing

$$I(0) = \int_0^1 \frac{1}{\sqrt{1-v^2}}dv = \frac{\pi}{2}$$

and substituting into (5.66), we find that equating lowest order terms gives $u_c'(0) = 0$. Our Taylor expansion for $u_c(u^*)$ now reads:

$$u_c(u^*) = \frac{1}{2}u_c''(0)(u^*)^2 + \mathscr{O}\left((u^*)^3\right)$$

We now consider (5.66) at first order in u^*. To first order in u^*, (5.64) is

$$\begin{aligned} I_1(u^*,u_c) &= \left(u_c'(0)+\frac{1}{2}u_c''(0)u^* + \mathscr{O}\left((u^*)^2\right)\right)\left(1+\frac{1}{6}u_c'(0)+\mathscr{O}(u^*)\right) \\ &= \left(\frac{1}{2}u_c''(0)u^* + \mathscr{O}\left((u^*)^2\right)\right)(1+\mathscr{O}(u^*)) \\ &= \frac{1}{2}u_c''(0)u^* + \mathscr{O}\left((u^*)^2\right) \end{aligned}$$

and to first order in u^*, (5.63) is

$$I(u^*) = I(0) + I'(0)u^* + \mathscr{O}\left((u^*)^2\right)$$

Using the above expansions for I and I_1 in (5.57), we get:

$$2I(0)+2I'(0)u^*-\frac{1}{2}u_c''(0)u^*+\mathscr{O}\left((u^*)^2\right)=\pi$$

Equating terms that are first order in u^*, we get:

$$u_c''(0)=4I'(0)$$

$I'(0)$ can be computed to be:

$$I'(0)=\int_0^1 \frac{1-v^3}{(1-v^2)^{\frac{3}{2}}}dv=2$$

so that we finally arrive at our Taylor expansion for $u_c(u^*)$:

$$\begin{aligned} u_c(u^*) &= 2I'(0)(u^*)^2+\mathscr{O}\left((u^*)^3\right) \\ &= 4(u^*)^2+\mathscr{O}\left((u^*)^3\right) \end{aligned}$$

We now have the lowest order behavior of $u_c(u^*)$. From it we will obtain the lowest order behavior of $\alpha_1(\varepsilon)$. First we substitute the above in (5.62) to get:

$$16\,(u^*)^4+\mathscr{O}\left((u^*)^6\right)=\varepsilon^2$$

Let

$$\bar{u}=(u^*)^2 \tag{5.67}$$

so that the above becomes:

$$16\bar{u}^2+\mathscr{O}\left(\bar{u}^3\right)=\varepsilon^2 \tag{5.68}$$

Substituting a Taylor expansion for $\bar{u}(\varepsilon)$:

$$\bar{u}(\varepsilon)=\bar{u}'(0)\varepsilon+\mathscr{O}(\varepsilon^2) \tag{5.69}$$

into (5.68), we get, to leading order:

$$16\bar{u}'(0)^2=1$$

So that (5.69) becomes:

$$\bar{u}(\varepsilon) = \frac{1}{4}\varepsilon + \mathscr{O}(\varepsilon^2)$$

Going to back to (5.67) and solving for u^* gives the lowest order behavior of $u^*(\varepsilon)$:

$$u^*(\varepsilon) = \frac{1}{2}\varepsilon^{\frac{1}{2}} + \mathscr{O}\left(\varepsilon^{\frac{3}{2}}\right)$$

Substituting the above in (5.61) and solving for $\sin\alpha_1$, we get, to leading order:

$$\sin\alpha_1 = 2\varepsilon^{\frac{1}{2}} + \mathscr{O}(\varepsilon)$$

Solving for $\alpha_1(\varepsilon)$ in the above and retaining only leading order terms, we finally get:

$$\alpha_1(\varepsilon) = \pi - 2\varepsilon^{\frac{1}{2}} + \mathscr{O}(\varepsilon) \tag{5.70}$$

Calculation of α_2 Next we turn to finding $\alpha_2(\varepsilon)$ for $\varepsilon \ll 1$. It will be useful to introduce the function $s_2(\varepsilon)$, defined as:

$$s_2(\varepsilon) = \left(\frac{\varepsilon}{\sin\alpha_2}\right)^2 \tag{5.71}$$

The integral (5.56) can now be viewed as a function of s_2 and ε instead of α_2 and ε. That is:

$$I_2(s_2,\varepsilon) = \int_0^{u_b(\varepsilon)} \frac{1}{\sqrt{s_2 - u^2(1-2u)}}du$$

Equation (5.58) now reads:

$$I_2(s_2,\varepsilon) = \pi \tag{5.72}$$

The above implicitly defines the function $s_2(\varepsilon)$. Substituting a Taylor expansion for $s_2(\varepsilon)$:

$$s_2(\varepsilon) = s_2(0) + \mathscr{O}(\varepsilon) \tag{5.73}$$

into (5.72) and using $u_b(\varepsilon) = \frac{1}{2} + \mathscr{O}(\varepsilon)$, we find, to leading order:

$$\int_0^{\frac{1}{2}} \frac{1}{\sqrt{s_2(0) - u^2(1-2u)}} du = \pi$$

The above can be solved numerically to give:

$$s_2(0) = 0.05033384242... \tag{5.74}$$

Substituting (5.73) into (5.71) and solving for $\sin\alpha_2$, we get, to leading order:

$$\sin\alpha_2 = \left(\frac{1}{s_2(0)^{1/2}}\right)\varepsilon + \mathscr{O}(\varepsilon^2)$$

so that to lowest order, $\alpha_2(\varepsilon)$ is given by:

$$\alpha_2(\varepsilon) = \pi - \alpha_2'(0)\varepsilon + \mathscr{O}(\varepsilon^2) \tag{5.75}$$

where $\alpha_2'(0)$ is given by

$$\alpha_2'(0) = \frac{1}{s_2(0)^{1/2}} = 4.457280417...$$

The value of $s_2(0)$ in the above was obtained from (5.74).

Horizon area increase from new generators

Using the formulae for $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$ in (5.70) and (5.75), we can find an approximate formula for $A_{ng}(\infty)$, where $A_{ng}(T)$ is the area at time T due to the joining of new horizon generators (this was defined in section 5.2.3). That is, according to our definitions of α_1 and α_2 from section 5.2.3, $A_{ng}(T)$ is the area of the portion of the horizon made up of generators with parameter α satisfying $\alpha_1 \leq \alpha \leq \alpha_2$ or $\alpha_1 \leq 2\pi - \alpha \leq \alpha_2$.

We want to find the $A_{ng}(T)$ in the limit that $T \to \infty$, and thus will first need the coordinates $r(T;\alpha)$

and $\phi(T;\alpha)$ of the generators in this limit. The formula (5.4) can then be used to find the area element, and the area element can be integrated from α_1 to α_2 to find $A_{ng}(T)$ in the limit that $T \to \infty$. Recall from (4.69) and (4.79) that we have:

$$\begin{aligned} r(\varepsilon) &= \frac{2}{H}\frac{1}{\varepsilon} + \mathscr{O}(1) \\ \phi(\varepsilon) &= \frac{\sin\alpha}{2}\varepsilon + \mathscr{O}(\varepsilon^2) \end{aligned}$$

where the limit $\varepsilon \to 0$ corresponds to $T \to \infty$. Substituting the above into (5.4), we obtain the following lowest order behavior of the area element, in the limit $T \to \infty$:

$$dA \sim \frac{\sin\alpha}{H^2} d\alpha d\beta \tag{5.76}$$

Integrating the above over the intervals $\alpha_1 \leq \alpha \leq \alpha_2$ and $\alpha_1 \leq 2\pi - \alpha \leq \alpha_2$, as well as $0 \leq \beta < \pi$, we obtain $A_{ng}(\infty)$:

$$A_{ng}(\infty) = \frac{2\pi}{H^2}\left(-\cos\alpha_2 + \cos\alpha_1\right)$$

Now we substitute the formulas for $\alpha_1(\varepsilon)$ and $\alpha_2(\varepsilon)$ in (5.70) and (5.75) into the above to get:

$$A_{ng}(\infty) = \frac{2\pi}{H^2}\left[\cos\left(\alpha_2'(0)\varepsilon + \mathscr{O}(\varepsilon^2)\right) - \cos\left(2\varepsilon^{\frac{1}{2}} + \mathscr{O}(\varepsilon)\right)\right]$$

Expanding the above and keeping only the lowest order terms, we get:

$$A_{ng}(\infty) = \frac{4\pi}{H^2}\left(\varepsilon + \mathscr{O}(\varepsilon^{3/2})\right) \tag{5.77}$$

It follows from the definition of new generators that $A_{ng}(-\infty) = 0$, so that the above is also the total area increase due to the new generators. That is:

$$\begin{aligned} \Delta A_{ng}(\infty) &= A_{ng}(\infty) - A_{ng}(-\infty) \\ &= \frac{4\pi}{H^2}\left(\varepsilon + \mathscr{O}(\varepsilon^{3/2})\right) \end{aligned} \tag{5.78}$$

Horizon area increase from expansion of existing generators

Recall from section 5.2.3 that $A_{eg}(T)$ was defined to be the area of the horizon due to generators existing on the horizon for all times. That is, $A_{eg}(T)$ is the area of the portion of the horizon made up of generators which are not new generators, and therefore with parameter α not satisfying $\alpha_1 \leq \alpha \leq \alpha_2$ or $\alpha_1 \leq 2\pi - \alpha \leq \alpha_2$. The two contributions $A_{ng}(\infty)$ and $A_{eg}(\infty)$ must add up to the total area at late times $A(\infty)$. That is, we must have:

$$A(\infty) = A_{ng}(\infty) + A_{eg}(\infty) \tag{5.79}$$

$A(\infty)$ can be found be integrating (5.76) over $0 \leq \alpha < 2\pi$ and $0 \leq \beta < \pi$. This gives:

$$A(\infty) = \frac{4\pi}{H^2}$$

This is precisely what one would expect, since it is the area of a sphere with radius $1/H$, which is the expected horizon shape at late times. Using the above and (5.77) in (5.79), we can obtain $A_{eg}(\infty)$:

$$A_{eg}(\infty) = \frac{4\pi}{H^2}\left(1 - \varepsilon + \mathscr{O}(\varepsilon^{3/2})\right) \tag{5.80}$$

At early times, we can also separate the area into two contributions and write:

$$A(-\infty) = A_{ng}(-\infty) + A_{eg}(-\infty) \tag{5.81}$$

From section 4.4.2 we know that all generators which make up the horizon start at either $r = r_b$ or $r = r_c$, so that the horizon at early times consists of the spheres $r = r_b$ and $r = r_c$. Therefore we have:

$$A(-\infty) = 4\pi\left(r_b^2 + r_c^2\right)$$

Furthermore, it follows from the definition of new generators that $A_{ng}(-\infty) = 0$. Using this and the above in (5.79), we get:

$$A_{eg}(-\infty) = 4\pi\left(r_b^2 + r_c^2\right)$$

Using the above and (5.80), we can calculate the total area increase due to the expansion of existing generators:

$$\begin{aligned}
\Delta A_{eg}(\infty) &= A_{eg}(\infty) - A_{eg}(-\infty) \\
&= \frac{4\pi}{H^2}\left(1 - (Hr_b)^2 - (Hr_c)^2 - \varepsilon + \mathscr{O}(\varepsilon^{3/2})\right) \\
&= \frac{4\pi}{H^2}\left(\varepsilon + \mathscr{O}(\varepsilon^{3/2})\right)
\end{aligned}$$

where we have used used (2.15) in the last step. Combining the above with (5.78), we see that we have:

$$\lim_{T\to\infty} \frac{\Delta A_{eg}(\infty)}{\Delta A_{ng}(\infty)} = 1$$

The above is the main result of this section.

Chapter 6

Concluding Remarks

6.1 Summary and discussion of results

By constructing a new coordinate system for SdS spacetime, and then solving the equations for the null generators of the horizon, we have been able to reveal that the spacetime can provide a rare example of an analytically known spacetime which has merging horizons. Such an example can then be used as a mathematical laboratory to investigate various questions about the structure and area merging horizons. Here we have focused on three mathematical aspects of the merging of horizons: the location and structure of the merger point, the shape of the horizon at late times, and the horizon area.

Our motivation for studying the location and structure of the merger point stems from the fact that by using an analytical metric, we have a unique opportunity to make analytical claims about the location and structure of the merger point. Furthermore, many of these claims can be formulated in a coordinate independent manner. We were able to find an analytical formula for the location of the caustic in the limit of infinitesimally small black hole mass (i.e. the limit $\varepsilon \to 0$). Although it has long been known from numerical simulations [23] that such a caustic point exists (for example, in binary black hole mergers) our study is only the second one which is analytical (the first being [14]), and the first one where the spacetime is known exactly. Furthermore, by focusing on the limit $\varepsilon \to 0$ and by using coordinate systems which have a simple geometric interpretation, we were

able to formulate results in a coordinate independent manner. This is something that would not be possible without the luxury of having analytical knowledge of the spacetime, as was the case in this thesis.

Our initial motivation for studying the late time behavior of the horizons was that it provided the initial conditions for the equations describing the coordinates of the null geodesic generators that make up the horizon. We were able to develop an asymptotic series for the late time behavior of the generators. Although this series is very accurate, it is unlikely that it is an exact solution, as was discussed in section 4.4.3. We hypothesized that the discrepancy between our asymptotic solution and the exact solution is a correction that is vanishingly small at all orders, and whose main importance is in determining the behavior at late times of the horizon area. If this hypothesis is true, it raises the natural question of whether it could be extended to other spacetimes where a horizon setttles down to a final shape, or whether it is specific to SdS spacetime. Therefore, although we initially analyzed the behavior of null generators at late times in order to accurately calculate the coordinates of these generators, we may have accidently stumbled upon a hidden mathematical richness in the behavior of these generators.

The third aspect of the mathematics of horizons that we considered was the horizon area. This is the most physically motivated category of mathematical questions, given that horizon area is directly proportional to horizon entropy, as well as being proportional to mass squared in the case of a Schwarzschild black hole. Our first main result involves an intriguing connection between the merger time and the time at which the rate of change of area is maximal. We provided compelling evidence that these times coincide, and that this coincidence is not merely numerical, nor an artifact of our choice of coordinates. Instead it stems from the fact that the caustic structure before merger results in a diverging positive second derivative of area with respect to time. When this is combined with a negative second derivative after merger, then it necessarily implies that the rate of change of area with respect to time must be maximal at merger. It seems to be the case that this second derivative is negative for a broad class of coordinate systems, so that the result is quite general. Our result reveals a deep connection between the rate of change of horizon area and the presence of a horizon caustic during the merger process. Notice that this connection is not immediately obvious,

since the area increase during the merging process occurs everywhere along the horizon (i.e. all the generators already on the horizon are expanding). Furthermore, the location of local maximal area increase does not occur along the caustic, so that this connection cannot be reduced to simply analyzing the area increase in the vicinity of the caustic. Nevertheless, the diverging discontinuity in the second derivative of $A(T)$ ensures that we can firmly link the time of maximal area increase with the merger time.

The second main result regarding the horizon area states that in the limit $\varepsilon \to 0$, the contribution to the area increase from expanding generators is precisely equal to the contribution to the area increase from the joining of new generators not previously on the horizon. This result is coordinate independent in the sense that there are canonical constant time slicings that one can use to measure area at early and late times. That is, the result holds in the broad class of coordinate systems which use these canonical slicings at early and late times. This canonical slicing at late times is the one which approximates the slicing of deSitter spacetime (recall from the discussion in section 2.6 that SdS spacetime reduces to deSitter spacetime in the $r \to \infty$ limit). As was discussed in section 4.4.3, in general our slicing does not approach this deSitter slicing at late times. However, in the limit that $\varepsilon \to 0$, our slicing does reduce to the deSitter slicing. This can be seen as a consequence of the 3-cone geometry reducing to an essentially flat geometry in the limit that $\varepsilon \to 0$ (see section 3.3.6 for a discussion of the 3-cone geometry). Whether our result regarding the equal contributions of generators to the area is a mathematical coincidence or is connected to a deeper principle is still unclear. One way to begin to answer this question would be to investigate this same question in other spacetimes with merging horizons, such as extreme-mass ratio binary black holes. This possibly is discussed below.

The third and final set of results regarding the area of horizons are those regarding the behavior of the $A(T)$ graph in the limit $\varepsilon \to 0$. More specifically, we investigated three questions in this limit: (i) What is the fractional area increase before merger? (ii) What is the maximal rate of area increase? (iii) What is the average rate of area increase over some suitable time interval?

Regarding (i), we were able to present numerical evidence that in the limit $\varepsilon \to 0$, all of the area

increase occurs before merger. Roughly speaking, this result is essentially stating that at merger time, the area of the two horizons is nearly equal to the area of a spherical cosmological horizon in the absence of a black hole. This suggests that in the limit $\varepsilon \to 0$, at the time of merger the small distortions that occur near the black hole are of negligible importance relative to the area contribution from the rest of the horizon. In light of this interpretation of the result, it is natural to expect that it would hold in other coordinate systems as well. That is, any coordinate system in which the slightly distorted cosmological horizon has area close to $4\pi H^{-2}$ would yield the same result. In this sense we expect our result to be quite general and not restricted to the LP time slicing chosen in this thesis. In terms of applications, our result could be useful when calculating cosmological horizon entropy.

Questions (ii) and (iii) yield the same qualitative answer: the maximum rate of change of area and the average rate of change of area both scale as the black hole mass. The constant of proportionality in the two cases is different of course, but the scaling is the same. Whether this scaling result holds in other coordinates is unclear. Suppose that it can be shown that in the limit $\varepsilon \to 0$, the area increase is concentrated along a specific part of the horizon. In this case it is conceivable that any coordinate system with constant time slices that cross this part of the horizon in a non-pathological manner would yield the same scaling relation. The specific details of the choice of time slices would only affect the proportionality constant in the scaling. We suspect that our result is generic in this sense, but the only way to definitively answer the question would be to find an analytical description of the horizon area in the limit $\varepsilon \to 0$. Such a study is a question we are currently investigating. As with question (i) above, questions (ii) and (iii) could have applications in calculating cosmological horizon entropy, or more specifically, the rate of change of cosmological horizon entropy.

6.2 Applications and future directions

In this section we discuss possible applications of the results in this thesis, as well as directions for future research.

6.2.1 Mathematical questions

One of the main themes of this thesis has been to use SdS spacetime as a mathematical laboratory for exploring questions related to the merging of horizons. By working with a spacetime where the geometry is known analytically, we have the opportunity to approach many of these questions using analytical methods instead of resorting to numerical methods. Furthermore, when using numerical methods we are solving a family of ordinary differential equations, and are able to confidently control the errors involved. Despite these advantages, we resorted to numerical methods instead of analytical methods on several occasions. In addition, these numerical results often had residual computational errors which will require further work to eliminate.

Numerical issues

The first issue one could address would be to reduce the magnitude of these computational errors. This could be done in one of three ways: either by improving the numerical accuracy of the solutions to the differential equations used in this thesis, or by using more null generators to approximate the horizon, or by improving the algorithm used to compute the area of the horizons. In order control the accuracy of the solution to the differential equations in this thesis, the key ingredient is to use an algorithm that can deal with the behavior of the generators near the "initial" time $T = \infty$ (or $\hat{T} = 1^-$ in compactified coordinates). We chose to solve for the function $r(\hat{T})$, and use the built-in algorithm of the MAPLE 14 software, with the option "stiff=true" to indicate that we are dealing with a stiff system near $\hat{T} = 1^-$. We found that this provided a more accurate solution than solving for $\hat{r}(\hat{T})$, where $\hat{r}$ is a compactification of the Schwarzschild radius r. It also provided a more efficient means of computing than solving for $r(T)$, since compactification of the time variable is essential to reducing the runtime of the integration. However, there may be better choices of variables and/or algorithms, in the sense that they would provide more efficient and/or accurate solutions. This is a possibility which we are currently exploring.

In addition to the choices made for both the algorithm and variables, there are two additional possible soures of error in computing the solution $r(T)$. The first is due to the fact that the differential equations are singular at $\hat{T} = 1$, so that we are forced to set initial conditions at $\hat{T}_0 = 1 - \varepsilon_0$ for some $\varepsilon \ll 1$. Such initial conditions must be set using an analytical approximation to the solution, and here we have used an asymptotic series truncated after a finite number of terms. There is obviously a choice that must be made regarding the number of terms in the series and the value of the constant ε_0. Here we have chosen to keep eleven terms in the series and use the value $\varepsilon_0 = 10^{-10}$. However, it may be possible to use a smaller value of ε_0 and keep more terms in the series. The number of terms in the series is limited only by the computation time involved in calculating the series using a symbolic manipulation software such as MAPLE. On the other hand, the value of ε_0 is limited to sufficiently large values for more fundamental reasons. Using a value of ε_0 which is too small would result in a loss of significance due to rounding errors. A careful choice of variables can help partly eliminate this problem. For example, compactifying the Schwarzschild radius r such that $r \in (0, \infty)$ becomes $\hat{r} \in (0, 1)$ would result in a loss of significance when rounding errors cause $\hat{r} \approx 1$ to become $\hat{r} = 1$. On the other hand, compactifying variables such that $r = \infty$ gets mapped to $\delta = 0$ would not have such a problem. In this thesis, we have chosen not to compactify the Schwarzschild radius r, for reasons which are discussed in the next paragraph. For such an uncompactified choice of variables, the initial condition for $r(\hat{T})$ is:

$$r(\hat{T}_0) = \frac{r_0}{\varepsilon_0} + r_1 + r_2\varepsilon_0 + \ldots$$

As we can see, the difficulty with the above is that for sufficiently small values of ε_0, rounding due to finite digits will cause the second and third term in the series to be ignored. Again, compactifying variables such that $r = \infty$ gets mapped to $\delta = 0$ would not have such a problem. Determining the ideal choice of variables in order to minimize these types of rounding errors is an issue we are currently working on.

The second additional source of error in computing the solution $r(T)$ comes from the compactification of the time variable T. It is necessary to compactify the time variable in order to reduce the runtime of the computation. The disadvantage of the compactification comes when attempting to

return to the original time variable T by applying the inverse of the mapping used to compactify the time variable. Such a mapping takes small differences in the compactified time variable $\hat{T}$ and magnifies them greatly. Similarly, small rounding errors in the variable $\hat{T}$ get magnified to large errors in the original variable T. This poses a problem when calculating derivatives with respect to T, since the error in a small interval of time $\Delta\hat{T}$ becomes a large error in the interval ΔT, which can affect the accuracy of the denominator in a discretized formula for the derivative (e.g. the derivative of area with respect to T). The same problem arises when compactifying the variable r and for this reason, we chose to only compactify the time variable. One possible solution to these difficulties is to always work with the compactified variables, and never return to the original variables. This is a possibility which we are currently exploring.

As mentioned above, a simpler way of achieving more accuracy in our approximation of the horizons would be to simply use more horizon generators. Apart from the issues discussed above in choosing an algorithm used to calculate the horizon generators, this comes down to our choice of programming language, as well as sheer computational power. Here we have used the MAPLE software, but there are clearly more efficient programming languages or software, such as Matlab. As well, we typically chose to compute four hundred or so horizon generators, but could compute many more by using a computer cluster instead of a personal computer.

The last improvement which could be made to our numerical computations would be a more accurate calculation of the horizon area. The horizon area formula can be well approximated and computed using Simpson's rule for integration, provided one has enough mesh points for integration (i.e. enough horizon generators). Thus the main source of error in the area computation is related to the number of generators used, as well as the error in the calculation of the horizon generators. By choosing more horizon generators, we would improve the area calculations. Furthermore, by choosing to compute more horizon generators at points on the horizon where the area is changing rapidly we would also achieve more accuracy. In this thesis we have used a uniform mesh of horizon generators, where the uniformity is over the parameter $\alpha \in [0, 2\pi)$. However, ideally we would use a non-uniform mesh with more generators near important points on the horizon, such as near the black hole event horizon or near the merger and caustic points. This possibility of using a

non-uniform mesh is an issue we are currently exploring.

Caustic structure

One of the avenues that was not fully explored analytically in this thesis is the caustic structure near the merger point of the horizons. Although we were able to illustrate this caustic structure using figure 4.5 and confirm its basic shape, we have not determined the quantitative properties of the merger point geometry, nor have we analytically proven that is has the shape shown in figure 4.5. Ideally one would want the description of the caustic structure to have as little dependence on the choice of coordinates as possible. One coordinate independent quantity that one can calculate is the curvature of the spacelike curve of caustic points (the "inner seam" of the surface in figure 4.5). To have a full description of the horizon structure near the merger point, one would ideally want to describe the surface in the neighborhood of the singular point in terms of a canonical normal form representation, as commonly used in the singularity theory of surfaces. Although such a normal form representation would inevitably have coordinate dependent aspects, the ambiguity associated with the choice of coordinates could be partly eliminated by using a canonical coordinate system, such as Riemann normal coordinates.

Horizon area

In terms of horizon area, the main direction for future work would be to confirm many of the numerical results in this thesis using analytical methods. Two results which may be amenable to analytical methods are the late time behavior of the horizon area and the horizon area in the limit $\varepsilon \to 0$. As was discussed in section 4.4.3, finding an analytical description of the horizon area at late times most likely involves uncovering vanishingly small terms in the series approximation for the late time behavior of the horizon generators.

In order to support our numerical results regarding the horizon area in the limit $\varepsilon \to 0$ presented in section 5.2.3, we would need an approximate analytical solution to the equations for the null generators in the limit $\varepsilon \to 0$. It is possible that such a solution could be obtained using a perturbation series. Such a perturbation series would most likely be a singular asymptotic series, since the case of $\varepsilon = 0$ (i.e. no black hole) the behavior of some of the generators is qualitatively different than in the case where $\varepsilon \neq 0$. Regardless of how small the black hole is, provided that $\varepsilon \neq 0$, there will always be generators that are strongly lensed by the black hole. Whether it is necessary to find an asymptotic expansion for the behavior of these strongly lensed generators depends partly on whether or not they make a significant contribution to the horizon area. With such an analytical formula, we could confirm many of the results from section 5.2.3, such as the idea that in the limit $\varepsilon \to 0$, all of the area increase occurs prior to merger. Furthermore, such results could potentially be generalized to other coordinate systems, provided one knows the explicit mapping between these other coordinates and the coordinates used in this thesis. Lastly, if the area increase occurs at a specific location on the horizon (ex. near the caustic points), then it may be possible to state the qualitative aspects of the results in a coordinate independent manner.

6.2.2 Binary black hole mergers

One of the primary motivations for the study of merging horizons undertaken in this thesis is its potential for stimulating questions about the merging horizons that occur in binary black hole mergers. It has long been known that the horizons in binary black hole mergers have the familiar "trousers" shape (see figure 4.1). In this thesis, we have created a similar picture (see figure 4.5). There are two important differences between our picture and the one for binary black holes. The first is that in the case of the cosmological horizon merging with a black hole, one horizon is contained within the other. Despite this difference, the two cases are similar topologically, in the sense that in both cases we have two disconnected surfaces of spherical topology joining to later form a single surface of spherical topology. The second main difference is that in the merging horizons considered in this thesis, the spacetime is known analytically. This is the crucial difference for the purpose of this thesis. Since the spacetime is known analytically, the only source of error comes directly from the

calculation of the horizon, and not from the knowledge of the spacetime. This error is only limited by the accuracy with which one can solve the ordinary differential equations for null generators, as well as the accuracy with which one can calculate quantities based on these coordinates, such as the area of the horizons. For both of these types of computations, the errors are well understood and easily controlled. This is in contrast to a case where one does not know the spacetime analytically, such as in the case of binary black holes. In these cases there are two methods for approximating the spacetime. One can use the initial value formulation of Einstein's equation, where one numerically solves a set of nonlinear coupled partial differential equations for the spacetime metric components. Alternatively, one can approximate the spacetime of an extreme mass ratio binary black hole by using a perturbation method. In both such computations one faces the difficulty of having to choose an appropriate gauge (i.e. coordinate system), and of controlling the sources of error in the spacetime metric.

Caustic location

All of the results about the caustic obtained in this thesis lead naturally to similar questions in binary black hole mergers. For example, we obtained the formula

$$r \approx 5.25M \tag{6.1}$$

for the Schwarzschild radius of the merger point in the limit $\varepsilon \to 0$. The same method that was used to find this formula could also be used to find a similar formula in the case of an extreme mass ratio binary black hole merger. The key to performing such a calculation is to realize that by the equivalence principle, the event horizon of the larger black hole can be approximated as a Rindler horizon. Thus, if we focus attention on the merger point, the merging of a black hole with the Rindler horizon of an accelerated observer has the same local caustic structure as for a binary black hole merger in the extreme-mass ratio limit. Furthermore, to lowest order we can ignore the tidal effects of the larger black hole and treat the spacetime as being only curved due to the smaller black hole. For the head-on of collision of non-rotating black holes, this leads to

considering the Rindler horizon of an accelerated observer in Schwarzchild spacetime. Using this Rindler approximation, we have performed the calculation leading to the location of the merger point and found the Schwarzschild radius to be:

$$r \approx 3.52M \tag{6.2}$$

One might have expected this value for the Schwarzshild radius to be the same as (6.1). This expectation would be based on the fact that one can use the equivalence principle to argue that the cosmological horizon is locally a Rindler horizon, much like in the case of extreme mass ratio binary black holes. Despite this similarity, we have the quantititavely different results (6.1) and (6.2).

Caustic structure

As discussed in section 6.2.1, one of the key quantities that one can calculate is the curvatuve of the spacelike curve of caustic points on the "inseam" of figure 4.5. Similarly, in the case of a head-on collision of black holes in the extreme mass ratio limit, one can use the Rindler approximation described in the previous paragraph to find the curvature of this "inseam" curve . We have performed such a calculation and found the value for the curvature to be:

$$k \approx \frac{118}{M} \tag{6.3}$$

Notice that the curvature has dimensions of inverse length, as one would expect. Another way of describing the curvature of the inseam is by using a length scale, which would be given by the inverse of the above curvature. One can also calculate the total proper length of the inseam. Given that the inseam extends all the way to $t = -\infty$, it is perhaps surprising that this proper length turns out to be a finite quantity. This can be understood as arising from the fact that although the inseam is a spacelike curve, it approaches a lightlike curve as one moves away from the merger point. In [14], the authors use a perturbation method to approximate the spacetime of an extreme mass ratio binary black hole system, and then use this approximate spacetime to analyze the caustic structure

and area of the horizons. One of the quantities they calculate is the total proper length of the inseam curve.

The results (6.2) and (6.3) are part of a publication we are currently preparing. This publication will focus on using the Rindler approximation to analyze the caustic structure near the merger point in a head on binary black hole merger. This Rindler approximation will compliment the work of [14], and will arrive at some of the same results using a different method. It should be mentioned that the idea of using a Rindler approximation to approximate the horizons of an extreme mass ratio binary black hole system has already been partly developed in the thesis [13]. However, our results (6.2) and (6.3) and the work that will follow will go beyond the analysis in that thesis.

The Rindler approximation is appealing for at least two reasons. First, the calculations which lead to an understanding of the caustic structure are simpler. This is because the spacetime in the Rindler approximation is simply Schwarzschild spacetime, which is known analytically, so that the only mathematical challenge is the integration of the equations for the null generators. Secondly, when using the Rindler approximation we have the possibility of easily generalizing the result to a rotating black hole. This is done by considering a Rindler horizon merging with a Kerr black hole. On the other hand, when using perturbation methods, there is no clearly established method for perturbing a large non-rotating black hole by a small rotating black hole, or perturbing a large rotating black hole by a small non-rotating black hole. In the Rindler approximation, it is easy to incorporate the spin of the smaller black hole. Incorporating the spin of the larger black hole is more difficult, but perhaps possible.

On the other hand, there are some advantages to using a perturbation method, as was done in [14]. The first is that one can naturally deal with non-radially plunging orbits. By contrast, when using the Rindler approximation, it is not as obvious how to deal with such orbits. The second advantage is that one can deal with finite but small ratios of black hole masses, whereas in the Rindler approximation one is essentially taking the mass of the larger black hole to be infinite. However, as will be explained below, it may be possible to introduce small corrections to the Rindler approximation to take into account a large but finite mass for the larger black hole. Lastly, it is easier to answer

questions about the area of the horizons when using a perturbation method. When using the Rindler approximation, the area is infinite at all times, so that without some additional assumptions about the horizons, questions about area increase are ill-posed.

We have stated that it is not immediately clear how the Rindler approximation can accomodate either the spin of the larger black hole, or the possibility that the smaller black hole moves on a non-radial trajectory. However, it should be mentioned that the challenge is technical and not conceptual. It is clear that the Rindler approximation is exploiting the equivalence principle, which will be valid regardless of the spin of the larger black hole, or the type of geodesic trajectory followed by the smaller black hole. Provided that we ignore the self-force of the smaller black hole, as well as any possible non-gravitational forces, the smaller black hole is in free fall. We can construct a local inertial frame in the vicinity of this freely falling black hole. In the limit that the mass of the larger black hole is infinite, tidal distortions of the spacetime curvature due to this larger black hole can be ignored. We are then left with Schwarzschild spacetime in the vicinity of the smaller black hole, or Kerr spacetime if the smaller black hole is spinning. By focusing on a neighborhood of the smaller black hole, the spacetime geometry is entirely understood and described analytically, at least to lowest order in the mass ratio of the black holes. Understanding how the event horizon of the larger black hole merges with the event horizon of the smaller black hole then reduces to understanding the null generators of the larger black hole, and how these are lensed by the presence of the smaller black hole. The spin of the larger black hole, or the effect of a non-radial trajectory for the smaller black hole, can both be taken into account by adjusting the "initial conditions" of the null generators which make up the large black hole horizon. That is, they are taken into account by adjusting how this family of generators behaves far away from the smaller black hole. For a head-on collision in the case of non-spinning larger black hole, the null generators are a plane of light rays when far from the smaller black hole, as one would normally expect for a Rindler horizon. This plane of light rays is modified when the spin of the larger black hole or the non-radial trajectory of the small black hole is taken into account. By using this Rindler approximation in the case where the spins of the smaller and larger black holes are taken into account, we could potentially reveal a rich caustic structure near the merger point of the horizons. It would be interesting to investigate the topology of this structure, as well as its dependence on the mass of the smaller black hole and spins of the

two black holes.

In addition to generalizing the Rindler approximation to deal with spin and non-radial trajectories, it would also be interesting to generalize it to finite mass ratios. One can imagine approximating the spacetime surrounding the small black hole using a series expansion, with the small parameter being the ratio of the small black hole mass to the large black hole mass. The Rindler approximation is then the lowest order approximation, where the mass ratio of the black holes is zero. At this level of approximation the spacetime surrounding the small black hole is given by the Schwarzschild or Kerr metric, and the horizon of the large black hole is the Rindler horizon of an observer accelerating uniformly away from the black hole. At the next order of approximation, there would be a lowest order correction to the Schwarzschild or Kerr metric which would incorporate the tidal distortions to the spacetime surrounding the small black hole, as caused by the larger black hole. There would also be a lowest order correction to the shape of the Rindler horizon due to the large but finite mass of the larger black hole. These small corrections to both the spacetime surrounding the small black hole and the shape of the larger horizon would result in small corrections to the shape of the caustic structure, and to small corrections with associated quantities such as the curvature of the inseam. Unlike the effect of taking into account spin or non-radial trajectories, these small lowest order corrections would be unlikely to affect the topology of the structure of the caustic near the merger point, and so would be of interest for their quantitative rather than qualitative effects. Understanding these quantitative corrections to the caustic structure is a question we are currently working on.

Notice that creating a perturbation series starting from the Rindler approximation, as was described in the previous paragraph, is quite different from the perturbation series that was created in [14]. There the authors use the Regge-Wheeler formalism for perturbing Schwarzschild spacetime and approximate the small black hole as a point mass. The resulting approximate spacetime is not valid near the small point mass. This can be understood as a consequence of the fact that the curvature of the spacetime near the small mass diverges as one gets closer to it. This is in contrast to the Rindler perturbation series discussed in the previous paragraph, which would in fact only be valid near the small black hole. At distance scales comparable to the Schwarzschild radius of the larger black hole, the tidal effects due to the large black hole would no longer be small corrections, and the series

would no longer be valid. These two perturbation approaches, with one valid near the small black hole and one valid far away, could be joined together into a matched asymptotic approximation, with a transition region between the two and a set of matching conditions in the transition region. Such approximations are well known in the context of black hole perturbation theory, but so far have not been used to study the event horizons of the black holes (to our knowledge).

Area of horizons

Recall the three main results from the chapter on the area of merging cosmological and black hole horizons (chapter 5): (i) The time at which the rate of change of horizon area is at its largest value is also the merger time of the horizons. (ii) In the limit $\varepsilon \to 0$, all the horizon area increase takes place before merger. (iii) In the limit $\varepsilon \to 0$, the area increase has equal contributions from the expansion of generators always on the horizon, and the joining of generators not previously on the horizon. These three results regarding horizon area naturally lead to three corresponding hypotheses in the context of binary black holes. The statement of these hypotheses in this new context are nearly identical to (i)-(iii) above, with the only difference being that for results (ii) and (iii), the limit $\varepsilon \to 0$ is replaced with the extreme mass ratio limit of binary black holes.

To our knowledge, only hypothesis (iii) has been shown to be true in the context of binary black holes. In [14], the authors use perturbation theory to approximate the spacetime of an extreme mass ratio binary black hole system. This approximate spacetime is then used to show that in the limit of infinitesimally small black hole mass, the total horizon area increase has equal contributions from both existing and new generators, as in (iii) above. The result in this thesis extends this result of [14] to the context of cosmological horizons. It is intriguing that although the spacetime considered in this thesis and the spacetime of an extreme mass ratio binary black hole are quite different, the horizons in these spacetimes both satisfy (iii). One might attribute this as being due to a similarity in the horizons near the caustic points, since both the cosmological horizon and the large black hole horizon can be locally approximated as a Rindler horizon in the limit of small black hole mass.

However, the increase in horizon area occurs not only near the caustics where new generators enter the horizon, but also far from the caustic points due to the expansion of existing generators. The fact that result (iii) is valid for the merging horizons considered in this thesis and for the horizons in binary black holes is even more surprising when we consider the fact that the horizons in binary black holes are observer independent event horizons, whereas the horizons in this thesis are observer dependent causal horizon. Despite this and other differences between the spacetimes, result (iii) holds for both spacetimes. This points to perhaps a deeper principle at work which may be valid more generally. It could be that whenever a small black hole mergers with a much larger horizon, the area increase can be split into equal contributions from two types of generators, as in (iii). However, we do not at the moment have any reason to believe that this is indeed the case.

It would be interesting to investigate the case of a charged and/or rotating black hole merging with a cosmological horizon, and see whether or not (iii) holds in those cases. The analysis would be similar to that performed in this thesis, but with Kerr-Newman-deSitter spacetime replacing Schwarzschild-deSitter spacetime. To investigate statement (iii) for rotating binary black holes would be more difficult, since there is no established perturbation technique that can be used to approximate the spacetime in those cases. Although the Rindler approximation can deal with rotating black holes, it is not well suited to answering questions about area, since the area of a Rindler horizon is effectively infinite. One can imagine getting around this difficulty of having an infinite area by first taking a finite Rindler horizon (for example, by ignoring the generators with large impact parameter), and then taking the limit as the Rindler horizon area goes to infinity. If one is answering questions about the rate of change of area, as in (i) above, or the relative area increase, as in (ii), then taking this limit could yield sensible answers. This is a direction that we are currently investigating. We are also currently investigating the possibility of extending the techniques in [14] to prove (i) and (ii) for binary black holes. Proving (i) may require the use of a matched asymptotic approximation, where an approximate spacetime near the small black hole is joined to an approximate spacetime valid far from the small black hole. This may be necessary because proving (i) requires taking into account the horizon area both in vicinity of the merger point and far away from the small black hole, so that we are required to have knowledge of the spacetime both close to and far from the small black hole.

Bibliography

[1] Thomas W Baumgarte and Stuart L Shapiro. *Numerical relativity: solving Einstein's equations on the computer*. Cambridge University Press, 2010.

[2] Jacob D Bekenstein. Black holes and entropy. *Physical Review D*, 7(8):2333, 1973.

[3] George David Birkhoff and Rudolph Ernest Langer. *Relativity and Modern Physics. By GD Birkhoff... with the co-operation of Rudolph E. Langer.* Harvard University Press, 1927.

[4] Scott A Caveny and Richard A Matzner. Adaptive event horizon tracking and critical phenomena in binary black hole coalescence. *Physical Review D*, 68(10):104003, 2003.

[5] Tamara M Davis and Charles H Lineweaver. Expanding confusion: common misconceptions of cosmological horizons and the superluminal expansion of the universe. *Publications of the Astronomical Society of Australia*, 21(6):97–109, 2004.

[6] Willem De Sitter. On the relativity of inertia. remarks concerning einstein's latest hypothesis. *Proc. Kon. Ned. Acad. Wet*, 19:1217–1225, 1917.

[7] Chas A. Egan and Charles H. Lineweaver. A larger estimate of the entropy of the universe. *The Astrophysical Journal*, 710:1825–1834, 2010.

[8] Albert Einstein. Die feldgleichungen der gravitation. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin), Seite 844-847.*, 1:844–847, 1915.

[9] Roberto Emparan and Nidal Haddad. Self-similar critical geometries at horizon intersections and mergers. *Journal of High Energy Physics*, 2011(10):1–21, 2011.

[10] Willy Fischler and Leonard Susskind. Holography and cosmology. *arXiv preprint hep-th/9806039*, 11, 1998.

[11] Gary W Gibbons and Stephen W Hawking. Cosmological event horizons, thermodynamics, and particle creation. *Physical Review D*, 15(10):2738, 1977.

[12] Ruth Gregory and Raymond Laflamme. Black strings and p-branes are unstable. *Physical review letters*, 70(19):2837, 1993.

[13] Ryan Hamerly. Event horizon deformations in extreme mass-ratio black hole mergers. Undergraduate honors thesis, Caltech, 2010.

[14] Ryan Hamerly and Yanbei Chen. Event horizon deformations in extreme mass-ratio black hole mergers. *Physical Review D*, 84(12):124015(1–24), 2011.

[15] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge University Press, 1973.

[16] Stephen W Hawking. Gravitational radiation from colliding black holes. *Physical Review Letters*, 26:1344–1346, 1971.

[17] Ted Jacobson and Renaud Parentani. Horizon entropy. *Foundations of Physics*, 33(2):323–348, 2003.

[18] Donald Marolf John T. Giblin Jr. and Robert Garvey. Spacetime embedding diagrams for spherically symmetric black holes. *General Relativity and Gravitation*, 36:83–99, 2004.

[19] David Kastor and Jennie Traschen. Cosmological multi-black-hole solutions. *Physical Review D*, 47(12):5370, 1993.

[20] Sung-Won Kim. Geometrical structure of schwarzschild desitter space-time. *Journal of the Korean Physical Society*, 24(2):118–124, 1991.

[21] M. A. Per L. J. Boya and A. J. Segu. Graphical and kinematical approach to cosmological horizons. *Physical Review D*, 66(6):064009(1–12), 2002.

[22] Georges Lemaitre. L'univers en expansion. *Annales de la Societe Scientifique de Bruxelles*, page 51, 1933.

[23] Richard A Matzner, HE Seidel, Stuart L Shapiro, L Smarr, W-M Suen, Saul A Teukolsky, and Jeffrey Winicour. Geometry of a black hole collision. *Science*, 270(5238):941–947, 1995.

[24] Saul Perlmutter, G Aldering, G Goldhaber, RA Knop, P Nugent, PG Castro, S Deustua, S Fabbro, A Goobar, DE Groom, et al. Measurements of ω and λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565, 1999.

[25] E. Poisson. *A Relativist's Toolkit*. Cambridge University Press, 2004.

[26] Adam G Riess, Alexei V Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M Garnavich, Ron L Gilliland, Craig J Hogan, Saurabh Jha, Robert P Kirshner, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009, 1998.

[27] Adam G Riess, Louis-Gregory Strolger, John Tonry, Stefano Casertano, Henry C Ferguson, Bahram Mobasher, Peter Challis, Alexei V Filippenko, Saurabh Jha, Weidong Li, et al. Type ia supernova discoveries at z> 1 from the hubble space telescope: Evidence for past deceleration and constraints on dark energy evolution. *The Astrophysical Journal*, 607(2):665, 2004.

[28] Kristin Schleich and Donald M. Witt. A simple proof of birkhoff's theorem for cosmological constant. *Journal of Mathematical Physics*, 51:112502(1–9), 2010.

[29] Karl Schwarzschild. Über das gravitationsfeld eines massenpunktes nach der einsteinschen theorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin), 1916, Seite 189-196*, 1:189–196, 1916.

[30] Kristin Schleich Sergei Slobodov and Donald M. Witt. Effects of a cosmological constant on the x-ray fluorescence spectra of black hole accretion disks, 2004.

[31] P C W Davies Tamara M Davis and Charles H Lineweaver. Black hole versus cosmological horizon entropy. *Classical and Quantum Gravity*, 20:2753–2764, 2003.

[32] R. M. Wald. *General Relativity*. University of Chicago Press, 1984.