On a Second Law of Black Hole Mechanics in a Higher Derivative Theory of Gravity

Diplomarbeit

vorgelegt von

Eric Morfa Morales

aus Eisenach

angefertigt am

Institut für Theoretische Physik

der

Georg-August-Universität zu Göttingen

September, 2008

Contents

Introduction

Black holes are perhaps the most fascinating objects in Einstein's general theory of relativity. The discovery of their properties initiated some of the most remarkable developments in theoretical physics in the last thirty five years, by revealing unforseen connections between otherwise distinct areas of physics, such as general relativity, quantum physics and statistical mechanics.

Bekenstein [5] was the first to point out that there might exist a close relationship between certain laws in black hole physics and the laws of thermodynamics. Hawking's area theorem in classical general relativity [12] asserts that the area of a black hole can never decrease in any process. This result bears a resemblance to the second law of thermodynamics, which states that the total entropy of a closed system can never decrease in any process. Furthermore, Bekenstein proposed that the area of a black hole (times a constant of order unity in Planck units) should be interpreted as its physical entropy. Shortly after that, this analogy was reinforced by a systematical analysis by Bardeen, Carter, and Hawking [3]. They found a mass variation formula for two nearby stationary black holes, which bears a resemblance to the first law of thermodynamics. Furthermore, the authors found that the surface gravity κ of a stationary black must be uniform over the event horizon. This result is very similar to the zeroth law of thermodynamics.

However, within the *classical* framework these analogies must be considered to be formal and of a pure mathematical nature. For instance, there is no physical relationship between the surface gravity κ and the classical temperature of a black hole, since it is a perfect absorber. Despite this difficulty, the analogy between the laws of thermodynamics and the laws of black hole mechanics gained a deep *physical* significance, when Hawking discovered [14] that black holes radiate all species of particles to infinity with a perfect black body spectrum at temperature $\kappa/2\pi$. This result left little doubt that a suitable multiple of the area of a black hole must represent the physical entropy of a black hole in general relativity.

Even though, general relativity is a physical theory which is experimentally confirmed to a very high precision, we are still far from a theory which describes the quantum aspects of the gravitational field. Many of the attempts to find such a theory consider modfications to the Einstein theory. In particular, within the context of perturbative quantization of general relativity [8], [26], [28], and the construction of an effective action for string theory [11], one is naturally led to the consideration of gravitational actions, which involve higher derivative terms. Of course, one can ask what the status of the laws of black hole mechanics in such gravitational theories is. To study black hole thermodynamics in such generalized theories of gravity might answer the question if the analogy between the ordinary laws of thermodynamics and the laws governing the behaviour of a black hole is a peculiar accident of general relativity or a robust feature of all generally covariant theories of gravity or something in between. Ultimately, the hope is that one can learn something about the possible nature of quantum gravity from this analysis [19].

In this thesis want to study a theory of gravity whose gravitational Lagrangean contains, in addition to the Einstein-Hilbert term, an additional Ricci-tensor squared contribution. In particular, we interested in the question if a second law can be established in such a theory.

This thesis is organized as follows. In part I, the basic results about black holes in general relativity are summarized. We do this in oder to introduce the necessary terminology and to fix our notation. Throughout this part we will always point out where the Einstein equations are used, such that it becomes transparent which features are peculiar to the Einstein theory. Chapter 1 introduces the basic notions from global lorentzian geometry along with some theorems about the properties of causal boundaries. This will be necessary in order to understand the definition of a black hole properly. Chapter 2 presents the definition of asymptotically flat spacetimes which is motivated by a calculation how Minkowski spacetime can be conformally embedded into the Einstein static universe. Chapter 3 is devoted to the geometry of null hypersurfaces. Special emphasis is placed on the construction of a particulary useful coordinate system (Gaussian null coordinates), which will be extensively used throughout this thesis. Furthermore, we will be concerned with null congruences and the Raychaudhuri equation, which governs their dynamical evolution. This equation will be the key ingedient in the proof of the area theorem. After a phenomenological introduction, the definition of a black hole is presented in chapter 4. General properties of stationary and nonstationary black holes will be discusssed and in particular a proof of the area theorem will be given. In chapter 5 we briefly recall the laws of black hole mechanics according to Bardeen, Carter, and Hawking and discuss their physical relevance.

Part II is devoted to higher derivative theories of gravity (HDTG) and the covariant phase space formalism of Wald and Zoupas. After we have discussed the relevance of such HDTG, we will present the particular theory which is investigated throughout this thesis in chapter 6. Furthermore, the status of the laws of black hole mechanics in such HDTG will be summarized. Chapter 7 is devoted to the covariant phase space formalism of Wald and Zoupas. A derivation of the first law of black hole mechanics within this framework as well as a computation of the black hole entropy in our HDTG will be presented. Chapter 8 summarizes the main results of our own work. We will present two ideas for a proof of a second law of black hole mechanics in the HDTG which we consider. In the first idea we will try to show that the rate of change of the entropy in our HDTG is positive along the null geodesic generators of the horizon. This idea is mainly based on the idea for the proof of the area theorem. The second idea for a proof applies the Wald-Zoupas formalism to the event horizon of a black hole, in order to define a "conserved quantity" on the horizon. This is done for the Einstein theory and the HDTG which we consider.

The appendices cover the following topics. Appendix A summarizes the notations and conventions which we use. Appendix B gives a list of the standard energy conditions in general relativity. Appendix C summarizes some useful relations from tensor analysis. Appendix D presents further details about GNC. The results for the Christoffel symbols are summarized here, and we present additional useful relations which will be extensively used. Furthermore, we prove that the null geodesic generators are hypersurface orthogonal and we show that the uu-component of the Einstein equation (in GNC) corresponds to the Raychaudhuri equation, if the null generators are affinely parametrized. These facts provide the main motivation for the first idea for a proof. Besides this, we compute the pullback of a certain tensor field to a horizon cross-section. This result will be needed in the last part of the analysis.

Part I.

Black Holes in General Relativity

1. Causal Structure of Spacetime

The causal structure of a Lorentzian manifold describes the causal relationships between points in the manifold. These causal relations are interpreted as describing which events can influence other events in the spacetime. In this section we will state the definitions and basic results concerning such causal relationships. This "vocabulary" will be further needed to state the definition of a black hole. For proofs of the theorems and propositions we refer the reader to [29], [15] and [4].

1.1. Preliminaries

Let (M, g_{ab}) be a spacetime. At each $p \in M$, the tangent space T_pM is isomorphic to Minkowski spacetime. As in special relativity, each lightcone, sitting inside T_pM, has two connected components which we arbitrarily label "future" and "past". If such a choice can be made in a continuous manner, as p varies over M, the spacetime is said to be *time orientable*. A timelike or null vector lying in the "future/past half" of the light cone will be called *future/past directed*. In the following we will only consider time orientable spacetimes. An important property satisfied by every time orientable spacetime is expressed in the following

Proposition 1. *Let (M, g_{ab}) be a time orientable spacetime. Then there exists a (nonunique) smooth nonvanishing timelike vector field t^a on M.*

Conversely, if a continuous nonvanishing timelike vector field can be chosen, then (M, g_{ab}) is time orientable.

Definition 1. A differentiable curve $c : I \to M$ is said to be

- *timelike* if its tangent vector is timelike for all $s \in I$
- *null* if its tangent vector is null for all $s \in I$
- *spacelike* if its tangent vector is spacelike for all $s \in I$
- *causal*(or *non-spacelike*) if it is timelike or null.

Definition 2. A causal curve is called

- *future directed* if, its tangent vector is future directed for all $s \in I$
- *past directed* if, its tangent vector is past directed for all $s \in I$.

Definition 3. Let $c : I \to M$ be a future directed causal curve. We say that $p \in M$ is a *future endpoint* of c if for every open neighborhood O of p there exists a t_0 such that $c(t) \in O$ for all $t > t_0$. The curve c is said to be *future inextendible* if it has no future endpoint.

Past inextendibility is defined similary.

1.2. Futures and Pasts

Now we will introduce two types of causal relations between spactime points.

Definition 4. Let $p, q \in M$, then we say that

- p *chronologically precedes* q, denoted $p \prec\!\!\prec q$, if there exists a future directed timelike curve $c : [a,b] \to M$ with $c(a) = p$ and $c(b) = q$
- p *causally precedes* q, denoted $p \prec q$, if there exists a future directed causal curve $c : [a,b] \to M$ with $c(a) = p$ and $c(b) = q$.

These relations are transitive, i.e.

- $p \prec\!\!\prec q$, $q \prec\!\!\prec r$ implies $p \prec\!\!\prec r$
- $p \prec q$, $q \prec r$ implies $p \prec r$

and the following implications hold:

- $p \prec\!\!\prec q$ implies $p \prec q$
- $p \prec\!\!\prec q$, $q \prec r$ implies $p \prec\!\!\prec r$
- $p \prec q$, $q \prec\!\!\prec r$ implies $p \prec\!\!\prec r$.

Definition 5. For $p \in M$ we define

- the *chronological future* of p, denoted $I^+(p)$, as the set of all points $q \in M$ such that p chronologically precedes q, i.e.
$$I^+(p) = \{q \in M \,|\, p \prec\!\!\prec q\} \tag{1.2.1}$$
- the *chronological past* of p, denoted $I^-(p)$, as the set of all points $q \in M$ such that q chronologically precedes p, i.e.
$$I^-(p) = \{q \in M \,|\, q \prec\!\!\prec p\} \tag{1.2.2}$$
- the *causal future* of p, denoted $J^+(p)$, as the set of all points $q \in M$ such that p causally precedes q, i.e.
$$J^+(p) = \{q \in M \,|\, p \prec q\} \tag{1.2.3}$$
- the *causal future* of p, denoted $J^-(p)$, as the set of all points $q \in M$ such that q causally precedes p, i.e.
$$J^-(p) = \{q \in M \,|\, q \prec p\}. \tag{1.2.4}$$

The sets $I^+(p), I^-(p), J^+(p), J^-(p)$ for all $p \in M$ are collectively called the *causal structure* of M.

Definition 6. For any subset $S \subset M$, we define

$$I^\pm(S) = \bigcup_{p \in S} I^\pm(p), \qquad J^\pm(S) = \bigcup_{p \in S} J^\pm(p). \tag{1.2.5}$$

As we see, the sets $I^+(S)$ and $J^+(S)$ represent events that could be influenced by a set S of events.

From the properties of the relations $\ll$ and $\prec$ we mentioned above clearly follows for any $p, q \in M$ and $S \subset M$

- $p \in I^-(q) \Leftrightarrow q \in I^+(p)$
- $p \prec q \Rightarrow I^-(p) \subset I^-(q)$
- $p \prec q \Rightarrow I^+(q) \subset I^+(p)$
- $I^+(S) = I^+(I^+(S)) \subset J^+(S) = J^+(J^+(S))$
- $I^-(S) = I^-(I^-(S)) \subset J^-(S) = J^-(J^-(S))$.

Furthermore we have the following important property:

Proposition 2. *If $q \in J^+(p) \backslash I^+(p)$ with $q \neq p$, then there exists a future directed null geodesic from p to q.*

Note that if p is a point in Minkowski spacetime $(\mathbb{R}^4, \eta_{ab})$, then $I^+(p)$ is open, $J^+(p)$ is closed and $\partial J^+(p) = J^+(p) \setminus I^+(p)$ is just the future null cone at p. $I^+(p)$ consists of all points inside the future null cone, and $J^+(p)$ consists of all points on and inside the future null cone. However, this picture can drastically change when curvature and topology come into play. For instance, if p is a point in a generic spacetime, then $J^+(p)$ does not need to be closed anymore, i.e. $J^+(p) \neq \overline{I^+(p)}$ in general.

Although the situation is more complicated in curved spacetimes, the following properties are still valid for all $p \in M$ and $S \subset M$:

- $I^\pm(p)$ is open
- $I^\pm(S)$ is open
- $\mathrm{int}(J^\pm(S)) = I^+(S)$
- $J^\pm(S) \subset \overline{I^\pm(S)}$

Note that from the above properties follows in particular $\partial J^\pm(S) = \partial I^\pm(S)$.

In section 4.3 we will define the event horizon of a black hole as the boundary of the causal past of a certain region in spacetime. The next theorem assures that this surface is in a certain sense "well behaved".

Definition 7. A subset $S \subset M$ is called *achronal* if there do not exist $p, q \in S$ such that $q \in I^+(p)$, i.e., if $I^+(S) \cap S = \varnothing$.

Theorem 1. *Let (M, g_{ab}) a time orientable spacetime, and let $S \subset M$. Then $\partial I^\pm(S)$ (if nonempty) is an achronal, 3-dimensional, embedded, C^0-submanifold of M.*

Furthermore, causal boundaries are generated by inextendible null geodesics:

Theorem 2. *Let C be a closed subset of the spacetime manifold M. Then every point $p \in \partial I^+(C)$ with $p \notin C$ lies on a null geodesic γ which lies entirely in $\partial I^+(C)$ and either is past inextentible or has a past endpoint on C.*

A similar statement holds for points $p \in \partial I^-(C)$ with "past" replaced by "future" in the above theorem

1.3. Causality Conditions

The Einstein equation admits solutions which contain closed causal curves. It is generally believed that such spacetimes are not physically realistic. If we consider a spacetime where causal curves exist which come arbitrarily close to intersecting themselves, an arbitrarily small perturbation of the spacetime metric could cause again causality violations. These spacetimes seem also physically unreasonable. The following two definition give conditions that assure that such a pathological, acausal behavior does not occur.

Definition 8. A spacetime (M, g_{ab}) is called *strongly causal* if for all $p \in M$ and every neighborhood O of p, there exists a neighborhood V of p contained in O such that no causal curve intersects V more than once.

A useful consequence of strong causality is expressed in the following lemma:

Lemma 1. *Let (M, g_{ab}) be strongly causal and $K \subset M$ compact. Then every causal curve γ confined within K must have past and future endpoints in K.*

There are certain examples of spacetimes which, even though they are strongly causal, are "on the verge" of displaying bad causal behavior in the sense that a small modification of g_{ab} in an arbitrarily small neighborhood of some point would cause causal curves to become closed. The following definition of "stably causal" spacetimes imposes stronger conditions, such that these causal pathologies are ruled out.

Definition 9. A spacetime (M, g_{ab}) is said to be *stably causal* if there exists a continuous nonvanishing timelike vector field t^a such that the spacetime $(M, \bar{g}_{ab})$, with $\bar{g}_{ab} = g_{ab} - t_a t_b$, possesses no closed timelike curves.

The following theorem shows that stable causality is equivalent to the existence of a "global time function".

Theorem 3. *A spacetime (M, g_{ab}) is stably causal if and only if there exists a differentiable function f on M such that $\nabla^a f$ is a past directed timelike vector field.*

As a corollary, we have:

Corollary 1. *Stable causality implies strong causality.*

1.4. Domain of Dependence, Global Hyperbolicity

The notion of global hyperbolicity is of fundamental importance in general relativity, since spacetimes with this property admit a well posed initial value formulation.

Definition 10. Let $S \subset M$ be closed and achronal. We define the *edge* of S as the set of points $p \in S$, such that every open neighborhood O of p contains a point $q \in I^+(p)$, a point $r \in I^-(p)$ and a timelike curve $c : [a, b] \to M$ with $c(a) = r$ and $c(b) = q$ which does not intersect S.

Proposition 3. *Let S be a (nonempty) closed, achronal set with edge$(S) = \varnothing$. Then S is a 3-dimensional, embedded, C^0-submanifold of M.*

Definition 11. Let S be a closed, achronal set (possibly with edge). We define the *future domain of dependence* of S, denoted $D^+(S)$, by

$$D^+(S) = \{p \in S \,|\, \text{every past inextendible causal curve through } p \text{ intersects } S\} \tag{1.4.1}$$

The *past domain of dependence* of S, denoted $D^-(S)$, is defined by interchanging "future" and "past" in (1.4.1). The *domain of dependence* of S, denoted $D(S)$, is defined as

$$D(S) = D^+(S) \cup D^-(S). \tag{1.4.2}$$

Definition 12. A closed, achronal set Σ for which $D(\Sigma) = M$ is called *Cauchy surface*.

Since Cauchy surfaces Σ are achronal, we may think of Σ as representing an "instant of time" throughout the universe.

Definition 13. A spacetime (M, g_{ab}) which possesses a Cauchy surface Σ is said to be *globally hyperbolic*.

A few basic consequences of global hyperbolicity are the following:

Proposition 4. *Let (M, g_{ab}) be a globally hyperbolic spacetime. Then,*

1. *The sets $J^{\pm}(A)$ are closed, for all compact $A \subset M$.*
2. *The sets $J^{\pm}(A) \cap J^{\pm}(B)$ are compact, for all compact $A, B \subset M$.*

Furthermore, we have the important property:

Theorem 4. *Let (M, g_{ab}) be a globally hyperbolic spacetime. Then (M, g_{ab}) is stably causal. Furthermore, a global time function, f, can be chosen such that each surface of constant f is a Cauchy surface. Thus M can be foliated by Cauchy surfaces and the topology of M is $\mathbb{R} \times \Sigma$, where Σ denotes any Cauchy surface.*

2. Conformal Infinity

In order to give a precise definition of a black hole, we need a concept of spacetimes that represent ideally isolated systems. Such systems are represented in general relativity by asymptotically flat spacetimes. Intuitively, such spacetimes have the property that the metric becomes flat at large distances from the source.

The notion of asymptotic flatness was first introduced by Penrose [23], [24] at "null infinity", i.e. as one goes to large distances along null geodesics. Seperately, Geroch [10] gave a definition of asymptotic flatness at "spatial infinity", which was based on earlier work of Arnowitt, Deser and Misner [1]. These two notions were combined into a single notion by Ashtekar and Hansen [2], and in the following we will follow their approach.

The key idea is to use a conformal transformation to bring "infinity" to a "finite distance", or more precisely, to attach suitable boundaries , which represent "points at infinity". This procedure has several advantages: Instead of imposing certain falloff conditions on the spacetime metric in a particular coordinate system, this notion is manifestly coordinate independent. Within this framework it is also possible to define quantities such as the total energy of a spacetime. Furthermore, this technique enables us to represent an entire spacetime in a compact region in a way that preserves the causal structure.

Note that the following exposition will be mainly based on [29].

2.1. Conformal Embedding of Minkowski Spacetime into the Einstein Static Universe

In order to illustrate the key idea, we will consider first of all Minkowski space $(\mathbb{R}^4, \eta_{ab})$. In spherical coordinates $\{t, r, \theta, \varphi\}$ the metric of Minkowski spacetime is given by

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\mathrm{d}\sigma^2, \tag{2.1.1}$$

where $\mathrm{d}\sigma^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2$ is the standard metric on the 2-sphere $\mathbb{S}^2$. We want to analyze the form of the metric "far out", i.e. for large lightlike distances, so it is convenient to introduce the advanced and retarded null coordinates

$$u = t + r, \qquad v = t - r. \tag{2.1.2}$$

In coordinates $\{u, v, \theta, \varphi\}$ the Minkowski metric is given by

$$\mathrm{d}s^2 = -\mathrm{d}u\,\mathrm{d}v + \frac{1}{4}(v-u)^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2). \tag{2.1.3}$$

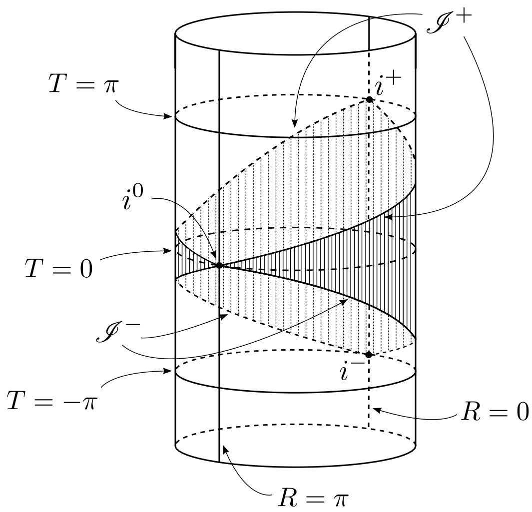

Figure 2.1.: Spacetime diagram of the Einstein static universe. Minkowski spacetime is isometric to the shaded region $O = I^+(i^-) \cap I^-(i^+)$. The (attached) boundary of O defines a precise notion of "infinity" for Minkowski spacetime.

In order to bring "null infinity"($|u|, |v| \to \infty$) to a finite place in our spacetime, we consider the following coordinate transformations[1]

$$V = T + R = \tan^{-1} v, \qquad U = T - R = \tan^{-1} u, \tag{2.1.4}$$

where T and R have ranges restricted by the inequalities

$$-\pi < T + R < \pi, \qquad -\pi < T - R < \pi, \qquad R \geq 0. \tag{2.1.5}$$

In the coordinates $\{T, R, \theta, \varphi\}$ the Minkowski metric is given by

$$\mathrm{d}s^2 = \Omega^{-2}\left[-\mathrm{d}T^2 + \mathrm{d}R^2 + \sin^2 R(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2)\right] =: \Omega^{-2}\mathrm{d}\tilde{s}^2 \tag{2.1.6}$$

with

$$\Omega^2 = \frac{4}{(1+v^2)(1+u^2)}. \tag{2.1.7}$$

Note that $\mathrm{d}\tilde{s}^2$ is the natural Lorentz metric on the manifold $\mathbb{S}^3 \times \mathbb{R}$, known as *Einstein static universe.*

Thus, we have found the following result: There exists a conformal isometry[2] of Minkowski spacetime $(\mathbb{R}^4, \eta_{ab})$ into the open region O of the Einstein static universe $(\mathbb{S}^3 \times \mathbb{R}, \tilde{g}_{ab})$ given by the coordinate restrictions (2.1.5).

Definition 14. Conformal infinity of Minkowski spacetime is defined as the boundary, ∂O, of O in the Einstein static universe as illustrated in Figure 2.1. This boundary can be devided into five parts

[1]This transformation is similar to the stereographic projection of the real line to the unit circle.
[2]A conformal isometry of (M, g_{ab}) into (M', g'_{ab}) is a diffeomorphism $\psi : M \to M'$ such that $(\psi^* g)_{ab} = \Omega^2 g'_{ab}$.

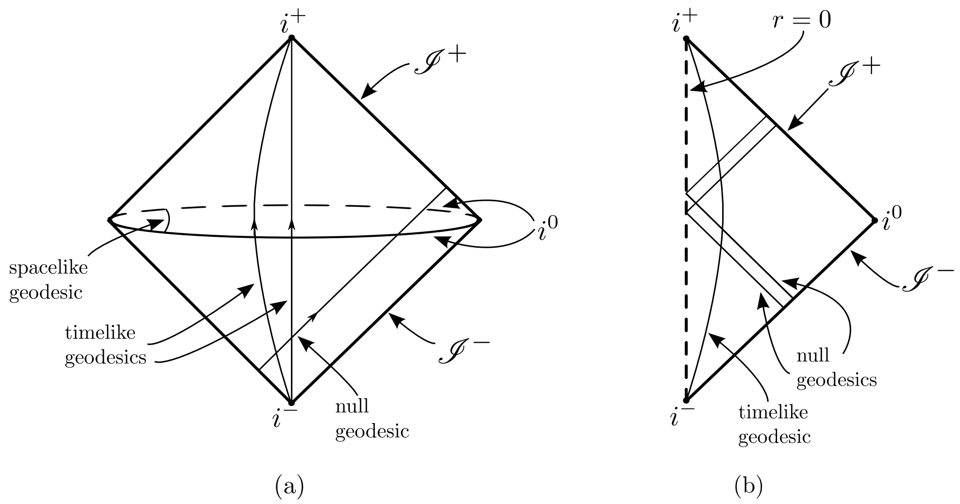

Figure 2.2.: (a) The shaded region of Figure 2.1 with only one spatial coordinate suppressed. (b) The Penrose diagram of Minkowski spacetime; each point represents a two-sphere, except for i^+, i^- and i^0, each of which is a single point, and points on the line $r = 0$.

- i^+ = future timelike infinity (given by $R = 0,\ T = \pi$)
- i^- = past timelike infinity (given by $R = 0,\ T = -\pi$)
- i^0 = spatial infinity (given by $R = \pi,\ T = 0$)
- $\mathscr{I}^+$ = future null infinity (given by $T = \pi - R$ for $0 < R < \pi$)
- $\mathscr{I}^-$ = past null infinity (given by $T = -\pi + R$ for $0 < R < \pi$)

Remark 1. All timelike geodesics of Minkowski spacetime begin at i^- and end at i^+, all spacelike geodesics begin and end at i^0, and all null geodesics begin at $\mathscr{I}^-$ and end at $\mathscr{I}^+$.

Since it can be quite difficult to draw spacetime diagrams on $\bar{O}$ as in Figure 2.1, and since two spatial dimensions are suppresed in this diagram, one often represents $\bar{O}$ as two null cones joined at their base as illustrated in Figure 2.2a. But this representation is somehow misleading, since i^0 is represented as a two-sphere rather than a point. One can also draw such spacetime diagrams as in Figure 2.2b known as *Penrose diagrams*, which still reflect the qualitative, causal structure of the spacetime.

2.2. Asymptotically Flat Spacetimes

Taking this construction of conformal infinity for Minkowski spacetime as a motivation, we will now turn to the definition of asymptotic flatness for arbitrary spacetimes. We would like to define a generic spacetime to be asymptotically flat if a similar construction, as in the Minkowski case, is possible. Therefore, we need to find a conformal isometry to map our physical spacetime (M, g_{ab}) into an "unphysical spacetime" $(\tilde{M}, \tilde{g}_{ab})$. Then, the boundary of the image of the physical spacetime under the conformal isometry would give us again a precise notion of infinity.

However, there are two important modifications which have to be made in order to include physically interesting spacetimes into the notion of asymptotic flatness. First, we do not impose any requirements on the presence of the points i^+ and i^-, since we would also like to describe isolated bodies which are present at early and late times. Second, although we require $\tilde{g}_{ab}$ to become flat at i^0, smoothness or even differentiability is too strong a requirement (see [29] for details).

Before we state the definition of asymptotic flatness for curved spacetimes we still need some technical definitions. Let $\{x^\mu, \mu = 1, \ldots, 4\}$ be a smooth coordinate system at i^0. We define the "radial function" ρ by

$$\rho^2 = \sum_{\mu=1}^{4} (x^\mu)^2 \tag{2.2.1}$$

and the angular functions ϕ^α, $(\alpha = 1, \ldots, 4)$ by the same formulas which are used to define the 3-sphere coordinates in 4-dimensional Euclidean space.

Definition 15. A function $f : M \to \mathbb{R}$ is said to have a *regular direction-dependent limit at* i^0 if the following three properties are satisfied:

1. For each C^1 curve γ ending at i^0, the limit of f along γ exists at i^0. Furthermore the value of limit only depends on the tangent directions of γ at i^0. We define $F(\phi^\alpha) = \lim_{i^0} f$, where the limit is taken along a curve whose tangent direction at i^0 is characterized by ϕ^α.

2. F is a smooth function on the 3-sphere.

3. Along every C^1 curve ending at i^0, we have for all $n \geq 1$

$$\lim_{i^0} \frac{\partial^n f}{\partial \phi^n} = \frac{\partial^n F}{\partial \phi^n}, \qquad \lim_{i^0} \rho^n \frac{\partial^n f}{\partial \rho^n} = 0. \tag{2.2.2}$$

   (Here $\partial^n / \partial \phi^n$ denotes the n-th partial derivative with respect to ϕ^α, where it is understood that the same partial derivative occurs on both sides of the equation.)

Definition 16. $\tilde{g}_{ab}$ is said to be of class $C^{>0}$ iff

1. $\tilde{g}_{ab}$ is continuous at i^0 and

2. all the first partial derivatives of the components of $\tilde{g}_{ab}$ in a smooth chart covering i^0 have regular direction-dependent limits at i^0.

Now we can state the definition of asymptotically flat curved spacetimes according to Ashtekar and Hansen [2].

Definition 17. A vacuum spacetime (M, g_{ab}) is called *asymptotically flat at null and spatial infinity* (or *asymptotically flat* for short) if there exists a spacetime $(\tilde{M}, \tilde{g}_{ab})$ - with $\tilde{g}_{ab}$ being C^∞ everywhere except possibly at a point i^0 where it is $C^{>0}$ - and a conformal isometry $\psi : M \to \psi[M] \subset \tilde{M}$ with conformal factor Ω (so that $\tilde{g}_{ab} = \Omega^2 (\psi^* g)_{ab}$ in $\psi[M]$) with the following conditions:

1. $\overline{J^+(i^0)} \cup \overline{J^-(i^0)} = \tilde{M} \setminus M$. (Here and in the following we write M instead of $\psi[M]$ for notational simplicity.) Thus i^0 is spacelike related to all points in M and the boundary, ∂M, of M consists of the union of i^0, $\mathscr{I}^+ = \partial J^+(i^0) \setminus i^0$ and $\mathscr{I}^- = \partial J^-(i^0) \setminus i^0$.

2. There exists an open neighborhood V of $\partial M = i^0 \cup \mathscr{I}^+ \cup \mathscr{I}^+$ such that the spacetime $(V, \tilde{g}_{ab})$ is strongly causal.

3. Ω can be extended to a function on all of $\tilde{M}$ which is C^2 at i^0 and C^∞ elsewhere.

4. a) On $\mathscr{I}^+$ and $\mathscr{I}^-$ we have $\Omega = 0$ and $\tilde{\nabla}_a \Omega \neq 0$. (Here $\tilde{\nabla}_a$ is the derivative operator associated with $\tilde{g}_{ab}$).
   b) We have $\Omega(i^0) = 0$, $\lim_{i^0} \tilde{\nabla}_a \Omega = 0$, and $\lim_{i^0} \tilde{\nabla}_a \tilde{\nabla}_b \Omega = 2\tilde{g}_{ab}(i^0)$. (We take limits at i^0 since $\tilde{g}_{ab}$ need not be C^1 there, and thus $\tilde{\nabla}_a$ need not be defined at i^0.)

5. a) The map of null directions at i^0 into the space of integral curves of $n^a = \tilde{g}^{ab}\tilde{\nabla}_b \Omega$ on $\mathscr{I}^+$ and $\mathscr{I}^+$ is a diffeomorphism.
   b) For a smooth function ω on $\tilde{M} \setminus i^0$ with $\omega > 0$ on $M \cup \mathscr{I}^+ \cup \mathscr{I}^-$ which satisfies $\tilde{\nabla}_a(\omega^4 n^a) = 0$ on $\mathscr{I}^+ \cup \mathscr{I}^-$, the vector field $\omega^{-1} n^a$ is complete on $\mathscr{I}^+ \cup \mathscr{I}^-$.

Remark 2. Note that since M and $\psi[M]$ are conformally isometric, they are in particular diffeomorphic. In general relativity, spacetimes which differ only by a diffeomorphism are identified as representing the same physical spacetime.[3] This is why we wrote M instead of $\psi[M]$ in the above definition.

Remark 3. According to the above definition we have $\tilde{M} = \psi[M] \cup \mathscr{I}^+ \cup \mathscr{I}^- \cup i^0$ with $\psi[M] = \mathrm{int}(\tilde{M})$. Since M and $\psi[M]$ represent the same spacetime we will think of M being embedded into $\tilde{M}$ in the same way as $\psi[M]$ is embedded into $\tilde{M}$. Having this in mind, the physical and unphysical metric are related by $\tilde{g}_{ab} = \Omega^2(\iota^* g)_{ab}$, where $\iota : \tilde{M} \hookrightarrow M$ is the inclusion map, i.e. we can think of $\tilde{g}_{ab}$ as an extension of g_{ab}. For notational simiplicity we omit the pullback of the inclusion map in the following.

Remark 4. Note that the causal structure is preseved as we proceed from the physical to the unphysical spacetime, i.e. timelike, null and and spacelike vector remain timelike, null and spacelike respectively under the conformal isometry. This follows from the fact that $\tilde{g}_{ab}$ and g_{ab} differ only by multiplication with a positive function, i.e. $\tilde{g}_{ab} = \Omega^2 g_{ab}$.

Remark 5. The association of an unphysical spacetime $(\tilde{M}, \tilde{g}_{ab})$ to an asymptotically flat physical spacetime (M, g_{ab}) is essentially arbitrary. If $(\tilde{M}, \tilde{g}_{ab})$ is an unphysical spacetime satisfying the properties of the definition with conformal factor Ω, then so is $(\tilde{M}, \omega^2 \tilde{g}_{ab})$ with conformal factor $\omega\Omega$, provided only that the function ω is strictly positive, is smooth everywhere except possibly at i^0, is $C^{>0}$, and satisfies $\omega(i^0) = 1$. Thus, there is considerable gauge freedom in the choice of the unphysical metric.

Remark 6. The definition of asymptotic flatness did not make any reference to a particular field equation. Therefore, it is also possible to use this definition to define such spacetimes in alternative theories of gravity, such as higher derivative theories of gravity. In particular, in the HDTG which we consider later on, it is assured that asymptotically flat solutions exist (see section 6.2).

[3] This means that a solution of the Einstein equation actually corresponds to an equivalence class of spacetimes, where spacetimes are identified if they differ by a diffeomorphism.

3. Null Geometry

After we have introduced the notion of a hypersurface of a manifold, we will derive a formula for the induced metric on a timelike, spacelike and null hypersurface. Furthermore, we will construct an adapted coordinate system in a neighborhood of non-null and null hypersurfaces, which will prove to be useful in subsequent calculations. Special emphasis will be placed on null congruences and the Raychaudhuri equation, which is the essential tool in the proof of the area theorem.

3.1. Geometry of Null Hypersurfaces

Consider two (topological) manifolds M and $\mathscr{S}$ with $\dim \mathscr{S} = r < n = \dim M$ and let $\phi : \mathscr{S} \to M$ be a map. If ϕ is locally one-to-one, i.e. for each $q \in \mathscr{S}$ there exists a neighborhood O such that $\phi|_O$ is one-to-one, and $\phi^{-1} : \phi(O) \to \mathscr{S}$ is smooth, then $\phi(\mathscr{S})$ is said to be an *immersed submanifold* of M. If in addition, ϕ is globally one-to-one, then $\phi(\mathscr{S})$ is said to be an *embedded submanifold* of M. An embedded submanifold of dimension $n-1$ is called a *hypersurface*.

Let S be a hypersurface of a spacetime (M, g_{ab}) and let $p \in S$. Each tangent space T_pS can be naturally viewed as a 3-dimensional subspace of T_pM. Thus, there exists a vector $\xi^a \in T_pM$, unique up to scaling, which is orthogonal (with respect to g_{ab}) to each vector in T_pS. The corresponding vector field ξ^a is said to be the *normal* of S. The hypersurface S is said to be spacelike (timelike, null), if ξ^a is timelike (spacelike, null). If S is spacelike (timelike), S is a Riemannian (Lorentzian) manifold with respect to the induced metric h_{ab}, i.e. g_{ab} restricted to tangent vectors of S. On the other hand, if S is null, then the induced metric is degenerate, and so does not define a pseudo-Riemannian metric on S. Despite this difficulty, null hypersurfaces are important in general relativity, since they represent horizons of various sorts, in particular the event horion of a black hole.

Consider now a smooth null hypersurface N of a spacetime (M, g_{ab}). As we mentioned before, N is a co-dimension one submanifold of M, such that $g_{ab} : T_pN \times T_pN \to \mathbb{R}$ is degenerate. The normal vector field k^a of N has the following properties:

- k^a is null and can be chosen future directed,
- $[k^a]^\perp = T_pN$,
- every vector in T_pN is either a multiple of k^a or spacelike.

Note that k^a is smooth if N is smooth. The following fact is fundamental.

Proposition 5. *Let N be a smooth null hypersurface and let k^a be a smooth future directed null vector field on N. Then the (affinely parameterized) integral curves of k^a are null geodesics.*

Proof. It suffices to show $k^b \nabla_b k^a = ck^a$ with $c \in \mathbb{R}$. (In the affine parametrization the geodesic equation follows.) To show this, it suffices to show that at each $p \in S$ we have $k^b \nabla_b k^a \perp T_pS$, i.e. $g_{ab}(k^c \nabla_c k^a)X^b = 0$ for all $X^a \in T_pS$.

We can extend each X^a, by making it invariant under the flow generated by k^a, i.e. we have

$$\mathcal{L}_k X^a = [k, X]^a = k^b \nabla_b X^a - X^b \nabla_b k^a = 0.$$

Clearly we have $g_{ab}k^a X^b = k_a X^a = 0$. Differentiation yields

$$0 = k^b \nabla_b (k_a X^a) = (k^b \nabla_b k_a) X^a + (k^b \nabla_b X^a) k_a,$$

and hence

$$(k^b \nabla_b k^a) X_a = -(k^b \nabla_b X^a) k_a = -(X^b \nabla_b k^a) k_a = -\frac{1}{2} X^b \nabla_b (k^a k_a) = 0$$

□

Remark 7. In the following we will refer to the integral curves of the vector field k^a as *null geodesic generators.*

Given the spacetime metric g_{ab}, we will now construct the induced metric h_{ab} on a hypersurface S by restricting the action of g_{ab} to tangent vectors of S. Consider first of all the case where S is either timelike or spacelike.

Non-null case: Let $S \subset M$ be a timelike hypersurface of M with unit normal vector field ξ^a. As we said before, each T_pS can be thought of as a subspace of T_pM. In each tangent space, we can define a projection P which maps vectors $X^a \in T_pM$ onto the orthogonal complement of ξ^a. Then, the induced metric $h_{ab} : T_pS \times T_pS \to \mathbb{R}$ on S can be defined by

$$h_{ab} X^a Y^b := g_{ab} (PX)^a (PY)^b, \qquad \forall X^a, Y^a \in T_pM. \tag{3.1.1}$$

For the construction of h_{ab} the following properties will be essential:

(i) $g_{ab}\xi^a (PX)^b = 0$, for all $X^a \in T_pM$

(ii) $(P^2 X)^a = (PX)^a$, for all $X^a \in T_pM$

(iii) $g_{ab} X^a (PY)^b = g_{ab} (PX)^a Y^b$, for all $X^a, Y^a \in T_pM$.

Properties (i) and (ii) are evident. In order to see that property (iii) holds, consider some $X^a \in T_pS$ and $Y^a \in T_pM$. We have

$$g_{ab} X^a (PY)^b = g_{ab} (PX)^a (PY)^b = g_{ab} (PY)^a (PX)^b, \tag{3.1.2}$$

since X^a remains unchanged under the projection P and g_{ab} is symmetric. By interchanging X^a and Y^b we find property (iii). For a timelike hypersurface S, the projector $P = P^{(t)}$ is given by

$$(P^{(t)} X)^a = X^a - \xi^a g_{bc} \xi^b X^c. \tag{3.1.3}$$

Insertion of (3.1.3) into (3.1.1) yields

$$\begin{aligned} h^{(t)}_{ab} X^a Y^b &= g_{ab}(P^{(t)}X)^a(P^{(t)}Y)^b = g_{ab}(P^{(t)}X)^a Y^b = g_{ab}[X^a - \xi^a g_{cd}\xi^c X^d]Y^b \\ &= g_{ab}X^aY^b - g_{ab}\xi^a Y^b g_{cd}\xi^c X^d = [g_{ab} - \xi_a\xi_b]X^aY^b, \end{aligned} \tag{3.1.4}$$

where we used properties (i)-(iii) from above. As we see, the induced metric on a timelike hypersurface is given by

$$h^{(t)}_{ab} = g_{ab} - \xi_a\xi_b. \tag{3.1.5}$$

If S is a spacelike hypersurface, then $P = P^{(s)}$ is given by

$$(P^{(s)}X)^a = X^a + \xi^a g_{bc}\xi^b X^c. \tag{3.1.6}$$

A similar calculation as in the timelike case yields

$$h^{(s)}_{ab} = g_{ab} + \xi_a\xi_b \tag{3.1.7}$$

for the induced metric on a spacelike hypersurface.

Null case: Let N be a smooth null hypersurface in (M, g_{ab}). If we restrict the metric g_{ab} to tangent vectors of N, the induced metric will be in general degenerate, i.e. $g_{ab}X^aY^a = 0$ for all X^a does not necessarily imply $Y^a = 0$. This property stems from the fact that the normal vector k^a of N is contained in T_pN, but is also orthogonal to every vector in T_pN. Due to this fact, it is not possible to define a unique projector onto the whole tangent space of a null hypersurface.[1]

However, we can overcome this difficulty by selecting an auxiliary null vector field l^a, normalized such that $l_ak^a = 1$. Then we can define a projector

$$(P^{(n)}X)^a = X^a - l^a g_{bc}k^bX^c - k^a g_{bc}l^bX^c, \tag{3.1.8}$$

which satisfies properties (i)-(iii) as in the non-null case. Since we have $(P^{(n)}k)^a = (P^{(n)}l)^a = 0$, the image of $P^{(n)}$ is a 2-dimensional subspace of T_pM which corresponds to the set of vectors which are orthogonal to both k^a and l^a. In the following, we will refer to this subspace as $\widehat{T_pN}$. Again, by inserting (3.1.8) into (3.1.1) we obtain the following form for the induced metric

$$\mu_{ab} := h^{(n)}_{ab} = g_{ab} - l_ak_b - k_al_b, \tag{3.1.9}$$

which corresponds to the metric on the (Riemannian) submanifold, specified by the two (normal) vector fields k^a and l^a. Note that the conditions $l_al^a = 0$ and $l_ak^a = 1$ do not determine l^a uniquely. Thus, (3.1.9) is not unique. However, as we shall see in the next paragraph, quantities of interest, like the expansion of a congruence, are the same for all choices of the auxiliary null vector field.

The inverse of h_{ab} (either in the non-null or null case) will be denoted by $h^{ab} = (h_{ab})^{-1}$, satisfying $h^{ab}h_{bc} = \delta^a{}_c$. Note that the projection operator P is given in terms of the induced metric by $h^a{}_c = h^{ab}g_{bc}$.

[1] From this follows in particular that it is not possible to define a Levi-Civita connection on a null hypersurface.

3.2. Gaussian Null Coordinates (GNC)

In this section, we will present the construction of a special coordinate system, known as *Gaussian null coordinates*. In order to illustrate the essential idea, we will consider first of all the case where S is a non-null hypersurface. In this case, the constructed coordinate system will be referred to as *Gaussian normal coordinates*.

Gaussian normal coordinates are defined for any non-null hypersurface S with normal vector field ξ^a in the following way: For each $p \in S$ we can construct a unique geodesic through p with tangent vector ξ^a. On (a portion of) the hypersurface S we choose arbitrary coordinates $\{x^1, x^2, x^3\}$. Each point in a neighbourhood of (that portion of) S may be labeled by the parameter t along the geodesic on which it lies and the coordinates $\{x^1, x^2, x^3\}$ of the point $p \in S$ from which the geodesic emanated. Thus, we have constructed a chart $p \mapsto \{t, x^1, x^2, x^3\}$ in a (sufficiently small) neighbourhood of S as we wished to do.

If S is a spacelike hypersurface, the spacetime metric may be written (in a neighborhood of S) as

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + h_{AB}\mathrm{d}x^A\mathrm{d}x^B \tag{3.2.1}$$

in the coordinate system $\{t, x^1, x^2, x^3\}$, with $A, B = 1, 2, 3$. Here, (h_{AB}) is a symmetric, positive definite 3×3 matrix which corresponds to the induced metric on the hypersurface S. The metric takes the special form (3.2.1), since we have $g_{tt} = -1$ due to the normalization of the vector field $\xi^a = (\partial/\partial t)^a$. Furthermore we have $g_{tA} = 0$, since ξ^a is orthogonal to any tangent vector $(\partial/\partial x^A)^a$ of the hypersurface S.

Now, we will proceed with the construction of such an adapted coordinate system in the case where S is a null hypersurface. The following exposition will be largely based on [9].

Let (M, g_{ab}) be a spacetime, let N be a smooth null hypersurface and let $\zeta \subset N$ be a smooth spacelike 2-dimensional submanifold. On an open subset $\tilde{\zeta}$ of ζ, we choose arbitrary coordinates $\{x^1, x^2\}$. On a neighbourhood of $\tilde{\zeta}$ in N, let k^a be a smooth nonvanishing normal vector field on N, such that the integral curves of k^a coincide with the null geodesic generators of N. Without loss of generality we choose k^a to be future directed. On an open neighbourhood R of $\tilde{\zeta} \times \{0\}$ in $\tilde{\zeta} \times \mathbb{R}$, let $\psi : R \to N$ be the map which takes each (q, u) into the point in N lying at parameter value u of the integral curve of k^a starting at q. Then, ψ is C^∞. From the inverse function theorem follows that ψ is one-to-one and onto from an open neighbourhood of $\tilde{\zeta}\times\{0\}$ onto an open neighbourhood $\tilde{N}$ of $\tilde{\zeta}$ in N. The functions x^1, x^2 can be extended from $\tilde{\zeta}$ to $\tilde{N}$, by keeping their values constant along the integral curves of k^a. Then, $\{u, x^1, \dots, x^{n-2}\}$ is a coordinate system on $\tilde{N}$. At each $p \in \tilde{N}$, let l^a be the unique null vector field, satisfying $l^a k_a = 1$ and $l^a X_a = 0$ for vectors X^a which are tangent to $\tilde{N}$ and satisfy $X^a \nabla_a u = 0$. On an open neighbourhood Q of $\tilde{N} \times \{0\}$ in $\tilde{N} \times \mathbb{R}$, let $\Psi : Q \to M$ be the map which takes each (p, r) into the point in M lying at parameter value r of the integral curve of l^a starting at p. Then, Ψ is C^∞. From the inverse function theorem follows that Ψ is one-to-one and onto from an open neighbourhood of $\tilde{N} \times \{0\}$ onto an open neighbourhood, O, of $\tilde{N}$ in M. The functions u, x^1, x^2 can be extended from $\tilde{N}$ to O, by keeping their values constant along the integral curves of l^a. Then, $\{u, r, x^1, x^2\}$ is a coordinate system of O, which will be referred to as *Gaussian null coordinates*. Note that on $\tilde{N}$ we have $k^a = (\partial/\partial u)$. By construction the vector field $l^a = (\partial/\partial r)$ is tangent to null geodesics in O, hence we have $g_{rr} = 0$. Furthermore we have

$$g_{ru} = g_{ab}l^a k^b = 1 \tag{3.2.2}$$

$$g_{rA} = g_{ab}l^a(\partial/\partial x^A)^b = 0 \tag{3.2.3}$$

for all $A = 1, 2$ throughout O and

$$g_{uu} = g_{ab}k^a k^b = 0 \tag{3.2.4}$$

$$g_{uA} = g_{ab}k^a(\partial/\partial x^A)^b = 0 \tag{3.2.5}$$

for all $A = 1, 2$ throughout $\tilde{N}$. From this follows that, within O, there exist smooth functions α and β_A, with $\alpha|_{\tilde{N}} = (\partial g_{uu}/\partial r)|_{r=0}$ and $\beta_A|_{\tilde{N}} = (\partial g_{uA}/\partial r)|_{r=0}$ such that the spacetime metric in O takes the form

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}u\mathrm{d}r + \mathrm{d}r\mathrm{d}u - 2r\alpha\mathrm{d}u^2 - r\beta_A\mathrm{d}u\mathrm{d}x^A - r\beta_A\mathrm{d}x^A\mathrm{d}u + \mu_{AB}\mathrm{d}x^A\mathrm{d}x^B, \tag{3.2.6}$$

where the $\mu_{AB}\mathrm{d}x^A\mathrm{d}x^B$ is a 2-dimensional Riemannian metric. Note that in the coordinate system $\{u, r, x^A\}$ the null hypersurface N is specified by $r = 0$. Using the abstract index notation, we can rewrite (3.2.6) as

$$g_{ab} = 2\Big[(\mathrm{d}r)_{(a}(\mathrm{d}u)_{b)} - r\alpha(\mathrm{d}u)_{(a}(\mathrm{d}u)_{b)} - r\beta_A(\mathrm{d}x^A)_{(a}(\mathrm{d}u)_{b)}\Big] + \mu_{AB}(\mathrm{d}x^A)_{(a}(\mathrm{d}x^B)_{b)}. \tag{3.2.7}$$

The construction of this coordinate system is of a very general nature, in the sense that we can construct such coordinates in a neighborhood of any null hypersurface - in particular the event horizon of a black hole (see figure 4.3).

In appendix D.4 we will prove $\alpha = \mathcal{O}(r)$, i.e. the function α vanishes on the null hypersurface N. Hence, we can make the replacement $\alpha \to r\alpha$, such that in the region O the metric takes the following form

$$g_{ab} = 2\Big[(\mathrm{d}r)_{(a}(\mathrm{d}u)_{b)} - r^2\alpha(\mathrm{d}u)_{(a}(\mathrm{d}u)_{b)} - r\beta_A(\mathrm{d}x^A)_{(a}(\mathrm{d}u)_{b)}\Big] + \mu_{AB}(\mathrm{d}x^A)_{(a}(\mathrm{d}x^B)_{b)}. \tag{3.2.8}$$

In the region O, the inverse metric takes the form

$$g^{ab} = 2\Big[(\partial_u)^{(a}(\partial_r)^{b)} + r\beta^A(\partial_A)^{(a}(\partial_r)^{b)}\Big] + r^2[\beta^2 + 2\alpha](\partial_r)^{(a}(\partial_r)^{b)} + \mu^{AB}(\partial_A)^{(a}(\partial_B)^{b)}, \tag{3.2.9}$$

where we introduced the shorthand notation

$$(\partial_u)^a := \left(\frac{\partial}{\partial u}\right)^a, \qquad (\partial_r)^a := \left(\frac{\partial}{\partial r}\right)^a, \qquad (\partial_A)^a := \left(\frac{\partial}{\partial x^A}\right)^a, \tag{3.2.10}$$

and we defined $\beta^A := \mu^{AB}\beta_B$, $\beta^2 := \beta^A\beta_A$.

3.3. Null Congruences

Consider a spacetime (M, g_{ab}) and some open subset $O \subset M$. A *congruence* in O is a family of curves γ_s, such that through every $p \in O$ there passes one and only one curve in this family.

The tangents to a congruence yield a vector field in O and conversely, every continuous vector field generates a congruence of curves. The congruence $(s,t) \mapsto \gamma_s(t)$ is said to be smooth, if the associated vector field is smooth.

Consider now a smooth congruence of null geodesics in a spacetime region O. We assume that the geodesics are affinely parametrized, with affine parameter u, i.e. the associated vector field k^a satisfies

$$k_a k^a = 0, \qquad k^a \nabla_a k^b = 0. \tag{3.3.1}$$

By choosing the parameters s, u as coordinates on the 2-dimensional submanifold, which is spanned by the curves γ_s, we may write $k^a = (\partial/\partial u)^a$. Furthermore, the congruence gives rise to a "deviation vector field" η^a which may be written as $\eta^a = (\partial/\partial s)^a$ in this coordinate system. The vector η^a represents the displacement to an infinitesimal nearby geodesic. Note that since k^a and η^a are coordinate vector field, they commute[2]:

$$[k,\eta]^a = k^b \nabla_b \eta^a - \eta^b \nabla_b k^a = \mathcal{L}_k \eta^a = 0. \tag{3.3.2}$$

In the following we are interested in how the congruence evolves with "time", i.e. we want to study if the individual geodesics start "winding around" each other or if the congruence starts to develope "focal points". In order to do so, we need to study the behavior of the deviation vector as a function of the affine parameter along some reference geodesic.

Since η^a is only supposed to represent the separation of two neighbouring curves, not the separation of particular points on these curves, there is an ambiguity in the specification of the deviation vector: η^a and $\eta'^a = \eta^a + ck^a$ respresent a displacement to the same geodesic, for some constant $c \in \mathbb{R}$. Thus, one is only interested in the equivalence class of deviation vectors, where vectors are said to be equivalent if they differ by a multiple of k^a.

In order to illustrate this problem, let us consider the case of a smooth congruence of timelike geodesics, with associated unit tangent vector field t^a. We can overcome the ambiguity we mentioned above by choosing $t_a \eta^a = 0$ at some initial proper time value τ_0. Since we have

$$t^a \nabla_a (t_b \eta^b) = \eta_b \underbrace{t^a \nabla_a t^b}_{=0} + t_b t^a \nabla_a \eta^b \overset{(3.3.2)}{=} t_b \underbrace{\mathcal{L}_t \eta_b}_{=0} + t_b \eta^a \nabla_a t^b = \frac{1}{2} \eta^a \nabla_a (t_b t^b) = 0, \tag{3.3.3}$$

$t_a \eta^a$ is constant along each geodesic. So if $t_a \eta^a$ is chosen to vanish at τ_0, it will do so for all other values of τ. As we see, the set of unambiguous deviation vectors corresponds to a 3-dimensional subspace in each $T_pM, \quad p = \gamma_s(\tau) \in O$, which is comprised by the vectors orthogonal to t^a. One can show that this space is isomorphic to the space of equivalence classes of vectors in T_pM which differ only by addition of a multiple of k^a (see [15] for details).

In the case of a smooth congruence of null geodesics, $k_a \eta^a$ is also constant along each geodesic, but the condition $k_a \eta^a = 0$ is not sufficient to remove the ambiguity, since we have

$$k_a \eta'^a = k_a (\eta^a + ck^a) = k_a \eta^a, \tag{3.3.4}$$

[2] This follows from the commutativity of mixed partial derivatives, since we have

$$\frac{\partial}{\partial s} \frac{\partial}{\partial \tau} f(\gamma_s(t)) = \frac{\partial}{\partial \tau} \frac{\partial}{\partial s} f(\gamma_s(t)),$$

for any $f \in C^\infty(M, \mathbb{R})$, in the coordinate system we described above.

i.e. we have $k_a \eta'^a = 0$ whenever $k_a \eta^a = 0$. This property stems from the fact that k^a is null, i.e. the 3-dimensional subspace of T_pM, consisting of all vectors orthogonal to k^a, contains k^a itself. In order to overcome this difficulty, we choose another vector l^a with the properties

$$l_a l^a = 0, \qquad l_a k^a = 1. \tag{3.3.5}$$

Furthermore, we choose l^a to be parallely transported along the geodesics, i.e. we have

$$k^a \nabla_a l^b = 0. \tag{3.3.6}$$

In the previous paragraph, we have shown that k^a and l^a are coordinate vector fields in an adapted coordinate system (Gaussian null coordinates). Hence, k^a and l^a commute:

$$k^a \nabla_a l^b = l^a \nabla_a k^b. \tag{3.3.7}$$

The ambiguity in the specification of the deviation vector field η^a in the case of a null congruence may be removed by requiring

$$\eta_a l^a = 0, \qquad \text{in addition to} \qquad \eta_a k^a = 0. \tag{3.3.8}$$

As we see, the set of unambiguous deviation vector field corresponds to a 2-dimensional subspace of T_pM, consisting of all vectors which are orthogonal to both k^a and l^a. In the previous paragraph we referred to this subspace as $\widehat{T_pN}$. In the following, we will always assume that deviation vectors are elements of $\widehat{T_pN}$. One can show that $\widehat{T_pN}$ is isomorphic to a subspace of T_pN, consisting of equivalence classes of vectors which differ only by a multiple of k^a (see [15] for details).

Now, let us define the tensor $B_{ab} := \nabla_b k_a$. It is orthogonal to k^a, in the sense that we have

$$\begin{aligned} B_{ab}k^a &= k^a \nabla_b k_a = \frac{1}{2}\nabla_b(k_a k^a) = 0, \\ B_{ab}k^b &= k^b \nabla_b k_a = 0, \end{aligned} \tag{3.3.9}$$

since k^a is geodesic and normalized to one. As we will see, this tensor determines the evolution of the deviation vector field. However, B_{ab} is not orthogonal to l^a and hence, B_{ab} has components in the "ambigious directions" of η^a. We can fix this problem by using the projector $\mu^a{}_b$ from the previous section to project B_{ab} onto $\widehat{T_pN}$. We have

$$\begin{aligned} \hat{B}_{ab} &= \mu^c{}_a \mu^d{}_b B_{cd} \\ &= (\delta^c{}_a - l^c k_a - k^c l_a)(\delta^d{}_b - l^d k_b - k^d l_b) B_{cd} \\ &= B_{ab} - k_a B_{cb} l^c - k_b B_{ac} l^c + k_a k_b B_{cd} l^c l^d \\ &= B_{ab} - k_a (\nabla_b k_c) l^c - k_b (\nabla_c k_a) l^c + k_a k_b (\nabla_d k_c) l^c l^d \\ &= B_{ab} - k_a \underbrace{\nabla_b (k_c l^c)}_{=0} + k_a k_c \nabla_b l^c - k_b \underbrace{l^c \nabla_c k_a}_{=0} + k_a k_b l^c \underbrace{l^d \nabla_d k_c}_{=0} \\ &= B_{ab} + k_a k_c \nabla_b l^c. \end{aligned} \tag{3.3.10}$$

One can check that $\hat{B}_{ab}$ is orthogonal to l^a, i.e. we have

$$\hat{B}_{ab}l^a = \hat{B}_{ab}l^b = 0. \tag{3.3.11}$$

Now, we introduce the following quantities:

Definition 18. The *expansion* ϑ, *shear* σ_{ab}, and *twist* ω_{ab} of a congruence are defined as follows

$$\vartheta := \hat{B}^{ab}\mu_{ab} \tag{3.3.12}$$

$$\sigma_{ab} := \hat{B}_{(ab)} - \frac{1}{2}\vartheta\mu_{ab} \tag{3.3.13}$$

$$\omega_{ab} := \hat{B}_{[ab]}. \tag{3.3.14}$$

Using these quantities, $\hat{B}_{ab}$ can be decomposed as follows

$$\hat{B}_{ab} = \frac{1}{2}\vartheta\mu_{ab} + \sigma_{ab} + \omega_{ab}. \tag{3.3.15}$$

The tensor $\hat{B}_{ab}$ has the following interpretation: The covariant derivative of some $\eta^a \in \widehat{T_pN}$ in the direction of k^a represents the relative velocity of two neighbouring geodesics. We have

$$\begin{aligned} k^b\nabla_b\eta^a &= k^b\nabla_b\mu^a{}_c\eta^c = k^b\nabla_b(\delta^a{}_c - k^a l_c - l^a k_c)\eta^c = \mu^a{}_c k^b\nabla_b\eta^c = \mu^a{}_c B^c{}_b\eta^b \\ &= \mu^a{}_c B^c{}_b \mu^b{}_d \eta^d = \hat{B}^a{}_d\eta^d, \end{aligned} \tag{3.3.16}$$

where we have used $k^b\nabla_b k^a = k^b\nabla_b l^a = 0$ for the third equality, $k^b\nabla_b\eta^a = \eta^b\nabla_b k^a = B^a{}_b\eta^b$ for the fourth equality and the fact that η^a remains unchanged under projection onto $\widehat{T_pN}$. As we see, $\hat{B}_{ab}$ measures the failure of η^a to be parallely transported along the congruence. From this follows that along any geodesic in the congruence, ϑ measures the average expansion of infinitesimally nearby surrounding geodesics; ω_{ab}, being the antisymmetric part of the linear map $\hat{B}_{ab}$, measures their rotation; and σ_{ab} measures their shear[3].

According to their definition, ω_{ab} and σ_{ab} are orthogonal to k^a and l^a, i.e. we have

$$\omega_{ab}k^a = \omega_{ab}k^b = \sigma_{ab}k^a = \sigma_{ab}k^b = \omega_{ab}l^a = \omega_{ab}l^b = \sigma_{ab}l^a = \sigma_{ab}l^b = 0. \tag{3.3.17}$$

Furthermore, congruences which are hypersurface orthogonal are characterized by the following

Proposition 6. *A congruence is hypersurface orthogonal, if and only if $\omega_{ab} = 0$.*

Proof. If $\omega_{ab} = 0$, then

$$0 = k_{[a}\omega_{bc]} = k_{[a}\hat{B}_{bc]} = k_{[a}B_{bc]} + k_{[a}k_b k_{|d|}\nabla_{c]}l^d = k_{[a}\nabla_b k_{c]}. \tag{3.3.18}$$

The last equality follows since we have

$$\begin{aligned} k_{[a}k_b k_{|d|}\nabla_{c]}l^d = \frac{1}{6}(&k_a k_b k_d\nabla_c l^d - k_b k_a k_d\nabla_c l^d + k_b k_c k_d\nabla_a l^d - k_c k_b k_d\nabla_a l^d \\ &+ k_c k_a k_d\nabla_b l^d - k_a k_c k_d\nabla_b l^d) = 0. \end{aligned} \tag{3.3.19}$$

[3] An inital sphere in a tangent space which is Lie transported along k^a will distort towards an ellipsoid where the principal axes given by the eigenvectors of $\sigma^a{}_b$ and the rate of change is given by the eigenvalues of $\sigma^a{}_b$.

By Frobenius's theorem follows that k^a is orthogonal to a family of hypersurfaces.

Conversely, if k^a is orthogonal to a family of hypersurfaces, then Frobenius's theorem implies $k_{[a}\nabla_b k_{c]} = 0$. Then, by reversing the previous steps we find that

$$0 = k_{[a}\omega_{bc]} = \frac{1}{3}(k_a\omega_{bc} + k_b\omega_{ca} + k_c\omega_{ab}).$$

Contraction with l^a yields $\omega_{ab} = 0$, since we have $l^a k_a = 1$ and $\omega_{ab}l^a = \omega_{ab}l^b = 0$. □

That the expansion does not depend on the choice of the auxiliary null vector field l^a can be seen in the following manner:

$$\begin{aligned}\vartheta &= \hat{B}^{ab}\mu_{ab} \\ &= (B^{ab} + k^a k_c \nabla^b l^c)(g_{ab} - k_a l_b - l_a k_b) \\ &= B^{ab}g_{ab} + k_b k_c \nabla^b l^c - \underbrace{(k^a k_a)}_{=0} k_c l_b \nabla^b l^c - \underbrace{(k^a l_a)}_{=1} k_b k_c \nabla^b l^c \\ &= B^{ab}g_{ab} \\ &= \nabla_a k^a.\end{aligned} \tag{3.3.20}$$

Since ϑ plays an essential role in the proof of the area theorem, we will further investigate its physical interpretation: Consider the null geodesic congruence which is generated by the normal vector field k^a of a null hypersurface N. The *extrinsic curvature* K_{ab} of N is defined as

$$K_{ab} = \hat{B}_{ba} = B_{ba} + k_b k_c \nabla_a l^c. \tag{3.3.21}$$

This tensor is orthogonal to k^a, i.e. we have $K_{ab}k^a = K_{ab}k^b = 0$. Since the congruence is hypersurface orthogonal, we have $\omega_{ab} = 0$. Therefore, from the definition of $\hat{B}_{ab}$ follows that $\hat{B}_{ab}$, and hence K_{ab}, are symmetric. The Lie derivative of the spacetime metric g_{ab} with respect to the vector field k^a is given by

$$\begin{aligned}\frac{1}{2}\mathcal{L}_k g_{ab} &= \frac{1}{2}(\nabla_a k_b + \nabla_b k_a) \\ &= \frac{1}{2}(K_{ab} - k_b k_c \nabla_a l^c + K_{ba} - k_a k_c \nabla_b l^c) \\ &= K_{ab} - k_{(a}k_{|c|}\nabla_{b)} l^c,\end{aligned} \tag{3.3.22}$$

where we used $K_{ab} = K_{ba}$ in the last equality. Consider now the induced metric μ^{ab} of the 2-dimensional submanifold ζ of N which is specified by the two (normal) vector fields k^a and l^a. Contraction of K_{ab} with μ^{ab} yields

$$K_{ab}\mu^{ab} = (\frac{1}{2}\mathcal{L}_k g_{ab} + k_{(a}k_{|c|}\nabla_{b)}l^c)\mu^{ab} = \frac{1}{2}(\mathcal{L}_k g_{ab})\mu^{ab}, \tag{3.3.23}$$

where we used $\mu_{ab}k^a = \mu_{ab}k^b$. Furthermore we have

$$\mathcal{L}_k \mu_{ab} = \mathcal{L}_k(g_{ab} - k_a l_b - l_a k_b) = \mathcal{L}_k g_{ab}, \tag{3.3.24}$$

since $\mathcal{L}_k k_a = \mathcal{L}_k l_a = 0$. In an adapted coordiante system $\{x^\alpha, \alpha = 0, \dots, 3\}$ (Gaussian null coordinates) the Lie derivative of μ_{ab} with respect to k^a can be expressed as

$$\mathcal{L}_k \mu_{\alpha\beta} = \frac{\mathrm{d}\mu_{\alpha\beta}}{\mathrm{d}u}, \tag{3.3.25}$$

where u is the affine parameter of the geodesics generated by the vector field k^a. Overall we obtain,

$$\vartheta = \hat{B}_{\alpha\beta}\mu^{\alpha\beta} = K_{\alpha\beta}\mu^{\alpha\beta} = \frac{1}{2}(\mathcal{L}_k \mu_{\alpha\beta})\mu^{\alpha\beta} = \frac{1}{2}\mu^{\alpha\beta}\frac{\mathrm{d}\mu_{\alpha\beta}}{\mathrm{d}u} = \sqrt{\mu}^{-1}\frac{\mathrm{d}}{\mathrm{d}u}\sqrt{\mu}, \tag{3.3.26}$$

where we have used the relation

$$\frac{\mathrm{d}}{\mathrm{d}u}\sqrt{\mu} = \frac{1}{2}\sqrt{\mu}\mu^{\alpha\beta}\frac{\mathrm{d}\mu_{\alpha\beta}}{\mathrm{d}u} \tag{3.3.27}$$

for $\mu = \det(\mu_{\alpha\beta})$. This calculation justifies the interpretation of ϑ as a measure for the average expansion of infinitesimally nearby geodesics. Equation (3.3.27) corresponds to the rate of change of the volume of submanifold ζ, which is generated by intersecting the null congruence with a spacelike hypersurface, with respect to the affine parameter.

3.4. The Raychaudhuri Equation

In the following, we will derive the *Raychaudhuri equation*, which determines the rate of change of ϑ, ω_{ab} and σ_{ab} along each geodesic in the congruence. Consider:

$$\begin{aligned}
k^c\nabla_c\hat{B}_{ab} &= k^c\nabla_c(B_{ab} + k_a k_d \nabla_b l^d) \\
&= k^c\nabla_c B_{ab} + k_a k_d k^c \nabla_c \nabla_b l^d \\
&= k^c\nabla_c\nabla_b k_a + k_a k_d k^c \nabla_c \nabla_b l^d \\
&= k^c\nabla_b\nabla_c k_a + R_{cba}{}^d k^c k_d + k_a k_d k^c \nabla_c \nabla_b l^d \\
&= \nabla_b(\underbrace{k^c\nabla_c k_a}_{=0}) - (\nabla_b k^c)(\nabla_c k_a) + R_{cba}{}^d k^c k_d + k_a k_d k^c \nabla_c \nabla_b l^d \\
&= -B^c{}_b B_{ac} + R_{cba}{}^d k^c k_d + k_a k_d k^c \nabla_c \nabla_b l^d.
\end{aligned} \tag{3.4.1}$$

By taking the trace of of the left hand side of (3.4.1) we obtain

$$\begin{aligned}
\mu^{ab}k^c\nabla_c\hat{B}_{ab} &= \mu^{ab}k^c\nabla_c\hat{B}_{ab} + \hat{B}_{ab}k^c\nabla_c(g^{ab} - l^a k^b - k^a l^b) \\
&= \mu^{ab}k^c\nabla_c\hat{B}_{ab} + \hat{B}_{ab}k^c\nabla_c\mu^{ab} \\
&= k^c\nabla_c(\hat{B}_{ab}\mu^{ab}) \\
&= k^c\nabla_c\theta \\
&= \frac{\mathrm{d}\vartheta}{\mathrm{d}u}
\end{aligned} \tag{3.4.2}$$

where we used the compatibility of the metric and $k^a\nabla_a k^b = k^a\nabla_a l^b = 0$ for the first equality. By taking the trace of the right hand side of (3.4.1) we obtain

$$-B^c{}_b B_{ac}\mu^{ab} + R_{cba}{}^d k^c k_d \mu^{ab} + k_a k_d k^c \nabla_c \nabla_b l^d = -B^c{}_b B_{ac}\mu^{ab} + R_{cba}{}^d k^c k_d \mu^{ab} \tag{3.4.3}$$

since μ_{ab} is orthogonal to k^a, and therefore

$$\begin{aligned}-B^c{}_bB_{ac}\mu^{ab} + R_{cba}{}^dk^ck_d\mu^{ab} &= -B^c{}_bB_{ac}(g^{ab} - l^ak^b - l^bk^a) + R_{cba}{}^dk^ck_d(g^{ab} - l^ak^b - l^bk^a)\\ &= -B^{ca}B_{ac} + R_{cba}{}^dg^{ab}k^ck_d - R_{cba}{}^dk^ck_dk^bl^a - R_{cba}{}^dk^ck_dk^al^b\\ &= -B^{ca}B_{ac} - \underbrace{R_{cbda}g^{ab}}_{=R_{cd}}k^ck^d - \underbrace{R_{abcd}k^al^bk^ck^d}_{=0} - \underbrace{R_{abcd}k^ak^bl^ck^d}_{=0}\\ &= -B^{ca}B_{ac} - R_{ab}k^ak^b,\end{aligned} \tag{3.4.4}$$

by using the symmetries of the Riemann tensor. One can rewrite the first term in (3.4.4) as follows:

$$\begin{aligned}B^{ca}B_{ac} &= (\hat{B}^{ca} - k^ck_d\nabla^al^d)(\hat{B}_{ac} - k_ak_e\nabla_cl^e)\\ &= \hat{B}^{ca}\hat{B}_{ac} + \underbrace{(k^c\nabla_cl^e)}_{=0}k_ek_dk_a\nabla^al^d\\ &= \left(\frac{1}{2}\vartheta\mu^{ca} + \sigma^{ca} + \omega^{ca}\right)\left(\frac{1}{2}\theta\mu_{ac} + \sigma_{ac} + \omega_{ac}\right)\\ &= \left(\frac{1}{2}\vartheta\mu^{ac} + \sigma^{ac} - \omega^{ac}\right)\left(\frac{1}{2}\vartheta\mu_{ac} + \sigma_{ac} + \omega_{ac}\right)\\ &= \frac{1}{4}\vartheta^2\underbrace{\mu^{ac}\mu_{ac}}_{=2} + \sigma^{ac}\sigma_{ac} - \omega^{ac}\omega_{ac} + \vartheta\mu^{ac}\sigma_{ac}.\end{aligned} \tag{3.4.5}$$

The last term in (3.4.5) vanishes, since we have

$$\begin{aligned}\mu^{ac}\sigma_{ac} &= \mu^{ac}\hat{B}_{(ac)} - \frac{1}{2}\vartheta\mu^{ac}\mu_{ac}\\ &= \mu^{ac}\hat{B}_{ac} - \vartheta\\ &= 0.\end{aligned} \tag{3.4.6}$$

Hence we obtain

$$\frac{\mathrm{d}\vartheta}{\mathrm{d}u} = -\frac{1}{2}\vartheta^2 - \sigma_{ab}\sigma^{ab} + \omega_{ab}\omega^{ab} - R_{ab}k^ak^b, \tag{3.4.7}$$

which is known as *Raychaudhuri equation* for a null geodesic congruence.

Let us investigate the nonpositivity of the right hand side of (3.4.7). If the congruence is hypersurface orthogonal, we have $\omega_{ab} = 0$. The terms, $-\vartheta^2$ and $-\sigma_{ab}\sigma^{ab}$, are manifestly nonpositive. Furthermore, if we assume that the null convergence condition (see appendix B), $R_{ab}k^ak^b \geq 0$ for all null vectors k^a, is satisfied, we obtain

$$\frac{\mathrm{d}\vartheta}{\mathrm{d}u} + \frac{1}{2}\vartheta^2 \leq 0, \tag{3.4.8}$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}u}\vartheta^{-1} \geq \frac{1}{2} \tag{3.4.9}$$

and hence

$$\vartheta^{-1}(u) \geq \vartheta_0^{-1} + \frac{1}{2}u, \tag{3.4.10}$$

where ϑ_0 is the initial value of ϑ. Suppose, that ϑ_0 is negative. Then (3.4.10) implies that ϑ^{-1} must pass through zero, i.e. $\vartheta \to -\infty$, within affine length $u \leq 2/|\vartheta_0|$. Thus we have proven the following lemma.

Lemma 2. *Let k^a be a tangent field of a hypersurface orthogonal congruence of null geodesics. Suppose $R_{ab}k^a k^b \geq 0$, as will be the case if the Einstein equation holds in the spacetime and the null energy condition is satisfied by the matter. If the expansion ϑ takes a negative value ϑ_0 at any point on geodesic in the congruence, then ϑ goes to $-\infty$ along that geodesic within affine length $u \leq 2/|\vartheta_0|$.*

This result can be heuristically interpreted as follows: $\vartheta_0 < 0$ states that the congruence is initially converging. The attractive nature of gravity then implies that the congruence must continue to converge which eventually leads to a "focal/conjugate point".

3.5. Conjugate Points

In order to understand the result of lemma 2 properly, we will need to introduce the notion of conjugate points: Let $\gamma : [a, b] \to M$ be a null geodesic with $\gamma(a) = p$ and $\gamma(b) = q$. An unambigious deviation vector field η^a is called *Jacobi field*, if it solves the geodesic deviation equation

$$k^a \nabla_a (k^b \nabla_b \eta^c) = -R_{abd}{}^c \eta^b k^a k^d \tag{3.5.1}$$

with $\eta^a|_p = \eta^a|_q = 0$. Two points p, q are called *conjugate* if there exists a Jacobi field connecting p and q. Together with Lemma 2, the next lemma states that q is conjugate to p if and only if the expansion of a null geodesic congruence emanating from p approaches $-\infty$ at q.

Lemma 3. *Let (M, g_{ab}) be a spacetime satisfying $R_{ab}k^a k^b \geq 0$ for all null vectors k^a. Let γ be a null geodesic and let $p \in \gamma$. Suppose the expansion ϑ of the null geodesic congruence emanating from p attains a negative value ϑ_0 at $r \in \gamma$. Then within affine paramerter length $u \leq 2/|\vartheta_0|$ from r, there exists a point q conjugate to p along γ, assuming that γ extends that far.*

For a proof of this Lemma, we refer the reader to [29].

A similar notion of conjugacy can be defined for a point and an 2-dimensional spacelike submanifold S. At each $q \in S$ there exists two future directed null vectors k_1^a, k_2^a which are orthogonal to S. If S is orientable, a continuous choice of k_1^a and k_2^a over S can be made and thereby we can define two families of null geodesics, which we will refer to as "outgoing" and "ingoing". Let γ be a null geodesic in one of these families. A point $p \in \gamma$ is said to be conjugate to S along γ, if there exists a deviation vector field satisfying (3.5.1) which is nonzero at S but vanishes at p. In analogy to lemma 3, we have

Lemma 4. *Let (M, g_{ab}) be a spacetime satisfying $R_{ab}k^a k^b \geq 0$ for all null vectors k^a. Let S be a smooth 2-dimensional spacelike submanifold such that the expansion ϑ of the "outgoing" null geodesics has a negative value ϑ_0 at $q \in S$. Then within affine parameter length $u \leq (n-2)/|\vartheta_0$, from q, there exists a point p conjugate to S along the outgoing null geodesic γ passing through q.*

The following result is the key technical lemma in the proof in the area theorem:

Lemma 5. *Let (M, g_{ab}) be a globally hyperbolic spacetime and let K be a compact, orientable, two-dimensional spacelike submanifold of M. Then every $p \in \partial I^+(K)$ lies on a future directed null geodesic starting from K which is orthogonal to K and has no point conjugate to K between K and p.*

For a proof of this Lemma, we refer the reader to [29].

4. Gravitational Collapse and Black Holes

In the following, we will summarize the essentials of the theory of black holes. After a phenomenological part, which explains under what circumstances the formation of a black hole occurs, we will see that, within general relativity, spherically symmetric gravitational collapse leads to the formation of a spacetime singularity. In the following, we will assume that these singularities cannot be seen from distant observers (cosmic cencorship conjecture). After that, we will state the mathematical definition of a black hole according to Hawking. We will discuss the properties of the event horizon and in particular the area theorem due to Hawking. Furthermore, we will talk about stationary black holes which are expected to represent the equilibrium configuration of a black holes at sufficiently late times.

4.1. Phenomenology

After a star has exhausted its nuclear fuel, it can no longer remain in equilibrium and must ultimately undergo gravitational collapse. Depending on the initial mass of the star, gravitational collapse will lead to the formation of a white dwarf, neutron star or black hole. In this section we will briefly review some of the physical processes that lead to the formation of these astrophysical objects.

When a star forms due to condensation of a gas cloud, it contracts and heats up until the central temperature and density is sufficiently high such that nuclear processes set in, which convert hydrogen to helium. The collapse of the star is then halted and an equilibrium configuration is obtained, since the total pressure due to nuclear reactions balances gravity. During this phase of the stellar evolution, a large core of helium is built up. If the star is sufficiently massive, this core will start to contract until helium reactions begin to occur which lead to the formation of heavier elements. This process may repeat itself until a large core of nickel and iron is produced.

When the star runs out of nuclear fuel it can no longer support itself against gravitational collapse. As the density of the star approaches values of nuclear matter ($\sim 10^{14}\text{g/cm}^{-3}$) quantum mechanical effect begin to play an important role. According to the Pauli exclusion principle no two electrons can be in the same state simultaneously, so not all electrons can be in the lowest energy level. Rather, electrons must occupy a band of energy levels. The interior of the star consists of a plasma, i.e. ions and free electrons. As the compression of the electron gas proceeds due to gravitational collapse, the number of electrons in a given volume increases as well. Thus, the maximum energy level is raised, i.e. the energy of the electrons increases upon compression. In order to compress the electron gas further, an additional compressing force is required, which manifests itself as a resisting pressure. This is the origin of the so called *electron degeneracy pressure*.

The fate of the collapsing star depends on whether the electron degeneracy pressure is sufficient to support the star against gravity. If the mass of the star is below the Chandrasekhar limit

$$m_C \approx 1,4\left(\frac{2}{\mu_N}\right)m_\odot, \tag{4.1.1}$$

where μ_N is the number of nucleons per electron and $m_\odot$ denotes the mass of the Sun, the star will approach an equilibrium configuration supported by electron degeneracy pressure. These bodies are known as *white dwarfs*. No further nuclear reactions will occur and the white dwarf slowly cools down as it radiates away its remaining thermal energy. If the mass of the star is greater than m_C, electron degeneracy pressure is not sufficient to support the star against gravity. The nickel and iron core will undergo gravitational collapse. If the mass of the collapsing part of the star is below the so called *cold matter upper mass limit* ($\sim 2M_\odot$), the *neutron degeneracy pressure* is sufficient to halt the collapse, resulting in the formation of a *neutron star*. If the mass of the star is above the cold matter upper mass limit, the star will eventually undergo complete gravitational collapse and it is believed that the result will be a *black hole*. It should be noted that black holes formed by stellar collapse are in the mass range $2m_\odot \leq m \leq 100m_\odot$ since stars with $m \leq 2m_\odot$ should not collapse, while stars with $m \geq 100m_\odot$ do not exist due to pulsational instabilities.

Besides the formation of black holes resulting from stellar collapse, there are also other physical processes which may lead to the formation of a black hole due to gravitational collapse. One can think for example of the collapse of an entire central core of a dense cluster of stars. The most likely site for the formation of such *massive black holes* is the center of a galaxy. Another, much more speculative, process by which black holes may have been produced is by gravitational collapse of regions of enhanced density in the early universe. These are commonly referred to as *primordial black holes*.

Concerning the detection of black holes: Due to the fact that black holes are extremely small objects (the Schwarzschild radius of a black hole of one solar mass would be ~ 3 km) and since they are literally “black”, it seems hopeless to detect these objects in any direct (optical) way. But if we consider a black hole resulting from stellar collapse, which is in close binary orbit with a star, the situation looks more promising. One would expect that matter would flow from the star to the black hole, thereby forming an *accretion disk* around the black hole. Viscous heating in the accretion disk could result in the production of X-rays. A number of X-ray sources with an ordinary star in a close binary orbit around an unseen companion have been found such as Cygnus X-1. In [22], a lower mass limit for the unseen companion of $\sim 9m_\odot$ was found. This is above the upper mass limit of neutron stars and white dwarfs, suggesting that the unseen companion of Cygnus X-1 is a black hole.

Furthermore, one would expect that a massive black causes a brightness enhancement as well as an increase of the average velocity very near the center of a galaxy. Exactly such a brightness enhancement and an increased “velocity dispersion” have been observed at the center of the galaxy M87 [39], [25], thus providing strong evidence for the existence of a black hole of mass $\sim 5 \cdot 10^9 m_\odot$.

4.2. Definition of a Black Hole

It is a well known fact from general relativity that the complete gravitational collapse of a spherical, non-rotating body, such as a star, always results in the production of a Schwarzschild spacetime as a finial equilibrium configuration. The Penrose diagram of the (extended)

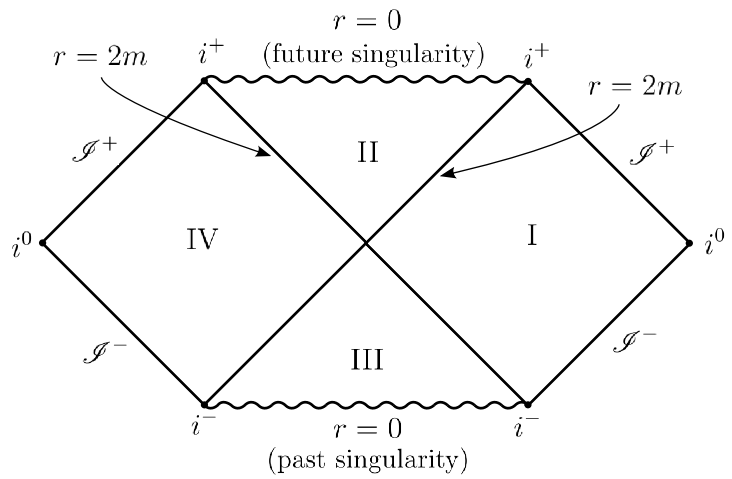

Figure 4.1.: Penrose diagram of the maximal extension of the Schwarzschild solution. Null lines are at $\pm 45°$

Schwarzschild solution is depicted in figure 4.1. Region I represents the exterior gravitational field of the spherical body. An interesting property of this spacetime is that any observer which enters region II can never escape from it. Once the the null surface $r = 2m$ is crossed, the observer will fall into the (future) singularity $r = 0$ within a finite proper time. Furthermore, any light signal which was sent by the observer will remain region II. Therefore, this region is called *black hole*. Region III is the time-reversed analog of the black hole: the *white hole*. Any observer in region III must have originated from the singularity and the observer must leave this region within a finite proper time. Region IV corresponds to another asymptotically flat spacetime with properties identical to those of region I.

The most interesting fact about the Schwarzschild solution is that is contains a singularity which is hidden within a "region of no escape" which we referred to as black hole. However, this solution to the Einstein equation is very special, because of its spherical symmetry. That the formation of a singularity is a genuine feature, even for non-spherical gravitational collapse, is guaranteed by the singularity theorems of Hawking and Penrose [15]: *For small deviations from spherical symmetry a spacetime singularity must neccesarily occur in gravitational collapse.* But the singularity theorems do not tell us whether or not this singularity is visible to distant observers or not. If the singularity is visible to far away observers, we say that the star has ended as a *naked singularity.* If the singularity is not visible to far away observers, i.e. it is hidden behind a spacetime region, we say that the star has ended as a black hole.

The Einstein equation admits solutions involving naked singularities . The presence of such naked singularities cause severe problems, since it is impossible to predict the behavior of spacetime in the causal future of the singularity. General relativity would therefore lose its predictive power in this spacetime region. Due to these problems, Penrose conjectured that naked singularities do not appear in physically reasonable spacetimes. This conjecture is commonly referred to as the

Cosmic Censorship Conjecture. *The complete gravitational collapse of a body always results in a black hole rather than a naked singularity, i.e. all singularities of gravitational collapse are "hidden" within black holes, where they cannot be "seen" by distant observers.*

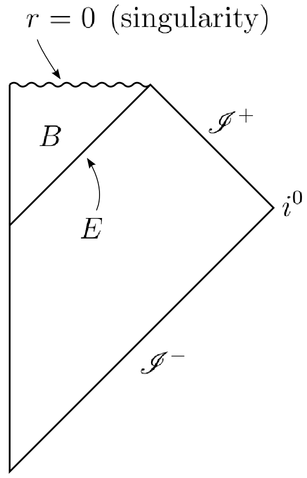

Figure 4.2.: Definition of the black hole region B and the event horizon E.

From now on we will assume that this conjecture is true. The notion of *strongly asymptotically predictable* spacetimes assures that spacetimes do not posses naked singularities.

Definition 19. Let (M, g_{ab}) be an asymptotically flat spacetime with associated unphysical spacetime $(\tilde{M}, \tilde{g}_{ab})$. We say that (M, g_{ab}) is *strongly asymptotically predictable* if in the unphysical spacetime there is an open region $\tilde{V} \subset \tilde{M}$ with $\overline{M \cap J^-(\mathscr{I}^+)} \subset \tilde{V}$ such that $(\tilde{V}, \tilde{g}_{ab})$ is globally hyperbolic.

Note that the closure of $M \cap J^-(\mathscr{I}^+)$ is taken in the unphysical spacetime $\tilde{M}$, so we have $i^0 \in \tilde{V}$. That this definition assures that singularities are not visible from infinity can be seen in the following manner: The requirement that $(\tilde{V}, \tilde{g}_{ab})$ is a globally hyperbolic region of the unphysical spacetime implies that $(M \cap \tilde{V}, g_{ab})$ is a globally hyperbolic region of the physical spacetime.[1] Furthermore, from theorem 4 we know that $M \cap \tilde{V}$ can be foliated by a family of Cauchy surfaces Σ_t. So, for all $p \in M \cap \tilde{V}$ and for all Σ_t with $p \in J^+(\Sigma_t)$, every past directed inextendible causal curve from p intersects Σ_t. This can be interpreted as saying that[2] no singularities are visible to any observer in $[M \cap \tilde{V}] \supset [M \cap \overline{J^-(\mathscr{I}^+}]$.

The following definition gives a precise meaning to the notion of a black hole as a "place of no escape". For asymptotically flat spacetimes, the crucial property that distinguishes the black hole region from the rest of the spacetime is the impossibility of escaping to future null infinity.

Definition 20. A strongly asymptotically predictable spacetime is said to contain a *black hole*, if M is not contained in $J^-(\mathscr{I}^+)$. The *black hole region* B of such a spacetime is defined as $B := M \setminus J^-(\mathscr{I}^+)$. The boundary of B in M, $E := \partial J^-(\mathscr{I}^+) \cap M = [\overline{J^-(\mathscr{I}^+)} \setminus J^-(\mathscr{I}^+)] \cap M$, is called the *event horizon* (see figure 4.2).

Note that since i^0 and $\mathscr{I}^-$ are contained in $J^-(\mathscr{I}^+)$, i^0 and $\mathscr{I}^-$ are not contained in E.

[1] According to property (1) of the definition of asymptotic flatness we have $M = \tilde{M} \setminus [J^+(i^0) \cup J^-(i^0)]$. Hence, a Cauchy surface for $(\tilde{V}, \tilde{g}_{ab})$ which passes through i^0 will be a Cauchy surface for $(M \cap \tilde{V}, g_{ab})$. The fact that g_{ab} and $\tilde{g}_{ab} = \Omega^2 g_{ab}$ have the same causal structure implies that $(M \cap \tilde{V}, g_{ab})$ is globally hyperbolic.

[2] apart from an initial singularity, such as a white hole

Furthermore, since $J^-(\mathscr{I}^+)$ is open in M (see section 1.2), the black hole B is closed in M. From this follows that the event horizont E is contained in B.

Remark 8. This definition does not make use of any field equation, and is therefore not limited to Einstein gravity. Thus, in alternative theories of gravity (such as a higher derivative theory of gravity) which admit strongly asymptotically predictable solutions, black holes can be defined in the same manner. In the HDTG which we consider later on, it is assured that there exists solutions which contain a black hole (see section 6.2).

4.3. General Properties of Black Holes

In the following we will list the properties of the event horizon E. The normal vector field of E will be denoted by n^a and its integral curves will be refered to as null geodesic generators.

(a) E is a global notion in the sense that one needs to know the entire future developement of the spacetime in order to determine if a black hole is present.

(b) E has all the properties of a past causal boundary as descricbed in section 1.2 (i.e. E is an achronal, three-dimensional, embedded C^0-submanifold of M).

(c) E is a null hypersurface.

(d) The null geodesic generators of E may have past endpoints (in the sense that their continuation into the past may leave E, e.g. $r = 0$ in the spherically symmetric case).

(e) The null geodesic generators of E have no future endpoints.

(f) The expansion of the null geodesic generators cannot become negative.

From properties (d) and (e) follows that geodesics may enter E but cannot leave it. This reflects the "intuitive notion" of a black hole as a "place of no escape". Properties (a)-(d) are evident. Property (e) follows from theorem 2. Property (f) will be further investigated in the proof of the area theorem (see below).

Remark 9. Let us discuss which of these properties are peculiar to general relativity. Properties (a)-(e) essentially follow from the definition of a black hole as a past causal boundary. Therefore, black hole solutions of other theories of gravity (such as higher derivative theories of gravity) also possess these characteristics. An exception is property (f). As we will see in the proof of the area theorem, the positivity of the expansion is established by means of the null convergence condition (see appendix B), which makes explicit use of Einstein's equation. Therefore, one cannot expect that property (f) is satisfied in other theories of gravity.

Remark 10. In section 3.2 we introduced the Gaussian null coordinate system which can be constructed in a neighborhood of any null hypersurface. Since E is a null hypersurface, this construction can also be applied to the event horizon of a black hole (see figure 4.3).

Now, we will define the notion of a black hole at an "instant of time".

Definition 21. Let (M, g_{ab}) be a strongly asymptotically predictable spacetime, with globally hyperbolic region $\tilde{V} \supset \overline{M \cap J^- \mathscr{I}^+}$ in the unphysical spacetime and let B be the black hole region of the spacetime. If Σ is a Cauchy surface for $\tilde{V}$, then we will call each connected component $\mathcal{B}$ of $\Sigma \cap B$ a *black hole at time* Σ. Furthermore, we will refer to the boundary $\partial\mathcal{B}$ of $\mathcal{B}$ as a *horizon cross-section* and we will denote it by $\mathcal{E}$.

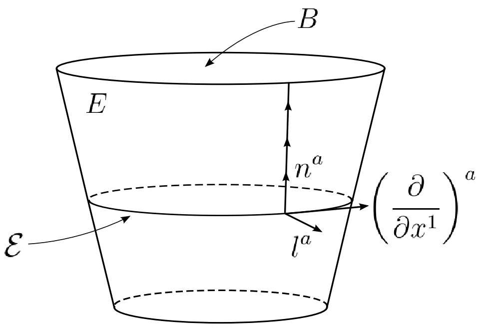

Figure 4.3.: Gaussian null coordinates $\{u, r, x^A\}$ for a black hole with one spatial dimension suppressed. The coordinates u and r correspond to the affine parameters of the integral curves of the vector fields n^a and l^a, respectively. The coordinates x^A, $A = 1, 2$ are arbitrarily chosen coordinates on a spatial cross-section $\mathcal{E}$.

The number of black holes in (M, g_{ab}) may vary with "time" (i.e. choice of Cauchy surface), since new black holes may form and black holes present at one time may merge at a later time. However, the next theorem states that black holes can neither disappear nor bifurcate.

Theorem 5. *Let (M, g_{ab}) be a strongly asymptotically predictable spacetime and let Σ_1 and Σ_2 be Cauchy surfaces for $\tilde{V}$ with $\Sigma_2 \subset I^+(\Sigma_1)$. Let $\mathcal{B}_1$ be a nonempty connected component of $B \cap \Sigma_1$. Then $J^+(\mathcal{B}_1) \cap \Sigma_2 \neq \varnothing$ and is contained within a single connected component of $B \cap \Sigma_2$.*

Proof. see [29] □

The next theorem concerns the evolution of the event horizon. Consider a horizon cross-section $\mathcal{E} = E \cap \Sigma$, where Σ is a spacelike Cauchy surface with respect to $\tilde{V}$. The following theorem, due to Hawking [12], states that the area of $\mathcal{E}$ never decreases with time.

Theorem 6 (The Area Theorem)**.** *Let (M, g_{ab}) be a strongly asymptotically predictable spacetime satisfying the null energy condition. Let Σ_1 and Σ_2 be spacelike Cauchy surfaces with respect to $\tilde{V}$ satisfying $\Sigma_2 \subset I^+(\Sigma_1)$ and let $\mathcal{E}_1 = E \cap \Sigma_1$, $\mathcal{E}_2 = E \cap \Sigma_2$. Then the area of $\mathcal{E}_1$ is greater than or equal to the area of $\mathcal{E}_2$*

Proof. The setup for the proof is summarized in figure 4.4a. In the following, we will consider a null geodesic congruence from $\mathcal{E}_1$ to $\mathcal{E}_2$ which is tangent to the affinely parametrized normal vector field $n^a = (\partial/\partial u)^a$ of E. This vector field gives rise to a one-parameter group of isometries $\phi_u : M \to M$. We define a family of two-surfaces $\mathcal{E}(u) := \phi_u(\mathcal{E}_1)$ by following the null geodesic generators from $\mathcal{E}_1$ by an amount u of the affine parameter. The parametrization of this isometry is chosen to be $\mathcal{E}(u_0) = \mathcal{E}_1$. On the horizon E we choose coordinates $\{u, x^1, x^2\}$, such that each $\mathcal{E}(u)$ is parametrized by $\{x^1, x^2\}$. As we have seen in section 3.3, the expansion of this congruence can be written as

$$\vartheta = \sqrt{\mu}^{-1} \frac{\mathrm{d}}{\mathrm{d}u} \sqrt{\mu},$$

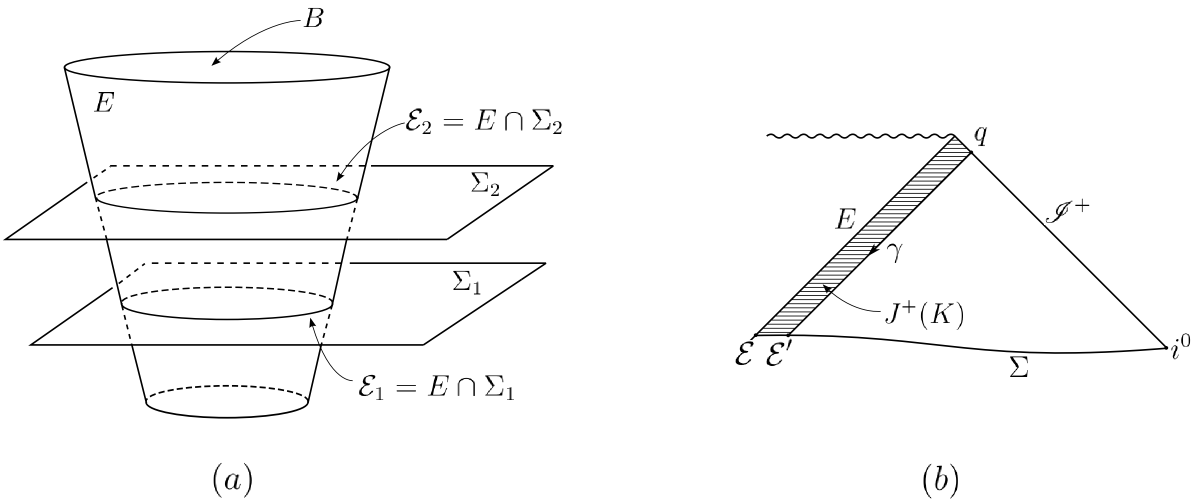

Figure 4.4.: Proof of the area theorem.

where μ is the determinant of the induced metric $\mu_{ab} = \mu_{ab}(u)$ on the cross-section $\mathcal{E}(u)$. From this follows

$$\mathcal{A}(\mathcal{E}(u)) - \mathcal{A}(\mathcal{E}_1) = \int_{u_0}^{u} \mathrm{d}u' \frac{\mathrm{d}}{\mathrm{d}u'} \mathcal{A}(\mathcal{E}(u')) = \int_{u_0}^{u} \mathrm{d}u' \left[\int_{\mathcal{E}(u')} \vartheta \sqrt{\mu}\, \mathrm{d}^2 x \right],$$

where $\mathcal{A}(\mathcal{E}(u))$ denotes the area of the cross-section $\mathcal{E}(u)$. If we could show $\vartheta \geq 0$ on E, then $\mathcal{A}(\mathcal{E}(u)) \geq \mathcal{A}(\mathcal{E}_1)$ would follow, and in particular $\mathcal{A}(\mathcal{E}_2) \geq \mathcal{A}(\mathcal{E}_1)$.

In order to show $\vartheta \geq 0$, we will derive a contradiction by assuming $\vartheta(p) = C < 0$ for some $p \in E$. Let Σ be a spacelike Cauchy surface for $\tilde{V}$ such that $p \in \Sigma$ and consider the two-surface $\mathcal{E} = E \cap \Sigma$. Since we have $\vartheta < 0$ at p, we can deform $\mathcal{E}$ outward in a neighborhood of p to obtain a surface $\mathcal{E}' \subset \Sigma$ which enters $J^-(\mathscr{I}^+)$ and has $\vartheta < 0$ everywhere in $J^-(\mathscr{I}^+)$. Let $K \subset \Sigma$ be the closed region lying between $\mathcal{E}$ and $\mathcal{E}'$ and let $q \in \partial J^+(K) \cap \mathscr{I}^+$. In the unphysical spacetime, let γ be the null geodesic generator of $\partial J^+(K)$, on which q lies (see figure 4.4b). According to lemma 5, γ must meet $\mathcal{E}'$ orthogonally with *no* conjugate point between q and Σ. On the other hand, since we have $\vartheta|_{\mathcal{E}'} = C < 0$, there *must* be a conjugate point in the causal future of $\mathcal{E}'$ after $u \leq 2/|C|$ according to lemma 4. Hence, we have a contradiction and $\vartheta \geq 0$ follows. □

Remark 11. The proof of this theorem crucially depends on lemma 4 which assumes that the null convergence condition (see appendix B) is satisfied. This condition is implied by the null energy condition together with the Einstein's equations. Therefore, this way of proving an area increase theorem cannot be applied to theories of gravity which satisfy other field equations (such as higher derivative theories of gravity), since it is tightly linked to the particular form of Einstein's equation.

Remark 12. Most discussions of the event horizon assume C^1 or even higher differentiability of E. Recently, this higher order differentiability assumption has been eliminated for the proof of the area theorem by [7].

4.4. Stationary Black Holes

In gravitational collapse that was strongly asymptotically predictable, i.e. no naked singularity evolved, one would expect the solution outside the horizon to become stationary for sufficiently late times. Therefore, it is interesting to study stationary solutions which contain a black hole, since these are expected to describe the final state of the collapsed system.

First of all, let us introduce the following terminology:

Definition 22. A black hole B is said to be

- *stationary* if there exists a one-paramerter group of isometries on (M, g_{ab}) generated by a Killing field t^a which is unit timelike at infinity.
- *static* if it is stationary and, in addition, t^a is hypersurface orthogonal.
- *axisymmetric* if there exists a one-parameter group of isometries on (M, g_{ab}) which correspond to rotations at infinity.[3]

Definition 23. Consider a Killing field K^a and the set of points on which K^a is null and not identically vanishing. Let $\mathcal{K}_i$ be a connected component of this set which is a null hypersurface. Any union $\mathcal{K} = \bigcup_i \mathcal{K}_i$ is called a *Killing horizon*.

Thus, $\mathcal{K}$ can be thought of as a null hypersurface whose null generators coincide with the orbits of a one-parameter group of isometries (so that there is a Killing field K^a which is normal to $\mathcal{K}$).

Definition 24. A *bifurcate Killing horizon* is a pair of null surfaces $\mathcal{K}_A$ and $\mathcal{K}_B$, which intersect in a spacelike two-surface $\mathcal{C}$, called *bifurcation surface*, such that $\mathcal{K}_A$ and $\mathcal{K}_B$ are Killing horizons with respect to the same Killing field K^a.

From this definition follows that K^a must vanish on $\mathcal{C}$, and conversely, if a Killing field K^a vanishes on a spacelike two-surface C, then $\mathcal{C}$ will be the bifurcation surface of a bifurcate Killing horizon associated with K^a .

In general relativity, a key result in the theory of black holes is a theorem due to Hawking [13] which relates the global concept of an event horizon to the local notion of Killing horizons. This result is commonly referred to as *rigidity theorem*. The result of this theorem will be stated in two steps:

Theorem 7 (Rigidity Theorem, part 1)**.** *The event horizon of a stationary black hole spacetime is a Killing horizon, provided that the spacetime is analytic, the present matter fields obey well behaved hyperbolic equations and the energy-momentum tensor fulfills the weak energy condition.*

For a proof of this theorem we refer the reader to [13] and [15]. A consequence of this theorem is that one of the following alternatives must hold:

Theorem 8 (Rigidity Theorem, part 2)**.** *The horizon Killing field K^a either coincides with the stationary Killing field t^a, or the spacetime admits at least one axial Killing field ϕ^a.*

[3] By convention, the associated axial Killing field ϕ^a is normalized such that its orbits have affine length 2π.

In the first case the black hole is said to be *nonrotating* (for this case it is known that the black hole must be static [31],[37]). In the second case, provided that there exists no third Killing field which would imply spherical symmetry, the black hole is said to be *rotating* and one has

$$K^a = t^a + \Omega_E \phi^a, \tag{4.4.1}$$

where the angular velocity of the horizon is denoted by the real constant Ω_E. In this case it can be shown that the black hole must be axisymmetric and stationary [13], [15]. This result is also refered to as rigidity theorem since it implies that the null geodesic generators of the horizon must rotate rigidly with respect to infinity.

Remark 13. Note that the proof of this theorem heavily relies on the fact that the event horizon cross-sections $\mathcal{E}$ are topologically 2-spheres (see topology theorem below). This is a nontrivial assumption which must not necessarily hold in HDTG. Therefore, the rigidity theorem does not readily extend to this context. As to our knowledge, such a theorem does not exists in a gravitational theory with an additional $R_{ab}R^{ab}$ contribution in the gravitational Lagrangean.

Another important result in the theory of black holes is the *topology theorem*, which is also due to Hawking [13]. This theorem asserts that, under suitable circumstances[4], horizon cross-sections $\mathcal{E}$ in asymptotically flat stationary black hole spacetimes obeying the dominant energy condition are spherical, i.e. all $\mathcal{E}$ are homeomorphic to the 2-sphere $\mathbb{S}^2$.

Remark 14. The proof of this theorem implicitly assumes that Einstein's equation is satisfied. Therefore, the topology theorem does not readily extend to the context of HDTG. As to our knowledge, such a theorem does not exists in a gravitational theory with an additional $R_{ab}R^{ab}$ contribution in the gravitational Lagrangean.

The rigidity and topology theorem are key results for the proof of the so called *black hole uniqueness theorems* which are due to Israel, Carter, Hawking and Robinson. These theorems were obtained between 1967 and 1975 and assure that all stationary black hole solutions are specified by a finite number of parameters, namely, in the vacuum case, their mass and angular momentum. This is why these theorems are sometimes also referred to as *no hair theorems*.[5] These results imply that 2-parameter Kerr family is the only possible stationary axisymmetric vacuum black hole solution to Einstein's equation.

Remark 15. The black hole uniqueness theorems do not readily extend to HDTG since they rely on the rigidity and topology theorem.

[4]The proof of the topology theorem requires the spacetime to be "regular predictable" (see [15], p.318 for a definition).

[5]Consider for example two bodies which differ greatly from each other in composition, shape and structure. If we assume that they undergo complete gravitational collapse, their final state will be the same provided only that their mass and angular momentum are the same. Therefore, black holes have no "individual features"(such as hairs) distinguishing them among each other, besides their mass and angular momentum.

5. Laws of Black Hole Mechanics

In the following we will state the laws of black hole mechanics which are due to Bardeen, Carter and Hawking [3]. As we will see, these laws have a remarkable similarity to the ordinary laws of thermodynamics. However, this similarity should only be considered to be a mathematical analogy within the classical framework. Only when when quantum effects are taken into account this analogy obtains a physical relevance.

This section will be concerned with the laws of black hole mechanics in general relativity, but we will also comment on the possible generalizations of these theorems. In section 6.3 we will discuss the status of these laws in higher derivative theories of gravity.

5.1. Zeroth Law

Consider a Killing horizon $\mathcal{K}$ (not necessarily the event horizon of a black hole) with normal Killing field K^a. On $\mathcal{K}$ we have $K^aK_a = 0$, so in particular K^aK_a is constant on $\mathcal{K}$. Hence $\nabla^a(K^bK_b)$ is normal to $\mathcal{K}$, so there exists a function κ, known as *surface gravity*, such that

$$\nabla^b(K^aK_a) = -2\kappa K^b. \tag{5.1.1}$$

By taking the Lie derivative of (5.1.1) with respect to K^a we obtain

$$\mathcal{L}_K\kappa = 0, \tag{5.1.2}$$

so κ is constant along the obits of K^a, i.e. κ is constant on each null geodesic generator of $\mathcal{K}$. In general, κ may vary from generator to generator but in the following we will show that κ is constant on the entire $\mathcal{K}$.

The surface gravity κ can be physically interpreted as follows: One can show (see [29]) that we have

$$\kappa = \lim(Va), \tag{5.1.3}$$

where $a = (a^ca_c)^{1/2}$, $a^c = (K^b\nabla_bK^c)/(-K^aK_a)$ is the magnitude of the acceleration of the orbits of K^a in the region off of $\mathcal{K}$ where K^a is timelike, $V = (-K^aK_a)^{1/2}$ is the "redshift factor" and the limit is taken as one approaches $\mathcal{K}$. Thus, Va is the force that must be exerted at infinity to hold a unit test mass in place near the horizon. This justifies the terminology surface gravity, since κ is the limiting value of this force.

Remark 16. Note that the surface gravity of a black hole is only defined when it is in "equilibrium", i.e. in the stationary case, so that its event horizon is a Killing horizon.

The following theorem asserts that κ is uniform over $\mathcal{K}$.

Theorem 9 (Zeroth Law of Black Hole Mechanics)**.** *Let $\mathcal{K}$ be a Killing horizon. Then the surface gravity κ is constant on $\mathcal{K}$, provided that Einstein's equation holds with matter satisfying the dominant energy condition.*

Proof. First of all, let us derive some useful formulas which will be needed in the proof. Equation (5.1.1) may be written as

$$\nabla^a(K^b K_b) = (\nabla^a K^b)K_b + K^b \nabla^a K_b = 2K^b \nabla_a K_b = -2\kappa K^a. \tag{5.1.4}$$

Since K^a is a Killing vector field, this implies

$$K^b \nabla^a K_b = -K^b \nabla_b K^a = -\kappa K^a. \tag{5.1.5}$$

Furthermore, since K^a is hypersurface orthogonal on the horizon, by Frobenius's theorem we have on the horizon

$$K_{[a} \nabla_b K_{c]} = 0. \tag{5.1.6}$$

Using Killing's equation $\nabla_b K_c = -\nabla_c K_b$, this implies

$$K_c \nabla_a K_b = -2K_{[a} \nabla_{b]} K_c. \tag{5.1.7}$$

Now, by applying $K_{[d} \nabla_{c]}$ to (5.1.5) we obtain

$$\begin{aligned} K_a K_{[d} \nabla_{c]} \kappa + \kappa K_{[d} \nabla_{c]} K_a &= K_{[d} \nabla_{c]} (K^b \nabla_b K^a) \\ &= (K_{[d} \nabla_{c]} K^b)(\nabla_b K_a) + K^b K_{[d} \nabla_{c]} \nabla_b K_a \\ &= (K_{[d} \nabla_{c]} K^b)(\nabla_b K_a) + K^b R_{ba[c}{}^e K_{d]} K_e, \end{aligned} \tag{5.1.8}$$

where we have used equation (C.0.3) was used in the last step. The first term in the last line of equation (5.1.6) may be written as

$$\begin{aligned} (K_{[d} \nabla_{c]} K^b)(\nabla_b K_a) &= -\frac{1}{2}(K^b \nabla_d K_c) \nabla_b K_a \\ &= -\frac{1}{2} \kappa K_a \nabla_d K_c \\ &= \kappa K_{[d} \nabla_{c]} K_a, \end{aligned} \tag{5.1.9}$$

where we used equation (5.1.7) for the first equality, equation (5.1.5) for the second equality and Killing's equation for the last equality. By inserting this result into (5.1.8) we find

$$K_a K_{[d} \nabla_{c]} \kappa = K^b R_{ab[c}{}^e K_{d]} K_e, \tag{5.1.10}$$

where the symmetries of the Riemann tensor were used.

On the other hand, if we apply $K_{[d} \nabla_{e]}$ to equation (5.1.7) we obtain

$$(K_{[d} \nabla_{e]} K_c) \nabla_a K_b + K_c K_{[d} \nabla_{e]} \nabla_a K_b = -2(K_{[d} \nabla_{e]} K_{[a}) \nabla_{b]} K_c - 2(K_{[d} \nabla_{e]} \nabla_{[b} K_{|c|}) K_{a]}. \tag{5.1.11}$$

By using (5.1.7) repeatedly, we find that the first term on the left-hand side of (5.1.11) cancels the the first term on the right hand side of the equation. Therefore, by using equation (C.0.3), we obtain

$$-K_c R_{ab[e}{}^f K_{d]} K_f = 2K_{[a} R_{b]c[e}{}^f K_{d]} K_f. \tag{5.1.12}$$

My multiplying this equation with g^{ce} and contracting over c and e, the left-hand side vanishes, and we find

$$-K_{[a}R_{b]}{}^{f}K_{f}K_{d} = K_{[a}R_{b]cd}{}^{f}K^{c}K_{f}. \tag{5.1.13}$$

Now, the term on the right-hand side of this equation is the same as the right-hand side of equation (5.1.10). Therefore, we have

$$K_{[d}\nabla_{c]}\kappa = -K_{[d}R_{c]}{}^{f}K_{f}. \tag{5.1.14}$$

In the following we will use Einstein's equation and the dominant energy condition (see appendix B) to show that the right-hand side of equation (5.1.14) vanishes. First of all, one can make the observation that the expansion ϑ, shear σ_{ab}, and twist ω_{ab} of the null geodesic generators of a Killing horizon vanish on the horizon (see [29]). Therefore, from the Raychaudhuri equation follows that we have

$$R_{ab}K^{a}K^{b} = 0. \tag{5.1.15}$$

Now, the dominant energy condition states that $-T^{a}{}_{b}K^{b}$ must be future directed timelike or null. Einstein's equation together with (5.1.15) implies $T^{a}{}_{b}K^{b}K_{a} = 0$. From this follows that $-T^{a}{}_{b}K^{b}$ points in the direction of K^{a}, i.e. $K_{[c}T_{a]b}K^{b} = 0$. By using Einstein's equation again we find that the right-hand side of (5.1.14) vanishes. Thus, we have found

$$K_{[d}\nabla_{c]}\kappa = 0, \tag{5.1.16}$$

which states that κ is constant on the horizon. □

Remark 17. Kay and Wald [33] have shown that it is also possible to establish the uniformity of κ over $\mathcal{K}$ without requiring the Einstein equation, if the Killing horizon is of a bifurcate type[1]. There is also another way to establish the above result in a purely geometrical manner, which neither relys on any field equations nor energy conditions of the matter. However, this derivation only works in the case of static or axisymmetric Killing horizons.

Remark 18. This law bears a resemblance to the zeroth law of thermodynamics, which states that the temperature T must be uniform over a body in thermal equilibrium. Stationary black holes represent equilibrium configurations in black hole physics. Theorem 9 asserts that a certain quantity, the surface gravity κ, must be constant over E. This mathematical analogy suggests that T and κ should represent the same physical quantity.

5.2. First Law

The Komar mass of a stationary, asymptotically flat spacetime which is a solution of the vacuum field equations near infinity is given by

$$m = -\frac{1}{8\pi}\int_{\mathbb{S}^2_\infty} \epsilon_{abcd}\nabla^{c}t^{d}, \tag{5.2.1}$$

where $\mathbb{S}^2_\infty$ is a two-sphere at spatial infinity and t^a is a stationary Killing field. It will turn out useful to rewrite this asymptotic integral as a volume integral. Consider a stationary,

[1] This is not a strong restriction, since, according to Rácz and Wald [35], [36], all "physically reasonable" Killing horizons are either bifurcate horizons or degenerate ($\kappa = 0$).

axisymmetric, asymptotically flat black hole solution to the vacuum Einstein equation and a spacelike hypersurface Σ which extends out to spatial infinity and intersects the event horizon E in a two-surface $\mathcal{E}$, such that we have $\partial\Sigma = \mathcal{E} \cup \mathbb{S}^2_\infty$. By defining the two-form $X_{ab} = \epsilon_{abcd}\nabla^c t^d$ we find

$$
\begin{aligned}
m - m_{BH} &= -\frac{1}{8\pi}\int_{\mathbb{S}^2_\infty} \epsilon_{abcd}\nabla^c t^d + \frac{1}{8\pi}\int_{\mathcal{E}} \epsilon_{abcd}\nabla^c t^d \\
&= -\frac{1}{8\pi}\int_{\mathbb{S}^2_\infty \cup \mathcal{E}} X_{ab} \\
&= -\frac{1}{8\pi}\int_{\Sigma} (\mathrm{d}X)_{abe} \\
&= -\frac{3}{8\pi}\int_{\Sigma} \nabla_{[e}(\epsilon_{ab]cd}\nabla^c t^d) \\
&= -\frac{1}{4\pi}\int_{\Sigma} R^d{}_f t^f \epsilon_{deab} \\
&= \frac{1}{4\pi}\int_{\Sigma} R_{ab} n^a t^b \mathrm{d}V \\
&= \frac{1}{4\pi}\int_{\Sigma} \left(T_{ab} - \frac{1}{2} T g_{ab}\right) n^a t^b \mathrm{d}V,
\end{aligned}
\tag{5.2.2}
$$

where m_{BH} corresponds to the Komar expression for the mass of a black hole. For the third equality we used Stokes' theorem[2] and for the fifth equality we used the Ricci identity for Killing fields $\nabla_a\nabla_b K_c = -R_{bca}{}^d K_d$. For the sixth equality we introduced n^a, the unit future pointing normal to Σ, so that $\epsilon_{abc} = n^d\epsilon_{dabc}$ is the natural volume form on Σ, represented by $\mathrm{d}V$. Finally, for the last equality we used the Einstein equation. By using (4.4.1) we find

$$
\begin{aligned}
m_{BH} &= -\frac{1}{8\pi}\int_{\mathcal{E}} \epsilon_{abcd}\nabla^c K^d + \frac{\Omega_E}{8\pi}\int_{\mathcal{E}} \epsilon_{abcd}\nabla^c \phi^d \\
&= -\frac{1}{8\pi}\int_{\mathcal{E}} \epsilon_{abcd}\nabla^c K^d + 2\Omega_E J,
\end{aligned}
\tag{5.2.3}
$$

where we interpreted $J = (1/16\pi)\int_{\mathcal{E}} \epsilon_{abcd}\nabla^c\phi^d$ as the angular momentum of the black hole. The first term of (5.2.3) may be evaluated as follows: The volume form on $\mathcal{E}$ may be written as $\epsilon_{ab} = \epsilon_{abcd} l^c K^d$, where l^a is the "ingoing" future directed null normal to $\mathcal{E}$, such that $l^a K_a = -1$. Thus, we have

$$
\epsilon^{ab}\epsilon_{abcd}\nabla^c t^d = l_e K_f \epsilon^{abef}\epsilon_{abcd}\nabla^c t^d = -4 l_c K_d \nabla^c K^d = -4\kappa,
\tag{5.2.4}
$$

where we used (C.0.2) for the second equality and (5.1.1) for the third equality. By using (C.0.1) we find

$$
-\frac{1}{8\pi}\int_{\mathcal{E}} \epsilon_{abcd}\nabla^c K^d = -\frac{1}{16\pi}\int_{\mathcal{E}} (\epsilon^{ef}\epsilon_{efcd}\nabla^c K^d)\epsilon_{ab} = \frac{1}{4\pi}\kappa\mathcal{A},
\tag{5.2.5}
$$

[2] In order to apply the Stokes' theorem, the orientations of S^2_∞ and $\mathcal{E}$ must be chosen appropriatly. Hence we have $m - m_{BH}$ instead of $m + m_{BH}$.

where $\mathcal{A} = \int_{\mathcal{E}} \epsilon_{ab}$ is the area of a horizon cross-section. Thus, we obtain

$$m = \frac{1}{4\pi}\int_{\Sigma}\left(T_{ab} - \frac{1}{2}Tg_{ab}\right)n^a t^a \mathrm{d}V + \frac{1}{4\pi}\kappa\mathcal{A} + 2\Omega_E J. \tag{5.2.6}$$

Remark 19. One should note that the Komar expression for the mass and angular momentum of a black hole only apply to black hole solutions of the Einstein equation which contain Killing fields t^a and ϕ^a which are stationary and axisymmetric, respectively.

In 1973 Bardeen, Carter and Hawking [3] derived a differential formula for m, i.e. a formula for how m changes when a small stationary, axisymmetric change is made in the solution. This differential formula is commonly refered to as *first law of black hole mechanics*. In the following we will only treat the vacuum case $T_{ab} = 0$. For a generalization, where the matter outside the black hole is modeled as a perfect fluid, see [3].

A formula for δm can be obtained by varying (5.2.6):

$$\delta m = \frac{1}{4\pi}(\mathcal{A}\delta\kappa + \kappa\delta\mathcal{A}) + 2(J\delta\Omega_E + \Omega_E\delta J). \tag{5.2.7}$$

But this is not the desired formula yet. A significantly longer computation shows (see [3]), that we can also express δm as

$$\delta m = -\frac{1}{4\pi}\mathcal{A}\delta\kappa - 2J\delta\Omega_E. \tag{5.2.8}$$

By adding (5.2.7) and (5.2.8) we obtain the following result:

Theorem 10 (First Law of Black Hole Mechanics)**.** *The variation of the total mass of two infinitesimally neighboring stationary, axisymmetric, vacuum black hole solutions can be expressed in terms of the horizon quantities κ, $\delta\mathcal{A}$, Ω_E and δJ_E by*

$$\delta m = \frac{\kappa}{8\pi}\delta\mathcal{A} + \Omega_E\delta J. \tag{5.2.9}$$

Remark 20. The original derivation of this law [3] required the perturbations to be stationary and made explicit use of the Einstein equation. This derivation can be generalized to hold for non-stationary perturbations [37], [32], provided that the change in area is evaluated on the bifurcation surface $\mathcal{C}$ of the unperturbed black hole. Furthermore, it has been shown [32] that the validity of this law does not depend on the details of the field equations. Specifically, a version of this law holds for field equations which were derived from a diffeomorphism covariant Lagrangian (see sections 6.3 and 7.2.1 for details).

Remark 21. This law stands in clear analogy with the first law of thermodynamics, $\delta E = T\delta S + p\delta V$. As we have already seen, the zeroth law indicates a relationship between the surface gravity κ and the temperature T. Thus, the variation formula (5.2.9) suggests an analogy between the horizon cross-section area $\mathcal{A}$ and the entropy S. This analogy is reinforced by the second law of black hole mechanics (see below), which asserts that $\mathcal{A}$ cannot decrease in any process.

5.3. Second Law

In section 4.4 we obtained a theorem about the event horizon of a strongly asymptotically predictable spacetime. The area theorem asserted that the horizon cross-section area is a nondecreasing quantitiy, i.e. if we consider two spacelike Cauchy surfaces Σ_1 and Σ_2 such that Σ_2 is contained in the chronological future of Σ_1, then we have $\mathcal{A}(\mathcal{E}_2) \geq \mathcal{A}(\mathcal{E}_1)$, where $\mathcal{E}_i = E \cap \Sigma_i$.

Thus, if we consider all black holes in the universe, their total cross-section area cannot decrease in any physically allowed process, $\delta\mathcal{A} \geq 0$. This implication of the area theorem is commonly refered to as the *second law of black hole mechanics*. It bears a resemblance to the second law of thermodynamics, which states that the total entropy S of all matter present in the universe cannot decrease in any physically allowed process, $\delta S \geq 0$.

At first sight, this resemblance seems to be a very superficial one, since the area theorem is a theorem in differential geometry whereas the second law of thermodynamics has a statistical origin. However, as we will see below, when quantum effects are taken into account, the mathematical analogy between $\mathcal{A}$ and S obtains physical significance. From this observation follows that $\mathcal{A}/4$ represents the physical entropy of a black hole.

Remark 22. In section 6.3 we will discuss why it is not possible to establish a second law by means of an area theorem in higher derivative theories of gravity.

5.4. Physical Relevance

As we have seen, there is a remarkable similarity between the physical laws governing the behavior of a thermodynamic system and the laws that describe the behavior of a black hole in general relativity. The zeroth law stated that κ is constant on E, the first law established the mass variation formula $\delta m = (\kappa/8\pi)\delta\mathcal{A} + \Omega_E \delta J$, and the second law asserted $\delta\mathcal{A} \geq 0$. These similarities with the laws of thermodynamics led Bekenstein [5] to propose the following identifications:

$$\mathscr{E} \leftrightarrow m, \qquad T \leftrightarrow \gamma\kappa, \qquad S \leftrightarrow \frac{1}{8\pi\gamma}\mathcal{A}, \tag{5.4.1}$$

where γ is an arbitrary, undetermined, real constant. Although $\mathscr{E}$ and M represent the same physical quantitiy, the other identifications remain on a formal level, since the temperature of a black hole (being a perfect absorber and emitting nothing) is absolut zero within the *classical* framework. Thus, it appears as if κ could not physically represent the temperature of a black hole. When quantum effects are taken into account this picture is drastically changed.

In 1975 Hawking [14] discovered that quantum effects cause the creation and emission of particles from a black hole with a blackbody spectrum at temperature $T = \kappa/2\pi$. Thus, $\kappa/2\pi$ does *physically* represent the thermodynamic temperature of a black hole, and is not merely a quantity playing a role mathematically analogous to the temperature of a black hole. This immediately suggests that $\mathcal{A}/4$ is the entropy of a black hole.

However, when black holes are discussed within a quantum context as above, the area theorem and the second law of black hole mechanics can be violated. For instance, the area of a black hole which evaporates due to Hawking radiation decreases to zero. But already the second law of ordinary thermodynamics fails in the presence of a black hole. When matter is dropped into a black hole, it will disappear into the spacetime singularity. All entropy initally present would be therefore lost and no compensating gain of entropy occurs. Therefore, the total entropy of the universe would decrease when matter falls into a black hole.[3] In 1974 Bekenstein [6] proposed a way to remedy these two problems simultaneously by introducing a *generalized entropy*

$$S_{\text{gen}} = S + \frac{\mathcal{A}}{4}, \tag{5.4.2}$$

where S represents the entropy of matter outside the black hole, and conjecturing that a *generalized second law* holds, i.e.

$$\delta S_{\text{gen}} \geq 0 \tag{5.4.3}$$

in any process. So when matter is dropped into a black hole, the decrease of S is accompanied by an increase of $\mathcal{A}$ (and vice versa), such that $\delta S_{\text{gen}} \geq 0$ remains valid. If this law turns out to be correct, the laws of black hole mechanics may be considered to be the ordinary laws of thermodynamics applied to a quantum system containing a black hole.

Remark 23. So far we have not mentioned if there exists an analog to the third law of thermodynamics, which states that $S \to 0$ (or a universal constant) as $T \to 0$, in black hole physics. The analog of this law fails in black mechanics since there exists extremal black holes ($\kappa = 0$) with finite $\mathcal{A}$. However, there do exist analogs of alternative versions of the third law which appear to hold for black holes [18].

[3] One could of course argue that one must still count the entropy of the matter after it fell into the black hole, as contributing to the total entropy of the universe. But then, the second law would have the status of being observationally unverifiable.

Part II.

HDTG and the Covariant Phase Space Formalism

6. Higher Derivative Theories of Gravity (HDTG)

6.1. General Relevance

On the classical level, the predictions of general relativity are in perfect accord with experiments, so there is no reason to modify this theory. However, since the experimental tests of general relativity only refer to the large-distance behaviour of the theory, we are free to add terms to the Einstein-Hilbert Lagrangean which leave this behaviour untouched. Such terms are for example R^2 or $R_{ab}R^{ab}$ (see below). Therefore, gravitational theories whose field equations contain derivatives of the metric of order greater than two were considered since the early days of general relativity, as possible other candidates for theories that describe the classical gravitational interaction.

Further motivation for the consideration of such HDTG is provided by attempts to quantize general relativity. The Einstein theory is perturbatively non-renormalizable at two loops in the vacuum case and at one loop for gravity interacting with matter [28], [8]. By adding suitable higher derivative terms to the gravitational Lagrangian, the ultraviolet behavior of the theory is improved [26]. Unfortunately, these modified theories contain, besides the usual massless spin-two excitation, an additional massive spin-two excitation with negative energy which leads to a breakdown of causality of the classical theory [27]. Furthermore, this additional excitation causes a loss of unitarity of the quantum theory [26]. Therefore, higher derivative theories have proven inadequate as a foundation for quantum gravity.

However, such theories might still be interesting within the context of effective field theories. It is expected that there exists a low energy effective action to a quantum theory of gravity. This action would yield field equations for a background metric field for sufficiently weak curvatures and sufficiently long distances. Presumably, this action will be generally covariant, and will consist of the Einstein-Hilbert action plus a series of higher curvature and higher derivative terms of the low energy matter fields. Such additional contributions naturally arise within the context of the renormalization of the stress-energy tensor of a quantum field propagating on a curved spacetime [17], and the constructions of an effective action for string theory [11]. Within this context, higher derivative theories of gravity such as Lovelock gravity, Gauss-Bonnet gravity for spacetimes with dimension $d > 4$, and polynomial-in-R gravity gained increased interest.

6.2. The Theory under Consideration

In the following we will consinder a vacuum HDTG which is given by the action[1]

$$I[g] = \frac{1}{16\pi}\int_M \mathrm{d}^4x\sqrt{-g}\left[R + \lambda R_{ab}R^{ab}\right] \tag{6.2.1}$$

where λ is a real constant with dimension length-squared. In contrast to the Einstein-Hilbert action, this action contains an additional Ricci tensor squared term. Therefore, the field equations will contain derivatives of the metric up to order four. In the following we will derive the equations of motion for this theory

Since the first term in (6.2.1) is the usual Einstein-Hilbert action, a variation[2] yields

$$\delta\left[\int_M \mathrm{d}^4x\sqrt{-g}\,R\right] = \int_M \mathrm{d}^4x\sqrt{-g}\left[R_{ab} - \frac{1}{2}g_{ab}R\right]\delta g^{ab}. \tag{6.2.2}$$

The variation of seconds may be written as

$$\delta\left[\int_M \mathrm{d}^4x\sqrt{-g}\,R_{ab}R^{ab}\right] = \int_M \mathrm{d}^4x(\delta\sqrt{-g})\,R_{ab}R^{ab} + \int_M \mathrm{d}^4x\sqrt{-g}\left[(\delta R_{ab})R^{ab} + R_{ab}\,\delta R^{ab}\right]. \tag{6.2.3}$$

By using the identities

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{ab}\,\delta g^{ab}, \tag{6.2.4}$$

and

$$\begin{aligned} R_{ab}\,\delta R^{ab} &= R_{ab}\,\delta(g^{ac}g^{bd}R_{cd}) \\ &= R_{ab}(\delta g^{ac})g^{bd}R_{cd} + R_{ab}g^{ac}(\delta g^{bd})R_{cd} + R_{ab}g^{ac}g^{bd}\,\delta R_{cd} \\ &= 2R_{ab}R_c{}^b\delta g^{ac} + R^{ab}\delta R_{ab}. \end{aligned} \tag{6.2.5}$$

we obtain

$$\delta\left[\int_M \mathrm{d}^4x\sqrt{-g}\,R_{ab}R^{ab}\right] = \int_M \mathrm{d}^4x\sqrt{-g}\left[-\frac{1}{2}g_{ab}R_{cd}R^{cd}\,\delta g^{ab} + 2R_{ac}R_b{}^c\,\delta g^{ab} + 2R^{ab}\,\delta R_{ab}\right]. \tag{6.2.6}$$

The variation of the Ricci tensor is given by the standard identity

$$\delta R_{ab} = \frac{1}{2}g^{cd}\left[\nabla_c\nabla_b\,\delta g_{ad} + \nabla_c\nabla_a\,\delta g_{bd} - \nabla_a\nabla_b\,\delta g_{cd} - \nabla_c\nabla_d\,\delta g_{ab}\right]. \tag{6.2.7}$$

Substitution of (6.2.7) into (6.2.6) yields

$$2\int_M \mathrm{d}^4x\sqrt{-g}\,R^{ab}\,\delta R_{ab} = \int_M \mathrm{d}^4x\sqrt{-g}\,R^{ab}g^{cd}\left[2\nabla_c\nabla_b\,\delta g_{ad} - \nabla_a\nabla_b\,\delta g_{cd} - \nabla_c\nabla_d\,\delta g_{ab}\right]. \tag{6.2.8}$$

[1] Actually, the prefactor $1/16\pi$ is not necessary for a vaccum theory. However, we still include it in order to obtain the correct form for the first law of black hole mechanics in this gravitational theory (see section 7.2.1). Furthermore, this prefactor assures that the entropy formula of Wald (equation (7.2.27)) yields $\mathcal{A}/4$ for $\lambda = 0$.

[2] See appendix A for a clarification of the notation in variational calculations.

In order to move the covariant derivatives onto the Ricci-tensor in (6.2.8), we perform a partial integration. Exemplary, for the first term in (6.2.8) we obtain

$$\begin{aligned}\int_M \mathrm{d}^4x\sqrt{-g}\, g^{cd}(\nabla_c\nabla_b\,\delta g_{ad})R^{ab} = &\int_M \mathrm{d}^4x\sqrt{-g}\,\nabla_c\Big[g^{cd}(\nabla_b\,\delta g_{ad})R^{ab}\Big] \\ &- \int_M \mathrm{d}^4x\sqrt{-g}\, g^{cd}(\nabla_b\,\delta g_{ad})(\nabla_c R^{ab}),\end{aligned} \tag{6.2.9}$$

where we used the Leibniz rule and the compatibility of the metric. By using Stokes theorem and the asymptotic boundary condition for the metric, we obtain

$$\int_M \mathrm{d}^4x\sqrt{-g}\,\nabla_c\Big[\underbrace{g^{cd}(\nabla_b\delta g_{ad})R^{ab}}_{=:w^c}\Big] = \int_{\partial M} w\cdot\boldsymbol{\epsilon} = 0, \tag{6.2.10}$$

where "·" denotes the contraction of the vector field w^a into the first index of the volume form $\boldsymbol{\epsilon}$ on M. By the integral over ∂M we mean a limiting process in which in the integral is first taken over the boundary, ∂K, of a compact region K in M (so that Stokes' theorem[3] can be applied) and then K approaches M in a suitable manner. By performing a partial integration twice in each term of (6.2.8) we obtain the following:

$$\begin{aligned}2\int_M \mathrm{d}^4x\sqrt{-g}\,R^{ab}(\delta R_{ab}) = \int_M \mathrm{d}^4x\sqrt{-g}\,g^{cd}\Big[&2(\delta g_{ad})\nabla_b\nabla_c R^{ab} - (\delta g_{cd})\nabla_b\nabla_a R^{ab} \\ &- (\delta g_{ab})\nabla_d\nabla_c R^{ab}\Big].\end{aligned} \tag{6.2.11}$$

Using the identity $\delta g_{ab} = -g_{ac}g_{bd}\,\delta g^{cd}$ yields

$$\begin{aligned}2\int_M \mathrm{d}^4x\sqrt{-g}\,R^{ab}(\delta R_{ab}) &= \int_M \mathrm{d}^4x\sqrt{-g}\,g^{cd}\Big[-2g_{aj}g_{dk}\nabla_b\nabla_c R^{ab} + g_{cj}g_{dk}\nabla_b\nabla_a R^{ab} \\ &\qquad + g_{aj}g_{bk}\nabla_d\nabla_c R^{ab}\Big]\delta g^{jk} \\ &= \int_M \mathrm{d}^4x\sqrt{-g}\,\Big[-2\nabla_b\nabla_k R_j{}^b + g_{jk}\nabla_b\nabla_a R^{ab} + \Box R_{jk}\Big]\delta g^{jk} \\ &= \int_M \mathrm{d}^4x\sqrt{-g}\,\Big[-2\nabla_c\nabla_b R_a{}^c + g_{ab}\nabla_d\nabla_c R^{cd} + \Box R_{ab}\Big]\delta g^{ab}.\end{aligned} \tag{6.2.12}$$

The first term in (6.2.12) can be simplified by using the identity

$$\nabla_d\nabla_c R^{cd} = \nabla^d\nabla_c R_d{}^c = \frac{1}{2}\nabla^d\nabla_d R = \frac{1}{2}\Box R. \tag{6.2.13}$$

From the commutator of covariant derivatives follows

$$\begin{aligned}\nabla_c\nabla_b R_a{}^c &= \nabla_b\nabla_c R_a{}^c + R_{cba}{}^d R_d{}^c - R_{cbd}{}^c R_a{}^d \\ &= \frac{1}{2}\nabla_b\nabla_a R - R_{acbd}R^{cd} + R_{bd}R_a{}^d.\end{aligned} \tag{6.2.14}$$

[3]We choose the orientation of ∂K to be the one specified by Stokes' theorem, i.e., we dot the first index of the orientation form on K into an outward pointing vector.

By substituting (6.2.13),(6.2.14) into (6.2.12) and the resulting expression into (6.2.6) we obtain

$$\begin{aligned}\delta\left[\int_M \mathrm{d}^4x\sqrt{-g}\,R_{ab}R^{ab}\right] &= \int_M \mathrm{d}^4x\sqrt{-g}\,\Big[-\nabla_b\nabla_a R + 2R_{acbd}R^{cd} - 2R_{bd}R_a{}^d \\ &\qquad + \frac{1}{2}g_{ab}\Box R + \Box R_{ab} + 2R_{ac}R_b{}^c - \frac{1}{2}g_{ab}R_{cd}R^{cd}\Big]\delta g^{ab} \\ &= \int_M \mathrm{d}^4x\sqrt{-g}\,\Big[-\nabla_b\nabla_a R + \Box R_{ab} + 2R_{acbd}R^{cd} \\ &\qquad - \frac{1}{2}g_{ab}(R_{cd}R^{cd} - \Box R)\Big]\delta g^{ab}.\end{aligned} \tag{6.2.15}$$

Since the variation δg^{ab} was chosen arbitrary, the equations of motion read as follows:

$$E_{ab} := R_{ab} - \frac{1}{2}g_{ab}R + \lambda\left[-\nabla_a\nabla_b R + \Box R_{ab} + 2R^{cd}R_{acbd} - \frac{1}{2}g_{ab}(R^{cd}R_{cd} - \Box R)\right] = 0. \tag{6.2.16}$$

The vacuum Einstein equation

$$R_{ab} - \frac{1}{2}Rg_{ab} = 0 \tag{6.2.17}$$

may be written as $R_{ab} = 0$, i.e. we have $R_{ab} = 0$ and $R = 0$ in the vacuum case. Therefore, any solution of the vacuum Einstein equation also solves the field equation (6.2.16) of our HDTG, so all vacuum spacetimes from the Einstein theory also appear the HDTG which we consider. However, may there may be an abundance of new solutions which are not present in general relativity. Among the vacuum spacetimes in Einstein gravity is the (maximally extended) Schwarzschild solution, which is asymptotically flat and contains a black hole region. Therefore, it is assured that black holes actually appear in the HDTG which we consider. However, our HDTG also has features which are not present in general relativity. In [27] it was shown that the static, linearized solutions of (6.2.16) are combinations of Newtonian and Yukawa potentials. Therefore, it is expected that the observational corrections of this theory set in at very small scales. Furthermore, we note that this theory possesses a a well posed inital value formulation (for a suitably defined initial data sets) [21].

6.3. Laws of Black Hole Mechanics in HDTG

A natural question to ask is what the status of the laws of black hole mechanics within the framework of effective field theories is. If one demands consistency of theses laws with the effective action, a preferred subclass of theories would be selected. In turn, this would place certain restrictions on the coefficients of the higher derivative contributions. From this analysis one might hope to learn something about the possible nature of quantum gravity. Furthermore, it is interesting to study black hole thermodynamics in such generalized gravitational theories in order to see whether the laws of black hole mechanics are a peculiar accident of Einstein gravity or a robust feature of all generally covariant theories of gravity, or something in between.

In the following we will summarize[4] the present status of the laws of black hole mechanics within the context of HDTG.

- The zeroth law of black hole mechanics states that the surface gravity κ is constant over the entire horizon. This statement has been proven for Einstein gravity with matter satisfying the dominant energy condition. The proof of this theorem heavily relies on Einstein's equation, so it does not readily extend to HDTG. If one assumes the existence of a bifurcate Killing horizon, then constancy of κ is easily seen to hold independently of the field equation [33]. Furthermore, in [20] a zeroth law is established for theories with gravitational Lagrangian $R + \lambda R^2$ without the assumption of a bifurcate Killing horizon.[5] If a zeroth law holds in general remains an open question.

- The first law of black hole mechanics (in the vacuum case) takes the form

$$\frac{\kappa}{2\pi}\delta S = \delta m - \Omega_E \delta J. \tag{6.3.1}$$

  For Einstein gravity, the black hole entropy S is given by one quarter of the horizon cross-section area, $S = \mathcal{A}/4$. A remarkable feature of (6.3.1) is that it relates variations in properties of the black hole as measured at asymptotic infinity to a variation of a geometric property of the horizon.

  The authors in [32] establish the result that, even though the precise expression of S is altered, the first law is still valid in an arbitrary diffeomorphism invariant theories of gravity. Such theories are given by diffeomorphism covariant Lagrangian densities of the form

$$L = L(g_{ab}, R_{abcd}, \nabla_a R_{bcde}, \dots, \psi, \nabla_a \psi, \dots), \tag{6.3.2}$$

  which depend on the metric g_{ab}, matter fields, collectively denoted by ψ, and a finite number of derivatives of the Riemann tensor and the matter fields (see chapter 7). Within this context, the black hole entropy is given by

$$S = -2\pi \int_{\mathcal{C}} \frac{\delta L}{\delta R_{abcd}} n_{ab} n_{cd}, \tag{6.3.3}$$

  where n_{ab} is the bi-normal to the bifurcation surface $\mathcal{C}$ and the functional derivative is evaluated by formally viewing the Riemann tensor as a field independent of the metric. As we see, the black hole entropy is still given by a local expression evaluated at the horizon, and so this aspect of the first law is preserved.

- In general relativity, the second law of black hole mechanics is established by the area theorem which states that the horizon cross-section area cannot decrease in any classical process, $\delta\mathcal{A} \geq 0$. An essential ingredient in the proof of this theorem is the null condition condition, $R_{ab}k^a k^b \geq 0$ for all null vectors k^a. This condition is implied by the Einstein's equation together with the null energy condition (see appendix B). In HDTG one can write the equations of motion in the form

$$R_{ab} - \frac{1}{2} g_{ab} R = 8\pi T_{ab}, \tag{6.3.4}$$

[4] as to our knowledge

[5] The idea of the proof is to relate the HDTG to a more conventional theory in which Einstein gravity is coupled to an auxiliary scalar field, using by a conformal field redefinition.

by absorbing the higher derivative terms in the energy-momentum tensor. Typically, these additional contributions spoil the null energy condtion, and so one cannot establish an area increase theorem in such theories using the standard techniques from general relativity.

However, this is not the relevant question for black hole thermodynamics. The relevant question is whether or not the quantity S, whose variation appears in the first law (6.3.1), satisfies an increase theorem. If so, one would have a second law of black hole thermodynamics for such a theory. This would further validate the interpretation of S as the black hole entropy.

In [20] a second law is established for quasistationary processes[6], independent of the details of the gravitational action. For such processes the second law is a direct consequence of the first law, as long as the matter stress-energy tensor satisfies the null energy condition. Furthermore, the authors proof a second law for theories whose gravitational Lagrangian is a polynomial in the Ricci scalar.

[6] These are dynamical processes where a small amount of matter enters from a great distance and falls into the (vacuum) black hole. The inital and final black holes are assumed to be stationary.

7. The Covariant Phase Space Formalism

7.1. Preliminaries

Let (M, g_{ab}) be a spacetime . In the following we will consider Lagrangian field theories on M of a vacuum type, so the only dynamical field that arises is the spacetime metric g_{ab}. By $\mathcal{F}$ we will denote the space of "kinematically allowed" metrics on the fixed manifold M. A precise definition of $\mathcal{F}$ would involve additional requirements on g_{ab} such as global hyperbolicity, the condition that a foliation of M is given by spacelike hypersurfaces and asymptotic fall-off conditions on the metric at spatial and/or null infinity. Therefore, the definition of $\mathcal{F}$ crucially depends of the theory under consideration and what is most suitable for one's purposes. In the following we will adopt a pragmatic point of view, in the sense that we assume that $\mathcal{F}$ has been chosen in a way that all integrals that occur below converge.

At the beginning of section 7.3 additional conditions[1] on $\mathcal{F}$ will be given which assure the convergence of all relevant integrals.

We will consider theories which are described by a Lagrangian density 4-form[2] locally constructed from the following quantities

$$\boldsymbol{L} = \boldsymbol{L}(g_{ab}, \overset{\circ}{\nabla}_{a_1} g_{bc}, \ldots, \overset{\circ}{\nabla}_{(a_1} \ldots \overset{\circ}{\nabla}_{a_k)} g_{bc}), \tag{7.1.1}$$

where $\overset{\circ}{\nabla}$ is an arbitrary, globally defined derivative operator and k is arbitrary but finite. The theories are assumed to be diffeomorphism invariant, i.e. the Lagrangian is diffeomorphism covariant in the sense that we have

$$\boldsymbol{L}(f^* \phi) = f^* \boldsymbol{L}(\phi), \tag{7.1.2}$$

for any diffeomorphism $f : M \to M$, where all variables appearing in (7.1.1) were collectively denoted by ϕ. The authors in [32] showed that condition (7.1.2) implies that $\boldsymbol{L}$ takes the form

$$\boldsymbol{L} = \boldsymbol{L}(g_{ab}, R_{abcd}, \nabla_{a_1} R_{bcde}, \ldots, \nabla_{(a_1} \ldots \nabla_{a_m)} R_{bcde}), \tag{7.1.3}$$

where ∇ is the derivative operator associated with g_{ab}, $m = k - 2$ and R_{abcd} is the Riemann tensor of g_{ab}.

A variation[3] of L can be expressed as

$$\delta \boldsymbol{L} = \boldsymbol{E} \cdot \delta g + \mathrm{d}\boldsymbol{\theta}, \tag{7.1.4}$$

[1] All asymptotically flat spacetimes at null infinity in vacuum general relativity satisfy these conditions.

[2] Actually it is more standard to consider Lagrangian density scalars in field theories. But we can use the Hodge dual, defined via the volume form $\boldsymbol{\epsilon}$, to convert the Lagrangian density scalar L to a Lagrangian density 4-form $\boldsymbol{L} = L\boldsymbol{\epsilon}$ and for our purposes it will be more convenient to view the Lagrangian density as a 4-form. See appendix A for an explanation of the boldface-notation.

[3] See appendix A for a clarification of the notation in variational calculation.

with

$$\boldsymbol{E} = \boldsymbol{E}(g), \qquad \boldsymbol{\theta} = \boldsymbol{\theta}(g, \delta g). \tag{7.1.5}$$

and

$$\boldsymbol{E} \cdot \delta g = (\boldsymbol{E})_{ab}\, \delta g^{ab} = E_{ab}(\delta g^{ab})\boldsymbol{\epsilon}. \tag{7.1.6}$$

The equations of motion of the theory are then simply $\boldsymbol{E} = 0$. The 3-form $\boldsymbol{\theta}$ is called *presymplectic potential*. Note that $\boldsymbol{\theta}$ corresponds to the boundary term that arises from the integration by parts in order to remove the derivatives from δg_{ab} if the variation is performed under an integral sign. Even though $\boldsymbol{E}$ is uniquely determined by (7.1.4), the symplectic potential is only unique up to addition of a closed 3-form. Since the symplectic potential is required to be locally constructed out of the metric g and the perturbation δg in a covariant manner, the freedom in the choice of $\boldsymbol{\theta}$ is limited to

$$\boldsymbol{\theta} \to \boldsymbol{\theta} + \mathrm{d}\boldsymbol{Y}, \tag{7.1.7}$$

where $\boldsymbol{Y}$ is locally constructed out of g and δg in a covariant manner.

The *presymplectic current* 3-form $\boldsymbol{\omega}$ is defined via the antisymmetrized variation[4] of $\boldsymbol{\theta}$, i.e.

$$\boldsymbol{\omega}(g, \delta_1 g, \delta_2 g) := \delta_1 \boldsymbol{\theta}(g, \delta_2 g) - \delta_2 \boldsymbol{\theta}(g, \delta_1 g). \tag{7.1.8}$$

Note that $\boldsymbol{\omega}$ is a local function of g and the two linearized perturbations $\delta_1 g$ and $\delta_2 g$ off of g. The ambiguity (7.1.7) in the choice of $\boldsymbol{\theta}$ leads to the ambiguity

$$\boldsymbol{\omega} \to \boldsymbol{\omega} + \mathrm{d}[\delta_1 \boldsymbol{Y}(g, \delta_2 g) - \delta_2 \boldsymbol{Y}(g, \delta_1 g)] \tag{7.1.9}$$

in the choice of $\boldsymbol{\omega}$.

Let Σ be a closed, embedded 3-dimensional submanifold without boundary; we will refer to Σ as a *slice*. The orientation of Σ is chosen to be $\tilde{\epsilon}_{a_1 a_2 a_3} = n^b \epsilon_{b a_1 a_2 a_3}$, where n^a is the future pointing normal to Σ and $\epsilon_{b a_1 a_2 a_3}$ is the positively oriented volume form on M. We can define a 2-form on $\mathcal{F}$ via

$$\Omega_\Sigma(g, \delta_1 g, \delta_2 g) := \int_\Sigma \boldsymbol{\omega}. \tag{7.1.10}$$

From the definition of $\boldsymbol{\omega}$ follows that Ω_Σ is antisymmetric in the perturbations. Therefore, Ω_Σ is *presymplectic form* on $\mathcal{F}$ associated with Σ. Although this definition depends, in general, on the choice of Σ, it can be shown that if $\delta_1 g$ and $\delta_2 g$ satisfy the linearized equations of motion and Σ is a Cauchy surface, then Ω_Σ does not depend on the choice of Σ, provided that Σ is compact or suitable asymptotic conditions are imposed on g (see [34]).

The ambiguity (7.1.9) in the choice of $\boldsymbol{\omega}$ gives rise to the ambiguity

$$\Omega_\Sigma(g, \delta_1 g, \delta_2 g) \to \Omega_\Sigma(g, \delta_1 g, \delta_2 g) + \int_{\partial\Sigma} [\delta_1 \boldsymbol{Y}(g, \delta_2 g) - \delta_2 \boldsymbol{Y}(g, \delta_1 g)] \tag{7.1.11}$$

in the presymplectic form Ω_Σ. By the integral over $\partial\Sigma$ above (Σ is assumed to have no boundary) we mean a limiting process in the sense that the integral is first taken over ∂K, of a compact region K, of Σ and then K approaches all of Σ. The orientation of ∂K is chosen to be $n^a \epsilon_{b a_1 a_2 a_3}$, where n^a is an outward pointing vector and $\epsilon_{b a_1 a_2 a_3}$ is the volume form on M, such that Stokes' theorem can be applied. Note that the right hand side of (7.1.11) is only

[4] Here and in the following we assume that all variations commute, i.e. $\delta_1 \delta_2 g - \delta_2 \delta_1 g = 0$.

well defined if the limit exists and is independent of the details of how K approaches Σ. (At the beginning of section 7.3 additional assumptions will be made which assure convergence of integrals over "$\partial\Sigma$".)

Given the presymplectic form Ω_Σ, it is possible to construct a phase space Γ by factoring out the orbits of the degeneracy subspaces of Ω_Σ (for detail of the construction see [34]). This phase space naturally acquires a genuine symplectic form from Ω_Σ. However, for our purposes it will be sufficient to work with the original field configuration space $\mathcal{F}$ and its (degenerate) presymplectic form Ω_Σ. In the following, the subspace of $\mathcal{F}$ where the equations of motions are satisfied will be denoted by $\bar{\mathcal{F}}$. The space $\bar{\mathcal{F}}$ is called *covariant phase space*.

Remark 24. A perturbation δg_0 off of g_0 is a tangent vector at $g_0 \in \mathcal{F}$ in the following sense: A one-parameter family of metrics g_t corresponds to a curve $\mathbb{R} \ni t \mapsto g_t = g(t) \in \mathcal{F}$, which gives rise the tangent vector $\delta g = \mathrm{d}g(t)/\mathrm{d}t|_{t=0} \in T_{g(0)}\mathcal{F}$ with $g = g(0) = g_0$.

Remark 25. A variation δg which is tangent to $\bar{\mathcal{F}}$ always satisfies the linearized equations of motion: The tangent vector $\delta g = \mathrm{d}g_t/\mathrm{d}t|_{t=0} = \mathrm{d}g(t)/\mathrm{d}t|_{t=0}$ defines a curve $t \mapsto g_t = g(t)$ such that $g(t) \in \bar{\mathcal{F}}$ for each $t \in \mathbb{R}$. Therefore we have $\boldsymbol{E}(g(t)) = 0$ for each $t \in \mathbb{R}$ and in particular $\mathrm{d}\boldsymbol{E}(g(t))/\mathrm{d}t|_{t=0} = 0$. These are the linearized equations of motion with solutions $\mathrm{d}g(t)/\mathrm{d}t|_{t=0} = \delta g$.

Remark 26. A complete vector field ξ^a on M naturally induces a field variation $\delta_\xi g = \mathcal{L}_\xi g$ in the following sense: The flow ϕ_t, generated by ξ^a, induces the action $g \to \phi_t^* g = g(t)$ on $\mathcal{F}$. The curve $t \mapsto g(t)$ gives rise to the tangent vector $\mathrm{d}g(t)/\mathrm{d}t|_{t=0} = \mathrm{d}(\phi_t^* g)/\mathrm{d}t|_{t=0} = \mathcal{L}_\xi g$. This vector is tangent to $\mathcal{F}$ if the flow Φ_s, generated by $\mathcal{L}_\xi g$, is a diffeomorphism which maps $\mathcal{F}$ into itself for each $s \in \mathbb{R}$ (see [34]).

Remark 27. The vector field $\mathcal{L}_\xi g$ on $\mathcal{F}$ always satisfies the linearized equations of motion if g satifies the equations of motion: Since $\boldsymbol{L}$ is diffeomorphism covariant, $\phi_t^* g$ satisfies the equations of motion, i.e. $\boldsymbol{E}(\phi_t^* g) = 0$, if g satisfies the equations of motion. Therefore we have $\mathrm{d}\boldsymbol{E}(\phi_t^* g)/\mathrm{d}t|_{t=0} = 0$ which are the linearized equations of motion with solutions $\mathrm{d}(\phi_t^* g/\mathrm{d}t)|_{t=0} = \mathcal{L}_\xi g$.

The vector field $\delta_\xi g = \mathcal{L}_\xi g$ may be viewed as a dynamical evolution vector field on $\mathcal{F}$, corresponding to the notion of "time translation" defined by ξ^a. Its role is analogous to the Hamiltonian vector field in classical mechanics and motivates the next definition.

Definition 25. Consider a diffeomorphism invariant theory as in the above framework with field configurations space $\mathcal{F}$ and solution subspace $\bar{\mathcal{F}}$. Let ξ^a be vector field on M, let Σ be a slice in M and let Ω_Σ be the presymplectic form defined by (7.1.10). (If the ambiguity in the choice of $\boldsymbol{\omega}$ gives rise to an ambiguity in Ω_Σ according to (7.1.11), then we assume that a particular choice of Ω_Σ has been made.) Furthermore, we assume that $\mathcal{F}$, ξ^a and Σ have been chosen in a way such that the integral $\int_\Sigma \boldsymbol{\omega}(g, \delta g, \mathcal{L}_\xi g)$ converges for all $g \in \bar{\mathcal{F}}$ and all tangent vectors δg to $\bar{\mathcal{F}}$ at g. Then, a function $H_\xi : \mathcal{F} \to \mathbb{R}$ is said to be a *Hamiltonian conjugate to* ξ^a on slice Σ, if for all $g \in \bar{\mathcal{F}}$ and field variations δg tangent to $\mathcal{F}$ we have

$$\delta H_\xi = \Omega_\Sigma(g, \delta g, \mathcal{L}_\xi g) = \int_\Sigma \boldsymbol{\omega}(g, \delta g, \mathcal{L}_\xi g). \tag{7.1.12}$$

Note that if there exists such a function H_ξ, its value on $\bar{\mathcal{F}}$ is only determined up to addition of an arbitrary constant by (7.1.12). This constant can be fixed by requiring that H_ξ vanishes for a reference solution, such as Minkowski spacetime. The value of H_ξ off of $\bar{\mathcal{F}}$ is essentially arbitrary.

Furthermore, there does not need to exist a function H_ξ at all which satisfies (7.1.12). For instance, this is the case in general relativity when ξ^a is an asymptotic time translation and the slice Σ extends to null infinity. It was shown in [38], that a necessary *and* sufficient condition for the existence of a Hamiltonian H_ξ conjugate to ξ^a on Σ is that for all solutions $g \in \bar{\mathcal{F}}$ and all pairs of perturbations $\delta_1 g, \delta_2 g$ tangent to $\bar{\mathcal{F}}$ we have

$$\int_{\partial\Sigma} \xi \cdot \boldsymbol{\omega}(g, \delta_1 g, \delta_2 g) = 0, \tag{7.1.13}$$

where "$\cdot$" denotes the contraction of ξ^a into the first index the differential form $\boldsymbol{\omega}$. There are two situation in which (7.1.13) is automatically satisfied:

(i) The asymptotic conditions on g are such that $\boldsymbol{\omega}(g, \delta_1 g, \delta_2 g)$ goes to zero sufficiently rapid such that the integral of $\xi \cdot \boldsymbol{\omega}$ over ∂K vanishes in the limit as K approaches Σ.

(ii) If ξ^a is such that K can always be chosen such that ξ^a is tangent to ∂K, since then the pullback of $\xi \cdot \boldsymbol{\omega}$ to ∂K vanishes.

The value of H_ξ provides a natural candidate for a conserved quantity associated with ξ^a at "time" Σ. In section 7.3 we will investigate the issue of defining "conserved quantities" even when no Hamiltonian exists.

7.2. Black Hole Entropy as Noether Charge

First of all, let us introduce some further useful quantities. The *Noether current* 3-form associated with ξ^a is defined by

$$\boldsymbol{J} = \boldsymbol{\theta}(g, \mathcal{L}_\xi g) - \xi \cdot \boldsymbol{L}. \tag{7.2.1}$$

The standard identity

$$\mathcal{L}_\xi \boldsymbol{\Lambda} = \mathrm{d}[\xi \cdot \boldsymbol{\Lambda}] + \xi \cdot \mathrm{d}\boldsymbol{\Lambda}, \tag{7.2.2}$$

which holds for any vector field ξ^a and differential form $\boldsymbol{\Lambda}$, together with (7.1.4) implies that we have

$$\mathrm{d}\boldsymbol{J} = -\boldsymbol{E} \cdot \mathcal{L}_\xi g. \tag{7.2.3}$$

Therefore, $\boldsymbol{J}$ is closed whenever the equations of motion are satisfied. Furthermore, $\boldsymbol{J}$ is not only closed but also exact if $\boldsymbol{E}(g) = 0$ holds [30]. From this follows that there exists a 2-form $\boldsymbol{Q}$, locally constructed from g and ξ^a, such that whenever $\boldsymbol{E}(g) = 0$, we have

$$\boldsymbol{J} = \mathrm{d}\boldsymbol{Q}. \tag{7.2.4}$$

When the equations of motion are not satisfied, the Noether current may be written as

$$\boldsymbol{J} = \mathrm{d}\boldsymbol{Q} + \xi^a \boldsymbol{C}_a, \tag{7.2.5}$$

where $\boldsymbol{C}_a$ are the "constraints" of the theory, i.e. we have $\boldsymbol{C}_a = 0$ whenever the equations of motion are satisfied. The quantity $\boldsymbol{Q} = \boldsymbol{Q}[\xi]$ appearing in (7.2.4) is the *Noether charge* 2-form.

In [32] it was shown that $\boldsymbol{Q}$ can always be written in the form

$$\boldsymbol{Q} = \boldsymbol{X}^{ab}(g)\nabla_{[a}\xi_{b]} + \boldsymbol{U}_a(g)\xi^a + \boldsymbol{V}(g, \mathcal{L}_\xi g) + \mathrm{d}\boldsymbol{Z}(g,\xi), \tag{7.2.6}$$

where $\boldsymbol{X}^{ab}$,$\boldsymbol{U}_a$, $\boldsymbol{V}$ and $\boldsymbol{Z}$ are covariantly constructed from the indicated quantities and their derivatives (with $\boldsymbol{V}$ linear in $\mathcal{L}_\xi g$ and $\boldsymbol{Z}$ linear in ξ). In particular, the first term in 7.2.4 is given by

$$\boldsymbol{X}^{cd} = (X^{cd})_{c_1c_2} = -E_R^{abcd}\epsilon_{abc_1c_2}, \tag{7.2.7}$$

with

$$E_R^{abcd} = \frac{\partial L}{\partial R_{abcd}} - \nabla_{a_1}\frac{\partial L}{\partial \nabla_{a_1} R_{abcd}} + \cdots + (-1)^m \nabla_{(a_1}\ldots\nabla_{a_m)}\frac{\partial L}{\partial\nabla_{(a_1}\ldots\nabla_{a_m)}R_{abcd}}. \tag{7.2.8}$$

In fact, (7.2.8) are the equations of motion for R_{abcd} *if* it were viewed as a field independent of the metric.

The quantities $\boldsymbol{J}$ and $\boldsymbol{Q}$ inherit the following ambiguities from (7.1.7):

$$\boldsymbol{J} \to \boldsymbol{J} + \mathrm{d}\boldsymbol{Y}(g, \mathcal{L}_\xi g) \tag{7.2.9}$$
$$\boldsymbol{Q} \to \boldsymbol{Q} + \boldsymbol{Y}(g, \mathcal{L}_\xi g) + \mathrm{d}\boldsymbol{W}, \tag{7.2.10}$$

where $\boldsymbol{W}$ is a 1-form locally constructued in a covariant manner.

7.2.1. Application to the First Law

Wald and Iyer [32] used the covariant phase space formalism to show that a version of the first law of black hole mechanics holds in every diffeomorphism invariant theory of gravity. In the following, we will illustrate their line of argument.

Consider some $g \in \bar{\mathcal{F}}$ and an arbitrary variation δg off of g (not necessarily tangent to $\bar{\mathcal{F}}$). Let ξ^a be a complete, fixed vector field on M. Then, we have

$$\begin{aligned}\delta\boldsymbol{J} &= \delta\boldsymbol{\theta}(g, \mathcal{L}_\xi g) - \xi\cdot\delta\boldsymbol{L}\\ &= \delta\boldsymbol{\theta}(g, \mathcal{L}_\xi g) - \xi\cdot\mathrm{d}\boldsymbol{\theta}(g,\delta g)\\ &= \delta\boldsymbol{\theta}(g, \mathcal{L}_\xi g) - \mathcal{L}_\xi\boldsymbol{\theta}(g,\delta g) + \mathrm{d}[\xi\cdot\boldsymbol{\theta}(g,\delta g)],\end{aligned} \tag{7.2.11}$$

where we used (7.1.4) and $\boldsymbol{E} = 0$ in the second line and the identity (7.2.2) in the third line. Since $\boldsymbol{\theta}$ is covariant, $\mathcal{L}_\xi\boldsymbol{\theta}$ is the same as the variation induced in $\boldsymbol{\theta}$ by the field variation $\delta' g = \mathcal{L}_\xi g$. Therefore, we have

$$\delta\boldsymbol{\theta}(g, \mathcal{L}_\xi g) - \mathcal{L}_\xi\boldsymbol{\theta}(g,\delta g) = \boldsymbol{\omega}(g, \delta g, \mathcal{L}_\xi g), \tag{7.2.12}$$

where we used the definition (7.1.8). From this follows that (7.2.11) reads as

$$\boldsymbol{\omega}(g,\delta g,\mathcal{L}_\xi g) = \delta\boldsymbol{J} - \mathrm{d}[\xi\cdot\boldsymbol{\theta}]. \tag{7.2.13}$$

By using (7.2.5), this can be rewritten as

$$\boldsymbol{\omega}(g,\delta g,\mathcal{L}_\xi g) = \xi^a\delta\boldsymbol{C}_a + \mathrm{d}\delta\boldsymbol{Q} - \mathrm{d}[\xi\cdot\boldsymbol{\theta}], \tag{7.2.14}$$

where we used the fact that $\delta \mathrm{d}\boldsymbol{Q} = \mathrm{d}\delta\boldsymbol{Q}$ holds. Therefore, if there exists a Hamiltonian conjugate to ξ^a on Σ, then it must satisfy

$$\delta H_\xi = \int_\Sigma \xi^a \delta \boldsymbol{C}_a + \int_{\partial\Sigma} (\delta\boldsymbol{Q} - \xi \cdot \boldsymbol{\theta}), \tag{7.2.15}$$

for all $g \in \bar{\mathcal{F}}$ and all δg. The integral over $\partial\Sigma$ has the meaning as described below equation (7.1.11). When δg satisfies the linearized equations of motion, i.e. δg is tangent to $\bar{\mathcal{F}}$, then (7.2.15) takes the form

$$\delta H_\xi = \int_{\partial\Sigma} (\delta\boldsymbol{Q} - \xi \cdot \boldsymbol{\theta}). \tag{7.2.16}$$

As we will show in the following, this equation can be used to define conserved quantities which are associated with asymptotic symmetries generated by ξ^a. If we can find a 3-form $\boldsymbol{B}$, such that

$$\delta \int_{\partial\Sigma} \xi \cdot \boldsymbol{B} = \int_{\partial\Sigma} \xi \cdot \boldsymbol{\theta}, \tag{7.2.17}$$

then the Hamiltonian H is given by

$$H = \int_{\partial\Sigma} (\boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{B}). \tag{7.2.18}$$

Now, let g be a solution which corresponds to an asymptotically flat spacetime and let Σ be a slice, which extends to spatial infinity, such that $\partial\Sigma = S^2_\infty$, where S^2_∞ is a two-sphere at spatial infinity. First of all, let us assume that the asymptotic conditions on g have been specified in such a way that ξ^a is an asymptotic time translation, $\boldsymbol{B}$ exists and the surface integral in (7.2.18) approaches a finite limit. Then, the *canonical energy* $\mathscr{E}$ of an asymptotically flat spacetime may be defined as

$$\mathscr{E} = \int_{S^2_\infty} (\boldsymbol{Q}[t] - t \cdot \boldsymbol{B}), \tag{7.2.19}$$

where t^a is an asymptotic time translation.

Consider now case where ξ^a is an asymptotic rotation ϕ^a. We can choose the surface S^2_∞ in such a way that ϕ^a is everywhere tangent to S^2_∞, such that the pullback of $\phi \cdot \boldsymbol{\theta}$ vanishes. Then, the *canonical angular momentum* J of an asymptotically flat spacetime can be defined as

$$J = -\int_{S^2_\infty} \boldsymbol{Q}[\phi]. \tag{7.2.20}$$

It is assumed that the asymptotic conditions on the metric g have been specified in such a way that this surface integral converges.

We will now apply equation (7.2.16) to the case of a stationary black hole solution with bifurcate Killing horizon. This will directly lead us to a generalized first law. Let ξ^a be a Killing field that vanishes on the bifurcation surface $\mathcal{C}$, normalized such that

$$\xi^a = t^a + \Omega_E \phi^a, \tag{7.2.21}$$

where t^a is a stationary Killing field with unit norm at infinity. Since we have $\mathcal{L}_\xi g = 0$, the left hand side of (7.2.13) vanishes as $\boldsymbol{\omega}(g, \delta_1 g, \delta_2 g)$ is linear in $\delta_2 g$ (see [34]). Therefore, equation

(7.2.16) reads as

$$0 = \int_{\partial\Sigma} (\delta \boldsymbol{Q} - \xi \cdot \boldsymbol{\theta}). \tag{7.2.22}$$

Furthermore, let Σ be an asymptotically flat hypersurface that extends from the bifurcation surface $\mathcal{C}$ to S^2_∞, such that $\partial\Sigma = \mathcal{C} \cup S^2_\infty$. Then, we have

$$\begin{aligned} 0 &= \int_{\partial\Sigma} (\delta \boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{\theta}) \\ &= \delta \int_{\partial\Sigma} (\boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{B}) \\ &= \delta \int_{\mathcal{C}} (\boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{B}) - \delta \int_{S^2_\infty} (\boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{B}) \\ &= \delta \int_{\mathcal{C}} \boldsymbol{Q}[\xi] - \delta \int_{S^2_\infty} (\boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{B}), \end{aligned} \tag{7.2.23}$$

where we used the fact that ξ^a vanishes on $\mathcal{C}$ for the fouth equality. From the form of the Noether current (7.2.6) follows

$$\boldsymbol{Q} = \boldsymbol{X}^{ab}(g) \nabla_{[a}\xi_{b]}, \tag{7.2.24}$$

since we have $\mathcal{L}_\xi g = 0$ and $\xi^a \restriction \mathcal{C} = 0$. Therefore, insertion of (7.2.21) into (7.2.23) yields

$$\begin{aligned} \delta \int_{\mathcal{C}} \boldsymbol{Q}[\xi] &= \delta \int_{S^2_\infty} (\boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{B}) \\ &= \delta \int_{S^2_\infty} (\boldsymbol{Q}[t] + \Omega_E \boldsymbol{Q}[\phi] - t \cdot \boldsymbol{B} - \Omega_E \phi \cdot \boldsymbol{B}) \\ &= \delta \int_{S^2_\infty} (\boldsymbol{Q}[t] - t \cdot \boldsymbol{B}) + \Omega_E \delta \int_{S^2_\infty} \boldsymbol{Q}[\phi] \\ &= \delta \mathscr{E} - \Omega_E \delta J. \end{aligned} \tag{7.2.25}$$

Here, we used the fact that S^2_∞ can be chosen in such a way that ϕ^a is tangent to S^2_∞, so that the pullback of $\phi \cdot \boldsymbol{\theta}$ to S^2_∞ vanishes for the third equality. One can show (see [32]), that the left hand side of (7.2.25) may be written as

$$\delta \int_{\mathcal{C}} \boldsymbol{Q}[\xi] = \kappa \delta \int_{\mathcal{C}} \boldsymbol{X}^{cd} n_{cd}, \tag{7.2.26}$$

where κ is the surface gravity of the black hole and n_{cd} is the binormal to $\mathcal{C}$, i.e. n_{cd} is the natural volume element on the tangent space perpendicular to $\mathcal{C}$, oriented so that $n_{cd} T^c R^d > 0$ when T^a is a future-directed timelike vector and R^a is a spacelike vector that points "towards infinity". This result establishes the following theorem, which is due Wald and Iyer [32].

Theorem 11 (Generalized First Law of Black Hole Mechanics)**.** *Let g be an asymptotically flat stationary black hole solution of an arbitrary diffeomorphism invariant theory of gravity with a bifurcate Killing horizon $\mathcal{C}$. Let δg be a (not necessarily stationary), asymptotically flat solution of the linearized equations of motion about g. If we define S by*

$$S = 2\pi \int_{\mathcal{C}} \boldsymbol{X}^{cd} n_{cd}, \tag{7.2.27}$$

then we have

$$\frac{\kappa}{2\pi}\delta S = \delta \mathscr{E} - \Omega_E \delta J. \tag{7.2.28}$$

Remark 28. In the above discussion we explicitely assumed that $\mathcal{C}$ is the bifurcation surface of a bifurcate Killing horizon. However, in [19] it was shown that for stationary black holes with a bifurcate horizon the integral of $\boldsymbol{Q}$ is independent of the choice of cross-section. Namely, if we define the entropy S, for an arbitrary cross-section $\mathcal{E} = E \cap \Sigma$, of a stationary black hole by

$$S[\mathcal{E}] = 2\pi \int_{\mathcal{C}'} \boldsymbol{X}^{cd} n'_{cd}, \tag{7.2.29}$$

where n'_{cd} is the binormal to $\mathcal{E}$, then S is independent of the choice of $\mathcal{E}$. In order to see this, one has to recognize that $\boldsymbol{X}^{cd}$ is invariant under the one-parameter group of isometries χ_t, generated by the vector field ξ^a. From this follows immediately $S[\chi_t(\mathcal{E})] = S[\mathcal{E}]$. Since we have $\chi_t(\mathcal{E}) \overset{t\to-\infty}{\longrightarrow} \mathcal{C}$, and since $\boldsymbol{X}^{cd}$ is smooth, we obtain $S[\mathcal{E}] = S[\mathcal{E}]$.

Therefore, for stationary perturbations, the first law holds when S taken to be the entropy of an arbitrary cross-section. When non-stationary perturbations are considered, it is essential to evaluate S on the bifurcation surface, in order to establish a first law.

7.2.2. Black Hole Entropy in our HDTG

In this section we will use the above framework to calculate the black hole entropy S_λ in our HDTG.

The key formula for this calculation is equation (7.2.27), namely

$$S_\lambda = 2\pi \int_{\mathcal{C}} \boldsymbol{X}^{cd} n_{cd} = -2\pi \int_{\mathcal{C}} E_R^{abcd} \epsilon_{abc_1c_2} n_{cd} = -2\pi \int_{\mathcal{C}} E_R^{abcd} n_{ab} n_{cd} \overset{(2)}{\boldsymbol{\epsilon}}, \tag{7.2.30}$$

where $\overset{(2)}{\boldsymbol{\epsilon}}$ denotes the induced volume-form on the 2-dimensional submanifold $\mathcal{C}$. The theories we are considering are given by the Lagrangian

$$\boldsymbol{L} = L\boldsymbol{\epsilon} = \frac{1}{16\pi}(R + \lambda R_{ab}R^{ab})\boldsymbol{\epsilon}. \tag{7.2.31}$$

Since this Lagrangian does not depend on derivatives of the Riemann tensor, we have

$$E_R^{abcd} = \frac{\partial L}{\partial R_{abcd}}. \tag{7.2.32}$$

A straightforward calculation yields

$$\frac{\partial L}{\partial R_{abcd}} = \frac{1}{16\pi}(g^{ac}g^{bd} + 2\lambda g^{bd}R^{ac}). \tag{7.2.33}$$

The binormal is chosen to be $n_{ab} = n_a l_b - l_a n_b$, where n^a and l^a are two linearly independent null vectors, which are normalized such that $n_a l^a = -1$. Therefore we have

$$\begin{aligned} n_{ab}n_{cd} &= (n_a l_b - l_a n_b)(n_c l_d - l_c n_d) \\ &= n_a l_b n_c l_d - n_a l_b l_c n_d - l_a n_b n_c l_d + l_a n_b l_c n_d. \end{aligned} \tag{7.2.34}$$

Insertion of (7.2.33) and (7.2.34) into (7.2.30) yields

$$\begin{aligned} S_\lambda &= -\frac{1}{8}\int_{\mathcal{C}} (g^{ac}g^{bd} + 2\lambda g^{bd}R^{ac}) n_{ab} n_{cd} \overset{(2)}{\boldsymbol{\epsilon}} \\ &= \frac{1}{4}\int_{\mathcal{C}} \overset{(2)}{\boldsymbol{\epsilon}} - \frac{\lambda}{2}\int_{\mathcal{C}} R_{ab}(n^a l^b + l^a n^b) \overset{(2)}{\boldsymbol{\epsilon}} \\ &= \frac{\mathcal{A}(\mathcal{C})}{4} - \lambda \int_{\mathcal{C}} R_{ur} \overset{(2)}{\boldsymbol{\epsilon}}, \end{aligned} \tag{7.2.35}$$

where we used the fact that the vectors n^a and l^a may be written as $n^a = (\partial/\partial u)^a$ and $l^a = (\partial/\partial r)^a$ in Gaussian null coordinates for the last equality.

As we see, the black hole entropy in our HDTG is given by the usual $\mathcal{A}/4$ term from general relativty, plus an additional contribution which is given by an integral of the ur-component of the Ricci-tensor over the bifurcation surface $\mathcal{C}$.

7.3. Generalized "Conserved Quantities"

At the end of section 7.1 we stated a condition, (7.1.13), which assured the existence of a Hamiltonian. However, in many cases of interest this condition is not satisfied, and therefore it is not possible to define conserved quantities. Wald and Zoupas [38] developed a technique for defining conserved quantities, even when no Hamiltonian exists. In the following, we will briefly summarize this method.

In section 8.2 this technique will be used for an attempts to establish a second law of black hole mechanics in our HDTG.

First of all, let us introduce some terminology and the basic assumptions of this framework. We consider a diffeomorphism invariant theory of gravity, whose asymptotic conditions are specified by attaching a boundary $\mathcal{B}$ to the spacetime manifold M and requiring a certain limiting behaviour of the metric g, as one approaches $\mathcal{B}$. The boundary $\mathcal{B}$ is assumed to be a 3-dimensional manifold , so that $M \cup \mathcal{B}$ is a 4-dimensional manifold with boundary. $M \cup \mathcal{B}$ will be equipped with additional non-dynamical structure - such as a conformal factor on $M \cup \mathcal{B}$ or other tensor fields - which will enter into the specification of the limiting behaviour of g, and will therefore be part of the definition of $\mathcal{F}$ and $\bar{\mathcal{F}}$. This additional non-dynamical structure will be refered to as *universal background structure* of $M \cup \mathcal{B}$.

The following two main assumptions are made:

1. $\mathcal{F}$ has been defined so that for all $g \in \bar{\mathcal{F}}$ and all $\delta_1 g, \delta_2 g$ tangent to $\bar{\mathcal{F}}$ the presymplectic current $\boldsymbol{\omega}(g, \delta_1 g, \delta_2 g)$ extends continuously to $\mathcal{B}$.

2. One only considers slices Σ, that extend smoothly to $\mathcal{B}$, such that the extended hypersurface intersects $\mathcal{B}$ in a smooth 2-dimensional submanifols, which will be denoted by $\partial\Sigma$. Furthermore, $\Sigma \cup \partial\Sigma$ is assumed to be compact.

From these two assumptions immediately follows that Ω_Σ is well defined, since it can be expressed as an integral of a continuous 3-form over the compact hypersurface $\Sigma \cup \partial\Sigma$.

Now, we turn to the definition of infinitesimal asymptotic symmetries.

Definition 26. Let ξ^a be a complete vector field on $M \cup \mathcal{B}$. ξ^a is called a *representative of an infinitesimal asymptotic symmetry* if its associated one-parameter group of diffeomorphisms maps $\bar{\mathcal{F}}$ into $\bar{\mathcal{F}}$, i.e. if it preserves the asymptotic conditions specified in the definition of $\bar{\mathcal{F}}$. Equivalently, ξ^a is a representative of an infinitesimal asymptotic symmetry if $\mathcal{L}_\xi g$ is tangent to $\bar{\mathcal{F}}$.

One can show (see [38]), that if ξ^a is a representative of an infinitesimal asympotic symmetry, then the right hand side of (7.2.16), namely

$$\Upsilon = \int_{\partial\Sigma} (\delta \boldsymbol{Q}[\xi] - \xi \cdot \boldsymbol{\theta}), \tag{7.3.1}$$

is always well defined and the integral only depends on the cross-section $\partial\Sigma$ of $\mathcal{B}$, not on Σ.

Now, let us introduce the following equivalence relation.

Definition 27. Two representatives of infinitesimal asymptotic symmetries ξ^a and ξ'^a are said to be equivalent if they coincide on $\mathcal{B}$ and if, for all $g \in \bar{\mathcal{F}}$, δg tangent to $\bar{\mathcal{F}}$, and all $\partial\Sigma$ on $\mathcal{B}$, we have $\Upsilon = \Upsilon'$. The *infinitesimal asymptotic symmetries* of the theory are then comprised by the equivalence class of representatives of the infinitesimal asymptotic symmetries.

Consider now an infinitesimal asymptotic symmetry, represented by the vector field ξ^a, and let Σ be a slice with boundary $\partial\Sigma$ on $\mathcal{B}$. Even though the asymptotic conditions, which we stated, assure that the right hand side of (7.2.16) is well defined, there does not, in general, exist a Hamiltonian H_ξ which satisfies this equation. Therefore, we have to consider the following two cases:

(I) Suppose that the continuous extension of $\boldsymbol{\omega}$ to $\mathcal{B}$ has vanishing pullback to $\mathcal{B}$. Then, the condition (7.1.13) implies that H_ξ exists for all infinitesimal asymptotic symmetries and is independent of the choice of representative ξ^a. Furthermore, one can show (see [38]) that in this case H_ξ truely corresponds to a conserved quantity, i.e. its value is independent of "time" Σ.

(II) Suppose that the continuous extension of $\boldsymbol{\omega}$ to $\mathcal{B}$ does not have vanishing pullback to $\mathcal{B}$. Then, in general, there does not exist an H_ξ which satisfies (7.2.16). One exception is the case when ξ^a is everwhere tangent to $\partial\Sigma$, such that the condition (7.1.13) is satisfied. In this case, if ξ^a is tangent to cross-sections $\partial\Sigma_1$ and $\partial\Sigma_2$ of $\mathcal{B}$, which bound a region $\mathcal{B}_{12} \subset \mathcal{B}$, we have

$$\delta H_\xi|_{\partial\Sigma_1} - \delta H_\xi|_{\partial\Sigma_2} = -\int_{\mathcal{B}_{12}} \boldsymbol{\omega}(g, \delta g, \mathcal{L}_\xi g), \tag{7.3.2}$$

where we used (7.2.16) and (7.2.14) in the case we are "on shell". As we see, even though H_ξ exists, it will, in general, not be conserved in this case.

The first case arises in general relativity for spacetimes which are asymptotically flat at spatial infinity. Then, equation (7.2.19) gives rise to the usual expression for the ADM mass (see [32]). The second case arises in general relativity for spacetimes which are asymptotically flat as null infinity.

Now, we will state the definition of a "conserved quantity" conjugate to an infinitesimal asymptotic symmetry ξ^a in case (II). This quantitiy will be denoted by $\mathcal{H}_\xi$ to distinguish it from the Hamiltonian H_ξ.

Remark 29. One should note that the "conserved quantitiy" $\mathcal{H}_n$ will, in general, not be conserved (as in the case of null infinity in general relativity), since symplectic current can be radiated away. This is due to the fact that no Hamiltonian exists which generates the asymptotic symmetry. Therefore, $\mathcal{H}_n$ should be rather interpreted as the energy which is radiated through the boundary $\mathcal{B}$.

On $\mathcal{B}$, let $\boldsymbol{\Theta}$ be the presymplectic potential for the pullback $\bar{\boldsymbol{\omega}}$ of the (extension of the) presymplectic current $\boldsymbol{\omega}$ to $\mathcal{B}$, so that on $\mathcal{B}$ we have

$$\bar{\boldsymbol{\omega}}(g, \delta_1 g, \delta_2 g) = \delta_1 \boldsymbol{\Theta}(g, \delta_2 g) - \delta_2 \boldsymbol{\Theta}(g, \delta_1 g), \tag{7.3.3}$$

for all $g \in \bar{\mathcal{F}}$ and all $\delta_1 g, \delta_2 g$ tangent to $\bar{\mathcal{F}}$. Furthermore $\boldsymbol{\Theta}$ is required to be a local quantity, to depend analytically on the metric when $\boldsymbol{L}$ is analytic, and to be independent of the choices made in the specification of the universal background structure. The quantity $\mathcal{H}_\xi$ is defined by the equation

$$\delta \mathcal{H}_\xi = \int_{\partial\Sigma} (\delta \boldsymbol{Q} - \xi \cdot \boldsymbol{\theta}) + \int_{\partial\Sigma} \xi \cdot \boldsymbol{\Theta}. \tag{7.3.4}$$

Note that the last term in this equation is an ordinary integral over the surface $\partial\Sigma$ of $\mathcal{B}$, whereas the first integral is understood as an asymptotic limit. Equation (7.3.4) satisfies the consistency check (7.1.13), and defines therefore a "conserved quantity" $\mathcal{H}_\xi$ up to an arbitrary constant. This constant can be fixed by requiring $\mathcal{H}_\xi$ to vanish on a reference solution $g_0 \in \bar{\mathcal{F}}$.

However, the above prescription does not define $\mathcal{H}_\xi$ uniquely. Equation (7.3.3) gives rise to the ambiguity

$$\boldsymbol{\Theta}(g, \delta g) \to \boldsymbol{\Theta}(g, \delta g) + \delta \boldsymbol{W}(g), \tag{7.3.5}$$

where $\boldsymbol{W}$ is a suitably (see [38]) constructed 3-form on $\mathcal{B}$. Therefore, an additional condition must be imposed, which selects a $\boldsymbol{\Theta}$ uniquely. We have seen above that the "conserved quantity" $\mathcal{H}_\xi$ will in general be not conserved, due to the possible presence of radiation at $\mathcal{B}$. Therefore, there should be a nonzero flux (3-form) $\boldsymbol{F}_\xi$ on $\mathcal{B}$, associated with $\mathcal{H}_\xi$. It is natural to demand that $\boldsymbol{F}_\xi$ vanishes on $\mathcal{B}$ in the case that g is stationary. One can show (see [38]) that this flux can be identified with $\boldsymbol{\Theta}$, i.e. we have

$$\boldsymbol{F}_\xi = \boldsymbol{\Theta}(g, \mathcal{L}_\xi g). \tag{7.3.6}$$

Therefore, if we require $\boldsymbol{\Theta}(g, \delta g)$ and $\delta \boldsymbol{W}(g)$ to vanish for all δg tangent to $\bar{\mathcal{F}}$ whenever $g \in \bar{\mathcal{F}}$ is stationary, a physically reasonable subset of admissible $\boldsymbol{\Theta}$'s is selected. Thus, if we write down an arbitrary $\boldsymbol{\Theta}$ of this subset and we cannot add a term $\delta \boldsymbol{W}$, such that this condition is preserved, a unique $\boldsymbol{\Theta}$ is selected.

Thus, if a unique $\boldsymbol{\Theta}$ is selected by the above condition and $\mathcal{H}_\xi$ is required to vanish on a reference solution g_0 (for all cross-sections and all ξ^a), then (7.3.4) determines a $\mathcal{H}_\xi$ uniquely.

However, there remains another difficulty in the specification of $\mathcal{H}_\xi$. The reference solutions g_0 and $\psi_* g_0$, where $\psi : M \cup \mathcal{B} \to M \cup \mathcal{B}$ is any diffeomorphism, cannot be distinguished in any meaningful way. Therefore, if we require $\mathcal{H}_\xi$ to vanish on g_0, we must also require $\mathcal{H}_\xi$ to vanish on $\psi_* g_0$. This overdetermines $\mathcal{H}_\xi$ (so that no solutions exists), unless the following condition is imposed: Let ξ^a, η^a be a representatives of infinitesimal asymptotic symmetries and consider a field variation $\delta g = \mathcal{L}_\eta g$ about g_0. Under this field variation we must have $\delta \mathcal{H}_\xi = 0$. Furthermore we have

$$\delta \boldsymbol{Q}[\xi] = \mathcal{L}_\eta \boldsymbol{Q}[\xi] - \boldsymbol{Q}[\mathcal{L}_\eta \xi]. \tag{7.3.7}$$

Since $\delta\mathcal{H}_\xi$ is determined by (7.3.4) and since $\boldsymbol{\Theta}$ is required to vanish at g_0, we obtain the following consistency requirement on g_0: For all representatives ξ^a, η^a of infinitesimal asymptotic symmetries and for all cross-sections $\partial\Sigma$ we must have

$$0 = \int_{\partial\Sigma} \{\mathcal{L}_\eta \boldsymbol{Q}[\xi] - \boldsymbol{Q}[\mathcal{L}_\eta \xi] - \xi \cdot \boldsymbol{\theta}(g_0, \mathcal{L}_\eta g_0)\}. \tag{7.3.8}$$

This is a nontrivial condition that must be satisfied by the reference solution g_0, such that $\mathcal{H}_\xi$ is uniquely defined. One can show (see [38]) that this condition is independent of the cross-section, and so, it must only be checked for one cross-section.

8. On a Second Law of Black Hole Mechanics in our HDTG

In this section we will present the main results of this thesis. We outline two ideas for a proof of a second law of black hole mechanics in our HDTG. Both approaches were not successful in establishing such a theorem.

The first idea we present is a “brute force” technique which adapts the essential idea from the proof of the area theorem in general relativity. By adopting Gaussian null coordinates as a local coordinate system in a neighborhood of the horizon, we will try find an evolution equation which implies that the rate of change of the black hole entropy in our HDTG is positive along the integral curves of the vector field n^a.

The second idea uses more sophisticated methods. We use the covariant phase space formalism from section 7.3 and apply it to the event horizon of a black hole. From this we obtain a quantity which corresponds, in analogy with the Einstein case (see section 8.2.1), to the rate of the change of the black hole entropy. However, the positivity of of this quantity is not investigated.

8.1. First Idea for a Proof

In section 4.3, we have seen that the crucial step in the proof of the area theorem was to show that the expansion ϑ is positive. This implied that we have

$$\partial_u \mathcal{A}(\mathcal{E}(u)) = \int_{\mathcal{E}(u)} \partial_u \sqrt{\mu}\, \mathrm{d}^2x = \int_{\mathcal{E}(u)} \vartheta \sqrt{\mu}\, \mathrm{d}^2x \geq 0, \tag{8.1.1}$$

from which $\mathcal{A}(\mathcal{E}_2) \geq \mathcal{A}(\mathcal{E}_1)$ followed. The positivity of ϑ was shown in the following way: The key ingredient was the Raychaudhuri equation, which may be written[1] in symbolic notation as

$$\partial_u \vartheta = \partial_u (\sqrt{\mu}^{-1} \partial_u \sqrt{\mu}) = \{\text{something negative}\} - \vartheta^2. \tag{8.1.2}$$

That the terms in the curly brackets are really negative made use of Einstein's equation, the null energy condition and the cosmic censorship conjecture. As we showed in appendix D.4, this equation corresponds to the uu-component of the field equations if they are written in GNC and restricted to the horizon. The Raychaudhuri equation implied that, if we have $\vartheta_0 < 0$ initially at any point of the congruence, we must have $\vartheta \to -\infty$ within finite affine length. After that, it was shown $\vartheta \to -\infty$ is equivalent to the existence of a conjugate point. However, since it is impossible that conjugate points exists on causal boundaries, the positivity of ϑ was established.

The idea for a proof of a second law in our HDTG is analogous to the idea of the proof of the area theorem, which we outlined above. In section 7.2.2 we found that the entropy in our

[1] in the case of a hypersurface orthogonal congrunce

HDTG is given by[2]

$$S_\lambda = \int_{\mathcal{E}} \left(\frac{1}{4} - \lambda R_{ur}\right)\sqrt{\mu}\,\mathrm{d}^2x, \tag{8.1.3}$$

where $\mathcal{E}$ is any cross-section of the event horizon. If we could show that we have

$$\partial_u S_\lambda = \int_{\mathcal{E}(u)} \partial_u \left(\frac{1}{4}\sqrt{\mu} - \lambda R_{ur}\sqrt{\mu}\right)\mathrm{d}^2x \geq 0, \tag{8.1.4}$$

we would have established a second law in our HDTG. In order to do so, we must find an evolution equation for the integrant

$$\mathcal{I} := \partial_u \left(\frac{1}{4}\sqrt{\mu} - \lambda R_{ur}\sqrt{\mu}\right), \tag{8.1.5}$$

which implies that is positive and monotonically increasing. In order to find such an equation, we will apply a strategy which is analogous to the Einstein case: We make the Ansatz (see section 8.1.2)

$$\partial_u(\sqrt{\mu}^{-1}\mathcal{I}) = \frac{1}{4}\partial_u\vartheta - \lambda\left(\partial_u^2 R_{ur} + (\partial_u\vartheta)R_{ur} + \vartheta\partial_u R_{ur}\right) \tag{8.1.6}$$

and we will substitute the uu-component of the field equations in our theory, when restricted to the horizon, via the $\partial_u^2 R_{ur}$-term into (8.1.6). Then we will try to bring the resulting equation in a form which implies the desired behaviour of the integrant. With this Ansatz, such a form for the evolution equation would be something like

$$\partial_u(\sqrt{\mu}^{-1}\mathcal{I}) = \sqrt{\mu}^{-1}\mathcal{I} + C, \tag{8.1.7}$$

where C is a positive constant, since the evolution equation must be linear for dimensional reasons (no λ^2-terms).

8.1.1. uu-Component of the Field Equations

In section 6.2 we have derived the field equations

$$E_{ab} = R_{ab} - \frac{1}{2}g_{ab}R + \lambda\left[-\nabla_a\nabla_b R + \Box R_{ab} + 2R^{cd}R_{acbd} - \frac{1}{2}g_{ab}(R^{cd}R_{cd} - \Box R)\right] = 0. \tag{8.1.8}$$

for our HDTG. Since the uu-component of the metric vanishes on the horizon (see appendix D), we have

$$E_{uu}|_{r=0} = R_{uu} + \lambda\left[-(\nabla_a\nabla_b R)_{(uu)} + (\Box R_{ab})_{(uu)} + 2R^{cd}R_{ucud}\right] = 0. \tag{8.1.9}$$

[2]Throughout this section we will only consider stationary perturbations.

A long and tedious calculation[3] shows that the terms in the curly brackets may be written as

$$
\begin{aligned}
(\nabla_a \nabla_b R)_{(uu)} &= \partial_u^2 \Big[2R_{ur} + \mu^{AB} R_{AB} \Big] \\[2ex]
(\Box R_{ab})_{(uu)} &= 2\partial_u \partial_r R_{uu} + \mu^{AB} \hat{D}_A \hat{D}_B R_{uu} - 2\Gamma^A_{ru} \hat{D}_A R_{uu} - 4\Gamma^A_{ru} \partial_u R_{uA} \\
&\quad - \mu^{AB} \Big[\Gamma^u_{AB} \partial_u R_{uu} + \Gamma^r_{AB} \partial_r R_{uu} + 4(\Gamma^C_{Au} \hat{D}_B R_{uC} + \Gamma^u_{Au} \hat{D}_B R_{uu}) \Big] \\
&\quad + 2R_{uu} \Big[3\Gamma^A_{ru} \Gamma^u_{Au} - \partial_r \Gamma^u_{uu} + \mu^{AB} (2\Gamma^u_{Au} \Gamma^u_{Bu} + \Gamma^C_{Au} \Gamma^u_{BC} - \hat{D}_B \Gamma^u_{Au}) \Big] \\
&\quad + 2R_{ur} \Big[-\partial_r \Gamma^r_{uu} + \mu^{AB} \Gamma^C_{Au} \Gamma^r_{BC} \Big] \\
&\quad + R_{uD} \Big[3\Gamma^A_{ru} \Gamma^D_{Au} - \partial_r \Gamma^D_{uu} - \partial_u \Gamma^D_{ru} + \mu^{AB} (\Gamma^r_{AB} \Gamma^D_{ru} + 3\Gamma^u_{Au} \Gamma^D_{Bu} - \hat{D}_B \Gamma^D_{Au}) \Big] \\
&\quad + 2R_{CD} \mu^{AB} \Gamma^C_{Au} \Gamma^D_{Bu} \\[2ex]
R^{cd} R_{ucud} &= R_{uu} \Big[\partial_r \Gamma^u_{uu} - \Gamma^A_{ru} \Gamma^u_{Au} \Big] \\
&\quad + 2R_{uA} \Big[\partial_r \Gamma^A_{uu} - \partial_u \Gamma^A_{ru} - \Gamma^B_{ru} \Gamma^A_{Bu} \Big] \\
&\quad - \mu^{AB} R_{BC} \Big[\partial_u \Gamma^C_{Au} + \Gamma^D_{Au} \Gamma^C_{Du} \Big].
\end{aligned}
\tag{8.1.10}
$$

[3] At first, we tried to implement this calculation with the computer algebra package GRTensor II. However, this attempt did not proove to be fruitful, since the computer program did not "know" how to collect the terms in a meaningful way. Therefore, the calculation was performed by hand, even though it involved about 150 handwritten pages.

in Gaussian null coordinates. By substituting this result into (8.1.9) and by making use of the explicit form of the Christoffel symbols from appendix (D.1), we obtain

$$\begin{aligned}0 = E_{uu}|_{r=0} = R_{uu} + \lambda\Big\{ & -2\partial_u^2 R_{ur} + \partial_u^2(\mu^{AB}R_{AB} + 2\partial_u\partial_r R_{uu}) + \hat{\Box}R_{uu} \\ & - \beta^A \hat{D}_A R_{uu} + 2\beta^A\partial_u R_{uA} + 2(\partial_u\mu^{AB})\hat{D}_A R_{uB} \\ & + \frac{1}{2}\mu^{AB}\Big[(\partial_r\mu_{AB})\partial_u R_{uu} + (\partial_u\mu_{AB})\partial_r R_{uu}\Big] \\ & + R_{uu}\Big[\frac{1}{2}(\partial_u\mu^{AB})\partial_r\mu_{AB} - \hat{D}_A\beta^A\Big] \\ & + \frac{1}{2}R_{ur}(\partial_u\mu^{AB})\partial_r\mu_{AB} \\ & + \frac{1}{2}R_{uA}\Big[\beta^C\mu^{AB}\partial_u\mu_{BC} + \frac{1}{2}\beta^A\mu^{BC}\partial_u\mu_{BC} + 5\partial_u\beta^A - \hat{D}^C(\mu^{AB}\partial_u\mu_{BC})\Big] \\ & - R_{AB}\mu^{BD}\partial_u(\mu^{AC}\partial_u\mu_{CD})\Big\},\end{aligned} \tag{8.1.11}$$

where we introduced the notation $\hat{\Box} := \mu^{AB}\hat{D}_A\hat{D}_B$. Again, this result required extensive calculations which we will not be presented at this point.

8.1.2. Evolution Equation

As we described at the beginning of this section, our Ansatz for the evolution equation is (8.1.6). Explicitly we have

$$\begin{aligned}\partial_u(\sqrt{\mu}^{-1}\mathcal{I}) =& \partial_u\left(\sqrt{\mu}^{-1}\partial_u\left[\frac{1}{4}\sqrt{\mu} - \bar{\lambda}R_{ur}\sqrt{\mu}\right]\right) \\ =& \frac{1}{4}\partial_u(\sqrt{\mu}^{-1}\partial_u\sqrt{\mu}) - \bar{\lambda}\partial_u\left(\sqrt{\mu}^{-1}\partial_u(R_{ur}\sqrt{\mu})\right) \\ =& \frac{1}{4}\partial_u\vartheta - \bar{\lambda}\left(\partial_u^2 R_{ur} + \partial_u(\sqrt{\mu}^{-1}R_{ur}\partial_u\sqrt{\mu})\right) \\ =& \frac{1}{4}\partial_u\vartheta - \bar{\lambda}\left(\partial_u^2 R_{ur} + \partial_u(\sqrt{\mu}^{-1}\partial_u\sqrt{\mu})R_{ur} + (\sqrt{\mu}^{-1}\partial_u\sqrt{\mu})\partial_u R_{ur}\right) \\ =& \frac{1}{4}\partial_u\vartheta - \bar{\lambda}\left(\partial_u^2 R_{ur} + (\partial_u\vartheta)R_{ur} + \vartheta\partial_u R_{ur}\right).\end{aligned} \tag{8.1.12}$$

If we solve the restricted uu-component of the field equations (8.1.11) for the $\partial_u^2 R_{ur}$-term, we obtain

$$\begin{aligned}
\partial_u^2 R_{ur} = & \frac{1}{2\bar{\lambda}} R_{uu} + \frac{1}{2}\partial_u^2(\mu^{AB}R_{AB} + \partial_u\partial_r R_{uu}) + \frac{1}{2}\hat{\Box} R_{uu} \\
& - \frac{1}{2}\beta^A \hat{D}_A R_{uu} + \beta^A \partial_u R_{uA} + (\partial_u \mu^{AB})\hat{D}_A R_{uB} \\
& + \frac{1}{4}\mu^{AB}\Big[(\partial_r \mu_{AB})\partial_u R_{uu} + (\partial_u \mu_{AB})\partial_r R_{uu}\Big] \\
& + \frac{1}{2} R_{uu}\Big[\frac{1}{2}(\partial_u \mu^{AB})\partial_r \mu_{AB} - \hat{D}_A \beta^A\Big] \\
& + \frac{1}{4} R_{ur}(\partial_u \mu^{AB})\partial_r \mu_{AB} \\
& + \frac{1}{4} R_{uA}\Big[\beta^C \mu^{AB}\partial_u \mu_{BC} + \frac{1}{2}\beta^A \mu^{BC}\partial_u \mu_{BC} + 5\partial_u \beta^A - \hat{D}^C(\mu^{AB}\partial_u \mu_{BC})\Big] \\
& - \frac{1}{2} R_{AB}\mu^{BD}\partial_u(\mu^{AC}\partial_u \mu_{CD}).
\end{aligned} \tag{8.1.13}$$

The next step would be to insert this result into our Ansatz (8.1.12).

At this point of the analysis we stopped to pursue this strategy. We inserted (8.1.13) into (8.1.12) and "played around" with it in order to see if it is possible to bring into the form (8.1.7). But since the number of terms which were involved was so overwhelmingly large, we could not see any structure in the resulting equation, so we decided to stop at this point and to pursue a more systematic approach.

8.2. Second Idea for a Proof

The formalism of Wald and Zoupas, which we summarized in section 7.3, is actually designed to define conserved quantities on the attached boundary $\mathcal{B}$ of the unphysical spacetime. However, if one carefully checks the assumptions which were made, one realizes that this formalism is not limited to spacetime boundaries, but can also be applied to the event horizon of a black hole. The quantity $\mathcal{H}_\xi$ which one obtains by this procedure, is a "conserved quantity" on E in the sense of section 7.3. It will, in general, not be conserved and it should be related to the rate of change of the black hole entropy.

We make the following modifications to the Wald-Zoupas formalism, in order to apply it to the event horizon of a black hole:

1. The event horizon E will play the role of the attached boundary $\mathcal{B}$. Since E is situated in a "finite region" of spacetime, one does not need to worry about "asymptotic conditions" on the metric g_{ab} as one approaches E.

2. The vector field ξ^a which generates the asymptotic symmetry will, in the case of a black hole, be the vector field $n^a = (\partial/\partial u)^a$. Furthermore, we assume that n^a is proportional to a Killing vector field.[4]

[4]This is a nontrivial assumption. One would have to prove a rigidity theorem for this gravitational theory, in order to justify this. However, this assumption seemed undispensable for the calculations which will follow.

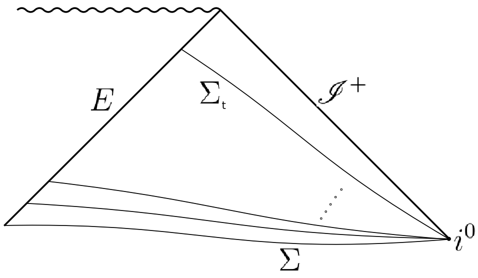

Figure 8.1.: Foliation of a black hole spacetime.

3. The universal background structure of the theory will be given by the requirement that a certain region in spacetime remains fixed under variations $g \to g+\delta g$. This region will be a neighborhood U_E of the event horizon E. What we mean by "fixed" is the requirement that if we have two neighborhoods U_E, $V_{E'}$, such that $E \subset U_E \subset (M, g)$ and $E' \subset V_{E'} \subset (M, g+\delta g)$, then U_E and $V_{E'}$ are required to be the same as topological manifolds, i.e. they are ought to be homeomorphic. Since we are free to apply diffeomorphisms to the spacetime, this can always be achieved after a variation was performed, i.e. we can find a diffeomorphism $\psi : M \to M$ such that $V_{E'} = \psi(U_E)$. Of course, this requirement selects a subset of admissible variations δg, such that E' can be "bend back" to E. However, this class of admissible variation is large enough for our purposes. What we gain from this requirement is that the coordinates r and u of the Gaussian null coordinate system do not change under variations, i.e. we have

$$\delta r = 0, \qquad \delta u = 0. \tag{8.2.1}$$

   This greatly simplifies the following calculations.

4. When the Wald-Zoupas formalism is applied to spacetime boundaries $\mathcal{B}$, the hypersurfaces Σ, over which one integrates in order to define quantities like the presymplectic form Ω_Σ, are assumed to be slices, i.e. closed embedded 3-dimensional submanifolds without boundary. So slices are not admitted to have a boundary in the physical spacetime.

   In the case of a black hole, the hypersurfaces of interest are those which extend from the event horizon to spatial infinity. Therefore, it seems more practicable to use, instead of a slice, spacelike hypersurfaces Σ_t which give rise to a foliation of the spacetime (see figure 8.1).

In order to see that these modifications to the formalism do indeed yield a meaningful result, we will show in the next section that, in the case of the Einstein theory, the "conserved quantity" $\mathcal{H}_n$ associated with the vector field n^a satisfies

$$\mathcal{H}_n \propto \mathcal{L}_n \mathcal{A}(\mathcal{E}). \tag{8.2.2}$$

So, since we have $S = \mathcal{A}/4$ in the Einstein theory, $\mathcal{H}_n$ is related to the rate of change of the black hole entropy along the integral curves of the vector field n^a.

The strategy for the computation of $\mathcal{H}_n$ on E is the following.

1. Calculate the presymplectic potential $\boldsymbol{\theta}$, the Noether current $\boldsymbol{J}$ and the Noether charge $\boldsymbol{Q}$.

2. Make the particular choice[5]
$$\boldsymbol{\Theta} = \iota^*\boldsymbol{\theta} + \delta\boldsymbol{W}, \tag{8.2.3}$$
where $\iota : E \to M$ is an embedding. Since we know that $\boldsymbol{\Theta}$ is required to vanish for stationary solutions, we can determine $\boldsymbol{W}$ by finding a decomposition
$$\iota^*\boldsymbol{\theta} = \{\text{part that vanishes for stationary metrics}\} - \delta\boldsymbol{W}. \tag{8.2.4}$$

3. Then, $\mathcal{H}_n$ satifies
$$\delta\mathcal{H}_n = \int_{\mathcal{E}}(\delta\boldsymbol{Q} - n\cdot\boldsymbol{\theta}) + \int_{\mathcal{E}} n\cdot\boldsymbol{\Theta} = \delta\int_{\mathcal{E}}(\boldsymbol{Q} + n\cdot\boldsymbol{W}), \tag{8.2.5}$$
i.e. we have
$$\mathcal{H}_n = \int_{\mathcal{E}}(\boldsymbol{Q} + n\cdot\boldsymbol{W}) + C. \tag{8.2.6}$$
The constant C can be fixed by requiring that $\mathcal{H}_n$ vanishes on a reference solution g_0, such as the Schwarzschild spacetime.

8.2.1. Calculation of $\mathcal{H}_n$ in Einstein Gravity

In this section we will apply that the strategy we outlined above to a black hole solution of Einstein gravity.

Our first task is to find a decomposition of the form (8.2.4). The presymplectic potential for Einstein gravity is given by (see the result of section 8.2.2 and set $\lambda = 0$)
$$\theta_{abc} = \frac{1}{16\pi}\epsilon_{dabc}(g^{fh}g^{de} - g^{he}g^{df})\nabla_f\delta g_{he}. \tag{8.2.7}$$
Since the pullback $\iota^*\theta_{abc}$ of θ_{abc} is only allowed to act on vectors which are tangent to E, the index d is fixed to the value $d = r$ due to the total antisymmetry of the tensor ϵ_{abcd}. Therefore, we will try to find the desired decomposition for
$$\iota^*\theta_{abc} = \frac{1}{16\pi}\epsilon_{rabc}(g^{fh}g^{re} - g^{he}g^{rf})\nabla_f\delta g_{he} \tag{8.2.8}$$
in the following. In GNC, the variation of the metric (3.2.8) is given by
$$\delta g_{ab} = (\delta\mu_{AB})(\mathrm{d}x^A)_a(\mathrm{d}x^B)_b =: (\delta\mu_{AB})\mathrm{d}x^A_a\mathrm{d}x^B_b =: \delta\mu_{ab}, \tag{8.2.9}$$

[5] However, this seems to be the most natural choice, since we have
$$\begin{aligned}\delta_1\boldsymbol{\Theta}(g,\delta_2 g) - \delta_2\boldsymbol{\Theta}(g,\delta_1 g) &= \delta_1\iota^*\boldsymbol{\theta}(g,\delta_2 g) + \delta_1\delta_2\boldsymbol{W}(g) - \delta_2\iota^*\boldsymbol{\theta}(g,\delta_1 g) - \delta_2\delta_1\boldsymbol{W}(g)\\ &= \iota^*[\delta_1\boldsymbol{\theta}(g,\delta_2 g) - \delta_2\boldsymbol{\theta}(g,\delta_1 g)]\\ &= \bar{\boldsymbol{\omega}}(g,\delta_1 g,\delta_2 g).\end{aligned}$$
Here, we have used the fact that mixed variations (being only partial derivatives) as well as variation and pullback commute.

where we have used $\delta r = 0$ and $r = 0$, since we restricted ourselves to the horizon after the variation was performed. The covariant derivative of this variation may be expressed as

$$\nabla_c \delta\mu_{ab} = (\nabla_c \delta\mu_{AB})\mathrm{d}x_a^A \mathrm{d}x_b^B + \delta\mu_{AB}(\nabla_c \mathrm{d}x_a^A)\mathrm{d}x_b^B + \delta\mu_{AB}\mathrm{d}x_a^A(\nabla_c \mathrm{d}x_b^B). \tag{8.2.10}$$

We have

$$\nabla_c \delta\mu_{AB} = \partial_c \delta\mu_{AB} - \Gamma^C_{cA}\delta\mu_{CB} - \Gamma^C_{cB}\delta\mu_{AC} \tag{8.2.11}$$

and

$$\begin{aligned} \nabla_c \mathrm{d}x_a^A &= -\sum_{\mu,\nu} \Gamma^A_{\mu\nu}\mathrm{d}x_a^\mu \mathrm{d}x_b^\nu \\ &= -\Big[\Gamma^A_{ur}\mathrm{d}u_a \mathrm{d}r_b + \Gamma^A_{uC}\mathrm{d}u_a\mathrm{d}x_b^C + \Gamma^A_{ru}\mathrm{d}r_a\mathrm{d}u_b + \Gamma^A_{rC}\mathrm{d}r_a\mathrm{d}x_b^C \\ &\quad + \Gamma^A_{Cu}\mathrm{d}x_a^C\mathrm{d}u_b + \Gamma^A_{Cr}\mathrm{d}x_a^C\mathrm{d}r_b + \Gamma^A_{CD}\mathrm{d}x_a^C\mathrm{d}x_b^D\Big], \end{aligned} \tag{8.2.12}$$

where we have used $\Gamma^A_{uu} = \Gamma^A_{rr} = 0$. Insertion of (8.2.10) together with (8.2.11),(8.2.12) into (8.2.9) yields

$$\iota^*\theta_{abc} = \frac{1}{16\pi}\epsilon_{rabc}\Big\{-\partial_u(\mu^{AB}\delta\mu_{AB})\Big\}. \tag{8.2.13}$$

This result required a longer, however straightforward calculation which would not be very enlightening at this point. At numerous occasions of the calculation we have used the form of the metric (D.0.1) and the fact that we have $g^{rr} = g^{rA} = 0$ on E, i.e. at $r = 0$. Furthermore, we have have used the explicit form of the Christoffel symbols (see D.1). Consider now the following expression

$$\begin{aligned} \delta\partial_u\sqrt{\mu} &= \delta(\sqrt{\mu}\sqrt{\mu}^{-1}\partial_u\sqrt{\mu}) \\ &= \frac{1}{2}\delta(\sqrt{\mu}\mu^{AB}\partial_u\mu_{AB}) \\ &= \frac{1}{2}(\delta\sqrt{\mu})\mu^{AB}\partial_u\mu_{AB} + \frac{1}{2}\sqrt{\mu}\Big[(\delta\mu^{AB})\partial_u\mu_{AB} + \mu^{AB}\delta\partial_u\mu_{AB}\Big] \\ &= \frac{1}{2}\Big[\frac{1}{2}\sqrt{\mu}\mu^{CD}\delta\mu_{CD}\Big]\mu^{AB}\partial_u\mu_{AB} + \frac{1}{2}\sqrt{\mu}\partial_u(\mu^{AB}\delta\mu_{AB}). \end{aligned} \tag{8.2.14}$$

Here we have used equation (D.1.20) at variaous stages, as well as the fact that δ and ∂_u commute, and the identity $(\delta\mu^{AB})\partial_u\mu_{AB} = (\delta\mu_{AB})\partial_u\mu^{AB}$. From this we obtain

$$\partial_u(\mu^{AB}\delta\mu_{AB}) = \frac{2}{\sqrt{\mu}}\delta\partial_u\sqrt{\mu} - \frac{1}{2}\mu^{AB}\mu^{CD}(\partial_u\mu_{AB})\delta\mu_{CD}. \tag{8.2.15}$$

Insertion of (8.2.15) into (8.2.13) yields

$$\begin{aligned} \iota^*\theta_{abc} &= \frac{1}{16\pi}\epsilon_{rabc}\Big\{\frac{1}{2}\mu^{AB}\mu^{CD}(\partial_u\mu_{AB})\delta\mu_{CD} - \frac{2}{\sqrt{\mu}}\delta\partial_u\sqrt{\mu}\Big\} \\ &= \frac{1}{16\pi}\epsilon_{rabc}\Big\{\frac{1}{2}\mu^{AB}\mu^{CD}(\partial_u\mu_{AB})\delta\mu_{CD} + 2(\partial_u\sqrt{\mu})\delta\sqrt{\mu}^{-1} - 2\delta(\sqrt{\mu}^{-1}\partial_u\sqrt{\mu})\Big\} \end{aligned} \tag{8.2.16}$$

The quantity $\epsilon_{rabc} = l^d \epsilon_{dabc}$ corresponds to the volumeform that can only act on vectors which are orthogonal to l^a, i.e. we have

$$\epsilon_{rabc} = (\mathrm{d}u)_a \wedge \overset{(2)}{\epsilon}_{bc} = 6\sqrt{\mu}(\mathrm{d}u)_{[a}(\mathrm{d}x^1)_b(\mathrm{d}x^2)_{c]} =: \overset{(3)}{\epsilon}_{abc}, \tag{8.2.17}$$

where $\overset{(2)}{\epsilon}_{ab}$ is the volumeform on a cross-section $\mathcal{E}$. Having this in mind, (8.2.16) can be written as

$$\begin{aligned}\iota^*\theta_{abc} = &\frac{1}{16\pi}\overset{(3)}{\epsilon}_{abc}\left\{\frac{1}{2}\mu^{AB}\mu^{CD}(\partial_u\mu_{AB})\delta\mu_{CD} + 2(\partial_u\sqrt{\mu})\delta\sqrt{\mu}^{-1} + (\sqrt{\mu}^{-1}\partial_u\sqrt{\mu})\mu^{AB}\delta\mu_{AB}\right\}\\ &- \frac{1}{8\pi}\delta\left[(\sqrt{\mu}^{-1}\partial_u\sqrt{\mu})\overset{(3)}{\epsilon}_{abc}\right].\end{aligned} \tag{8.2.18}$$

The additional $(\sqrt{\mu}^{-1}\partial_u\sqrt{\mu})\mu^{AB}\delta\mu_{AB}$ terms appears in the curly brackets in order to compensate the variation of $\overset{(3)}{\epsilon}_{abc}$, since we have

$$\begin{aligned}\delta(\overset{(3)}{\epsilon}_{abc}) =&(\delta\sqrt{\mu})(\mathrm{d}u)_a \wedge (\mathrm{d}x^1)_b \wedge (\mathrm{d}x^2)_c\\ =&\frac{1}{2}\sqrt{\mu}\mu^{AB}(\delta\mu_{AB})(\mathrm{d}u)_a \wedge (\mathrm{d}x^1)_b \wedge (\mathrm{d}x^2)_c\\ =&\frac{1}{2}\mu^{AB}(\delta\mu_{AB})\overset{(3)}{\epsilon}_{abc}\end{aligned} \tag{8.2.19}$$

Equation (8.2.18) is the decomposition which we tried to find. The terms in the curly brackets clearly vanish in the stationary case, since we have $\partial_u\mu_{AB} = 0$, and therefore $\partial_u\sqrt{\mu}$, for stationary spacetimes. From this we find

$$W_{abc} = \frac{1}{8\pi}\vartheta\overset{(3)}{\epsilon}_{abc}, \tag{8.2.20}$$

where $\vartheta = \sqrt{\mu}^{-1}\partial_u\sqrt{\mu}$ is the expansion of the null geodesic generators of the horizon.

Now, we come to the calculation of the "conserved quantity" $\mathcal{H}_n$. First of all, we will show that we have

$$\int_{\mathcal{E}} \boldsymbol{Q} = 0. \tag{8.2.21}$$

Since the 2-form $\boldsymbol{Q}$, which is defined on the entire manifold, appears under an integral sign which is evaluated on the cross-section $\mathcal{E}$, we have have to consider the pullback of $\boldsymbol{Q}$ to $\mathcal{E}$, i.e. we will show

$$\int_{\mathcal{E}} \psi^*\boldsymbol{Q} = 0 \tag{8.2.22}$$

where $\psi : \mathcal{E} \to M$ is an embedding. For Einstein gravity, the Noether charge is given by

$$Q_{ab} = -\frac{1}{16\pi}\epsilon_{abcd}\nabla^c n^d. \tag{8.2.23}$$

We have

$$\psi^*\boldsymbol{Q} = (\psi^*Q)_{ab} = -\frac{1}{16\pi}\psi^*\epsilon_{abcd}\nabla^c n^d = -\frac{1}{16\pi}\psi^*\epsilon_{abcd}g^{ce}g^{df}\nabla_e n_f. \tag{8.2.24}$$

The action of the pullback has the effect that the 2-form Q_{ab} can only act on vectors tangent to $\mathcal{E}$. From this, and the fact that ϵ_{abcd} is totally antisymmetric, follows that the indices c, d are fixed to the values u, r, i.e. we have

$$(\psi^* Q)_{ab} = -\frac{1}{16\pi} \psi^* \epsilon_{abur} g^{ue} g^{rf} \nabla_e n_f. \tag{8.2.25}$$

Consider now

$$n_a = g_{ab} n^b = c(x^A) \Big[(\mathrm{d}r)_a - r^2 \alpha (\mathrm{d}u)_b - r\beta_A (\mathrm{d}x^A)_a \Big]. \tag{8.2.26}$$

Since ϵ_{abcd} is totally antisymmetric, the expression $\nabla_a n_b$ is (8.2.25) antisymmetrised, i.e. we have

$$\epsilon_{abcd} g^{ce} g^{df} \nabla_e n_f = -\epsilon_{abcd} g^{ce} g^{df} \nabla_f n_e. \tag{8.2.27}$$

Consider therefore only the antisymmetric part of $\nabla_a n_b$:

$$\begin{aligned} \nabla_{[a} n_{b]} = (\nabla_{[a} c) \Big[(\mathrm{d}r)_{b]} - r^2 \alpha (\mathrm{d}u)_{b]} - r\beta_A (\mathrm{d}x^A)_{b]} \Big] \\ + c \Big[-\nabla_{[e} (r^2 \alpha)(\mathrm{d}u)_{b]} - \nabla_{[e} (r\beta_A)(\mathrm{d}x^A)_{b]} \Big]. \end{aligned} \tag{8.2.28}$$

From this follows

$$g^{ue} g^{rf} \nabla_e n_f = g^{ue} g^{rf} \nabla_{[e} n_{f]} = rc(\beta^2 - 2\alpha). \tag{8.2.29}$$

Therefore, we have

$$\int_{\mathcal{E}} (\psi^* Q)_{ab} = \int_{\mathcal{E}} (\psi^* Q)_{ab} = -\frac{1}{16\pi} \int_{\mathcal{E}} rc(\beta^2 - 2\alpha) \overset{(2)}{\epsilon}_{ab} = -\frac{1}{16\pi} \int_{\mathcal{E}} f(r, x^A) \overset{(2)}{\epsilon}_{ab}. \tag{8.2.30}$$

Since the function $f(r, x^A)$ vanishes on $\mathcal{E}$ ($\mathcal{E}$ is defined by $r = 0$), it follows $\int_{\partial\Sigma} Q = 0$.

As $\boldsymbol{Q}$ does not contribute to $\mathcal{H}_n$, it is now an easy task the "conserved quantity" $\mathcal{H}_n$. We find

$$\mathcal{H}_n = \int_{\mathcal{E}} n^a W_{abc} = \frac{1}{8\pi} \int_{\mathcal{E}} \vartheta n^a \overset{(3)}{\epsilon}_{abc} = \frac{1}{8\pi} \int_{\mathcal{E}} \vartheta \overset{(2)}{\epsilon}_{bc} = \frac{1}{8\pi} \partial_u \int_{\mathcal{E}} \overset{(2)}{\epsilon}_{bc} = \frac{1}{8\pi} \partial_u \mathcal{A}(\mathcal{E}) = \frac{1}{8\pi} \mathcal{L}_n \mathcal{A}(\mathcal{E}), \tag{8.2.31}$$

where we used the fact that the Lie derivative $\mathcal{L}_n$ can be locally written as ∂_u in an adapted coordinate system. As we see, $\mathcal{H}_n$ is equal to a mupltiple of $\mathcal{L}_n \mathcal{A}(\mathcal{E})$, which corresponds to the rate of change of the cross-section area along the flow lines of the vector field n^a.

8.2.2. Calculation of θ in our HDTG

The presymplectic potential $\boldsymbol{\theta}$ was determined by the equation

$$\delta \boldsymbol{L} = \boldsymbol{E} \cdot \delta g + \mathrm{d}\boldsymbol{\theta}. \tag{8.2.32}$$

So all we have to do in order to determine $\boldsymbol{\theta}$ is to pick up all the total divergence terms which we dropped out in the derivation of the field equations (see section 6.2).

The total divergence terms for the Einstein-Hilbert action arise from the variation of the

Ricci-tensor, which is given by the standard identity

$$\delta R_{ab} = \nabla_c(\delta\Gamma^c_{ab}) - \nabla_b(\delta\Gamma^c_{ac}). \tag{8.2.33}$$

If we introduce the vector field

$$w^a = g^{cd}\delta\Gamma^a_{cd} - g^{ca}\delta\Gamma^d_{cd}, \tag{8.2.34}$$

we may write

$$\int_M g^{ab}\delta R_{ab}\boldsymbol{\epsilon} = \int_M (\nabla_a m^a)\boldsymbol{\epsilon}, \tag{8.2.35}$$

since ∇_a is compatible with the metric. This is the term that drops out due to the asymptotic conditions on the metric in the derivation of the Einstein field equations by using Stokes theorem. The vector field m^a gives the first contribution to the presymplectic potential. We have

$$\begin{aligned} m^a &= g^{cd}\left[\frac{1}{2}g^{ae}(\nabla_c\delta g_{de} + \nabla_d\delta g_{ce} - \nabla_e\delta g_{cd})\right] - g^{ca}\left[\frac{1}{2}g^{de}(\nabla_c\delta g_{de} + \nabla_d\delta g_{ce} - \nabla_e\delta g_{cd})\right] \\ &= \frac{1}{2}\left[g^{cd}g^{ae}\nabla_c\delta g_{de} + g^{cd}g^{ae}\nabla_d\delta g_{ce} - g^{cd}g^{ae}\nabla_e\delta g_{cd} - g^{ca}g^{de}\nabla_c\delta g_{de}\right] \\ &= g^{cd}g^{ae}\nabla_c\delta g_{de} - g^{cd}g^{ae}\nabla_e\delta g_{cd} \\ &= g^{cd}g^{ae}\left[\nabla_c\delta g_{de} - \nabla_e\delta g_{cd}\right]. \end{aligned} \tag{8.2.36}$$

The other total divergence terms in the derivation of the field equations arise from the term (6.2.8), namely

$$2\int_M R^{ab}\delta R_{ab}\boldsymbol{\epsilon} = \int_M g^{cd}\left[2(\nabla_c\nabla_a\delta g_{bd})R^{ab} - (\nabla_c\nabla_d\delta g_{ab})R^{ab} - (\nabla_b\nabla_a\delta g_{cd})R^{ab}\right]\boldsymbol{\epsilon}. \tag{8.2.37}$$

The first term in (8.2.37) may be rewritten as

$$\begin{aligned} \int_M g^{cd}(\nabla_c\nabla_a\delta g_{bd})R^{ab}\boldsymbol{\epsilon} &= \int_M \nabla_c \underbrace{\left[g^{cd}(\nabla_a\delta g_{bd})R^{ab}\right]}_{=:u^c}\boldsymbol{\epsilon} - \int_M g^{cd}(\nabla_a\delta g_{bd})\nabla_c R^{ab}\boldsymbol{\epsilon} \\ &= \int_M \nabla_c u^c\boldsymbol{\epsilon} - \int_M \nabla_a \underbrace{\left[g^{cd}(\delta g_{bd})\nabla_c R^{ab}\right]}_{=:-v^a}\boldsymbol{\epsilon} + \int_M g^{cd}(\delta g_{bd})\nabla_a\nabla_c R^{ab}\boldsymbol{\epsilon}. \end{aligned} \tag{8.2.38}$$

So, the vector fields u^a and v^a give another contribution to $\boldsymbol{\theta}$. By proceeding in a similar manner with the remaining two terms in (8.2.37), we obtain the following further contributions:

$$w^a = g^{ab}(\nabla_b \delta g_{cd}) R^{cd} \tag{8.2.39}$$
$$x^a = -g^{ab}(\delta g_{cd}) \nabla_b R^{cd} \tag{8.2.40}$$
$$y^a = g^{cd}(\nabla_b \delta g_{cd}) R^{ab} \tag{8.2.41}$$
$$z^a = -g^{cd}(\delta g_{cd}) \nabla_b R^{ab}. \tag{8.2.42}$$

Adding up all the contributions yields

$$r^a := m^a + \lambda(2u^a + 2v^a - w^a - x^a - y^a - z^a). \tag{8.2.43}$$

By performing some index gymnastics, this vector can be rewritten as

$$\begin{aligned} r^a = &(g^{fh} g^{ae} - g^{he} g^{af}) \nabla_f \delta g_{he} \\ &+ \lambda \Big\{ (2 g^{ae} R^{fh} - g^{af} R^{he} - g^{he} R^{af}) \nabla_f \delta g_{he} \\ &\quad + (g^{af} \nabla_f R^{eh} + g^{eh} \nabla_f R^{af} - 2 g^{fh} \nabla_f R^{ae}) \delta g_{eh} \Big\}. \end{aligned} \tag{8.2.44}$$

Now, since the presymplectic potential corresponds to the boundary terms that arises in the derivation of the field equations, $\boldsymbol{\theta} = \theta_{abc} = \epsilon_{dabc} r^d$ is given by

$$\begin{aligned} \theta_{abc} = &\frac{1}{16\pi} \epsilon_{dabc} \Big[(g^{fh} g^{de} - g^{he} g^{df}) \nabla_f \delta g_{he} \\ &+ \lambda \Big\{ (2 g^{de} R^{fh} - g^{df} R^{he} - g^{he} R^{df}) \nabla_f \delta g_{he} \\ &\quad + (g^{df} \nabla_f R^{eh} + g^{eh} \nabla_f R^{df} - 2 g^{fh} \nabla_f R^{de}) \delta g_{eh} \Big\} \Big], \end{aligned} \tag{8.2.45}$$

where the prefactor $1/16\pi$ from (8.2.32) was taken into account.

8.2.3. Calculation of J and Q in our HDTG

In order to calculate the Noether current $\boldsymbol{J} = J_{abc}$ we will use the formula

$$J_{abc} = \theta_{abc}(g, \mathcal{L}_n g) - n^d L_{dabc}. \tag{8.2.46}$$

By substituting $\mathcal{L}_n g_{ab} = \nabla_a n_b + \nabla_a n_b$ into (8.2.45) we obtain

$$\begin{aligned} \theta_{abc}(g, \mathcal{L}_n g) = \epsilon_{dabc} \Big[&\Box n^d + \nabla^e \nabla^d n_e - 2 \nabla^d \nabla^e n_e \\ &+ 2\lambda \Big\{ R^{fh} \nabla_f \nabla_h n^d + R^{fh} R^d{}_{fjh} n^j - R^{df} \nabla_f \nabla^h n_h \\ &\quad + (\nabla^d R^{fh}) \nabla_f n_h + (\nabla_f R^{df}) \nabla^h n_h \\ &\quad - (\nabla^f R^{dh}) \nabla_h n_f - (\nabla^f R^{dh}) \nabla_f n_h \Big\} \Big]. \end{aligned} \tag{8.2.47}$$

Furthermore we have

$$n^d L_{dabc} = \frac{1}{16\pi} n^d \epsilon_{dabc}(R + \lambda R_{fh} R^{fh}). \tag{8.2.48}$$

Substitution of (8.2.47) and (8.2.48) into (8.2.46), together with the replacements

$$R^{fh}\nabla_f \nabla_h n^d = \nabla_f(R^{fh}\nabla_h n^d) - \nabla_h[(\nabla_f R^{fh})n^d] + \frac{1}{2} n^d \Box R \tag{8.2.49}$$

$$R^{fh} R^d{}_{fjh} n^j = n^k g^{dj} R^{fh} R_{jfkh} \tag{8.2.50}$$

$$(\nabla_f R^{df})\nabla^h n_h = \nabla^h[(\nabla_f R^{df})n_h] - \frac{1}{2} n^k \nabla^d \nabla_k R \tag{8.2.51}$$

$$(\nabla^f R^{dh})\nabla_f n_h = \nabla_f[(\nabla^f R^{dh})n_h] - n^k g^{dj} \Box R_{jk} \tag{8.2.52}$$

yields

$$J_{abc} = \frac{1}{8\pi}\epsilon_{dabc}\Bigg[\nabla_e \nabla^{[e} n^{d]} + \lambda\Bigg\{\frac{1}{2} n^k \nabla^d \nabla_k R - n^k R^{fh} R^d{}_{fkh} + n^k R^{df} R_{fk} + (\nabla^d R^{fh})\nabla_f n_h$$
$$+ \nabla_f\{R^{fh}\nabla_h n^d\} - \nabla_h\{(\nabla_f R^{fh})n^d\} + \nabla^h\{(\nabla_f R^{df})n_h\} - \nabla_f\{(\nabla^f R^{dh})n_h\}\Bigg\}\Bigg], \tag{8.2.53}$$

where we have used the field equations

$$E_{ab} = R_{ab} - \frac{1}{2} g_{ab} R + \lambda\left[-\nabla_a \nabla_b R + \Box R_{ab} + 2R^{cd} R_{acbd} - \frac{1}{2} g_{ab}(R^{cd} R_{cd} - \Box R)\right] = 0 \tag{8.2.54}$$

to further simplify the resulting expression. Equation (8.2.53) can be further simplified to

$$J_{abc} = \frac{1}{8\pi}\epsilon_{dabc}\left[\nabla_e \nabla^{[e} n^{d]} + \lambda \nabla_e X^{ed}\right], \tag{8.2.55}$$

with

$$X^{ab} = R^{ac}\nabla_c n^b - R^{bc}\nabla_c n^a + (\nabla^b R^{ac})n_c - (\nabla^a R^{bc})n_c + (\nabla_c R^{cb})n^a - (\nabla_c R^{ca})n^b. \tag{8.2.56}$$

One should notice that X^{ab} is antisymmetric.

The Noether charge $\boldsymbol{Q} = Q_{ab}$ is given by

$$Q_{ab} = -\frac{1}{16\pi}\epsilon_{abcd}(\nabla^c n^d + \lambda X^{cd}). \tag{8.2.57}$$

One can check that we have

$$(\mathrm{d}Q)_{abc} = 3\nabla_{[a} Q_{bc]} = J_{abc}, \tag{8.2.58}$$

by using the identity

$$-2\nabla_b(\epsilon_{ca_1a_2a_3} T^{[bc]}) = 3\nabla_{[a_1}\epsilon_{a_2a_3]bc} T^{bc} \tag{8.2.59}$$

which holds for arbitrary type (2,0)-tensors T^{ab}.

8.2.4. Calculation of $\mathcal{H}_n$ in our HDTG

In this section, we will use the same strategy as in section 8.2.1 to calculate the "conserved quantity" $\mathcal{H}_n$. Therefore, our first task is to find a decomposition of $\iota^*\theta_{abc}$ in a part that vanishes in the stationary case, and a total variation. In section 8.2.2 we found

$$\iota^*\theta_{abc} = \frac{1}{16\pi}\iota^*\epsilon_{dabc}\Big\{Z_1^d + \lambda(Z_2^d + Z_3^d)\Big\} = \frac{1}{16\pi}\overset{(3)}{\epsilon}_{abc}\Big\{Z_1^d + \lambda(Z_2^d + Z_3^d)\Big\}(\mathrm{d}r)_d, \tag{8.2.60}$$

with

$$Z_1^d = (g^{de}g^{fh} - g^{df}g^{eh})\nabla_f\delta\mu_{eh} \tag{8.2.61}$$
$$Z_2^d = (2g^{de}R^{fh} - g^{df}R^{eh} - g^{eh}R^{df})\nabla_f\delta\mu_{eh} \tag{8.2.62}$$
$$Z_3^d = (g^{df}\nabla_f R^{eh} + g^{eh}\nabla_f R^{df} - 2g^{fh}\nabla_f R^{de})\delta\mu_{eh}. \tag{8.2.63}$$

For the term $Z_1^d(\mathrm{d}r)_d$ we already found the desired decomposition in section 8.2.1, so we do not need to worry about this contribution anymore. For the second part we find

$$\begin{aligned} Z_2^d(\mathrm{d}r)_d = &-R^{AB}\partial_u\delta\mu_{AB} - \mu^{AB}\Big[R^{ru}\partial_u\delta\mu_{AB} + R^{rr}\partial_u\delta\mu_{AB} + R^{rC}\hat{D}_C\delta\mu_{AB}\Big] \\ &- 2\delta\mu_{AB}\Big[\Gamma^A_{ur}R^{rB} + \Gamma^A_{uC}R^{CB}\Big] \\ &+ 2\mu^{AB}(\delta\mu_{BC})\Big[\Gamma^C_{uA}R^{ru} + \Gamma^C_{rA}R^{rr}\Big]. \end{aligned} \tag{8.2.64}$$

This result required a fair amount of index manipulations and we used most of the relations of appendix D. In a similar manner we find

$$\begin{aligned} Z_3^d(\mathrm{d}r)_d = &(\delta\mu_{AB})\Big[\partial_u R^{AB} + 2\Gamma^A_{ur}R^{rB} + 2\Gamma^A_{uC}R^{CB}\Big] \\ &+ \mu^{AB}(\delta\mu_{AB})\Big[\partial_u R^{ur} + \partial_r R^{rr} + \hat{D}_C R^{rC} + 3\Gamma^r_{rC}R^{rC} \\ &\quad + \Gamma^C_{Cu}R^{ur} + \Gamma^C_{Cr}R^{rr} + \Gamma^r_{CD}R^{CD}\Big] \\ &+ \mu^{AB}(\delta\mu_{BC})\Big[\hat{D}_A R^{rC} + \Gamma^C_{Au}R^{ur} + \Gamma^C_{Ar}R^{rr} + \Gamma^r_{Ar}R^{rC} + \Gamma^r_{AD}R^{CD}\Big]. \end{aligned} \tag{8.2.65}$$

From this follows, after some further simplifications,

$$\begin{aligned} [Z_1^d + Z_2^d](\mathrm{d}r)_d = &(\partial_u R^{AB})\delta\mu_{AB} - R^{AB}\partial_u\delta\mu_{AB} \\ &+ \mu^{AB}\Big[(\partial_u R^{ur})\delta\mu_{AB} + (\partial_r R^{rr})\delta\mu_{AB} + (\hat{D}_C R^{rC})\delta\mu_{AB} \\ &\quad - R^{ur}\partial_u\delta\mu_{AB} - R^{rr}\partial_r\delta\mu_{AB} - R^{rC}\hat{D}_C\delta\mu_{AB}\Big] \\ &+ \mu^{AB}(\delta\mu_{BC})\Big[R^{ru}\mu^{CD}\partial_u\mu_{AD} + R^{rr}\mu^{CD}\partial_r\mu_{AD} - \hat{D}_A R^{rC}\Big]. \end{aligned} \tag{8.2.66}$$

The terms which involve $\partial_u R^{AB}$, $\partial_u R^{ur}$ and $\partial_u \mu_{AB}$ vanish in the stationary case, so we do not need to worry about them anymore. Since ∂_u and δ commute, we have

$$R^{AB}\partial_u\delta\mu_{AB} = \delta(R^{AB}\partial_u\mu_{AB}) - (\delta R^{AB})\partial_u\mu_{AB} \tag{8.2.67}$$

$$R^{ur}\mu^{AB}\partial_u\delta\mu_{AB} = \delta(R^{ur}\mu^{AB}\partial_u\mu_{AB}) - \delta(R^{ur}\mu^{AB})\partial_u\mu_{AB}, \tag{8.2.68}$$

so these terms are also fine, since they can be decomposed as we wish. Furthermore, we have

$$(\hat{D}_C R^{rC})\mu^{AB}\delta\mu_{AB} - \mu^{AB}R^{rC}\hat{D}_C\delta\mu_{AB} = 2(\hat{D}_C R^{rC})\mu^{AB}\delta\mu_{AB} - \hat{D}_C(R^{rC}\mu^{AB}\delta\mu_{AB}), \tag{8.2.69}$$

since $\hat{D}_A$ is compatible with μ_{AB}, i.e. $\hat{D}_A\mu_{BC} = 0$. The total divergence term can be omitted since it drops out after using Stokes' theorem and the boundary conditions. The remaining terms may be treated as follows: We have

$$\begin{aligned}
&(\partial_r R^{rr})\mu^{AB}\delta\mu_{AB} - R^{rr}\mu^{AB}\partial_r\delta\mu_{AB} + \mu^{AB}\mu^{CD}(\delta\mu_{BC})R^{rr}\partial_r\mu_{AD} \\
&\quad = (\partial_r R^{rr})\mu^{AB}\delta\mu_{AB} - R^{rr}\mu^{AB}\delta\partial_r\mu_{AB} - R^{rr}(\delta\mu^{AB})\partial_r\mu_{AB} \\
&\quad = (\partial_r R^{rr})\mu^{AB}\delta\mu_{AB} - \delta(R^{rr}\mu^{AB}\partial_r\mu_{AB}) + (\delta R^{rr})\mu^{AB}\partial_r\mu_{AB},
\end{aligned} \tag{8.2.70}$$

where we have used $[\partial_r, \delta] = 0$. At $r = 0$ (on the horizon), the first term in (8.2.70) may be written as

$$\begin{aligned}
\partial_r R^{rr} &= \partial_r R_{uu} \\
&= -\partial_r\left[\frac{1}{2}\mu^{AB}\partial_u^2\mu_{AB} + \frac{1}{4}(\partial_u\mu^{AB})\partial_u\mu_{AB}\right].
\end{aligned} \tag{8.2.71}$$

Here we have used the results from appendix D.2. As we see, it vanishes in the stationary case, since we have $[\partial_u, \partial_r] = 0$. The second term in (8.2.70) is a total variation. The last term in (8.2.70) may be written as (see appendix D.2)

$$\begin{aligned}
\delta R^{rr}\mu^{AB}\partial_r\mu_{AB} &= \delta R_{uu}\mu^{AB}\partial_r\mu_{AB} \\
&= -\bigg\{\frac{1}{2}\mu^{CD}\delta\partial_u^2\mu_{CD} + \frac{1}{2}(\delta\mu^{CD})\partial_u^2\mu_{CD} \\
&\quad + \frac{1}{4}\Big[(\delta\partial_u\mu^{CD})\partial_u\mu_{CD} + (\partial_u\mu^{CD})\delta\partial_u\mu_{CD}\Big]\bigg\}\mu^{AB}\partial_r\mu_{AB}.
\end{aligned} \tag{8.2.72}$$

Except for the first term in (8.2.72), all other terms vanish in the stationary case. By omitting these, we find

$$\begin{aligned}
(\delta R^{rr})\mu^{AB}\partial_r\mu_{AB} &= -\frac{1}{2}\mu^{CD}(\delta\partial_u^2\mu_{CD})\mu^{AB}\partial_r\mu_{AB} \\
&= -\frac{1}{2}\delta\Big[\mu^{CD}(\partial_u^2\mu_{CD})\mu^{AB}\partial_r\mu_{AB}\Big] \\
&\quad + \frac{1}{2}(\partial_u^2\mu_{CD})\delta\Big[\mu^{CD}\mu^{AB}\partial_r\mu_{AB}\Big].
\end{aligned} \tag{8.2.73}$$

Again, the last term vanishes in the stationary case. The remaining terms in (8.2.66), which still need to be treated, are the following:

$$2(\hat{D}_C R^{rC})\mu^{AB}\delta\mu_{AB} - (\hat{D}_A R^{rC})\mu^{AB}\delta\mu_{BC}. \tag{8.2.74}$$

We have (see appendix D.2)

$$\begin{aligned}
\hat{D}_A R^{rC} &= \hat{D}_A(\mu^{CD} R_{uD}) \\
&= \mu^{CD}\hat{D}_A R_{uD} \\
&= \mu^{CD}\hat{D}_A\left[\frac{1}{2}\partial_u\beta_D + \frac{1}{4}\beta_D\mu^{EF}\partial_u\mu_{EF} - \frac{1}{2}\hat{D}_D(\mu^{EF}\partial_u\mu_{EF}) + \frac{1}{2}\hat{D}_E(\mu^{EF}\partial_u\mu_{DF})\right] \\
&= \mu^{CD}\left[\frac{1}{2}\partial_u\hat{D}_A\beta_D + \frac{1}{4}(\hat{D}_A\beta_D)\mu^{EF}\partial_u\mu_{EF}\right].
\end{aligned} \tag{8.2.75}$$

For the first equality we used the results for the Ricci tensor components in GNC from appendix D.2 and for the second one we used $\hat{D}_A\mu_{BC} = 0$. For the third equality we used, again, the results from appendix D.2. For the fourth equality we used $[\hat{D}_A, \partial_u] = 0$ and and the fact that we have $\hat{D}_A\mu_{BC} = 0$ on each cross-section of the horizon, so this property does not change along the flowlines of n^a, i.e. we have $\partial_u\hat{D}_A\mu_{BC} = 0$. Therefore, all the terms in (8.2.75) vanish in the stationary case, so (8.2.74) does not contribute to W_{abc}.

By putting everything together, we find the following decomposition of (8.2.66):

$$\begin{aligned}
[Z_1^d + Z_2^d](\mathrm{d}r)_d = Z_{\text{stat}} \\
- \delta\left[R^{AB}\partial_u\mu_{AB} + R^{ur}\mu^{AB}\partial_u\mu_{AB} + R^{rr}\mu^{AB}\partial_r\mu_{AB} + \frac{1}{2}\mu^{AB}\mu^{CD}(\partial_r\mu_{AB})\partial_u^2\mu_{CD}\right],
\end{aligned} \tag{8.2.76}$$

where Z_{stat} denotes all the terms that vanish in the stationary case. We have

$$\begin{aligned}
Z_{\text{stat}} = {} & (\partial_u R^{AB})\delta\mu_{AB} + (\partial_u R^{ur})\mu^{AB}\delta\mu_{AB} - R^{ur}(\delta\mu^{AB})\partial_u\mu_{AB} \\
& + (\delta R^{AB})\partial_u\mu_{AB} + \delta(R^{ur}\mu^{AB})\partial_u\mu_{AB} \\
& + \frac{1}{2}\mu^{AB}(\delta\mu_{AB})\Big\{(\partial_r\mu^{CD})\partial_u^2\mu_{CD} + \mu^{CD}\partial_u^2\partial_r\mu_{CD} + \frac{1}{2}(\partial_u\partial_r\mu^{CD})\partial_u\mu_{CD} \\
& \qquad + \frac{1}{2}(\partial_u\mu^{CD})\partial_u\partial_r\mu_{CD} + 2\mu^{CD}\partial_u\hat{D}_C\beta_D \\
& \qquad + \mu^{CD}\mu^{EF}(\hat{D}_C\beta_D)\partial_u\mu_{EF}\Big\} \\
& + \frac{1}{2}(\delta\mu^{AB})\Big\{\partial_u\hat{D}_A\beta_B + \frac{1}{2}\mu^{CD}(\hat{D}_A\beta_B)\partial_u\mu_{CD}\Big\} \\
& - \frac{1}{2}\mu^{AB}(\partial_r\mu_{AB})\Big\{(\delta\mu^{CD})\partial_u^2\mu_{CD} - (\partial_u^2\mu_{CD})\delta[\mu^{CD}\mu^{EF}\partial_r\mu_{EF}] \\
& \qquad + \frac{1}{2}\Big[(\delta\partial_u\mu^{CD})\partial_u\mu_{CD} + (\partial_u\mu^{CD})\delta\partial_u\mu_{CD}\Big]\Big\}.
\end{aligned} \tag{8.2.77}$$

By combining this result with the results from section 8.2.1 we find

$$W_{abc} = \frac{1}{16\pi}\Big\{2\vartheta + \lambda\Big[R^{AB}\partial_u\mu_{AB} + R^{ur}\mu^{AB}\partial_u\mu_{AB} + R^{rr}\mu^{AB}\partial_r\mu_{AB} + \frac{1}{2}\mu^{AB}\mu^{CD}(\partial_r\mu_{AB})\partial_u^2\mu_{CD}\Big]\Big\}\overset{(3)}{\epsilon}_{abc}. \tag{8.2.78}$$

in our HDTG. On the horizon we have (see appendix D.2)

$$\begin{aligned} R^{rr} &= R_{uu} \\ &= -\frac{1}{2}\mu^{AB}\partial_u^2\mu_{AB} - \frac{1}{4}(\partial_u\mu^{AB})\partial_u\mu_{AB}. \end{aligned} \tag{8.2.79}$$

Therefore, equation (8.2.78) can be rewritten as

$$W_{abc} = \frac{1}{16\pi}\Big\{2\vartheta + \lambda\Big[R^{AB}\partial_u\mu_{AB} + R^{ur}\mu^{AB}\partial_u\mu_{AB} - \frac{1}{4}\mu^{AB}(\partial_r\mu_{AB})(\partial_u\mu^{CD})\partial_u\mu_{CD}\Big]\Big\}\overset{(3)}{\epsilon}_{abc}. \tag{8.2.80}$$

This expression can be made covariant in the following manner: One has to make the observations that, in an adapted coordiante system, the Lie derivative with respect to the vector field n^a can be written (locally) as ∂_u and the vector fields $(\partial/\partial u)^a$ and $(\partial/\partial x^A)^a$ clearly commute since they are coordinate vector fields. A similar statement holds for the Lie derivative with respect to the vector field l^a. Locally it may be written as ∂_r and it commutes with the vector fields $(\partial/\partial x^A)^a$ since they coordinate vector fields. Furthermore, the Lie derivative with respect to the vector field n^a (l^a) of the 1-form $(\mathrm{d}x^A)_a$ vanishes. From this follows that we can make the following replacements

$$\begin{aligned} \mu_{AB} &\to \mu_{ab} = \mu_{AB}(\mathrm{d}x^A)_a(\mathrm{d}x^B)_b \\ \mu^{AB} &\to \mu^{ab} = \mu^{AB}(\partial_A)^a(\partial_B)^b \\ R^{AB} &\to R^{ab} = R^{AB}(\partial_A)^a(\partial_B)^b \\ \partial_u &\to \mathcal{L}_n \\ \partial_r &\to \mathcal{L}_l \end{aligned} \tag{8.2.81}$$

Therefore, we find

$$W_{abc} = \frac{1}{16\pi}\Big\{2\vartheta + \lambda\Big[R^{de}\mathcal{L}_n\mu_{de} + R^{ur}\mu^{de}\mathcal{L}_n\mu_{de} - \frac{1}{4}\mu^{de}(\mathcal{L}_l\mu_{de})(\mathcal{L}_n\mu^{fg})\mathcal{L}_n\mu_{fg}\Big]\Big\}\overset{(3)}{\epsilon}_{abc}. \tag{8.2.82}$$

Now we will use the decompositions $\mathcal{L}_n\mu_{ab} = 2\sigma_{ab} + \vartheta\mu_{ab}$ and $\mathcal{L}_n\mu^{ab} = g^{ac}g^{bd}\mathcal{L}_n\mu_{cd} = 2\sigma^{ab} + \vartheta\mu^{ab}$. The second decompotion follows from the first one under the assumption that n^a is a Killing vector field. We obtain

$$W_{abc} = \frac{1}{16\pi}\Big\{2\vartheta + \lambda\Big[2R^{de}\sigma_{de} + \vartheta R^{de}\mu_{de} + 2R^{ur}\vartheta - \mu^{de}(\mathcal{L}_l\mu_{de})(\sigma_{fg}\sigma^{fg} + \frac{1}{2}\vartheta^2)\Big]\Big\}\overset{(3)}{\epsilon}_{abc}. \tag{8.2.83}$$

Furthermore, since we have $g^{ab} = \mu_{ab} + 2(\partial_u)^{(a}(\partial_r)^{b)}$ we find

$$\begin{aligned} R^{ur} = R_{ur} &= R_{ab}(\partial_u)^a(\partial_r)^b \\ &= R^{ab}\Big[(\partial_u)^{(a}(\partial_r)^{b)} + \frac{1}{2}\mu^{ab} - \frac{1}{2}\mu^{ab}\Big] \\ &= \frac{1}{2}R_{ab}\Big[2(\partial_u)^{(a}(\partial_r)^{b)} + \mu_{ab}\Big] - \frac{1}{2}R_{ab}\mu^{ab} \\ &= \frac{1}{2}R_{ab}g^{ab} - \frac{1}{2}R_{ab}\mu^{ab} \\ &= \frac{1}{2}(R - R_{ab}\mu^{ab}). \end{aligned} \tag{8.2.84}$$

Insertion of this result into (8.2.83) yields

$$W_{abc} = \frac{1}{16\pi}\Big\{2\vartheta + \lambda\Big[2R^{de}\sigma_{de} + \vartheta R - \mu^{de}(\mathcal{L}_l\mu_{de})(\sigma^2 + \frac{1}{2}\vartheta^2)\Big]\Big\}\overset{(3)}{\epsilon}_{abc}, \tag{8.2.85}$$

where we defined $\sigma^2 := \sigma_{ab}\sigma^{ab}$.

Now, we can write down for the conserved quantity $\mathcal{H}_n$. We have

$$\begin{aligned} \mathcal{H}_n = &\int_{\mathcal{E}} \boldsymbol{Q} + n\cdot\boldsymbol{W} \\ =&\frac{1}{8\pi}\mathcal{L}_n\mathcal{A}(\mathcal{E}) + \frac{\lambda}{16\pi}\Big\{\int_{\mathcal{E}}\Big[2R^{cd}\sigma_{cd} + \vartheta R - \mu^{cd}(\mathcal{L}_l\mu_{cd})(\sigma^2 + \frac{1}{2}\vartheta^2)\Big]\overset{(2)}{\epsilon}_{ab} - \int_{\mathcal{E}}\epsilon_{abcd}X^{cd}\Big\}, \end{aligned} \tag{8.2.86}$$

where X^{ab} is given by (8.2.56). In appendix D.5 we computed the pullback of the 2-form $\epsilon_{abcd}X^{cd}$ to a horizon cross-section $\mathcal{E}$. Our result is

$$\begin{aligned} \int_{\mathcal{E}}\epsilon_{abcd}X^{cd} = \int_{\mathcal{E}}\Big\{&\mathcal{L}_nR - \mathcal{L}_n(R_{cd}\mu^{cd}) + \mu^{cd}\hat{D}_c(n^eR_{ed}) - n^c\beta^dR_{cd} \\ &+\frac{1}{2}\Big[\vartheta R + 2R_{cd}\sigma^{cd} + \mu^{cd}(\mathcal{L}_l\mu_{cd})R_{ef}n^en^f\Big]\Big\}\overset{(2)}{\epsilon}_{ab}. \end{aligned} \tag{8.2.87}$$

As we see, the "conserved quantity" $\mathcal{H}_n$ involves the Lie derivative with respect to the vector field l^a of the metric μ_{ab} on the spatial cross-sections $\mathcal{E}$. Since the behaviour of μ_{ab} off of the horizon is essentially arbitrary, we strongly doubt that it is possible to make any statement about the positvity of $\mathcal{H}_n$.

Furthermore, we clearly have $\mathcal{H}_n[g_0] = 0$, where g_0 is the reference solution (Schwarzschild spacetime): We have $R_{ab}[g_0] = 0$ and therefore $R[g_0]$ since g_0 is a solution of the vacuum Einstein equations. Furthermore, for g_0 it is known that the horizon generators have vanishing expansion and shear. Therefore, all horizon cross-sections are isometric and we have $\mathcal{L}_n\mathcal{A}(\mathcal{E}) = 0$. Hence, the constant C which was involved in the ambiguity of $\mathcal{H}_n$ can be really fixed to zero.

8.2.5. Calculation of $\boldsymbol{F}_n$ in our HDTG

An interesting byproduct of the calculation from the previous section is that it is now very easy to write down the flux $\boldsymbol{F}_n = \boldsymbol{\Theta}(g, \mathcal{L}_n g)$ which appeared in equation (7.3.6). In order to obtain this quantity, we need to collect the terms in the decomposition $\iota^*\boldsymbol{\theta} = \boldsymbol{\Theta} + \delta\boldsymbol{W}$ that vanish in the stationary case and replace $\delta g \to \mathcal{L}_n g$. Again, the expression for the flux can be made covariant by making the replacements

$$\mu_{AB} \to \mu_{AB}(\mathrm{d}x^A)_a(\mathrm{d}x^B)_b =: \mu_{ab} \tag{8.2.88}$$

$$\mu^{AB} \to \mu^{AB}(\partial_A)^a(\partial_B)^b =: \mu^{ab} \tag{8.2.89}$$

$$R_{AB} \to R_{AB}(\mathrm{d}x^A)_a(\mathrm{d}x^B)_b =: R_{ab} \tag{8.2.90}$$

$$\partial_u \to \mathcal{L}_n \tag{8.2.91}$$

$$\partial_r \to \mathcal{L}_l \tag{8.2.92}$$

As we explained below equation (8.2.80), this procedure is consistent. In addition, we will make the replacements

$$\hat{D}_A \to (\mathrm{d}x^A)_a\hat{D}_A =: \hat{D}_a \tag{8.2.93}$$

$$\beta^A \to (\partial_A)^a\beta^A =: \beta^a. \tag{8.2.94}$$

As a result we obtain

$$\begin{aligned}(F_n)_{abc} = \frac{1}{16\pi}\overset{(3)}{\epsilon}_{abc}\bigg\{&2\vartheta^2 + \lambda\bigg[2(\mathcal{L}_nR^{de})\mathcal{L}_n\mu_{de} + \frac{1}{2}(\mathcal{L}_n\mu^{de})\mathcal{L}_n\hat{D}_d\beta_e \\ &+ \vartheta\bigg\{2\mathcal{L}_nR - 2\mathcal{L}_n(R^{de}\mu_{de}) + R^{de}\mathcal{L}_n\mu_{de} + 2\mu^{de}\mathcal{L}_n\hat{D}_d\beta_e \\ &\qquad + \mathcal{L}_l(\mu^{de}\mathcal{L}_n\mathcal{L}_n\mu_{de}) + (\mathcal{L}_n\mathcal{L}_l\mu^{de})\mathcal{L}_n\mu_{de}\bigg\} \\ &+ \vartheta^2\bigg\{2\mu^{de}\hat{D}_d\beta_e + \frac{1}{2}(\mathcal{L}_n\mu^{de})\mathcal{L}_n\hat{D}_d\beta_e + R - R^{de}\mu_{de}\bigg\} \\ &- \frac{1}{2}\mu^{de}(\mathcal{L}_l\mu_{de})\bigg\{2(\mathcal{L}_n\mu^{fg})\mathcal{L}_n\mathcal{L}_n\mu_{fg} + 2(\mathcal{L}_n\vartheta\mu^{fg})\mathcal{L}_n\mathcal{L}_n\mu_{fg} \\ &\qquad - \mu^{fg}\mathcal{L}_n\mathcal{L}_n\mu_{fg} + 2(\mathcal{L}_n\vartheta + \frac{1}{2}\vartheta^2 + \sigma^2)\bigg\}\bigg]\bigg\}.\end{aligned} \tag{8.2.95}$$

Conclusion and Outlook

In this thesis we presented two attempts for a proof of a second law of black hole mechanics in a theory of gravity which an additional $R_{ab}R^{ab}$-contribution in its gravitational Lagrangean. Neither of these approaches were successful in the sense that we were not able to answer the question whether it is possible to establish such a theorem in this gravitational theory or not.

The first idea for a proof was not further pursued on the level when we inserted the uu-component of the field equations into our Ansatz for the evolution equation. The amount of terms that were involved was so overwhelmingly large, that it was not possible to see any structure in the resulting equation or to bring it in the desired form.

One way to further pursue this strategy would be to insert the components of the Ricci-tensor (rewritten in Gaussian null corrdinates) that were involved and to use a computer algebra program to simplify the resulting equation. Maybe on this level it is possible to see what the structure of the evolution equation is.

One should of course also note, that the Ansatz for the evolution equation which we took, was in analogy with the Einstein-case, and therefore more or less ad-hoc.

But the (laborious) work that was done in this approach should not be considered to be completely in vain. The result for the uu-component of the field equations can be also used for other purposes, such as an attempt to prove a rigidity theorem in the HDTG which we considered.

The second idea for a proof was not further pursued on the level when the conserved quantity $\mathcal{H}_n$ was computed. As we already mentioned, we strongly doubt that it is possible to make any statement about the positivity of $\mathcal{H}_n$, since the Lie derivative with respect to the vector field l^a of the metric μ_{ab} on a cross-section $\mathcal{E}$ is involved in the explicit expression for the "conserved quantity". However, this issue was not analyzed in detail, due to the amount of time that was already spent for the first approach.

The main achievement of this part is that we succesfully applied the Wald-Zoupas-formalism for the definition of conserved quantities to the horizon of a black hole. To our knowledge, this has not been done so far in the literature. The result for $\mathcal{H}_n$ in the Einstein-case suggests that our modifications to this formalism yield indeed a meaningful result, since it is related to the rate of change of the black hole entropy along the null geodesic generator of the event horizon.

One should note that the results for $\mathcal{H}_n$ (in the Einstein-case and in our HDTG) are not unique, since we made a particular choice for the quantity $\mathbf{\Theta}$. However, this choice seems to be the most natural one.

Even though we were not able answer the (ambitious) question if a second law exists in our HDTG, the results of this thesis should not be considered to be useless, since they can be used in other contexts as well. Furthermore, we gained the insight that our first idea might not be the most elegant method to prove a second law. Finally, we give indications how the Wald-

Zoupas formalism should be modified in order to define conserved quantities on the horizon of a black hole. Further investigation of this issue might answer the question how to define the entropy of a nonstationary black hole.

Appendices

A. Notation and Conventions

Throughout this thesis we will use the conventions from [29]. Lowercase latin indices $a, b, c, \ldots$ will denote abstract tensor indices. Lowercase greek indices $\alpha, \beta, \gamma, \ldots$ will denote tensor components in a particular coordinate system. In section 8, we will use uppercase latin indices $A, B, C, \ldots$ to denote components (in the Gaussian null corrdinate system $\{u, r, x^A\}$) of the induced metric $\mu_{ab} = \mu_{AB}(\mathrm{d}x^A)_a(\mathrm{d}x^B)_b$ of the 2-dimensional submanifold, which is generated by intersecting the event horizon with a spacelike hypersurface. Furthermore, throughout chapters 7 and 8 we will use boldface letters to denote differential forms on the spacetime manifold and, when we do so, the spacetime indices of the forms will be suppressed.

The spacetime manifold (or spacetime for short) will be denoted by the pair (M, g_{ab}), where M is a smooth 4-dimensional[1] connected paracompact oriented manifold, and g_{ab} is a Lorentzian metric with signature $(-, +, +, +)$. Furthermore, the spacetime (M, g_{ab}) is assumed to be time orientable. The canonical volume 4-form on M will be denoted by

$$\boldsymbol{\epsilon} = \epsilon_{abcd} = \sqrt{-g}\,(\mathrm{d}x^0)_a \wedge (\mathrm{d}x^2)_b \wedge (\mathrm{d}x^2)_c \wedge (\mathrm{d}x^3)_d =: \sqrt{-g}\,\mathrm{d}^4x, \tag{A.0.1}$$

where $\{x^0, x^1, x^2, x^3\}$ is right handed, and $\sqrt{-g}$ is the square root of minus the determinant of the metric g_{ab}. Throughout section 7, the abstract indices of the metric will be suppressed in order to simplify the notation. From the context it should be clear what is meant. The covariant derivative ∇_a on M is chosen to be torsion-free and compatible with the metric, i.e. we have $\nabla_a g_{bc} = 0$. Similary, we chose a torsion-free derivative operator $\hat{D}_A$ associated with μ_{AB}, i.e. $\hat{D}_A \mu_{BC} = 0$. Furthermore we define the operators $\Box := \nabla^a \nabla_a = g^{ab} \nabla_a \nabla_b$ and $\hat{\Box} := \hat{D}^a \hat{D}_a = g^{ab} \hat{D}_a \hat{D}_b$.

Abstract tensor indices will be raised and lowered with the metric g_{ab} and its inverse g^{ab}, i.e. we have $T_a{}^b = g^{bc} T_{ac}$ and $T^a{}_b = g_{bc} T^{ac}$. In a similar manner boldface latin indices will be raised and lowered with μ_{AB} and its inverse μ^{AB}.

For a diffeomorphism $\psi : M \to N$ between manifolds M and N, we denote the pullback of a tensor field $T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$ by $(\psi^* T)^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} = \psi^* T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$. The push-forward of a tensor field $S^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$ will be denoted by $(\psi_* S)^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} = \psi_* S^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$.

We define the Riemann tensor $R_{abc}{}^d$ by $(\nabla_a \nabla_b - \nabla_b \nabla_a)\omega_c = R_{abc}{}^d \omega_d$, the Ricci tensor R_{ab} by $R_{ab} = R_{acb}{}^c$ and the Ricci scalar R by $R = R_a{}^a$.

[1] Most of the results in this thesis can be easily formulated in arbitrary dimensions $d \geq 4$. However, since certain theorems in black hole physics (such as the rigidity, black hole uniqueness, and topology theorem) do not readily extend to arbitrary dimensions, we decided to stick to $d = 4$ in order keep their presentation as simple as possible. Furthermore, since we are primarily concerned gravitational theories with higher derivative contributions, whose presence already causes severe problems, we decided to place the focus on higher derivative instead higher dimensional theories of gravity.

Symmetrization of tensor indices will be denoted by parenthesis and antisymmetrization will be denoted by brackets.

Throughout this thesis we will use the standard shorthand notation in variational calculations: $\mathrm{d}g^{(t)}_{ab}/\mathrm{d}t|_{t=0}$ will be denoted by δg_{ab}, where, $g^{(t)}_{ab}$ is a one-parameter family metrics, such that $g^{(t)}_{ab}$ depends differentiably on t and $g^{(0)}_{ab}$ satisfies appropriate boundary conditions. Similary, the variation δI of a functional $g_{ab} \mapsto I[g_{ab}]$ is understood in the following way: replace g_{ab} by $g^{(t)}_{ab}$ and differentiate $I[g^{(t)}_{ab}]$ with respect to t and set $t = 0$ afterwards.

In chapter 8 we will use the shorthand notation ∂_u and ∂_r for the derivative operators with respect to the coordinates u and r, respectively, in the Gaussian null coordinate system $\{u, r, x^A\}$. By these operators we actually mean the Lie derivatives $\mathcal{L}_n$ and $\mathcal{L}_l$ with respect to the vector fields $(\partial/\partial u)^a$ and $(\partial/\partial r)^a$, respectively, such that covariance is preserved at all steps. This identification is justified, since one can always write a Lie derivative (locally) as a coordinte dervative, in a suitably adapted coordiate system.

Furthermore, we will work in units with $G = \hbar = c = k_B = 1$.

B. Energy Conditions in General Relativity

(WEC) **Weak Energy Condition**
The scalar $T_{ab}\xi^a\xi^b$ represents the energy density of matter in a frame defined by the timelike vector field ξ^a. If the energy density is positive in all frames, we should have

$$T_{ab}\xi^a\xi^b \geq 0 \tag{B.0.1}$$

for all future directed timelike vectors ξ^a. Condition (B.0.1) is known as *weak energy condition*.

(SEC) **Strong Energy Condition**
This condition states that we have

$$T_{ab}\xi^a\xi^b \geq -\frac{1}{2}T \tag{B.0.2}$$

for all future directed unit timelike vectors ξ^a.

(DEC) **Dominant Energy Condition**
The vector field $-T^a{}_b\xi^a$ represents the energy- momentum 4-current density in a frame defined by the timelike vector field ξ^a. It is believed that the current density flux should always have velocity smaller than the speed of light. Hence we have

$$-T^a{}_b\xi^a \text{ is future directed timelike or null} \tag{B.0.3}$$

for all future directed timelike vectors ξ^a. Condition (B.0.3) is known an *dominant energy condition*.

(NEC) **Null Energy Condition**
This condition states that we have

$$T_{ab}k^ak^b \geq 0 \tag{B.0.4}$$

for all future directed null vectors k^a.

Note that (DEC) implies (WEC), but (SEC) does *not* imply (WEC). Furthermore, (NEC) is implied by (WEC) and (SEC) using continuity arguments. Of particular interest is the

(NCC) **Null Convergence Condition**
This condition states that we have

$$R_{ab}k^ak^b \geq 0 \tag{B.0.5}$$

for all future directed null vectors k^a.

By using Einstein's equation, (NCC) is implied by (NEC).

C. Useful Relations

On an n-dimensional manifold (M, g_{ab}) the totally antisymmetric tensor $\epsilon_{a_1 \dots a_n}$ satisfies the following relations

$$\epsilon^{a_1 \dots a_n} \epsilon_{b_1 \dots b_n} = (-1)^s \; n! \; \delta^{[a_1}{}_{b_1} \dots \delta^{a_n]}{}_{b_n} \tag{C.0.1}$$

$$\epsilon^{a_1 \dots a_j a_{j+1} \dots a_n} \epsilon_{a_1 \dots a_j b_{j+1} \dots b_n} = (-1)^s \; (n-j)! \; j! \; \delta^{[a_{j+1}}{}_{b_{j+1}} \dots \delta^{a_n]}{}_{b_n}, \tag{C.0.2}$$

where s is the number of minuses appearing in the signature of g_{ab}.

Let K^a be a Killing vector field. From the definition of the Riemann tensor, together with Killing's equation follows

$$\nabla_a \nabla_b K_c = -R_{bca}{}^d K_d. \tag{C.0.3}$$

D. More on Gaussian Null Coordinates

The metric (3.2.8), and its inverse, may be written in matrix notation as

$$(g)_{\mu\nu} = \begin{pmatrix} -2r^2\alpha & 1 & -r\beta^A \\ 1 & 0 & 0 \\ -r\beta^A & 0 & \mu_{AB} \end{pmatrix}, \qquad (g^{-1})^{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & r^2(\beta^E\beta_E + 2\alpha) & r\beta^A \\ 0 & r\beta^A & \mu^{AB} \end{pmatrix}, \tag{D.0.1}$$

where $\mu^{AB} = (\mu^{-1})^{AB}$ is the inverse matrix of μ_{AB} and where $\beta^A = \mu^{AB}\beta_B$.

D.1. Christoffel Symbols

By introducing $\bar{\alpha} = -2r^2\alpha$ and $\bar{\beta}_A = -r\beta_A$ we find

$$\Gamma^u_{uu} = -\frac{1}{2}\partial_r\bar{\alpha} \tag{D.1.1}$$

$$\Gamma^u_{uA} = -\frac{1}{2}\partial_r\bar{\beta}_A \tag{D.1.2}$$

$$\Gamma^u_{AB} = -\frac{1}{2}\partial_r\mu_{AB} \tag{D.1.3}$$

$$\Gamma^u_{ur} = \Gamma^u_{rr} = \Gamma^u_{rA} = 0 \tag{D.1.4}$$

$$\Gamma^r_{ru} = \frac{1}{2}(\partial_r\bar{\alpha} - \bar{\beta}^C\partial_r\bar{\beta}_C) \tag{D.1.5}$$

$$\Gamma^r_{rA} = \frac{1}{2}(\partial_r\bar{\beta}_A - \bar{\beta}^C\partial_r\mu_{CA}) \tag{D.1.6}$$

$$\Gamma^r_{uu} = -\frac{1}{2}(\bar{\beta}^E\bar{\beta}_E - \bar{\alpha})\partial_r\bar{\alpha} + \frac{1}{2}\partial_u\bar{\alpha} + \frac{1}{2}\bar{\beta}^C\hat{D}_C\bar{\alpha} - \bar{\beta}^C\partial_u\bar{\beta}_C \tag{D.1.7}$$

$$\Gamma^r_{AB} = -\frac{1}{2}\{\partial_u\mu_{AB} + (\bar{\beta}^E\bar{\beta}_E - \bar{\alpha})\partial_r\mu_{AB}\} + \frac{1}{2}(\hat{D}_A\bar{\beta}_B + \hat{D}_B\bar{\beta}_A) \tag{D.1.8}$$

$$\Gamma^r_{uA} = -\frac{1}{2}(\bar{\beta}^E\bar{\beta}_E - \bar{\alpha})\partial_r\bar{\beta}_A + \frac{1}{2}\hat{D}_A\bar{\alpha} - \frac{1}{2}\bar{\beta}^B(\partial_u\mu_{AB} + \hat{D}_A\bar{\beta}_B - \hat{D}_B\bar{\beta}_A) \tag{D.1.9}$$

$$\Gamma^r_{rr} = 0 \tag{D.1.10}$$

$$\Gamma^A_{BC} = \frac{1}{2}\bar{\beta}^A\partial_r\mu_{BC} + \hat{\Gamma}^A_{BC} \tag{D.1.11}$$

$$\Gamma^A_{Bu} = \frac{1}{2}\bar{\beta}^A\partial_r\bar{\beta}_B + \frac{1}{2}\mu^{CA}\partial_u\mu_{BC} + \frac{1}{2}\mu^{CA}(\hat{D}_B\bar{\beta}_C - \hat{D}_C\bar{\beta}_B) \tag{D.1.12}$$

$$\Gamma^A_{Br} = \frac{1}{2}\mu^{CA}\partial_r\mu_{BC} \tag{D.1.13}$$

$$\Gamma^A_{uu} = \frac{1}{2}\bar{\beta}^A\partial_r\bar{\alpha} - \frac{1}{2}\mu^{CA}\hat{D}_C\bar{\alpha} + \mu^{AC}\partial_u\bar{\beta}_C \tag{D.1.14}$$

$$\Gamma^A_{ur} = \frac{1}{2}\mu^{CA}\partial_r\bar{\beta}_C \tag{D.1.15}$$

$$\Gamma^A_{rr} = 0, \tag{D.1.16}$$

where $\hat{D}_A$ is the derivative operator associated with the matrix μ_{AB}, i.e. we have

$$\hat{D}_A\omega_B = \partial_A\omega_B - \hat{\Gamma}^C_{AB}\omega_C, \tag{D.1.17}$$

where

$$\hat{\Gamma}^C_{AB} = \frac{1}{2}\mu^{CD}(\partial_A\mu_{BD} + \partial_B\mu_{AD} - \partial_D\mu_{AB}). \tag{D.1.18}$$

Furthermore, if we introduce $\mu = \det(\mu_{AB})$ we have

$$\partial_*\mu^{AB} = -\mu^{AC}\mu^{BD}\partial_*\mu_{CD} \tag{D.1.19}$$

$$\sqrt{\mu}^{-1}\partial_*\sqrt{\mu} = \frac{1}{2}\mu^{AB}\partial_*\mu_{AB} \tag{D.1.20}$$

$$\partial_u(\sqrt{\mu}^{-1}\partial_r\sqrt{\mu}) = -\frac{1}{2}\mu^{AC}\mu^{BD}(\partial_u\mu_{CD})\partial_r\mu_{AB} + \frac{1}{2}\partial_u\partial_r\mu_{AB}, \tag{D.1.21}$$

where $_*$ stands for u, r or A.

D.2. Ricci Tensor

Here we will give a list of useful relations between components of the Ricci tensor in GNC. Each of these relations only holds when we are restricted to the horizon $r = 0$.

$$R^{uu} = R_{rr} \tag{D.2.1}$$

$$R^{rr} = R_{uu} \tag{D.2.2}$$

$$R^{AB} = \mu^{AC}\mu^{BD}R_{CD} \tag{D.2.3}$$

$$R^{ur} = R_{ur} \tag{D.2.4}$$

$$R^{uA} = \mu^{AB}R_{rB} \tag{D.2.5}$$

$$R^{rA} = \mu^{AB}R_{uB}. \tag{D.2.6}$$

and

$$
\begin{aligned}
R^A{}_u &= \mu^{AB} R_{uB} && \text{(D.2.7)}\\
R^u{}_u &= R_{ur} && \text{(D.2.8)}\\
R^r{}_u &= R_{uu} && \text{(D.2.9)}\\
R^r{}_A &= R_{uA} && \text{(D.2.10)}\\
R^A{}_B &= \mu^{AC} R_{CB} && \text{(D.2.11)}\\
& && \text{(D.2.12)}
\end{aligned}
$$

Furthermore, we have.

$$R_{uu} = -\frac{1}{2}\mu^{AB}\partial_u^2\mu_{AB} - \frac{1}{4}(\partial_u\mu^{AB})\partial_u\mu_{AB} + \mathcal{O}(r) \tag{D.2.13}$$

$$R_{uA} = \frac{1}{2}\partial_u\beta_A + \frac{1}{4}\beta_A\partial_u\mu_{BC} - \hat{D}_{[A}(\mu^{BC}\partial_u\mu_{B]C}) + \mathcal{O}(r^2) \tag{D.2.14}$$

The result for R_{uu} comes from appendix D.4, and the result for R_{uA} is taken from [16].

D.3. $(\partial/\partial u)^a$ is Hypersurface Orthogonal on E

In the following we will show that the vector field $n^a = (\partial/\partial u)^a$ is hypersurface orthogonal at $r = 0$, i.e. on the event horizon E. In section 3.3 we have seen that hypersurface orthogonality is equivalent to the condition $\omega_{ab} = 0$. We have

$$\omega_{ab} = \hat{B}_{[ab]} = B_{[ab]} - n_{[a}l^c B_{|c|b]} - n_{[b}l^c B_{a]c} + n_{[a}n_{b]}B_{cd}l^c l^d \tag{D.3.1}$$

The last term in (D.3.1) clearly vanishes since we have $n_{[a}n_{b]} = 0$. First of all, we will show that we have $B_{[ab]} = 0$ at $r = 0$:

$$
\begin{aligned}
B_{[ab]} &= \nabla_{[b}g_{a]c}\left(\frac{\partial}{\partial u}\right)^c\\
&= \nabla_{[b}\Big\{(\mathrm{d}u)_{a]}(\mathrm{d}r)_c + (\mathrm{d}r)_{a]}(\mathrm{d}u)_c - 2r^2\alpha(\mathrm{d}u)_{a]}(\mathrm{d}u)_c\\
&\qquad - r\beta_A(\mathrm{d}u)_{a]}(\mathrm{d}x^A)_c - r\beta_A(\mathrm{d}x^A)_{a]}(\mathrm{d}u)_c + \mu_{AB}\mathrm{d}x^A)_{a]}(\mathrm{d}x^B)_c\Big\}\left(\frac{\partial}{\partial u}\right)^c\\
&= \nabla_{[b}\Big\{(\mathrm{d}r)_{a]} - 2r^2\alpha(\mathrm{d}u)_{a]} - r\beta_A(\mathrm{d}x^A)_{a]}\Big\}\\
&= \underbrace{\nabla_{[b}\nabla_{a]}r}_{=0} - 2\Big\{2r\underbrace{(\nabla_{[b}r)}_{=0}\alpha(\mathrm{d}u)_{a]} + r(\nabla_{[b}\alpha)(\mathrm{d}u)_{a]} + r\alpha\underbrace{\nabla_{[b}\nabla_{a]}u}_{=0}\Big\}\\
&\qquad - \Big\{\underbrace{(\nabla_{[b}r)}_{=0}\beta_A(\mathrm{d}x^A)_{a]} + r(\nabla_{[b}\beta_A)(\mathrm{d}x^A)_{a]} + r\beta_A\underbrace{\nabla_{[b}\nabla_{a]}x^A}_{=0}\Big\}\\
&= -r\Big\{2(\nabla_{[b}\alpha)(\mathrm{d}u)_{a]} + (\nabla_{[b}\beta_A)(\mathrm{d}x^A)_{a]}\Big\}\\
&= 0.
\end{aligned} \tag{D.3.2}
$$

Here we have used the torsion freeness of the connection, i.e. $\nabla_a \nabla_b f = \nabla_b \nabla_b f$ for all $f \in C^\infty(M)$, and the fact that $\nabla_a r = 0$ since $r = 0 = const$ on E. The second term in (D.3.1) may be rewritten as

$$\begin{aligned} n_{[a} l^c B_{|c|b]} + n_{[b} l^c B_{a]c} &= \frac{1}{2}\Big\{ n_a l^c B_{cb} - n_b l^c B_{ca} + n_b l^c B_{ac} - n_a l^c B_{bc} \Big\} \\ &= \frac{1}{2}\Big\{ n_a l^c (B_{cb} - B_{bc}) - n_b l^c (B_{ca} - B_{ac}) \Big\} \\ &= n_a l^c B_{[cb]} - n_b l^c B_{[ca]}. \end{aligned} \tag{D.3.3}$$

Since we have $B_{[ab]} = 0$, it follows that the twist ω_{ab} of the congruence, defined by the vector field $n^a = (\partial/\partial u)^a$, vanishes on the horizon E. Therefore, n^a is hypersurface orthogonal on E.

D.4. Connection Between Sections 3.2 and 3.4

In the following we will show, that if one rewrites the vacuum Einstein equation $R_{ab} = 0$ in GNC and restricts them to the event horizon , then uu-component of the resulting equation corresponds to the Raychaudhuri equation. This result will place an additional condition on the function α, appearing in the metric.

The uu-component of the vacuum Einstein equation is calculated as follows: We will use the standard representation of the Ricci tensor in a coordinate system $\{x^\alpha, \alpha = 0, \ldots, 3\}$

$$R_{\mu\rho} = \partial_\nu \Gamma^\nu_{\mu\rho} - \partial_\mu \Gamma^\nu_{\nu\rho} + \Gamma^\alpha_{\mu\rho}\Gamma^\nu_{\alpha\nu} + \Gamma^\alpha_{\nu\rho}\Gamma^\nu_{\alpha\mu}. \tag{D.4.1}$$

We will use the Gaussian null coordinate system $\{u, r, x^A\}$ from section 3.2. By writing out all the internal summations in the above equation, setting $\mu = \rho = u$, inserting the Christoffel-symbols[1] and ommiting those (except for the ones which appear under an ∂_r derivative) which vanish on the horizon ($r = 0$) we obtain[2]

$$R_{uu}|_{r=0} = R_{ab} n^a n^b|_{r=0} = -\frac{1}{2}\mathcal{L}_n(\mu^{ab}\mathcal{L}_n \mu_{ab}) - \frac{1}{4}\mu^{ac}\mu^{bd}(\mathcal{L}_n\mu_{ab})\mathcal{L}_n\mu_{cd} + \frac{1}{2}\alpha\mu^{ab}\mathcal{L}_n\mu_{ab} = 0. \tag{D.4.2}$$

Again, we have used the replacements $\partial_u \to \mathcal{L}_n$, $\mu_{AB} \to \mu_{ab}$ and $\mu^{AB} \to \mu^{ab}$ in order to make the expression covariant (see section 8.2.4).

On E the shear is equal to the trace free part of $\mathcal{L}_n \mu_{ab}$ while the expansion ϑ is equal to the trace of this quantity. Together with equation (3.3.25) follows that we have the decomposition

$$\mathcal{L}_n \mu_{ab} = 2\sigma_{ab} + \vartheta \mu_{ab}. \tag{D.4.3}$$

For the inverse metric μ^{ab} we will use the decomposition $\mathcal{L}_n \mu^{ab} = g^{ac} g^{bd} \mathcal{L}_n \mu_{cd} = 2\sigma^{ab} + \vartheta \mu^{ab}$, since the vector field n^a is Killing, according to the result of the Rigidity theorem (see section 4.4). Furthermore, the vector field n^a is hypersurface orthogonal on E (see appendix D.3), so we will have $\omega_{ab} = 0$ in the following. From the decomposition (D.4.3), we can derive the

[1] For this calculation we have used the Christoffelsymbols from section D.1 with $\bar{\alpha} = -2r\alpha$.
[2] This result differs from the the one obtained in [16]. However, the consistency check that the $R_{uu}|_{r=0}$ yields the Raychaudhuri equation shows that our result is correct.

following identities

$$\frac{1}{2}\mu^{ab}\mathcal{L}_n\mu_{ab} = \frac{1}{2}\mu^{ab}(2\sigma_{ab} + \vartheta\mu_{ab}) = \vartheta, \tag{D.4.4}$$

where we have used $\mu^{ab}\mu_{ab} = 2$ and $\mu^{ab}\sigma_{ab} = (g^{ab} - n^a l^b - l^a n^b)\sigma_{ab} = 0$. Furthermore we have

$$\begin{aligned} \mu^{ac}\mathcal{L}_n\mu_{cb} &= 2\mu^{ac}\sigma_{cb} + \vartheta\mu^{ac}\mu_{cb} \\ &= 2\sigma^a{}_b + \vartheta(\delta^a{}_b - n^a l_b - l^a n_b), \end{aligned} \tag{D.4.5}$$

from which

$$\begin{aligned} \mu^{ac}\mu^{bd}(\mathcal{L}_n\mu_{ab})\mathcal{L}_n\mu_{cd} &= \mu^{ca}(\mathcal{L}_n\mu_{ab})\mu^{bd}\mathcal{L}_n\mu_{dc} \\ &= \left[2\sigma^c{}_b + \vartheta(\delta^c{}_b - n^c l_b - l^c n_b)\right]\left[2\sigma^b{}_c + \vartheta(\delta^b{}_c - n^b l_c - l^b n_c)\right] \\ &= 4\sigma_{ab}\sigma^{ab} + \frac{\vartheta^2}{4}(\delta^c{}_b - n^c l_b - l^c n_b)(\delta^b{}_c - n^b l_c - l^b n_c) \\ &= 4\sigma_{ab}\sigma^{ab} + \frac{\vartheta^2}{4}(4 - 1 - 1 - 1 + 1 + 0 - 1 + 0 + 1) \\ &= 4\sigma_{ab}\sigma^{ab} + 2\vartheta^2. \end{aligned} \tag{D.4.6}$$

follows. Insertion of these results into (D.4.2) yields

$$\begin{aligned} R_{ab}n^a n^b|_{r=0} &= -\mathcal{L}_n\vartheta - \frac{1}{4}(4\sigma_{ab}\sigma^{ab} + 2\vartheta^2) + \alpha\vartheta \\ &= -\frac{\mathrm{d}\vartheta}{\mathrm{d}u} - \sigma_{ab}\sigma^{ab} - \frac{1}{2}\vartheta^2 + \alpha\vartheta \end{aligned} \tag{D.4.7}$$

As we see, this is the Raychaudhuri equation

$$\frac{\mathrm{d}\vartheta}{\mathrm{d}u} = -\frac{1}{2}\vartheta^2 - \sigma_{ab}\sigma^{ab} - R_{ab}n^a n^b \tag{D.4.8}$$

for a hypersurface orthogonal congruence defined by the affinely parametrized vector field n^a, up the additional factor which involves α. Since the Raychaudhuri equation is a general identity between geometric objects, and not particular to any field equations, $R_{uu}|_{r=0}$ must yield the Raychaudhuri equation only. Therefore, we obtain the additional condition on the function α that it must vanish on the horizon. This justifies the replacement $\alpha \to r\alpha$ which we made at the end of section 3.2.

D.5. Pullback of $\epsilon_{abcd}X^{cd}$ to $\mathcal{E}$

In the following we will compute the quantity

$$\psi^*\epsilon_{abcd}X^{cd} = \overset{(2)}{\epsilon}_{ab}X^{cd}(\mathrm{d}u)_c(\mathrm{d}r)_d, \tag{D.5.1}$$

where $\psi : \mathcal{E} \to M$ is an embedding. In section 8.2.3 we found

$$X^{ab} = R^{ac}\nabla_c n^b - R^{bc}\nabla_c n^a + (\nabla^b R^{ac})n_c - (\nabla^a R^{bc})n_c + (\nabla_c R^{cb})n^a - (\nabla_c R^{ca})n^b. \tag{D.5.2}$$

From this we find

$$\begin{aligned} X^{ab}(\mathrm{d}u)_a(\mathrm{d}r)_b =& R^{ac}[\nabla_c n^b](\mathrm{d}u)_a(\mathrm{d}r)_b - R^{bc}[\nabla_c n^a](\mathrm{d}u)_a(\mathrm{d}r)_b \\ &+ [\nabla^b R^{ac}]n_c(\mathrm{d}u)_a(\mathrm{d}r)_b - [\nabla^a R^{bc}]n_c(\mathrm{d}u)_a(\mathrm{d}r)_b \\ &+ [\nabla_c R^{cb}]n^a(\mathrm{d}u)_a(\mathrm{d}r)_b. \end{aligned} \tag{D.5.3}$$

The last term from (D.5.2) does not appear, since we have $n^a(\mathrm{d}r)_a = (\partial/\partial u)^a(\mathrm{d}r)_a = 0$.

Before we calculate each term in (D.5.3), let us collect some formulas which will be needed in what follows. From the form of metric g_{ab} and its inverse g^{ab} in GNC (see equations (3.2.9) and (3.2.10)), we find (when restricted to the horizon $r = 0$)

$$g_{ab}(\partial_u)^a = (\mathrm{d}r)_b \tag{D.5.4}$$

$$g^{ab}(\mathrm{d}u)_a = (\partial_r)^b \tag{D.5.5}$$

$$g_{ab}(\partial_r)^a = (\mathrm{d}u)_b \tag{D.5.6}$$

$$g^{ab}(\mathrm{d}r)_a = (\partial_u)^b. \tag{D.5.7}$$

We remind the reader that we have $n^a = (\partial/\partial u)^a$.

For the first term in (D.5.3), we find

$$\begin{aligned} R^{ac}[\nabla_c n^b](\mathrm{d}u)_a(\mathrm{d}r)_b &= R^{ac}\left[\partial_c(\partial_u)^b + \Gamma^b_{cd}(\partial_u)^d\right](\mathrm{d}u)_a(\mathrm{d}r)_b \\ &= R^{ac}\Gamma^r_{cu}(\mathrm{d}u)_a \\ &= \left[R^{au}\Gamma^r_{uu} + R^{ar}\Gamma^r_{ru} + R^{au}\Gamma^A_{Au}\right](\mathrm{d}u)_a \\ &= 0, \end{aligned} \tag{D.5.8}$$

For the first equality we used the standard formular for the covariant derivative acting on a vector field. For the second equality we used the fact that we have $\partial_a(\partial_\mu)^b$ for any coordinate vector field $(\partial_\mu)^a = (\partial/\partial x^\mu)^a$. For the third equality we used $\Gamma^r_{uu} = \Gamma^r_{ru} = \Gamma^r_{Au} = 0$ (at $r = 0$).

For the second term in (D.5.3), we find

$$\begin{aligned} R^{bc}[\nabla_c n^a](\mathrm{d}u)_a(\mathrm{d}r)_b &= R^{bc}\left[\partial_c(\partial_u)^a + \Gamma^a_{cd}(\partial_u)^d\right](\mathrm{d}u)_a(\mathrm{d}r)_b \\ &= R^{bc}\Gamma^u_{cu}(\mathrm{d}r)_b \\ &= R^{bA}\Gamma^u_{Au}(\mathrm{d}r)_b \\ &= R_d{}^A\Gamma^u_{Au}g^{db}(\mathrm{d}r)_b \\ &= R_d{}^A\Gamma^u_{Au}(\partial_u)^d \\ &= R_u{}^A\Gamma^u_{Au} \\ &= \frac{1}{2}R_u{}^A\beta_A \\ &= \frac{1}{2}\beta^A R_{uA} \end{aligned} \tag{D.5.9}$$

For the third equality we used $\Gamma^u_{uu} = \Gamma^u_{ru} = 0$ (at $r = 0$). For the fifth equality we used equation (D.5.7). For the sixth equality we used the explicit form of the Christoffel symbols from appendix D.1. For the last equality we used the results for the Ricci tensor from appendix D.2.

For the third term in (D.5.3), we find

$$\begin{aligned}
[\nabla^b R^{ac}]n_c(\mathrm{d}u)_a(\mathrm{d}r)_b &= [\nabla_b R_{ac}](\partial_u)^c(\partial_r)^a(\partial_u)^b \\
&= [(\partial_u)^b\nabla_b R_{ac}](\partial_u)^c(\partial_r)^a \\
&= (\partial_u)^b\nabla_b[R_{ac}(\partial_u)^c(\partial_r)^a] - R_{ac}[(\partial_u)^b\nabla_b(\partial_u)^c](\partial_r)^a \\
&\quad - R_{ac}[(\partial_u)^b\nabla_b(\partial_r)^a](\partial_u)^c \\
&= \partial_u R_{ru} - R_{rc}\Gamma^c_{uu} - R_{au}\Gamma^a_{ur} \\
&= \partial_u R_{ur} - R_{Au}\Gamma^A_{ur} \\
&= \partial_u R_{ur} - \beta^A R_{uA}.
\end{aligned} \tag{D.5.10}$$

For this term we used the same techniques (formulas etc.) as for the second term, and we will not go through each line explicitly.

For the fourth term in (D.5.3), we find

$$[\nabla^a R^{bc}]n_c(\mathrm{d}u)_a(\mathrm{d}r)_b = \partial_r R_{uu} + \beta^A R_{uA}. \tag{D.5.11}$$

The computation is analogous to (D.5.10).

For the fifth term in (D.5.3), we find

$$\begin{aligned}
[\nabla_c R^{cb}]n^a(\mathrm{d}u)_a(\mathrm{d}r)_b &= [\nabla_c R^{cb}](\mathrm{d}r)_b \\
&= [\nabla_c R^c{}_b](\partial_u)^b \\
&= \nabla_c[R^c{}_b(\partial_u)^b] - R^c{}_b\nabla_c(\partial_u)^b \\
&= \nabla_c R^c{}_u - R^c{}_b\Gamma^b_{cu} \\
&= \partial_c R^c{}_u + \Gamma^c_{cd}R^d{}_u - R^c{}_b\Gamma^b_{cu}
\end{aligned} \tag{D.5.12}$$

We have

$$\partial_c R^c{}_u = \partial_u R_{ur} + \partial_r R_{uu} + \partial_A R^A{}_u \tag{D.5.13}$$

$$\Gamma^c_{cd}R^d{}_u = \frac{1}{2}\mu^{AB}(\partial_u\mu_{AB})R_{ur} + \frac{1}{2}\mu^{AB}(\partial_r\mu_{AB})R_{uu} + \hat{\Gamma}^A_{AB}R^B{}_u \tag{D.5.14}$$

$$R^c{}_b\Gamma^b_{cu} = -\frac{1}{2}(\partial_u\mu^{AB})R_{AB}. \tag{D.5.15}$$

Insertion of these results into (D.5.12) yields

$$\begin{aligned}
[\nabla_c R^{cb}]n^a(\mathrm{d}u)_a(\mathrm{d}r)_b &= \partial_u R_{ur} + \partial_r R_{uu} + \mu^{AB}\hat{D}_A R_{Bu} + \frac{1}{2}\mu^{AB}(\partial_u\mu_{AB})R_{ur} \\
&\quad + \frac{1}{2}\mu^{AB}(\partial_r\mu_{AB})R_{uu} + \frac{1}{2}(\partial_u\mu^{AB})R_{AB}.
\end{aligned} \tag{D.5.16}$$

By combining all of these results we obtain

$$\begin{aligned} X^{ab}(\mathrm{d}u)_a(\mathrm{d}r)_b = 2\partial_u R_{ur} + \mu^{AB}\hat{D}_A R_{uB} - \beta^A R_{uA} + \frac{1}{2}\mu^{AB}(\partial_u \mu_{AB})R_{ur} \\ + \frac{1}{2}\mu^{AB}(\partial_r \mu_{AB})R_{uu} + \frac{1}{2}(\partial_u \mu^{AB})R_{AB}. \end{aligned} \tag{D.5.17}$$

This expression can be made covariant by making, again, the replacements

$$\mu_{AB} \to \mu_{AB}(\mathrm{d}x^A)_a(\mathrm{d}x^B)_b =: \mu_{ab} \tag{D.5.18}$$

$$\mu^{AB} \to \mu^{AB}(\partial_A)^a(\partial_B)^b =: \mu^{ab} \tag{D.5.19}$$

$$\partial_u \to \mathcal{L}_n \tag{D.5.20}$$

$$\partial_r \to \mathcal{L}_r \tag{D.5.21}$$

$$R_{AB} \to R_{AB}(\mathrm{d}x^A)_a(\mathrm{d}x^B)_b =: R_{ab}. \tag{D.5.22}$$

As we explained below equation (8.2.80), this procedure is consistent. In addition, we will make the replacements

$$\hat{D}_A \to (\mathrm{d}x^A)_a \hat{D}_A =: \hat{D}_a \tag{D.5.23}$$

$$\beta^A \to (\partial_A)^a \beta^A =: \beta^a. \tag{D.5.24}$$

Furthermore, from equation (8.2.84) we have

$$R_{ur} = \frac{1}{2}(R - R_{ab}\mu^{ab}). \tag{D.5.25}$$

Putting all this together we find

$$\begin{aligned} \int_{\mathcal{E}} \epsilon_{abcd}X^{cd} = \int_{\mathcal{E}} \Bigg\{ \mathcal{L}_n R - \mathcal{L}_n(R_{cd}\mu^{cd}) + \mu^{cd}\hat{D}_c(n^e R_{ed}) - n^c\beta^d R_{cd} \\ + \frac{1}{2}\left[\vartheta R + 2R_{cd}\sigma^{cd} + \mu^{cd}(\mathcal{L}_l \mu_{cd})R_{ef}n^e n^f\right]\Bigg\} \overset{(2)}{\epsilon}_{ab}. \end{aligned} \tag{D.5.26}$$

Bibliography

[1] R. Arnowitt, S. Deser, and C.W. Misner, *The dynamics of general relativity*, Gravitation: An Introduction to Current Research (L. Witten, ed.), Wiley, New York, U.S.A., 1962, pp. 227–265.

[2] A. Ashtekar and R. O. Hansen, *A unified treatment of null and spatial infinity in general relativity. i. universal structure, asymptotic symmetries, and conserved quantities at spatial infinity*, Journal of Mathematical Physics **19** (1978), 1542–1566.

[3] J. M. Bardeen, B. Carter, and S. W. Hawking, *The four laws of black hole mechanics*, Commun. Math. Phys. **31** (1973), 161–170.

[4] J. K. Beem, P. E. Ehrlich, and K. L. Easley, *Global lorentzian geometry*, Marcel Dekker Ltd, 1996.

[5] J. D. Bekenstein, *Black holes and entropy*, Phys. Rev. D **7** (1973), 2333–2346.

[6] ———, *Generalized second law of thermodynamics in black-hole physics*, Phys. Rev. D **9** (1974), 3292–3300.

[7] P. T. Chruściel, E. Delay, G. J. Galloway, and R. Howard, *Regularity of horizons and the area theorem*, Ann. Inst. Henri Poincare **2** (2001), 109–178.

[8] S. Deser, H. S. Tsao, and P. van Nieuwenhuizen, *One-loop divergences of the einstein-yang-mills system*, Phys. Rev. D **10** (1974), 3337–3342.

[9] H. Friedrich, I. Racz, and R. M. Wald, *On the Rigidity Theorem for Spacetimes with a Stationary Event Horizon or a Compact Cauchy Horizon*, Commun. Math. Phys. **204** (1999), 691–707.

[10] R. Geroch, *Structure of the gravitational field at spatial infinity*, Journal of Mathematical Physics **13** (1972), 956–968.

[11] M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory*, Cambridge, UK: Univ. Pr. (1987), Cambridge Monographs On Mathematical Physics.

[12] S. W. Hawking, *Gravitational radiation from colliding black holes*, Phys. Rev. Lett. **26** (1971), 1344–1346.

[13] ———, *Black holes in general relativity*, Commun. Math. Phys. **25** (1972), 152–166.

[14] ———, *Particle creation by black holes*, Commun. Math. Phys. **43** (1975), 199–220.

[15] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time*, Cambridge Universtity Press, 1973.

[16] S. Hollands, A. Ishibashi, and R. M. Wald, *A higher dimensional stationary rotating black hole must be axisymmetric*, Commun. Math. Phys. **271** (2007), 699–722.

[17] G. T. Horowitz and R. M. Wald, *Dynamics of einstein's equation modified by a higher-order derivative term*, Phys. Rev. D **17** (1978), 414–416.

[18] W. Israel, *Third law of black-hole dynamics: Formulation and proof*, Phys. Rev. Lett. **57** (1986), 397–399.

[19] T. Jacobson, G. Kang, and R. C. Myers, *On black hole entropy*, Phys. Rev. D **49** (1994), 6587–6598.

[20] ______, *Increase of black hole entropy in higher curvature gravity*, Phys. Rev. D **52** (1995), 3518–3528.

[21] D. R. Noakes, *The initial value formulation of higher derivative gravity*, J. Math. Phys. **24** (1983), 1846–1850.

[22] B. Paczyński, *Mass of cygnus x-1*, Astronomy and Astrophysics **34** (1974), 161–162.

[23] R. Penrose, *Asymptotic properties of fields and space-times*, Phys. Rev. Lett. **10** (1963), 66–68.

[24] ______, *Zero rest-mass fields and space-times*, Proc.Roy.Soc.Lond. **A284** (1965), 159–203.

[25] W. L. Sargent, P. J. Young, A. Boksenberg, K. Shortridge, C.R. Lynds, and F. D. A. Hartwick, *Dynamical evidence for a central mass concentration in the galaxy m87*, Astrophysical Journal **221** (1978), 731–744.

[26] K. Stelle, *Renormalization of higher derivative quantum gravity*, Phys. Rev. D **16** (1976), 953–969.

[27] ______, *Classical gravity with higher derivatives*, General Relativity and Gravitation **9** (1977), 353–371.

[28] G. 't Hooft and M. Veltman, *One-loop divergencies in the theory of gravitation*, Ann. Inst. Henri Poincaré **20** (1974), 69–94.

[29] R. M. Wald, *General relativity*, The University of Chicago Press, 1984.

[30] ______, *On identically closed forms locally constructed from a field*, J. Math.Phys. **31** (1990), 2378–2384.

[31] R. M. Wald and P. T. Chruściel, *Maximal hypersurfaces in stationary asymptotically flat spacetimes*, Commun. Math. Phys. **163** (1994), 561–604.

[32] R. M. Wald and V. Iyer, *Some properties of noether charge and a proposal for dynamical black hole entropy*, Phys. Rev. D **50** (1994), 846–864.

[33] R. M. Wald and B. S. Kay, *Theorems on the uniqueness and thermal properties of stationary, nonsingular, quasifree states on spacetimes with a bifurcate killing horizon*, Phys. Rep. **207** (1991), 49–136.

[34] R. M. Wald and J. Lee, *Local symmetries and contraints*, J. Math.Phys. **31** (1990), 725–743.

[35] R. M. Wald and I. Racz, *Extensions of spacetimes with killing horizons*, Class. Quantum Grav. **9** (1992), 2643–2656.

[36] ———, *Global extensions of spacetimes describing asymptotic final states of black holes*, Class. Quantum Grav. **13** (1996), 539–552.

[37] R. M. Wald and D. Sudarsky, *Extrema of mass, stationarity and staticity, and solutions to the einstein–yang–mills equations*, Phys. Rev. D **46** (1992), 1453–1474.

[38] R. M. Wald and A. Zoupas, *A General Definition of "Conserved Quantities" in General Relativity and Other Theories of Gravity*, Phys. Rev. **D61** (2000), 084027.

[39] P. J. Young, J. A. Westphal, J. Kristian, C. P. Wilson, and F. P. Landauer, *Evidence for a supermassice object in the nucleus of the galaxy m87 from sit and ccd area photometry*, Astrophysical Journal **221** (1978), 721–730.

Acknowledgements

First of all, I would like to thank Prof. S. Hollands for giving me the opportunity to work in this interesting field of physics. His ideas and insistence on improvements taught me about physical intuition and constitute the basis of this thesis. I would also like to thank Prof. K.-H. Rehren for writing the additional report. I am grateful to Prof. H. Goenner for the numerous discussions about the topic of my thesis. He was a guide through the rough terrain of conventions in general relativity. Furthermore, I highly appreciate that Prof. F. Müller-Hoissen was willing to check some of my calculations.

During my time in the QFT group as a lonesome young relativist, many people helped me not to loose confidence. Especially, I would like to thank S. Jäger for the countless discussions and his critical mind. As a companion in misfortune, he kept me from panicing many times and without him this thesis would have probably not been completed. Furthermore, I would like to thank J. Slowik and C. Marx for creating such an enjoyable working atmosphere in our office. I am also grateful to the Diploma and PhD students of the QFT group for revitalizing the DiDo-Seminar. Their iniative created the necessary platform for the exchange of knowledge in a student-oriented manner.

I am deeply indepted to my mother, who always trusted in my decisions and supported me in every possible way.

Finally, the financial support from the Alexander von Humboldt-Stiftung is gratefully acknowledged.