Canonical Energy and Black Hole Stability

Enya Hsiao

August 2016

Abstract

Black hole stability is itself an interesting subject of both mathematical and physical interest. For the former, stability of solutions to Einstein equation is a legitimate PDE problem, whereas for the latter, it is crucial that solutions are stable under general perturbations to be considered physical. This paper investigates the stability of linear, axisymmetric perturbations to stationary, asymptotically flat black hole backgrounds using canonical energy.

Contents

1 Introduction 1
2 Lorenztian Geometry 3
3 Black hole spacetime 6
4 Killing horizon 10
5 Canonical Energy 12
  5.1 Gauge Choices at the Horizon 12
  5.2 Canonical Energy 14
  5.3 Canonical Energy Flux 18
  5.4 Hollands-Wald Stability Criterion 19

1 Introduction


For black hole analysis in general relativity, we consider a $D$-dimensional Lorentzian manifold $(\mathcal{M}, g)$ called spacetime, such that the metric $g$ is a solution to the vacuum Einstein equation,

$$\nabla^2 g = 0,$$
where \(Ric\) is the Ricci curvature tensor, and such that \(\mathcal{M}\) admits a trapped surface called the event horizon.

The Einstein equation is a system of coupled nonlinear partial differential equations that are nontrivial to solve. In this paper, we do not attempt to describe methods to obtain a solution. Instead, given a stationary solution \(g\) written as \(g_{ab}\) in coordinates, we investigate its stability properties under linear axisymmetric perturbations. Analysis for the linear case is usually the first step before generalizing to nonlinear analysis.

Explicitly, linear perturbations of \(g\) are obtained as follows. Let \(\mathcal{G}\) denote the space of metrics, then given a background spacetime \(g_{ab}\), an element \(\gamma_{ab}\) in the tangent space \(T_g \mathcal{G}\) that satisfies the linear Einstein equation is a linear perturbation. That is, for each one-parameter family of metrics \(g_{ab}(\lambda)\) with \(g_{ab}(0) = g_{ab}\), a perturbation is its derivative with respect to \(\lambda\) at \(\lambda = 0\).

**Definition 1.1.** A linear perturbation \(\gamma_{ab}\) is defined as,

\[
\gamma_{ab} \equiv \delta g_{ab} = \frac{d}{d\lambda} g_{ab}(\lambda) \bigg|_{\lambda=0}.
\]

Note. From this point on, we adopt the usual convention and use \(\delta g_{ab}\) to denote taking the derivative at \(\lambda = 0\), \(\frac{d}{d\lambda} g_{ab}(\lambda) \bigg|_{\lambda=0}\), and \(\frac{d}{d\lambda}\) for equations that hold for all \(\lambda\).

In addition, we require the perturbations to also satisfy the linearized Einstein equation, which can be derived from the nonlinear equation. Let \(g\) be a solution to

\[
Ric(g(\lambda)) = R_{ab}(g(\lambda)) = 0.
\]

Then we have

\[
\frac{d}{d\lambda} R_{ab}(g(\lambda)) \bigg|_{\lambda=0} = \mathcal{L}(\gamma) = 0,
\]

where \(\mathcal{L}\) is a linear operator. Explicitly calculated, we obtain the following:

**Proposition 1.2.** Under the transverse traceless gauge (a choice in coordinates that we are free to make) for \(\gamma\),

\[
\begin{align*}
\nabla^a \gamma_{ab} &= 0 \\
g^{ab} \gamma_{ab} &= 0.
\end{align*}
\]

The linear Einstein equation is

\[
\nabla^b \nabla_b \gamma_{ac} - 2 R^b_{\quad ac} d \gamma_{bd} = 0. \tag{1}
\]

Equation (1) is manifestly linear, whereby perturbations satisfying the above form a linear subspace\(^1\). For the purposes of Hollands-Wald, we require our linear perturbations to be also axisymmetric and to "fall off asymptotically" as the

\(^1\)Note that in this paper we do not assume the transverse traceless gauge, since it may not be compatible with our gauge conditions. It is used here simply to display linearity.
coordinates approach infinity. At this point, note that this set of perturbations also forms a linear subspace.

A black hole is stable if the following can be shown: (1) global boundedness of solution, and (2) sufficient decay in time (see Theorem 3.1 and 4.1 of [9]). Intuitively, this means that small perturbations in spacetime (e.g. a spaceship falling into the blackhole) eventually die down and the universe settles to the original blackhole or another stationary solution (e.g. with mass increased by that of a spaceship). However, this problem is, even for the simplest non-rotating, charge-free Schwarzschild solution, of which the linear case had just been solved [8] in early 2016, still very much an open.

In 2012, Hollands and Wald published a new criterion [1] for the stability of stationary, asymmetric black holes that did not limit its validity to any particular dimension or require explicit analysis of the background metric, as has been the major machinery so far for several fortunate cases. This turned out to be a remarkably simple energy argument: "positivity of canonical energy leads to dynamical stability", in which they defined a symmetric bilinear form called canonical energy $\mathcal{E}$, and showed that if $\mathcal{E}$ is semi-positive definite on the space of perturbations $\mathcal{T}$, we have global boundedness and decay and vice versa.

We begin by reviewing tools in Lorentzian geometry, then proceed to examine causality structures and Killing fields. We define asymptotic flatness, black hole regions and state a few crucial theorems regarding stationary axisymmetric blackhole solutions. Finally, we investigate canonical energy and the Hollands-Wald stability criteria. The aim is for an intuitive understanding of the ideas behind, hence we have omitted technicalities and refer the curious reader to the original work.

2 Lorenztian Geometry


We begin by a brief overview of geometry. Note that since calculations are carried out in index notation, we will express everything in coordinates, which may not be in the familiar form seen in mathematical text. See Wald [2] ch.3 for detailed review.

**Definition 2.1.** Given a Lorenztian manifold $\mathcal{M}$, the *metric*, $g$ is a symmetric $(0,2)$-tensor field (i.e. a smooth assignment such that for each $p \in \mathcal{M}$ there is a $(0,2)$-tensor) on $\mathcal{M}$ with signature $(- + \cdots +)$.

Written in local coordinates,

$$g = \sum_{\mu, \nu=0}^{D-1} g_{\mu\nu} dx^\mu \otimes dx^\nu.$$
We adopt the Einstein summation notation, that repeated indices are automatically summed. The inverse of \( g \) we denote with upper indices, \( g^{\mu\nu} \equiv (g_{\mu\nu})^{-1} \). Further, we shall refer to \( g_{\mu\nu} \) as the metric from this point on.

The standard \( D \)-dim Minkowski metric takes the form,
\[
\eta = -dt^2 + \sum_i dx_i^2
\]
where instead of \( dt \otimes dt \), we write squares as shorthand.

For each point \( p \in \mathcal{M} \), the map \( T_p M \to T^*_p M, v \mapsto g(v, -) \) is an isomorphism. Thus, we use \( g^{\mu \nu} \) or \( g_{\mu \nu} \) to raise and lower the indices of a tensor.

**Example 2.2.** Let \( T_{abc} \) be a (0,3)-tensor field. At \( p \in M \) we have that,
\[
T_{abc} : T_p M \times T_p M \times T_p M \to \mathbb{R}
\]
Then multiplying by \( g^{ad} \) raises the first index.
\[
g^{ad}T_{abc} = T^d_{bc} : T^*_p M \times T_p M \times T_p M \to \mathbb{R}.
\]

**Definition 2.3.** For a \((k,l)\)-tensor \( T \), the contraction of the \( i \)th, \( j \)th component is defined by identifying the \( i \)th input with the dual element of the \( j \)th input and summing over that repeated index, i.e.
\[
T^{a_1 \cdots a_{i-1} b_j a_{i+1} \cdots a_k}_{b_1 \cdots b_j \cdots b_l} : (\ldots, v_i-1, w_j^*, v_i+1, \ldots, w_j, \ldots) \mapsto \mathbb{R}.
\]
The output is a \((k - 1, l - 1)\) tensor,
\[
T^{a_1 \cdots a_{i-1} a_{i+1} \cdots a_k}_{b_1 \cdots b_j-1 b_j+1 \cdots b_l} : (\ldots, v_i-1, v_i+1, \ldots, w_j-1, w_j+1, \ldots) \mapsto \mathbb{R}.
\]

For a \((k,l)\)-tensor \( T \), we may symmetrize any \( n \leq k \) upper indices or \( m \leq l \) lower indices by defining a new tensor that is a sum of all permutations of those indices. Similarly, to antisymmetrize, we define a new tensor that is a sum of all even permutations minus the sum of all odd permutations of the specified indices. These new tensors are then divided by a normalization factor.

**Definition 2.4.**
\[
T_{a_1 \cdots (a_i \cdots a_j) \cdots a_k} \equiv \frac{1}{n!} \sum_{\sigma} T_{a_1 \cdots a_{\sigma(i)} \cdots a_{\sigma(j)} \cdots a_k}
\]

---

2 We have used both Greek and English letters for the index of tensors. Notation-wise, if we choose to work with our tensors in a particular set of basis, e.g. \( \{\beta_0, \beta_1, \beta_2, \beta_3\} \), we denote the indices of our tensor with Greek letters, e.g. \( T_{\mu_1 \cdots \mu_n} \). Meanwhile, we use English alphabets for the indices if we do not impose a specific basis. The indices are now markers for rank, e.g. \( T_{a_1 \cdots a_n} \). There is essentially no difference in calculations.
\[ T_{a_1 \cdots [a_i \cdots a_j] \cdots a_k} = \sum_{\sigma} \frac{(-1)^d}{n!} T_{a_1 \cdots a_{\sigma(i)} \cdots a_{\sigma(j)} \cdots a_k} \]

where \( \sigma \in S_n \), and \( d = +1 \) when \( \sigma \) is an odd permutation, and -1 otherwise. The notation for general \((k,l)\)-tensors follows analogously.

Next we define quantities that are relevant to the geometry of spacetime.

**Definition 2.5.** The Christoffel symbol, \( \Gamma^i_{jk} \), characterizes the failure of partial derivatives to be covariant,

\[ \Gamma^i_{jk} = \frac{1}{2} g^{il} (\partial_k g_{lj} + \partial_j g_{lk} - \partial_l g_{jk}) \]

where \( \partial_i = \partial/\partial x^i \).

**Definition 2.6.** The covariant derivative associated to \( g_{ab} \), is a unique graded derivation \( \nabla_a \) of degree +1 on tensor fields, such that

\[ \nabla_a g_{bc} = 0. \]

Written in local coordinates for vector field \( V^a \), this is

\[ \nabla_a V^a = \partial_a V^a + \Gamma^b_{ac} V^c. \]

*Note.* \( \Gamma^a_{bc} \) is not tensorial. This make sense since \( \Gamma^a_{bc} \) appends a non-tensorial term \( \partial_c \), to create a tensorial covariant derivative \( \nabla_a \).

Given the covariant derivative, we can define geodesics of \( M \). Let \( \lambda \) be a curve in \( M \). Then \( \lambda \) is a geodesic if for each point \( p \in \lambda \), the vector at \( p \) tangent to \( \lambda \), \( \gamma^a \), satisfies:

\[ \gamma^a \nabla_a \gamma^b = 0. \]

Geodesics can be intuitively thought of as the straightest possible path or the path taken by free falling particles, particles that are subject to zero net force.

**Definition 2.7.** The Riemann curvature tensor \( R^a_{bcd} \) defined in coordinates is

\[ R^a_{bcd} = \partial_c \Gamma^a_{db} - \partial_d \Gamma^a_{cb} + \Gamma^a_{ce} \Gamma^e_{db} - \Gamma^a_{de} \Gamma^e_{cb}. \]

**Definition 2.8.** The Ricci curvature tensor \( R_{bd} \) is the contraction of the Riemann curvature tensor,

\[ R_{bd} \equiv g^{ac} R_{abcd} = R^a_{bad}. \]

**Definition 2.9.** The Ricci scalar \( R \) is a further contraction of the Ricci curvature,

\[ R \equiv g^{bd} R_{bd} = R^b_b. \]

\[ \]
3 Black hole spacetime


Consider the Minkowski spacetime \((M, \eta_{ab})\), which is flat (i.e. \(R^{abcd} = 0\)). For any \(p \in M\), note that there exists \(v \in T_pM\), \(v \neq 0\), e.g. \((1,1,0,0)\), such that the metric is degenerate, \(\eta(v,v) = 0\). In fact, we may split the vectors in \(T_pM\) into three types.

\[
\begin{align*}
\eta(v,v) > 0 & \quad \text{spacelike} \\
\eta(v,v) = 0 & \quad \text{null} \\
\eta(v,v) < 0 & \quad \text{timelike}
\end{align*}
\]

The set of null vectors form a \((D-1)\)-dim surface called the light cone. Further, depending on the sign of time coordinates, we split the light cone into the future light cone and past light cone, where timelike and null vectors with positive time coordinate are called future directed, and lay within the future light cone.

![Figure 1: Light cone. Reference:[13].](image)

One may have suspected at this point, that the light cone gives the maximal set that \(p\) can influence and depend on. Indeed, the Einstein equations are shown to be hyperbolic, owning a finite speed of propagation. The causal structure given by the light cone categorizes pairs of points (events) in \(M\), where events related by spacelike vectors not reachable within the speed of light, and events related by timelike or null vectors can be causally affected.

It is worthy to note that the causal structure of Minkowski spacetime is preserved under action of the Poincare group \(\mathbb{R}^4 \rtimes SO(3,1)\). That is, timelike, spacelike and null vectors remain timelike, spacelike and null under translation (action of \(\mathbb{R}^4\)), rotation and boosts (action of \(SO(3,1)\)). These transformations comprise the symmetries, specifically, Killing fields (see section 4) of Minkowski spacetime, and shall be useful later on to analyze the infinity of spacetimes that...
are asymptotically flat.

Generalizing to curved space, where \( \mathcal{M} \) is no longer flat, we have:

**Definition 3.1.** A curve \( \lambda \subset \mathcal{M} \) is timelike, spacelike or null if for each point \( p \in \lambda \), the vector tangent to the curve \( v^a \), is timelike, spacelike or null. Further, a vector field \( v^a \) is timelike, spacelike or null if for \( p \in \mathcal{M} \), \( v|_p \) is timelike, spacelike or null.

In the case of curved space, the situation becomes tricky since there may be pathologies such as closed timelike curves or the tipping of the light cone. We therefore require \( \mathcal{M} \) be time-orientable \(^4\) and stably causal \(^5\) such that we may divide tangent vectors of \( p \in \mathcal{M} \) into timelike, spacelike and null vectors unambiguously with respect to \( g_{ab} \).

**Definition 3.2.** A curve \( \lambda \subset \mathcal{M} \) is causal if \( v^a \) is timelike or null at each point. Further, \( \lambda \) is inextendible if it does not have a future/past end point, i.e. for every \( p \in \mathcal{M} \), \( \lambda \) does not stay in every open neighborhood \( \mathcal{O} \) of \( p \) for all time in the future/past.

**Definition 3.3.** The causal future/past of the set \( S \subset \mathcal{M} \), \( J^+(S)/J^-(S) \) is the set of \( p \in \mathcal{M} \) such that there exists a \( q \in S \) with a future/past directed causal curve from \( q \) to \( p \). The timelike future/past \( I^+(S)/I^-(S) \), is defined analogously for timelike curves.

**Definition 3.4.** Let \( D^+(S) \) be the set of elements \( p \in \mathcal{M} \) such that every past inextendible causal curve through \( p \) intersects \( S \). Similarly, \( D^-(S) \) is defined with past replaced by future.

The *domain of dependence* of \( S \in \mathcal{M} \) is

\[
D(S) = D^+(S) \cup D^-(S).
\]

Next, we have the notion of a set being spacelike or null such that no two points are timelike related. A set \( S \subset \mathcal{M} \) is *achronal* if there do not exist \( q, r \in S \) such that \( r \in I^+(q) \).

**Definition 3.5.** A *Cauchy surface* is a closed achronal hypersurface \( \Sigma \) such that its domain of dependence \( D(\Sigma) = \mathcal{M} \).

Therefore, given a suitable \(^6\) set of *initial data* \( (\Sigma, h_{ab}, K_{ab}) \), consisting of, respectively, a smooth hypersurface surface \( \Sigma \), a spatial metric \( h_{ab} \) on \( \Sigma \) and a

---

\(^4\)A *time-orientation* on \( (\mathcal{M}, g) \) is defined by an equivalence class \([K]\) of timelike vector fields, such that \( K_1 \sim K_2 \) if \( g(K_1, K_2) < 0 \).

\(^5\)Stable causality prevents \( g_{ab} \) from the verge of creating closed timelike curves(see Wald[2] ch.8).

\(^6\)The initial data should satisfy the evolution constraints given by the Einstein equation. Let \( D_a \), \( (D^{-1})R \) be the spatial derivative operator and Ricci scalar associated to \( h_{ab} \) and \( n^a \) be the unit timelike normal at each \( p \in \Sigma \). Then these constraints are,

\[
D_a K_b^a - D_b K_a^b = 0
\]

\[
\frac{1}{2} \left\{ (D^{-1})R + (K_a^a)^2 - K_{ab} K^{ab} \right\} = 0.
\]
symmetric 2-tensor field $K_{ab}$, we should be able to obtain all information about $M$. One may ask, given a set of initial data, is the Einstein equation well-posed in the sense that the solution is unique and depends continuously on initial data. The answer is affirmative. In fact, we have the following theorem:

**Theorem 3.6.** Given a set of initial data $(\Sigma, h_{ab}, K_{ab})$ satisfying the constraint equations given in the footnotes, there is a unique spacetime $(M, g_{ab})$, called the maximal Cauchy development, such that $(M, g_{ab})$ satisfies the Einstein equations and $\Sigma, h_{ab}, K_{ab}$ are a Cauchy surface, the induced spatial metric by $g_{ab}$ and the extrinsic curvature respectively. $M$ is maximal in the sense that every other spacetime $(M', g'_{ab})$ satisfying the above properties can be mapped isometrically into $(M, g_{ab})$. Further, $g_{ab}$ on $M$ depends continuously on initial data.

Now when we speak of a physical spacetime $(\tilde{M}, \tilde{g}_{ab})$ it is not meaningful to think of $\tilde{M}$ as embedded in some larger spacetime, since it is not clear how to physically interpret this larger spacetime. However, similar to the one-point compactification of the Riemann sphere, there are instances when we can conformally compactify spacetime by first embedding $(M, g_{ab})$ into an unphysical spacetime $(\tilde{M}, \tilde{g}_{ab})$ via a conformal transformation, then adding boundaries and points that represent infinity. This gives us a way to study objects, in particular radiation, that travels to infinity.

**Definition 3.7.** A conformal transformation is a map $\psi : (M, g_{ab}) \rightarrow (\tilde{M}, \tilde{g}_{ab})$ such that $\tilde{g}_{ab} = \Omega^2 g_{ab}$, where $\Omega$ is strictly positive smooth function on $M$.

Note that conformal transformations preserve causal structure, i.e. suppose $v^a$ is null, timelike or spacelike with respect to $g_{ab}$. Then it is null, timelike or spacelike with respect to $\tilde{g}_{ab}$ as well.

**Example 3.8.** We may conformally compactify Minkowski space in $D = 4$ such that the unphysical metric $\tilde{\eta}$ takes the form

$$\tilde{\eta} = -dT^2 + dR^2 + \sin^2 R (d\theta^2 + \sin^2 \theta d\phi^2)$$

for $\theta, \phi$ variables of the unit round metric and

$$-\pi < T + R < \pi$$

$$-\pi < T - R < \pi$$

$$0 \leq R.$$

Schematically, this can be put into the diagram:

\[\text{Diagram}\]

---

\[\text{Footnote:} \text{Let } n^a \text{ be the unit timelike normal on each point. Then } K_{ab} \text{ is given by } \frac{1}{2} \mathcal{L}_n h_{ab}.\]
The boundaries that represent infinity may be partitioned into:
(1) Future/past null infinity $\mathcal{I}^\pm$, for $T = \pm(\pi - R), 0 < R < \pi$.
(2) Future/past timelike infinity $i^\pm$, for $T = \pm\pi, R = 0$.
(3) Spatial infinity: $i^0$, for $R = 0, T = \pi$.

Suppose that for our spacetime, matter and radiation is sufficiently local such that far away influences are negligible. We would like to say that as one approaches infinity, $g_{ab}$ decays to $\eta_{ab}$, i.e. spacetime becomes flat; this is made precise in the definition of asymptotic flatness.

**Definition 3.9.** A spacetime $(\mathcal{M}, g_{ab})$ is asymptotically simple if there exists an unphysical spacetime $(\tilde{\mathcal{M}}, \tilde{g}_{ab})$ such that there is a conformal transformation $\Phi$ satisfying,
(1) $\Phi(\mathcal{M}) \subset \tilde{\mathcal{M}}$ is open, and $\tilde{\mathcal{M}}$ has boundary $\partial \tilde{\mathcal{M}}$
(2) $\Omega = 0, \partial_\mu \Omega \neq 0$ on $\partial \tilde{\mathcal{M}}$.
(3) Every null geodesic begins and ends on $\partial \tilde{\mathcal{M}}$.

**Definition 3.10.** A spacetime $(\mathcal{M}, g_{ab})$ is asymptotically flat if the vacuum Einstein equation is satisfied at a neighborhood of $\partial \tilde{\mathcal{M}}$.

In particular, $\partial \mathcal{M} = \mathcal{I}^+ \cup i^0 \cup \mathcal{I}^-$, where the topology on each of the null infinities is $S^{D-2} \times \mathbb{R}$. Therefore, for asymptotically flat spacetime that admits a conformal compactification, its Penrose diagram resembles that of Minkowski as one travels to infinity.

**Example 3.11.** The Schwarzschild metric for a charge free, non-rotating black-hole of mass $M$ is asymptotically flat:

$$g = (1 - \frac{2M}{r})dt^2 + (1 - \frac{2M}{r})^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2 + d\phi^2).$$

It can be easily seen by taking $r \to \infty$ that $g \to \eta$. 

Figure 2: Penrose diagram for Minkowski spacetime, note that $r$ is from the original metric written in polars. Edited from [14].
Definition 3.12. Given $(\mathcal{M}, g_{ab})$ the black hole region $BH$ is defined as

$$BH = \mathcal{M} - J^-(\mathcal{I}+)$$

The future event horizon is defined as, $\mathcal{H}^+ = \partial J^-(\mathcal{I}+) \cap \mathcal{M}$.

In words, things that did not end up at null infinity fell into the black hole. The blackhole region is defined to be the trapped surface such that ingoing objects do not have a chance to reach null infinity. It is important to realize that the definition of a blackhole depends strongly on the spacetime being asymptotically flat, where null infinity is well-defined.

4 Killing horizon


For a smooth manifold $M$, consider a one-parameter group of diffeomorphisms, $\phi : \mathbb{R} \times M \to M$. Let $\phi^*_t : T^*_{\phi^t(p)}M \to T^*_pM$ be the pullback map of cotangent vectors and let $v$ be the vector field that generates $\phi$; that is, for each point $p \in M$, $v|_p$ is tangent to the orbits of $\phi_t$ at $t = 0$.

Definition 4.1. The Lie derivative of a covariant tensor $T$ in the direction of $v$ is defined as,

$$\mathcal{L}_v T = \lim_{t \to 0} \frac{\phi^*_t T - T}{t}.$$ 

Note. Given $i_X$ the interior product with respect to $X^a$ and $d$ the exterior derivative, we have the Cartan formula for differential forms,

$$\mathcal{L}_X = i_X d + di_X.$$ 

Further, for $X^a, Y^a$ vector fields we have that

$$\mathcal{L}_X Y^a = [X, Y]^a = X^b \nabla_b Y^a - Y^b \nabla_b X^a.$$ 

The Lie derivative is the infinitesimal difference of $T$ in the direction $v$. Suppose now $v$ is in a direction such that $\phi$ is a symmetry transformation, i.e. $\phi^* g_{ab} = g_{ab}$, then $v$ is a Killing field.

Definition 4.2. Let $v$ be a vector field on $\mathcal{M}$ then $v$ is a Killing field, if $\mathcal{L}_v g_{ab} = 0$.

Note. $v$ is a Killing field if and only if the flow that it generates is an isometry. Recall that Minkowski space has a 10-dimensional symmetry group, the Poincare. In light of asymptotically flat spacetime, asymptotic symmetries are $\xi^a$ that preserve infinity. At infinity, these are translations, boosts and rotations,
but the ways to asymptotically approach these symmetries are infinite. That is, \( \xi^a \) satisfies,
\[
\mathcal{L}_{\xi} \tilde{g}_{ab} \big|_{\mathcal{I}} = 0 \\
\mathcal{L}_{\xi} \tilde{g}_{ab} \big|_{\mathcal{I}} = 0
\]
These form two infinite-dimensional groups called the Bondi-Metzner-Sachs (BMS) group and the Spi (Spatial infinity) group respectively.

**Definition 4.3.** A spacetime is **stationary** if it admits a timelike killing field. A spacetime is **axisymmetric** if it admits a rotational Killing field, i.e. Killing fields whose generated flows form closed orbits.

**Example 4.4.** The Kerr metric for a charge-free rotating blackhole with mass \( M \) and angular momentum \( J = aM \), is stationary and axisymmetric:
\[
g = -dt^2 + \left( r^2 + a^2 \cos^2 \theta \right) \left( \frac{dr^2}{r^2 + 2Mr + a^2} + d\theta^2 \right) + (r^2 + a^2) \sin^2 \theta d\phi^2 + \frac{2Mr}{r^2 + a^2 \cos^2 \theta} (a \sin^2 \theta d\phi - dt)^2.
\]
It is clear that \( g \) is independent of \( t \) and \( \phi \).

**Definition 4.5.** Let \( \xi^a \) be a Killing field on \((M, g_{ab})\). The **Killing horizon** is a null hypersurface \( \mathcal{N} \) where \( \xi^a \xi_a = 0 \).

A **bifurcate Killing horizon** \( \mathcal{N} \) is comprised of two Killing horizons, \( \mathcal{N}_a \) and \( \mathcal{N}_b \) which intersect on a codimension 2 bifurcation surface \( B \). The Killing field \( K^a \big|_{\mathcal{N}_a} \) and \( K^a \big|_{\mathcal{N}_b} \) essentially point in different directions, but for \( K^a \) not to take on two different directions when \( \mathcal{N}_a, \mathcal{N}_b \) intersect, it must be zero on \( B \).

**Theorem 4.6** (Rigidity Theorem). For stationary axisymmetric black holes, we have the decomposition,
\[
t^a = K^a - \sum_{A=1}^{N} \Omega_A \phi^a_A,
\]
where \( K^a \) is the horizon killing field, and \( \phi^a_A \) rotational killing fields, with \( \Omega_A \) associating constants called angular frequency.

This theorem is discussed in [3]. The result proved to be crucial to our analysis. Note that since \( K^a \big|_B = 0 \), \( t^a \) is proportional to \( \phi^a_A \) at \( B \).

**Theorem 4.7.** Stationary, axisymmetric black holes admit a bifurcate Killing horizon that corresponds with the event horizon.

This can be shown by using the Rigidity theorem and [5][6].
5 Canonical Energy


We consider a stationary, axisymmetric black hole \((\mathcal{M}, g_{ab})\) with bifurcate Killing horizon \(\mathcal{H} = \mathcal{H}^+ \cup \mathcal{H}^-\) and bifurcation surface \(B \in \mathcal{H}\). Let \(K^a, t^a\) and \(\phi^a\) be the horizon, timelike and rotational Killing fields for \(g_{ab}\). Further, let \(A, M, \{J_A\}_{A=1}, \{P_i\}_{i=1}\) be the area of \(B\), ADM\(^8\) mass, angular momentum and linear momentum defined at spatial infinity \(i^0\). For the asymptotically flat condition we require existence of Minkowskian coordinates \(x^\mu\) such that \(g_{\mu\nu}(\lambda)\) asymptotically approach \(\eta_{\mu\nu}\) by \(O(\rho^{-(D-3)})\) as \(\rho = (x_1^2 + \cdots x_{D-1}^2)^{1/2}\).

![Figure 3: Schematic presentation of our spacetime. On the right, we have asymptotic; on the left we have the bifurcate Killing horizon and blackhole region. Edited from [14](#)](image)

We limit our discussion to even dimensions, since it is shown that conformal compactification with null infinity does not exist for odd dimensions [10], but we will need conformal infinity when calculating the canonical energy flux.

5.1 Gauge Choices at the Horizon

General Relativity obeys the **principle of general covariance**, i.e. it is invariant under diffeomorphisms of both the spacetime \(\mathcal{M}\) and within local charts. This implies that any choice of local coordinates, as long as it exists, is valid to work in. We call a suitable choice a **gauge**, and this freedom to choose, a **gauge freedom**. In this language, General Relativity is gauge invariant under coordinate transformations.

\(^8\)The ADM formalism provides a gauge invariant way of defining total mass, angular momentum and linear momentum for stationary, asymptotically flat spacetimes. Since these quantities are not well-defined locally, this notion is adopted instead. We will not be using this formalism explicitly, but here is a reference to its construction [12].
Clearly, certain gauge choices are easier and more intuitive to work with, as we know from using polar over rectangular coordinates in some situations in calculus. Therefore the natural scheme in our analysis is as follows:

1. We impose certain gauge conditions on the background that are shown to be applicable for all $g_{ab}$ that we consider.
2. We work under these convenient coordinates.
3. We prove our result to be invariant under gauge transformations that preserve the imposed gauge conditions.

By theorem 4.7, the event horizon corresponds to the Killing horizon for $g_{ab}$.

In this subsection we choose the desired gauge at the horizon such that for any perturbation the event horizon corresponds to the Killing horizon up to first order. We further have the freedom to require that the first order change in volume form $\epsilon$ relates properly to the first order change in $A$. The proofs are rather technical, therefore we state only the results.

First, near the horizon $\mathcal{H}^+$ we choose gaussian normal coordinates.\(^9\) This restricts $\mathcal{H}^+$ to a null surface. For a one-parameter family of metrics, they can be put in the form

$$g_{ab}(\lambda) = 2\nabla_{(a}u_{b)}r - r^2\alpha(\lambda)\nabla_{b}u - r\beta_{b}(\lambda) + \mu_{ab}(\lambda) \quad (3)$$

where $r, u$ are functions independent of $\lambda$, and $r = 0$ corresponds to $\mathcal{H}^+$, $u = 0$ corresponds to $\mathcal{H}^+$. $\mu_{ab}(\lambda)$ and $\beta(\lambda)$ are orthogonal to the normal bundle of $B(u, r)$, which are joint level sets of $u, r$.

![Figure 4: Gaussian null coordinates with $u, r$ near $\mathcal{H}^+$, and the Bondi gauge with $\tilde{u}$ near $\mathscr{I}^+$. The arrows on the horizon represent the Killing field. Edited from [14].](image)

In these coordinates, the Killing field $K^a$ takes the form,

$$K^a = \kappa \left[ u \left( \frac{\partial}{\partial u} \right)^a - r \left( \frac{\partial}{\partial r} \right)^a \right]$$

where $\kappa$ is a constant called surface gravity.

\(^9\)Proof of its existence can be found in [3].
Let 
\[ n^a = \frac{\partial}{\partial u}^a, \quad l^a = \frac{\partial}{\partial r}^a. \]
Then \( K^a = \kappa n^a \) on \( \mathcal{H}^+ \).

Further, we have the gauge freedom to set 
\[ \delta \Theta \big|_B = 0, \]
where \( \delta \Theta \) is the first order expansion of null geodesics on the horizon. In fact, it turns out that imposing this condition at \( B \) gives \( \delta \Theta = 0 \) on the entire horizon.

Second, given the gauge conditions above, let \( \epsilon \) be the background volume form on \( B \) and \( \delta A \) the perturbed area. Then the following gauge condition can be imposed on \( B \),
\[ \delta \epsilon \big|_B = \frac{\delta A}{A} \epsilon \big|_B. \]
Equivalently, this is,
\[ \mu^{ab} \delta \mu_{ab} \big|_B = \text{const}, \]
where \( \mu_{ab} \) is from (3).

**Lemma 5.1.** The gauge freedom remaining for any perturbation \( \gamma_{ab} \) after imposing the first and second conditions is that \( \gamma_{ab} \) is invariant under \( \gamma_{ab} \mapsto \gamma_{ab} + \mathcal{L}_X g_{ab} \), where \( X^a \) is a smooth vector field that is an asymptotic symmetry near infinity, and is tangent to \( \mathcal{H}^+ \).

### 5.2 Canonical Energy

Fix \((\mathcal{M}, g_{ab})\) with the conditions above, and let \( g_{ab}(\lambda) \) be a one-parameter family of metrics also satisfying the conditions above.

**Definition 5.2.** The Lagrangian density \( L(g) \) is a \( D \)-form on \( \mathcal{M} \) defined as
\[ L(g) = R(g) \epsilon_{1...D}, \]
where \( \epsilon_{1...D} \) is the volume form and \( R(g) \) is the Ricci scalar.

The Lagrangian density is rather important since its first variation returns the Einstein equations plus boundary terms, i.e.
\[ \frac{d}{d\lambda} L(g) = E(g) \cdot \frac{d}{d\lambda} g + d\theta(g; \frac{d}{d\lambda} g), \]
with \( E(g) = 0 \) the Einstein equation.

Explicitly calculated, \( \theta \) is a \( D - 1 \) form on \( \mathcal{M} \)
\[ \theta_{a_1...a_{D-1}} = \frac{1}{16\pi} \epsilon^{ca_1...a_{D-1}}, \]
where $v^a$ is given by

$$v^a = g^{ac} g^{bd} \left( \nabla_d \frac{d}{d\lambda} g_{ec} - \nabla_c \frac{d}{d\lambda} g_{bd} \right).$$

**Definition 5.3.** For two parametrizations $\lambda_1, \lambda_2$, define a $(D - 1)$-form $w$

$$w(g; \frac{d}{d\lambda_1} g, \frac{d}{d\lambda_2} g) = \frac{d}{d\lambda_1} \theta(g; \frac{d}{d\lambda_2} g) - \frac{d}{d\lambda_2} \theta(g; \frac{d}{d\lambda_1} g).$$

**Proposition 5.4.** $w$ is a symplectic bilinear form on the space of linear perturbations that satisfy the linear Einstein equation.

**Proof.** It is clear from the explicit formula of $\theta$ that $w$ is bilinear and antisymmetric in $(\frac{d}{d\lambda_1} g, \frac{d}{d\lambda_2} g)$. Further, let the first parametrization be $\lambda_1$. Since full derivatives commute, we may take a second antisymmetrized variation of the Lagrangian $L$ with respect to $\lambda_2$,

$$\left[ \frac{d}{d\lambda_1} \frac{d}{d\lambda_2} - \frac{d}{d\lambda_2} \frac{d}{d\lambda_1} \right] L(g) = \left[ \nabla_d \frac{d}{d\lambda_1} g \cdot \frac{d}{d\lambda_2} - \nabla_d \frac{d}{d\lambda_2} g \cdot \frac{d}{d\lambda_1} \right] g + \frac{d}{d\lambda_1} \left[ \frac{d}{d\lambda_2} \theta(g, \frac{d}{d\lambda_2} g) - \frac{d}{d\lambda_2} \theta(g, \frac{d}{d\lambda_1} g) \right] = dw(g; \frac{d}{d\lambda_1} g, \frac{d}{d\lambda_2} g) = 0.$$

Taking $\frac{d}{d\lambda_i} g|_{\lambda=0} = \gamma_i$, and $\Sigma$ any Cauchy surface, let

$$W_\Sigma(g; \gamma_1, \gamma_2) \equiv \int_\Sigma w(g; \gamma_1, \gamma_2).$$

This integral converges by our asymptotic conditions at spatial infinity for $D > 4$ and for $D = 4$ by the Regge-Teitelboim parity conditions [7] that we will assume to hold.

The following lemma identifies those vector fields $\xi^a$ whose corresponding infinitesimal gauge transformations (specifically, $\gamma_{ab} \mapsto \gamma_{ab} + L_\xi g$) leave $W_\Sigma(g; \gamma_1, \gamma_2)$ invariant for general $\gamma_1$ and $\gamma_2$. Equivalently, $W_\Sigma(g; \gamma, L_\xi g) = 0$ for general $\gamma$.

**Lemma 5.5.** Let $\gamma$ be a perturbation of $g$ satisfying the linearized Einstein equation and the gauge and asymptotically flat conditions in the previous section. Suppose we also have that $\delta A = 0$ and $\delta H = 0$ for some asymptotic symmetry $X^a$. Then for smooth $\xi$ such that

1. $\xi$ is tangent to the generators of $H^+$ at $B$,
2. $\xi^a \rightarrow c X^a$ for some constant $c \neq 0$ as $\rho = (x_1^2 + \cdots + x_{D-1}^2)^{1/2} \rightarrow \infty$,

we have

$$W_\Sigma(g; \gamma, L_\xi \gamma) = 0.$$

The converse is also true.
Definition 5.6. For any spacelike or null hypersurface $\Sigma$, and a pair of perturbations $\gamma_1, \gamma_2$, define the canonical energy $E(\gamma_1, \gamma_2) \equiv W_\Sigma(g; \gamma_1, L_t \gamma_2)$, where $t^a$ is the timelike Killing field of the background $g_{ab}$. Further, we write $E \equiv E(\gamma, \gamma)$.

Proposition 5.7. Let $\gamma_1, \gamma_2$ be two asymptotically flat perturbations that satisfy the linear Einstein equation. Then

$$E(\gamma_1, \gamma_2) = E(\gamma_2, \gamma_1).$$

Proof. Using the antisymmetry of $w$ we have,

$$w(L_t \gamma_1, \gamma_2) - w(L_t \gamma_2, \gamma_1) = w(L_t \gamma_1, \gamma_2) + w(\gamma_1, L_t \gamma_2) = L_t w(\gamma_1, \gamma_2).$$

Note that $L_t$ factors out because for $g_{ab}$ a stationary background, $t^a$ is a Killing field, and thus the Lie derivative commutes with the covariant derivative, $\nabla_a$, which $\theta$ and consequently $w$ is defined in terms of.

Now using the identity $L_t = dt + it d$, and noting that $dw = 0$, integrate the last expression over $\Sigma$,

$$\int_\Sigma dt w(\gamma_1, \gamma_2) = \int_\infty^t it w(\gamma_1, \gamma_2) - \int_B it w(\gamma_1, \gamma_2) = E(\gamma_1, \gamma_2) - E(\gamma_2, \gamma_1).$$

The first equality is by Stokes theorem, whereas the second term vanish by asymptotic properties. On the other hand, since $K^a = 0$ at $B$, we have $t^a$ in the direction of $\phi^a_A$ by the rigidity theorem (see discussion after Theorem 4.6). Therefore $t^a$ is tangent to $B$, and $it w(\gamma_1, \gamma_2)$ gives contributions normal to $B$, so the second term vanishes. Thus the final difference is zero, and $E$ is symmetric in $(\gamma_1, \gamma_2)$.

Proposition 5.8. $E(\gamma)$ is conserved in the sense that for any two Cauchy surfaces $\Sigma, \Sigma'$ extending from $B$ to $i^0$, we have

$$W_\Sigma(g; \gamma_1, \gamma_2) = W_{\Sigma'}(g; \gamma_1, \gamma_2).$$

Proof. Let $D$ be the $D$-dim space enclosed by $\Sigma$ and $\Sigma'$. Since $dw = 0$, by Stokes theorem,

$$\int_D dw = \int_\Sigma w(g; \gamma_1, \gamma_2) - \int_{\Sigma'} w(g; \gamma_1, \gamma_2) = 0.$$

Proposition 5.9. $E(\gamma)$ is invariant under the map $\gamma_{ab} \mapsto \gamma_{ab} + \mathcal{L}_X g_{ab}$ where $X^a$ preserve the gauge conditions in Lemma 5.1 and are axisymmetric, $\mathcal{L}_{\phi^a_A} X^a = 0$, perturbations that satisfy the linear Einstein equation. In addition, we require $\delta P_i = 0$ and $\delta A = 0$.

Proof. By Proposition 5.7, $E$ is symmetric, it suffices to prove that

$$E(\gamma, \mathcal{L}_X g) = E(\gamma, \mathcal{L}_{\mathcal{L}_X} g) = 0.$$

Since $t^a$ is a Killing field, $\mathcal{L}_t g_{ab} = 0$, we may write,

$$\mathcal{L}_t \mathcal{L}_X g = (\mathcal{L}_t \mathcal{L}_X - \mathcal{L}_X \mathcal{L}_t) g = \mathcal{L}_t [t, X] g.$$
Now $X^a$ is an asymptotic symmetry by assumption, and $[t, X]^a$ is an asymptotic translation because the commutator of a time translation generator with any other Poincare transformation generator must be a spatial translation generator. This can be shown in multiple ways including the Lie algebra of the Poincare group. Furthermore, by the rigidity theorem and axisymmetry of $g$, we have,

$$[t, X]^a|_B = [K, X]^a|_B - \left[ \sum A \phi_A, X \right]^a|_B = [K, X]^a|_B = 0.$$  

Thus $[t, X]^a$ is normal to $B$. Now given $\delta P_i = 0$ we have $\delta H_{[t,X]} = 0$ and $\delta A = 0$, we apply Lemma 5.5 with $\xi^a = [t, X]^a$.

**Definition 5.10.** A *perturbation towards a stationary black hole* is a $\gamma$ such that

$$\mathcal{L}_t \gamma_{ab} = \mathcal{L}_\xi g_{ab},$$

where $\xi^a|_B$ is normal to $\mathcal{H}^+$ and near infinity takes the form

$$\xi^a = ct^a + \sum A b_A \phi_A^a + \sum A a_i \left( \frac{\partial}{\partial x^i} \right)^a.$$

Note first that since $\xi$ satisfies the conditions in Lemma 5.1, by Proposition 5.9, a perturbation is gauge invariant under the addition of $\mathcal{L}_\xi g_{ab}$. Further, the set of perturbations towards a stationary blackhole forms a subspace $\mathcal{P}$. The motivation for considering perturbations towards stationary blackholes is that there exists a well-known blackhole spacetime such that $E < 0$. Moreover, there exist $\gamma_1, \gamma_2$ such that $E$ is degenerate:

**Proposition 5.11.** Let $\gamma_1$ be a perturbation with $\delta M = \delta J_A = \delta P_i = 0$ and $\gamma_2$ a perturbation towards a stationary black hole. Then

$$E(\gamma_1, \gamma_2) = 0.$$  

**Proof.** Let $\xi^a$ be the vector field associated to $\gamma_2$. We have,

$$E(\gamma_1, \gamma_2) = W_\Sigma(g; \gamma_1, \mathcal{L}_t \gamma_2) = W_\Sigma(g; \gamma_1, \mathcal{L}_\xi g) = 0.$$  

However, if we quotient our space of perturbations $\mathcal{T}$ by $\mathcal{P}$ to get $\mathcal{T}'$, $E$ then defines a conserved non-degenerate bilinear form on $\mathcal{T}' \times \mathcal{T}'$ that is, explicitly calculated, quadratic in $\nabla \gamma$. In other words, $E$ is precisely degenerate on elements in $\mathcal{T}/\mathcal{T}'$, that is, on perturbations towards stationary blackholes. It takes some analysis to get there, but we have the following result that is stronger than Proposition 5.11.

**Proposition 5.12.** Let $\gamma \in \mathcal{T}$. Then $E(\gamma, \gamma) = 0$ if and only if $\gamma$ is a perturbation towards a stationary black hole.
5.3 Canonical Energy Flux

Previously we’ve worked with $E$ on a Cauchy surface from $B$ to $i^0$. In this subsection we consider a slice $S(t)$ parametrized by $t$ that extends from $C(t) \in \mathcal{I}^+$ to $B(t) \in \mathcal{H}^+$ such that the union of $S(t)$ with $C(t)$ and $B(t)$ extended to $i^0$ and $B$ along $\mathcal{I}^+$ and $\mathcal{H}^+$ is a Cauchy surface $\Sigma_t$.

Given two such surfaces $\Sigma_{t_1}, \Sigma_{t_2}$ with $t_1 < t_2$, let $\mathcal{I}_{12} \subset \mathcal{H}^+$ be the surface from $B(t_1)$ to $B(t_2)$ along $\mathcal{H}^+$, $\mathcal{I}_{12} \subset \mathcal{I}^+$ the surface from $C(t_1)$ to $C(t_2)$, and $S(t)$ the slices extending from $B(t)$ to $C(t)$. Define $E(\gamma, S(t)) \equiv W_{S(t)}(\gamma, L_t \gamma)$. By Proposition 5.8, $E$ is conserved on $\Sigma_{t_1}, \Sigma_{t_2}$. Therefore we have that

$$E(\gamma, S(t_1)) = E(\gamma, S(t_2)) = W_{\mathcal{I}_{12}}(\gamma, L_t \gamma) - W_{\mathcal{H}_{12}}(\gamma, L_t \gamma).$$

Let $\tilde{g}_{ab}$ be the unphysical metric of our conformal compactification. Similar to our choice of Gaussian null coordinates near $\mathcal{H}^+$, near $\mathcal{I}^+$ we impose the Bondi gauge that puts $\tilde{g}$ into the following form:

$$\tilde{g} = 2\tilde{\nabla}_a \Omega \tilde{\nabla}_b \tilde{u} + \tilde{\mu}_{ab} + O(\Omega),$$

where $\tilde{\nabla}$ is the covariant derivative corresponding to $\tilde{g}$, $\tilde{\mu}_{ab}$ is the unit round metric on $\mathcal{I}^+$ and $\tilde{u}$ is future directed and tangent to the generator of $\mathcal{I}^+$.

![Diagram](image)

Figure 5: Canonical energy is conserved for the yellow, red, blue Cauchy surfaces, but decreases on the slices as $\mathcal{I}(t) \to i^+$. Edited from [14].

**Proposition 5.13.** Let $\gamma$ be an axisymmetric asymptotically flat perturbation satisfying the linearized Einstein equation, then for $t_1 < t_2$. We have that

$$E(\gamma, \mathcal{I}(t_1)) - E(\gamma, \mathcal{I}(t_2)) \leq 0.$$  

The difference is strictly less than zero for $E(\gamma, \mathcal{I}(t)) \neq 0$, and zero for $E(\gamma, \mathcal{I}(t)) = 0$. That is, $E$ decreases on $\mathcal{I}(t)$ as $t \to \infty$. 

18
5.4 Hollands-Wald Stability Criterion

The following is a summary of the propositions proven or stated.

**Proposition 5.14.** Given $g_{ab}$ and gauge conditions stated at the beginning of this section and $E$ defined in Definition 5.6, we have that

1. $E$ is conserved for any two Cauchy surfaces. (Proposition 5.8)
2. $E$ has positive flux exiting $\mathcal{H}_{12}$ and $\mathcal{I}_{12}$. (Proposition 5.13)
3. $E$ is degenerate if and only if $\gamma$ is a perturbation towards a stationary black hole. (Proposition 5.12)

Finally, we come to the Hollands-Wald stability criterion. This criterion can be enlarged to hold for general axisymmetric asymptotically flat perturbations, but we here show the result only for a restricted space $\mathcal{T}$ of perturbations with $\delta M = \delta P_3 = \delta J_A = 0$.

**Theorem 5.15** (Hollands-Wald, 2012). Let $(\mathcal{M}, g_{ab})$ be a spacetime satisfying background and gauge conditions, $\mathcal{T}'$ be the space of axisymmetric asymptotically flat perturbations with $\delta M = \delta P_3 = \delta J_A = 0$ quotient the space of perturbations toward stationary blackholes. We have,

$$E(\gamma, \gamma) \geq 0, \quad \forall \gamma \in \mathcal{T}' \iff \text{Stability}.$$

**Proof.** Suppose $E \geq 0$ for all perturbations in $\mathcal{T}$. Then by Proposition 5.14(3), $E$ defines a conserved norm on $\mathcal{T}' \times \mathcal{T}'$. Further, elements $\gamma'$ in $\mathcal{T}/\mathcal{T}'$ have $E(\gamma', \gamma') = 0$. Thus, if the canonical energy of a perturbation evaluated on the initial Cauchy surface is finite, by Proposition 5.14(2), it has to decay and remains finite and semi-positive definite for all time $t > 0$. This shows global boundedness and decay. $g_{ab}$ is then stable. Conversely, suppose there is a $\gamma \in \mathcal{T}'$ such that $E(\gamma, \gamma) < 0$. By Proposition 5.14(3), $\gamma$ cannot be a perturbation towards a stationary blackhole. Then by Proposition 5.14(2), it can only decrease and become more negative. In fact, the solution blows up exponentially for a subclass of perturbations [11]. This is instability.

**Acknowledgements**

I would like to give my sincere thanks to my mentor Kartik Prabhu from the Uchicago physics department, who, although was not an obligated participant of the math REU program, provided generous and insightful guidance on the subject. I would also like to express gratitude to my mentors Subhadip Chowdhury and Joel Specter from the Uchicago math department, whose demand for mathematical rigor made the paper possible to read; fellow student Indraneel Tambe for great discussions, and Professor Peter May, for organizing the REU and providing assistance whenever I am in need. Lastly, to all physicists and mathematicians working to solidify our understanding of General Relativity, I am excited and grateful for the hard works that are done, and for the hard works to come.
References


