Lecture XXVIII: Horizons and the area theorem

Christopher M. Hirata
*Caltech M/C 350-17, Pasadena CA 91125, USA**
(Dated: March 7, 2012)

I. OVERVIEW

In this lecture, we prove one of the simplest but most profound theorems in GR – that the horizon area in black holes can increase but not decrease. In terms of the entropy picture you derived in the previous homework, this is equivalent to the second law of thermodynamics.

Reading:

- MTW Ch. 34.
- If you like this subject, I recommend Hawking & Ellis.

II. THE LOCAL NATURE OF THE HORIZON

We begin by reconsidering the nature of the horizon. As we found in the Schwarzschild case, when properly considered in terms of a set of coordinates that are nonsingular at the horizon $(\tilde{u}, \tilde{v}, \theta, \phi)$, the horizon (in that case, $\tilde{u} = 0$) is actually a 3-dimensional submanifold $\mathcal{H}$ of the overall 4D manifold $\mathcal{M}$. This is not surprising since it represents a *boundary* between the region observable by an observer far away if he/she waits long enough (the "exterior" of the hole) and an invisible, interior region that cannot send signals to the outside. In fact, so long as spacetime is asymptotically flat, we may **define** a black hole by this property.

A. The null surface

The horizon in Schwarzschild has the further peculiar property that it is a surface formed by a bundle of outgoing null geodesics ($\tilde{u} = 0$, θ,ϕ constant, $\tilde{v}$ increasing). It turns out that this is **not** a coincidence but is in fact a general property of horizons. To explain why, we begin with a few simple realizations.

First, if we consider a point $\mathcal{P}$ on the horizon, there is a 3D vector space $T_\mathcal{P}\mathcal{H}$ of vectors tangent to the horizon, which is a subspace of the full space of vectors $T_\mathcal{P}\mathcal{M}$ tangent to the manifold. Such a subspace can always be described by its surface normal, i.e. we may say that $\boldsymbol{v}$ lies in the subspace $T_\mathcal{P}\mathcal{H}$ if and only if $\boldsymbol{v} \cdot \boldsymbol{l} = 0$, where $\boldsymbol{l}$ is a normal (defined up to an overall constant). We may also define a particular sign (though not magnitude) for the normalization constant of $\boldsymbol{l}$ by saying that a vector $\boldsymbol{v}$ points to the exterior if $\boldsymbol{v} \cdot \boldsymbol{l} > 0$ and to the interior if $\boldsymbol{v} \cdot \boldsymbol{l} < 0$.

We immediately see that $\boldsymbol{l}$ has to be null. This is because if it were timelike and forward directed, a particle with instantaneous velocity $\boldsymbol{u} = \boldsymbol{l}$ would pass from the interior of the hole to the exterior, which is impossible. If it were timelike and backward directed, then in some local Lorentz frame $l^\alpha = (-1, 0, 0, 0)$, and we could (locally) choose a gauge where the horizon is $t = 0$. Then a particle with $\boldsymbol{u} = -\boldsymbol{l}$ would pass from the exterior to the interior (which is allowed) but at a point $\mathcal{Q}$ on its trajectory ϵ before horizon crossing would have all information radiated from that point cross $t = 0$ and enter the hole's interior, so there is no escape to future infinity. Thus $\mathcal{Q}$ should have been inside the hole. Finally, if $\boldsymbol{l}$ were spacelike, we could choose a local Lorentz frame with $\boldsymbol{l} = (0, 1, 0, 0)$ and then a particle with negative 1-component of the velocity in that frame could pass from the interior to the exterior.

Arguments similar to the above also show that $\boldsymbol{l}$ is future directed (i.e. on the future null cone, not the past null cone). Thus we conclude that **the horizon $\mathcal{H}$ is a 3D surface with normal vector $\boldsymbol{l}$ given by a forward-directed null vector**.

Since $\boldsymbol{l} \cdot \boldsymbol{l} = 0$, it follows that $\boldsymbol{l}$ is tangent to $\mathcal{H}$.

---

*Electronic address: chirata@tapir.caltech.edu

B. Null geodesics

A further realization is that since $\boldsymbol{l}$ is the surface normal to $\mathcal{H}$, it is of the form $l_\alpha = q f_{;\alpha}$, where f is some function with $f = 0$ on the horizon, $f > 0$ in the exterior, and $f < 0$ in the interior, and q is a positive scalar. We then find

$$\nabla_{\boldsymbol{l}} l_\alpha = l^\beta l_{\alpha;\beta} = q f^{;\beta}(q f_{;\alpha\beta} + q_{;\beta} f_{;\alpha}) = \frac{1}{2} q^2 (f^{;\beta} f_{;\beta})_{;\alpha} + q f^{;\beta} q_{;\beta} f_{;\alpha}. \tag{1}$$

Defining $\chi = f^{;\beta} f_{;\beta}$, we thus have

$$\nabla_{\boldsymbol{l}} \boldsymbol{l} = \frac{1}{2} q^2 \boldsymbol{\nabla} \chi + (\boldsymbol{\nabla} f \cdot \boldsymbol{\nabla} q) \boldsymbol{l}. \tag{2}$$

But $\chi = 0$ on $\mathcal{H}$, i.e. when $f = 0$, so it follows that we may write $\chi = cf$ where c is a well-behaved function near $\mathcal{H}$. We thus find that

$$\nabla_{\boldsymbol{l}} \boldsymbol{l} = \frac{1}{2} q^2 (f \boldsymbol{\nabla} c + c \boldsymbol{\nabla} f) + (\boldsymbol{\nabla} f \cdot \boldsymbol{\nabla} q) \boldsymbol{l}. \tag{3}$$

In this equation, on $\mathcal{H}$ we have $f = 0$ and so we find that $\nabla_{\boldsymbol{l}} \boldsymbol{l}$ is a scalar β times $\boldsymbol{l}$.

Note that we have not yet said anything about the normalization of $\boldsymbol{l}$. We may fix this as follows. Suppose we were to define a new vector field $\boldsymbol{L} = \alpha \boldsymbol{l}$, where α is a scalar. Then

$$\nabla_{\boldsymbol{L}} \boldsymbol{L} = \alpha \nabla_{\boldsymbol{l}} (\alpha \boldsymbol{l}) = \alpha^2 \nabla_{\boldsymbol{l}} \boldsymbol{l} + \alpha (\nabla_{\boldsymbol{l}} \alpha) \boldsymbol{l} = \alpha^2 (\beta + \nabla_{\boldsymbol{l}} \ln \alpha) \boldsymbol{l}. \tag{4}$$

This may be set to zero if we choose α to be defined by

$$\nabla_{\boldsymbol{l}} \ln \alpha = -\beta, \tag{5}$$

which may be integrated independently along the integral curves of $\boldsymbol{l}$. Thus we may set the normalization of $\boldsymbol{l}$ by replacing it with $\alpha \boldsymbol{l}$. We then have that $\nabla_{\boldsymbol{l}} \boldsymbol{l} = 0$.

We thus conclude that **the horizon normal $\boldsymbol{l}$ describes a family of null geodesics tangent to the horizon,** which are called the *generators*.

III. THE GLOBAL NATURE OF THE HORIZON

We have thus considered the local nature of the horizon. But this leaves open the more interesting question of its global structure. What happens to the generators – where do they begin and end, and are there any irregular points (or kinks) on the surface? Can the generators ever intersect each other?

It is clear that the answers to these questions are related to each other. If two generators intersect, then at the point $\mathcal{P}$ of intersection there are at least two surface normals to the horizon $\boldsymbol{l}$ and $\boldsymbol{l}'$, which implies that the surface becomes nondifferentiable there. Conversely if the surface is to become nondifferentiable then the smooth set of geodesics with tangent vector $\boldsymbol{l}$ must either become singular, or intersect such that we cannot define a unique normal. We cannot dismiss this possibility; indeed it is precisely what happens at the center of a collapsing massive spherically symmetric star at the instant the black hole first forms. The essential subject of global horizon structure is to understand the location of these so-called *caustics*.

A. The nature of caustics

Let us consider a horizon generator and follow it into the future. If it does not hit a singularity, it may either continue to be tangent to $\mathcal{H}$ forever, or it may leave $\mathcal{H}$. The latter causes a problem, however. If the generator leaves $\mathcal{H}$ and enters the exterior of the hole, reaching a point $\mathcal{P}$ in the exterior, then we conclude that $\mathcal{P}$ can receive information from a point $\mathcal{Q} \in \mathcal{H}$. By slight perturbation of this trajectory, a point $\mathcal{Q}'$ near $\mathcal{Q}$ but in the interior of the hole could have sent information to a neighboring point $\mathcal{P}'$ outside the hole. Therefore the horizon generator, extended into the future, cannot go to the exterior of the hole.

[Note: a horizon generator extended into the past may go the exterior of the hole, e.g. by passing through a caustic at which a horizon is created.]

We now ask: can a horizon generator extended into the future pass through a caustic and then enter into the interior of the hole? The answer is no. For suppose that there were a point $\mathcal{P}$ where such a generator left the horizon and plunged inward, and consider a previous point $\mathcal{O} \in \mathcal{H}$ and a future point $\mathcal{Q}$ inside the hole on the generator. Then since the generator in question is the only null geodesic passing through $\mathcal{O}$ that does not immediately go into the hole (for Lorentz signature, any future-directed null vector $\boldsymbol{w}$ that could be tangent to a null geodesic at $\mathcal{O}$ has $\boldsymbol{w} \cdot \boldsymbol{l} < 0$), we conclude that all null geodesics projected forward from $\mathcal{O}$ land strictly inside the hole at some future affine parameter. Thus in some neighborhood of $\mathcal{O}$, all null geodesics projected forward in time go into the hole, and none escape to infinity. We are thus left with the conclusion that there is a point $\mathcal{O}'$ near $\mathcal{O}$, from which all null geodesics projected into the future land in the hole.

But: since $\mathcal{O}'$ is outside the horizon, there are causal curves emanating forward from it and reaching future null infinity. Among this set one can always find the curve that reaches a distant observer $\mathcal{D}$ first. We can show that this curve must be null by considering a path $\mathcal{C}(\lambda)$ originating at $\mathcal{O}'$ and terminating at some point on $\mathcal{D}$'s trajectory, with tangent vector $\boldsymbol{s}$. Then everywhere $\boldsymbol{s} \cdot \boldsymbol{s} \leq 0$. We can also consider a parameterized perturbed path $\mathcal{C}(\lambda, \kappa)$ with $\mathcal{C}(\lambda, 0) = \mathcal{C}(\lambda)$ being the original curve. Now consider a perturbation to the path $\boldsymbol{\xi} = \partial\mathcal{C}/\partial\kappa$. We immediately see that

$$\frac{\partial}{\partial\kappa}(\boldsymbol{s} \cdot \boldsymbol{s}) = \nabla_{\boldsymbol{\xi}}(\boldsymbol{s} \cdot \boldsymbol{s}) = 2\boldsymbol{s} \cdot \nabla_{\boldsymbol{\xi}}\boldsymbol{s} = 2\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{\xi}, \tag{6}$$

where the last equality holds since $[\boldsymbol{s}, \boldsymbol{\xi}] = 0$ (try computing the latter in a coordinate system where λ and κ are two of the coordinates). Thus a perturbed curve remains causal if $\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{\xi} < 0$ along null portions of the curve, and for arbitrary $\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{\xi}$ in timelike parts.

Now take the 4-velocity $\boldsymbol{u}$ of the distant observer at $\mathcal{D}$, and form the vector field $\boldsymbol{t}$ along the curve $\mathcal{C}(\lambda)$ by parallel transport ($\boldsymbol{t} = \boldsymbol{u}$ at $\mathcal{D}$, and $D\boldsymbol{t}/d\lambda = 0$ along the curve). Then if the curve has a timelike segment, we may define the perturbed curve

$$\boldsymbol{\xi} = f(\lambda)\boldsymbol{t}, \tag{7}$$

where $df/d\lambda < 0$ in a portion of the timelike segment, $df/d\lambda > 0$ in the null segments, $f = 0$ at $\mathcal{O}'$ and $f < 0$ at $\mathcal{D}$. This perturbed curve reaches $\mathcal{D}$ a small amount of time earlier than the original – by an amount $\kappa|f(\mathcal{D})|$. Also since

$$\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{\xi} = \boldsymbol{s} \cdot \frac{D}{d\lambda}(f\boldsymbol{t}) = \frac{df}{d\lambda}\boldsymbol{s} \cdot \boldsymbol{t} + f\frac{D\boldsymbol{t}}{d\lambda} = \frac{df}{d\lambda}\boldsymbol{s} \cdot \boldsymbol{t} \tag{8}$$

and $\boldsymbol{s} \cdot \boldsymbol{t} < 0$ (a forward-directed null vector dotted with a forward timelike vector), it satisfies the causality conditions. Therefore the first curve to reach $\mathcal{D}$ is null.

It can then be shown that this curve must be a null geodesic. This is because if it is not a geodesic, and there is an acceleration $\boldsymbol{a} \equiv \boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{s} \neq 0$ somewhere, we have that:

$$\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{a} = \boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\nabla_{\boldsymbol{s}}\boldsymbol{s} = \nabla_{\boldsymbol{s}}(\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{s}) - \nabla_{\boldsymbol{s}}\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{s} = \frac{1}{2}\nabla_{\boldsymbol{s}}\nabla_{\boldsymbol{s}}(\boldsymbol{s} \cdot \boldsymbol{s}) - \nabla_{\boldsymbol{s}}\boldsymbol{s} \cdot \nabla_{\boldsymbol{s}}\boldsymbol{s} = -\boldsymbol{a} \cdot \boldsymbol{a}, \tag{9}$$

where we have used that $\boldsymbol{s} \cdot \boldsymbol{s} = 0$. Furthermore, since $\boldsymbol{s} \cdot \boldsymbol{a} = 0$, we find that

$$\boldsymbol{s} \cdot \frac{D}{d\lambda}[f(\lambda)\boldsymbol{a}] = -f(\lambda)\boldsymbol{a} \cdot \boldsymbol{a} \tag{10}$$

for any f. Now if the curve has a spacelike acceleration anywhere, $\boldsymbol{a} \cdot \boldsymbol{a}$, then by taking $f(\lambda)$ to be positive somewhere in that region and zero elsewhere, and then taking the perturbation $\boldsymbol{\xi} = f\boldsymbol{a}$, we can construct a curve with a timelike segment that is tied with the fastest curve to reach $\mathcal{D}$. This would contradict our earlier argument, so $\boldsymbol{a}$ must be null or timelike.

But then we see also that

$$\frac{d}{d\lambda}(\boldsymbol{s} \cdot \boldsymbol{s}) = 2\boldsymbol{s} \cdot \frac{D\boldsymbol{s}}{d\lambda} = 2\boldsymbol{s} \cdot \boldsymbol{a} = 0, \tag{11}$$

so $\boldsymbol{a}$ is orthogonal to $\boldsymbol{s}$. If we now choose a local Lorentz frame where $s^\mu = (1, 1, 0, 0)$ (which can always be done by boosts/rotations), we find that this restricts $a^\mu = (a^0, a^0, a^2, a^3)$. Then $\boldsymbol{a} \cdot \boldsymbol{a} = (a^2)^2 + (a^3)^2$, so the non-spacelike condition requires $a^2 = a^3 = 0$ and hence $\boldsymbol{a}$ is a scalar times $\boldsymbol{s}$. This leaves us with the conclusion that the first curve to reach $\mathcal{D}$ is a null geodesic – a contradiction to our finding that all the null geodesics from $\mathcal{O}'$ end up in the hole.

Thus: **generators cannot leave the horizon if projected into the future. Caustics appear only in the past as generators enters the horizon.**

IV. HORIZON AREA

We now turn our attention to another topic: how to determine the area of a horizon. The notion of area applies to a 2-dimensional surface: we imagine a spacelike hypersurface Σ intersecting the horizon $\mathcal{H}$, and we may consider the area of $\Sigma \cap \mathcal{H}$. We imagine that the horizon generators are described by $\mathcal{P}(\alpha, \beta; \lambda)$ where λ is the affine parameter and α, β are the 2 parameters indicating a particular generator.

Now we may compute the area element $d^2S/d\alpha\, d\beta$. A section of horizon with sides $\Delta\alpha$ and $\Delta\beta$ form a parallelogram with vector lengths

$$\bar{\boldsymbol{a}} = \partial_\alpha \mathcal{P} + c_\alpha \partial_\lambda \mathcal{P} \quad \text{and} \quad \bar{\boldsymbol{b}} = \partial_\beta \mathcal{P} + c_\beta \partial_\lambda \mathcal{P}, \tag{12}$$

where c_α and c_β are chosen so as to make these vectors lie tangent to Σ. The area they trace out is given by a determinant,

$$\frac{d^2S}{d\alpha\, d\beta} = \sqrt{(\bar{\boldsymbol{a}} \cdot \bar{\boldsymbol{a}})(\bar{\boldsymbol{b}} \cdot \bar{\boldsymbol{b}}) - (\bar{\boldsymbol{a}} \cdot \bar{\boldsymbol{b}})^2}. \tag{13}$$

Now since (i) $\boldsymbol{l} = \partial_\lambda \mathcal{P}$ is null and hence orthogonal to itself; and (ii) $\boldsymbol{a} \equiv \partial_\alpha \mathcal{P}$ and $\boldsymbol{b} \equiv \partial_\beta \mathcal{P}$ are tangent to $\mathcal{H}$ and hence orthogonal to $\boldsymbol{l}$, it follows that we may drop $\boldsymbol{l}$ from the dot products and equally well write:

$$\frac{d^2S}{d\alpha\, d\beta} = \sqrt{(\boldsymbol{a} \cdot \boldsymbol{a})(\boldsymbol{b} \cdot \boldsymbol{b}) - (\boldsymbol{a} \cdot \boldsymbol{b})^2}. \tag{14}$$

A. Example: Kerr hole

A simple example can be found for the Kerr hole, if we take $r = r_+$ and fix t (we should be a bit more careful here since t is singular, but since it is associated with a Killing vector no trouble arises by doing this), and consider the above determinant for θ and ϕ as the parameters of the horizon. Then

$$\frac{d^2S}{d\theta\, d\phi} = \sqrt{g_{\theta\theta} g_{\phi\phi} - (g_{\theta\phi})^2} = \sqrt{\Sigma \frac{(r_+^2 + a^2)^2}{\Sigma} \sin^2\theta} = (r_+^2 + a^2)\sin\theta = 2Mr_+ \sin\theta. \tag{15}$$

It thus follows that integrating over the horizon we get an area

$$S = 8\pi M r_+ = 8\pi M^2 (1 + \sqrt{1 - \chi^2}). \tag{16}$$

V. THE LAW OF AREA INCREASE

We are finally ready for one of the great theorems of black hole physics – the area increase theorem. We begin with a look at the dynamics of the area element $J \equiv d^2S/d\alpha\, d\beta$, with a view toward understanding its rate of increase (or decrease).

A. Horizon kinematics

We begin by extending $\boldsymbol{l}$ off of $\mathcal{H}$ in an arbitrary way (subject to still being null geodesics). Then define $A_{\alpha\beta} = \nabla_\alpha l_\beta$. By nullness,

$$A_{\alpha\beta} l^\beta = \frac{1}{2} \nabla_\alpha (l^\beta_\beta) = 0, \tag{17}$$

and we already know that

$$A_{\alpha\beta} l^\alpha = \nabla_{\boldsymbol{l}} l_\beta = 0. \tag{18}$$

Then we find that

$$\frac{1}{2}\frac{d(J^2)}{d\lambda} = (\boldsymbol{a} \cdot \boldsymbol{a})(\boldsymbol{b} \cdot \nabla_{\boldsymbol{l}} \boldsymbol{b}) + (\boldsymbol{a} \cdot \nabla_{\boldsymbol{l}} \boldsymbol{a})(\boldsymbol{b} \cdot \boldsymbol{b}) - (\boldsymbol{a} \cdot \boldsymbol{b})(\boldsymbol{a} \cdot \nabla_{\boldsymbol{l}} \boldsymbol{b} + \boldsymbol{b} \cdot \nabla_{\boldsymbol{l}} \boldsymbol{a}). \tag{19}$$

Now we can see that

$$\nabla_{\boldsymbol{l}} a^\mu - \nabla_{\boldsymbol{a}} l^\mu = [\boldsymbol{l}, \boldsymbol{a}]^\mu = l^\nu a^\mu{}_{,\nu} - a^\nu l^\mu{}_{;\nu} = \partial_\lambda \partial_\alpha x^\mu - \partial_\alpha \partial_\lambda x^\mu = 0. \tag{20}$$

Therefore using the commutation relation we have

$$\frac{1}{2}\frac{d(J^2)}{d\lambda} = (\boldsymbol{a}\cdot\boldsymbol{a})(\boldsymbol{b}\cdot\nabla_{\boldsymbol{b}}\boldsymbol{l}) + (\boldsymbol{a}\cdot\nabla_{\boldsymbol{a}}\boldsymbol{l})(\boldsymbol{b}\cdot\boldsymbol{b}) - (\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\nabla_{\boldsymbol{b}}\boldsymbol{l} + \boldsymbol{b}\cdot\nabla_{\boldsymbol{a}}\boldsymbol{l}). \tag{21}$$

Now in considering the right-hand side, let us build a frame in which the coordinate basis at a point $\mathcal{P} \in \mathcal{H}$ is $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{l}, \boldsymbol{n}\}$, where we have chosen $\boldsymbol{n}$ to be a null vector orthogonal to $\boldsymbol{a}$ and $\boldsymbol{b}$, and to have $\boldsymbol{n}\cdot\boldsymbol{l} = 1$. [Exercise: prove that this choice is always legal, in fact there are infinitely many choices since one may add a multiple of $\boldsymbol{l}$ to $\boldsymbol{n}$.] Then in this basis we have metric tensor

$$g_{\mu\nu} = \begin{pmatrix} \boldsymbol{a}\cdot\boldsymbol{a} & \boldsymbol{a}\cdot\boldsymbol{b} & 0 & 0 \\ \boldsymbol{a}\cdot\boldsymbol{b} & \boldsymbol{b}\cdot\boldsymbol{b} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{22}$$

The inverse metric is

$$g^{\mu\nu} = J^{-2} \begin{pmatrix} \boldsymbol{b}\cdot\boldsymbol{b} & -\boldsymbol{a}\cdot\boldsymbol{b} & 0 & 0 \\ -\boldsymbol{a}\cdot\boldsymbol{b} & \boldsymbol{a}\cdot\boldsymbol{a} & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{23}$$

From this:

$$\boldsymbol{\nabla}\cdot\boldsymbol{l} = g^{\mu\nu}\nabla_\mu l_\nu = J^{-2}[(\boldsymbol{b}\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\nabla_{\boldsymbol{a}}\boldsymbol{l}) - (\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\nabla_{\boldsymbol{b}}\boldsymbol{l} + \boldsymbol{b}\cdot\nabla_{\boldsymbol{a}}\boldsymbol{l}) + (\boldsymbol{a}\cdot\boldsymbol{a})(\boldsymbol{b}\cdot\nabla_{\boldsymbol{b}}\boldsymbol{l}) - \boldsymbol{n}\cdot\nabla_{\boldsymbol{l}}\boldsymbol{l} - \boldsymbol{l}\cdot\nabla_{\boldsymbol{n}}\boldsymbol{l}]. \tag{24}$$

The last two terms are zero since $\boldsymbol{l}$ is a geodesic field and is null. The remaining terms are those in Eq. (21) and give

$$\boldsymbol{\nabla}\cdot\boldsymbol{l} = \frac{1}{2J^2}\frac{d(J^2)}{d\lambda} = \frac{d\ln J}{d\lambda}. \tag{25}$$

We thus see that **the fractional rate of increase of the area element is simply $\boldsymbol{\nabla}\cdot\boldsymbol{l}$.**

B. Horizon dynamics

The next step is to understand the evolution of $\boldsymbol{\nabla}\cdot\boldsymbol{l}$ as we follow a generator. Using the definition of the Riemann tensor, we find

$$\frac{d}{d\lambda}(\boldsymbol{\nabla}\cdot\boldsymbol{l}) = l^\mu\nabla_\mu(\nabla_\nu l^\nu) = l^\mu l^\nu{}_{;\nu\mu} = l^\mu(l^\nu{}_{;\mu\nu} + R^\nu{}_{\gamma\mu\nu}l^\gamma) = l^\mu l^\nu{}_{;\mu\nu} - R_{\gamma\mu}l^\mu l^\gamma. \tag{26}$$

Moreover, we know that $l^\mu l^\nu{}_{;\mu} = 0$, so taking its derivative and using the product rule we find

$$l^\mu l^\nu{}_{;\mu\nu} + l^\mu{}_{;\nu} l^\nu{}_{;\mu} = 0. \tag{27}$$

Also we use the null nature of $\boldsymbol{l}$ to replace the Riemann tensor with the Einstein tensor:

$$\frac{d}{d\lambda}(\boldsymbol{\nabla}\cdot\boldsymbol{l}) = -l^\mu{}_{;\nu} l^\nu{}_{;\mu} - G_{\gamma\mu}l^\mu l^\gamma. \tag{28}$$

The left-hand side is now the second derivative of $\ln J$. The right-hand side contains one term associated with first derivatives of $\boldsymbol{l}$, and one term associated with the Einstein tensor (and hence matter content). Our task is to say something useful about each.

First consider $-A_{\mu\nu} = l_{\mu;\nu}$. This in general has 16 components, but many are simple. For example, we know that $A_{\mu\nu}l^\mu = A_{\mu\nu}l^\nu = 0$ by the nullness and geodesic conditions, thereby reducing it to 9 components. We switch to a basis $\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{l}, \boldsymbol{n}\}$ basis with metric

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{29}$$

In this case, we find that for $A_{\mu\nu}$ the lower-l indices vanish, and

$$A_{\mu\nu}A^{\nu\mu} = (A_{xx})^2 + (A_{yy})^2 + 2(A_{xy})(A_{yx}). \tag{30}$$

But there are restrictions on the form of $A_{\mu\nu}$. In particular:

$$A_{[\mu;\nu]} = l_{[\nu;\mu]} = (qf_{[;\nu})_{;\mu]} = q_{[;\mu}f_{;\nu]} = \frac{q_{;[\mu}l_{\nu]}}{q}, \tag{31}$$

and so the antisymmetric part $A_{[xy]} = 0$. It thus follows that $A_{xy} = A_{yx}$ and so

$$A_{\mu\nu}A^{\nu\mu} = (A_{xx})^2 + (A_{yy})^2 + 2(A_{xy})^2. \tag{32}$$

This is to be compared with

$$\boldsymbol{\nabla}\cdot\boldsymbol{l} = A^{\mu}{}_{\mu} = A_{xx} + A_{yy}. \tag{33}$$

We then see that

$$A_{\mu\nu}A^{\nu\mu} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{l})^2 = (A_{xx})^2 + (A_{yy})^2 + 2(A_{xy})^2 - \frac{1}{2}(A_{xx} + A_{yy})^2 = \frac{1}{2}[(A_{xx})^2 - (A_{yy})^2] + 2(A_{xy})^2 \geq 0. \tag{34}$$

Thus Eq. (28) reduces to

$$\frac{d}{d\lambda}(\boldsymbol{\nabla}\cdot\boldsymbol{l}) \leq -\frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{l})^2 - G_{\gamma\mu}l^{\mu}l^{\gamma}. \tag{35}$$

The inequality may not seem like an improvement, but note that we have isolated $\boldsymbol{\nabla}\cdot\boldsymbol{l}$, without any of the more complicated metric machinery.

We now use Einstein's equations to get:

$$\frac{d}{d\lambda}(\boldsymbol{\nabla}\cdot\boldsymbol{l}) \leq -\frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{l})^2 - 8\pi T_{\gamma\mu}l^{\mu}l^{\gamma}. \tag{36}$$

Note that this remains valid even in the presence of a cosmological constant.

C. Null energy condition

To proceed, we must make a physical assumption about the matter fields present. We define the *null energy condition* as the condition that $T_{\gamma\mu}l^{\mu}l^{\gamma} \geq 0$ for any null $\boldsymbol{l}$. This condition is easily verified for normal matter – e.g. perfect fluids with $|p| < \rho$, or for the electromagnetic field. In this case, Eq. (36) gives

$$\frac{d}{d\lambda}(\boldsymbol{\nabla}\cdot\boldsymbol{l}) \leq -\frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{l})^2. \tag{37}$$

Therefore, if $\boldsymbol{\nabla}\cdot\boldsymbol{l} = -C < 0$, we can integrate this ODE and find that $\boldsymbol{\nabla}\cdot\boldsymbol{l}$ reaches $-\infty$ within a finite affine parameter $1/(2C)$. Thus a caustic forms in the future, which is impossible. Thus $\boldsymbol{\nabla}\cdot\boldsymbol{l} \geq 0$, $dJ/d\lambda \geq 0$, and it follows that the area of a parcel of horizon with some width $\Delta\alpha\,\Delta\beta$ never decreases. Since new horizon can form in caustics, and the parcel can grow, the total area of horizons may increase – but never decrease!