Deriving the Gibbons-Hawking-York Boundary Term

Abdel Elabd

April 2023

1 Acknowledgements

This closely follows an MsC thesis by a certain *Simone S. Bavera* from ETH Zurich, titled "The Boundary Terms of the Einstein-Hilbert Action". Link here: `https://www.physik.uzh.ch/dam/jcr:1c2ca173-a3e8-41a1-b26e-0cda64debc6c/master_thesis_bavera.pdf`

Also, chapter 4 of Eric Poisson's *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics*, and chapter 2 of Hand & Finch's *Analytical Mechanics*.

2 Conventions

Conventions

$$R^{\alpha}{}_{\mu\beta\nu} = \Gamma^{\alpha}_{\mu\nu,\beta} - \Gamma^{\alpha}_{\mu\beta,\nu} + \Gamma^{\alpha}_{\lambda\beta}\Gamma^{\lambda}_{\mu\nu} - \Gamma^{\alpha}_{\lambda\nu}\Gamma^{\lambda}\mu\beta \tag{1}$$

Metric signature: (-, +, +, +)

$$c = G = \hbar = k_B = 1 \tag{2}$$

3 Motivation

"The starting point of any field theory is the action principle."

In GR, this is given by the Einstein-Hilbert action

$$S_{EH} = \frac{c^4}{16\pi G}\int_{\mathcal{M}} d^4x\sqrt{-g}R \tag{3}$$

When this was first formulated, the contribution of the boundary term (to the variation of the action) was forced to vanish by fixing the metric at infinity. However, when dealing with spacetime manifolds with (finite) boundaries, doing so would implicitly assume a boundary condition.

When we explicitly consider the contribution of the boundary term, without fixing the metric at infinity, it doesn't trivially vanish. In order to make the boundary term vanish, then, it is necessary to add an extra term to the Einstein-Hilbert action: a counter boundary term. There are many ways to do this, and Einstein himself did it, albeit with a non-covariant counter term.

Generally, we do want this counter-term to be covariant. "The most-popular covariant counter-term to the EH action is the Gibbons-Hawking-York boundary term."

Digression/Questions

The reason we consider boundary-terms in the first place is because the Einstein-Hilbert action contains second-order derivatives of the dynamical variable (in this case, the metric).

See Hand & Finch 2.2 for a more rigorous explanation of this, but here's my hand-wavey understanding:

For a function with only a first-derivative, you can constrain it with 2 pieces of information: slope and intercept. The single definite integral in the usual action-formalism, with two endpoints, provides this.

For a function with a second-derivative, you now need more information to constrain it: slope, intercept, and the slope of the slope. The usual definite integral constrains the slope-of-slope and the slope (the higher-order equivalents of the slope and intercept, respectively); you then need to also consider the *boundary* integral in order to constrain the "intercept".

Then what about a third-order integrand? Is it not solveable? Or is there another integral you can perform, over the boundary of the boundary? Can a boundary even have its own boundary?

It's not clear that we can just extrapolate this boundary-integral formulation indefinitely, which makes me wonder if my hand-wavey logic is correct to begin with...

I wonder if this has something to do with Galois Theory

4 First, some hypersurface maths

"... the induced metric $h_{\alpha\beta}$ has to be held fixed on the boundary, rather than the full metric $g_{\mu\nu}$"

Holding the metric fixed at infinity *might* be a valid method for dealing with the entire universe/all of spacetime (i.e. the FRW metric), but it's probably not valid when dealing with a specific, bounded submanifold of all spacetime (e.g. a black hole), for the aforementioned issue with the implied assumed boundary conditions.

We're therefore interested in performing algebra/calculus on a *submanifold* of a spacetime. We're especially interested in performing operations on a 3D *hypersurface*, Σ, embedded in 4D spacetime.

Usually, a hypersurface is defined by a parametrization:

$$x^\alpha = x^\alpha(y^\alpha) \tag{4}$$

But here, instead, they define it using a scalar function Φ such that the direction of increasing Φ points normal to the hypersurface, i.e.

$$\Phi(x^\alpha) = 0 \tag{5}$$

4.1 Unit normal vector to a hypersurface

If the hypersurface is non-null (i.e. $ds^2 \neq 0, g^{\mu\nu}\Phi_{,\mu}\Phi_{,\nu} \neq 0$), the *unit normal vector* to Σ is

$$n_\alpha = \frac{\varepsilon \Phi_{,\alpha}}{|g^{\mu\nu}\Phi_{,\mu}\Phi_{\nu}|^{\frac{1}{2}}} \tag{6}$$

Where

$$\varepsilon = n_\mu n^\mu = \begin{cases} +1 \, timelike \\ -1 \, timelike \end{cases} \tag{7}$$

i.e. n^α points in the direction of increasing Φ

4.2 Length of intervals on a hypersurface

"The three-metric induced on Σ is obtained by restricting the line element to displacements confined to the hypersurface

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta = g_{\alpha\beta}\left(\frac{\partial x^\alpha}{\partial y^a}dy^a\right)\left(\frac{\partial x^\beta}{\partial y^b}dy^b\right) = g_{\alpha\beta}e^\alpha_a e^\beta_b dy^a dy^b \equiv h_{ab}dy^a dy^b \tag{8}$$

Where $e^\alpha_a := \frac{\partial x^\alpha}{\partial y^a}$ are vectors tangent to curves contained in Σ, i.e. $n_\alpha e^\alpha_a = 0$. ”

4.3 Induced metric on a hypersurface

“h_{ab} is the *induced metric* of the hypersurface

$$h_{ab} := g_{\alpha\beta} e^\alpha_a e^\beta_b \tag{9}$$

Note that h_{ab} transforms as a scalar under spacetime coordinate transformations $x^\alpha \to x'^\alpha$, and as a tensor under hypersurface coordinate transformations $y^a \to y'^a$.

Furthermore, given the induced metric and the normal to Σ, one can recover the metric:

$$g_{\alpha\beta} = h_{ab} e^a_\alpha e^b_\beta + \varepsilon n_\alpha n_\beta \equiv h_{\alpha\beta} + \varepsilon n_\alpha n_\beta \tag{10}$$

”

4.4 Covariant derivative on a hypersurface

First, define the projection operator

$$f^\alpha{}_\beta = \delta^\alpha{}_\beta - \varepsilon n^\alpha n_\beta \tag{11}$$

This allows us to define the *induced* covariant derivative

$$D_\alpha A_\beta := f^\mu{}_\alpha f^\nu{}_\beta \nabla_\mu A_\nu \tag{12}$$

$$D_\alpha A^\beta := f^\mu{}_\alpha f^\nu{}_\beta \nabla_\mu A^\nu \tag{13}$$

Side note

In the text they use ”$h^\alpha{}_\beta$” to denote the projection operator, but I can’t fathom why you would use the same letter as the induced metric when the two are definitely not the same thing.

4.5 Divergence of a tangent vector (field) on a hypersurface

Imagine two tangent vector fields to Σ, $\mathbf{u}$ and $\mathbf{v}$. Consider the (induced) covariant derivative of $\mathbf{u}$ along $\mathbf{v}$.

$$(D_\mathbf{v}\mathbf{u})^\alpha = v^\lambda D_\lambda u^\alpha = v^\lambda f^\mu{}_\lambda f^\alpha{}_\nu \nabla_\mu u^\nu = v^\mu(\delta^\alpha{}_\nu - \varepsilon n^\alpha n_\nu)\nabla_\mu u^\nu = v^\mu \nabla_\mu u^\alpha - \varepsilon v^\mu n^\alpha n_\nu \nabla_\mu u^\nu \tag{14}$$

By the definition of normal and tangent vectors, we know $n_\nu u^\nu = 0$, which allows us to write $n_\nu \nabla_\mu u^\nu = -u^\nu \nabla_\mu n_\nu$. In words: the derivative of the tangent vector along the normal vector is equal and opposite in magnitude to the derivative of the normal vector along the tangent vector.

Question

Why must the two vectors be orthogonal in order for this to be true?

Guess: Has something to do with commutation of orthogonal and non-orthogonal vectors, and the fact that the free indeces are between n_α and u^α, not ∇.

I’m actually not entirely sure how to interpret this.

Plugging $n_\nu \nabla_\mu u^\nu = -u^\nu \nabla_\mu n_\nu$ into (14) gives:

$$(D_\mathbf{v}\mathbf{u})^\alpha = v^\lambda D_\lambda u^\alpha = ... = v^\mu \nabla_\mu u^\alpha + \varepsilon v^\mu n^\alpha u^\nu \nabla_\mu n_\nu \tag{15}$$

Since λ is not a free-index on the LHS

$$v^\mu D_\mu u^\alpha = v^\mu \nabla_\mu u^\alpha + \varepsilon v^\mu n^\alpha u^\nu \nabla_\mu n_\nu \tag{16}$$

And because ε is a scalar and commutes with v^μ, we can take v^μ out from both sides and rearrange to get

$$\nabla_\beta u^\alpha = D_\beta u^\alpha - \varepsilon n^\alpha u^\nu \nabla_\beta n_\nu \tag{17}$$

Pretty handy: we've related the full covariant derivative in 4D spacetime to the *induced* covariant derivative on the hypersurface and the full covariant derivative of the unit normal to the hypersurface.

Defining $a_\nu := n^\alpha \nabla_\alpha n_\nu$, we get the following expression for the divergence

$$\nabla_\alpha u^\alpha = D_\alpha u^\alpha - \varepsilon a_\nu u^\nu \tag{18}$$

Even handier: we've defined the divergence of a tangent vector (u^a) to a hypersurface (Σ), in terms of the normal to the hypersurface (n^α), the induced metric on the hypersurface (h_{ab}), the induced covariant derivative on the hypersurface (D_α), and some information about whether the hypersurface is spacelike or timelike (ε) .

4.6 Infinitesimal element of a hypersurface

$$d\Sigma = \sqrt{|h|} d^3 y \tag{19}$$

Where $|h| \equiv det(h_{ab})$

The *directed* surface element that points normal to the surface is $n_\alpha d\Sigma$, and we might also like to distinguish between spacelike or timelike surface elements:

$$d\Sigma_\alpha := \varepsilon n_\alpha d\Sigma = \varepsilon n_\alpha \sqrt{|h|} d^3 y \tag{20}$$

4.7 Gauss' Law

"For any vector field A^α defined within a finite region of spacetime manifold $\mathcal{M}$, bounded by a closed hypersurface $\partial\mathcal{M}$, the following holds [and is the equivalent of Gauss' Law]:

$$\int_{\mathcal{M}} d^4 x \sqrt{-g} \nabla_\alpha A^\alpha = \int_{\partial\mathcal{M}} d\Sigma_\alpha A^\alpha \tag{21}$$

"

4.8 Stokes' Theorem

"For any antisymmetric tensor field $B^{\alpha\beta}$ in a three-dimensional region of the hypersurface Σ, bounded by a closed two-surface $\partial\Sigma$, the following holds [and is the equivalent of Stokes' Theorem, at least for antisymmetric tensors e.g. Maxwell tensor. Don't know if Stokes' theorem generally applies for non-antisymmetric tensors]

$$\int_\Sigma d\Sigma_\alpha \nabla_\beta B^{\alpha\beta} = \frac{1}{2} \int_{\partial\Sigma} dS_{\alpha\beta} B^{\alpha\beta} \tag{22}$$

"

5 Deriving the Gibbons-Hawking-York Boundary Term

Cracks knuckles

5.1 Variation of the Einstein-Hilbert Action

The starting point is the variation of the Einstein-Hilbert action. For a derivation of this, see *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics* by Eric Poisson (chapter 4, section 1). Apparently it's also in Weinberg's book.

The variation of the Einstein-Hilbert action is given by:

$$16\pi \delta S_{EH} = \int_{\mathcal{M}} d^4 x \sqrt{-g} \left[G_{\mu\nu} \delta g^{\mu\nu} + \nabla_\mu (g^{\alpha\beta} \delta \Gamma^\mu_{\alpha\beta} - g^{\alpha\mu} \delta \Gamma^\beta_{\beta\alpha}) \right] \tag{23}$$

Define V^μ

$$V^\mu = g^{\alpha\beta}\delta\Gamma^\mu_{\alpha\beta} - g^{\alpha\mu}\delta\Gamma^\beta_{\beta\alpha} \tag{24}$$

And (23) becomes

$$16\pi\delta S_{EH} = \int_\mathcal{M} d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} + \int_\mathcal{M} d^4x\sqrt{-g}\nabla_\mu V^\mu \tag{25}$$

Applying Gauss' Law (21) to the second term on the RHS, we get a boundary term!

$$16\pi\delta S_{EH} = \int_\mathcal{M} d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} + \int_{\partial\mathcal{M}} d\Sigma_\alpha V^\alpha \tag{26}$$

Using the definition of $d\Sigma_\alpha$, (20), we get the following expression for the variation of the Einstein-Hilbert action

$$16\pi\delta S_{EH} = \int_\mathcal{M} d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} + \int_{\partial\mathcal{M}} d^3y\varepsilon\sqrt{|h|}n_\mu V^\mu \tag{27}$$

5.2 Evaluating the boundary term

We're interested in the boundary term on the RHS of (27). We ignore the scalar coefficients for now and focus just on $n_\mu V^\mu$.

5.2.1 Expanding the boundary term

Start by invoking the following relations:

$$\delta(\nabla_\mu n_\nu) = \nabla_\mu\delta n_\nu - \delta\Gamma^\lambda_{\mu\nu}n_\lambda \tag{28}$$

$$\delta(\nabla_\mu n^\mu) = \nabla_\mu\delta n^\mu - \delta\Gamma^\mu_{\mu\lambda}n^\lambda \tag{29}$$

Side note

Try to work these out later.

Somehow, the variation of the covariant derivative along n_ν only pulls out a single Christoffel symbol.

Free-indices commute, and with some rearrangment (28) becomes

$$n_\lambda\delta\Gamma^\lambda_{\mu\nu} = \nabla_\mu\delta n_\nu - \delta(\nabla_\mu n_\nu) \tag{30}$$

And (29) becomes

$$n^\lambda\delta\Gamma^\mu_{\mu\lambda} = -\nabla_\mu\delta n^\mu + \delta(\nabla_\mu n^\mu) \tag{31}$$

This allows us to simplify the term $n_\mu V^\mu$. Substituting the definition of V^μ, (24)

$$n_\mu V^\mu = n_\mu(g^{\alpha\beta}\delta\Gamma^\mu_{\alpha\beta} - g^{\alpha\mu}\delta\Gamma^\beta_{\beta\alpha}) = g^{\alpha\beta}n_\mu\delta\Gamma^\mu_{\alpha\beta} - n^\alpha\delta\Gamma^\beta_{\beta\alpha} \tag{32}$$

Substituting (30) and (31)

$$n_\mu V^\mu = g^{\alpha\beta}(\nabla_\alpha\delta n_\beta - \delta(\nabla_\alpha n_\beta)) + \nabla_\alpha\delta n^\alpha - \delta(\nabla_\alpha n^\alpha) \tag{33}$$

Let's inspect the first term on the RHS of (33) more closely...

$$g^{\alpha\beta}(\nabla_\alpha\delta n_\beta - \delta(\nabla_\alpha n_\beta)) = g^{\alpha\beta}\nabla_\alpha\delta n_\beta - g^{\alpha\beta}\delta(\nabla_\alpha n_\beta)$$

$$= \left[\nabla_\alpha(g^{\alpha\beta}\delta n_\beta) - (\nabla_\alpha g^{\alpha\beta})\delta n_\beta\right] - g^{\alpha\beta}\delta(\nabla_\alpha n_\beta)$$

$$= \nabla_\alpha (g^{\alpha\beta}\delta n_\beta) - g^{\alpha\beta}\delta(\nabla_\alpha n_\beta)$$

$$= \nabla_\alpha (g^{\alpha\beta}\delta n_\beta) - \delta(g^{\alpha\beta}\nabla_\alpha n_\beta) + (\delta g^{\alpha\beta})\nabla_\alpha n_\beta \tag{34}$$

Where we used $\nabla_\alpha g^{\alpha\beta} = 0$ in the third line of (34). Then (33) becomes

$$n_\mu V^\mu = \nabla_\alpha (g^{\alpha\beta}\delta n_\beta) - \delta(g^{\alpha\beta}\nabla_\alpha n_\beta) + (\delta g^{\alpha\beta})\nabla_\alpha n_\beta + \nabla_\alpha \delta n^\alpha - \delta(\nabla_\alpha n^\alpha)$$

$$= \nabla_\alpha (\delta n^\alpha + g^{\alpha\beta}\delta n_\beta) - 2\delta(\nabla_\alpha n^\alpha) + \nabla_\alpha n_\beta \delta g^{\alpha\beta} \tag{35}$$

Now define $\delta u^\alpha := \delta n^\alpha + g^{\alpha\beta}\delta n_\beta$. We find

$$n_\mu V^\mu = \nabla_\alpha \delta u^\alpha - 2\delta(\nabla_\alpha n^\alpha) + \nabla_\alpha n_\beta \delta g^{\alpha\beta} \tag{36}$$

And for the sake of bookkeeping, let's check back on the variation of the Einstein-Hilbert action:

$$16\pi\delta S_{EH} = \int_{\mathcal{M}} d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} + \int_{\partial\mathcal{M}} d^3y\varepsilon\sqrt{|h|}\left[\nabla_\alpha \delta u^\alpha - 2\delta(\nabla_\alpha n^\alpha) + \nabla_\alpha n_\beta \delta g^{\alpha\beta}\right] \tag{37}$$

5.2.2 Evaluating the first part of the boundary term: $\nabla_\alpha \delta u^\alpha$

First, we show that δu^α is orthogonal to n^α

$$\delta u^\alpha n_\alpha = (\delta n^\alpha + g^{\alpha\beta}\delta n_\beta)n_\alpha = (\delta n^\alpha)n_\alpha + (g^{\alpha\beta}\delta n_\beta)n_\alpha$$

$$= (\delta n^\alpha)n_\alpha + n_\alpha(g^{\alpha\beta}\delta n_\beta) = (\delta n^\alpha)n_\alpha + n^\beta\delta n_\beta$$

$$= n_\alpha \delta n^\alpha + n^\alpha \delta n_\alpha = \delta(n^\alpha n_\alpha) = 0 \tag{38}$$

Meaning δu^α lies entirely on the boundary, i.e. it's a tangent vector to the hypersurface Σ. This means we can use the definition (18) to find its divergence:

$$\nabla_\alpha \delta u^\alpha = D_\alpha \delta u^\alpha - \varepsilon a_\beta \delta u^\beta \tag{39}$$

We can rewrite the second term on the RHS as

$$a_\alpha \delta u^\alpha = a_\alpha(\delta n^\alpha + g^{\alpha\beta}\delta n_\beta)$$

$$= a_\alpha \delta n^\alpha + a^\beta \delta n_\beta$$

$$= a_\beta \delta n^\beta + a^\beta \delta n_\beta$$

$$= a_\beta \delta(n_\alpha g^{\alpha\beta}) + a^\beta \delta n_\beta$$

$$= a_\beta[n_\alpha \delta g^{\alpha\beta} + (\delta n_\alpha)g^{\alpha\beta}] + a^\beta \delta n_\beta$$

$$= a_\beta n_\alpha \delta g^{\alpha\beta} + 2a^\beta \delta n_\beta \tag{40}$$

Looking at the definition of the unit normal n^α, given by (6), we see that $n_\alpha \propto \Phi(X^\mu)_{,\alpha}$, so the variation δn_α can ultimately be written as $\delta n_\alpha = C(X^\mu)n_\alpha$, where $C(X^\mu)$ is a scalar function. This means

$$a^\beta \delta n_\beta \propto a^\beta n_\beta \tag{41}$$

Recalling $a_\beta := n^\mu \nabla_\mu n_\beta$, we find that (41) is zero. So the second term on the RHS of (40) vanishes:

$$a^\beta \delta n_\beta = 0 \tag{42}$$

Confusion

Actually, (42) doesn't make much sense to me, for several reasons:

- By the same logic, the entire expression $a_\alpha \delta u^\alpha$ should vanish, by inspection of the third line of (40). If $a^\beta \delta n_\beta = 0$ then surely $a_\beta \delta n^\beta$ must also be equal to zero?
- The *specific reason* why a_β is orthogonal to δn_β is not clear.
  - I first guessed that $a_\beta := n^\mu \nabla_\mu n_\beta$ is somehow proportional to the curl of n_β, but upon closer inspection that doesn't make sense because μ is not a free index.
  - The text explains "we used... the identity $a^\beta \delta n_\beta = 0$ due to the fact that a^β and n_α are orthogonal because using eq. (6) one can show that $\delta n_\alpha = C(X^\mu) n_\alpha$ where $C(X^\mu)$ is a scalar function"
    * It seems what they're saying is that $a_\beta := n^\mu \nabla_\mu n_\beta$ is orthogonal to n_α because the variance of n_α is proportional to n_α
    * Need some more time to convince myself why this is true

And (40) becomes:

$$a_\alpha \delta u^\alpha = a_\beta n_\alpha \delta g^{\alpha\beta} \tag{43}$$

Which gives the following expression for the divergence of δu^α

$$\nabla_\alpha \delta u^\alpha = D_\alpha \delta u^\alpha - \varepsilon n_\alpha a_\beta \delta g^{\alpha\beta} \tag{44}$$

And $n_\mu V^\mu$ becomes:

$$n_\mu V^\mu = D_\alpha \delta u^\alpha - 2\delta(\nabla_\alpha n^\alpha) + (\nabla_\alpha n_\beta - \varepsilon n_\alpha a_\beta)\delta g^{\alpha\beta} \tag{45}$$

5.2.3 Substituting the Extrinsic Curvature Tensor

It turns out that the coefficient in front of $\delta g^{\alpha\beta}$ in (45) is known as the *extrinsic curvature tensor*, which quantifies how the submanifold/hypersurface Σ bends and curves within 4D spacetime:

$$K_{\alpha\beta} \equiv \nabla_\alpha n_\beta - \varepsilon n_\alpha a_\beta \tag{46}$$

The following properties are declared without proof:

$$K_{\alpha\beta} = K_{\beta\alpha} \tag{47}$$

$$n^\alpha K_{\alpha\beta} = n^\beta K_{\alpha\beta} = 0 \tag{48}$$

$$K = \nabla_\alpha n^\alpha \tag{49}$$

Recall (10), which allows us to recover the metric from the induced metric and the unit normal, restated here in contravariant form:

$$g^{\alpha\beta} = h^{ab} e^\alpha_a e^\beta_b + \varepsilon n^\alpha n^\beta \equiv h^{\alpha\beta} + \varepsilon n^\alpha n^\beta \tag{50}$$

Combined with (48), this allows us to write:

$$\begin{aligned} K_{\alpha\beta}\delta g^{\alpha\beta} &= K_{\alpha\beta}\delta(h^{\alpha\beta} + \varepsilon n^\alpha n^\beta) \\ &= K_{\alpha\beta}\delta h^{\alpha\beta} \end{aligned} \tag{51}$$

$n_\mu V^\mu$ becomes

$$n_\mu V^\mu = D_\alpha \delta u^\alpha - 2\delta(\nabla_\alpha n^\alpha) + K_{\alpha\beta}\delta g^{\alpha\beta}$$

$$= D_\alpha \delta u^\alpha - 2\delta(\nabla_\alpha n^\alpha) + K_{\alpha\beta}\delta h^{\alpha\beta}$$

$$= D_\alpha \delta u^\alpha - 2\delta K + K_{\alpha\beta}\delta h^{\alpha\beta} \tag{52}$$

Where we used (49) in the last line of (52).

Then we employ the "well known relation for the variation of the square root of the determinant of the metric"

$$\delta\sqrt{|h|} = -\frac{1}{2}\sqrt{|h|}h_{\alpha\beta}\delta h^{\alpha\beta} \tag{53}$$

Side note

Maybe you should track this one down

Which allows us to write

$$\sqrt{|h|}n_\mu V^\mu = \sqrt{|h|}D_\alpha \delta u^\alpha - 2\sqrt{|h|}\delta K + \sqrt{|h|}K_{\alpha\beta}\delta h^{\alpha\beta}$$

$$= \sqrt{|h|}D_\alpha \delta u^\alpha - \delta(2\sqrt{|h|}K) - \sqrt{|h|}(Kh_{\alpha\beta} - K_{\alpha\beta})\delta h^{\alpha\beta} \tag{54}$$

And we get this final expression for the integrand of our boundary term in the variation of the Einstein-Hilbert action

$$\varepsilon\sqrt{|h|}n_\mu V^\mu = \varepsilon\left[\sqrt{|h|}D_\alpha \delta u^\alpha - \delta(2\sqrt{|h|}K) - \sqrt{|h|}(Kh_{\alpha\beta} - K_{\alpha\beta})\delta h^{\alpha\beta}\right] \tag{55}$$

5.3 Final expression for the variation of the Einstein-Hilbert Action

Recall the expression for the variation of the Einstein-Hilbert variation given by (27), restated here:

$$16\pi\delta S_{EH} = \int_{\mathcal{M}} d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} + \int_{\partial\mathcal{M}} d^3y\varepsilon\sqrt{|h|}n_\mu V^\mu \tag{56}$$

Substituting our expression for the boundary term, (55), this becomes

$$16\pi\delta S_{EH} = \int_{\mathcal{M}} d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} \tag{57}$$

$$- \int_{\partial\mathcal{M}} d^3y\varepsilon\sqrt{|h|}(Kh_{\alpha\beta} - K_{\alpha\beta})\delta h^{\alpha\beta} \tag{58}$$

$$- \int_{\partial\mathcal{M}} d^3y\varepsilon\delta(2\sqrt{|h|}K) \tag{59}$$

$$+ \int_{\partial\mathcal{M}} d^3y\varepsilon\sqrt{|h|}D_\alpha \delta u^\alpha \tag{60}$$

This is of the form

$$\delta S_{EH} = \int_{\mathcal{M}} d^4x(\text{Equation of Motion Term})\delta(\text{Dynamical Variable}) \tag{61}$$

$$+ \int_{\partial\mathcal{M}} d^3x(\text{Conjugate Momentum})\delta(\text{Variables to be fixed}) \tag{62}$$

$$+ \int_{\partial\mathcal{M}} d^3x\delta(\text{Boundary Term}) \tag{63}$$

$$+ \int_{\partial\mathcal{M}} d^3x(\text{Total Divergence Term}) \tag{64}$$

5.4 Imposing a fixed induced metric on the boundary

Setting all of the boundary terms below (57) equal to zero would again implicitly assume some boundary conditions. Instead, we can impose that the induced metric $h_{\alpha\beta}$ is *fixed* on the boundary, i.e. that

$$\delta h^{\alpha\beta}|_{\partial\mathcal{M}} = 0. \tag{65}$$

This automatically takes care of the second term, (58), but how about the fourth term?

If we employ Stokes' theorem, (22), then the fourth term, (60), "becomes a boundary term on the two dimensional boundary $\partial^2\mathcal{M}$ and is usually ignored" $\Longleftarrow$ Don't ask me why.

Complaint

We came all this way just to hand-wave away one of three terms we're trying to get rid of?

I *would* like to track this down but I don't know where else this derivation exists.

Side note

This also seems to answer an earlier question of mine, in the box labeled "Digression/Questions" on page 2: apparently a boundary *can* have a boundary.

5.5 Introducing the Gibbons-Hawking-York Boundary Term

And so, with the second and fourth terms in the variation of the Einstein-Hilbert action taken care of, all that remains is the third term, (59).

To tackle this, let's assume that the action is *actually* given by

$$S = S_{EH} + S_{GHY} + S_0 \tag{66}$$

where S_0 is a constant of integration and S_{GHY} is the GIBBONS-HAWKING-YORK BOUNDARY TERM

$$S_{GHY} := \frac{1}{16\pi}\int_{\partial\mathcal{M}} d^3y\varepsilon\sqrt{|h|}2K \tag{67}$$

At last.

Defining the action in this way, we find that the variation of the action is indeed

$$16\pi\delta S \equiv 16\pi\delta(S_{EH} + S_{GHY}) = \int_{\mathcal{M}} d^4x\sqrt{-g}G_{\mu\nu}\delta g^{\mu\nu} \tag{68}$$