Why is the Generalized Second Law True?

by Aron Wall,
UC Santa Barbara

(based on arXiv:1105.3445, "A proof of the generalized second law for rapidly changing fields and arbitrary horizon slices")

This work was aided by discussions with
Ted Jacobson, Rafael Sorkin, and William Donnelly.

The Second Law of Thermodynamics

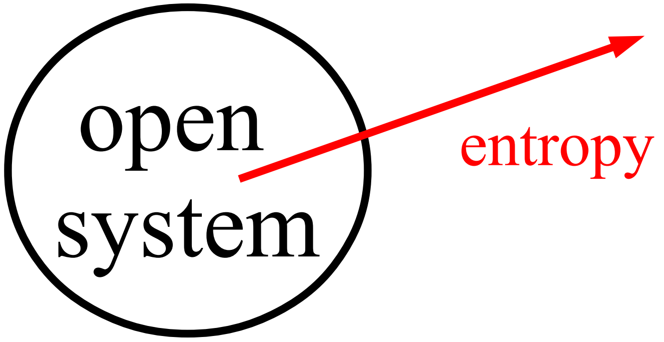
$$\frac{dS}{dt} \geq 0$$

for any CLOSED system.

The Second Law of Thermodynamics

$$\frac{dS}{dt} \geq 0$$

for any CLOSED system.

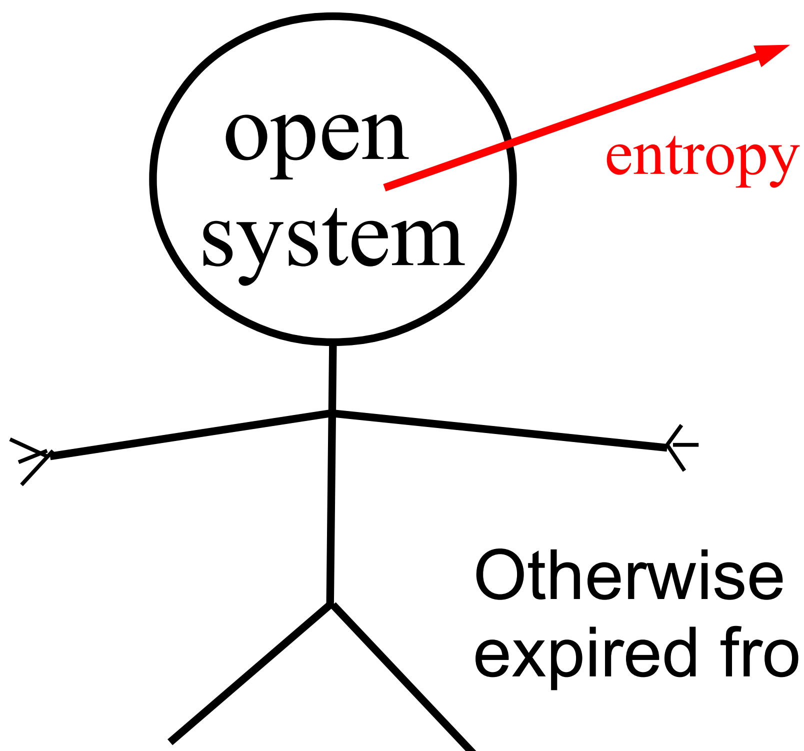


In an OPEN system entropy can normally just exit, lowering S(system)

The Second Law of Thermodynamics

$$\frac{dS}{dt} \geq 0$$

for any CLOSED system.



In an OPEN system entropy can normally just exit, lowering S(system)

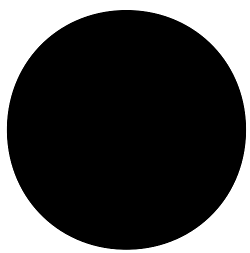
Otherwise we would already have expired from personal heat death...

Causal Horizons

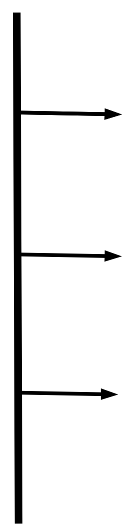
Causal horizon = boundary of past of any future infinite worldline.
(an "observer")

"Outside" means the side with the worldline.

Examples:



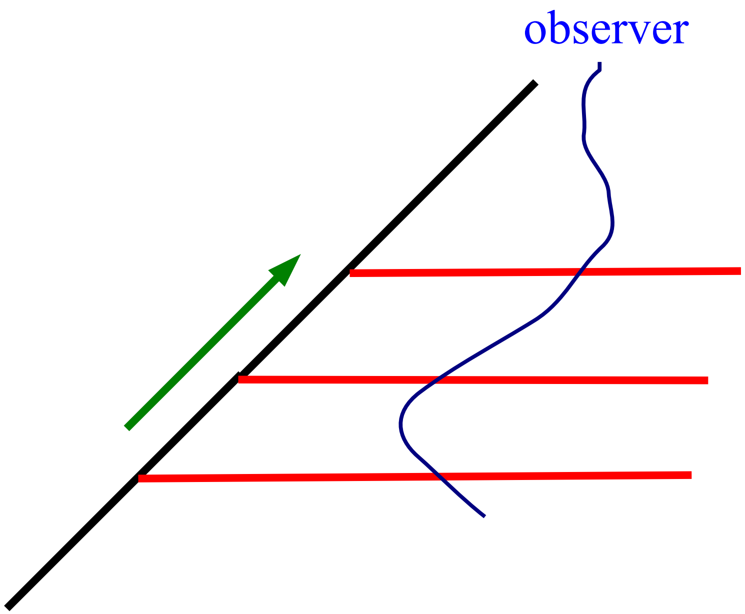
exterior of
black hole



Rindler-like horizons
of accelerating observers

Horizon Thermodynamics

The outside of a horizon is an OPEN system—
info can leave (but not enter).



But the generalized entropy

$$S_{\text{gen}} = \frac{A}{4\hbar G} + S_{\text{out}}$$

still increases. Area A of horizon contributes to entropy.

$$\frac{dS_{gen}}{dt} \geq 0$$

Generalized Second Law (GSL).

Why study the GSL?

Thermodynamics
of Matter

Fact

Statistical Mechanics
of Atoms

Analogy

Thermodynamics
of Horizons

Hope

Statistical Mechanics
of Spacetime
(i.e. quantum gravity!)

My view: Too early to take sides about correct quantum gravity theory.

I want to find out what horizon thermodynamics implies about micro degrees of freedom without being constrained to a specific model.

Semiclassical approximation

1. Pick a classical background and look at QFT state ρ.

2. Expand out metric in powers of $\hbar$:

$$g_{ab} = g^{0}_{ab} + g^{1/2}_{ab} + g^{1}_{ab} + \mathcal{O}(\hbar^{3/2})$$

classical background metric

quantized gravitons

gravitational fields of matter/gravitons calculated using expectation of Einstein equation

$$\langle G_{ab} \rangle = 8\pi G \langle T_{ab} \rangle$$

3. See if the generalized entropy of the state ρ increases:

$$\frac{dS_{\text{gen}}}{dt} = \frac{d}{dt}\left[\frac{\langle A \rangle}{4\hbar G} - \text{tr}(\rho \ln \rho)\right] \geq 0$$

Entanglement entropy divergence must be renormalized!
This will be done by considering only entropy *differences*.

Why try to prove the semiclassical GSL?

1) If we know *why* the GSL is true semiclassically,
may be able to figure out what is required at the Planck scale.

2) We want to know *when* the GSL applies:

* Can it be proven in a way which is local on the horizon?

* Does it hold in a differential form at each horizon point,
or only globally for stationary —> stationary processes?

* What does the GSL assume (if anything) about
matter (e.g. number of fields) and gravity (e.g. derivative couplings)

* Is there a generalization to null surfaces which aren't horizons?

3) We understand QFT and GR better than quantum gravity.

Don't Read This Slide

Instead read “Ten Proofs of the Generalized Second Law” (09):
A.C. Wall, arXiv:1007.1493, JHEP 06 (2009) 021

PROOF BY	REGIME	PERTURB.	EXTRA CONDITIONS / DIFFICULTIES
Hawking	classical	any	null energy condition, cosmic censorship
Zurek & Thorne	semiclassical	quasi-steady	entropy localization, renormalizability
Wald	hydrodynamic	quasi-steady	adiabaticity (fixable)
Frolov & Page	semiclassical	quasi-steady	CPT insufficient for charged BH (fixable)
Sorkin 1	quantum gravity	any	inconsistent assumptions
Sorkin 2	semiclassical	quasi-steady	thermality undefined, not superradiant
Mukohyama	semiclassical	quasi-steady	free scalar field, not superradiant
Flanagan et al.	hydrodynamic	any	null energy condition, Bekenstein bound
Bousso et al.	hydrodynamic	any	entropy gradient, isolation assumption
Fiola et al.	semiclassical	any	RST model, large N, apparent horizon

classical: assumes a classical background and neglects the outside entropy term.
hydrodynamic: the entropy is approximated by a fully localizable four-vector.
semiclassical: neglects fluctuations in the metric.
quasi-steady: a small, slowly changing perturbation to a stationary black hole.

Quasi-Steady Approximation

Before my work, the semiclassical proofs only showed either that

1. $\Delta S_{\text{gen}} \geq 0$, where the fields may be rapidly varying but the entropy is only shown to increase from the asymptotically infinite past to the asymptotically infinite future, OR

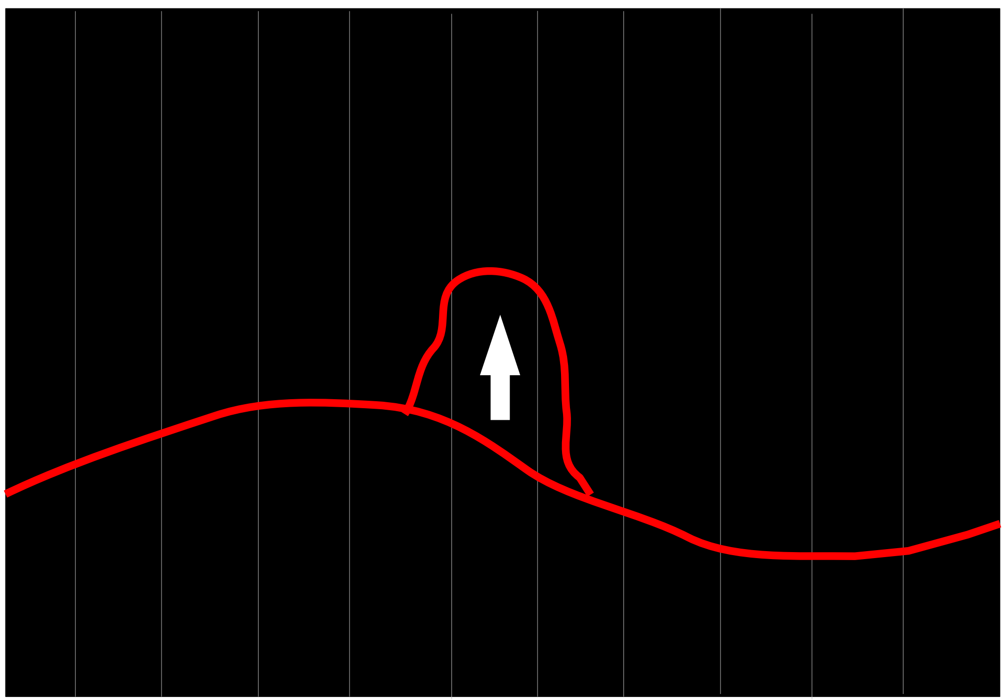
2. $\delta S \geq 0$, where the fields are in a nearly steady state
   (this can be obtained from case 1 by linear interpolation)

Either choice permits use of the "First Law": $dM = \frac{\kappa}{8\pi} dA + \Omega dJ$

but then one can't check whether the GSL holds at each instant of time—need to check this with rapidly changing fields. E.g. throwing a tea-cup into a black hole. First Law not useful in this case.

If we want a differential form of the GSL, we have to be able to show that the entropy can increase for arbitrary initial and final slices:

Then $\Delta S \geq 0 \longrightarrow \delta S_{\mathrm{gen}} \geq 0$



A section of a horizon with horizon generating lightrays shown in grey and slices shown in red.

A *local* version of GSL, if true, gives hints about the *local* statistical mechanics of spacetime. Maybe constrains microscopic QG physics?

Proof of the GSL for rapidly changing fields

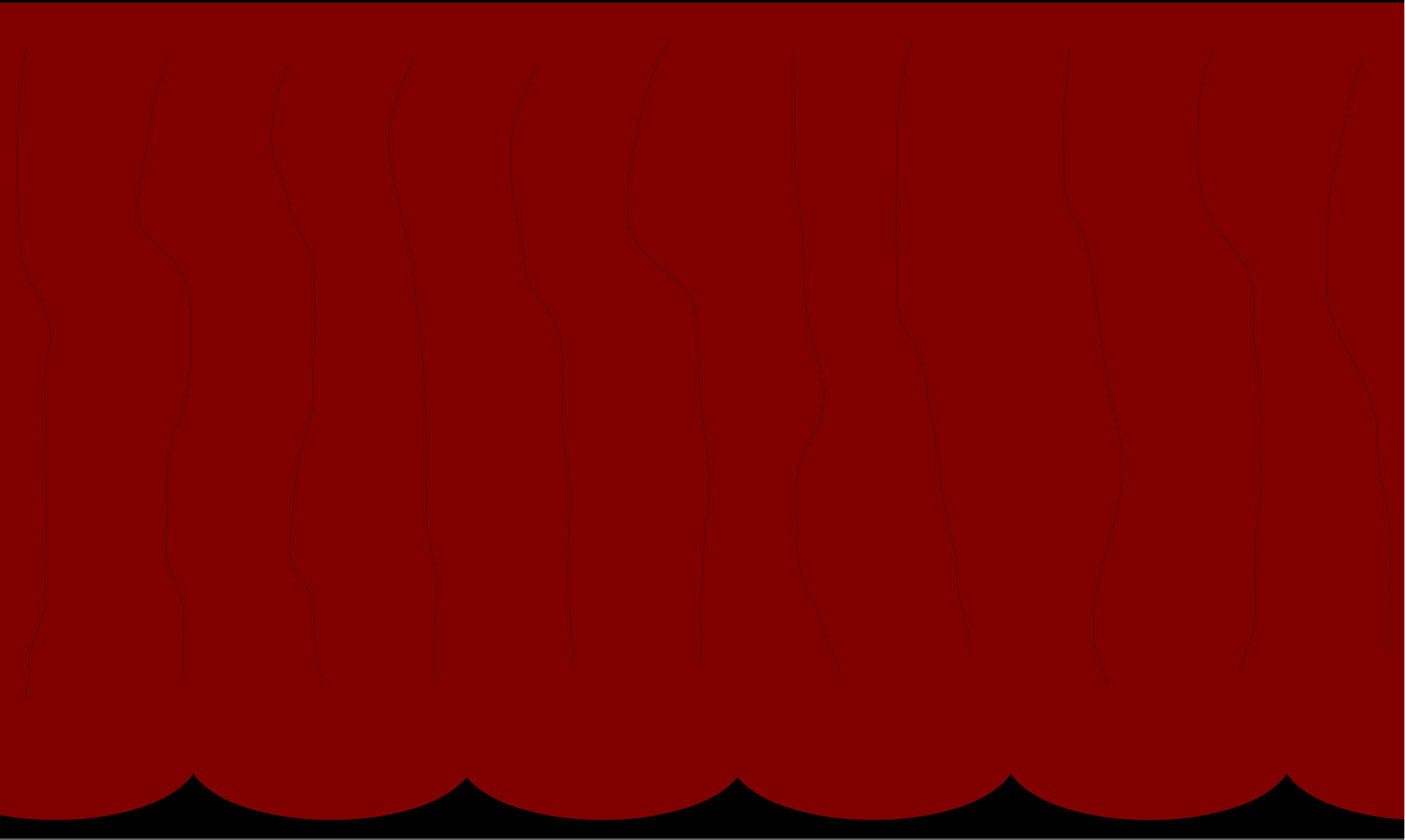
1. Applies to rapidly evolving semiclassical perturbations to any stationary background horizon (e.g. Rindler, de Sitter, Kerr).

2. Proves $\Delta S_{\text{gen}} \geq 0$ for arbitrary initial and final slices of the horizon.

3. Works for free fields of any spin. Can also accommodate certain superrenormalizable interactions.

END OF PART 1.

PART 2 WILL EXPLAIN THE PHYSICAL REASONING USED TO PROVE THAT BLACK HOLES OBEY THE SECOND LAW...



INTERMISSION

Proof of the GSL for rapidly changing fields

1. Applies to rapidly evolving semiclassical perturbations to any stationary background horizon (e.g. Rindler, de Sitter, Kerr).

2. Proves $\Delta S_{\text{gen}} \geq 0$ for arbitrary initial and final slices of the horizon.

3. Works for free fields of any spin. Can also accommodate certain superrenormalizable interactions.

start with special case: Rindler horizons.

Rindler GSL proof summary

"A proof of the generalized second law
for rapidly evolving Rindler horizons"
arXiv:1007.1493

Basic ideas:

1. Relate generalized entropy to free (boost) energy in wedge.
2. Relate free energy to a quantity known as "relative entropy"
3. Apply theorem that says relative entropy can't increase.

free boost energy

relative entropy

$$S_{\text{gen}} = (TS_{\text{out}} - K)/T = -S(\rho\,|\,\sigma)$$

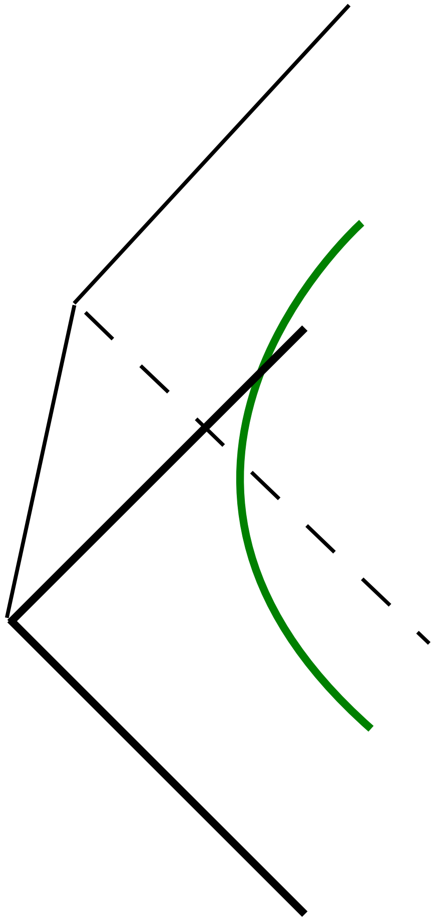
(up to additive constants)

Previous proofs of the GSL implicitly used the concept of relative entropy (Casini 08).

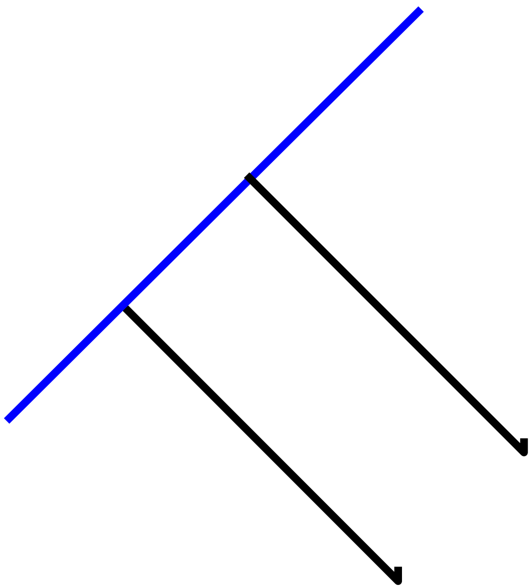
Rindler Wedges

A Rindler wedge is the intersection of the past & future of uniformly accelerating worldline.

1-parameter family of Rindler wedges share same future horizon & fit inside each other.



perspective drawing of wedge
& accelerating observer

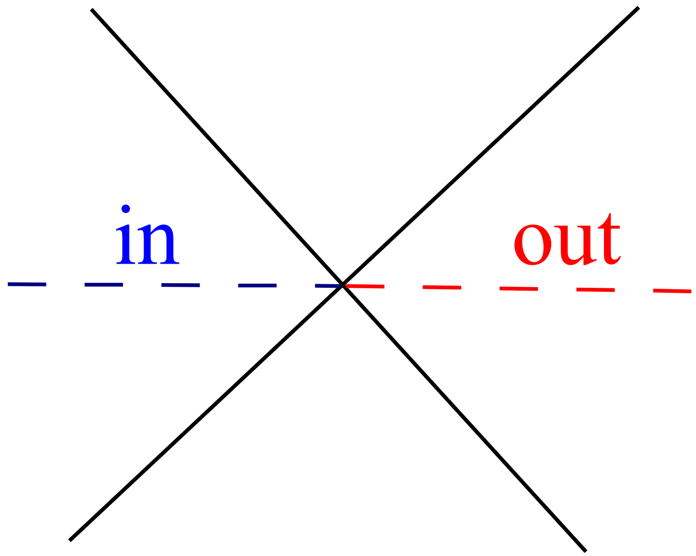


Go from bigger wedge to smaller wedge by restriction..

Rindler Wedges are Thermal

QFT vacuum always KMS (i.e. thermal) in boost energy *K* when restricted to Rindler wedge, at temperature $T = \hbar/2\pi$.

Consequence of the wedge's *boost symmetry* (Bisognano-Wichmann 75).



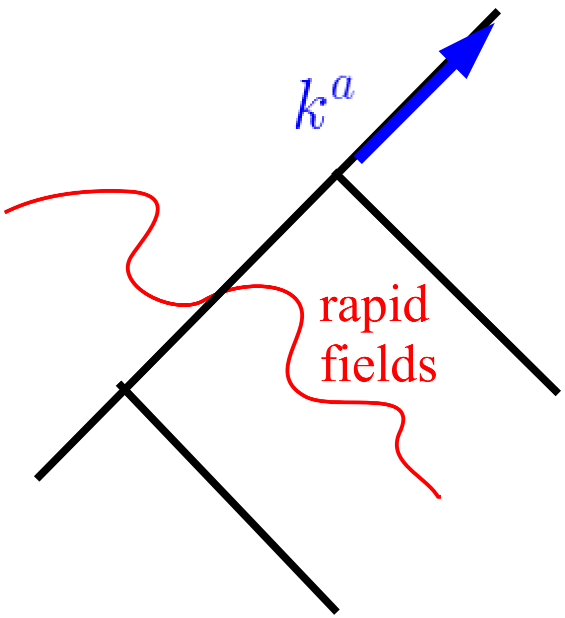
$$\mathrm{tr}_{\mathrm{in}}(|0\rangle\langle 0|) \propto e^{-2\pi K_{\mathrm{out}}/\hbar}$$

(formally) where the boost Killing energy on a slice Σ is:

$$K_{\mathrm{out}} = \int_{\mathrm{out}} T_{ab}\,\xi^a d\Sigma^b$$

So the vacuum is thermal. Next we will perturb it with quantum fields.

Area deficit ~ Boost Energy



Along each horizon generating lightray, the Raychaudhuri & Einstein equations hold:

$$\dot{\theta} = -\theta^2/2 - \sigma_{ab}\sigma^{ab} - 8\pi G T_{kk}$$

where $\theta = (1/A)(dA/d\lambda)$ is the expansion w.r.t. an affine parameter λ.

Linearize and integrate to get expression in terms of $K(\lambda)$, the boost energy of the wedge, up to constants.

$$A(\lambda) = A(\infty) - 8\pi G \int_\lambda^\infty T_{kk}(\lambda' - \lambda) d\lambda' = -8\pi G[K(\lambda) - \underbrace{K_{\rm rad}] + A(\infty)}_{\text{constants}}$$

Horizon area canonically conjugate to boost time: generalizes Carlip & Teitelboim (95), Massar & Parentani (00) to dynamical situations.

Relative Entropy

Information theory property of two mixed states ρ and σ.

$$\boxed{S(\rho\,|\,\sigma) = \mathrm{tr}(\rho \ln \rho) - \mathrm{tr}(\rho \ln \sigma)}$$

(definition can be extended to arbitrary algebras of observables)

Properties:

* Range is $[0, +\infty]$. Finite for nice enough states (no renormalization).

* $S(\rho\,|\,\rho) = 0$

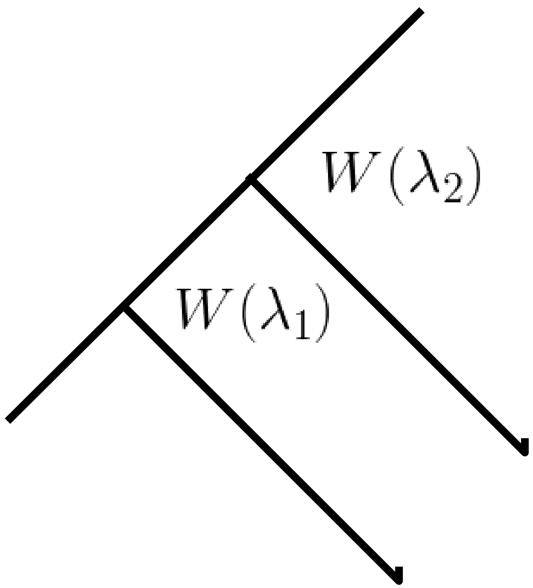
* If σ is a KMS (thermal) state, proportional to free energy difference:

$$S(\rho\,|\,\sigma) = [T^{-1}E - S]_\rho - [T^{-1}E - S]_\sigma$$ (Araki & Sewell 77)

* Monotonicity: Always nonincreasing under restriction to subsystems:

$$S(\rho\,|\,\sigma)_M \geq S(\rho\,|\,\sigma)_{M'}$$ when $M' \subset M$ (Araki 75)

Proof of the Rindler GSL



Let ρ be the state we are interested in proving the GSL for.

Let σ be the Minkowski vacuum state.

Since σ is thermal in each wedge, $S(\rho\,|\,\sigma)$ is the free boost energy up to terms constant in each wedge:

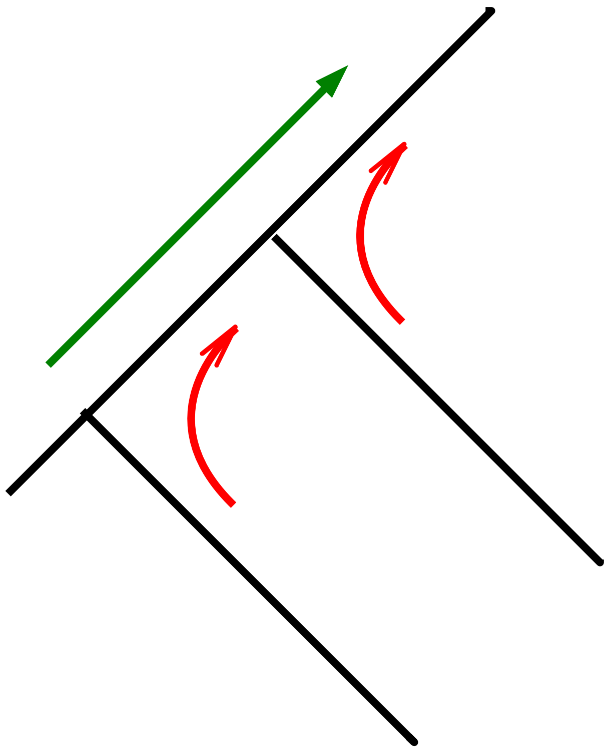
$$S(\rho\,|\,\sigma) = [\tfrac{2\pi}{\hbar}K - S_{\mathrm{out}}]_\rho = -\frac{A}{4\hbar G} - S_{\mathrm{out}} = -S_{\mathrm{gen}}$$

$S_{\mathrm{out}} = S_\rho - S_\sigma$ is the *renormalized* entropy.

Relative entropy is monotonic under restriction, so the GSL holds!

The GSL comes from Horizon Symmetry

Rindler Symmetry

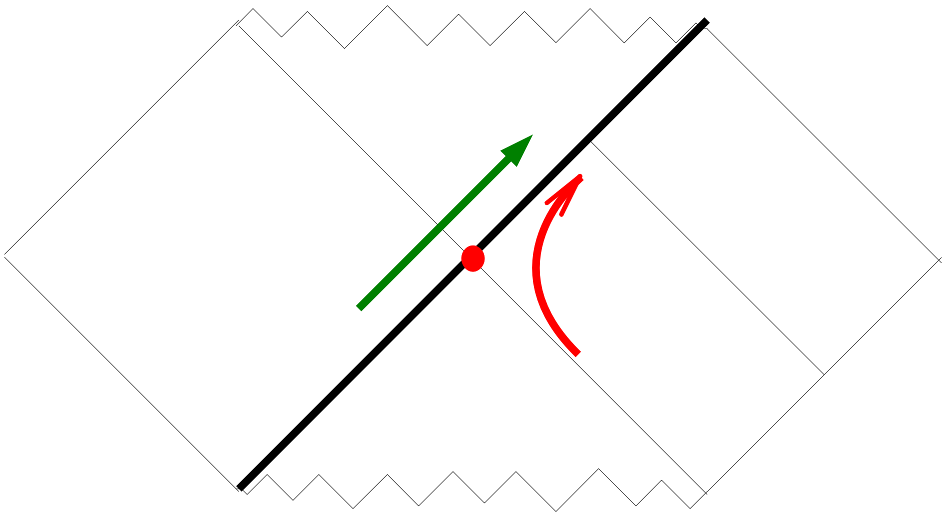


Argument just given requires each wedge to have a boost symmetry so that the vacuum state σ is thermal.

Commutator of two boosts is a null translation symmetry. Vacuum state σ invariant under this too.

Rindler horizon invariant under 2d Lie group.
That's why it works.

Black Holes have less symmetry



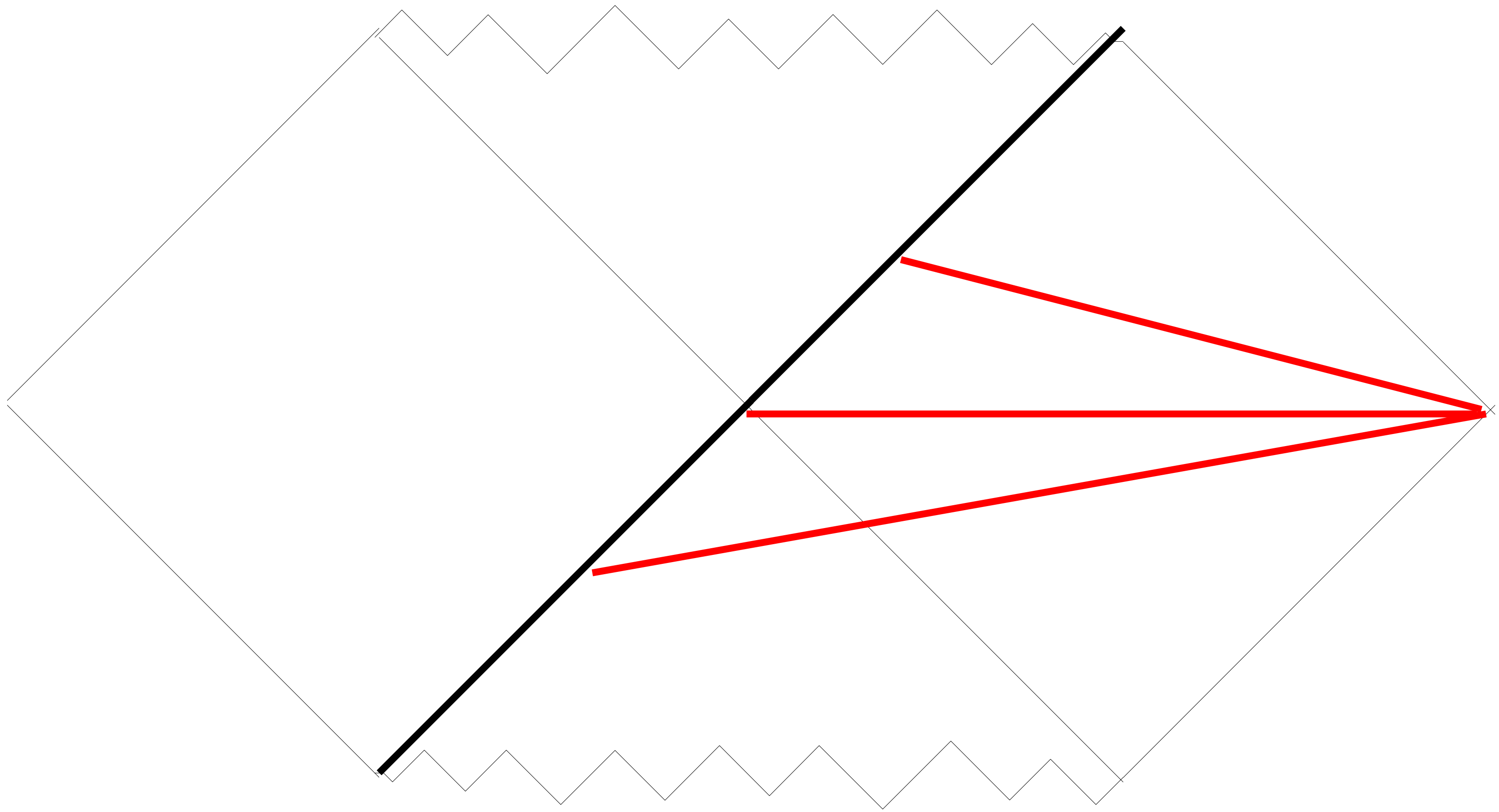
Spacetime has a Killing boost symmetry only about the bifurcation surface.

No null translation Killing symmetry.

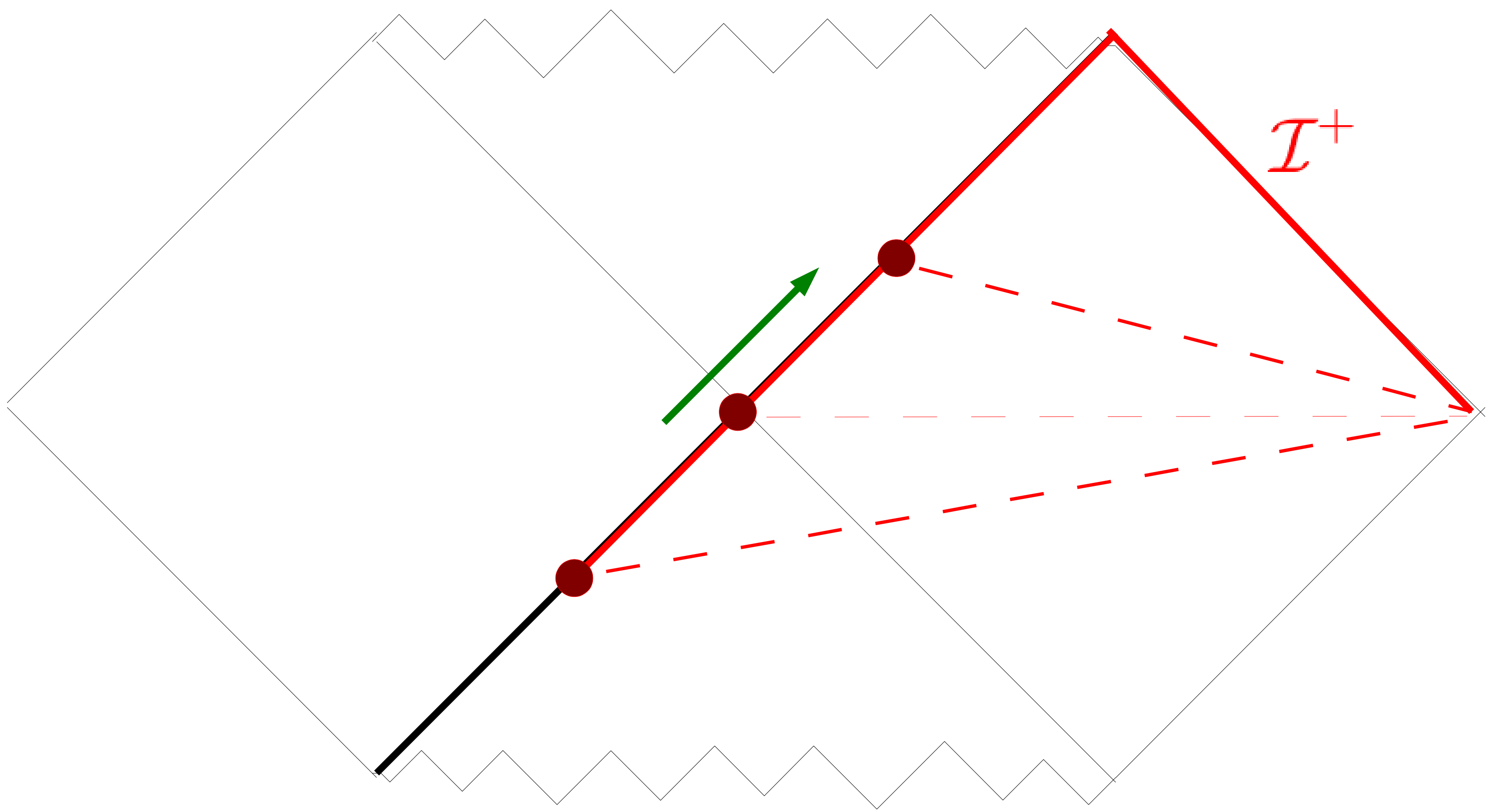
Kerr is even worse because no thermal Hartle-Hawking state exists at all (angular momentum is unbounded below).

Proof does not work—take near-horizon limit?

Instead of using spacelike slices:



Push forward to the horizon itself



The horizon has translation symmetry,
even though the full spacetime does not.

Restrict fields to horizon algebra

Possible to restrict free fields operators to the event horizon itself.
Tricky since fields must be smeared only across the horizon.
One finds that:

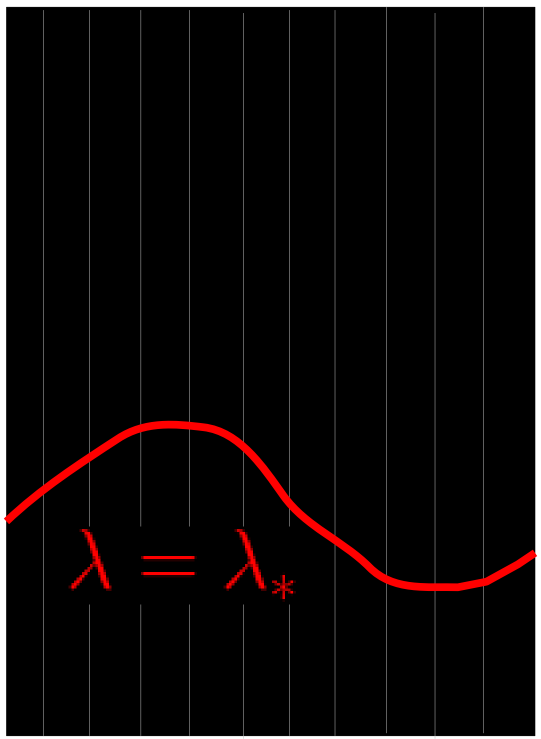
* Φ can't be restricted, but $\nabla_k \Phi$ can be. Normally derivatives hurt, but the field is already smeared in the k direction, and null mass shell tells us that $\nabla_+ \Phi \sim p_+ \Phi \sim \frac{1}{p_-} \Phi$ so it actually helps.

* The horizon algebra is ultralocal; each horizon generator is independent.

* There is an infinite dimensional symmetry group: translations and dilations of each horizon generator *independently*. (boosts = dilations on the horizon)

* Can accommodate arbitrary (nonderivative) potentials $V(\Phi)$ or Yang Mills at the level of naïve Fock space perturbation theory. Does not affect horizon algebra or null energy. But the RG flow may introduce derivative couplings, unless theory is superrenormalizable.

Arbitrary horizons, arbitrary slices:



a piece of a horizon, with grey generators

Because each horizon generator can be independently translated, can translate to wiggly slices.

Can define a canonical vacuum state w.r.t. all null translation symmetries (Sewell 81).

σ is KMS above *any* slice w.r.t. dilations about that slice. Formally,

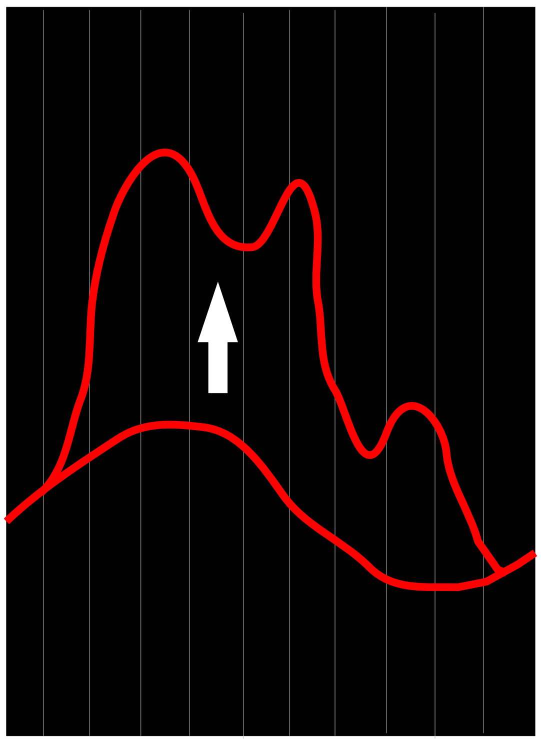
$$\sigma_{\lambda>\lambda_*} = e^{-2\pi K(\lambda_*)/\hbar}$$

where

$$K(\lambda_*) = \int_{\lambda_*}^{\infty} T_{kk}(\lambda - \lambda_*)\, d\lambda$$

This works on any background with a stationary horizon even when no Hartle-Hawking state can be defined on the bulk spacetime (e.g. Kerr).

Adapting the Proof of the GSL



a piece of a horizon, with grey generators

GSL can now be proven analogously
to Rindler case for semiclassical perturbations
to any stationary horizon:

Let ρ be the bulk state we are interested in,
Let σ be any bulk state which restricts to
the vacuum state on the horizon, and is
otherwise arbitrary. For each slice:

* σ is thermal with respect to $K(\lambda_*)$, which is proportional to the area A of the slice λ_*.

* $S(\rho\,|\,\sigma) = -S_{\text{gen}}$ up to additive constants.

* Thus the GSL holds by monotonicity of relative entropy.

Black Hole Microstates

Why is the GSL true in full quantum gravity?

1. How does this relate to quantum gravity statistical mechanics?
Two different ways to approach issue:

A) Look for this infinite dimensional symmetry of the horizon in a quantum gravity theory to get the GSL for similar reasons,

Or try to reverse the argument...

B) Find some other way to derive GSL in quantum gravity, and then try to see Lorentz symmetry emerge from it.

2. Can (interacting) gravity be quantized directly on the horizon?
An ultraviolet scaling limit...
Poisson brackets of GR have been analyzed (Reisenberger 08).

Quantum Gravity Microstates

What does this tell us about microstates in quantum gravity?

To explain the GSL in terms of a statistical mechanics, one must:

1. Identify the degrees of freedom whose state counting gives rise to the area term,

2. Identify what kind of coarse-graining (if any) is needed to get the semiclassical generalized entropy,

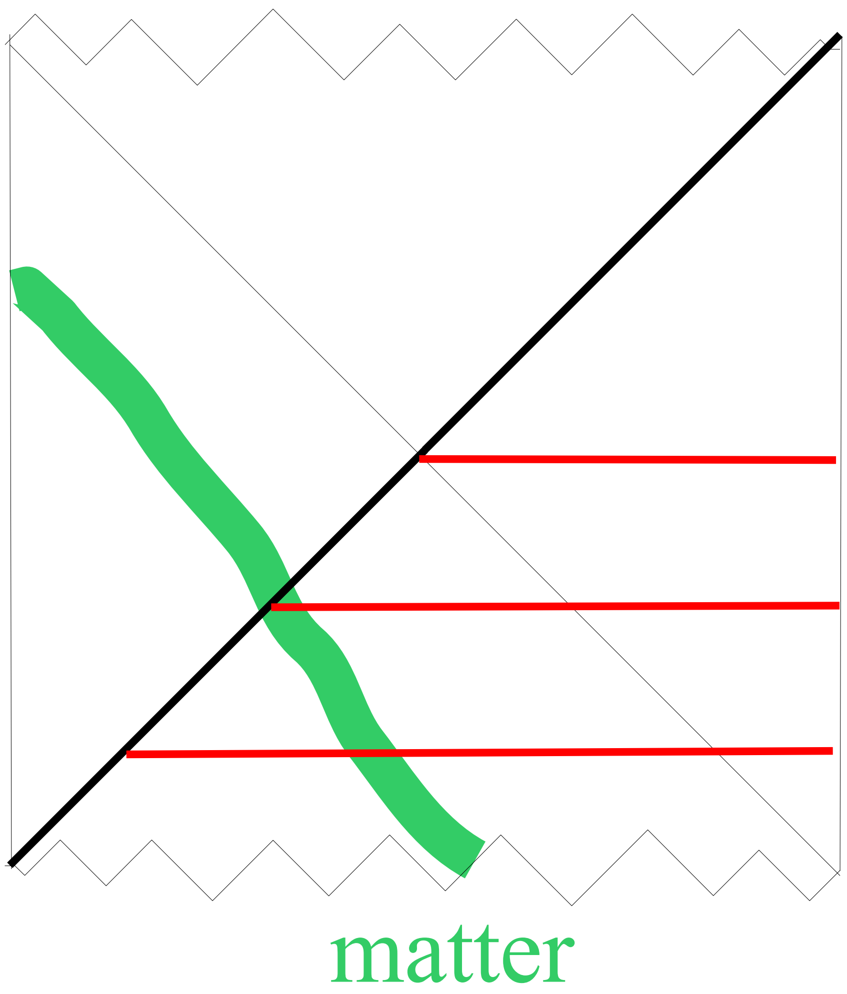
3. Explain why this entropy *increases*, for causal horizons, &

4. Explain why it can *decrease*, for null surfaces that aren't horizons.

If info is lost across horizons, (2) can be trivial but (3) is surprising!

If info is preserved outside, (3) seems easy, but (2) and (4) are hard!

Holographic Principle does not fully explain GSL

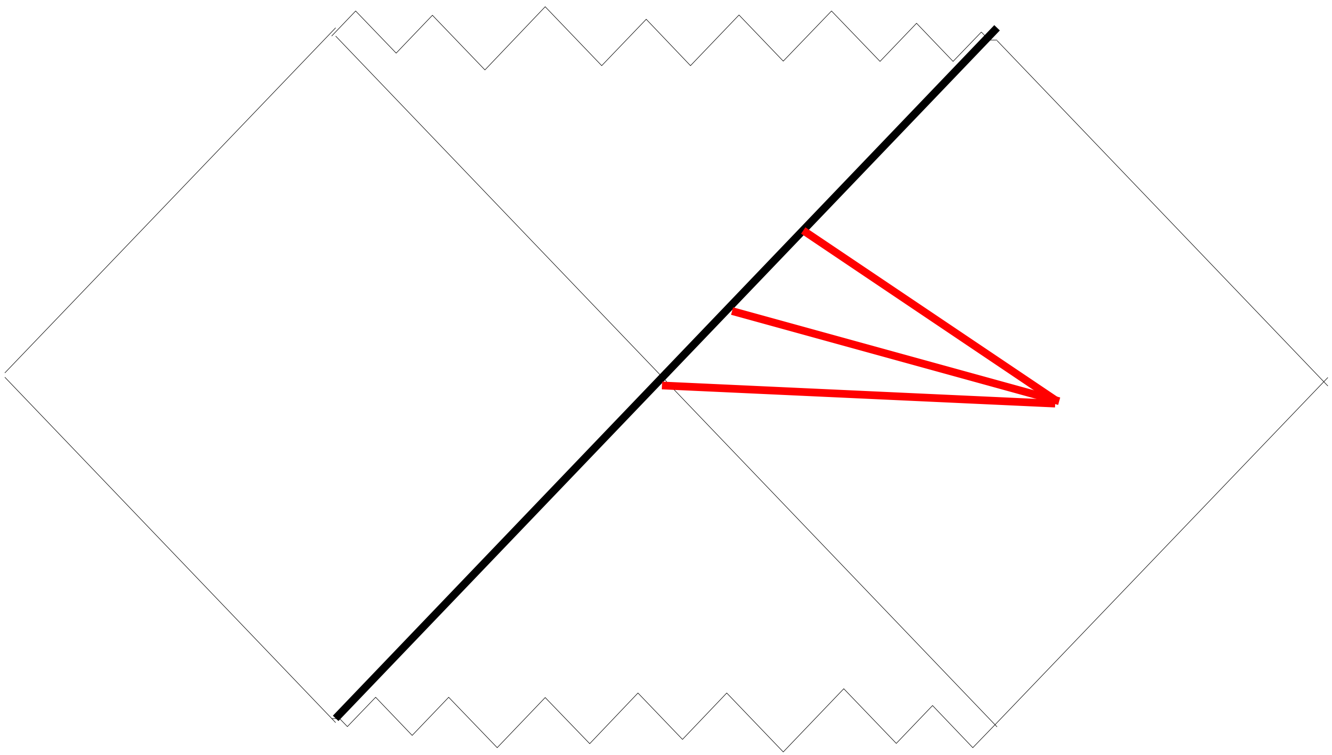


Even in the context of AdS/CFT,
info can be lost across a future
horizon if it falls across
prior to a past horizon.
(info is in the other CFT)

And yet,
the generalized entropy
is still increasing!

Why?

Not because info gets “stuck on the horizon”.



Causal horizon on left-hand side of incomplete slices.
No info escapes on right-hand side since slices meet.

And yet, generalized entropy is decreasing in Hartle-Hawking state!

(inconsistent with firewalls/fuzzballs?)

Conclusions

1. The semiclassical GSL holds because the horizon has more symmetry than the rest of the spacetime.

2. The proof works for free or superrenormalizable matter theories, but rigorous interactions may require a more delicate near-horizon limit.

3. Except in the case of nested Rindler wedges, where arbitrary interactions may be accommodated.