Black Holes and Thermodynamics

Stefan Hollands

Leipzig University

IHP Paris, 24 September 2015

European Research Council

Established by the European Commission

Supported by ERC-grant QC&C 259562

Outline

General relativity

Phase space of GR

It is fair to say that 100 years after the invention of GR, its phase space is still poorly understood. One knows e.g.

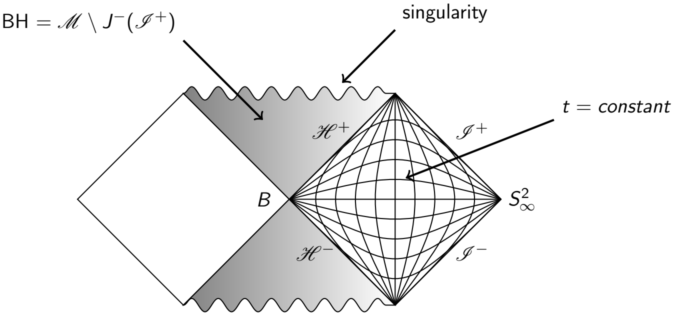
1. Some static/stationary solutions
2. Some dynamical solutions with high symmetry
3. The vicinity (in phase space) of such solutions
4. Restricted dynamical setups in which singularities can form

Stationary solutions

Stationary solutions usually describe black holes, or other types of black objects, such as black branes, broadly meaning objects with non-compact event horizons and BCs different from asymptotic flatness.

4*d* Schwarzschild (no rotation)

$$g = -\left(1 - \frac{r_0}{r}\right) dt^2 + \left(1 - \frac{r_0}{r}\right)^{-1} dr^2 + r^2 \, d\sigma^2_{S^2} \ .$$



Q: Are such solutions stable in exterior region? (This talk)

Higher dimensional BH's/branes etc.:

This talk:

The question of stability is of obvious physical relevance in $d = 4$. Interesting also for higher dimensional spacetimes (e.g. $d = 10$) and unconventional boundary conditions (e.g. "AdS"-type).

Motivations:

- Unified theories often require higher dimensions
- Holographic descriptions of strongly coupled CFTs require AdS-type BCs
- Interesting mathematical questions relating geometric analysis, PDEs, numerics

What is different?

Unusual boundary conditions can lead to unexpected new instabilities: Recent example: "Turbulent instability of AdS-spacetime" (aka "Gravity in a box") [Bizon & Rostoworski]

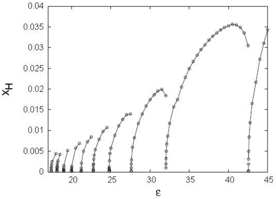


FIG. 1: Horizon radius vs amplitude for initial data (9). The number of reflections off the AdS boundary before collapse varies from zero to nine (from right to left).

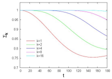
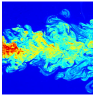


FIG. 3: Evolution of the fraction of the total energy contained in the first $(k+1)$ modes $\Sigma_k := \frac{1}{M}\sum_{j=0}^{k} E_j$ for the two-mode initial data $\phi(0,x) = \varepsilon\left(\frac{1}{d_0}e_0(x) + \frac{1}{d_1}e_1(x)\right)$ with $\varepsilon = 0.088$.



What is different?

Higher dimensions lead to zoo of unexpected new solutions such as black rings, saturns, ...

[Emparan-Reall, Pomeranski-Senkov,

Elvang-Figueras,... (many people)]

Real black saturn as photographed by NASA

Example: d-dimensional Schwarzschild

$$g = -\left[1 - \left(\frac{r_0}{r}\right)^{d-3}\right] dt^2 + \left[1 - \left(\frac{r_0}{r}\right)^{d-3}\right]^{-1} dr^2 + r^2 \, d\sigma^2_{S^{d-2}} \, .$$

Stability of black holes

Stability

It is natural to ask how the existence of higher dimensions affects the stability properties of black objects.

In fact, there exist many different approaches to analyze the (in)stability of a black hole or a black 'brane':

Different methods

- "Direct approach" (analytical/numerical treatment of linearized field equations)
- "Thermodynamic stability" (entropic/thermodynamic considerations)

These methods have a rather different flavor but they are related to each other, as I will explain.

Direct method, dynamical stability

Consider perturbations off a given black hole spacetime background g_{ab} which is a vacuum solution. Write the perturbed metric as $g_{ab}(\lambda) = g_{ab} + \lambda \delta g_{ab} + O(\lambda^2)$, and impose the vacuum Einstein equations. Then the (linear) perturbation has to satisfy

$$0 = \nabla^c \nabla_c \delta g_{ab} + (\text{other terms})$$

which is a "wave equation".

Definition dynamical instability

A background is said to be dynamically unstable if there are perturbations which do not settle down to a pure gauge transformation, or to a perturbation towards another black hole.

Dynamical stability

These issues are non-trivial already for Schwarzschild background:

- Mode analysis/decoupling [Regge-Wheeler, Zerilli, Moncrief, Ishibashi-Kodama, ...]
- Uniform decay (4d scalar case)
  [Kay-Wald,Andersson-Blue,Dafermos-Rodnianski, ...]

In the 4d Kerr background, one has Teukolsky formalism [Teukolsky], and mode-by-mode stability has been estabilshed [Whiting], but

- No formalism of comparable power known for more general Einstein-matter systems.
- No similar formalism in higher d (e.g. Myers-Perry BH).
- Decoupling/separation of variables has been achieved only in highly special cases.

Thermodynamical method

Stationary black holes/branes are known to have mathematical properties that are analogous to the ordinary laws of thermodynamics of laboratory-type systems:

Laws of BH mechanics

1. Thermodynamic equilibrium $\leftrightarrow$ stationary BH.
2. Temperature $\leftrightarrow$ surface gravity $\kappa =$ constant over $\mathscr{H}$.
3. Energy $\leftrightarrow$ mass $m = r_0/2$, entropy $\leftrightarrow$ $A = \mathrm{Area}(B) = 4\pi r_0^2$,

$$\delta m = \frac{\kappa}{8\pi}\delta A - \sum \Omega^I \delta J_I$$

   where Ω^I is the angular velocity and J_I the angular momentum associated to I-th rotational symmetry of g_{ab}.
4. Entropy (area) never decreases in any physical process.
5. Thermodynamical (in)stability $\Rightarrow$ dynamical instability???

Thermodynamical (in)stability

Consider laboratory system in equilibrium characterized by state parameters (energy, charge,...).

Suppose I can increase the entropy keeping conserved quantities fixed. $\Longrightarrow$ Thermodynamical instability!

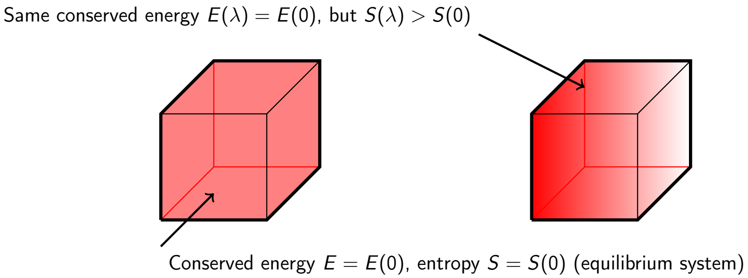


Figure: A homogeneous equilibrium system (left), and an unstable finite inhomogeneous perturbation.

Thermodynamical (in)stability

In BH-context, state parameters are $m, J_I, \ldots$.

Definition thermodynamical (in)stability

A homogeneous system is a BH with non-compact spatial isometry. A system is said to be dynamically unstable if there exists a perturbation such that $\delta^2 A > 0$ for $m(\lambda) = J_I(\lambda) = \cdots = const.$ to second order.

Note: By first law,

$$\delta^2 A \equiv \frac{d^2}{d\lambda^2} A(0)$$

does not depend on particular choice of second order perturbation $\delta^2 g_{ab}$! For homogeneous black object (e.g. brane) $\delta^2 A > 0 \Leftrightarrow -$ve eigenvalue of Hessian of A.

Canonical energy method

The canonical energy method [SH-Wald] provides a link between dynamical and thermodynamical notions of stability. It also provides a useful (in practice!) variational approach to stability problem.

Definition of canonical energy

$$\mathscr{E} = \int_\Sigma (\text{2nd order Einstein tensor})_{ab} n^a K^b + (\text{bndy terms})$$
$$= \int_\Sigma n^a j_a(\delta g, \mathscr{L}_K \delta g)$$

Here j^a is the conserved "symplectic current" of GR.

$$j^a(\delta_1 g, \delta_2 g) = g^{abcdef}(\delta_1 g_{bc} \nabla_d \delta_2 g_{ef} - \delta_2 g_{bc} \nabla_d \delta_1 g_{ef})$$

Canonical energy method

The usefulness of $\mathscr{E}$ rests on following properties:

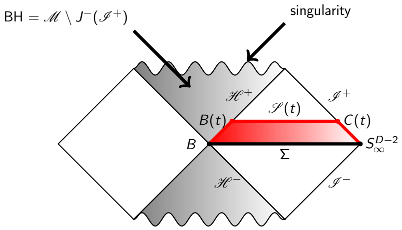
Key properties

- $\mathscr{E}$ satisfies a 'balance equation'
- $\mathscr{E}$ satisfies a 'thermodynamic identity'
- $\mathscr{E}$ is gauge-invariant
- $\mathscr{E}$ vanishes for stationary perturbations/perturbations to other stationary BH's

All of these properties can ultimately be traced back to the symplecitic structure of GR, diffeomorphism invariance, and the geometric nature of the boundaries (scri/horizon). I now explain the first two properties.

Canonical energy: balance law

The balance law applies to the following situation:



Balance formula

For axi-symmetric perturbations:

$$\mathscr{E}(t_2) - \mathscr{E}(t_1) = \int_{\text{infinity}} (\delta\text{news})^2 + \int_{\text{horizon}} (\delta\text{shear})^2 > 0$$

A **corollary** is that $\mathscr{E}$ is monotonically decreasing towards future.

Canonical energy: magic formula

The following formula connects the canonical energy to the second variation of thermodynamic quantities:

"Magic" formula

$$\mathscr{E} = \delta^2 m - \frac{\kappa}{8\pi}\delta^2 A - \sum \Omega^I \delta^2 J_I$$

The sign of $\mathscr{E}$ can thereby be related to the second order change in entropy $\delta^2 A$ (assuming e.g. that m, J_I remain unchanged to second order). The sign of $\mathscr{E}$ can also be related to dynamical instabilities. Thus:

Summary

- Balance formula gives monotonicity of $\mathscr{E}$
- Magic formula gives link to thermodynamics

Outline of argument:

(In-)stability argument (SH and Wald 2013)

1. Show that $\mathscr{E} < 0$ for some linear perturbation. Do this either
   - by exploiting the relationship between $\mathscr{E}$ and thermodynamic quantities ("magic formula") or
   - by making an educated guess for initial data of unstable "mode". This typically involves advanced tools such as "initial data engineering", "blow-up constructions", etc.
2. Note that $\mathscr{E}$ cannot increase towards future by "balance formula".
3. Note that, in order for the mode to settle down, you must have $\mathscr{E} \to 0$, a contradiction
4. [Parbhu et al.] even show $\mathscr{E} < 0 \Rightarrow$ exponential growth.

The remainder of this talk illustrates this general strategy in some examples.

Overview of applications of method

Different example systems

1. Gubser-Mitra conjecture about instability of black branes [SH and Wald].
2. Einstein-fluid system: Stability of stars (no BHs) [Friedman & Schutz 1972, Green et al. 2014]. Timescale of instability very long.
3. Instabilities of extremal higher dimensional BHs (Durkee-Reall conjecture) [SH and Ishisbashi]
4. Super radiant instabilities ("black hole bombs") [Green, SH, Ishisbashi and Wald].
5. Ultra-spinning BHs, ...

Perturbations of higher dimensional black hole/brane spacetimes

Issue:

Perturbations of generic higher dimensional BHs not tractable by 'Teukolsky equation'.

But:

Explicit criteria for instabilities can be identified via canonical energy method.

[SH-Ishibashi 14, Durkee-Reall 11]

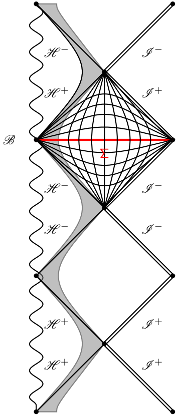


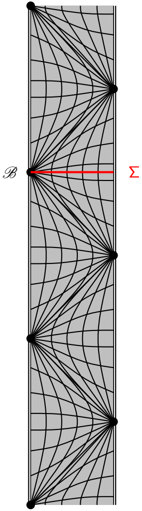
Figure: Conformal diagram of extremal Myers-Perry (MP) black hole. Near horizon region is shaded

Extremal black holes

One may simplify the situation by zooming in onto the horizon. Technically speaking, we take near horizon limit, i.e. "blow up" of shaded region.

Gain:

Limit has "algebraically special" properties [Pravda et al., Milson et al., Durkee-Reall]. $\Rightarrow$ algebraic simplification of perturbed Einstein eqn.s



Conformal diagram of the NH limit of the extremal MP spacetime, i.e. AdS_2. This should be thought of as corresponding to the shaded region in previous of the extreme MP black hole, to be taken "infinitely thin". The Cauchy surface Σ in that conformal diagram corresponds to the surface Σ drawn here. The curvy upward lines show the orbits of K^a.

Durkee-Reall conjecture [Durkee-Reall 2011]

Tool: Gaussian Null Coordinates:

$$g = 2du(dr + r^2 h du + r f_A dx^A) + q_{AB} dx^A dx^B$$

The algebraically special properties of NH-geometries naturally lead to the construction of a natural stability operator $\mathscr{A}$:

Definition of $\mathscr{A}$:

$\mathscr{A}$ = second order elliptic operator on horizon S^{d-2} ("squashed sphere"), depending on geometric data of NH.

Based partly on numerical evidence up to $d < 15$ [Dias et al. 2010], Durkee-Reall conjectured:

Durkee-Reall conjecture:

If smallest eigenvalue λ of $\mathscr{A}$ is $< -1/4 \Longrightarrow$ BH unstable!

Proof of Durkee-Reall conjecture [SH-Ishibashi 2014]

Near horizon metric: [Durkee-Reall, Hubeny et al., ...]

$$g_{\rm NH} = L^2 \, d\hat{s}^2 + g_{IJ}(d\phi^I + k^I \, \hat{A})(d\phi^J + k^J \, \hat{A}) + d\sigma^2_{d-n-2} \, ,$$

is a fibration over AdS_2. Has algebraically special character [Coley, Milson, Pravda, Pradova, Reall, ...] with preferred null fields n^a, l^a. Generalized Newman-Penrose formalism possible. Hertz potential ansatz for metric perturbation

$$\delta g_{ab} = l_a l_b (C_{cedf} l^e n^f U^{cd}) - 2 l_{(a} \text{þ} \eth^c U_{b)c} - 2 l_{(a}(2\tau^c + [l,n]^c) \text{þ} U_{b)c} + \text{þ}^2 U_{ab}$$

Solves EOM iff [Kegeles-Cohen; Chrzanowski, Wald, M Godazgar, SH-Ishibashi]:

Decoupling equation:

$$(\hat{D}^2 - q^2 - \mathscr{A}) U_{ab} = 0.$$

$\mathscr{A}$ the stability operator. Make separation of variables ansatz $U_{ab} = \phi(AdS_2) Y_{ab}(angles)$ with separation constant λ.

Proof of Durkee-Reall conjecture [SH-Ishibashi 2014]

- Show that $\mathscr{E} < cst.(\lambda + \frac{1}{4})(\lambda^2 + \frac{7}{2})$ for suitable perturbation ϕ on AdS_2. Thus, if $\lambda < -\frac{1}{4}$ then $\mathscr{E} < 0$ for some perturbation on NH.
- Modify initial data of NH-geometry perturbation to one on extremal BH ["initial data engineering" Corvino-Schoen, Chrusciel-Delay]
- Make sure that still $\mathscr{E} < 0$ [Sobolev-space technology]
- Argue from the monotonicity ($\mathscr{E}$ decreasing) that $\mathscr{E} \nrightarrow 0$!

Theorem:

If lowest eigenvalue of $\mathscr{A}$ in axi-symmetric subsector is $\lambda < -\frac{1}{4}$, then there is a perturbation such that $\mathscr{E} \nrightarrow 0$, i.e. this perturbation cannot settle down to pure gauge or perturbation to other stationary BH. $\Longrightarrow$ BH unstable!

Generalizations to AdS-type BHs and near extremal BHs possible.

Black hole bombs

Superradiance

If a wave of the form $e^{i\omega t + im\phi}$ is incident upon a rotating object with angular velocity Ω, then provided

$$\omega < m\Omega,$$

the scattered wave is amplified. [Zel'dovich,Penrose,Christodoulou,...].

Feeding back the scattered wave (e.g. by "mirror"), one expects that one can render the system unstable. Particularly interesting for rotating BHs [Press-Teukolsky 1972, Cardoso et al. 2008,...] $\rightarrow$ "black hole bomb".


LETTERS TO NATURE

PHYSICAL SCIENCES

Floating Orbits, Superradiant Scattering and the Black-hole Bomb

Penrose[1] and Christodoulou[2] have shown how, in principle, rotational energy can be extracted from a black hole by orbiting and fissioning particles. Recently Misner[3] has pointed

We now consider a wa hole. Normally, a part o potential barrier $W(r^*)$, wh down the hole, so that the ingoing wave. If, howeve range $0 < \omega < m\omega_{\text{horizon}}$, the coordinate outgoing, it re outside of the barrier, and energy than ingoing. The

Black hole bombs

Superradiance

If a wave of the form $e^{i\omega t+im\phi}$ is incident upon a rotating object with angular velocity Ω, then provided

$$\omega < m\Omega,$$

the scattered wave is amplified. [Zel'dovich,Penrose,Christodoulou,...].

Feeding back the scattered wave (e.g. by "mirror"), one expects that one can render the system unstable. Particularly interesting for rotating BHs [Press-Teukolsky 1972, Cardoso et al. 2008,...] $\rightarrow$ "black hole bomb".

AdS BH-bombs = superradiant instability

AdS-boundary conditions provide a natural 'mirror' and hence natural candidate for this type of instability [Hawking-Reall 99].

Superradiance

Geometric optics explanation of super radiance in GR: If γ worldline of material observer, T^a Killing field which is time-like near infinity, then locally measured energy by observer with momentum $p_a \propto \dot{\gamma}_a$:

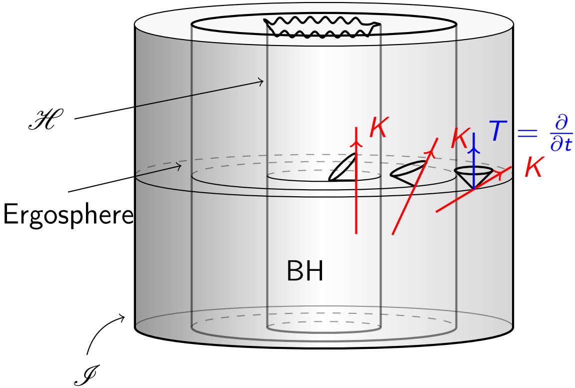
$$E = -T^a p_a < 0$$

possible in ergo region (where T^a is space-like). $\Longrightarrow$ energy extraction from BH possible. Ergo regions typically exist in a neighborhood of rotating BHs. (For waves propagating on BH one argues that one can arrange transmitted wave to have negative energy, and reflected wave to have positive energy.)
A new situation can arise in asymptotically AdS-BHs, because waves can be reflected off conformal infinity. Let

$$K^a = T^a + \sum \Omega^I \phi_I^a$$

be the Killing field which is normal to the horizon. If the BH is rotating sufficiently fast, then K^a can be space like at the conformal boundary. $\Longrightarrow$ $E = -K^a p_a < 0$ possible.

AdS superradiant instability



$-K^a p_a < 0$ is possible in asymptotic region, conformal boundary is reflecting, energy flux across horizon $> 0 \implies$ spacetime unstable.

Detailed analysis

Difficulties

Detailed/quantitative analysis of superradiant instabilities using explicit mode solutions runs into many difficulties:

- No Teukolsky-formalism in $d > 4 \longrightarrow$ no decoupled equations/separation of variables.
- Explicit treatment of modes not possible in most interesting cases. Numerical results: [Dias-Santos-Way 2015]

"Proof" of super radiant instability is possible using the canonical energy method. Key point: balance formula is now

$$\mathscr{E}(t_2) - \mathscr{E}(t_1) = \int_{\text{horizon}} (\delta\text{shear})^2 > 0$$

for arbitrary (including non-axisymmetric) perturbation, because here is no flux term at infinity in AdS-type spacetimes ("mirror")!

AdS superradiant instability

Argument [Green, SH, Ishibashi, Wald]

- One constructs a perturbation of the form $\delta g_{ab} = A_{ab} \exp(iS/h)$, where A_{ab} is a tensor field supported in a tubular neighborhood around a worldline in asymptotic region and where S is a phase function chosen such that $-K^a p_a = K^a \nabla_a S < 0$ around this worldline. h is a WKB parameter.
- The linearized Einstein equation is imposed in the WKB sense (formal series expansion in h), giving transport/eikonal equations on A_{ab}, S. Solve order-by-order. [Choquet-Bruhat]
- One verifies $\mathscr{E} \sim -K^a p_a + O(h) < 0$ for this perturbation.
- One modifies WKB initial data by order h to get solution to linearized constraints ["initial data engineering" Corvino-Schoen,Chrusciel-Delay]
- One makes sure that still $\mathscr{E} < 0$ [Sobolev-space technology]
- Argue from the balance formula ($\mathscr{E}$ decreasing) that $\mathscr{E} \nrightarrow 0$!

Instabilities of 'black strings'

Given a stationary asymptotically flat BH with metric ds_d^2, then corresponding "black string" $ds_{d+1} = ds_d^2 + dz^2$ with $z \cong z + 2\pi L$ (spacetime $\mathscr{M} \times T^1$) is automatically a solution to Einstein's equations in $d+1$ dimensions. L is the length of the string. An example is

Schwarzschild black string

$$g = -\left[1 - (\tfrac{r_0}{r})^{d-3}\right] dt^2 + \left[1 - (\tfrac{r_0}{r})^{d-3}\right]^{-1} dr^2 + r^2 \, d\sigma^2_{S^{d-2}} + dz^2 \, .$$

To surprise of many researchers it was found by [Gregory-Laflamme] that Schwarzschild black string is unstable for sufficiently large L/r_0 ('long, thin strings'). A full numerical simulation of this instability has recently been carried out by [Lehner-Pretorius]. Simulation reveals an interesting 'self-similar' structure:

Black string instability

Simulation of black string instability [Lehner-Pretorius]:

(lehner.mpg)

The z-coordinate of the string goes up, the r-coordinate goes left. Simulation zooms in onto thin necks. No singularities in finite affine time

[Maeda-Horowitz,Marolf]

The instability of black string seems strikingly similar to the Plateau-Rayleigh instability of a thin jet of water.

Black string instability

Simulation of black string instability [Lehner-Pretorius]:

(lehner.mpg)

The z-coordinate of the string goes up, the r-coordinate goes left. Simulation zooms in onto thin necks. No singularities in finite affine time

[Maeda-Horowitz,Marolf]

The instability of black string seems strikingly similar to the Plateau-Rayleigh instability of a thin jet of water.

Gubser-Mitra conjecture

Gubser-Mitra conjectured that there is a **simple criterion** for when a black string has an instability.

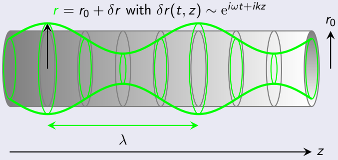
Suppose we have a family of stationary, asymptotically flat BH's parameterized by m, J_I (and possibly further 'charges'). Then $A = Area(m, J_1, ..., J_N)$. Now form the Hessian of the area of the BH,

$$\mathrm{Hess}_A = \begin{pmatrix} \frac{\partial^2 A}{\partial m^2} & \frac{\partial^2 A}{\partial J_I \partial m} \\ \frac{\partial^2 A}{\partial m \partial J_K} & \frac{\partial^2 A}{\partial J_I \partial J_K} \end{pmatrix} .$$

Gubser-Mitra

If the hessian or BH has a **positive eigenvalue**, then the corresponding black string $ds^2 + dz^2$ is **dynamically unstable**.

Black string vs. Plateau-Rayleigh instability



Dispersion relation $\omega^2 = \ldots (1 - k^2 r^2)$, implying instability if $kr_0 < 1$ (that is $\lambda > 2\pi r_0$), i.e. sufficiently long wavelength perturbation compared to radius r_0. (Plateau experiment 1873: $\lambda \geq 6.2\ldots\, r_0$)

Numerical studies [Gregory-Laflamme]:

Schwarzschild black string instability occurs for sufficiently long, thin strings, i.e. $L/r_0 \geq c_*$. Generic feature of *all* black strings?

Proof of Gubser-Mitra conjecture

1. Magic formula shows existence of a perturbation with $\mathscr{E} < 0$.
2. Balance formula shows that $\mathscr{E} \nrightarrow 0$ since $\mathscr{E}$ decreasing.
3. Gauge-invariance of $\mathscr{E}$ shows that $\mathscr{E} \nrightarrow 0$ pure gauge mode or perturbation towards other stationary black object. $\implies$ Instability!

Details [SH-Wald 2012]

Local Penrose inequality method

Penrose inequality ($d = 4$):

If $A =$ area of outermost 'horizon' of initial data set $\Rightarrow A \leq 16\pi m^2$.

[Penrose,Huisken-Ilmanen,Bray]

Heuristic argument:

1. Area A increases with time
2. Mass m decreases with time
3. Equality in Schwarzschild ($r_0 = 2m$)
4. If solution settles down to Schwarzschild $\Rightarrow$ ineq. must hold!

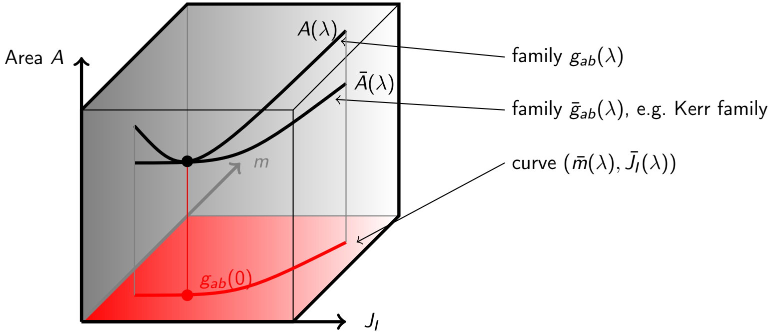
Linearized version of Penrose-inequality:

For axi-symmetric perturbations of Kerr: $\delta^2 A \leq \delta^2 \bar{A}$ for any perturbation such that $\delta^2 J = \delta^2 \bar{J}, \delta^2 m = \delta^2 \bar{m}$ (overbar = background Kerr quantities)

One can turn the logic around [Figueras-Murata-Reall] and propose that

Penrose inequality vs. instability:

If for axi-symmetric perturbations of stationary BH: $\delta^2 A > \delta^2 \bar{A}$ for some perturbation such that $\delta^2 J_I = \delta^2 \bar{J}_I, \delta^2 m = \delta^2 \bar{m}$ (overbar = background quantities) $\Rightarrow$ instability!



Graphical representation: Two families of metrics $g_{ab}(\lambda)$ and $\bar{g}_{ab}(\lambda)$ project onto same curve in space of conserved parameters.

Local Penrose inequality method

The local Penrose inequality proposal was tested numerically in [Murata et al. 2010] by:

1. Construct numerically initial data such that $\delta^2 A > \delta^2 \bar{A}$ (constraints!)
2. Compare with time evolution (also numerical) to see if instability really occurs.
3. Agreement found.

Proof of local Penrose inequality method:

Local Penrose inequality violation is directly related to $\mathscr{E} < 0$ via magic formula. $\Rightarrow$ direct relationship with canonical energy method.

Conclusions

In this talk, I have:

1. Explained that the laws of BH mechanics should be supplemented by the statement that thermodynamic instability implies dynamical instability.
2. Explained canonical energy method
3. Explained how this is related to thermodynamic instability
4. Explained how it can be used to understand instabilities in several situations of interest, e.g. (near) extremal black holes, black branes, super radiant instabilities in asymptotically AdS BH spacetimes.