The instability of Anti-de Sitter spacetime for the Einstein–scalar field system

Georgios Moschidis

Princeton University

Mathematical and Numerical Aspects of Gravitation

IPAM, 25^{th} *October* 2021

Introduction: Anti-de Sitter spacetime

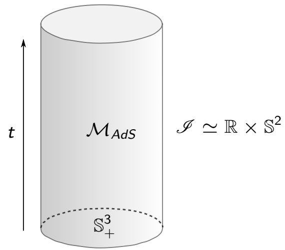
Introduction: Anti-de Sitter spacetime

Simplest solution of the **vacuum** Einstein equations with **negative** cosmological constant Λ: *Anti-de Sitter* spacetime $(\mathbb{R}^{3+1}, g_{AdS})$,

$$g_{AdS} = -\left(1 - \frac{1}{3}\Lambda r^2\right) dt^2 + \left(1 - \frac{1}{3}\Lambda r^2\right)^{-1} dr^2 + r^2 g_{\mathbb{S}^2}.$$

- Of central importance for high energy physics in the context of the holographic principle.

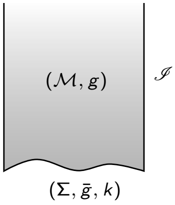
AdS spacetime can be conformally identified with the interior of the cylinder $\mathbb{R} \times \mathbb{S}^3_+$ equipped with the natural product Lorentzian metric.



- Conformal boundary $\mathscr{I}$ at infinity: Of timelike character.
- Initial data at $t = 0$ are **not** sufficient to uniquely determine a solution to a hyperbolic equation on AdS: Boundary conditions should also be imposed on $\mathscr{I}$.

The initial-boundary value problem

In view of the timelike character of the conformal boundary $\mathscr{I}$, the right setting to study the Einstein equations on *asymptotically AdS* spacetimes: **Initial-boundary value problem**.



- Initial data $(\Sigma^3, \bar{g}, k)$ satisfying the *constraint equations*

$$R[\bar{g}] + (\mathrm{tr}k)^2 - |k|^2 = 2\Lambda,$$
$$\mathrm{div}(k - \mathrm{tr}k \cdot \bar{g}) = 0.$$

- Conformal boundary conditions on $\mathscr{I}$, plus compatibility conditions at the "corner" $\partial\Sigma = \mathscr{I} \cap \Sigma$.

The initial-boundary value problem

The initial-boundary value problem

Identifying the right asymptotic boundary conditions is non-trivial!

Theorem (Friedrich, 1995)

For any prescribed smooth conformal structure on $\mathscr{I}$ and any asymptotically AdS initial data set $(\Sigma^3, \bar{g}, k)$ such that $r^{-2}\bar{g}$, $r^{-1}k$ extend smoothly to $\partial\Sigma$: $\exists!$ smooth solution of the vacuum equations.

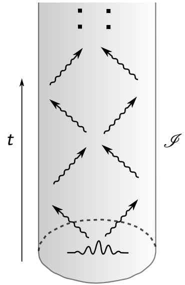
- *Reflecting* boundary condition in this class: $g|_{\mathscr{I}} \sim g_{\mathbb{R}\times\mathbb{S}^2}$.
- Geometric uniqueness for the IBVP in the case of a regular boundary: Fournodavlos–Smulevici.

Confinement and weak turbulence

Confinement and weak turbulence

The well-posedness of the initial-boundary value problem allows the study of the long time dynamics of asymptotically AdS solutions.

Question: *What are the stability properties of small initial perturbations of AdS under reflecting boundary conditions at $\mathscr{I}$?*

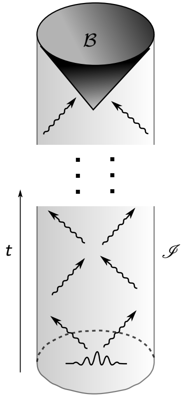


- In the case of the linear toy model $\Box_g \phi = 0$: The energy $\mathcal{E}[\phi](t) \sim \int_{\Sigma_t} |\partial\phi|^2 \, dx$ does *not* decay as $t \to \infty$ when reflecting conditions are assumed on $\mathscr{I}$.

- Non-linear effects can accumulate over long timescales, possibly precipitating cascading effects.

The AdS instability conjecture

The AdS instability conjecture



In 2006, DAFERMOS–HOLZEGEL conjectured the following scenario:

AdS instability conjecture

Assuming a *reflecting* boundary condition on $\mathscr{I}$ for the vacuum equations, there exist arbitrarily small perturbations of the AdS initial data which lead to the formation of a **black hole** region after sufficiently long time. In particular, $(\mathcal{M}_{AdS}, g_{AdS})$ is *non-linearly unstable*.

- Black hole formation: Concentration of energy at small scales.

The AdS instability conjecture

The AdS instability conjecture

Remarks:

- Smallness of the perturbations: With respect to a suitable norm for which well-posedness holds.
  - Example: Higher order, weighted Sobolev spaces $\mathcal{H}^k$.
  - The conserved total ADM mass M_{ADM} is not suitable: The equations are supercritical with respect to it.
- The choice of reflecting boundary conditions on $\mathscr{I}$ is important. Holzegel–Luk–Smulevici–Warnick: Superpolynomial decay at the linearized level for maximally dissipative boundary conditions.
- The conjecture is not restricted to the vacuum case; it also applies to any "reasonable" matter model for which the stability of Minkowski spacetime holds (in the case $\Lambda = 0$).

Spherically symmetric models: Einstein–scalar field

Spherically symmetric models: Einstein–scalar field

It is natural to seek an unstable family of initial data having additional symmetries. In $3+1$ dimensions, the only surface symmetry for the initial data which is compatible with the AdS asymptotics: Spherical symmetry.

- Birkhoff's theorem: The only spherically symmetric solution of the vacuum equations with a regular center of symmetry is AdS
  $\Rightarrow$ Trivial dynamics for the vacuum equations in this class.

Spherically symmetric models: Einstein–scalar field

It is natural to seek an unstable family of initial data having additional symmetries. In $3+1$ dimensions, the only surface symmetry for the initial data which is compatible with the AdS asymptotics: Spherical symmetry.

- Birkhoff's theorem: The only spherically symmetric solution of the vacuum equations with a regular center of symmetry is AdS
$\Rightarrow$ Trivial dynamics for the vacuum equations in this class.

A simple matter model admitting non-trivial dynamics in spherical symmetry: The **Einstein–scalar field** system

$$\begin{cases} Ric_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}[\phi], \\ \Box_g \phi + \frac{2}{3} \Lambda \alpha \phi = 0. \end{cases}$$

The initial-boundary value problem for a scalar field

The initial-boundary value problem for a scalar field

On AdS spacetime, solutions to the linear scalar field equation have an asymptotic expansion near $\mathscr{I}$ of the form

$$\phi = r^{-\lambda_-} \underbrace{\phi^-_{\mathscr{I}}}_{\text{Dirichlet data}} + r^{-\lambda_+} \underbrace{\phi^+_{\mathscr{I}}}_{\text{Neumann data}} + \mathcal{O}(r^{-2-\lambda_-}), \quad \lambda_\pm = \frac{3}{2} \pm \sqrt{\frac{9}{4} + 2\alpha}.$$

- When $\alpha \neq -1$: Solutions are conformally singular at $\mathscr{I}$.
- Well-posedness of the *linear* scalar field equation on asymptotically AdS spacetimes with homogeneous Dirichlet conditions when $\alpha > -\frac{9}{8}$: VASY.

In the case of the *non-linear* Einstein–scalar field system in *spherical symmetry*:

- HOLZEGEL–SMULEVICI: (Homogeneous) Dirichlet boundary conditions when $\alpha > -\frac{9}{8}$.
- HOLZEGEL–WARNICK: More general boundary conditions (including Neumann) for $-\frac{9}{8} < \alpha < -\frac{5}{8}$.

AdS Instability: The Einstein–scalar field system

AdS Instability: The Einstein–scalar field system

First numerical and heuristic study of the instability of AdS in the setting of the spherically symmetric Einstein–scalar field system with Dirichlet conditions at $\mathscr{I}$: Bizon–Rostworowski (2011).

- Proposed instability mechanism: Perturbative analysis of the effective scalar field equation

$$\Box_{(AdS)}\phi + \frac{2}{3}\Lambda\alpha\phi \simeq \frac{2m[\phi]}{r^3}\phi = \mathcal{N}^{(3)}[\phi]$$

  suggests that energy is transferred to high frequency modes of

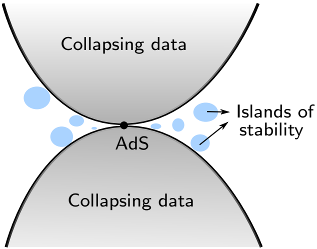
$$\phi(t,x) = \sum_{k\in\mathbb{Z}} e^{i\omega_k t}\phi_k(t;x)$$

  through *resonant* interactions (when $\omega_k = \omega_l - \omega_m + \omega_n$).

AdS Instability: The Einstein–scalar field system

AdS Instability: The Einstein–scalar field system

Subsequent numerical and heuristic works also explored the distribution of unstable perturbations in the space of initial data, as well as different boundary conditions:
BUCHEL–LEHNER–LIEBLING, DIAS–HOROWITZ–MAROLF–SANTOS, BALASUBRAMANIAN–BUCHEL–GREEN–LEHNER–LIEBLING, BIZON–MALIBORSKI, CRAPS–EVNIN–VANHOOF, DIMITRAKOPOULOS–FREIVOGEL–LIPPERT–YANG...



Major questions:

- Do *all* perturbations of AdS spacetime collapse into black holes? Are there "islands of stability", i.e. open sets in the moduli space of initial data close to AdS giving rise to quasiperiodic, non-collapsing solutions?
- Once a black hole is formed, what are its long time dynamics? Does its exterior become asymptotically stationary?

An alternative approach

An alternative approach

The resonant mode mixing mechanism: relevant for the first stage of the instability, where perturbation theory is still valid.

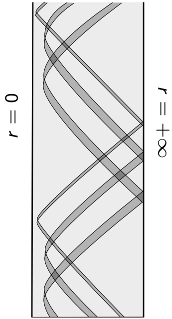
- No rigorous proof so far using this approach!

An alternative approach for a rigorous proof of the AdS instability conjecture: Study the interaction of short pulses in *physical space* and use the *monotonicity properties* of the Einstein equations.

Instability via beam interactions

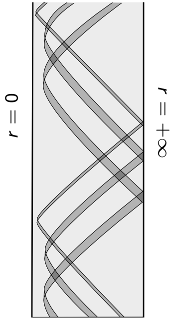
Instability via beam interactions

Let us assume that the initial perturbation is chosen so that it gives rise to a number of spherically symmetric, narrow beams which are initially ingoing.



Instability via beam interactions

Let us assume that the initial perturbation is chosen so that it gives rise to a number of spherically symmetric, narrow beams which are initially ingoing.



- Away from $r = 0$: Narrow beams approximately satisfy the *Einstein–null dust* system (geometric optics approximation).
- Near $r = 0$: Each beam turns from ingoing to outgoing through a self-interaction process.

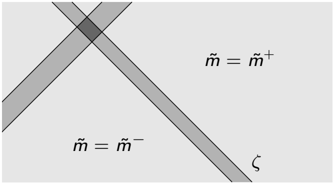
AdS Instability via beam interactions

AdS Instability via beam interactions

The region between the beams of matter is approximately *vacuum*. The energy of each beam ζ can be expressed in terms of the renormalised *Hawking mass* $\tilde{m}$:

$$\mathcal{E}[\zeta] \doteq \tilde{m}^{+} - \tilde{m}^{-},$$

where $\tilde{m} \doteq m - \frac{1}{3}\Lambda r^3$.



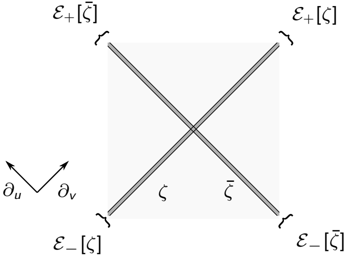
The nearly empty regions between the beams have approximately constant renormalised Hawking mass.

- Trapped surface at sphere of symmetry p if $\frac{2m}{r}(p) > 1$.
- $\mathcal{E}[\zeta]$ changes each time ζ is intersected by another beam, but is preserved at each reflection off $\mathscr{I}$.

AdS Instability via beam interactions

AdS Instability via beam interactions

Let ζ, $\bar{\zeta}$ be a pair of intersecting beams, so that the intersection lies in the regime where the geometric optics approximation holds.



In double null coordinates (u, v), the relation

$$\partial_u \partial_v \tilde{m} \simeq \frac{2}{r\left(1 - \frac{2m}{r}\right)} (-\partial_u \tilde{m}) \partial_v \tilde{m}$$

yields the approximate energy exchange formulas:

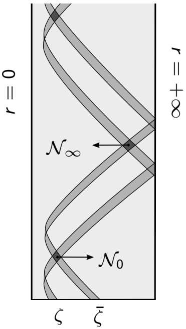
$$\mathcal{E}_+[\bar{\zeta}] = \mathcal{E}_-[\bar{\zeta}] \cdot \exp\left(\frac{2}{r} \frac{\mathcal{E}_-[\zeta]}{1 - \frac{2m}{r}} + \mathfrak{Err} \right),$$

$$\mathcal{E}_+[\zeta] = \mathcal{E}_-[\zeta] \cdot \exp\left(- \frac{2}{r} \frac{\mathcal{E}_-[\bar{\zeta}]}{1 - \frac{2m}{r}} + \mathfrak{Err} \right).$$

AdS Instability via beam interactions

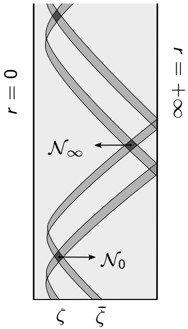
AdS Instability via beam interactions

Assume, for a moment, that our configuration consists of only two beams $\bar{\zeta}$ and ζ:



AdS Instability via beam interactions

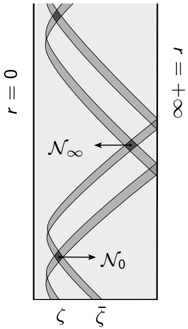
Assume, for a moment, that our configuration consists of only two beams $\bar{\zeta}$ and ζ:



As long as $r|_{\mathcal{N}_0} \ll r|_{\mathcal{N}_\infty}$: $\mathcal{E}[\bar{\zeta}]$ increases and $\mathcal{E}[\zeta]$ decreases after each successive reflection.

AdS Instability via beam interactions

Assume, for a moment, that our configuration consists of only two beams $\bar{\zeta}$ and ζ:

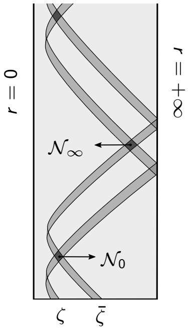


As long as $r|_{\mathcal{N}_0} \ll r|_{\mathcal{N}_\infty}$: $\mathcal{E}[\bar{\zeta}]$ increases and $\mathcal{E}[\zeta]$ decreases after each successive reflection.

If $\bar{\zeta}$ is narrower than ζ: energy is carried at smaller scales.

AdS Instability via beam interactions

Assume, for a moment, that our configuration consists of only two beams $\bar{\zeta}$ and ζ:



As long as $r|_{\mathcal{N}_0} \ll r|_{\mathcal{N}_\infty}$: $\mathcal{E}[\bar{\zeta}]$ increases and $\mathcal{E}[\zeta]$ decreases after each successive reflection.
If $\bar{\zeta}$ is narrower than ζ: energy is carried at smaller scales.

This is a non-linear instability mechanism!

AdS instability: the Einstein–scalar field system

AdS instability: the Einstein–scalar field system

The previous heuristic mechanism can in fact lead to a rigorous proof of the conjecture for certain matter models.

- Einstein–massless Vlasov system $(\mathrm{M.}\ 2018)$: Allows for perfectly localized matter beams.

It can also be applied to the case of the Einstein–scalar field system:

Theorem (M.):

There exists a family of spherically symmetric characteristic initial data $\mathcal{D}_\epsilon(\Omega_\epsilon, r_\epsilon, \phi_\epsilon)$ on $\{u=0\}$ for the the **conformally coupled** (i.e. $\alpha=-1$) Einstein–scalar field system such that:

- $||\mathcal{D}_\epsilon||_{BV} \doteq \int_{u=0} \left| \frac{\partial_v}{\partial_v r}\left(\frac{\partial_v(r\phi_\epsilon)}{\partial_v r}\right)\right| dv \xrightarrow{\epsilon\to 0} 0,$
- For any $\epsilon>0$, the evolution of $\mathcal{D}_\epsilon$ with **Dirichlet** or **Neumann** bc's on $\mathscr{I}$ leads to the creation of a **black hole region** after sufficiently long time.

AdS instability: the Einstein–scalar field system

Remarks:

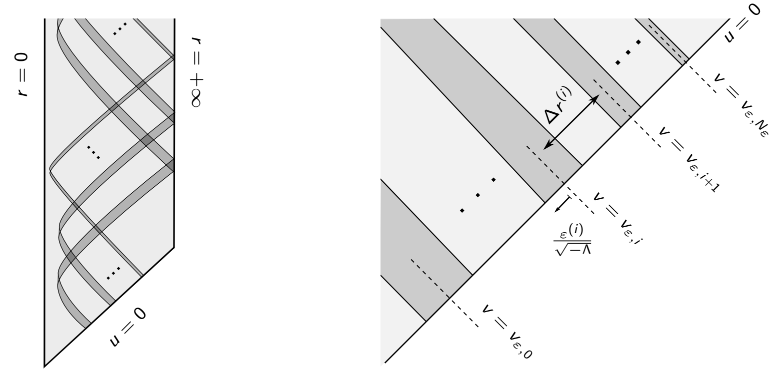
- Well-posedness of the initial-boundary value problem in the $||\cdot||_{BV}$ topology when $\alpha = -1$ follows by a simple modification of the work of CHRISTODOULOU.

- When $\Lambda = 0$: Minkowski spacetime is *stable* under spherically symmetric perturbations which are initially small with respect to $||\cdot||_{BV}$.

Main ideas of the proof

Main ideas of the proof

The proof proceeds by arranging the scalar field initially into a large number of ingoing narrow beams, with each successive beam being much narrower than the previous one.

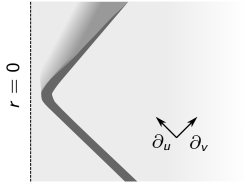


- Narrower beams are exected to approach closer to $r = 0$.
- Energy is expected to flow towards the narrowest beam through the mechanism sketched before.

Main ideas of the proof

Main ideas of the proof

A serious technical obstacle to implementing the heuristics: Beams lose coherence over time.



- Decoherence is most severe when a beam reaches close to $r = 0$, due to non-linear self-interactions:

$$\partial_u \partial_v (r\phi) + V(\phi)\phi = 0,$$

where

$$V(\phi) \doteq -2\frac{(\partial_v r)(\partial_u r)}{1 - \frac{2m}{r}}\left(\frac{\tilde{m}}{r^2} - \frac{4}{3}\pi\Lambda r\phi^2\right)$$

and $\tilde{m}$ is determined by

$$\partial_v \tilde{m} = 2\pi r^2 \left(1 - \frac{2m}{r}\right)\frac{(\partial_v \phi)^2}{\partial_v r}.$$

Main ideas of the proof

Main ideas of the proof

In order to keep decoherence under control for sufficiently long time under low regularity assumptions: The scalar field is split as

$$\phi = \phi_1 + \ldots + \phi_N + \phi_{Err},$$

where ϕ_i satisfies initially $\partial\phi_i \lesssim \delta$, $\partial^2\phi_i \lesssim \delta\epsilon_i^{-1}$ and solves the simpler equation

$$\partial_u\partial_v(r\phi_i) = 0,$$

for which stronger coherence estimates can be established.

- Regularity at $r = +\infty$ is also important here.

The "error" term ϕ_{Err} measures the "total decoherence" of the beams and solves:

$$\partial_u\partial_v(r\phi_{Err}) + V(\phi)\phi_{Err} = -\sum_{i=1}^{N} V(\phi)\phi_i$$

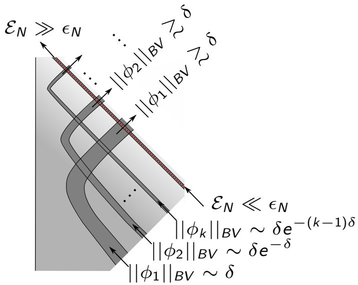
As long as $||\phi_i||_{BV} \lesssim \delta \ll 1$: $||\phi_{Err}||_{BV} \lesssim \delta$, *even if* $1 \ll ||\phi||_{BV} \ll \delta^{-1}$.

- Choosing the hierarchy of scales ϵ_i, Δr_i carefully is crucial for this step.

Main ideas of the proof

Main ideas of the proof

At the last step before trapped surface formation:



- Fine tuning of the initial data: Before the final step, the beams ϕ_i create a profile of exponential form, resembling a discretely self-similar background.
- Trapped surface formation after final interaction: CHRISTODOULOU.
- Control of the decoherence of the beams becomes more difficult at this stage.

Open questions and future directions

Open questions and future directions

Generic spherically symmetric perturbations: Do all perturbations of AdS eventually collapse into black holes?

- Probably no; see also CHATZIKALEAS–SMULEVICI.
- Do islands of stability exist for timescales beyond the ones provided by scaling considerations?

Moving beyond spherical symmetry: In 3+1 dimensions, the vacuum equations **cannot** be reduced under symmetry to a 1+1 dimensional system. However, in principle, a similar physical space approach could be followed in this case as well.

Major challenges:

- Well-posedness in a class of initial data lying in a scale-invariant topology, with additional regularity in the "angular directions".
- Boundary effects are expected to be highly non-trivial outside surface symmetry.

Thank you for your attention!