Retiring the third law of black hole thermodynamics

Ryan Unger

Department of Mathematics, Princeton University

EPF Lausanne analysis seminar
February 2024

joint work with Christoph Kehle (ETH Zürich)

Outline

1. What is black hole thermodynamics and what are extremal black holes?
2. The third law is false
3. Broad strokes of the proof: characteristic gluing
4. Extremal black hole formation as a critical phenomenon

Classical thermodynamics

Law	Thermodynamics
Zeroth	T constant throughout body in thermal equilibrium
First	$dE = TdS +$ work terms (PdV)
Second	$dS \geq 0$ in any process
Third	Impossible to achieve $T = 0$ in finite time

Black hole thermodynamics

Commun. math. Phys. 31, 161–170 (1973)

The Four Laws of Black Hole Mechanics

J. M. Bardeen*
Department of Physics, Yale University, New Haven, Connecticut, USA

B. Carter and S. W. Hawking
Institute of Astronomy, University of Cambridge, England

Received January 24, 1973

Black hole thermodynamics is a proposed close mathematical **analogy** between black hole dynamics and classical thermodynamics.

- What is a black hole?
- What is the analogue of temperature T?
- What is the analogue of entropy S?

Crash course on black holes: Spacetimes and Einstein's equations

The setting is Einstein's theory of general relativity. A **spacetime** consists of a 4-manifold $\mathcal{M}^{3+1}$ and a Lorentzian metric g satisfying the **Einstein field equations**

$$Ric(g) - \tfrac{1}{2}R(g)g = 2T. \tag{EFE}$$

Example: Minkowski space. $\mathcal{M}^{3+1} = \mathbb{R}^{3+1}_{t,x,y,z}$ and

$$g = -dt^2 + dx^2 + dy^2 + dz^2$$

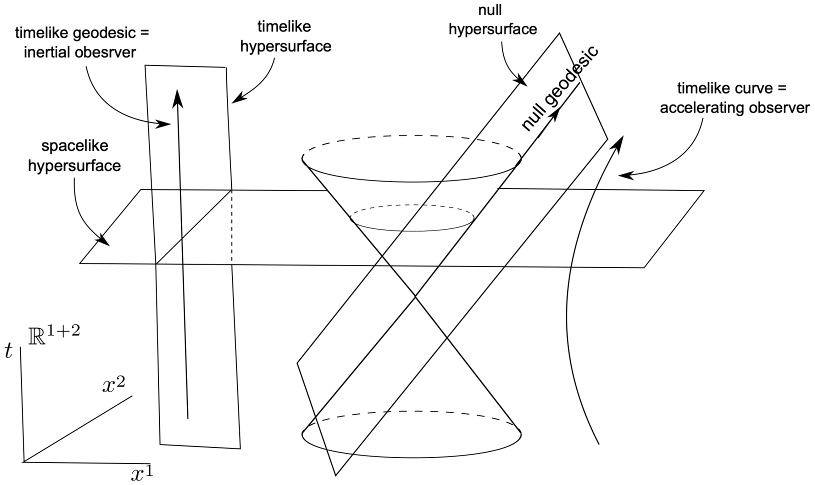
This metric describes the geometry of special relativity with speed of light $c = 1$.

Given a vector $v \in T_p\mathcal{M}$:

- $g(v, v) < 0$, v is *timelike*
- $g(v, v) = 0$, v is *null*
- $g(v, v) > 0$, v is *spacelike*

Curves with null or timelike tangent vector let us define **causality** on a spacetime.

Crash course on black holes: Geometry of Minkowski space



Crash course on black holes: What is a black hole?

Asymptotic flatness: g approaches the Minkowski metric "at infinity."

An asymptotically flat spacetime is said to contain a **black hole region**

$$\mathcal{BH} \subset \mathcal{M}$$

if observers and signals originating in $\mathcal{BH}$ cannot reach infinity.

The **event horizon** of the black hole is its topological boundary

$$\mathcal{H}^+ := \partial(\mathcal{BH}).$$

The event horizon $\mathcal{H}^+$ is a **null hypersurface**: it is ruled by geodesics with null tangent vectors.

Examples to come shortly!

Crash course on black holes: Teleology

The black hole region $\mathcal{BH}$ and event horizon $\mathcal{H}^+$ of a spacetime are **teleological** notions—they can't be known to exist, much less be located, *without knowing the entire future of the spacetime*.

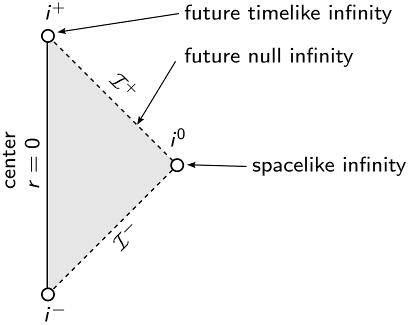
The event horizon $\mathcal{H}^+$ is not a singular object.

The formation, dynamics, and stability of black holes can be rigorously studied using the tools of Lorentzian geometry and hyperbolic PDE.

Crash course on black holes: Penrose diagrams

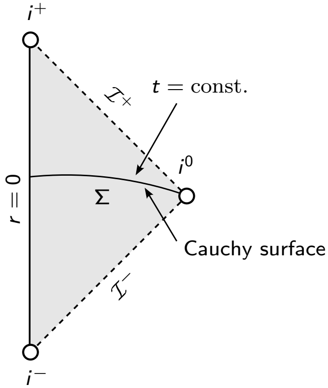
Penrose diagrams: A $(1+1)$-dimensional diagrammatic representation of spacetimes. This is particularly useful for visualizing causality.

- 45 degree lines correspond to null hypersurfaces in $(3+1)$ dimensions
- steeper than 45: timelike
- shallower than 45: spacelike

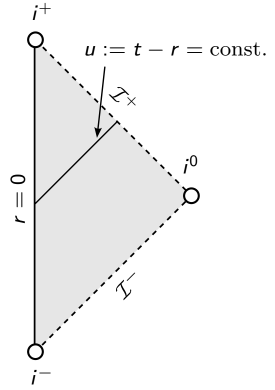


Penrose diagram of Minkowski space

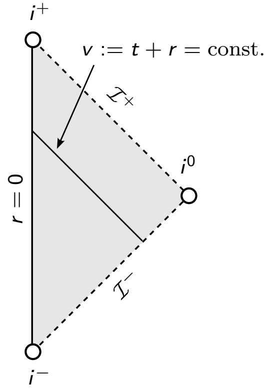
Crash course on black holes: Minkowski space



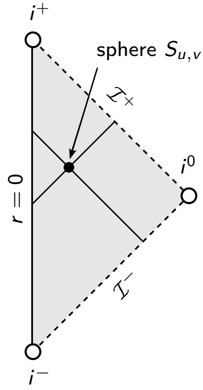
Crash course on black holes: Minkowski space



Crash course on black holes: Minkowski space



Crash course on black holes: Minkowski space

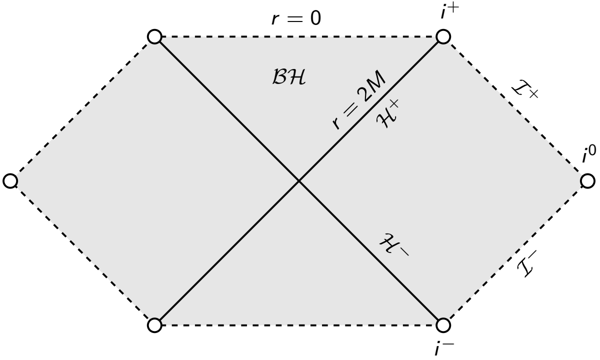


Crash course on black holes: The Schwarzschild solution

- Mass $M > 0$

$$g_M = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

- Solves the **Einstein vacuum** equations
- Describes a static, nonrotating black hole

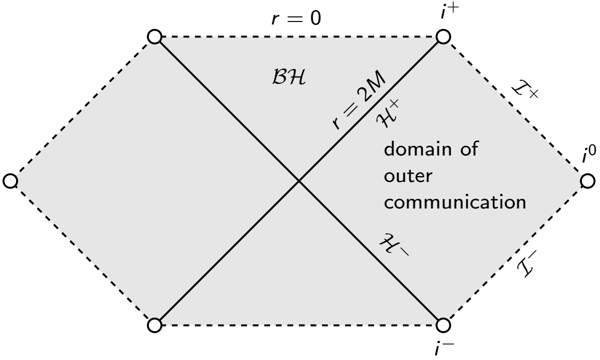


Crash course on black holes: The Schwarzschild solution

- Mass $M > 0$

$$g_M = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

- Solves the **Einstein vacuum** equations
- Describes a static, nonrotating black hole

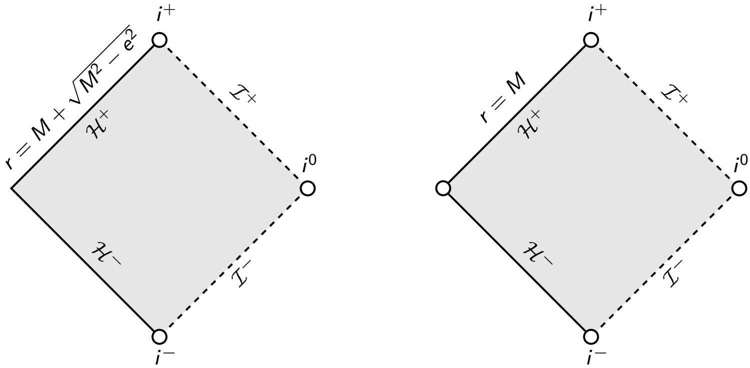


Crash course on black holes: Reissner–Nordström

- Mass $M > 0$ and charge $e \in \mathbb{R}$

$$g_{M,e} = -\left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right) dt^2 + \left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta \, d\phi^2)$$

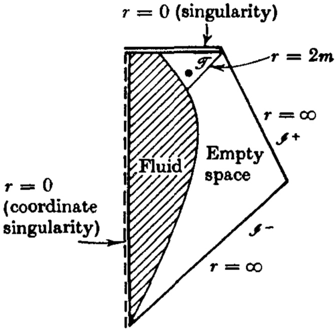
- Solves the **Einstein–Maxwell** equations
- Describes a static, charged, nonrotating black hole



Subextremal $|e| < M$ **Extremal** $|e| = M$

Superextremal $|e| > M$ does not contain a black hole!

Crash course on black holes: Gravitational collapse



Penrose diagram of gravitational collapse Oppenheimer–Snyder '39.
One-ended Cauchy data!

Crash course on black holes: Surface gravity and extremality

- Associated to a horizon is a function $\kappa \geq 0$, called the **surface gravity**
- Physically, κ measures the force required to hold a test mass "at the horizon" "from infinity"
- It is also related to the celebrated **(local) redshift effect**
- For Reissner–Nordström,

$$\kappa(g_{M,e}) = \frac{\sqrt{M^2 - e^2}}{(M + \sqrt{M^2 - e^2})^2}$$

- **Subextremal horizon:** $|e| < M$, $\kappa > 0$
- **Extremal horizon:** $|e| = M$, $\kappa = 0$

Black hole thermodynamics

Law	Classical thermodynamics	Black holes
Zeroth	T constant in equilibrium	surface gravity κ constant on stationary horizon
First	$dE = TdS + \cdots$	$dM = \kappa dA + \cdots$
Second	$dS \geq 0$	$dA \geq 0$
Third	$T \not\to 0$ in finite process	surface gravity $\kappa \not\to 0$ in finite advanced time

- Proposed by Bardeen–Carter–Hawking '73
- Hawking radiation—interpret $\kappa \propto T$
- Bekenstein entropy—interpret $A \propto S$
- These identifications are further justified once quantum effects are considered, but are **formal analogies** in classical GR.
- Laws 0, 1, and 2 proved by Hawking, Carter, Bardeen–Carter–Hawking, Wald, ...

Retiring the third law

Original formulation of Bardeen–Carter–Hawking:

Extending the analogy even further one would postulate:

The Third Law

It is impossible by any procedure, no matter how idealized, to reduce κ to zero by a finite sequence of operations.

Conjecture (The third law, BCH '73, Israel '86).

A subextremal black hole cannot become extremal in finite time by any continuous process, no matter how idealized, in which the spacetime and matter fields remain regular and obey the weak energy condition.

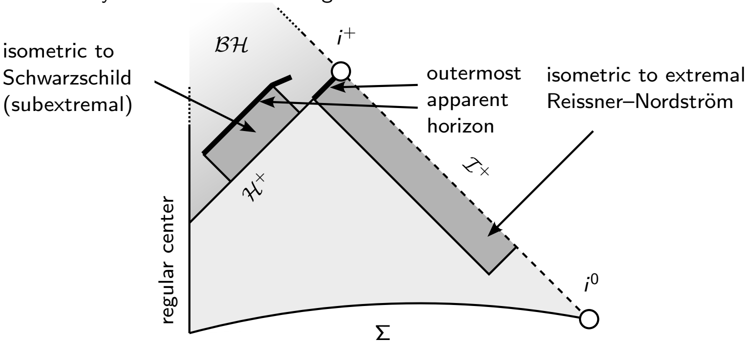
Theorem (Kehle–U. '22).

Subextremal black holes can become extremal in finite time, evolving from regular initial data. In particular, the "third law of black hole thermodynamics" is false.

More precisely, there exist regular solutions of the Einstein–Maxwell-charged scalar field system (which satisfies the weak energy condition) with the following behavior:

Counterexample to the third law: detailed statement of our theorem

There exist regular spherically symmetric solutions of the Einstein–Maxwell-charged scalar field system with the following behavior:



- Spherically symmetric Cauchy data for the Einstein–Maxwell-charged scalar field system which undergo gravitational collapse.
- Forms an exactly Schwarzschild "apparent horizon."
- Forms an exactly extremal Reissner–Nordström event horizon at a later advanced time.
- These solutions are arbitrarily regular: for any $k \in \mathbb{N}$, there exists a C^k example.

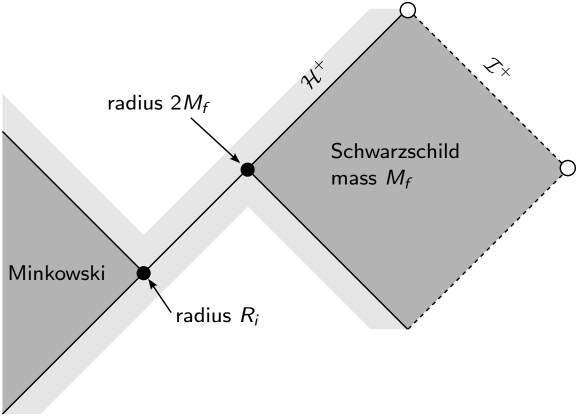
Einstein–Maxwell-charged scalar field system

- Lorentzian manifold $(\mathcal{M}^{3+1}, g)$
- 2-form $F = dA$ (electromagnetism)
- Charged (complex) scalar field ϕ
- Equations:

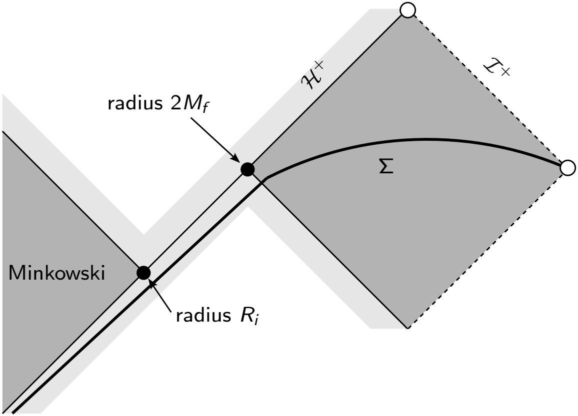
$$
\begin{aligned}
R_{\mu\nu}(g) - \tfrac{1}{2} R(g) g_{\mu\nu} &= 2\left(T^{\mathrm{EM}}_{\mu\nu} + T^{\mathrm{CSF}}_{\mu\nu}\right) && (1)\\
\nabla^{\mu} F_{\mu\nu} &= 2\mathfrak{e}\,\mathrm{Im}(\phi\overline{D_\nu\phi}) && (2)\\
g^{\mu\nu} D_\mu D_\nu \phi &= 0 && (3)\\
T^{\mathrm{EM}}_{\mu\nu} &= g^{\alpha\beta} F_{\alpha\nu} F_{\beta\mu} - \tfrac{1}{4} F^{\alpha\beta} F_{\alpha\beta} g_{\mu\nu} && (4)\\
T^{\mathrm{CSF}}_{\mu\nu} &= \mathrm{Re}(D_\mu\phi\overline{D_\nu\phi}) - \tfrac{1}{2} g_{\mu\nu} g^{\alpha\beta} D_\alpha\phi\overline{D_\beta\phi} && (5)
\end{aligned}
$$

- In an appropriate gauge, this is a nonlinear hyperbolic system for (g, F, ϕ).
- Can be thought of as an **initial value problem**.
- Spherical symmetry! Scalar field ϕ degree of freedom required to break the rigidity of pure Einstein–Maxwell in spherical symmetry.
- Scalar field carries charge in order to charge up the black hole to extremality.
- Model satisfies the dominant energy condition ($\Longrightarrow$ weak energy condition).

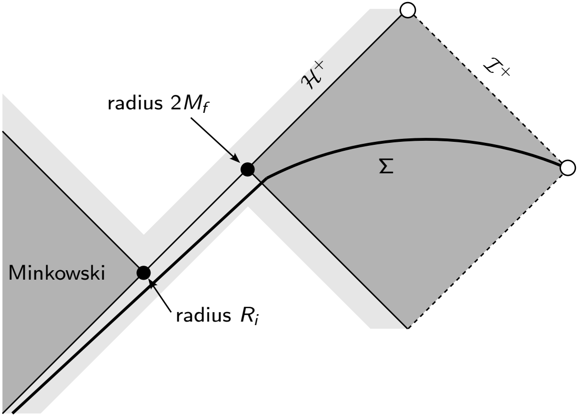
Prototype: Minkowski to Schwarzschild gluing



Prototype: Minkowski to Schwarzschild gluing



Prototype: Minkowski to Schwarzschild gluing



Characteristic gluing for the vacuum Einstein equations was initiated in fundamental work of ARETAKIS–CZIMEK–RODNIANSKI '21, but in a **perturbative** regime around Minkowski space which is inapplicable here.

The present problem is fundamentally **nonlinear** and **nonperturbative**, but is tractable because of the symmetry assumption.

Toy model: gluing for the linear wave equation

We now fix a spacetime $(\mathcal{M}^{3+1}, g)$ and consider the **linear wave equation**

$$\Box_g \phi = 0 \qquad (*)$$

Question.
Given two disjoint domains $\mathfrak{R}_1, \mathfrak{R}_2 \subset \mathcal{M}$, and functions

$$\phi_1 : \mathfrak{R}_1 \to \mathbb{R}$$
$$\phi_2 : \mathfrak{R}_2 \to \mathbb{R}$$

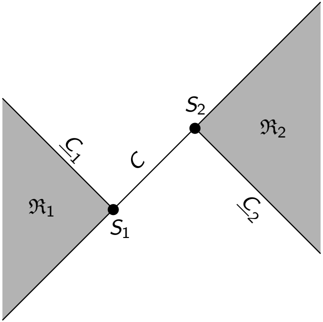
solving $(*)$, when does there exist a function

$$\phi : \mathcal{M} \to \mathbb{R}$$

solving $(*)$, such that

$$\phi|_{\mathfrak{R}_1} = \phi_1$$
$$\phi|_{\mathfrak{R}_2} = \phi_2\,?$$

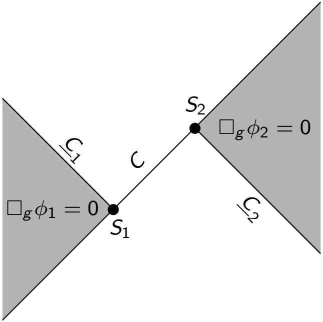
Characteristic/null gluing for the linear wave equation



Analogy:

- $\mathfrak{R}_1$ is Minkowski
- $\mathfrak{R}_2$ is Schwarzschild
- C will be the event horizon $\mathcal{H}^+$

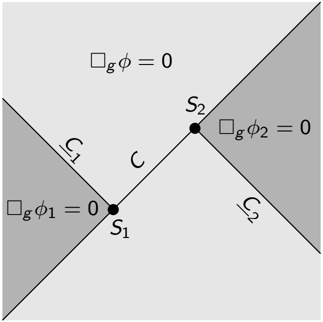
Characteristic/null gluing for the linear wave equation



Analogy:

- $\mathfrak{R}_1$ is Minkowski
- $\mathfrak{R}_2$ is Schwarzschild
- C will be the event horizon $\mathcal{H}^+$

Characteristic/null gluing for the linear wave equation



Analogy:

- $\mathfrak{R}_1$ is Minkowski
- $\mathfrak{R}_2$ is Schwarzschild
- C will be the event horizon $\mathcal{H}^+$

The characteristic initial value problem

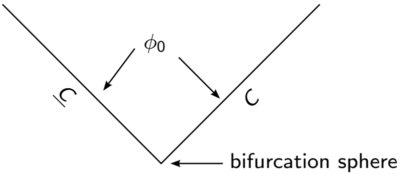
Bifurcate null hypersurface: $C \cup \underline{C}$ where C and $\underline{C}$ meet transversally, $C \cap \underline{C} \approx S^2$

Proposition (Characteristic IVP for the wave equation).

Given a bifurcate null hypersurface $C \cup \underline{C}$ and $\phi_0 : C \cup \underline{C} \to \mathbb{R}$, there exists a unique function ϕ defined to the future of $C \cup \underline{C}$ such that

$$\Box_g \phi = 0 \qquad (*)$$

and $\phi|_{C \cup \underline{C}} = \phi_0$.



The characteristic initial value problem

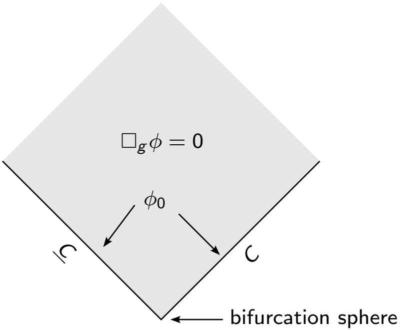
Bifurcate null hypersurface: $C \cup \underline{C}$ where C and $\underline{C}$ meet transversally, $C \cap \underline{C} \approx S^2$

Proposition (Characteristic IVP for the wave equation).

Given a bifurcate null hypersurface $C \cup \underline{C}$ and $\phi_0 : C \cup \underline{C} \to \mathbb{R}$, there exists a unique function ϕ defined to the future of $C \cup \underline{C}$ such that

$$\Box_g \phi = 0 \qquad (*)$$

and $\phi|_{C \cup \underline{C}} = \phi_0$.



The characteristic initial value problem

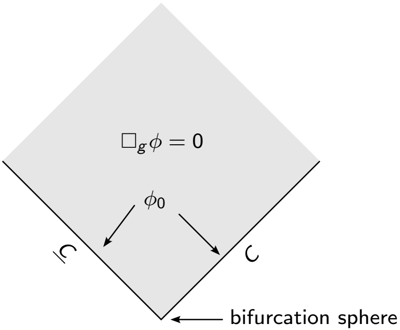
Bifurcate null hypersurface: $C \cup \underline{C}$ where C and $\underline{C}$ meet transversally, $C \cap \underline{C} \approx S^2$

Proposition (Characteristic IVP for the wave equation).

Given a bifurcate null hypersurface $C \cup \underline{C}$ and $\phi_0 : C \cup \underline{C} \to \mathbb{R}$, there exists a unique function ϕ defined to the future of $C \cup \underline{C}$ such that

$$\Box_g \phi = 0 \qquad (*)$$

and $\phi|_{C \cup \underline{C}} = \phi_0$.



Characteristic data for the wave equation does not involve derivatives.

The null "constraints"

In any well posed initial value problem, there needs to be a mechanism to compute the full solution on the initial data hypersurface. Toy problem for the toy problem:

- Wave equation on Minkowski $\mathbb{R}^{3+1}$
- $u = t - r$, $v = t + r$ double null coordinates
- $\phi = \phi(t, r, \vartheta, \varphi) = \phi(u, v, \vartheta, \varphi)$ satisfies

$$\Box\phi = -\partial_t^2\phi + r^{-2}\partial_r(r^2\partial_r\phi) + \not{\Delta}\phi = 0$$

$$\partial_u\partial_v(r\phi) = \not{\Delta}(r\phi) \qquad (*)$$

- Integrate $(*)$ on C:

$$\partial_u(r\phi)|_{v_2} - \partial_u(r\phi)|_{v_1} = \int_{v_1}^{v_2} \not{\Delta}(r\phi)\, dv'$$

- $\implies \partial_u\phi|_C$ can be calculated from (1) $\partial_u\phi$ at the bifurcation sphere and (2) $\phi|_C$
- $(*)$ lets us compute $\partial_u^i\phi$ on C to all orders

The null "constraints"

In general, the wave equation can be interpreted as a transport equation for the transverse derivative of ϕ, sourced by ϕ and tangential derivatives (which are a part of the characteristic data)

On C, the wave equation is equivalent to

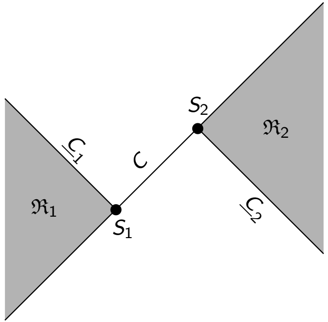
$$(\partial_v + \text{known function})\partial_u \phi = \text{known function},$$

where "known function" depends on underlying geometry and data.

Commuting the wave equation with transverse derivatives gives an inductive system of transport equations for all orders!

These equations have become known as the *null "constraints"*.

In what sense are these constraints?



The C^k characteristic gluing problem reduces to:

- Given two spheres S_1 and S_2 along an outgoing null cone C
- Given k ingoing and outgoing derivatives of ϕ at S_1 and S_2
- Prescribe ϕ along the part of C between S_1 and S_2 so that the outgoing derivatives agree with the given ones and the solutions of the transport equations for the ingoing derivatives have the specified initial and final values.
- Initial/boundary value problem for a system of coupled ODEs where you are allowed to (almost) freely prescribe the potential
- For the Einstein equations, we have additional constraints *tangent to C*

Obstructions to characteristic gluing

- For the linear wave equation, not all solutions can be characteristically glued.
- Example: Minkowski space

$$\partial_v \int_{S^2} \partial_u(r\phi)\, d\vartheta = \int_{S^2} \not{\Delta}(r\phi)\, d\vartheta = 0$$

$$\int_{S^2} \partial_u(r\phi)\, d\vartheta \text{ is } \textbf{conserved} \text{ for any solution of the wave equation}$$

- In general the solutions at S_1 and S_2 have to satisfy compatibility conditions.
- A complete geometric characterization of first order conservation laws for $\Box_g \phi = 0$ was given by Aretakis '17.
- These linear obstructions are a fundamental difficulty if one wishes to approach the nonlinear problem via linearization (Aretakis–Czimek–Rodnianski '21)
- The linearized Einstein equations around Minkowski space and extremal Reissner–Nordström have conservation laws of the above type.
- By exploiting "high frequency" perturbations, Czimek–Rodnianski '22 have shown how to avoid some of the conservation laws in the perturbative regime around Minkowski space (see also Mao–Oh–Tao '23).
- Remarkably, the presence of such conservation laws at the linear level is completely circumvented by our fully nonperturbative mechanism.

Einstein-scalar field system in spherical symmetry

$$\mathrm{Ric}(g) = 2\, d\phi \otimes d\phi \qquad \Box_g \phi = 0$$

- $\mathcal{M}^{3+1} = \mathcal{Q}^{1+1} \times S^2$

$$g = -\Omega^2 du dv + r^2 g_{S^2}$$

- $\Omega(u, v) > 0$ lapse, $r(u, v) > 0$ area-radius
- Wave equations

$$\partial_u \partial_v \phi = -\frac{\partial_u \phi \partial_v r}{r} - \frac{\partial_u r \partial_v \phi}{r} \tag{6}$$

$$\partial_u \partial_v r = -\frac{\Omega^2}{4r} - \frac{\partial_u r \partial_v r}{r} \tag{7}$$

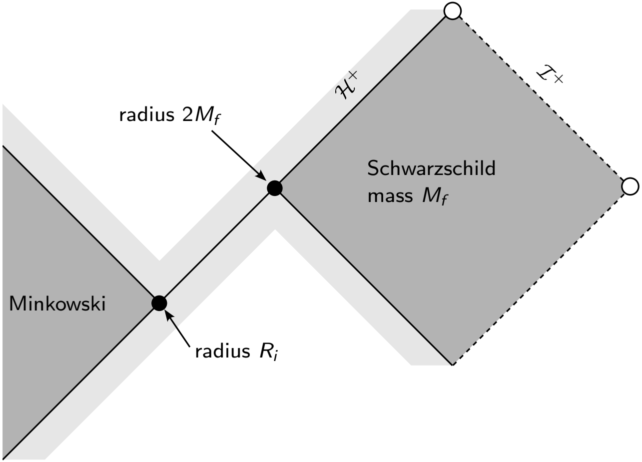
$$\partial_u \partial_v \log \Omega^2 = \frac{\Omega^2}{2r^2} + 2\frac{\partial_u r \partial_v r}{r^2} \tag{8}$$

- Raychaudhuri's equations (constraints)

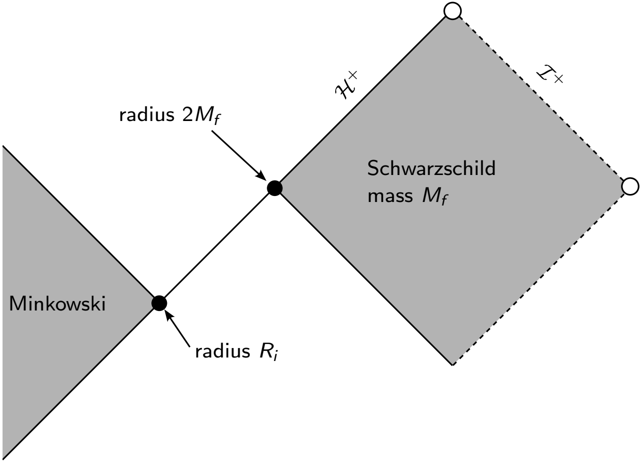
$$\partial_u \left(\frac{\partial_u r}{\Omega^2} \right) = -\frac{r}{\Omega^2} (\partial_u \phi)^2 \tag{9}$$

$$\partial_v \left(\frac{\partial_v r}{\Omega^2} \right) = -\frac{r}{\Omega^2} (\partial_v \phi)^2 \tag{10}$$

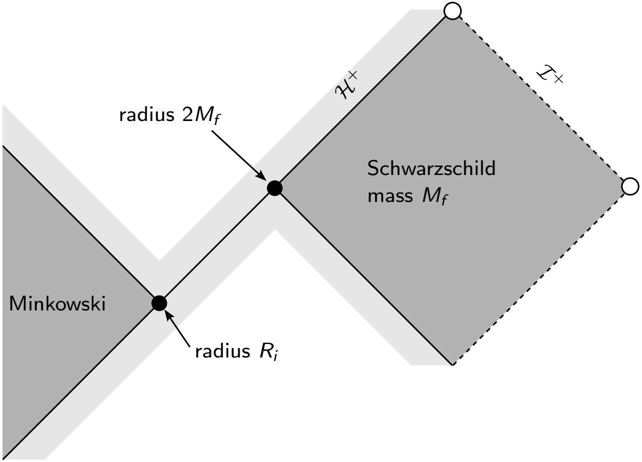
Minkowski to Schwarzschild gluing



Minkowski to Schwarzschild gluing

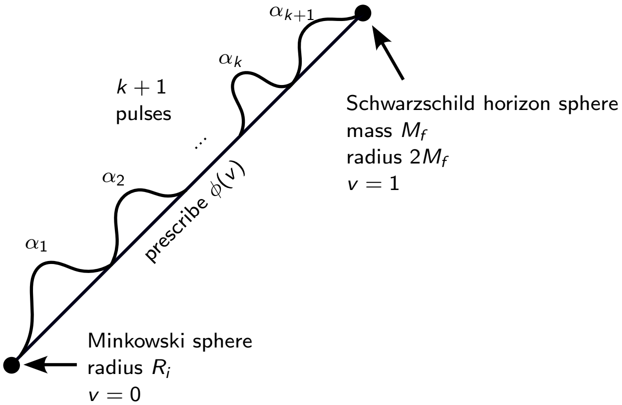


Minkowski to Schwarzschild gluing



Minkowski to Schwarzschild gluing

Theorem (Kehle–U. '22).
For any $k \in \mathbb{N}$ and $0 < R_i < 2M_f$, the Minkowski sphere of radius R_i can be characteristically glued to the Schwarzschild event horizon sphere with mass M_f to order C^k within the Einstein-scalar field model in spherical symmetry.



Idea of the proof

- Null constraint system becomes a coupled nonlinear system of ODEs sourced by $\phi(v)$
- Scalar field ansatz

$$\phi_\alpha(v) = \sum_{1\le j\le k+1} \alpha_j \chi_j(v), \quad \alpha \in \mathbb{R}^{k+1}$$

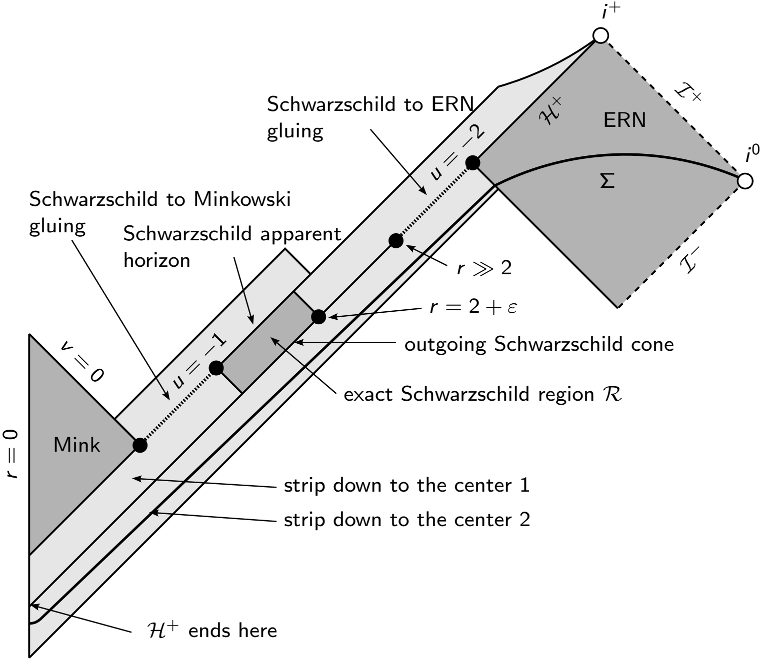
- Initialize $\partial_u \phi_\alpha(0) = \cdots = \partial_u^k \phi_\alpha(0) = 0$
- Initialize $r(v)$ teleologically at $v = 1$ ($\partial_v r(0)$ is a gauge choice, but $\partial_v r(1) = 0$!)
- For each α, generate $r(v)$, $\partial_u r(v)$, $\partial_u \phi_\alpha(v)$, etc.
- The antipodal map $\alpha \mapsto -\alpha$ leaves geometric quantities fixed (r, Ω^2, and derivatives)
- The antipodal map $\alpha \mapsto -\alpha$ changes the sign of the scalar field and its derivatives
- In particular, the map

$$\alpha \mapsto \left(\partial_u \phi_\alpha(1), \ldots, \partial_u^k \phi_\alpha(1)\right)$$

  is odd
- **Borsuk–Ulam theorem:** there exists α_* such that

$$\left(\partial_u \phi_{\alpha_*}(1), \ldots, \partial_u^k \phi_{\alpha_*}(1)\right) = 0.$$

Blueprint to disproving the third law



The vacuum case

We believe that the charged scalar field is not necessary to provide a counterexample to the third law:

Conjecture.
There exist Cauchy data for the Einstein vacuum equations

$$R_{\mu\nu} = 0$$

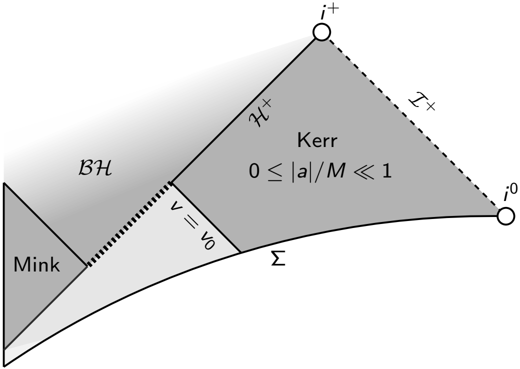
which undergo gravitational collapse and form an exactly Schwarzschild apparent horizon, only for the spacetime to form an exactly extremal Kerr event horizon at a later advanced time. In particular, already in vacuum, the "third law of black hole thermodynamics" is false.

Very related:

Conjecture (Thorne).
In an astrophysical accretion process, $|a|/M \leq 0.998$.

The very slowly rotating case

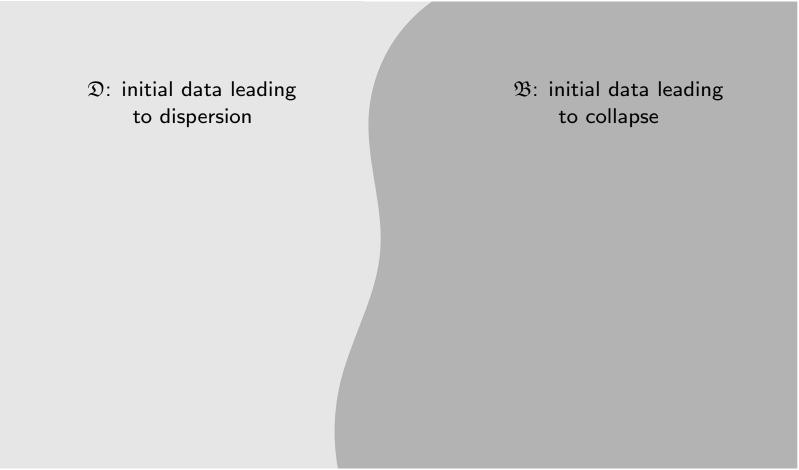
Theorem (Kehle–U. '23).
For any mass M and specific angular momentum $0 \leq |a| \ll M$, there exists a solution of the Einstein vacuum equations with the following Penrose diagram:



This provides an alternative, and fundamentally different, proof of CHRISTODOULOU's seminal theorem that *black holes can form due to focusing of gravitational waves alone.*

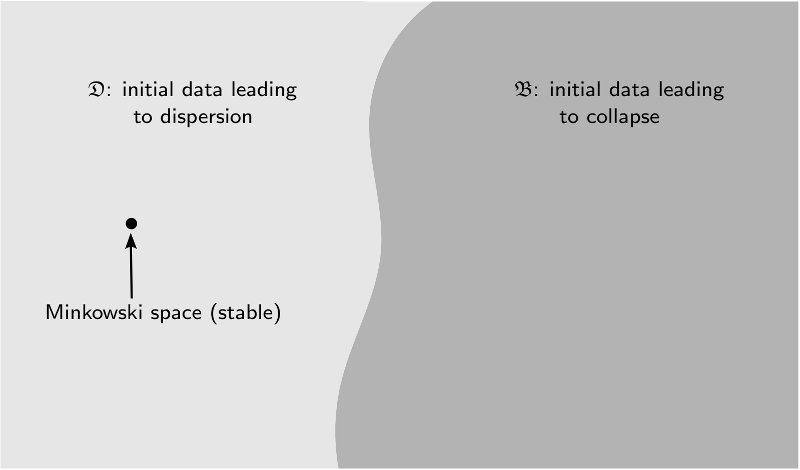
Critical collapse

$\mathfrak{M}$: the **moduli space** of one-ended asymptotically flat initial data
vacuum, scalar field, Vlasov, ..., spherical symmetry, ...



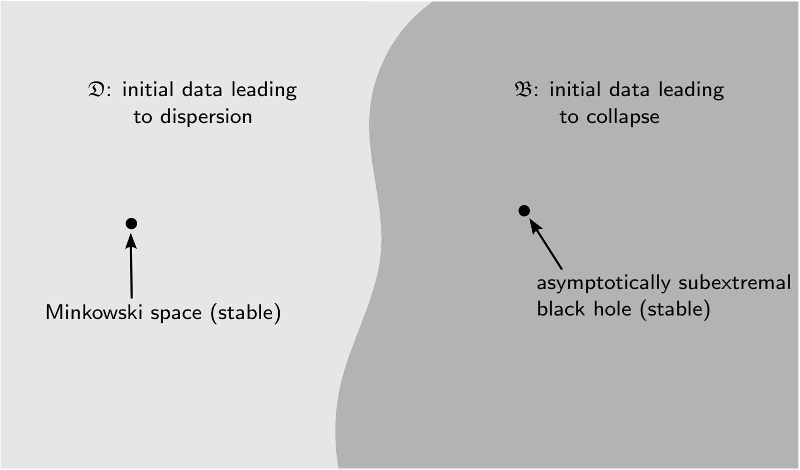
Critical collapse

$\mathfrak{M}$: the **moduli space** of one-ended asymptotically flat initial data
vacuum, scalar field, Vlasov, ..., spherical symmetry, ...



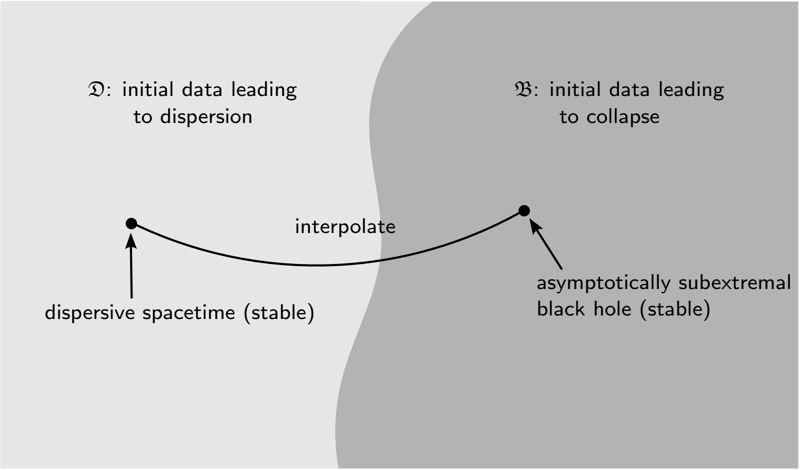
Critical collapse

$\mathfrak{M}$: the **moduli space** of one-ended asymptotically flat initial data
vacuum, scalar field, Vlasov, ..., spherical symmetry, ...



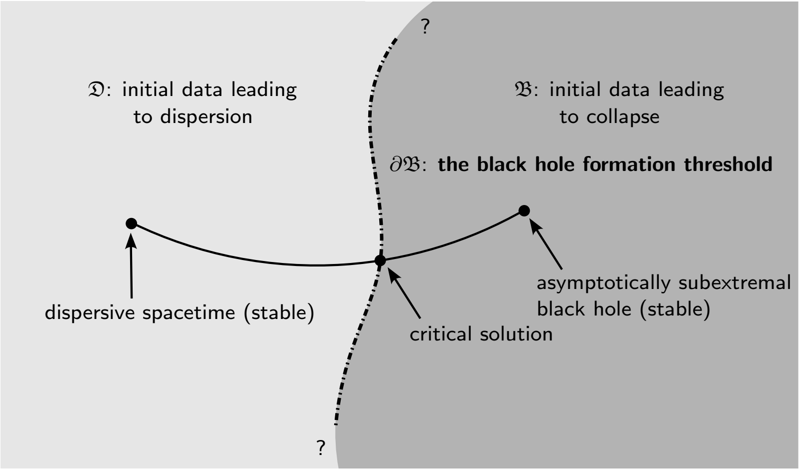
Critical collapse

$\mathfrak{M}$: the **moduli space** of one-ended asymptotically flat initial data
vacuum, scalar field, Vlasov, ..., spherical symmetry, ...



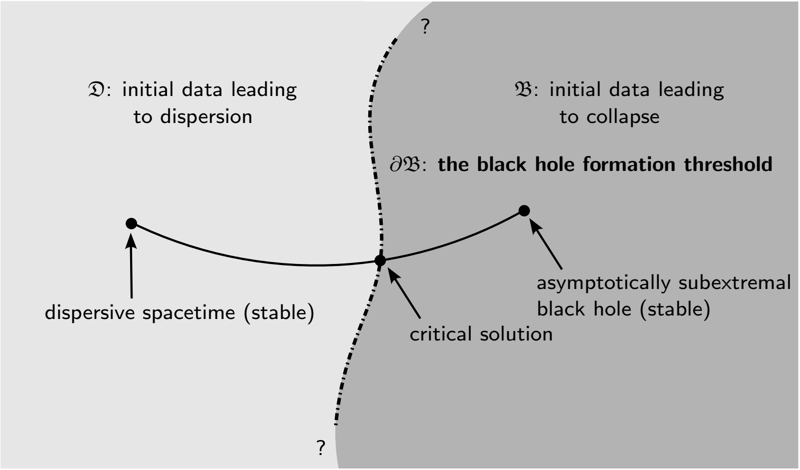
Critical collapse

$\mathfrak{M}$: the **moduli space** of one-ended asymptotically flat initial data
vacuum, scalar field, Vlasov, ..., spherical symmetry, ...



Critical collapse

$\mathfrak{M}$: the **moduli space** of one-ended asymptotically flat initial data
vacuum, scalar field, Vlasov, ..., spherical symmetry, ...

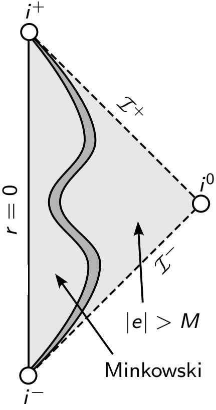


Critical collapse is the study of $\partial\mathfrak{B}$:
the **phase transition** between dispersion and gravitational collapse
pioneered in numerical work of CHOPTUIK '93 (**naked singularities**)

A new piece of the puzzle: extremal black holes as critical solutions

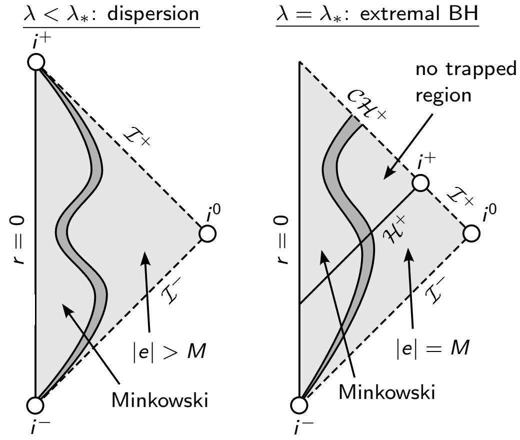
Theorem (Kehle–U. posted today).
Extremal critical collapse *occurs in the Einstein–Maxwell–Vlasov model. There exist smooth one-parameter families of solutions* $\{\mathcal{D}_\lambda\}_{\lambda\in[0,1]}$ *with the following behavior:*

$\lambda < \lambda_*$: dispersion



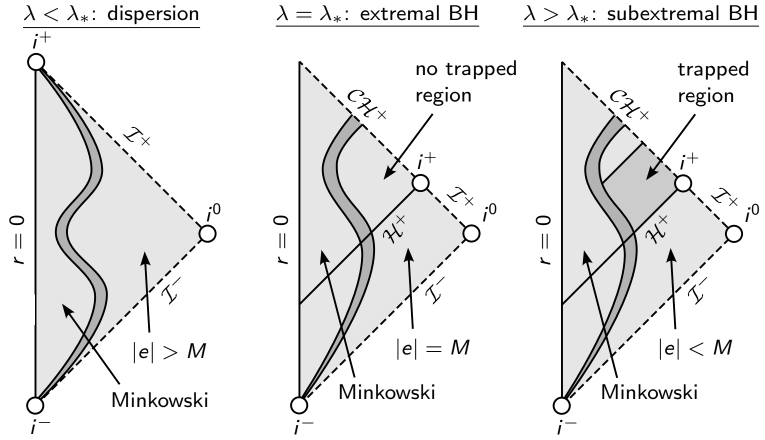
A new piece of the puzzle: extremal black holes as critical solutions

Theorem (Kehle–U. posted today).
Extremal critical collapse *occurs in the Einstein–Maxwell–Vlasov model. There exist smooth one-parameter families of solutions* $\{\mathcal{D}_\lambda\}_{\lambda\in[0,1]}$ *with the following behavior:*

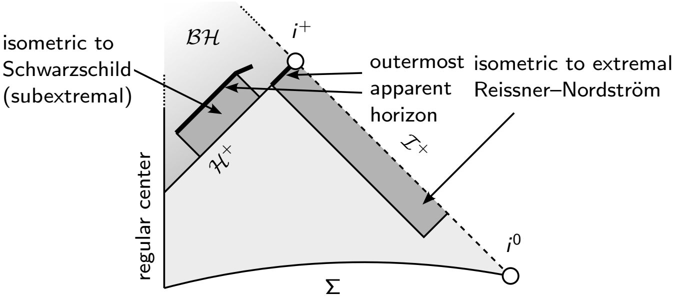


A new piece of the puzzle: extremal black holes as critical solutions

Theorem (Kehle–U. posted today).
Extremal critical collapse *occurs in the Einstein–Maxwell–Vlasov model. There exist smooth one-parameter families of solutions* $\{\mathcal{D}_\lambda\}_{\lambda\in[0,1]}$ *with the following behavior:*



- Extremal black holes can form in gravitational collapse
- Subextremal black holes can become extremal in finite time: **the third law of black hole thermodynamics is false**
- There is new **critical behavior** lurking near dynamical extremal black holes: this suggests a revised conjectural picture for the black hole threshold



Thank you!