A Survey of Foundational Problems about Classical Black Holes and the Hawking Effect

Erik Curiel

Munich Center For Mathematical Philosophy
Ludwig-Maximilians-Universität

Black Hole Initiative
Harvard University

Smithsonian Astrophysical Observatory
Radio and Geoastronomy Division

erik@strangebeautiful.com

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 312032894, CU 338/1-2. (Funded by the German Research Foundation (DFG) – project number 312032894, CU 338/1-2.)

Outline

What Is a Black Hole?

The Many Definitions of a Black Hole

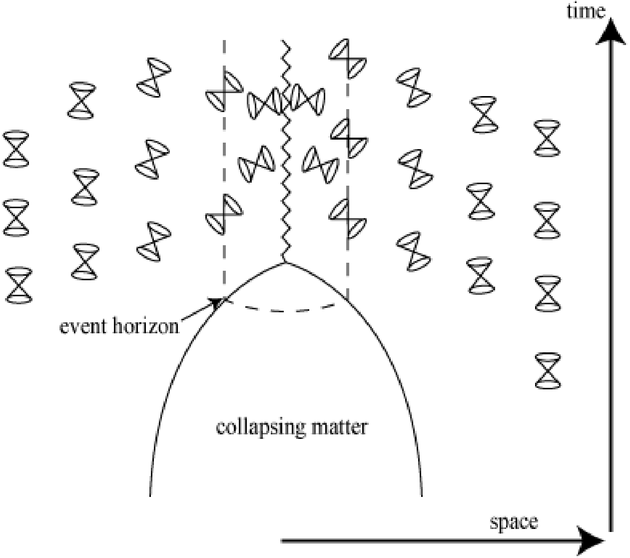
The Hawking Effect and Black Holes as Thermodynamical Systems

what is a black hole (crudely speaking)

- region of "no escape" (does <u>not</u> imply "gravity is strong")
- No Hair: completely characterized by mass, angular momentum, electric charge—like a fundamental quantum particle
- Chandrasekhar: "most perfect objects in the universe" (pure geometry)

Gravitational Collapse

- gravitational instability
  - collapse increases density $\rightarrow$ increases attraction $\rightarrow$ further collapse $\rightarrow$ ...
  - forces that counteract collapse: fluid stiffness, kinetic energy, angular momentum, thermal convection, pressure from nuclear reactions, ...
- stellar collapse
  - large stars: too much mass for forces to halt collapse after nuclear fuel consumed
  - when mass compressed into small enough volume, "intensity of gravity" stops even light from escaping
  - $\rightarrow$ formation of black hole


time
event horizon
collapsing matter
space

Observational Evidence

- black holes from stellar collapse
  - x-ray emitting binary systems (*e.g.*, Cygnus X-1—indirect but strong evidence)
- supermassive black holes
  - located in center of galaxies
  - quasars: energy produced as matter falls into black hole shoots out in large-scale jets (indirect but strong evidence)
  - our galaxy (SgrA*): motion of stars near galactic center (indirect but strong evidence)
  - Event Horizon Telescope: image of "shadow" of SgrA* and M87—DIRECT EVIDENCE!!!
- other black holes?
  - LIGO gravitational wave detection: binary black hole coalescence (indirect but strong evidence)

Event Horizon Telescope: M87

Why Black Holes Matter

1. they're incredibly cool
2. they're incredibly weird: can teach us much about conceptual structure of general relativity, what is (in some sense) physically possible in the world
3. black hole thermodynamics: one of most important, central, and fruitful fields of study in theoretical physics, closely connecting disciplines once seen as largely independent: cosmology, general relativity, quantum field theory, thermodynamics, statistical mechanics, particle physics, fluid dynamics, condensed matter, quantum optics, ...
4. widely considered best source of clues to a more fundamental theory of quantum gravity (simplest "purely gravitational" system to try to quantize; indirect access to micro-degrees of freedom and their dynamics by treating thermodynamics as arising from statistical mechanics)

Black hole thermodynamics and results concerning quantum fields in the presence of strong gravitational fields more generally are without a doubt the most widely accepted, most deeply trusted set of conclusions in theoretical physics in which general relativity and quantum field theory work together in seemingly fruitful harmony.

This is especially remarkable when one reflects on the fact that we have absolutely no experimental or observational evidence for any of it, nor hope of gaining empirical access any time soon to the regimes where such effects may appreciably manifest themselves.

Field	Core Concepts
astrophysics	► compact object ► region of no escape ► engine for enormous power output
classical relativity	► causal boundary of the past of future null infinity (event horizon) ► apparent horizon (all outgoing light rays "get turned around") ► quasi-local horizon
mathematical relativity	► apparent horizon ► singularity

Table: the core concepts common to different fields for characterizing black holes

(Curiel 2019, "The Many Definitions of a Black Hole", *Nature Astronomy*, 3:27–34)

Field	Core Concepts
semi-classical gravity	▶ same as classical relativity ▶ thermodynamical system of maximal entropy
quantum gravity	▶ particular excitation of quantum field ▶ ensemble or mixed state of maximal entropy ▶ no good definition to be had
analogue gravity	▶ region of no escape for finite time ("long" compared to characteristic time) ▶ same for low energy modes ("low" compared to characteristic energies)

Table: the core concepts common to different fields for characterizing black holes, cont.

(Curiel 2019, "The Many Definitions of a Black Hole", *Nature Astronomy*, 3:27–34) serious methodological and epistemological (and ontological?) problems:

1. event horizon is "global": makes implicit reference to "all future time"
2. that, and fact that nothing locally distinguishable about event horizon $\Rightarrow$ *no local measurements can ever determine its location, much less whether there is one*
3. the definitions used in classical general relativity do not match those in astrophysics
4. in particular, our universe is not "asymptotically flat", so it cannot have anything like an event horizon
5. *all* properties of black holes, all theorems, relied on by astrophysics assume event horizon—still applicable in real world?
6. how is SgrA* similar to and how different from a Kerr black hole?

different properties one may demand of a "black hole"

- possesses a horizon that satisfies the four laws of black hole mechanics
- possesses a locally determinable horizon
- possesses a horizon that is, in a suitable sense, vacuum
- is vacuum with a suitable set of symmetries
- defines a region of no escape, in some suitable sense, for some minimum period of time
- defines a region of no escape for all time
- is embedded in an asymptotically flat spacetime
- is embedded in a topologically simple spacetime
- encompasses a singularity
- satisfies the No-Hair Theorem

different properties one may demand of a "black hole" (cont.)

- is the result of evolution from initial data satisfying an appropriate Hadamard condition (stability of evolution)
- allows one to predict that final, stable states upon settling down to equilibrium after a perturbation correspond, in some relevant sense, to the classical stationary black hole solutions (Schwarzschild, Kerr, Reissner-Nordström, Kerr-Newman)
- agrees with the classical stationary black hole solutions when evaluated in those spacetimes
- allows one to derive the existence of Hawking radiation from some set of independent principles of interest
- allows one to calculate in an appropriate limit, from some set of independent principles of interest, an entropy that accords with the Bekenstein entropy (*i.e.*, is proportional to the area of a relevant horizon, with corrections of the order of $\hbar$)
- possesses an entropy that is, in some relevant sense, maximal
- has a lower-bound on possible mass
- is relativistically compact.

different subsets of these properties are used in different contexts in different investigations, often in the same field

but they are jointly inconsistent

$\Rightarrow$ no definition can accommodate all actual uses of the concept in contemporary physics

some philosophical issues

semantics of theoretical terms in scientific theories:

- we do not know what objects the name 'black hole' may appropriately pick out ⇒ no fixed relation of "designation" between words and objects (*e.g.*, *à la* Tarski)
- inconsistencies in its use ⇒ current usage cannot fix such a univocal relation
- ⇒ rigidity of reference based on "essential structure" (Kripke/Putnam cannot get off the ground
- no causal chains to black holes when concept was developed, nor even when term was coined in early 1960s, and no guarantee that those descriptions latched onto properties of any real systems ⇒ causal theory of reference (Kripke/Putnam) cannot work

Lessons:

1. difficult to see how standard truth-conditional semantics can work
2. meaning comes before reference!

(Curiel, "Schematizing the Observer and the Epistemic Content of Theories", forthcoming 2019, *SHPMP*; arXiv:1903.02182)

What Is a Black Hole?

The Many Definitions of a Black Hole

The Hawking Effect and Black Holes as Thermodynamical Systems

black hole thermodynamics, based on the Hawking effect, arises from the combination of our 3 best physical theories: general relativity, thermodynamics, and quantum field theory

general relativity

- best theory of space, time and gravity, classically conceived
- regime of applicability: intermediate energy and large spatiotemporal scales

thermodynamics

- study of relationship between work and heat for ordinary matter classically conceived, and how that constrains gross properties and behavior
- regime of applicability: low energy and intermediate spatiotemporal scales

quantum field theory

- best theory of fundamental properties and behavior of matter (ignoring gravity)
- regime of applicability: high (or *extremely* low) energy and small spatiotemporal scales

but how do they fit together?

- each conceptually and physically independent of the other two
- each finds application in regimes well separated from those of the others
- the most characteristic features of QFT—superposition, entanglement, uncertainty principle, non-locality—are in manifest tension if not outright contradiction with the principles of general relativity

first hint at black hole thermodynamics

No Hair $\Rightarrow$ black hole "macrostate"
completely independent of how it formed
(compatible with *any* "microstructure")!

just like ordinary thermodynamical systems

Zeroth Law

Thermodynamics

The temperature T is constant throughout a body in thermal equilibrium.

Black Holes

The surface gravity κ is constant over the event horizon of a stationary black hole.

First Law (Energy Conservation)

Thermodynamics

change in energy = (temperature × change in entropy) + work done

$$(\mathsf{d}E = T\mathsf{d}S + p\mathsf{d}V + \Omega\mathsf{d}J)$$

Black Holes

change in mass = (surface gravity × change in area) + "rotational work"

$$(\delta M = \frac{1}{8\pi}\kappa\delta A + \Omega_{\mathsf{H}}\delta J)$$

Second Law

Thermodynamics

entropy never decreases ($\delta S \geq 0$) in any physical process

Black Holes

area of event horizon never decreases ($\delta A \geq 0$) in any physical process

Third Law (Nernst Theorem)

Thermodynamics

$T = 0$ is not achievable by any physical process

Black Holes

$\kappa = 0$ is not achievable by any physical process

“Minus First Law” (Brown and Uffink)

Thermodynamics

isolated thermodynamical systems tend to approach a unique equilibrium state

Black Holes

isolated, non-stationary black holes tend to settle down to a unique stationary state (Kerr-Newman spacetime)

Laws of Black Holes Mechanics. . . or Thermodynamics?

classically, the analogy is purely formal, because temperature = 0 irrespective of value of κ (according to orthodoxy)

1. black holes are perfect absorbers, they emit nothing
2. infamous Geroch thought-experiment: use a classical black hole as heat sink to transform heat into work with 100% efficiency

(Alain Connes: "Laws of Asshole Mechanics")

but take quantum effects into account. . .

Hawking discovered that a black hole behaves like a perfect black body in the sense of ordinary statistical thermodynamics: in the presence of a quantum field in its vacuum state, a stationary black hole emits thermal radiation with the Planckian power spectrum characteristic of a perfect blackbody at a fixed temperature equal to κ. It glows like a lump of smoldering coal even though nothing, not even light, should be able to escape from it!

$\Rightarrow$ now we can take the formal analogy to be a truly physical one

Generalized Second Law

- seems easy to violate standard Second Law when black holes are around:
  1. throw favorite highly entropic system into black hole
  2. the entropy of the world outside the event horizon—a causally isolated system—spontaneously decreases
- $\Rightarrow$ Bekenstein proposed Generalized Second Law: *total* entropy, black hole (area) + ordinary matter outside, never decreases
- many powerful (purely theoretical!) arguments supporting it

overwhelming consensus today is that black holes truly are thermodynamical objects, and the laws of black hole mechanics just are the laws of ordinary thermodynamics extended into a new regime, to cover a new type of physical system

(compare extension of thermodynamics to encompass electromagnetic blackbody radiation at end of 19th Century) but what on Earth can that mean?!

how can an empty, locally undistinguished region of spacetime have thermodynamical properties?

- difficult to think of two more different quantities than entropy and spatial area. . .
- unless it be temperature and surface gravity
- how can these possibly be *identical*?
- $\Rightarrow$ *deep* problem for conceptual understanding of inter-theory relations: "same" quantity as represented in different theories

also:

- laws of ordinary thermodynamics are empirical generalizations
- laws of black hole mechanics are theorems of differential geometry
- how can they possibly be "the same"?

requires radical changes to picture of ontology of spacetime and matter?

quantum field theory

- **matter** excitations in fields
- **spacetime** fixed, not affected by matter (Minkowski spacetime)

general relativity

- **matter** non-zero mass-energy ($T_{ab} \neq 0$, Ricci curvature)
- **spacetime** dynamical, affected by matter (zero mass-energy $\Rightarrow$ only Weyl curvature)

black hole thermodynamics spacetime curvature *transformed into* matter and matter *transformed into* spacetime curvature conceptual problem with Hawking radiation: *it is not standard blackbody radiation!*

it is not generated by micro-degrees of freedom of the event horizon, like electromagnetic blackbody radiation of hot iron is caused by jiggling of its atoms and electrons, but rather is excited modes of external quantum field

so why should we take it as a proxy for the *black hole's* temperature?

perhaps most puzzling of all

How did a theoretically predicted phenomenon (Hawking radiation), derived by combining seemingly incompatible theories in a novel way so as to extend their reach into regimes that we have no way of testing in the foreseeable future, constrained only by principles based on physical intuition, become the most important touchstone for testing novel ideas in theoretical physics? Can it play that role? What epistemic warrant do we or can we have for it in the end?