Two Paths to the Einstein Field Equation from Horizon Thermodynamics

Erik Curiel

Munich Center For Mathematical Philosophy
Ludwig-Maximilians-Universität

Black Hole Initiative
Harvard University

Smithsonian Astrophysical Observatory
Radio and Geoastronomy Division

email: `erik@strangebeautiful.com`

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 312032894.
(Funded by the German Research Foundation (DFG) – project number 312032894.)

Outline

thermodynamics	**black holes**
isolated	asymptotically flat
equilibrium	stationary
M, Ω, Q	p, T
surface gravity	temperature
area	entropy

Zeroth Law surface gravity κ constant over event horizon of stationary black hole

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

Third Law $\kappa = 0$ not achievable by any physical process

Motivation

The four laws of black hole mechanics, which are analogous to those of thermodynamics, were originally derived from the classical Einstein equation. With the discovery of the quantum Hawking radiation, it became clear that the analogy is, in fact, an identity. How did classical general relativity know that the horizon area would turn out to be a form of entropy, and that surface gravity is a temperature?

Ted Jacobson (1995)

Aim

> *[The Einstein field equation] is born in the thermodynamic limit as a relation between thermodynamic variables, and its validity is seen to depend on the existence of local equilibrium conditions.*
>
> Ted Jacobson (1995)

resolve puzzle: derive Einstein field equation from thermodynamics, not vice-versa that's how general relativity "knows": Einstein field equation as "equation of state", just like ordinary thermodynamics "knows" equilibrium maximizes entropy and minimizes free energy, that entropy is $\frac{\delta Q}{T}$, that temperature mediates coupling and characterizes equilibrium, and so on

Assume: Background Spatiotemporal Structure

1. causal horizons (conformal structure)
2. accelerated timelike curves (affine structure)
3. $\Rightarrow$ full metric (Weyl's Theorem)

all classical

Assume: "Local" Rindler Horizons

1. "locally flat"
2. "static" on the horizon itself
3. generated by "almost Killing fields" near horizon (timelike curves of "constant acceleration")
4. $\Rightarrow$ dynamical and thermodynamical equilibrium

all classical

Assume: Unruh Temperature

semi-classical effect

Assume: Functional Form of Entropy

proportional to "area":
classical or semi-classical?

Recall Ordinary Thermodynamics

1. entropy as function of energy and volume ($S = f(E, V)$)
2. Clausius relation ($\delta Q = T\mathrm{d}S$)
3. First Law ($\delta Q = \mathrm{d}E + p\mathrm{d}V$)
4. $\Rightarrow$ equation of state ($p = T\partial S/\partial V$)

   (e.g., $pV \propto T$ from $S \propto \ln V + f(E)$)

Basic Structure of Argument

1. assume form of entropy (proportional to area)
2. assume form of heat (matter flux of particular sort, $= 0$ when $T_{ab} = 0$)
3. assume Clausius relation (temperature from Unruh effect)
4. $\Rightarrow$ Einstein field equation as consistency condition on thermodynamical relations ("equation of state")

matter flux

stress-energy tensor null-component *tangent* to horizon—Killing energy

conserved along horizon $\Rightarrow$ equilibrium, $\Delta S = 0$, as should be since horizon is "static"

Basic Structure of Argument

1. assume form of entropy (proportional to area)
2. assume form of heat (matter flux of particular sort, $\neq 0$ when $T_{ab} = 0$, quasi-local "energy" $\propto {}^{(2)}R$)
3. assume form of work ("force integral" of particular sort, associated with matter flux)
4. $\Rightarrow$ Einstein field equation "as equation of state" equivalent to First Law

matter flux

- stress-energy tensor null-component *transverse* to horizon—thus “work” term absent in Jacobson's derivation
- includes ${R^a}_{bcd}$ terms: conformal structure, so “gravitational energy”?

Perhaps not surprising: just a special case of a strengthened form of Lovelock's Theorem in disguise (G_{ab} is only symmetric, covariantly divergence-free concomitant of $R^a{}_{bcd}$ with units of stress-energy)

Existence of Local Rindler Horizons

justified by invocation of the Equivalence Principle ("locally, spacetime looks like Minkowski spacetime"), but that isn't a good argument: in required form, Equivalence Principle is either false or beside the point

1. can't "transform away" $R^a{}_{bcd}$
2. to derive non-trivial relation involving R_{ab}, one can't ignore $R^a{}_{bcd}$
3. non-trivial T_{ab} flux $\Rightarrow$ generically *not static*
4. ignores shear and twist

Well Defined Temperature?

- Unruh temperature defined for unique KMS state invariant under action of isometries generating the horizon, which is the (relevant) vacuum state
- but both derivations assume non-trivial matter-flux: not a vacuum!
- can't assume it's "test matter", otherwise wouldn't get $G_{ab} \propto T_{ab}$, only $R_{ab} = 0$

Possible Response

matter is "weak enough" and "static enough" that the local state is "close enough" to vacuum that the deviations from thermality are negligible

is there a way to make this precise, while accounting for non-linearity of Einstein field equation?

First Stab (Not Very Precise!)

1. horizon entropy is huge (entanglement of UV degrees of freedom?)
2. view as thermal entropy: integral of T^3, where T is temperature measured locally by "almost static" observers, diverging as $1/\xi$ for observers ξ distance from horizon
3. excitation of state corresponding to stress-energy flux is "IR" phenomenon, negligible entropy and temperature
4. $\Rightarrow$ to system outside horizon, excitation enters thermal bath at the horizon and is "thermalized"

Away from Equilibrium?

two possibilities

1. calculate non-equilibrium corrections (higher-order "transport" terms) required for consistency with Einstein field equation
2. constrain non-equilibrium corrections based on other physical principles (*e.g.*, GSL), as observable deviations from Einstein field equation

both $\Rightarrow$ possible clues to underlying statistical theory, and so to quantum gravity?

deviations would be *miniscule* in non-extreme FLRW—explains why deviations not yet observed?

fascinating project: work through differences between the two, see why they still both work—but here I'm focusing on similarities, shared problems and issues

Jacobson: Matter Flux as "Heat Flow"?

pro *really* care only about *stress-energy* flux, what "couples to curvature"—exactly like heat, independent of idiosyncracies of particular forms of matter

con "energy", especially gravitational, ill defined in general relativity—why should Killing energy and particular, special form of quasi-local energy (First Law) be relevant, and even more the relation between them?

Padmanabhan: Matter Flux as “Doing Work”?

1. “gravitational force” not unambigously defined in general relativity, especially for accelerated matter (“gravity as attractive force” for geodesics from Rauchaudhuri equation is already problematic, and not applicable to accelerated curves)
2. interpreting term multiplying dV as “pressure” in First Law works only for particular form of stress-energy tensor (Hawking-Ellis type I)—not all matter fields of interest are of this type
3. why is *volume* relevant? all problems with defining volume for interior of black holes...

deriving a classical field equation from a semi-classical premise?

- in regime where Unruh temperature physically significant, we don't expect classical Einstein field equation to hold for "non-quantum" matter
- where Einstein field equation holds for "non-quantum" matter, how can Unruh temperature be physically significant?

Black-Hole Laws *versus* Ordinary Thermodynamical Laws

theorems of differential geometry

versus

empirical generalizations (albeit *profoundly* entrenched), theoretically "justified" by SM and likelihood of initial states based on dynamical arguments (*very* complex evidential relations)

presumably, "thermodynamics" underlying gravitation arises itself from underlying statistical mechanics of micro-degrees of freedom, same as ordinary thermodynamics deriving a non-thermodynamical relation from thermodynamics and interpreting it as thermodynamical

versus

deriving a thermodynamical relation from thermodynamics—which, and why?

in part depends on what one assumes (fixed principles, interpretation of quantities, *etc.*), and what one derives and how one jusitifies interpretation: has to be more than formal identification from relations that look like thermodynamical ones

physical significance of assumptions

Clausius Relation vacuous? just definitive of "entropy"?

First Law vacuous? "heat" usually defined as all the energy not account for in work (see Eddington's marvelous remarks)

physical significance of Einstein field equation?

"equation of state":

- "consistency condition" is *not* status of equation of state in ordinary thermodynamics
- neither is First Law

gets at something physically substantive or just interpretive move, giving new conceits based on identities for old quantities?

Limits of Propriety

“near equilibrium”

same as ordinary thermodynamics

what about wildly non-static spacetime regions, *e.g.*, AGNs?
we have good evidence that Einstein field equation obeyed there

quantum gravity all misguided?

if correct, it's a category mistake to "quantize" Einstein field equation—would be like quantizing ideal gas law

but what about phonons?

in sum. . .

enticingly suggestive, seems to get at something deep

BUT WHAT?