Black Hole Thermodynamics: More Than an Analogy?

John Dougherty and Craig Callender

October 4, 2016

Analogies can be tremendously fruitful in physics. The history of physics is filled with one successful analogy after another: the atom as plum pudding, electric current as water flowing through a pipe, and so on. One of the most famous instances of this method is Maxwell's use of an analogy between the streamlines of incompressible fluid flow and Faraday's magnetic field lines of force. With the magnetic flux density playing the role of the fluid velocity, Maxwell was able to find fluid counterparts for the six known empirical "laws" of electric and magnetic phenomena known in the 1850s. This analogy motivated many of his later discoveries.

As physics today searches for insights into a theory of quantum gravity, another analogy has become so entrenched that it has taken on a life of its own: black hole thermodynamics (BHT). Originating in parallels between the area of a black hole's event horizon and the thermodynamic entropy established by Bekenstein, Hawking, and others in the 1970s, counterparts of thermodynamics' famous four laws were found for classical black holes, much as Maxwell had earlier found fluid counterparts for electric and magnetic phenomena. However, according to the standard lore, the 1974 discovery of Hawking radiation significantly tightened the analogy. Now BHT is widely proclaimed to be "more than a formal analogy," unlike Maxwell's use of fluid. In particular, the relationship is said to be one of identity: "the laws of black hole thermodynamics … simply *are* the ordinary laws of thermodynamics applied to a black hole" (Wald, 1994, 174). Relationships among various black hole variables are understood as the manifestations of deep thermodynamic principles operating in the universe. What causes your tea to come to room temperature is said also to cause the area of black holes to increase. Not only is this surprising analogy said to be more than formal, but it forms the basis of some very popular speculations in quantum gravity.

We want to pour a little cold water on the claim that BHT is more than a formal analogy. Analogies can be good or bad. When good they can motivate new ideas and progress, but when bad they can lead astray. Indeed, Maxwell's analogies, pushed to an extreme, led to an aether composed of vortex tubes connected by frictionless "idle wheels." The history of science is littered with good and bad analogies. BHT may turn out to be a useful and deep guide to future physics. Only time will tell. Conventional wisdom in physics seems entirely behind the idea that the analogy is as strong as it gets. In the face of such unanimity, philosophers have a duty to play the "gadfly," challenging not merely societal assumptions as Socrates did, but also widely shared assumptions in physics. By showing that the analogy is not as strong as is commonly supposed, we wish to recommend caution.

In what follows we report a few problems with the analogy between thermodynamics and black holes and the main argument that seeks to join the two together as one. We focus on three worries. The first is that BHT is often based on a kind of caricature of thermodynamics. The second is an even more basic worry, namely, that it's unclear what the systems in BHT are supposed to be, and furthermore, it's unlikely that any resolution of this question leaves the conventional wisdom surrounding the subject intact. The third, and perhaps worst, is a challenge to the main argument that identifies black hole entropy with thermodynamic entropy. The combined effect of these worries dramatically decreases our confidence that BHT rests on more than a formal analogy.

Readers already familiar with BHT may jump directly to section 2; but for those new to the topic, we present an introduction in the next section.

1 Black Hole Thermodynamics, Information Loss, Holography, and All That

In 1782, the astronomer John Michell argued that a sufficiently massive star would have such a strong gravitational pull that its light could never escape. One hundred and fifty years later, physicists showed that such astronomical bodies could form as heavy stars collapse at the end of their lives. The singularity theorems of the 1960s, by Stephen Hawking, Roger Penrose, Robert Geroch, and others, proved that we ought to expect such collapsed objects in our universe. By the middle of the 1970s, these collapsed objects were standard astronomical fare, named "black holes" after their characteristic ability to prevent all light from escaping.

Investigations into the evolution of black holes showed that their behavior can be described by simple rules involving a handful of quantities. This is much the same as the situation in equilibrium thermodynamics, so these rules are analogously called the laws of black hole thermodynamics. There are some immediately obvious disanalogies. The laws of equilibrium thermodynamics are phenomenological laws, and when they are underwritten by statistical mechanics they only hold with very high probability. The laws of BHT, by contrast, are theorems of differential geometry, and many of them are independent of the specific laws of general relativity. Nevertheless, the analogy is striking, and it has played a large role in the development of quantum theories of gravity.

1.1 The laws of black hole thermodynamics

The laws of BHT are mathematical facts about certain kinds of geometric spaces, given physical significance by their use in general relativity. Each of the laws makes different assumptions about the system under consideration, but they are all concerned with a particular kind of gravitationally isolated system. That is, one thinks of the system as the only thing in the universe, so that far away from the object spacetime is flat, like the spacetime of special relativity. BHT assumes further that there is an event horizon, so that no light can ever escape from some region of the system. The event horizon is a two-dimensional surface which can be characterized geometrically, so from a mathematical point of view the laws of BHT are theorems about certain two-dimensional surfaces in four-dimensional spaces—one of time and three of space.

The geometrical character of an event horizon can be seen with the help of a Penrose diagram, like the one shown in Fig. 1. This diagram is a two-dimensional

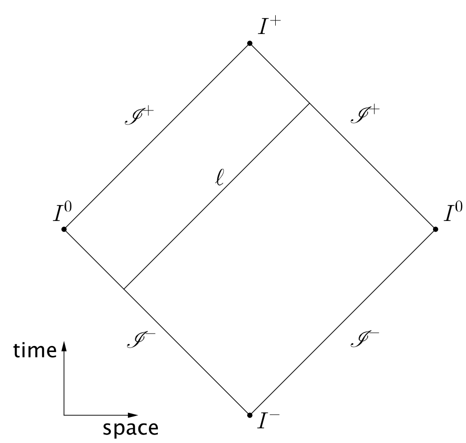


Figure 1: A Penrose diagram

slice of the spacetime. The spacetime itself is represented by the region inside the borders of the diagram; the borders represent theoretical "points at infinity". In particular, the lines $\mathscr{I}^+$ are the future endpoints of light rays that escape to infinity. The worldline of any possible light ray is given by a diagonal line from $\mathscr{I}^-$ to $\mathscr{I}^+$, like the line ℓ in the diagram. This represents a light ray that starts infinitely far away from the system in the past, then passes by the system and travels infinitely far away in the future.

In Fig. 1 every diagonal line ends on $\mathscr{I}^+$. This means that every ray of light eventually escapes to infinity, so there are no event horizons. Fig. 2, by contrast, is a Penrose diagram for a collapsing star. Here we have suppressed the two angular spatial coordinates rather than two rectilinear spatial coordinates as in Fig. 1. In the far past, at the bottom of the diagram, the star is at rest in the center

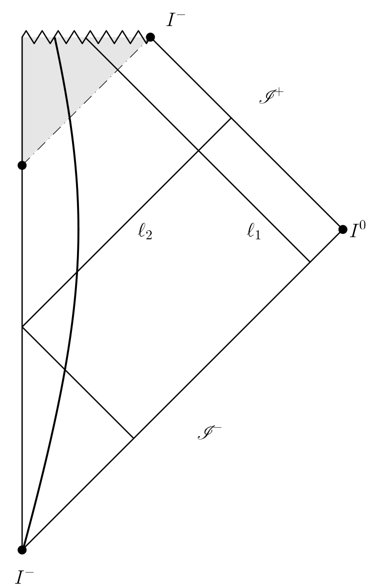


Figure 2: A collapsing star

of the spacetime. As time passes and we travel up the diagram, the radiation source powering the star depletes, and the distance from the center of the star to the curved line representing its boundary shrinks. The pressure from gravity eventually overwhelms the star and an event horizon forms at the spacetime point marked in the diagram. A singularity forms, represented by the zig-zag at the top. Some light rays, like ℓ_1, now terminate at the singularity instead of $\mathscr{I}^+$. So only some light, like ℓ_2, escapes to infinity. In fact, any light that enters the shaded region will terminate at the singularity. So the shaded region is the black hole region, and the border of the black hole region, marked by a dashed line, is the event horizon. Since we can describe it geometrically, we can apply the tools of differential geometry to make predictions about it.

The event horizon at a fixed time is (topologically) a sphere with a well-defined area. The Second Law of BHT says that this area will not decrease in any physical process—analogous to entropy's nondecreasing nature in thermodynamics. In this context, a process is "physical" if it satisfies two conditions: the null energy condition and the cosmic censorship hypothesis. The first of these can be interpreted as an assumption about matter. Roughly, it says that gravity is attractive; two infinitesimally close light rays subject only to gravitational forces will be drawn toward one another. The second condition is a form of determinism. It asserts that the only singularities in spacetime are hidden behind event horizons. There is no telling what will happen near a singularity. If the singularity is safely hidden behind an event horizon, then nothing that happens there can affect the spacetime outside. The first proof of the Second Law was published by Hawking in 1971. The proof is a *reductio*: supposing that the event horizon shrinks, it follows from the null energy condition that light near the event horizon will be focused into a point outside the black hole. Focusing all this light into one point results in a singularity as the light beams collide. But by the cosmic censorship condition, this is not allowed. So the event horizon cannot shrink.

The discussion to this point has been essentially global. In order to identify the black hole region, and hence the event horizon, we have to know what happens infinitely far in the future and in space. This is counterintuitive. The picture of a collapsing star alone in the universe seems entirely local: the star shrinks until an event horizon forms, then it eventually collapses behind the event horizon, leaving a black hole alone in the universe. Since the radius of the event horizon can be calculated from the properties of the star, and we know where the star collapsed, we should be able to find the event horizon any time after it forms, without looking far into the future.

Like most, this divination rests on a trick. We know what will happen in the future because the black hole is assumed stationary; it may rotate, but it will not otherwise move. Moreover, it is spatially symmetric; either it rotates around an axis of symmetry, or it is a still sphere. A spacetime with these two symmetries is rigidly constrained, so any spacetime ripples far away in the future can be felt throughout the spacetime. This is how we are able to know where the event horizon is shortly after a stellar collapse—the symmetries of the spacetime allow the event horizon to be picked out by local properties. They allow us to say which observers are stationary with respect to the spacetime structure. Outside of the event horizon, stationary observers move through time at a fixed distance from the black hole. Inside, they travel toward its center. Geometrically, these symmetries are represented by Killing vector fields. In the case of stellar collapse, the event horizon is a Killing horizon, which is a surface picked out by the Killing vector fields. A collection of results called the rigidity theorems show that for any stationary black hole, the event horizon will also be a Killing horizon. These theorems tie the global characterization of the event horizon to local properties in spacetime and give further geometrical information about the surfaces we're interested in.

The rest of the laws of BHT are theorems about Killing horizons in stationary spacetimes. The Zeroth Law of thermodynamics implies that temperature is constant everywhere in a body in equilibrium. The Zeroth Law of BHT states that the surface gravity of a Killing horizon is constant everywhere on the horizon. Physically, the surface gravity κ can be characterized by the following experiment. Suppose that you are a stationary observer far from the Killing horizon. Attaching a rope to a box of mass m, you lower the box toward the black hole until it is on the event horizon and keep it stationary there. The box will feel two counteracting forces, from you and the black hole, and because it is stationary these forces will cancel out. Numerically, it feels you applying a force with magnitude $m\kappa$. In Newtonian gravity, a body with mass m lowered to a spherical object feels a gravitational force of magnitude mg, where the gravitational acceleration g depends only on the mass of the spherical object. So, κ is naturally interpreted as the gravitational acceleration felt by a body at the event horizon.[1]

The Third Law of BHT is also an analogy between temperature and surface gravity, but it has a rather different status from the other laws in the analogy. The Third Law of thermodynamics states that the entropy of a system approaches a universal constant as its temperature approaches absolute zero. The analogous claim fails in BHT—a black hole with vanishing surface gravity may have any area. However, it is a consequence of the Third Law of thermodynamics that no finite series of operations may reduce a system to zero temperature. Bardeen et al. (1973) offered the analogous claim as the Third Law of BHT: no finite series of operations may reduce the surface gravity of a black hole to zero. Unlike the other laws, Bardeen et al. did not offer a proof of this Third Law, only giving plausibility arguments. By and large, the Third Law has second-class status. But since it is a problematic law even in the thermodynamic context, finding an adequate statement and proof of its counterpart in BHT has attracted less attention than the other laws.

Finally, the First Law of thermodynamics is a statement of conservation of energy for thermodynamic systems. In any thermodynamic process, the change in a system's internal energy is given by the difference between heat added to the system and work produced by the system. If the system is in equilibrium, then its thermodynamic state can be characterized by a collection of intensive quantities: the temperature T, the pressure p, the rotational velocity Ω, and others. Assuming the Second Law, the First Law of thermodynamics for equilibrium systems may be expressed as the Gibbs relation, which gives the change in internal energy U in terms of these intensive quantities and their extensive partners

$$dU = T\,dS + p\,dV + \Omega\,dJ + \phi\,dQ$$

Here S is the entropy of the system, V its volume, J its angular momentum, ϕ its electric potential, and Q its charge. If the system has further intensive properties, like charge or chemical potentials, these have their own extensive partners and contribute further terms to the Gibbs relation. Note the assumptions required to obtain this relation. First, the system was assumed to be in equilibrium. We

[1]The Zeroth Law has many proofs, but perhaps the most general is by Kay and Wald (1991).

then had to assume the Zeroth Law of thermodynamics to count T an intensive quantity and the Second Law of thermodynamics to write the quantity of heat added to the system as $T\,dS$.

Like thermodynamic systems in equilibrium, stationary charged black holes can be characterized by a handful of parameters: their mass, angular momentum, and electromagnetic charge. In other words, black holes have no "hair" that can be styled to distinguish two with the same mass, area, etc.[2] The First Law of BHT states that changes in the mass M of a neutral black hole are related to changes in the area A and angular momentum J as

$$\delta M = \frac{1}{8\pi}\kappa\,\delta A + \Omega\,\delta J + \phi\,\delta Q$$

where κ is the surface gravity, Ω the angular velocity, and Q the charge of the black hole, and ϕ is the value of the electromagnetic potential on the horizon, which is constant. Taking M to be the analogue of internal energy and A the analogue of entropy, this equation is analogous to the Gibbs relation. Like the Gibbs relation, the First Law of BHT assumes that the system is in equilibrium, namely, stationary. Also like the Gibbs relation, derivations of the First Law of BHT assume that the surface gravity is constant over the body—in this case, the event horizon. As always, the proof of the First Law of BHT is a theorem of differential geometry, and the quantities involved are geometric invariants associated with the symmetries of the spacetime. Unlike the other laws of BHT, however, the proof of the First Law does not require any assumptions about the matter—at least not directly. Because the Zeroth Law is assumed in the derivation of the first, and some proofs of the Zeroth Law rely on an energy condition, the First Law would inherit this dependence.

In sum, the laws of BHT are four geometric facts about spacetimes with black holes that satisfy cosmic censorship and some energy condition. The surface gravity κ is analogous to temperature: it is constant on the horizon, cannot be reduced to zero by a finite process, and multiplies the analogue of entropy in the First Law. The area of the event horizon plays entropy's role in the First Law and is time-asymmetric. And generally, as evinced by the First Law, black hole mechanics can be described by the evolution of just a handful of macroscopic quantities, just like phenomenological thermodynamics.[3]

1.2 Generalized entropy, quantum effects, and holography

There is certainly some degree of formal similarity between thermodynamics and black hole mechanics. Two reasons are taken to push this beyond mathematical coincidence. First, attempts to salvage the thermodynamic Second Law when

[2] This is established by the no-hair theorems. Note that this is no longer true when one introduces Yang–Mills fields other than the electromagnetic field. See Chruściel et al. (2012).

[3] For fuller accounts of BHT see Wald (2001), Wald (1994, ch. 6), and Heusler (1996, ch. 7).

dealing with black hole phenomena led to the conjecture that thermodynamic entropy can be converted into black hole area, giving a Generalized Second Law. Second, quantum effects near the black hole horizon suggest that the black hole radiates like a blackbody with temperature proportional to the surface gravity. These results suggest that the surface gravity and area of the black hole really can be considered the physical temperature and entropy. And this motivates a search for a statistical mechanical underpinning for black hole entropy.

Bardeen et al. famously questioned the analogy. They argued that the temperature of a black hole must be absolute zero, because it is a perfect absorber. If any thermodynamic system is brought near a black hole, it will radiate some of its heat into the black hole—yet the black hole cannot transfer any heat to the thermodynamic system. Since heat will flow from any system to a black hole and never the reverse, the effective temperature of a black hole is absolute zero, independent of its surface gravity, which may be any value. Bardeen et al. further argue that black holes transcend the Second Law of thermodynamics. Because a black hole is effectively at absolute zero, dumping in a box of entropy will not change the black hole. They do not say what they mean by "transcend", but Bekenstein (1972) explains it in epistemic terms. If an external observer tosses a box of entropy into a black hole without changing the black hole's parameters, then there will be no difference in external observables before and after the box is sent through the event horizon. This means that the external observer "cannot exclude the possibility that the total entropy of the universe may have decreased" (1972, 737). Additionally, if one only accounts for the exterior of the black hole, then it is possible to convert heat into work with perfect efficiency via a "Geroch process."[4] Lower a box of mass to the event horizon, extracting work from it. When it reaches the black hole, allow some mass to radiate away. From a distance, the gravitational potential energy of the box vanishes at the event horizon, so this radiation makes no difference to the energy of the box. Drawing the box back up will take some work, but less than the work extracted while lowering it to the event horizon. At the completion of this cycle, the only difference outside the black hole is that mass has been converted to work at perfect efficiency. Since the interior of the black hole is off the books, this is apparently a perpetual motion machine of the second kind, a violation of the Second Law.

These considerations led Bekenstein to formulate the Generalized Second Law, which says that the total generalized entropy of the universe never decreases. Generalized entropy includes both the usual thermodynamic entropy in the exterior region and black hole entropy, which is proportional to the area of the black hole. If the Generalized Second Law holds, then whatever entropy is "lost" to the interior of the black hole will be compensated by an increase in the area of the black hole. This also prevents the Geroch process from violating the

[4]This argument is due to Geroch, who gave it during the question-and-answer session of a colloquium he gave at Princeton.

Second Law: when one lowers the box toward the black hole, the gravitational interactions between the box and the black hole will cause the area of the horizon to increase.

Establishing the Generalized Second Law has proven difficult.[5] In fact, it is easily violated. Consider some system with very small mass and very large entropy. Dropping it into the black hole will only increase its area by a very small amount, by the First Law of BHT, but the thermodynamic entropy in the exterior region will greatly decrease. To block this counterexample (and for no other reason), Bekenstein (1981) proposed that the entropy of any system must be bounded by the surface area of the system. Such a bound would prevent counterexamples along the lines of the Geroch process, but it does not appear to be sufficient for establishing the Generalized Second Law.

A further thermodynamic analogy lurks in the quantum realm. In 1974, Hawking argued that a black hole appears to radiate like a blackbody with temperature $\kappa/2\pi$. Intuitively, the picture is this: near the event horizon, particle-antiparticle pairs are created by fluctuations of the quantum field. One of these pairs passes through the event horizon and is ignored, the other radiates away to infinity. The relative frequency of particle creation depends on the energy of the created particles, and this depends on the background geometry. The particles that make it to infinity replicate radiation from a blackbody at the Hawking temperature.[6]

Bekenstein's arguments and Hawking radiation are taken to show that the analogy between the laws of thermodynamics and the laws of BHT is more than formal. If this is true, it suggests a natural research program. In the late nineteenth century, phenomenological thermodynamics was theoretically reduced to the statistical mechanics of some microphysical degrees of freedom. If the laws of BHT are the laws of phenomenological thermodynamics, and the laws of phenomenological thermodynamics really are reducible to the statistical mechanics of microphysical degrees of freedom, then so are the laws of BHT. In 1996, Strominger and Vafa calculated the entropy of a limited class of black holes by counting string-theoretic microstates, finding agreement with the entropy of BHT. This calculation is a feather in string theory's cap, and other quantum theories of gravity are tasked with replicating it in their own frameworks. So the analogy of BHT plays a crucial role in judging the success of physical theories that look to reconcile quantum mechanics and gravitation.

Combining Bekenstein's and Hawking's proposals leads to another puzzle. According to Bekenstein's entropy bound and its descendants, both kinds of generalized entropy are bounded by surface area—black hole entropy because it is proportional to the area, and thermodynamic entropy by fiat. However, the entropy of a quantum-mechanical system is proportional to its volume. If BHT is an expression of the quantum character of black holes, as Hawking's

[5] For a review of attempts and their shortcomings, see Wall (2009).

[6] The problematic particle-talk of this heuristic story can be eliminated; see Wald (1994, §7.1).

argument suggests, then the volume of the black hole as a quantum system must be its area. These can't both be right if there are three dimensions of space, so one apparently has to go.

The popular response is instead to hold fast to the analogy, concluding that BHT teaches us that the world has fewer dimensions than we thought; it is a kind of "hologram" (Bousso, 2002). The degrees of freedom in some region live on the surface bounding it, and the physics in the region is "projected" from the surface, much like a hologram is a three-dimensional projection of two-dimensional data. So, starting from BHT one infers that space and time are not fundamental. It is no exaggeration to say that the analogy underlying BHT is regarded as the most important clue we have to a theory of quantum gravity, and the holographic principle as one of the best cases to build on this clue.

2 Analogies and Thermodynamics

Analogies can be strong or weak, and arguments based on them correspondingly good or fallacious. Any two objects are similar and dissimilar in an infinity of ways. What matters is relevant similarity, and relevance is in part determined by background information and context. The reliance on relevance entails that the difference between good and bad analogies is inherently a bit fuzzy. But in science we typically have a pretty good sense of similarity because we enjoy a lot of background information. The wave equation is ubiquitous in physics, but we don't try to identify everything that obeys it: quantum fields aren't little masses on Hookean springs, even if it's sometimes useful to think like this.

Thermodynamic analogies can be found throughout science. Some are strong, some weak. Interesting analogies exist, for instance, between thermodynamics, on the one hand, and electricity, acoustics, and mechanics, on the other.[7] Take electricity. Let current play the role of heat flow and voltage play the role of temperature difference. Then definitions of charge and power straightforwardly have counterparts in heat and thermodynamic power, respectively, and a host of electrical laws, e.g., Kirchhoff's Current and Loop Laws, hence share close formal features with famous thermodynamic identities. Similar claims can be made for acoustics and mechanics. Here, given background knowledge, the superficiality of the similarity, and disanalogies when one peers closely, no one understands all of these identities holding due to some grand thermodynamic principle at work.

Outside physics one finds thermodynamic analogies in economics, finance, ecology, and more. Thermoeconomics is a good cautionary example. The economist Fisher (1892), whose advisor was none other than J. Willard Gibbs, sought a kind of parallelism between economics and thermodynamics. On its face, the similar role of equilibrium in the two theories invites such speculation,

[7]For a discussion of many such analogies, see Karnopp et al., 1990.

and the early twentieth century saw many attempts to develop economics as a thermodynamic theory (as opposed to economics merely taking into account thermodynamic processes, as it should). While some researchers continue pressing the analogy, probably most economists now agree that the search is no longer worth pursuing. Any isomorphism between economics and physics is merely a coincidence or due to using similar mathematical tools. Here we have a set of functional relationships, many of which parallel the thermodynamic laws, but where the disanalogies eventually stood out, the lack of a common mechanism was plain, and the absence of fruit borne became an obstacle.

BHT could turn out like thermoeconomics. Currently there is no consensus on a common "statistical mechanical" underpinning of the Generalized Second Law, so there may be no common mechanism at all. That is a major concern. We don't wish to focus on this possible disanalogy, important though it may be. Rather, we want to direct attention to the prior question of whether the analogy that motivates the search for a statistical mechanical explanation is really all that tight.

3 Pale Shadows of Thermodynamic Laws: The Zeroth Law and Equilibrium

As we saw, under mild assumptions, one can prove that the surface gravity κ will be constant on the event horizon of a black hole. BHT takes $\kappa/2\pi$'s constancy over the event horizon to be the counterpart of the thermodynamic Zeroth Law.

No thermodynamicist, however, would regard mere constancy of a quantity to express the Zeroth Law. The Zeroth Law is a foundational piece of the theory's edifice. It does not primarily claim that temperature in equilibrium is constant. The Law is commonly expressed as stating that thermal equilibrium is transitive: that if a system A is in equilibrium with a system B, and B with C, then A is in thermal equilibrium with C. But even this proposition undersells the Law. If we peer under the hood we see that the claim of transitivity presupposes an awful lot. In particular, it assumes at least

(a) that there is a such a state as thermal equilibrium,

(b) that systems will spontaneously approach this state—sometimes called the Minus First Law (Brown and Uffink, 2001),

(c) that the 'equilibrium with' relation exists, and

(d) that this relation is transitive.

The Zeroth Law in fact sets up the equilibrium state space for the theory, asserts a general tendency toward these states, and then imposes a structural relationship amongst these states. With all of these assumptions in place one can then prove that there must exist a state function having the same value for all systems that are in thermodynamic equilibrium with each other. This state function then allows the creation of an empirical temperature scale. From the intensive nature of temperature, it is also possible to prove that a system in equilibrium has parts that are mutually in equilibrium, and hence enjoy the same empirical temperature.

The claim that we find the Zeroth Law in κ's constancy therefore seems a bold one. At best it mistakes a consequence of the Zeroth Law for the law itself. We have a counterpart of the Zeroth Law only if we replace the real thing with an impoverished, hollowed out version of the real law. The Zeroth Law has analogies everywhere if this version counts.

The reader may respond by trying to find BHT counterparts for the real Zeroth Law. This is an important project if the analogy is to be a strong one. After all, as indicated above, the Zeroth Law is really foundational in thermodynamics, not something to be confused with mere constancy of a function. The state space and fundamental concepts all hang on it.

A natural place to begin would be with thermal equilibrium. In BHT one commonly finds the claim that stationarity plays the role of thermal equilibrium. Stationary spacetimes have the nice property that the vacuum region outside a stationary black hole has a Kerr–Newman geometry. Thus, reminiscent of thermodynamic equilibrium, the vacuum region is given by a few numbers: the mass, angular momentum, and electromagnetic charges. Everything else of interest is swallowed up by the black hole. Furthermore, although there is no proof here, one often assumes that after a body has collapsed, eventually it will "settle down" to a stationary state.

Perhaps we have found analogues of (a) and (b)? No doubt this judgement will be based on holistic considerations, ones considering whether stationarity plays anything like the same role equilibrium does in thermodynamics. In that science, for instance, isolated systems in equilibrium minimize their internal energy. Is the BHT counterpart of internal energy, the mass, minimized when isolated systems are stationary? It's not clear what that even means. So there are many questions that would need to be answered before we agree that counterparts of (a) and (b) have been achieved.

Even if we ignore all that, we submit that a significant and fundamental disanalogy will persist: there is no 'equilibrium with' relation in BHT. Without this relation, one will never get anything like the Zeroth Law (it will leave out (c) and (d)), nor will one recover much of thermodynamics (because most of it rests on assumptions made in the Zeroth Law). The point is very simple. One iron bar can be in equilibrium with another iron bar. Can one black hole be in equilibrium with another black hole? What is the counterpart of 'equilibrium with'? Suppose equilibrium is identified with stationarity. Then the counterpart is one black hole being in the 'stationary with' relation with another. Does that make sense? Certainly not if the thermodynamic system is the whole spacetime. Nor is it clear what this means on any "local" understanding of BHT. Maybe we can evade this worry by relying on the intensivity of temperature? In thermodynamics we can take one iron bar in equilibrium and mentally divide it into two subsystems, each with the same temperature. Because temperature doesn't depend upon the "amount of stuff" (i.e., is intensive), we can do this kind of division without ascribing different temperature values to the subsystems. Here too we get nonsense. The black hole counterpart is dividing the black hole in two and regarding each as having the same surface gravity. If we could overcome this, still we face the fact that the surface gravity is not intensive (discussed below).

Although we've focused on the Zeroth Law, that is only because it is foundational and first. Worries could be raised about all the laws of BHT. Here, quickly, are a few other disanalogies:

- The First Law of thermodynamics asserts that a function of state, internal energy, exists. Internal energy in thermodynamics is distinct from total energy, but in BHT it is identified with total energy, and ultimately, total mass. And the lack of a local notion of energy in general relativity can only compound trouble.

- Suppose we can make sense of two black holes coalescing into one. Then by the area theorem the resulting black hole will have an area larger than the sum of the original two. The black hole entropy thereby increases. But that is so even if the two black holes were originally of the same temperature, i.e., surface gravity, contrary to thermodynamics.

- The volume is important in many thermodynamic relationships, e.g., the ideal gas law. It is of course definitionally tied for a cube with sides of length l to be $l \times$ area. Area is proportional to black hole entropy. This relationship forces awkward questions. Can we substitute black hole entropy wherever we find $l \times$ area (i.e., volume) in thermodynamic laws? That will make a mess of any thermodynamic relationship that includes volume as a variable, such as the ideal gas law. Or do we claim that volume is not the counterpart of volume in BHT? Neither option is attractive.

- Peter Landsberg once called the fact that thermodynamic variables sort themselves into such a neat set of relationships between the intensive and extensive variables the "fourth law" of thermodynamics (1978) . Loosely put, extensive variables are variables that scale with the amount of stuff in a system or its size, whereas intensive variables are independent of the amount of stuff or size of a system. Entropy and internal energy are famously extensive. Double the system and you thereby double the entropy and energy. Temperature, by contrast, is intensive. Double the system and it will still be the same temperature. It's well-known that the scaling of typical thermodynamics systems doesn't even approximately hold in

> BHT and that this has important effects: e.g., the impossibility of deriving the Gibbs-Duhem relation. If we restrict attention to Schwarzschild black holes, now when we double the system, we don't double the entropy, but we do halve the temperature! No less than the most paradigmatic intensive variable, the temperature $\kappa/2\pi$, varies inversely with system size.[8] As the "amount of stuff" changes, nothing is holding steady and all the relationships amongst variables are shifting.

By getting picky or a little deeper into thermodynamics we can find more disanalogies. But there is no need to get picky or deep, as the lack of an 'equilibrium with' relation will always haunt the analogy, insofar as the Zeroth Law sets up the state space and basic concepts of the theory.

4 Entropy of What?

The previous point focused on whether the analogy is all that strong. Here we focus on a fundamental tension in the analogy itself (see also Corichi and Sudarsky, 2002). If the laws of BHT are the laws of thermodynamics, then they must describe the behavior of thermodynamic systems. These systems are the bearers of black hole entropy, internal energy, and so on. They must interact with normal thermodynamic systems to obey the Generalized Second Law. But identifying these systems is difficult. On the one hand, BHT defines black holes as regions circumscribed by event horizons. However, event horizons are global beasts and can't be hunted by finite physicists. They are also inadequate to the task of fully working out the BHT analogy. On the other hand, there are no local horizons that can play all of the roles required. Resolving this tension requires giving up some part of BHT. For example, the reasoning behind Bekenstein's Generalized Second Law applies only to global event horizons, but Hawking radiation is a local phenomenon. Sticking to the global definition means rampant nonlocality, and moving to a local horizon means keeping only a fragment of the already distorted thermodynamic analogy.

On the most intuitive picture, the thermodynamic system of BHT is, or is bounded by, an event horizon. Event horizons are intrinsically global. In §1.1, we made the move to local considerations like the surface gravity and Killing horizons by restricting attention to a special class of symmetric spacetimes. For general black hole spacetimes, identifying a black hole region requires knowing what happens infinitely far in space and time:

> The location of the event horizon, or even its existence, is known only after the universe has ended, or, depending on one's religious beliefs, to the gods looking down on space-time as a vast Penrose

[8] Classical self-gravitation introduces many subtleties regarding scaling (see Callender, 2011); even so, the scaling here is not analogous.

> diagram. It cannot be known to mere mortals in the here and now (Hayward, 2002, 569).

In fact, our ignorance of event horizons is as complete as can be. There is no relationship between the location of an event horizon and the curvature at that location. In particular, event horizons can exist in flat space. Consider a large collapsing shell of matter. At any point inside the shell space is flat, and remains so until the shell arrives. Send out a photon from the center of the shell. If you do so early enough, that photon will escape to infinity. But if the shell has sufficiently collapsed, the photon will be trapped. The trajectories of the earliest emitted trapped photons form an event horizon inside the shell, and the horizon expands at the speed of light until it is the size of a black hole with the same mass as the shell. And it does all of this in completely flat spacetime. The event horizon forms and expands even though no matter passes through it. You could be passing through an event horizon right now and be none the wiser.

Event horizons can "react" to matter that will fall through them at some point in the future, which makes them seem clairvoyant. But they are also importantly ignorant. Anything behind an event horizon is invisible at infinity. For example, suppose that a star falls through the event horizon of a large black hole and then collapses, as in Fig. 3. Intuitively, a black hole should form, just as

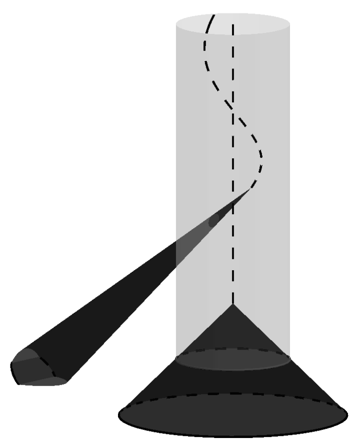
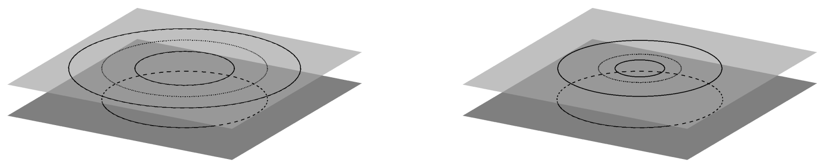


Figure 3: Collapsing inside a black hole

it would outside of the large black hole. Not so. Event horizons are the *outermost* surfaces of the region of no escape. So event horizons cannot form within event horizons. As a consequence, black hole thermodynamics does not apply to collapsed objects behind an event horizon.

This nonlocality has incubated a cottage industry in locally-defined aftermarket replacements for event horizons.[9] None is clearly adequate. Many are

[9]For further details on locally-defined horizons, see the review by Nielsen (2009).



(a) Flat space (b) Curved space

Figure 4: A flashing loop at two instants

foliation-dependent, a major drawback. And it's not clear that the foliation-independent notions are really local in a relevant way: foliation-independent horizons can exhibit a problematic clairvoyance similar to event horizons, for example (Bengtsson and Senovilla, 2011). Without a better handle on what counts as an objectionably non-local property it's hard to see how to corral them into any sort of order. And this makes it all the more difficult to get on to the real work, which is determining which things are supposed to be thermodynamic. The local horizons are all distinct and are furthermore distinct from event horizons. Indeed, one can choose a slicing of Schwarzschild spacetime such that some local horizons don't appear at all, despite the presence of an event horizon (Wald and Iyer, 1991)! Are the laws of thermodynamics meant to govern the behavior of the event horizon, or one (or more) of the local horizons?

As an example of the difficulties, we consider one of the most successful notions of local horizons: trapping horizons, defined by the following local no-escape condition. Consider any spacelike two-surface in spacetime—e.g., a sheet of paper—and draw a flashing loop on it, as on the lower slices of Fig. 4. At every point of the loop, there are two orthogonal directions that the light it emits might travel across the surface: directly away from the loop, or toward the inside. Shortly after a flash, the ingoing light will form one copy of the loop, and the outgoing light will form another. In flat spacetime, the ingoing loop will be smaller than the original, and the outgoing loop with be larger. But if there is a strong gravitational attractor inside the loop, then both loops might be contracted into smaller copies of the original loop. If this happens, light near the loop will be trapped, unable to escape to observers. For this reason, the surface traced out by the loop as it evolves through time is called a trapping horizon.

Trapping horizons can replace event horizons in some applications. Assuming the null energy condition, the area of a trapping horizon must increase through time, giving a version of the Second Law for trapping horizons. Though there is no univocal definition of surface gravity or equilibrium for a trapping horizon, analogues of the Zeroth and First Laws of BHT also hold. The Third

Law, ever-neglected, has no proof for locally-defined horizons, and has likely counterexamples. However, the symmetric black hole spacetimes above can be foliated so as to make the event horizon a trapping horizon. So the laws of BHT apply to some trapping horizons, and in cases where an event horizon and trapping horizon coincide, one must decide which properties to attribute to the event horizon and which to the trapping horizon. Is the entropy of the BHT system given by the area of the trapping horizon, or of the event horizon? This problem is most pressing in non-stationary cases, where the trapping horizon will not coincide with the event horizon.

One of the most important choices is whether Hawking radiation ought to be associated with locally-defined horizons or the event horizon. Recall that Hawking radiation is supposed to substantiate the identification of surface gravity with the temperature of the black hole. In the usual narrative, the analogy of BHT is merely formal in the classical regime—as a perfect absorber, the black hole's temperature is absolute zero. Hawking radiation gives a radiation mechanism, making the surface gravity physically analogous to temperature as well as mathematically. So the surface gravity ought to follow the Hawking radiation, and by the First Law of BHT so should the other BHT quantities.

It seems clear that Hawking radiation is associated with locally-defined horizons, not event horizons. As Visser (2003) shows, Hawking radiation is essentially a kinematical effect: any theory of gravitation that uses a Lorentzian metric will have Hawking radiation associated with horizons in it. By contrast, the Generalized Second Law must rely on dynamical principles, like the cosmic censorship conjecture and the Einstein field equation. Furthermore, the existence of an event horizon is inessential to the effect; all that is required is a local horizon. When Hawking radiation is associated with an event horizon, this is because it coincides with a local horizon. If Hawking radiation is a sign of a cosmological surface's temperature, then cosmological temperature may exist even when no event horizon does. This strongly suggests that temperature is borne by the locally defined surface, not the event horizon with which it may coincide.

Despite all this, trapping surfaces have their drawbacks. They are foliation-dependent, so whether they exist or not depends on how one carves up spacetime. Foliation dependence is the cardinal relativistic sin. Some of our deepest intuitions about space and time—absolute simultaneity, perfect rigidity, and more—have been committed to the flames on this charge. Foliation independence is tantamount to reality, in a relativistic world. The laws of BHT are foliation-independent, so it would be strange, to say the least, if the physics they describe is bound to a choice of spacetime slicing. Most dramatically, black hole spacetimes can be foliated so that they contain no trapping horizons at any time, despite the existence of an event horizon. In this situation, the laws of BHT hold for the event horizon, but there are no trapping horizons for them to describe. So there seems no choice but to attribute the quantities of BHT to the event horizon, on whose existence all foliations agree. But this brings us right back to the problems that opened this section.

5 Out of Sight, Out of World: Entropy and Epistemicism

Despite the many problems with the analogy, what motivates many is the thought that when objects fall into a black hole compensation must be paid for lost entropy. The reasoning evokes nineteenth century experiments by Joule and others that found the heat equivalents of mechanical energy, electrical energy, and so on. The First Law, and by extension the Second Law, was widened with each new kind of energy that might disappear. BHT then seems the next step in this grand tradition.[10]

Here is the reasoning Bekenstein uses in his classic article:

> Suppose that a body carrying entropy S goes down a black hole. ... The S is the uncertainty in one's knowledge of the internal configuration of the body. So long as the body was still outside the black hole, one had the option of removing this uncertainty by carrying out measurements and obtaining information up to the amount S. But once the body has fallen in, this option is lost; the information about the internal configuration of the body becomes truly inaccessible. We thus expect the black hole entropy, as the measure of the inaccessible information, to increase by an amount S (1973, 2339).

The worry, in other words, is that when matter falls into a black hole the total entropy decreases, violating the standard Second Law. It is then noticed that in these processes the black hole area goes up. Assuming that entropy is additive then leads to the Generalized Second Law, the statement that the gain in black hole area will compensate the entropy lost when matter falls into black holes. Arguments along these lines are repeated throughout the literature (e.g., Wald, 1994, 417).

The key assumption throughout BHT is that the entropy of a body that has fallen into a black hole is lost. Why accept that? After all, steam engines, boxes of gas, and so on all have definite thermodynamic efficiencies regardless of whether or not anyone is looking, and it's certainly possible that such systems cross event horizons without any strong tidal effects much affecting them. Thinking of the information loss paradox, one might gesture toward the worry that information is lost if the black hole eventually evaporates away. However, the identification of entropy with black hole area preceded the threat of evaporation. One may also worry about the matter eventually falling into a singularity, although almost

[10]The first chip in this intuition comes from noticing that this is not what's being done in BHT. If the analogy with Joule held, then we would expect to incorporate gravitational energy into the standard thermodynamic story. BHT doesn't do anything like this, though see Curiel (2014) for an attempt in this direction.

everyone expects unknown quantum gravitational effects to be relevant in that regime.

No, the reason this premise is universally adopted is no mystery. BHT seems committed to an information-theoretic understanding of the thermodynamic entropy. Evidence for this claim can be found throughout the field. The original Bekenstein paper (1973) explicitly embraces an information-theoretic conception of entropy, both thermodynamic and black hole. Using Shannon entropy and Brillouin's identification of information with negative entropy, Bekenstein writes that "the entropy of a thermodynamic system which is not in equilibrium increases because information about the internal configuration of the system is being lost" (6–7).[11] And black hole entropy is understood "as the measure of the inaccessibility of information (to an exterior observer) as to which internal configuration of the black hole is actually realized" (6).

This understanding is ubiquitous, from research articles to distinguished textbooks:

> Indeed, given that the entropy represents your lack of knowledge about a system, once matter goes into a black hole one can say that our knowledge about it completely vanishes. (Sethna, 2006, 94)

In short, as part of the growing infiltration of physics by information theory, it is held that the entropy vanishes when it passes behind the event horizon because we can't gain access to it. The system itself doesn't vanish; indeed, it had better not because its mass is needed to drive area increase. But for the ordinary entropy, when it crosses the event horizon it's "out of sight, out of world," or at least, out of physics.

To our knowledge, no attention has been drawn to this assumption in the black holes literature. Those familiar with the foundations of statistical mechanics, however, will recognize it as a massively controversial assumption, one worthy of scrutiny.

If we grossly simplify a long and subtle history, we find in the foundations of statistical mechanics two very different understandings of entropy, one "objective," the other "epistemic." Everyone agrees that gases expand throughout their available volumes because that is the most likely behavior for a mechanical system in that state. But what entropy should we use, and what is the physical justification of the claim that that entropy most likely rises? Debate ensues.[12]

The objective understanding is found in the pioneering work of Boltzmann and Gibbs, and it is now carried on by modern physicists such as Sinai, Khinchin, Lanford, Ruelle, and Lebowitz. Using one or more of the entropies devised by Boltzmann or Gibbs, ultimately the justification for thermodynamic behavior lies in the detailed dynamics of the microphysics. Boltzmann hoped that the

[11] Note that if one switches to trapping horizons instead of event horizons, this is no longer the case. Light may cross a trapping horizon but still escape to infinity.

[12] For a more detailed discussion of this history and debate, see Sklar (1993).

dynamics is ergodic, Gibbs that it is mixing, and others that it is quasi-ergodic. These are strong conditions on the dynamics. Some modern physicists such as Lanford and Lebowitz hope for something weaker, namely, that "typical" (in a precise sense) initial conditions follow "thermodynamic" trajectories. Note that the probabilities invoked are interpreted objectively, as a reflection of the way ensembles of real systems actually behave. The differences amongst all these programs are hotly contested and describing them could fill volumes. However, they all have in common an objective conception of the entropy and an explanation of thermodynamic phenomena that hangs on the detailed microphysical dynamics. Degrees of belief and information have nothing essentially to do with why physical systems spontaneously head toward thermal equilibrium on this view.

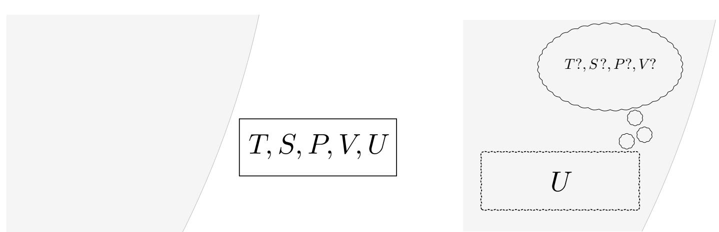
The epistemic understanding of entropy and thermodynamic behavior, by contrast, offers a quite different explanation. Using the Shannon entropy, one understands entropy in terms of how much information is conveyed in a signal. The powerful perspective and methods of information theory were then brought to bear on statistical mechanics primarily by E. T. Jaynes in the 1960s. Statistical mechanics, on Jaynes' view, is not a theory about the world, per se, but a theory of inference. The Shannon entropy is a function whose value is maximized with maximum uncertainty and which vanishes when uncertainty disappears. The probability distributions in the formula for entropy are interpreted as rational degrees of belief; in particular, how uncertain we are of the microstate of the system, given its macrostate. So-called subjectivists hold that these degrees of belief are your own subjective ones. Objective Bayesians, by contrast, hold that there is a uniquely rational set of degrees of belief for one to have given the evidence. Either way, so understood entropy becomes a feature of one's epistemic state and its increase is a matter of uncertainty increasing.

We have two broad schools of thought, one objectivist and one epistemic. According to the former, systems behave thermodynamically because the microdynamics makes it objectively likely. According to the latter, statistical mechanics becomes a branch of epistemology, not physics; it is a theory of inference. What leads to great confusion, we submit, is that the Shannon entropy, used by the latter approach, and the Gibbs fine-grained entropy, used by the former approach, are formally identical. In fact, outside of the foundational literature, one commonly finds expressions dubbed the "Boltzmann–Gibbs–Shannon entropy" despite the three names each referring to distinct objects. While physicists would not be fooled by a pun on the English word "entropy," it is very easy to conflate the Shannon and Gibbs entropies, interpreted as we have here. But the two entropies can have different values even though they are formally the same. The reason is simple: the probability distribution of one entropy refers to subjective degrees of belief whereas in the other it may refer to objective frequencies or propensities in the external world. If a Laplacian demon told you the exact microstate of a gas, that would affect the value of the Shannon entropy (driving it to zero) whereas it wouldn't affect the value of the Gibbs entropy, as understood here.

Although we won't resolve this long-standing debate here, many believe—and we concur—that it is a big mistake to identify the thermodynamic entropy with any information-theoretic entropy like Shannon's. Given the formal identity between the Gibbs and Shannon entropies, it's natural to slip between the two. Arguably, however, only the "objective" entropies are suitable for identification with the thermodynamic entropy. The thermodynamic entropy seems not to have anything to do with people or beliefs. Perhaps the general thought is best put by David Albert's incredulity:

> Can anybody seriously think that it is somehow *necessary* ... that the particles that make up the material world must arrange themselves in accord with *what we know*, with *what we happen to have looked into?* Can anybody seriously think that our merely being *ignorant* of the exact microconditions of thermodynamic systems plays some part in *bringing it about*, in *making it the case*, that (say) *milk dissolves in coffee?* How could that *be?* (2000, 64)

The thermodynamic entropy is directly connected to efficiency and the amount of work a system can do. These are perfectly objective facts about boxes of gas, steam engines, and the like. A steam engine's efficiency doesn't care about whether anyone is looking or not, our uncertainty, or our beliefs. Since Shannon entropy does, it therefore cannot be identified with the thermodynamic entropy.



(a) Before (b) After

Figure 5: A thermodynamic system, before and after entering the black hole

If this is correct, then there is no reason to believe that a body slipping past an event horizon would lose its entropy. As pointed out earlier, the event horizon itself is not a local observable. Minus whatever effects tidal forces cause (which may be minimal or none), a steam engine falling behind an event horizon is just as efficient as before and still has the same entropy—after all, you could follow it in and check. It's a mistake to believe that the entropy of the body changes in any way. In particular, there is then no reason to believe that entropy decreases, and therefore, no compensation is necessary.

Finally, note that a potentially objectionable philosophy of science sometimes creeps in with the epistemic approach. Why does mass inside a black hole get treated differently than entropy? One answer, just discussed, is that entropy but not mass is epistemic in nature. Another is more philosophical. Bekenstein writes:

> Entropy lost into black holes cannot be kept track of, and so one should not, in ordinary circumstances, discuss entropy inside black holes. Thus the ordinary Second Law must be given a generalised form. (2004, 33)

This claim, which is hardly unique to Bekenstein, appeals to a kind of philosophical operationalism. The potential double-standard in the treatment of mass and entropy is avoided by appeal to what we can measure: we can measure the mass of the body that has fallen into a black hole—one can measure the gravitational acceleration of objects around the black hole—but one cannot measure the entropy. The idea, familiar to physicists from Mach, is that physics is only about what can be observed or measured. This position is a hard one to maintain philosophically. By the argument that one should not discuss what one cannot keep track of it follows that one of the authors (CC) should not discuss his teenage children! More seriously, it is notoriously difficult to precisely specify what observation or measurement come to, and then it is equally hard to justify using our contingent measuring abilities to determine the limits of our knowledge. Quarks cannot be observed singly, but arguably we can still discuss and even know about them. For these reasons and others the view known as operationalism, which says that science is restricted to what we can measure, is "nowadays commonly regarded as an extreme and outmoded position" (Chang, 2009).

No matter one's philosophical predilections, the operationalism necessary in the present case is a particularly odd and stringent version. The reason is that we could observe the entropy of steam engines and the like that fall behind event horizons. Just jump in with them! The fact that we would have few volunteers for such a measurement shouldn't matter epistemologically. Why prefer some observers over others with no difference between the two besides what side of the unobservable event horizon they are on?

There may be reasons to endorse the compensation argument that don't involve mixing up the objective with the epistemic. But if there are—and they don't already presuppose that BHT is more than a formal analogy—we haven't yet encountered them. We leave this as a challenge to the reader.

6 Conclusion

BHT may well be a useful analogy and a clue to quantum gravity. We are not saying it isn't. Nor are we saying that all the hundreds of articles on this topic have been a waste of time. Clearly, if nothing else, it has produced some very interesting and significant results in classical general relativity and inspired novel speculations. Nor are we saying that an information-theoretic understanding of the Generalized Second Law isn't true. All known proofs have serious problems (see Wall, 2009), but this generalization about information outside event horizons might still be true.

What we are arguing is that the analogy may not be more than formal. First, we've shown that the analogy is not that tight. Similarity is in the eye of the beholder. Yet one only sees similarity between the two theories if one very selectively focuses on some parts while ignoring others. Second, we've pointed to a major tension in the analogy itself. Is the object globally defined system based on the event horizon, or a quasi-local entity based on trapped surfaces? Either option has costs. Third, we've highlighted how a major motivation linking black hole entropy and ordinary entropy is based on massively contentious reasoning in the foundations of statistical mechanics. If we now add the fact that BHT is only defined for a small sector of the space of solutions of general relativity and the lack of a statistical mechanical grounding of the theory, we believe that there is no urgency in explaining why black holes behave thermodynamically. To the extent that there is similarity, given the heavy dose of geometry present, one might suggest an alternative narrative where much of differential geometry begins to look "thermodynamic" (see Baez, 2012). We leave such speculations to another day.

Physical discovery is a game of bets. It seems to us that most of the physics community is "all in" on the analogy being more than formal. Physical hypotheses are proposed almost daily whose basis lay in no theoretical or empirical advance but merely the extension of this analogy. But the analogy is not as strong as is commonly supposed. BHT may go the way of thermoeconomics. That is why we recommend hedging one's bets.[13]

References

Albert, D. Z. (2000). *Time and Chance*. Harvard University Press.

Baez, J. (2012). Classical mechanics versus thermodynamics. `https://johncarlosbaez.wordpress.com/2012/01/23/classical-mechanics-versus-thermodynamics-part-2/`.

Bardeen, J., Carter, B., and Hawking, S. (1973). The four laws of black hole mechanics. *Communications in Mathematical Physics*, 31:161–170.

Bekenstein, J. D. (1972). Black holes and the second law. *Lettere Al Nuovo Cimento*, 4(15):737–740.

[13] Many thanks to Erik Curiel, John Norton, Chip Sebens, Daniel Sheehan, and Bob Wald for comments that helped improve this paper.

Bekenstein, J. D. (1973). Black holes and entropy. *Physical Review D*, 7:2333–2346.

Bekenstein, J. D. (1981). A universal upper bound on the entropy to energy ratio for bounded systems. *Physical Review D*, 27:2262.

Bekenstein, J. D. (2004). Black holes and information theory. *Contemporary Physics*, 45(1):31–43.

Bengtsson, I. and Senovilla, J. M. M. (2011). Region with trapped surfaces in spherical symmetry, its core, and their boundaries. *Phys. Rev. D*, 83:044012.

Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics*, 74:825–874.

Brown, H. R. and Uffink, J. (2001). The origins of time-asymmetry in thermodynamics: The minus first law. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(4):525–538.

Callender, C. (2011). Hot and heavy matters in the foundations of statistical mechanics. *Foundations of Physics*, 41(6):960–981.

Chang, H. (2009). Operationalism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2009 edition.

Chruściel, P. T., Costa, J. L., and Heusler, M. (2012). Stationary black holes: Uniqueness and beyond. *Living Reviews in Relativity*, 15(7).

Corichi, A. and Sudarsky, D. (2002). When is $S = A/4$? *Modern Physics Letters A*, 17:1431–1443.

Curiel, E. (2014). Are classical black holes hot or cold? PhilSci Archive: 11675.

Fisher, I. (1892). Mathematical investigations in the theory of value and prices. *Transactions of the Connecticut Academy*, 9. Doctoral Thesis.

Hawking, S. W. (1971). Gravitational radiation from colliding black holes. *Physical Review Letters*, 26:152–166.

Hawking, S. W. (1975). Particle creation by black holes. *Communications in Mathematical Physics*, 43(3):199–220.

Hayward, S. (2002). Black holes: New horizons. In Jantzen, R. T., Ruffini, R., and Gurzadyan, V. G., editors, *Proceedings of the Ninth Marcel Grossman Meeting on General Relativity*, pages 568–580. World Scientific.

Heusler, M. (1996). *Black Hole Uniqueness Theorems*. Cambridge University Press.

Karnopp, D. C., Margolis, D. L., and Rosenberg, R. C. (1990). *System Dynamics: A Unified Approach*. John Wiley & Sons, Inc.

Kay, B. S. and Wald, R. M. (1991). Theorems on the uniqueness and thermal properties of stationary, nonsingular, quasifree states on spacetimes with a bifurcate Killing horizon. *Physics Reports*, 207(2):49–136.

Landsberg, P. T. (1978). *Thermodynamics and Statistical Mechanics*. Oxford University Press.

Maxwell, J. C. (1855). On Faraday's lines of force. *Transactions of the Cambridge Philosophical Society*, 10(1):155–229.

Nielsen, A. (2009). Black holes and black hole thermodynamics without event horizons. *General Relativity and Gravitation*, 41(7):1539–1584.

Sethna, J. P. (2006). *Statistical Mechanics: Entropy, Order Parameters, and Complexity*. Oxford University Press.

Sklar, L. (1993). *Physics and Chance*. Cambridge University Press.

Strominger, A. and Vafa, C. (1996). Microscopic origin of the Bekenstein–Hawking entropy. *Physics Letters*, B379:99.

Visser, M. (2003). Essential and inessential features of Hawking radiation. *International Journal of Modern Physics D*, 12(04):649–661.

Wald, R. M. (1994). *Quantum field theory in curved spacetime and black hole thermodynamics*. University of Chicago Press.

Wald, R. M. (2001). The thermodynamics of black holes. *Living Reviews in Relativity*, 4(6).

Wald, R. M. and Iyer, V. (1991). Trapped surfaces in the schwarzschild geometry and cosmic censorship. *Phys. Rev. D*, 44:R3719–R3722.

Wall, A. C. (2009). Ten proofs of the generalized second law. *Journal of High Energy Physics*, 2009(06):021.