

# Chapter 14

## Black Holes and Analogy



Alex Mathie

**Abstract** It is generally accepted that science sometimes involves reasoning with analogies. Often, this simply means that analogies between disparate objects of study might be used as heuristics to guide theory development. Contemporary black hole physics, however, deploys analogical reasoning in a way that seems to overreach this traditional heuristic role. In this chapter, I describe two distinct pieces of analogical reasoning that are quite central to the contemporary study of black holes. The first underpins arguments for the existence of astrophysical Hawking radiation, and the second underpins arguments for black holes being ‘genuinely’ thermodynamical in nature. I argue that while these are distinct analogical arguments, they depend on one another in an interesting way: the success of the second analogical argument presupposes the success of the first. This induces a tension for those who wish to take black hole thermodynamics seriously, but who are sceptical of the evidence provided for astrophysical Hawking radiation by the results of analogue gravity. I consider three ways to resolve this tension, and show that each fails.

### 14.1 Introduction

To put it flippantly, the trouble with black holes is that they are black, and that they are holes. This makes direct empirical contact with black hole systems an intrinsically troublesome business, since by definition the (classical) black hole necessarily has no optical signature and the fact that its conventional definition—in terms of an event horizon—makes reference to future null infinity means that it has no well-defined location in spacetime. But the experimental astrophysicist’s loss is the philosopher’s gain: the difficulties that hinder straightforward empirical access to black holes have necessitated the use of less conventional epistemic techniques

---

A. Mathie (✉)  
Munich Center for Mathematical Philosophy, Munich, Germany  
e-mail: [alex.mathie@lmu.de](mailto:alex.mathie@lmu.de)

© The Author(s) 2023  
N. Mills Boyd et al. (eds.), *Philosophy of Astrophysics*, Synthese Library 472,  
[https://doi.org/10.1007/978-3-031-26618-8\\_14](https://doi.org/10.1007/978-3-031-26618-8_14)

by physicists, and it is precisely these difficulties that make the epistemology of black hole physics such a fertile ground for philosophical scrutiny. Some of these techniques are comparatively quotidian: indirect detection of black holes by the observation of their interaction (gravitational or electromagnetic) with ordinary matter, for instance.<sup>1</sup> But others are more philosophically interesting, and perhaps controversial. One example is the use of analogical reasoning in black hole physics.

That there is analogical reasoning in contemporary black hole physics is, by now, an uncontroversial point. Indeed, there is a burgeoning literature on the epistemology of analogue experiments, and analogue gravity (as we shall examine in greater detail below) is a central case study in this literature (Dardashti et al. 2017, 2019; Thébault 2019; Evans and Thébault 2020; Crowther et al. 2021; Field 2022). Similarly, black hole thermodynamics (BHT)—which takes as its departure point the striking analogy between the laws of black hole mechanics and those of ordinary thermodynamics—has been of great interest to physicists since its discovery by Bardeen et al. (1973). Recent years have seen a growing interest in—and dispute over—how we should interpret this unexpected correspondence between two pillars of modern physics by philosophers (Curiel 2014; Wallace 2018, 2019; Wüthrich 2019; Prunkl and Timpson 2019; Dougherty and Callender 2016).

Relatively little attention, however, has been paid to BHT *qua* a piece of analogical reasoning. The literature that does exist focuses almost exclusively on the conceptual difficulties that arise from attempting to interpret various black hole mechanical quantities as genuinely equivalent to their thermodynamical analogues, and the attendant difficulties for the claim that black holes are capable of genuinely thermodynamical behaviour. But it does not explicitly apply the philosophical literature on analogical reasoning *to* the BHT analogy. This leaves two blind spots in the literature on analogical reasoning in black hole physics. First, there is the need for precisely this kind of focused analysis of BHT *qua* a piece of analogical reasoning. Second, although philosophers have scrutinised both analogue gravity and BHT in isolation, scant philosophical work has been done on the connection between them.<sup>2</sup>

And yet such a connection clearly exists: on the standard view, it is Hawking's (1974) prediction of radiative flux from a black hole's event horizon that “removes the blemishes in BHT and transforms it from a suggestive analogy to a full equivalence” (Wallace 2018, p. 60).<sup>3</sup> Because direct empirical detection of the Hawking flux borders on the impossible, however, it is analogue gravity research

<sup>1</sup> See Rees (1998) for a helpful—if dated—survey, and Bambi (2018) for a more up-to-date review of recent developments.

<sup>2</sup> The connection between Hawking radiation and BHT is no doubt a subtext to much of the literature on analogue gravity epistemology, but the connection is rarely considered explicitly.

<sup>3</sup> Wallace is certainly not alone here, nor am I suggesting that he is incorrect. I single him out because this is the most lucid—and perhaps most provocative—statement of the role that Hawking radiation is purported to play in BHT. See also various comments to this effect in Wald (1984), Wald (1994), Jacobson (1996), Lüst and Vleeshouwers (2019), and cf. Wüthrich (2019, p. 203): “[U]ntil Stephen Hawking offered a persuasive semi-classical argument that black holes radiate,

that provides the closest thing to empirical evidence for the Hawking effect. Here, then, the two instances of analogical reasoning clearly make contact. I shall argue that their doing so creates an interesting tension for those who are compelled by BHT and yet remain skeptical about analogue gravity epistemology.

This chapter is primarily intended to clarify the *structure* of the instances of analogical reasoning in contemporary black hole physics, and to clarify how they make contact with one another, rather than intended to directly evaluate whether those instances of analogical reasoning succeed. In a sense, then, this chapter plays the role of prelude to both the existing philosophy of physics literature that seeks to defend or downplay the physical significance of the analogue gravity or BHT analogies, and the existing philosophy of science literature that examines the possibility of confirmation via analogue experimentation in black hole physics.

The structure of this piece is as follows: The first half of the paper (Sect. 14.2) provides a survey of the two central instances of analogical reasoning in contemporary black hole physics: analogue gravity and black hole thermodynamics. I begin with some groundwork on analogical reasoning (Sect. 14.2.1), and then survey, and explicitly reconstruct analogical arguments for, analogue gravity (Sect. 14.2.2) and BHT (Sect. 14.2.3). The second half of the paper (Sect. 14.3) examines what the relationship between these two analogical arguments might be, arguing that they are distinct but nevertheless importantly interdependent, and teasing out an important tension that this implies. I then consider three possible ways to resolve this tension and argue that each will fail. In Sect. 14.4 I conclude.

## 14.2 Two Analogies in Contemporary Black Hole Physics

### 14.2.1 Analogical Reasoning

Before we examine the ways in which analogies and analogical reasoning are invoked in contemporary black hole physics, we should make explicit exactly what is meant by these terms in the round. In this chapter, I adopt a broadly Hessean picture of analogy, following in particular Keynes (1921), Hesse (1966), and Bartha (2010), since these works build, sequentially, on one another and I take them collectively to constitute the mainstream orthodoxy in the analogical reasoning literature. It is worth noting, however, that plenty of dissenting accounts of analogical reasoning exist.<sup>4</sup> My concern, then, is with specifically *Hessean* analogical reasoning in contemporary black hole physics. Whether the problems

---

and so exhibit thermodynamic behaviour like a body with a temperature, most physicists were not moved by Bekenstein's earlier case for black hole entropy:"

<sup>4</sup> See Bartha (2019) for an excellent overview.

I describe in this chapter are remedied or exacerbated by adopting an alternative account of analogy is a question that must be left for another time.<sup>5</sup>

Bartha (2010, p. 1) takes an analogy to be “a comparison between two objects, or systems of objects, that highlights respects in which they are thought to be similar”. The conventional way of formalising this idea appeals to the notion of a mapping. Given a source domain,  $S$ , and a target domain,  $T$ , an analogy is a one-to-one mapping,  $\phi$ , that maps elements of  $S$  to elements of  $T$ . Here, ‘elements’ can be construed quite broadly: elements could be properties, relations, objects, or functions depending on the nature of the domains in question (p. 13).

$$a \in S \leftrightarrow_{\phi} a^* \in T$$

$$b \in S \leftrightarrow_{\phi} b^* \in T$$

$$c \in S \leftrightarrow_{\phi} c^* \in T$$

Starred and unstarred elements are ‘analogues’ of one another under  $\phi$ . From here on I drop the ‘ $\in$ ’ notation and take it as understood that elements on either side of the horizontal relation are from different domains (as a convention, I shall stick to ‘source’ being left, and ‘target’ being right).

It is worth making two useful terminological refinements to our account of analogy:

1. Following Keynes (1921), we can distinguish between ‘positive’, ‘negative’ and ‘neutral’ parts of an analogy. For expositional simplicity, let the elements of our source and target domains be propositions. The positive analogy is the set  $P$  of propositions in  $S$  whose images under  $\phi$  hold in  $T$ . The negative analogy is the set  $N$  of propositions in  $S$  whose images under  $\phi$  do *not* hold in  $T$ . The neutral analogy is the set of propositions in  $S$  for whom it is not known whether their images under  $\phi$  hold in  $T$ , i.e.  $S \setminus P \cup N$ .
2. Following Hesse (1966), we can make a further distinction between ‘horizontal’ and ‘vertical’ relations in an analogy. Horizontal relations are, according to Hesse (1966, p. 59), those concerned with the similarity of counterparts determined by the mapping,  $\phi$ . The horizontal relation is the relation of identity or difference between an element of  $S$  and its image under  $\phi$  in  $T$ . What makes  $a$  and  $a^*$  analogues is the fact that  $\phi$  picks them out as identical with respect to,

---

<sup>5</sup> In any case, I draw only lightly on the specifics of the Hessean account, and so it seems plausible that much of what I say here will generalise to other accounts of analogy. One account for which this may not be true, however, is that of John Norton, who eschews the pursuit of “some universal schema that separates the good from the bad analogical inferences” (2021, p. 120) in favour of a case-by-case analysis of whether there are appropriate empirical facts to warrant specific analogical inferences. Formal analogies, which do not take into account these empirical facts, cannot, on Norton’s account be sufficient to underpin analogical inference. On those grounds, it would seem that much of what I say in Sects. 14.3.1–14.3.3 will not be compelling to someone that adopts Norton’s account of analogy. I thank an anonymous reviewer for this point.

e.g., function, property, mathematical structure, and so on. By contrast, vertical relations, for Hesse are the *causal* relations that obtain between elements of a single domain. For example, how  $a$  relates to  $b$ , and so forth.

It is worth noting that this definition sets the bar trivially low for the existence of *analogies* in black hole physics. One can always propose a comparison between any two systems (black holes and thermodynamical systems, for example) that moots certain similarities, and as such one can always trivially assert the existence of analogies more or less anywhere. This makes the question of whether there are analogies in black hole physics somewhat uninteresting. The answer is yes, for the same reason that there are analogies everywhere—we are pattern-seeking beings, and we are apt to propose patterns or similarities wherever we like. Our concern, rather than with analogy *simpliciter*, is with the earnest deployment of *analogical reasoning* in black hole physics.

Analogical reasoning, Bartha suggests, is “any type of thinking that *relies* upon an analogy” (2010, p. 1, my emphasis). In formal terms, this usually comprises conjecturing that an element currently in the neutral analogy belongs in the positive analogy, with the chosen element of the neutral analogy being referred to by Keynes as the ‘hypothetical analogy’. An ‘analogical argument’ is simply an explicit representation, by way of premise and conclusion, of a piece of analogical reasoning, conforming roughly to the general schema below:

### General Schema

(i) $P \leftrightarrow_{\phi} P^*$	[Positive Analogy]
(ii) $Q \leftrightarrow_{\phi} Q^*$	[Positive Analogy]
(iii) $R \leftrightarrow_{\phi} R^*$	[Positive Analogy]
<hr style="width: 50%; margin: 0 auto;"/>	
$\therefore S \leftrightarrow_{\phi} S^*$	[Hypothetical Analogy]

The rest of this section applies this schema to the two cases of analogical reasoning in contemporary black hole physics, identifying the salient positive analogies, the types of horizontal relationships involved, and any critical disanalogies that appear to block the conclusion. Section 14.2.2 discusses and formalises the analogical argument that underpins analogue gravity research; Sect. 14.2.3 does the same for BHT.

## 14.2.2 Analogue Gravity

Analogue gravity research exploits a precise mathematical isomorphism, first discovered by Unruh (1981), between the behaviour of sound waves in a convergent fluid flow, and the behaviour of light in black hole spacetimes. Because our understanding of the propagation of sound in a fluid is so much better than our understanding of quantum field theory in curved spacetime, the primary benefit of analogue gravity is that a system whose microphysics is well understood can be

used as an empirical surrogate for a system whose microphysics is incredibly poorly understood (since we expect a complete description to require an adequate theory of quantum gravity), and almost entirely inaccessible. To return to the comment I made at the start of this chapter, analogue gravity’s ingenuity stems from the fact that it seems to offer a way to study systems that are not black, and not holes, but are—in some sense—still black holes.

### 14.2.2.1 The Positive Analogy

Crudely put, the positive analogy between analogue black holes and astrophysical black holes is that in both cases we can describe them well mathematically by “[taking] some sort of ‘excitation’, travelling on some sort of ‘background’, and [analysing] its propagation in terms of the tools and methods of differential geometry” (Visser 2013, p. 31). For present purposes, it is not necessary to delve too much further into the technical details of semiclassical or analogue gravity.<sup>6</sup> Rather, we can simply note (as Unruh 1981 originally did) that sound waves in a converging fluid flow can be modelled as the excitations of a minimally-coupled massless scalar field propagating in a (3 + 1)-dimensional Lorentzian geometry with the covariant acoustic metric:

$$g_{\mu\nu}^{\text{acoustic}} = \frac{\rho_0}{c_{\text{sound}}} \begin{bmatrix} -(c_{\text{sound}}^2 - v_0^2) & \vdots & -v_0^j \\ \dots & \cdot & \dots \\ -v_0^i & \vdots & \delta_{ij} \end{bmatrix} \quad (14.1)$$

Which bears a striking resemblance to the Painlevé-Gullstrand form of the Schwarzschild metric:

$$g_{\mu\nu}^{\text{Schwarzschild}} = \begin{bmatrix} -(c_0^2 - \frac{2GM}{r}) & \vdots & -\sqrt{\frac{2GM}{r}} \mathbf{r}_j \\ \dots & \cdot & \dots \\ -\sqrt{\frac{2GM}{r}} \mathbf{r}_i & \vdots & \delta_{ij} \end{bmatrix} \quad (14.2)$$

Thus, the Schwarzschild spacetime and the converging fluid flow share the same effective geometry, encoded by these isomorphic metrics. Using the tools of quantum field theory of curved spacetime, we approximate quantum gravity by propagating quantum fields across this curved but classical spacetime background. In much the same way, sound waves moving in the fluid medium described by (14.1) can be modelled as scalar fields propagating across this geometry.

<sup>6</sup> Readers interested in such details should consult the excellent and comprehensive living review by Barceló et al. (2011), or the philosophical analysis thereof by Thébault (2019, Sects. 11.3.1–11.3.3).

Despite the numerous ways in which analogue black holes remain radically dissimilar to their astrophysical counterparts,<sup>7</sup> the thought is that this underlying mathematical isomorphism is nevertheless strong enough to capture at least *kinematic* similarity between acoustic black holes and their astrophysical counterparts. Thus, the thinking goes, the positive analogy between the mathematics of analogue black holes and astrophysical black holes can support certain inferences about the behaviour of astrophysical black holes from the behaviour of their analogue counterparts.

In principle, there could be many different candidates for what these inferences might be. To put it in Keynes' terms, any property of analogue black holes in the 'neutral analogy' could, in principle, be proposed as the 'hypothetical analogy' (i.e. the property that we are suggesting astrophysical black holes might have, based on the strength of the existing 'positive analogy'). This flexibility is simultaneously analogical reasoning's greatest strength and greatest weakness. But in practice, the inference most commonly made from analogue gravity research has to do with the Hawking effect.

The Hawking effect is the result that at asymptotically late times, particles of a quantum field in a stationary black hole spacetime are radiated from the event horizon out to infinity in precisely the same way as radiation from a perfect blackbody at the *Hawking temperature*,  $T = \frac{\kappa}{2\pi}$ .<sup>8</sup> But because surface gravity scales inversely with mass, and because black holes are typically very massive objects, the Hawking temperature for even solar mass black holes is minuscule (of the order  $10^{-8}$  K) which makes astrophysical Hawking radiation nigh on impossible to detect empirically. The inference that analogue gravity research is commonly taken to support is that direct observation of the analogue Hawking effect in analogue black holes allows us to infer (at least analogically) that astrophysical Hawking radiation does, indeed, exist.<sup>9</sup>

We might worry that because there remain so many differences between analogue and astrophysical black holes, this inference is a spurious one. There are two ways to allay this concern, which both appeal to the robustness of the Hawking effect. First, we can appeal to the fact that various universality arguments have been given to support the idea that the Hawking effect will be robust under changes in the

---

<sup>7</sup> Many of these dissimilarities will be benign. But others seem plausibly more serious. For example, the analogue gravity case involves *two* relevant metrics: acoustic black holes obey the relevant acoustic metric, which is the one we are interested in, but they also obey—purely by dint of being physically realised systems on Earth—the physical spacetime metric, which for us is approximately Minkowski (Barceló et al. 2011, p. 16). Discussing these kinds of issues is regrettably beyond the scope of this chapter—I mention them only to flag that one should perhaps be wary of waving away 'dissimilarities' too hastily.

<sup>8</sup> I will say slightly more about the technical details of the Hawking effect in Sect. 2.3.1.2, but for a more rigorous treatment, see Wald (1994, ch. 7), and for a more thorough discussion, see Wallace (2018, Sects. 4.1–4.2).

<sup>9</sup> Indeed, the title of Unruh's seminal (1981) paper, '*Experimental black hole evaporation?*' attests to this being the central inference of analogue gravity research.

microphysics of the system in which it appears.<sup>10</sup> It is worth noting, however, that Gryb et al. (2020) argue forcefully that none of the available universality arguments for Hawking radiation are entirely convincing (especially in light of the fact that none provide a fully satisfactory response to the trans-Planckian problem, which we shall see in more detail in Sect. 14.3.2). Field (2022) goes one step further, questioning whether the epistemic situation in analogue gravity is even one that can be ameliorated by universality arguments. On Field’s view, “our state of knowledge with respect to Hawking radiation is not currently universality-argument-apt” (2022, p. 25). A safer option, then, is the second: we can note that because the Hawking effect is entirely kinematic and depends only on the existence of a Lorentzian metric and an appropriate horizon (Visser 2003), it would seem to be the kind of thing that *is* determined by the isomorphism we saw above. Thus, analogue gravity advocates suggest, the existence of the Hawking effect in analogue black hole systems would support the existence of the Hawking effect in astrophysical black hole systems. And furthermore, since researchers have demonstrated the existence of both stimulated and spontaneous phononic Hawking radiation from acoustic horizons in fluids (Weinfurter et al. 2011), and from optical horizons in Bose-Einstein condensates (Steinhauer 2016; de Nova et al. 2019; Drori et al. 2019; Kolobov et al. 2021), the antecedent in this conditional seems to be satisfied.<sup>11</sup> Thus, the results of analogue gravity experiments may be considered to indirectly confirm Hawking’s (1974; 1975) prediction of thermalised radiation from astrophysical black holes.

#### 14.2.2.2 Formalisation

Here is a reasonable first pass at formalising the analogical argument for astrophysical Hawking radiation:

##### Analogue Gravity (Mathematical Similarity)

- (i) Mathematics of analogue black holes  $\leftrightarrow_{\phi}$  Mathematics of astrophysical black holes

---

$\therefore$  Radiation from analogue black holes  $\leftrightarrow_{\phi}$  Radiation from astrophysical black holes

This is, of course, not deductively valid. No analogical arguments are. But it is also not sufficiently detailed. For a start, we need to be more precise about the nature of  $\phi$ . In what sense is the mathematics of the two systems similar? After all, the mathematics used to describe analogue black holes is not *entirely* similar to the mathematics used to describe astrophysical black holes—gravitation is not fluid flow.<sup>12</sup> So which parts of the mathematics for the two types of systems are we

<sup>10</sup> See Gryb et al. (2020, Sect. 4) for a review.

<sup>11</sup> See also Unruh (2014) for an appraisal (and endorsement) of early attempts to measure Hawking radiation in black hole analogues, and Leonhardt (2018) for a dissenting view.

<sup>12</sup> See also my fn. 4, above.



interested in, and in what sense are these parts connected? Whichever mathematical commonality (i) is picking out will have to be appropriately related to the proposed physical commonality in the conclusion.

If Visser is correct that Hawking radiation requires only the existence of a Lorentzian geometry and a suitable horizon, then the relevant parts of the mathematics are the  $(3+1)$ -dimensional Lorentzian metric we saw for both systems in Sect. 14.2.2.1, and a horizon within that geometry (i.e. some region that excitations of the scalar field may enter but not leave). This also takes care of the question of vertical relations: Visser's argument explains why these two parts of the mathematics are sufficient to license an inference to the existence of Hawking radiation. And since we are dealing with mathematical structures, the relevant type of mapping is an isomorphism such that ' $\leftrightarrow_{\phi}$ ' denotes 'is isomorphic to'. Here is a second pass:

### **Analogue Gravity (Isomorphic Kinematics)**

- (i) Lorentzian geometry for analogue black holes  $\leftrightarrow_{\phi}$  Lorentzian geometry for astrophysical black holes
- (ii) Horizons in analogue black holes  $\leftrightarrow_{\phi}$  Horizons in astrophysical black holes
- (iii) Hawking radiation requires only a suitable Lorentzian geometry and the existence of a horizon

---

$\therefore$  Radiation from analogue black holes  $\leftrightarrow_{\phi}$  Radiation from astrophysical black holes

But this is too quick, because we can actually read off a deductive argument for astrophysical Hawking radiation by only looking at the right hand side: if astrophysical black holes have a suitable Lorentzian geometry and suitable horizon structure, then the addition of (iii) entails that astrophysical Hawking radiation exists. But if this were the case, there would be no need for analogue gravity! What has gone wrong here is that Visser's argument, i.e. (iii), is fundamentally an argument about differential geometry, not about empirical physics. What I mean here is that Visser's assertion about the 'essential' features of Hawking radiation concerns the *derivation* of the Hawking effect from within a particular theoretical framework—either semiclassical gravity, or hydrodynamics, or condensed matter. It is not—nor does it purport to be—an argument about the existence of Hawking radiation in any empirical sense. This version therefore states Visser's argument too strongly. A third pass:

### **Analogue Gravity (Derivation from Isomorphic Kinematics)**

- (i) Lorentzian geometry for analogue black holes  $\leftrightarrow_{\phi}$  Lorentzian geometry for astrophysical black holes
- (ii) Horizons in analogue black holes  $\leftrightarrow_{\phi}$  Horizons in astrophysical black holes
- (iii) The *derivation* of Hawking radiation requires only a suitable Lorentzian geometry and the existence of a horizon

---

$\therefore$  Radiation from analogue black holes  $\leftrightarrow_{\phi}$  Radiation from astrophysical black holes

This seems to capture the thinking behind analogue gravity without overreaching it.

## 14.2.3 Black Hole Thermodynamics

### 14.2.3.1 The Positive Analogy

BHT hinges upon the recognition that there is a remarkable mathematical similarity between the laws of classical black hole mechanics derived by Bardeen et al. (1973) and the ordinary laws of thermodynamics (Table 14.1).

The formal analogy between (a) surface gravity and thermodynamical temperature, and (b) horizon area and thermodynamical entropy suggests, as Wald (1994, p. 133) puts it, “a close—and, undoubtedly, deep—relationship between the laws of black hole physics and the laws of ordinary thermodynamics”. The basis for this suggestion is that a black hole’s surface gravity *behaves*, mathematically, as if it were the temperature of a thermodynamical system, and the area of a black hole’s event horizon *behaves*, mathematically, as if it were the entropy of a thermodynamical system. This follows from the fact that these quantities occupy structurally equivalent positions in each set of laws. But it does not follow from the formal analogy alone that surface gravity and horizon area *physically are* a black hole’s thermodynamical temperature and entropy respectively. Indeed, there are many good reasons to resist this conclusion in the classical regime, with two of the most obvious being precisely those that make empirical access to black holes so difficult in the first place: (i) there are severe problems when it comes to even delineating where a black hole *is*, which makes it difficult to see how to distinguish ‘the system’ in a thermodynamical context; and (ii) in classical relativity, a black hole is a perfect absorber, and thus, if it has a thermodynamical temperature at all, that temperature could only ever be absolute zero. These are seemingly critical disanalogies between the two domains that prevent the analogy from being taken seriously.<sup>13</sup>

---

<sup>13</sup> These are not the only problems with the analogy between black hole mechanics and ordinary thermodynamics. For instance, there are several deep conceptual puzzles around equating horizon area with entropy—and the concomitant need for a ‘generalised second law’—that I set aside here for reasons of space. For a helpful review of these issues, see Curiel (2019b, especially Sects. 5.3–5.4). I also omit: the *prima facie* lack of a microphysical basis for phenomenological BHT; the puzzling fact that the laws of black hole mechanics are theorems in differential geometry, yet the laws of thermodynamics are bulk empirical generalisations and do not admit of analytical proof (Curiel 2014); the fact that in its current guise, BHT has been accused of recovering only a ‘pale shadow’ of ordinary thermodynamics (Dougherty and Callender 2016). I focus on (i) and (ii) because they seem to pose the most severe challenge to taking BHT seriously, and because the resolution to (ii)—the prediction of Hawking radiation in semiclassical gravity—in particular sets up an interesting tension that provides the basis for discussion in Sect. 14.3.

**Table 14.1** The formal analogy between the laws of black hole mechanics and ordinary thermodynamics

	Black hole mechanics	Thermodynamics
Zeroth law	$\kappa$ constant across the event horizon	$T$ constant throughout the system
First law	$dM = \kappa dA + \Omega dJ + \Phi dQ$	$dU = T dS + p dV + \Omega dJ + \Phi dQ$
Second law	$dA \geq 0$	$dS \geq 0$
Third law	$\kappa \rightarrow 0$ not physically realisable in finite steps	$T \rightarrow 0$ not physically realisable in finite steps

The fundamental idea of BHT is that the analogy between the laws of black hole mechanics and the laws of ordinary thermodynamics is more than just a formal coincidence *in spite of* these disanalogies. According to BHT, the analogy in Table 14.1 is physically significant, which is to say that it broadly indicates that black holes are thermodynamical systems in the fullest sense of the term. At least part of the case for this physical significance comes from addressing critical disanalogies like those above, and thereby extending the positive analogy such that black holes and thermodynamical systems are more plausibly equivalent in some sense. In the next section, I consider the disanalogies (i) and (ii), and show how the responses to these problems bolster the analogical argument for BHT.

### 14.2.3.2 The Negative Analogy

#### System Boundaries

One critical part of the negative analogy between the physics of classical black holes and the physics of ordinary thermodynamical systems is that while ordinary thermodynamical systems are *local* systems that exist in some region of spacetime, classical black holes are not. Classical black holes are typically defined as the region bounded by an event horizon, and the standard definition of an event horizon is the boundary of the causal past of future null infinity (Hawking and Ellis 1973, pp. 311–312; Wald 1984, pp. 299–300).<sup>14</sup> But as Erik Curiel remarks:

This definition is global in a strong and straightforward sense: the idea that nothing can escape the interior of a black hole once it enters makes implicit reference to all future time—the thing can never escape no matter how long it tries. Thus, in order to know the location of the event horizon in spacetime, one must know the entire structure of the spacetime, from start to finish, so to speak, and all the way out to infinity. As a consequence, no local measurements one can make can ever determine the location of an event horizon. That

<sup>14</sup> This is the canonical definition in classical general relativity, though as Curiel (2019a) notes, the precise definition of a black hole varies widely across subdisciplines of physics. I lack the space to do justice to the heterogeneity of these definitions, so I assume the canonical definition for the exposition here.

feature is already objectionable to many physicists on philosophical grounds: one cannot operationalize an event horizon in any standard sense of the term.

(Curiel 2019a, p. 29)

It is extraordinarily difficult to see how global spacetime structures like the event horizon could ever number among the things treated by the theory of thermodynamics, itself a fundamentally operational theory that makes crucial use of the distinction between system and environment. *Prima facie*, this poses a severe challenge for any argument to the effect that we should take BHT seriously.

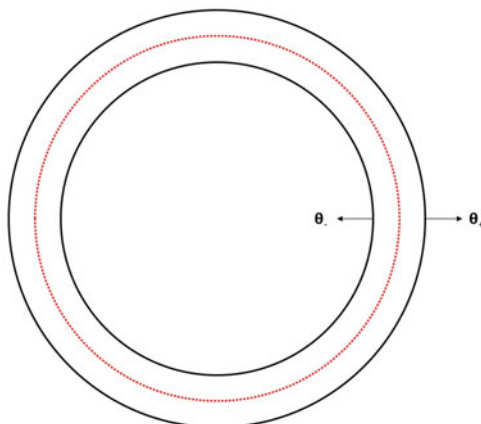
Any practically-minded physicist would rightly balk at the idea that the global nature of event horizons in classical relativity somehow prohibits us from believing BHT, however. For example: Astrophysicists routinely speak about the ‘location’ of black holes (Sagittarius-A\* being at the centre of the Milky Way for instance); the LIGO collaboration are credited with detecting black hole mergers for the first time (Abbott et al. 2016); and scientists from the Event Horizon Telescope were lauded for producing the first direct image of the shadow of what is thought to be a  $6.5 \times 10^9 M_{\odot}$  black hole at the centre of the elliptical galaxy M87 (The EHT Collaboration et al. 2019). There is clearly a sense in which these fabulous achievements presuppose that black holes are entities occupying a well-defined region of spacetime—how else would the EHT team have known where to point their telescope? Thus, we might object, that to argue that black holes cannot be local, dynamical entities because their classical definition is non-local is to fixate on a technicality. If anything, the objection would go, this problematic global nature of the event horizon speaks to a deficiency of classical relativity as a tool for describing the universe, not to an intrinsic deficiency of astrophysical black holes that prevents them from being thermodynamical systems.<sup>15</sup> But an objection such as this, that suggests that classical relativity should not be the final arbiter of the physical nature of black holes, also suggests, a fortiori, that an analogy between classical relativity and ordinary thermodynamics is equally deficient as a guide to the physical nature of black holes.

As such, there are good pragmatic arguments against allowing the global nature of the event horizon to block the argument for BHT. But there are also more principled arguments, which draw from the wide variety of locally-definable alternatives to the global event horizon to argue that BHT can be made suitably ‘local’. The most well-known of these is the trapping horizon, which takes as

---

<sup>15</sup> Many other physical theories will face similar problems. An adiathermal barrier in ordinary thermodynamics is arguably global in much the same sense: how long and how closely do we have to watch to convince ourselves that a system is ‘truly’ isolated? Much as we are happy to say that a thermodynamical system can be considered isolated *for practical purposes*, we can say the same for the event horizon. As Ramesh Narayan remarks in a comment quoted in Curiel (2019a, p. 31, Box 1): “for all intents and purposes we *are* at future null infinity with respect to Sagittarius-A\*”. The ubiquity of strictly global concepts in physics that are nevertheless operationalisable is grist to the practically-minded physicist’s mill.

**Fig. 14.1** A closed spacelike surface (dotted) with two future-directed null vectors. The surface is *trapped* when both  $\theta_+$  and  $\theta_-$  are negative (Color figure online)



its departure point the notion of a ‘trapped surface’ (Penrose 1965).<sup>16</sup> A trapped surface is a closed, spacelike surface with normal, future-directed null vectors that have an expansion parameters,  $\theta_+$  and  $\theta_-$ , which are both negative (Nielsen 2009). Loosely put, this means that not just inward but also outward lightlike vectors are infalling, toward the interior of the trapped surface (Fig. 14.1). A trapping horizon is an extremal trapped surface such that  $\theta_+ = 0$ .<sup>17</sup>

Whether BHT can be localised in terms of trapping horizons is a matter of dispute. Nielsen (2009) identifies them as the forerunner in a sizeable group of locally-defined alternatives to the event horizon, for a number of reasons whose technical details go beyond the scope of this paper (pp. 39–40).<sup>18</sup> While the suitability of the trapping horizon for truly ‘localising’ BHT has been questioned by Dougherty and Callender (2016, Sect. 4), others have come to its defence (Wallace 2018, A.2). Insofar as the jury remains out on this question, it seems reasonable to treat it as a live and promising strategy for reformulating BHT without the global event horizon. For our purposes, all that matters is that there *are* viable local alternatives to the event horizon that take the sting out the critical analogy discussed above.

<sup>16</sup> See Nielsen (2009) for a more detailed discussion of why we might want to expel the globally-defined event horizon from BHT, and for a nice review of a wider range of locally-defined alternatives to the event horizon.

<sup>17</sup> I gloss over some technical subtleties here, to do with distinguishing trapping horizons from marginally trapped tubes. See Hayward (1994) for further details.

<sup>18</sup> Furthermore, he notes (p. 36) that in the case of analogue gravity, it is trapping horizons (not event horizons) that we are dealing with. As such, there are reasons to adopt the trapping horizon as a local alternative to the event horizon if one wishes to further substantiate the connection between astrophysical black holes and their analogue counterparts we discussed in Sect. 14.2.2.

## Hawking Radiation

Taking the mathematical similarity between surface gravity and thermodynamic temperature seriously would mean assigning a black hole a thermodynamic temperature of  $T = \alpha\kappa$ , where  $\alpha$  is some constant to be determined. In classical general relativity this stands in direct contradiction to the fact that a black hole is a perfect absorber, prevented by its very definition from emitting radiation. The only thermodynamic temperature that it makes sense to assign to a black hole is zero, and yet such black holes generically have nonzero surface gravity. This rules out taking the formal analogy between surface gravity and temperature seriously.

In the semiclassical framework, however, this is no longer the case. In semiclassical gravity, we approximate quantum corrections to classical relativity by propagating quantum fields on a classical but curved spacetime background (a modelling framework that will be familiar from Sect. 14.2.2). And in this framework, as Hawking's (1975) celebrated result shows, black holes emit radiation with a perfect thermal spectrum corresponding to the Hawking temperature (setting  $G = c = \hbar = 1$ ):

$$T_H = \frac{\kappa}{2\pi} \quad (14.3)$$

Which is exactly what the formal analogy suggests.

The precise role of Hawking radiation in bolstering the analogy between black hole physics and ordinary thermodynamics is not always entirely clear, but *prima facie* it achieves at least the following two things. First, it remedies the numerical disagreement between the classical black hole's surface gravity and its temperature, since if black holes are no longer perfect absorbers, they need not have zero temperature. Second, it provides a way to make sense of the notion of thermal contact between black holes. Let two black holes of surface gravity  $\kappa_1$  and  $\kappa_2$ , respectively, occupy a box that is sufficiently large so as to allow us to neglect their gravitational interaction with one another. Radiation from each will cause the box to reach a temperature  $\kappa_1/2\pi > T_{\text{Box}} > \kappa_2/2\pi$ , resulting in the movement of heat from the hotter black hole to the colder (Wallace 2018). Thus, Hawking radiation moves two important elements of the negative analogy (numerical disagreement between  $T$  and  $\kappa$ ; the lack of a well-defined notion of thermal contact between black holes) into the positive analogy.

Two problems hamper Hawking radiation's role in substantiating the BHT analogy, however. The first problem is that for a solar mass black hole,  $T_H$  is of the order  $10^{-8}$  K, making it some 100 million times smaller than the 2.7 K cosmic microwave background temperature, and a thousand times smaller than the  $\sim 10^{-5}$  K fluctuations in the cosmic microwave background alone (Smoot et al. 1992). The prospect of direct detection of astrophysical Hawking radiation is therefore incredibly poor. As Thébault (2019, pp. 185–186) vividly puts it: “Trying to detect astrophysical Hawking radiation in the night sky is thus like trying to see the heat from an ice cube against the background of an exploding nuclear bomb.”

Indeed, as we saw in Sect. 14.2.2, the entire field of analogue gravity research can be understood as a reaction to the astonishingly dim prospects for the direct detection of astrophysical Hawking radiation. The second problem is that there are a battery of fairly severe conceptual problems with Hawking's theoretical derivation. Some of these have been resolved in more modern derivations of the Hawking effect by other means,<sup>19</sup> but some remain unresolved and are therefore cause for concern over the existence of astrophysical Hawking radiation. I discuss both of these problems in more detail in Sect. 14.3, but for now we have enough to sketch a formalisation of the analogical argument for BHT.

### 14.2.3.3 The Hypothetical Analogy?

Together, local alternatives to the event horizon, the existence of astrophysical Hawking radiation and the generalised second law substantially strengthen the analogical argument for BHT. But towards which conclusion? There are two plausible candidates. On the one hand, the conclusion might be that the formal analogy between the specific mathematics under comparison in Table 14.1 is now endowed with physical significance. On the other hand, the conclusion might be that the formal analogy between black holes and ordinary thermodynamical systems is *complete*—that is, that it extends far beyond what is already encoded in the similarities of Table 14.1, to *all* the physical quantities of black holes and thermodynamical systems. The first conclusion can be understood as attempting to *deepen* the purported connection between black holes and ordinary thermodynamical systems; the second as attempting to *broaden* that connection.

There are reasons both for and against each interpretation. If the connection between black holes and thermodynamical systems is to be deepened by an appeal to 'physical significance', then we owe an account of what this mysterious property might be.<sup>20</sup> But similarly, if the scope of this connection is to be broadened by extending the formal analogy, we soon run into problems. For instance: How are we to account for the missing  $pdV$  term in the first law of black hole mechanics (Dolan 2012)? What about the fact that the Hawking temperature generically disagrees with the Tolman temperature, which is the canonical measure of temperature in relativistic thermodynamics, since the two only coincide at future null infinity?<sup>21</sup> And how are we to understand the fact that the laws of black hole mechanics are

---

<sup>19</sup> See the useful discussion and provided by Wallace (2018, Sect. 4.2), and references therein, for more detail.

<sup>20</sup> This is a notoriously vexed question that cuts right to the heart of the relationship between mathematics and physics. Some recent accounts attempt to develop the connection between physical significance and topological stability (Fletcher 2014, 2016), but this seems to miss the mark for our purposes: the formal analogy in Table 14.1 does not seem to be the kind of thing that could be considered 'topologically stable'.

<sup>21</sup> I thank Erik Curiel for useful discussions on this point.

*theorems*, while the laws of ordinary thermodynamics admit of no analytic proof (Curiel 2014)?

I shall adopt the ‘deepening’ interpretation, going forward, since the problems with it are problems of clarification rather than problems of contradiction. That is to say: it seems more feasible to provide a satisfactory account of physical significance than it does to provide an exhaustive argument resolving all of the myriad possible mathematical disanalogies that the broadening approach would require.<sup>22</sup>

#### 14.2.3.4 Formalisation

The most coarse-grained version of the analogical argument for BHT is simply to state that the positive analogy between the laws in Table 14.1 suggests that the systems described by these laws are really and truly the same kinds of systems. A schematisation of this coarse-grained argument would then be:

##### **Black Hole Thermodynamics (Nomic Isomorphism)**

(i) Laws of black hole mechanics  $\leftrightarrow_{\phi_1}$  Laws of ordinary thermodynamics

---

$\therefore$  Black holes  $\leftrightarrow_{\phi_2}$  Thermodynamical systems

Modulo some minor discrepancies—notably the absence of a  $pdV$  term— $\phi_1$  is a horizontal relationship of isomorphism, since we are comparing structurally identical mathematical equations.  $\phi_2$ , following the discussion in Sect. 14.2.3.3, has two plausible interpretations, but the better of the pair is that black hole quantities are *physically*, not just mathematically, equivalent to thermodynamical quantities.<sup>23</sup> But as we have seen, this argument seems nevertheless to fail due to the critical negative analogy between various aspects of black hole physics and various aspects of thermodynamics, namely the global nature of event horizons, and the lack of radiation from classical black holes. As we saw above, however, local alternatives to the event horizon, coupled with the derivation of Hawking radiation in the semiclassical framework allow us to smooth over two of the most pressing disanalogies. The result is a stronger and more detailed analogical argument:

##### **Black Hole Thermodynamics (Nomic Isomorphism + Boundaries and Radiation)**

- (i) Laws of black hole mechanics  $\leftrightarrow_{\phi_1}$  Laws of ordinary thermodynamics
- (ii) System boundaries for astrophysical black holes  $\leftrightarrow_{\phi_1}$  System boundaries for thermodynamical systems

---

<sup>22</sup> Developing a satisfactory account of physical significance goes far beyond the scope of this paper. I endorse the deepening account here only insofar as it seems more plausible than the broadening account.

<sup>23</sup> Cashing this out with greater precision lies regrettably beyond the scope of the paper. I am sympathetic to the broadly functionalist position espoused by Curiel (2014) and later augmented by Prunkl and Timpson (2019), upon which ‘physical equivalence’ should be understood as a kind of functional equivalence, and will touch upon this briefly in Sect. 14.3.3.



(iii) Radiation from astrophysical black holes  $\leftrightarrow_{\phi_1}$  Radiation from thermodynamical systems

---

$\therefore$  Black holes  $\leftrightarrow_{\phi_2}$  Thermodynamical systems

### 14.3 What Is the Relationship Between Them?

We have arrived at plausible formalisations for the analogical arguments for astrophysical Hawking radiation and for the physical significance of BHT. These arguments are distinct from one another. Primarily, because they appeal to different premises to draw different conclusions. But also because the conclusions they seem to license are of a fundamentally different kind. On the one hand, analogue gravity concerns the ability of one type of physical system to serve as an experimental surrogate for another type of physical system. This is an epistemic matter: it concerns the validity of certain kinds of experiments that scientists might conduct in order to gain indirect access to empirically inaccessible systems. On the other hand, BHT concerns the physical nature of black holes. The upshot of taking BHT seriously is, to quote the physicist Robert Wald, “that the laws of black hole mechanics *truly are* the ordinary laws of thermodynamics applied to a system containing a black hole” (1994, p. 163, my emphasis). Whereas nobody appears to think that analogue gravity evinces some profound and hitherto unappreciated connection between waterfalls and black holes, precisely that kind of connection is being claimed in the case of BHT—in its more extreme form, the claim is that “gravity [...] is a fundamentally thermodynamical phenomenon” (Curiel 2014, p. 3, see also Curiel 2019b, Sect. 5.5). Not only do the two analogical arguments invoked in contemporary black hole physics concern different premises in support of different conclusions, those conclusions also constitute fundamentally different kinds of claim.

It would be too hasty, however, to conclude that the two analogical arguments can be divorced from one another completely. Because a large part of our evidence for astrophysical Hawking radiation comes from the impressive results of analogue gravity research, and because the existence of astrophysical Hawking radiation in turn constitutes such a crucial part of the analogical argument for BHT, there is a clear sense in which these two analogical arguments bear on one another. Indeed, if the analogical argument for Hawking radiation fails, so too must the analogical argument for BHT. One way to see this is to note that one side of the conclusion in the final version of the analogical argument for Hawking radiation appears as one side of premise (iii) in the final version of the analogical argument for BHT. Consequently, the analogical argument for astrophysical Hawking radiation appears as a component of the analogical argument for BHT.

Why is this interesting? Because it appears to tie together, in a very specific way, the epistemic legitimacy of the BHT analogy, the basis of a research programme that is “as widely accepted an idea in theoretical physics as an idea with no direct

empirical substantiation can be” (Curiel 2019b), with that of analogue gravity, which some have dismissed as little more than an “amusing feat of engineering [that] won’t teach us anything about black holes” (Daniel Harlow, quoted in Wolchover 2016). If the analogical argument for BHT presupposes the success of the analogical argument for astrophysical Hawking radiation, then it is difficult, *prima facie*, to see how these two attitudes can be reconciled. That is, it is hard to see how it is consistent to (a) interpret the connection between black hole mechanics and ordinary thermodynamics as having genuine physical significance while (b) remaining skeptical about the epistemic warrant that analogue gravity gives us for astrophysical Hawking radiation. Yet it is exactly this combination of attitudes that some wish to adopt: Harlow himself, for instance, is actively engaged in research on BHT while remaining openly skeptical about the results of analogue gravity. Similarly, Wüthrich (2019, p. 221) claims that BHT certainly has “some support”, but elsewhere concludes starkly that analogue confirmation “does not work” (Crowther et al. 2021, p. 3723).

I can see three plausible strategies for resolving this tension. In the rest of this section, I shall examine each in turn.

### 14.3.1 *Naïve Formalism*

I take it to be uncontroversial that there is a clear distinction to be drawn between saying “that is a black hole” while pointing one’s finger at the Schwarzschild metric, and saying “that is a black hole” while pointing one’s finger into the night sky in the direction of Sagittarius-A\*. The former identifies a black hole as a type of mathematical entity, while the latter identifies a black hole as a type of astrophysical entity. Of course, we should expect that there is some connection between the two things (indeed, the use of mathematics in physics presupposes such a connection) but the fact remains that they are distinct: the mathematics is not the physical system it describes.

One way to resolve the tension for those who wish to combine belief in the physical significance of BHT with scepticism about analogue experimentation might be to try to interpret the conclusions of the two analogical arguments as being conclusions about *mathematical*—rather than astrophysical—entities. On this view, the term ‘black holes’ is to be understood as referring to certain classes of spacetime models in general relativity, rather than as referring to real physical systems in our universe, which may be reasonably well described by certain mathematical models under certain idealisations. Call this strategy ‘naïve formalism’, in recognition of its wholesale focus on the formalism of relativity.

When the naïve formalist says ‘black holes are thermodynamical systems’, roughly what they mean is that black holes (*qua* particular solutions to the Einstein field equation) have certain mathematical properties in common with

thermodynamical systems.<sup>24</sup> One of these mathematical properties will be that which corresponds to their tendency to produce a thermally distributed radiative flux in certain situations: blackbody radiation in the case of ordinary thermodynamical systems, and Hawking radiation from black holes.<sup>25</sup> But this mathematical property can obtain for black holes—and so can be shared with thermodynamical systems—completely irrespective of any empirical confirmation of astrophysical Hawking radiation. Matters of *mathematical* similarity are to be settled by the mathematics alone.

For the naïve formalist, then, the claim that BHT presupposes the epistemic legitimacy of analogue experimentation fails. This is because if we are only interested in black holes as mathematical entities, we need only be interested in Hawking radiation as a mathematical entity, too. And since Hawking radiation can be derived for black hole spacetimes in semiclassical gravity without any assumptions about whether acoustic black holes can serve as legitimate empirical surrogates for astrophysical black holes, black holes (*qua* mathematical entities) can be said to have the correct mathematical properties so as to be considered ‘thermodynamical’ without any recourse to analogue gravity. To put the same point slightly differently, this objection would say that claims about the thermodynamical nature of black holes and claims about the existence of Hawking radiation are, at bottom, claims about the mathematical machinery of our physical theories. The same claims could be made about *astrophysical* black holes, and indeed then one would need a story about the empirical warrant for *astrophysical* Hawking radiation. But, the objection might go, this is a separate claim.

There are two arguments against the naïve formalist, however. The first is a straightforward sociological observation: nobody actually seems to interpret BHT this way. Presumably, we are interested in formal properties of mathematical entities only insofar as we expect those properties to manifest in the physical world—that is the essence of ‘physical significance’. It should not be surprising that nobody interprets BHT in this extremely narrow sense, since to do so is perhaps to fixate on the mathematics at the expense of the physics. While naïve formalism could plausibly resolve the tension for those willing to adopt it, the cost is an unfamiliar and arguably myopic interpretation of BHT.

A second argument is that naïve formalism is simply too weak to capture the central thesis of BHT. If the central idea of BHT is supposed to be that “black holes are thermodynamical systems *in the fullest sense*” (Wallace 2018, p. 52, emphasis added), it stands to reason that this sense includes not just shared mathematical properties of the two kinds of systems, but shared non-mathematical properties too. Black holes must actually *be* thermodynamical systems (whatever that means),

---

<sup>24</sup> This is still terribly imprecise, but the general idea should be clear enough: the two mathematical entities share certain features that renders them ‘equivalent’ in the sense that BHT has in mind.

<sup>25</sup> Strictly speaking, in the case of black holes, it is not quite right to attribute this mathematical property to the spacetime geometry alone, since one requires appropriate quantum fields propagating on the geometry to produce this flux. I thank an anonymous reviewer for this point.

not just be mathematically equivalent to them. Thus, we might worry, the naïve formalist fails to actually demonstrate anything beyond formal analogy between black hole mechanics and ordinary thermodynamics. Rather, they smuggle in the physical significance of that analogy by simply moving the goalposts, redefining what physical significance means in flat-footed, mathematical terms. It would, of course, be entirely possible to parry this second objection to naïve formalism by adopting a kind of Pythagoreanism about physics, such that there is nothing to *being* a certain kind of thing over and above having the appropriate mathematical properties. But this kind of radical Pythagoreanism is surely an inordinately heavy price for the naïve formalist to pay.

### 14.3.2 *Sophisticated Formalism*

A second strategy is to break with the naïve formalist approach by maintaining that ‘black holes’ are to be understood as astrophysical entities, and therefore to maintain that astrophysical Hawking radiation must exist in more than just a mathematical sense, but to argue that the theoretical basis for astrophysical Hawking radiation is already strong enough to justify this without the results of analogue gravity. Broadly stated, the animating idea for this second strategy is that the absence of empirical evidence for the thermodynamical behaviour of black holes need not be a severe problem so long as we can fall back on non-empirical evidence for that behaviour, including non-empirical evidence for Hawking radiation.

Call this view ‘sophisticated formalism’. Like naïve formalism, it retains a focus on the importance of *theory*, in the sense that it takes a sufficiently strong theoretical basis for astrophysical Hawking radiation to render confirmation of astrophysical Hawking radiation by analogue experimentation redundant. Unlike naïve formalism, however, sophisticated formalism does not limit attention to theory completely. Rather, it maintains that theory can be strong enough to support inferences about the empirical world, even in the absence of experiment. Sophisticated formalism can be seen as a refinement of the ideas of naïve formalism: both identify the importance of theory for decoupling the two analogical arguments, but where the naïve formalist attempted to achieve this by radically reinterpreting BHT as a thesis about mathematical entities, their sophisticated counterpart attempts to achieve this by instead lowering the evidential bar for the existence of astrophysical Hawking radiation.

Something like sophisticated formalism is what underpins pre-analogue gravity attempts to endow BHT with physical significance. This extends to black hole entropy, too: Bekenstein’s (1972) arguments for black hole entropy on the basis of information theory, for instance, are one such attempt, and these are (trivially) independent of any analogue gravity results, predating Unruh’s seminal work by a

whole decade.<sup>26</sup> As Wüthrich (2019, p. 203) notes, to the extent that non-empirical theory confirmation is legitimate, “the fact that thermodynamic behaviour of black holes has not been observed to date may not be a worry as long as we have convincing non-empirical reasons for believing the Bekenstein-Hawking formula”. This is an elegant statement of the sophisticated formalist position. However, Wüthrich’s own conclusion is that Bekenstein’s information-theoretic reasons fail to provide such convincing non-empirical reasons.<sup>27</sup>

Although Bekenstein’s original arguments for the physical significance of BHT have failed to convince contemporary philosophers of physics, Wallace (2018) has argued forcefully that the contemporary theoretical arguments for the existence of astrophysical Hawking radiation are, in fact, “very powerful” (p. 61), and have only been strengthened in the 50 years since Hawking’s original work. Now, we have at least five independent means of deriving the Hawking effect, each with their own strengths and weaknesses. This consilience is, itself, evidence for the robustness of the Hawking effect, but the theoretical case is not entirely watertight, as we shall see in a moment. It is, however, nevertheless true that the *prima facie* strength of the arguments for the Hawking effect make the prospects for the sophisticated formalist seem initially promising, and this is bolstered further by the sociological observation that many physicists seem to espouse this view (one need only look at the huge literature on the information loss paradox to be persuaded that physicists are sufficiently convinced of the existence of Hawking radiation to be concerned about its consequences for unitarity).

But there are two problems for the sophisticated formalist. The first is that despite the theoretical basis for astrophysical Hawking radiation from black holes being very strong, it is not perfect. For example, it is a matter of some delicacy determining *what* the Hawking temperature should actually be attributed to. Giddings (2016) has argued that Hawking radiation originates not from the black hole horizon but from a region “well outside the horizon” (p. 40).<sup>28</sup> If Giddings is correct, then it is not quite right to say that the *horizon* is the source of the thermalised radiation. This poses a challenge for the sophisticated formalist’s claim that the theory unequivocally militates in favour of Hawking radiation *from black holes*, since that radiation in fact comes from a region in the black hole exterior. This is a pedantic point, of course—we might think that the quantum region Giddings describes is

---

<sup>26</sup> I thank an anonymous reviewer for this example.

<sup>27</sup> It is worth noting that Wüthrich appears to equivocate on the broader question of whether these kinds of (what I would call) ‘sophisticated formalist’ arguments could ever be sufficient to establish the physical significance of BHT. On the one hand, he notes that compelling thought experiments can and do “lend some support to the idea that black holes are thermodynamic in nature” (p. 221), but on the other hand he maintains a few lines later that “only the usual kind of experimental and observational work can establish that black holes are thermodynamic objects”. I read Wüthrich as rejecting the possibility of sophisticated formalism, but the fact remains that this kind of attitude prevails among the physics community, for whom recovering the Bekenstein-Hawking formula has “become something of a *sine qua non* for programs of quantum gravity” (Curiel 2019b).

<sup>28</sup> See also Unruh (1977), who makes a similar point.

close enough for the sophisticated formalist's purposes, or that the production of Hawking radiation in that region can nevertheless be causally attributed to the presence of the black hole. But even if these objections succeed, there are other technical challenges beyond that posed by Giddings. The most notorious is the so-called 'trans-Planckian problem', which refers to the fact that the finite wavelengths of the Hawking radiation particles measured at future null infinity, which have been gravitationally redshifted to a drastic extent by their journey away from the black hole, must correspond to particles originating from the black hole with wavelengths arbitrarily shorter than the Planck length. Intuitively, this strongly suggests that the physics giving rise to the Hawking effect near the black hole is beyond Planck scale, i.e. from a regime where we should expect quantum gravitational corrections to relativity to become non-negligible.<sup>29</sup> Polchinski (1995) has argued that the trans-Planckian problem can be finessed by the so-called 'nice slice argument', which roughly says that the Hawking effect can be derived using only 'nice' spacetime slices upon which *only* low-energy (i.e. cis-Planckian) physics is happening. Polchinski's argument is echoed by Wallace (2018), but Gryb et al. (2020, Sect. 2.3) argue that certain assumptions in the nice slice argument make it essentially question-begging. The extent to which the trans-Planckian problem remains an issue for the sophisticated formalist is therefore unclear, but generally, it seems prudent to tread carefully. As Wallace himself rightly notes at the end of his defence of the theoretical basis for the Hawking effect, "as good scientists we should remind ourselves that [the Hawking effect] remains purely theoretical, and that tests of quantum field theory itself in the curved-spacetime regime to date have been much less precise and numerous than in the flat-spacetime regime" (2018, p. 62).

Technical worries about the theoretical basis for Hawking radiation aside, there is a second argument against sophisticated formalism. Namely, that upon closer inspection, the position seems inconsistent. At bottom, to be a sophisticated formalist is to try to rid BHT of any dependence on analogue gravity, and to be motivated to do so by a skepticism about the epistemology of analogue experimentation in general. But this seems necessarily to involve replacing one dubious epistemic technique with another. Is bare theory really more capable than analogue gravity when it comes to providing evidence for the existence of Hawking radiation? In thinking so, the sophisticated formalist seems to be trying to use one hand to set the bar for empirical confirmation so high that analogue experimentation cannot reach it, while using the other hand to remove the bar completely. If theory is strong enough to confirm, on its own, then so too should analogue experimentation be, a fortiori. While it would be consistent to maintain either (a) full-blooded empiricism, such that we have reason to believe that astrophysical Hawking radiation exists if and only if we have suitable empirical evidence of astrophysical Hawking radiation; or (b) or full-blooded anti-empiricism, such that a belief in the existence of astrophysical Hawking radiation can be justified even by non-empirical (e.g.

---

<sup>29</sup> See Jacobson (2005) and references therein for more detail on the trans-Planckian problem.

theoretical or analogical) evidence, the sophisticated formalist tries to be both and arguably achieves neither.<sup>30</sup> It seems plainly inconsistent for the sophisticated formalist to demand that the existence of astrophysical Hawking radiation requires, empirically, *more* than the evidence provided by analogue gravity, but no more than the evidence provided by bare theory alone.

### 14.3.3 *Classicalism*

The third strategy is the least fiddly, but the most radical. Rather than entering the treacherous ground of attempting to describe how non-empirical evidence of astrophysical Hawking radiation may nevertheless justify belief in the empirical existence of that radiation, the third strategy de-couples the two analogical arguments by severing all ties between the physical significance of BHT and the existence of astrophysical Hawking radiation. It does so by arguing that *classical* BHT—BHT without, *inter alia*, Hawking radiation—is already physically significant. Call this strategy ‘classicalism’.

The clearest statement of the classicalist position is given in Curiel (2014). For a black hole to be considered a thermodynamical system, Curiel tells us, what surely matters is only that “surface gravity and area couple to ordinary thermodynamical systems in the same way as temperature and entropy, respectively, do”, and that counterpart quantities are introduced using into the theory using the same “constructions and arguments” (2014, p. 3). Curiel goes on to argue that since one can, indeed, construct a Carnot cycle operating between a black hole and a thermodynamic fluid, and since  $\kappa$  appears in place of the black hole’s ‘temperature’ when one defines the efficiency of that Carnot cycle, there is little room to doubt that surface gravity really *is* a temperature. On the classicalist view, one need not wait for Hawking radiation to establish the physical significance of the analogy between the two quantities. Insofar as Curiel’s arguments succeed, they establish that BHT need not presuppose the epistemic legitimacy of analogue gravity in the way that I have described, simply by virtue of the fact that premise (ii) in the second formalisation of the analogical argument for BHT above is doing no work. The conclusion—the classicalist would have it—follows immediately from (i) and (ii).

There are at least two reasons to think that the classicalist response fails, however. The first is that Curiel’s Carnot cycle fails to be reversible in the standard thermodynamical sense: it results in an increase in the black hole’s irreducible mass,

---

<sup>30</sup> One example of a consistent anti-empiricist attitude would be considering empirical evidence for astrophysical Hawking radiation to be supererogatory. On this view, empirical confirmation is a bonus, but is not necessary for justified belief in astrophysical Hawking radiation. Such an anti-empiricist might consistently maintain that analogue gravity research fails to constitute such ‘bonus’ evidence, but that the theoretical arguments alone were sufficient to warrant belief in astrophysical Hawking radiation. They would still have to respond to the technical problems laid out above, however. I thank an anonymous reviewer for pressing me on this point.

and thus an increase in its horizon area, which in classical relativity cannot be reversed on pain of violating the second law of BHT—Hawking’s area theorem (Curiel 2014, pp. 16–17). One could reasonably question whether this cycle can be considered sufficiently similar to a thermodynamical Carnot cycle to justify the very strong conclusion of the analogical argument for BHT.<sup>31</sup>

The second is that while the classicalist response may be viable for establishing the physical significance of the analogy between surface gravity and temperature, it seems impossible—in principle—to establish the physical significance of the analogy between horizon area and entropy with the same logic. This is because, as Curiel notes (2014, p. 10, fn. 20) thermodynamic entropy mediates no known physical processes, and so there is, ipso facto, no process one could construct so as to demonstrate that it is mediated by horizon area in the same way it would have been mediated by entropy in the thermodynamical case. Thus, while classicalism may succeed in removing the reliance of BHT on the analogical argument for astrophysical Hawking radiation, it seems to do so at the price of only recovering, at best, a partial equivalence between black holes and thermodynamical systems.

## 14.4 Conclusion

Contemporary black hole physics is an interesting case study for the epistemology of science because it deploys analogical reasoning in a way that seems to overreach the traditional heuristic role of analogy. There are two distinct ways in which it does so. First, it relies on analogue experimentation and the observation of analogue Hawking radiation in acoustic and optical black holes to justify the *existence* of astrophysical Hawking radiation. Second, it relies on the analogy between the laws of black hole mechanics and the laws of ordinary thermodynamics to justify the claim that black holes are genuinely thermodynamical systems.

The analogical arguments in favour of these two claims are distinct. Each one appeals to different premises in order to draw different conclusions, and those conclusions seems to be fundamentally different in kind. But because the BHT analogy is typically only considered physically significant with the inclusion of Hawking radiation, there is an important sense in which the two analogical arguments are linked. Indeed, the first can be nested inside the other in such a way that the analogical argument for BHT cannot succeed without the analogical argument for the existence of astrophysical Hawking radiation. This connection, I argued in Sect. 14.3, leads to a tension for those who wish to combine an optimism about the physical significance of BHT with a pessimism about the epistemic warrant provided by analogue experiments.

---

<sup>31</sup> Interestingly, the cycle Curiel describes can effectively be *made* reversible when one takes Hawking radiation into account (Prunkl and Timpson 2019), but this rather defeats the point of the classicalist strategy we are considering.



I considered three strategies for resolving this tension, and showed that each one fails. The naïve formalist (Sect. 14.3.1) and the classicalist (Sect. 14.3.3) strategies may succeed in decoupling the two analogical arguments, but they do so at the cost of settling for impoverished versions of BHT: one nakedly mathematical; the other only partially complete. The sophisticated formalist (Sect. 14.3.2), on the other hand, seems to replace the tension between optimism about BHT and skepticism about analogue gravity with another, new tension. The sophisticated formalist is committed to the contradictory idea that analogical evidence is inadequate to the task of justifying the existence of astrophysical Hawking radiation, and yet that same justification may be amply provided by bare theory—something *weaker*, not stronger, than analogical evidence.

If these three strategies exhaust the space of possibilities, then the tension I pointed out at the start of Sect. 14.3 remains: the analogical argument for BHT presupposes the success of the analogical argument for the existence of astrophysical Hawking radiation. Thus, it presupposes the epistemic legitimacy of analogue gravity. It is difficult to see how Harlow's assertion that analogue gravity is no more than an amusing feat of engineering can be reconciled with the belief that black holes are, nevertheless, genuinely thermodynamical systems.

Perhaps the best prospects for resolving this tension come from a retreat to the 'traditional' role of analogy: as heuristic. If there are reasons to believe that these analogical arguments will, in time, be superseded by robust non-analogical arguments, then there are, *eo ipso*, reasons to believe that this tension will, in time, dissipate. It is far from clear, however, that the analogical arguments discussed here are even capable of being replaced by non-analogical arguments. Insofar as the magnitude of Hawking radiation from an astrophysical black hole and the magnitude of CMB fluctuations are fixed by the nature of physical law, the impossibility of direct detection of Hawking radiation seems fixed with at least nomological necessity. One way around this, which we have not considered in this chapter, would be the discovery of primordial black holes, whose low mass would result in a Hawking temperature sufficiently high to admit direct detection. Until such a discovery, it seems we should make peace with the fact that black hole physics seems destined to continue depending upon analogical reasoning in this unusually strong way.

## References

- Abbott, B.P., R. Abbott, T.D. Abbott, M.R. Abernathy, F. Acernese, K. Ackley, et al. 2016. Binary black hole mergers in the first advanced LIGO observing run. *Physical Review X* 6: 041015. <https://link.aps.org/doi/10.1103/PhysRevX.6.041015>.
- Bambi, C. 2018. Astrophysical black holes: A compact pedagogical review. *Annalen der Physik* 530: 1700430.
- Barceló, C., S. Liberati, and M. Visser. 2011. Analogue gravity. *Living Reviews in Relativity* 14. <http://www.livingreviews.org/lrr-2011-3>.

- Bardeen, J.M., B. Carter, and S.W. Hawking. 1973. The four laws of black hole mechanics. *Communications in Mathematical Physics* 31(2): 161–170.
- Bartha, P. 2010. *By parallel reasoning: The construction and evaluation of analogical arguments*. Oxford: Oxford University Press.
- Bartha, P. 2019. Analogy and analogical reasoning. In *The Stanford encyclopedia of philosophy*, E.N. Zalta, ed., Spring 2019 edn. Metaphysics Research Lab, Stanford University.
- Bekenstein, J.D. 1972. Black holes and the second law. *Lettere al Nuovo Cimento (1971–1985)* 4(15): 737–740. <https://doi.org/10.1007/BF02757029>.
- Crowther, K., N.S. Linnemann, and C. Wüthrich. 2021. What we cannot learn from analogue experiments. *Synthese* 198: 3701–3726. <https://doi.org/10.1007/s11229-019-02190-0>.
- Curiel, E. 2014. Classical black holes are hot. arXiv:1408.3691 [gr-qc].
- Curiel, E. (2019a), The many definitions of a black hole. *Nature Astronomy* 3(1): 27–34. <http://www.nature.com/articles/s41550-018-0602-1>.
- Curiel, E. (2019b). Singularities and black holes. In *The Stanford encyclopedia of philosophy*, E.N. Zalta, ed. Spring 2019 edn. Metaphysics Research Lab, Stanford University.
- Dardashtii, R., K.P.Y. Thébault, and E. Winsberg. 2017. Confirmation via analogue simulation: What dumb holes could tell us about gravity. *The British Journal for the Philosophy of Science* 68: 55–89. <https://doi.org/10.1093/bjps/axv010>
- Dardashtii, R., S. Hartmann, Thébault, K. P.Y. & E. Winsberg. 2019. Hawking radiation and analogue experiments: A Bayesian analysis. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 67: 1–11.
- de Nova, J.R.M., K. Golubkov, V.I. Kolobov, and J. Steinhauer. 2019. Observation of thermal Hawking radiation and its temperature in an analogue black hole. *Nature* 569: 688–691.
- Dolan, B.P. 2012. Where is the PdV in the first law of black hole thermodynamics? In *Open questions in cosmology*, G.J. Olmo, ed. Rijeka: IntechOpen, chapter 12.
- Dougherty, J., and C. Callender. (2016). Black hole thermodynamics: More than an analogy? In *Guide to philosophy of cosmology*, B. Loewer, ed. Oxford University Press.
- Drori, J., Y. Rosenberg, D. Bermudez, Y. Silberberg, and U. Leonhardt. 2019. Observation of stimulated Hawking radiation in an optical analogue. *Physical Review Letters* 122: 010404.
- Evans, P.W., and K.P.Y. Thébault. 2020. On the limits of experimental knowledge. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 378(2177): 20190235.
- Field, G. 2022. Putting theory in its place: The relationship between universality arguments and empirical constraints. *The British Journal for the Philosophy of Science* . <https://doi.org/10.1086/718276>
- Fletcher, S.C. 2014. Similarity and spacetime: Studies in intertheoretic reduction and physical significance, PhD thesis, UC Irvine.
- Fletcher, S.C. 2016. Similarity, topology, and physical significance in relativity theory. *British Journal for the Philosophy of Science* 67(2): 365–389.
- Giddings, S.B. 2016. Hawking radiation, the Stefan-Boltzmann law, and unitarization. *Physics Letters B* 754: 39–42.
- Gryb, S., P. Palacios, and K. Thébault. 2020. On the universality of Hawking radiation. *The British Journal for the Philosophy of Science* 72(3): 809–837.
- Hawking, S.W. 1974. Black hole explosions? *Nature* 248: 30–31.
- Hawking, S.W. 1975. Particle creation by black holes. *Communications in Mathematical Physics* 43: 199–220.
- Hawking, S.W., and G.F.R. Ellis. 1973. *The large scale structure of space-time*. Cambridge: Cambridge University Press.
- Hayward, S.A. 1994. General laws of black-hole dynamics. *Physical Review D* 49(12): 6467.
- Hesse, M.B. 1966. *Models and analogies in science*. Notre Dame: University of Notre Dame Press.
- Jacobson, T. 1996. Introductory lectures on black hole thermodynamics. *Institute for Theoretical Physics University of Utrecht*.

- Jacobson, T. 2005. Introduction to quantum fields in curved spacetime and the Hawking effect. In *Lectures on quantum gravity*, A. Gomberoff, and D. Marolf, eds., 39–89. New York: Springer, chapter 2.
- Keynes, J.M. 1921. *A treatise on probability*. London: Macmillan.
- Kolobov, V.I., K. Golubkov, J.R.M. de Nova, and J. Steinhauer. 2021. Observation of stationary spontaneous Hawking radiation and the time evolution of an analogue black hole. *Nature Physics* 17: 362–367.
- Leonhardt, U. 2018. Questioning the recent observation of quantum Hawking radiation. *Annalen der Physik* 530: 1700114.
- Lüst, D., and W. Vleeshouwers. 2019. *Black hole information and thermodynamics*. Springer Briefs in Physics. Springer.
- Nielsen, A.B. 2009. Black holes and black hole thermodynamics without event horizons. *General Relativity and Gravitation* 41: 1539–1584.
- Norton, J.D. 2021. *The material theory of induction*, BSPS Open.
- Penrose, R. 1965. Gravitational collapse and space-time singularities. *Physical Review Letters* 14: 57.
- Polchinski, J. 1995. String theory and black hole complementarity. arXiv preprint hep-th/9507094.
- Prunkl, C. E.A., and C.G. Timpson. 2019. Black hole entropy is thermodynamic entropy. arXiv:1903.06276 [physics.hist-ph].
- Rees, M.J. 1998. Astrophysical evidence for black holes. In *Black holes and relativistic stars*, R.M. Wald, ed., 79–102. Chicago: The University of Chicago Press, chapter 4.
- Smoot, G.F., C.L. Bennett, A. Kogut, E.L. Wright, J. Aymon, N.W. Boggess, E.S. Cheng, G. De Amici, S. Gulkis, M.G. Hauser, G. Hinshaw, P.D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S.S. Meyer, S.H. Moseley, T. Murdock, L. Rokke, R.F. Silverberg, L. Tenorio, R. Weiss, and D.T. Wilkinson. 1992. Structure in the COBE differential microwave radiometer first-year maps. *The Astrophysical Journal* 396: L1–L5.
- Steinhauer, J. 2016. Observation of quantum Hawking radiation and its entanglement in an analogue black hole. *Nature Physics* 12: 959–965.
- The EHT Collaboration et al. 2019. First M87 event horizon telescope results. IV. Imaging the central supermassive black hole. *The Astrophysical Journal Letters* 875: 4. <https://iopscience.iop.org/article/10.3847/2041-8213/ab0e85>
- Thébaud, K. 2019. What can we learn from analogue experiments? In *Why trust a theory? Epistemology of fundamental physics*, R. Dardashti, R. Dawid, and K. Thébaud, eds., 184–201. Cambridge University Press, Cambridge, chapter 11.
- Unruh, W.G. 1977. Origin of the particles in black hole evaporation. *Physical Review D* 15: 365.
- Unruh, W.G. 1981. Experimental black-hole evaporation? *Physical Review Letters* 46 (21): 1351–1353.
- Unruh, W.G. 2014. Has Hawking radiation been measured? *Foundations of Physics* 44: 532–545.
- Visser, M. 2003. Essential and inessential features of Hawking radiation. *International Journal of Modern Physics D* 12: 649–661.
- Visser, M. 2013. Survey of analogue spacetimes. In *Analogue gravity phenomenology: Analogue spacetimes and horizons, from theory to experiment*. Lecture Notes in Physics, D. Faccio, F. Belgiorno, S. Cacciatori, V. Gorini, S. Liberati, and U. Moschella, eds., vol. 870, 31–50. Springer, chapter 2.
- Wald, R.M. 1984. *General relativity*. Chicago: The University of Chicago Press.
- Wald, R.M. 1994. *Quantum field theory in curved spacetime and black hole thermodynamics*. Chicago: University of Chicago Press.
- Wallace, D. 2018. The case for black hole thermodynamics, Part I: Phenomenological thermodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 64: 52–67.
- Wallace, D. 2019. The case for black hole thermodynamics, Part II: Statistical mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 66: 103–117.

- Weinfurter, S., E.W. Tedford, M.C.J. Penrice, W.G. Unruh, and G.A. Lawrence. 2011. Measurement of stimulated Hawking emission in an analogue black hole system. *Physical Review Letters* 106: 021302.
- Wolchover, N. 2016. What sonic black holes say about real ones. *Quanta Magazine*. <https://www.quantamagazine.org/what-sonic-black-holes-say-about-real-ones-20161108/>.
- Wüthrich, C. 2019. Are black holes about information? In *Why trust a theory? Epistemology of fundamental physics*, R. Dardashti, R. Dawid, and K. Thébault, eds., 202–223. Cambridge: Cambridge University Press, chapter 11.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

