

On the Universality of Hawking Radiation

Sean Gryb^{*1,2}, Patricia Palacios^{†3}, and Karim Thébault^{‡4}

¹*Department of Philosophy, University of Groningen*

²*Department of Physics, University of Groningen*

³*Department of Philosophy, University of Salzburg*

⁴*Department of Philosophy, University of Bristol*

April 11, 2019

Abstract

A physically consistent semi-classical treatment of black holes requires universality arguments to deal with the ‘trans-Planckian’ problem where quantum spacetime effects appear to be amplified such that they undermine the entire semi-classical modelling framework. We evaluate three families of such arguments in comparison with Wilsonian renormalization group universality arguments found in the context of condensed matter physics. Our analysis is framed by the crucial distinction between robustness and universality. Particular emphasis is placed on the quality whereby the various arguments are underpinned by ‘integrated’ notions of robustness and universality. Whereas the principal strength of Wilsonian universality arguments can be understood in terms of the presence of such integration, the principal weakness of all three universality arguments for Hawking radiation is its absence.

*email: sean.gryb@gmail.com

†email: patricia.palacios@sbg.ac.at

‡email: karim.thebault@bristol.ac.uk

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Hawking Radiation and the Trans-Planckian Problem | 4 |
| 2.1 | What is Hawking Radiation? | 4 |
| 2.2 | Red-Shift and Robustness | 5 |
| 2.3 | Ultraviolet Catastrophe | 7 |
| 3 | Universality Arguments in Condensed Matter Systems | 10 |
| 3.1 | The Wilsonian Approach to Critical Phenomena | 10 |
| 3.2 | The Six Qualities | 12 |
| 4 | Universality Arguments for Hawking Radiation | 15 |
| 4.1 | Arguments from the Equivalence Principle | 15 |
| 4.2 | Arguments from Horizon Symmetries | 17 |
| 4.3 | Arguments from Modified Dispersion Relations | 19 |
| 5 | Conclusion | 22 |

1 Introduction

Universality arguments have been developed in the context of black holes physics to resolve a problem in Hawking’s famous prediction that black holes produce thermal radiation ([Hawking 1975](#)). As was recognised soon after Hawking’s original paper ([Gibbons 1977](#)), the derivation of Hawking radiation makes essential use of a breakdown in the separation between micro- and macro-scales. In particular, due to an exponential redshift, the Hawking radiation that is detected far away from the black hole, at relatively large wavelengths and late times, is sensitive to the near horizon physics at ultra-short wavelengths — even much smaller than the Planck scale of 10^{-35} m. This ‘trans-Planckian problem’ is not just a curious side-note. Rather it implies that quantum spacetime effects can be amplified such that they undermine the semi-classical framework for modelling black holes.

The strongest lines of response to the trans-Planckian problem that are available in the literature are all based upon ‘universality’ arguments. What each of these sets of arguments have in common is that they aim to establish that the phenomenon of Hawking radiation is both ‘universal’ and ‘robust’. Universality in this context is insensitivity to variation of the macroscopic details that characterise the *type* of black hole considered (e.g. stationary vs. non-stationary spacetime). Robustness in this context is insensitivity to variation of the microphysical details that characterise the particular *token* of black hole considered (e.g. different ultraviolet physics

at the horizon). In this paper we consider three sets of arguments designed to establish the universality and robustness of Hawking radiation. These are based upon respectively: i) the Unruh effect and equivalence principle (Agullo et al. 2009; Wallace 2017, §4.2); ii) horizon symmetries (Birmingham et al. 2001; Banerjee and Kulkarni 2008; Iso et al. 2006b); and iii) modified dispersion relations (Unruh and Schützhold 2005; Himemoto and Tanaka 2000; Barcelo et al. 2009).

The structure of the highly successful ‘Wilsonian’ universality arguments used in condensed matter physics is the starting point for our present analysis.¹ Our goal is to use these arguments to establish a general framework within which other universality arguments, including those for Hawking radiation, can be evaluated. Our framework is based upon six qualities that characterize Wilsonian universality arguments:

1. *Degree of Robustness.* The range of single-type token-level variation across which the invariance of the relevant phenomena can be established;
2. *Physical Plausibility.* The applicability of the robustness and universality arguments to de-idealised target systems;
3. *Degree of Universality.* The range of inter-type variation across which the invariance of the relevant phenomena can be established;
4. *Comprehensiveness.* The size of the set of observables over which the robustness argument can be applied;
5. *Empirical Support.* Experimental evidence supporting instantiation of effects in different types and/or relevant methods of approximation;
6. *Integration.* Feature whereby the theoretical basis behind the invariance found in the universality arguments and the robustness arguments are mutually consistent.

It will be argued that whereas Wilsonian universality arguments have a high score in all these six qualities, universality arguments for Hawking radiation fail to measure up to their Wilsonian counterparts. In particular, it will be emphasized the strength of the Wilsonian arguments is predicated upon their integration: they can combine such a high degree of robustness and universality because the theoretical basis behind the invariance are mutually consistent. Furthermore, it is this combination that allows the arguments to also be highly physically plausible (c.f. Batterman (2000), Batterman (2002), Batterman and Rice (2014)). The converse is found to hold in the Hawking case: the universality arguments that do well on degree of robustness do poorly on degree of universality and physical plausibility, and arguments that do well on degree of universality do poorly on degree of robustness. We argue that the lack of

¹For philosophical discussion see (Batterman 2000; Batterman 2002; Mainwood 2006; Butterfield 2011; Ruetsche 2011; Franklin 2017; Palacios 2019; Fraser 2018; Shech 2018; Saatsi and Reutlinger 2018).

integration found in these arguments is a plausible reason for the existence of these trade-offs between the relevant subset of qualities. This negative conclusion provides a possible basis for future scientific research aimed at developing integrated universality arguments for Hawking radiation.

2 Hawking Radiation and the Trans-Planckian Problem

2.1 What is Hawking Radiation?

In quantum field theory on curved spacetime the ground state of a free quantum field can evolve into a state with excitations when the geometry of the background spacetime is non-trivial. In an asymptotically flat spacetime, this means that there can be a flux of ‘out-particles’ at future-null infinity \mathcal{I}^+ even if there is no flux of ‘in-particles’ at past-null infinity \mathcal{I}^- (Jacobson 2005, §5). Most famously when the spacetime in question features a black hole, this flux can have a characteristic thermal form that depends only upon intrinsic properties of the black hole spacetime (Kay and Wald 1991). This thermalised flux is the Hawking radiation of a black hole.

The original treatment of Hawking (1975) proceeds as follows. Model an astrophysical black hole spacetime as the formation of a Schwarzschild black hole from the collapse of a matter shell.² The conformal diagram for the spacetime exterior to the collapsing matter is given in Fig 1. Within the spacetime assume that there is a free quantized Klein-Gordon field $\hat{\phi}$. Further assume that there is no back-reaction between the quantum field and the background spacetime. The asymptotic flatness of the spacetime before and after the collapse selects unique ground state vacua and Fock representations for the Klein-Gordon field on \mathcal{I}^\pm . Assuming that there is no incoming radiation from \mathcal{I}^- , Hawking used a frame-independent procedure to show that the evolution of such a state would appear, to first order, as a thermal state at the *Hawking temperature* $T_H = \frac{\hbar\kappa}{2\pi}$ for an observer near \mathcal{I}^+ at asymptotically late times – i.e., near future time-like infinity i^+ – where κ is the surface gravity of the black hole horizon.

Various proposals have been made to provide a local physical mechanism for the production of Hawking radiation. The different proposals vary significantly in terms of where and how the thermal radiation is produced and are largely mutually inconsistent. The most significant possible mechanisms include: splitting of entangled modes as the horizon forms (Unruh 1977; Gibbons 1977); tidal forces pulling apart virtual particle-anti-particle pairs (Hawking and Israel 1979; Adler et al. 2001; Dey et al. 2017); entangled radiation quantum tunnelling through the horizon (Parikh and Wilczek 2000); the effects of non-stationarity of the background metric field (Fredenhagen and Haag 1990; Jacobson 2005) and anomaly cancellation (Banerjee and Kulkarni 2008). The formal rigour of these proposals varies greatly, and none is entirely satisfactory from a physical perspective.

²See (DeWitt 1975) for the Kerr case.

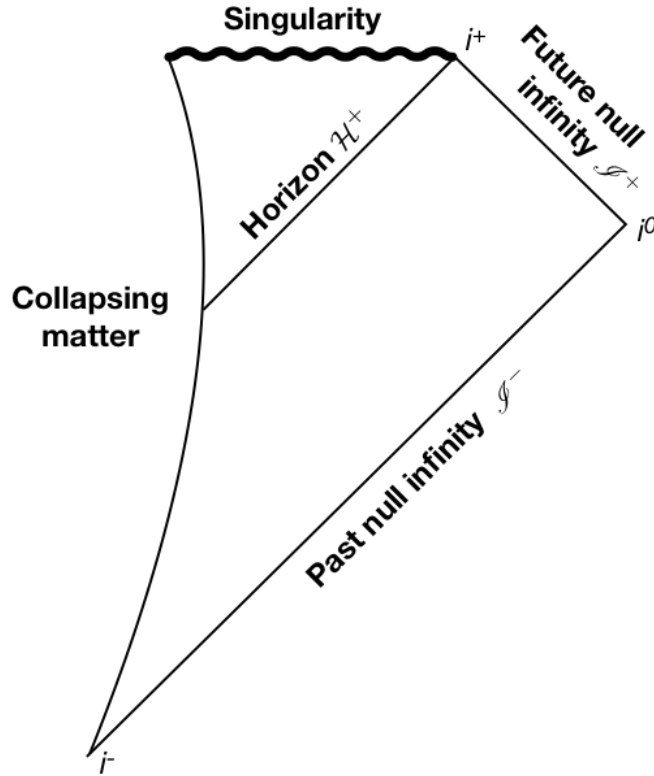


Figure 1: The conformal diagram of a spacetime external to a spherically symmetric distribution of collapsing matter.

Ideally, what we want to find is a relationship between the way in which the gravitational degrees of freedom interact with $\hat{\phi}$ during the evaporation process and the production of the radiation itself. An explicit such demonstration is currently lacking for most, if not all, of the putative mechanisms (Curiel 2018). At the very least, to be physically plausible, a mechanism for Hawking radiation must be explicitly demonstrated to be generalisable from the highly idealised cases of eternal black holes to physically realistic astrophysical models in which time translation symmetry is broken by the collapse phase leading up to the formation of the horizon (Hollands and Wald 2015).

Notwithstanding this lack of unique physically plausible mechanism or region of origin associated with Hawking radiation, it is certainly significant that the formal expression for Hawking flux has proved ‘remarkably robust’ under the inclusion of various complicating factors (Leonhardt and Philbin 2008; Thompson and Ford 2008) and formal clarifications (Freidenhagen and Haag 1990). As noted by Wallace (2017), such consistency between various theoretical derivations strongly suggests that Hawking radiation really is a consequence of semi-classical gravity and not simply an artefact of a particular (potentially flawed) argument. That said, robustness of an effect between different theoretical models offers little comfort if there is a problematic feature that all the various derivations share.

2.2 Red-Shift and Robustness

An important insight into the physics behind Hawking radiation, which is independent of the mechanism of production, can be obtained by considering the possible equilibrium states of the quantum field long after the collapse process has taken place. By this time we can assume that the geometry will have effectively reached Schwarzschild form (assuming zero angular momentum). Since this geometry is not maximally symmetric, there is no unique ground state singled-out by the global symmetries. A variety of physically motivated principles can be used to resolve this underdetermination. These principles lead to different choices of privileged observers, who are required to see the quantum field in a ground state. One choice, the *Boulware vacuum* (Boulware 1975), is a ground state as seen by the static observers of the exterior Schwarzschild geometry. In this state no fluxes are observed and there is thus no Hawking radiation. The Boulware vacuum is irregular on the horizon and thus unphysical in this context. Two vacua that are regular on the horizon are the *Hartle–Hawking vacuum* (Hartle and Hawking 1976) and the *Unruh vacuum* (Unruh 1976; Dappiaggi, Moretti, Pinamonti, et al. 2011).

The Hartle–Hawking vacuum is a ground state as seen by in- and out-going null observers of the entire maximally-extended Schwarzschild geometry. When restricted to the exterior geometry, this state appears as a thermal state near future time-like infinity i^+ for modes out-going from near the horizon (Kay and Wald 1991). Under similar conditions, the Unruh vacuum also appears as a thermal state on i^+ but differs from the Hartle–Hawking vacuum away from i^+ (Dappiaggi et al. 2011). The Unruh vacuum is defined to be a ground state for null out-going observers on the past horizon \mathcal{H}^- of maximally-extended Schwarzschild and for static observers near \mathcal{I}^- . It is better physically motivated than the Hartle–Hawking vacuum as a description of the vacuum state of a collapsed black hole since any state evolving on a background that becomes Schwarzschild at late times will approach the Unruh vacuum near i^+ , provided that the state satisfies certain regularity conditions in the ultraviolet limit and has no in-coming radiation from near spatial infinity i^0 (Fredenhagen and Haag 1990; Hollands and Wald 2015). In this sense, the Hawking spectrum is seen to be insensitive to the details of the collapse process and to the initial state of $\hat{\phi}$ on \mathcal{I}^- (but away from i^0). The main constraint on the state can be formulated, for example, in terms of the Hadamard condition (Hollands and Wald 2015) or via a scaling limit in the ultraviolet (Fredenhagen and Haag 1990). These constraints amount to regularity conditions in the ultraviolet limit and can be understood intuitively as enforcing high-energy modes to be effectively in a ground state. This avoids non-local divergences in the 2-point functions (Wald 1994; Hollands and Wald 2015). Furthermore, in conventional effective field theory treatments (Candelas 1980), which assume no unusual effects in the UV due to quantum-gravity, violation of the regularity conditions can be shown to lead to pathological behaviour near the horizon.

To better understand the role of the regularity conditions in establishing the robustness of

the thermal spectrum we can analyse the relative frequency shift between the Killing frequency and affine frequency of a particular Hawking mode as seen respectively by static and null-geodesic observers (Jacobson 1996). The Killing frequency, which is constant along geodesics, is particularly useful for describing the frequency of free-falling modes of $\hat{\phi}$ as they approach future time-like infinity i^+ , where the static observers become inertial. The affine frequency is suitable for describing the frequency of free-falling $\hat{\phi}$ modes near the horizon, where the affine coordinates describe a local patch of Minkowski space. A key formal property of the Schwarzschild geometry, which can be generalized to arbitrary Killing horizons (Kay and Wald 1991), is that the affine frequency is exponentially related to the Killing frequency near the horizon. Given the properties outlined above, this fact implies that free-falling modes near the horizon will be red-shifted exponentially as they approach i^+ . Furthermore, if the Killing modes of $\hat{\phi}$ are expanded in terms of in-going and out-going affine components, then as the horizon is approached more and more of the in-going affine modes will disappear behind the horizon, which provides a causal barrier to their escape. Thus a static detector approaching i^+ will observe both an exponentially red-shifted spectrum and one entirely dominated by the out-going components. This means that the extreme red-shift combined with the presence of the horizon erases all information about how the horizon itself was formed and the details of the initial state of the radiative mode. The extreme red-shift further implies that in studying a moderate frequency mode at late times a static observer is effectively probing the ultraviolet structure of the 2-point functions of the mode near the horizon.

It is in this context that the regularity conditions imposed on the 2-point function in the ultraviolet become vitally important. As noted above, the conditions require that the modes observed near i^+ be in their ground state near the horizon. Locally, this ground state is approximately the vacuum of Minkowski space split by the horizon into right and left Rindler-like regions. Tracing out the in-going modes (which disappear behind the horizon) therefore leaves the characteristically thermal state expected of the Rindler vacuum.

The robustness of the thermal spectrum detected by late-time observers thus results from three ingredients:

1. *Exponential red-shift* (which implies that late-time observers probe ultra-high frequency modes near the horizon).
2. *Regularity conditions* (which enforce that a state in the ultraviolet is approximately a Minkowski vacuum state).
3. *Event horizon* (which creates a causal barrier between the in-going modes and the observers confined to the exterior region and therefore modifies the character of the state of the UV modes from ‘Minkowski-like’ to ‘Rindler-like’).

2.3 Ultraviolet Catastrophe

With some irony, the mechanisms that are responsible for the ‘remarkable robustness’ of Hawking radiation are simultaneously a cause for scepticism. As we have just seen, the exponential red-shift implies that the spectrum of radiation at late times is dominated by the characteristics of the state of the field near the horizon at energies that well exceed the Planck scale. In fact, *any* modes measured as *cis*-Planckian by stationary observers near i^+ , must have originated as *trans*-Planckian modes from the point of view of free-falling observers less than a Planck unit of proper time before falling through the horizon. The Hawking radiation incident on a finite, stationary detector far away from the black hole can therefore be traced back to what are, for free-falling observers, trans-Planckian energies at the horizon. The above formulation is due to (Helfer 2003) and demonstrates that this *trans-Planckian problem* can be stated in an observer-independent way and thus cannot be ameliorated by redescription in a more fortunate coordinate system.³ The energy of the near-horizon modes in question is in a regime where quantum gravity effects, such as those due to entanglement with the horizon, would be expected to be relevant. Thus, the implication of the trans-Planckian problem is that the semi-classical framework is not straightforwardly valid when applied to any effect, such as Hawking radiation, that is taken to be sensitive to near horizon physics.⁴

As was aptly put by Jacobson (2005), the trans-Planckian problem amounts to ‘a breakdown in the usual separation of scales invoked in the application of effective field theory’ (p.79).⁵ In response to this breakdown, and in particular the role played by back-reaction and the evaporation process, Fredenhagen and Haag (1990) conclude that ‘a full understanding of the [Hawking] phenomenon including a self-consistent description of the causal structure needs some elements of a quantum theory of gravity’ (pp.282-3). As noted by Unruh (2014), in typically robust fashion, ‘if one examines Hawking’s original calculation, there are some severe problems with his derivation. While mathematically unimpeachable, they are nonsense physically’ (p.533). To highlight the significance of the scales involved Unruh estimates that the ‘frequencies which are needed to explain the radiation produced even one second after a solar mass black hole forms, correspond to energies which are e^{10^5} times the energy of the whole universe.’ (Unruh 2014, p.533). Plausibly, such considerations throw into doubt the idealisations and approximations upon which the entire semi-classical model of black holes is built.

An influential attempt to resolve the trans-Planckian problem is the ‘nice slice’ argument due to Polchinski (1995) (c.f. Wallace (2017)). Polchinski defines a particular slicing of a black hole spacetime that is ‘nice’ in the sense that the slices are smooth, have small extrinsic

³We will return to this point explicitly in the context of Polchinski’s ‘nice-slice’ argument shortly.

⁴The key historical papers on the trans-Planckian problem are (Gibbons 1977; Unruh 1981; Jacobson 1991; Jacobson 1993; Unruh 1995; Brout, Massar, Parentani, and Spindel 1995). Accessible introductions are give in (Jacobson 2005, §7) and (Harlow 2016, pp.36-8). For discussion of the trans-Planckian problem in cosmology see (Martin and Brandenberger 2001).

⁵For a more general argument that effective field theory methods may breakdown near horizons see (Burgess et al. 2018).

curvature, and are such that in-falling particles are seen to have modest velocities. Due to the small extrinsic curvature the geometry of the slices changes slowly from slice to slice. Polchinski then argues that the adiabatic theorem should apply to modes in this slice and therefore that only very low-energy degrees of freedom can be excited from their ground state in the Hawking emission process (by whatever mechanism it takes place). According to this argument, the entire process can be described using local low-energy physics near the horizon and will therefore be independent of Planck scale effects.

The ‘nice slice’ argument is unconvincing as a response to the trans-Planckian problem as we have formulated it. In particular, the assumption that the relevant modes are in a genuine ground state is precisely the assumption that is in question. As emphasised by Jacobson in the quote above, the trans-Planckian problem occurs precisely because we have good reasons to expect the separation of energy scales to breakdown. Thus, whether the adiabatic theorem, or local arguments from effective field theory in general, should apply to near-horizon modes, which may be entangled or even interacting with the geometric degrees of freedom of the horizon, is exactly the issue at hand. Plausibly, some form of non-linear gravity-matter interaction is required for the evaporation process to take place at all. Furthermore, as noted by [Harlow \(2016\)](#), the adiabatic theorem applies only to the global conserved energy not to the centre of mass energy of localized excitations. Finally, the nice-slice argument does nothing to ameliorate the exponential redshift: ‘it does not get rid of the fact that projecting onto possible final states of the late-time Hawking radiation produces states with a genuine high energy collision in the past’ ([Harlow 2016](#), p.37-38). Thus, *pace* [Wallace \(2017\)](#), Polchinski’s argument offers no definitive means to rebut the force of the trans-Planckian problem.⁶

A different response to the trans-Planckian problem is based upon the inconsistency of Hawking radiation *not* existing. The argument is that the physics of the Planck scale cannot alter the Hawking spectrum since to do so would violate the semi-classical field equations ([Candelas 1980](#); [Sciama et al. 1981](#)). However, once more, such a line of response is based upon an assumption that is itself in question. We might reasonably assume that due to the relatively mild curvature, the semi-classical field equations must apply to any local observers description of vacuum fluctuations near the horizon of an astrophysical black hole. However, in order for us to extend this assumption to near-horizon fluctuations as seen by a late-time observer, due to the exponential red-shift, we must assume that the semi-classical field equations continue to hold up to and beyond the Planck scale. However, it is at precisely at the Planck scale that we expect violations of the semi-classical field equations to occur.

What we take to be the essential lesson is that absent a well-trusted theory of quantum gravity, any derivation of Hawking radiation as a phenomena that depends on near-horizon physics must be supplemented with an argument for the insensitivity of the effect to ultra short-distance

⁶The possibility for quantum gravity effects to undermine the nice slice argument is noted in ([Polchinski 1995](#), §2) and in subsequent discussions regarding firewalls ([Almheiri, Marolf, Polchinski, and Sully 2013](#)).

physics.⁷ This supplement may take the form of an additional argument or a modification of the derivation. In either case, the insensitivity is required to both token-level and type-level. The first, since we need to establish the insensitivity of the effect to different possible ultraviolet physics for a given type of black hole. The second, since to justify the idealisations (such as no-backreaction) which the trans-Planckian problem gives us *prima facie* cause to doubt, we need to establish the insensitivity of the effect to the type of black hole being considered (e.g. astrophysical vs eternal).⁸

In the terminology introduced above what we are looking for is an argument for both the *robustness* and *universality* of Hawking radiation. It would be unreasonably restrictive to take the precise form of the radiative spectrum to be insensitive to such token-level and type-level modifications. Rather, what we take to be the necessary, and arguably defining, property of the Hawking effect is the *thermality* of the detected radiation (at least to first order in \hbar). A token-level or type-level variation that destroys the thermality whilst preserving the particle flux is then a variation that undermines the Hawking effect (not least since it would allow the radiation to carry information).

3 Universality Arguments in Condensed Matter Systems

3.1 The Wilsonian Approach to Critical Phenomena

A phase transition occurs when there is an abrupt change in the macroscopic parameters that uniquely specify the equilibrium states of a system. A first-order phase transition is characterized by the existence of discontinuities in the first derivatives of the free energy. Continuous phase transitions, in contrast, involve divergence of the response functions. An important feature of continuous phase transitions is that in the vicinity of the critical point measurable quantities depend upon one another in a power-law fashion. For example, in the ferromagnetic-paramagnetic transition, the net magnetization, M , the magnetic susceptibility, χ , and the specific heat, C , all depend upon the reduced temperature $t = \frac{T-T_c}{T_c}$ (the temperature of the system with respect to the critical temperature T_c) as:

$$M \sim |t|^\beta, C \sim |t|^{-\alpha}, \chi \sim |t|^{-\gamma}, \quad (1)$$

where β , α , γ are the *critical exponents*. Another remarkable feature of these phenomena is the existence of cooperative behaviour at the transition or critical point, which means that the correlations between particles extend to very large distances even if the microscopic interac-

⁷To keep our discussion within reasonable constraints we have chosen to excluded approaches to Hawking radiation that do not feature near horizon sensitivity. See (Parentani 2010; Giddings 2016; Hod 2016; Dey, Liberati, and Pranzetti 2017).

⁸As shall be discussed in Section 4.3, a further degree of type-level insensitivity relates to derivations of Hawking radiation that also apply to condensed matter analogue black holes.

tions between them remain short range. This implies the divergence of the correlation length ξ , which is a quantity that measures, for example, the distance over which spins in are correlated. The divergence of the correlation length is perhaps the most important feature of continuous phase transitions. In particular, it involves the loss of a characteristic scale at the transition point and thus provides a basis for the explanation of universal behaviour. That is, the divergence of the correlation length explains why system types with physically distinct micro-structure, such as ferromagnets, antiferromagnets and fluids, display the same macro-behaviour.

Landau's theory of continuous phase transitions (Landau 1936) was one of the first attempts to give a rigorous explanation for the behaviour of physical variables close to the critical point and anticipated the development of renormalization group approaches. In this theory, physical variables, such as magnetization, are replaced by their average values and non-linear fluctuation contributions are neglected. One can then use Landau theory to estimate the importance of fluctuations close to the critical point. For space dimension $d > 4$, the theory makes adequate predictions of the order parameter and the critical exponents. Unfortunately, for $d \leq 4$, the Landau's theory of continuous phase transitions predicts strong infrared singularities in the lowest order fluctuation contributions, which means that fluctuations dominate the behaviour close to the critical point (Goldenfeld 1992, §6).

In analogy with the trans-Planckian problem, the problem of infrared singularities requires us to find a means to screen observable quantities in our theory from a breakdown in the separation of scales. One strategy is to explicitly renormalize the ultraviolet divergences, which involves expressing the weight of fluctuation contributions (amplitudes), in terms of physical coupling constants without assuming any particular cut-off in the calculation. Wilson (Wilson 1971; Wilson and Kogut 1974) suggested a different strategy, now called momentum shell renormalization group (RG), that consists in integrating out short-range wavelength modes up to a finite cut-off in momentum Λ .⁹ In this approach, one starts by defining a field theory in which the degrees of freedom are represented in terms of Fourier modes, $S(q)$. The partition function is then expressed as an integral over the full range of Fourier components. Each field theory, as defined by a particular local Hamiltonian (or Lagrangian), will then be characterized by the set of coupling constants g_i that measure the strength of the various interactions. The core idea of RG approach is to examine how the coupling constants change as one varies the length scale of interest, which is achieved by changing the value of the cut-off Λ . For continuous phase transitions one is interested in the limit of large length scales thus, for such theories, one analyses the behaviour of the coupling constants as the length scale increases.

This RG procedure involves the following three steps: First, one carries out the partition integral over all Fourier components $S(q)$ with wave vectors residing in the momentum shell $\Lambda/b \leq |q| \leq \Lambda$, where $b > 1$. This step effectively eliminates the short-wavelength modes and thus corresponds to a coarse graining. Second, one relabels the control parameters by

⁹Kadanoff's real-space renormalization approach has been neglected for reasons of space. Plausibly our analysis will apply mutatis mutandis. See (Fisher 1998; Goldenfeld 1992; Mainwood 2006; Franklin 2017).

performing a scale transformation: $x \rightarrow x' = x/b$ and $q \rightarrow q' = bq$. Third, one relabels the field degrees of freedom by performing a scale transformation:

$$S(x) \rightarrow S'(x') = b^\zeta S(x), S(q) \rightarrow S'(q') = b^{\zeta-d} S(q), \quad (2)$$

where ζ must be chosen so as to assure that the rescaled residual Hamiltonian recovers the original form. The three steps result in an ‘effective’ Hamiltonian, which has different values of coupling constants than the original Hamiltonian but in successful cases has the same form. The RG transformations, given by repeated application of the three steps, are associated with a flow on theory space that tells us how the coupling constants change. Depending on their behaviour under repeated iterations of the coarse graining transformations, the coupling constants can be classified as:

- i *relevant coupling constants*, which grow and ultimately tend to infinity as the number of iterations tends to infinity. These parameters allow one to define the *critical surface*;
- ii *irrelevant coupling constants*, which ultimately approach zero in the RG procedure and do not affect the critical behaviour;
- iii *marginal coupling constants*, which approach an infrared-stable fixed point that is associated to scale-invariant behaviour.

The disappearance of the irrelevant couplings and the existence of non-trivial infrared fixed points is precisely what resolves the problem of infrared singularities in the RG approach. Crucially for our analysis, it is also this feature that establishes both the robustness and universality of the critical phenomena.

The Wilsonian RG argument for universality takes the following form. Given the disappearance of the irrelevant couplings and the existence of a non-trivial infrared fixed point, the critical behaviour will depend only on the spatial dimension and the symmetries of the original Hamiltonians and not on the strength of the nonlinear couplings or other non-universal parameters. All distinct Hamiltonians whose trajectories converge toward the same infrared fixed point, i.e., the basin of attraction of the fixed point, will then exhibit identical behaviour at the critical point. This establishes universality since the critical phenomena in question have been shown to be insensitive to an inter-type variation between systems described by the set of distinct Hamiltonians (i.e., variation across the relevant universality class). Furthermore, the same arguments also establish robustness. This is because the critical phenomena in question have been shown to be insensitive to a variation between possible microphysical realisations of a single type of system as described by distinct Hamiltonians. It is thus the very same mathematical properties of the basin of attraction of the fixed point that establish both robustness and universality.

3.2 The Six Qualities

Drawing upon the Wilsonian exemplar we can identify six key qualities of successful universality arguments. First, the degree of robustness, which is the range of single-type token-level variation across which the invariance of the macroscopic phenomena can be established. A Wilsonian treatment of RG transformations comes with a well-defined set of restrictions regarding the type of possible micro-interactions that can be shown to be irrelevant, and these limitations restrict the degree of robustness that the arguments can establish. One should also keep in mind that the theory space contains only theories within the broad framework of quantum field theory. When we are dealing with discrete systems, this assumption can obscure the connection to the microphysics of the systems. Further restrictions are then imposed within the family of field theories. In particular, the interactions must be short range (Wilson and Kogut 1974, p. 161) and expressible in terms of a convergent set of constants and differential operators. Finally, in constructing theory space one is required to make assumptions about the number of spacetime dimensions and the irrelevance of, as yet unknown, fundamental spacetime structure. Not least in assuming a smooth $3 + 1$ spacetime model we are implicitly also assuming that we can rule out macro-level effects that have their origin in an unknown number of compactified extra dimensions or even fundamental dimensional heterogeneity (Täuber 2012) due for example to quantum modifications to the spectral dimension of spacetime. That said, we have good reason to expect that the ‘separation of scales’ should mean that even if spacetime is fundamentally dimensionally heterogeneous or matter is not described by a quantum field theory, these assumptions will not undermine the effectiveness of the RG approach. These limits on the degree of robustness are thus very mild when considered in the relevant context and we can conclude that the Wilsonian universality arguments have a high degree of robustness.

The second quality of Wilsonian arguments regards the degree of universality. This is the range of inter-type variation across which the invariance of the phenomena can be established. This is one of the core virtues of Wilsonian renormalization group methods. In particular, the Wilsonian approach provides a means to characterise both quantitative and qualitative aspects of the ‘universal phenomena’ across a wide range of system types such as fluids, ferromagnets, and antiferromagnets (Goldenfeld 1992; Täuber 2012).

The third quality is physical plausibility, which is closely connected to the first two. Physical plausibility is the applicability of the robustness and universality arguments to de-idealised target systems. The Wilsonian framework does particularly well on this front. In applying a renormalization group transformation, one coarse-grains out physically irrelevant detailed information until one reaches a non-trivial fixed point, which implies that the phenomenon under investigation is robust with regard to different possible microphysics *and* different possible complicating factors, both micro and macro (Batterman and Rice 2014; Palacios 2019). Moreover, it is the same arguments for robustness (insensitivity of single-type token level vari-

ation) that explain why scientific idealised models can be used to describe the behaviour of real target systems. In fact, one can use renormalization group arguments to demonstrate, for example, that real fluids and ferromagnets are in the same universality class as the idealised 2-dimensional Ising model, so that the details that distinguish idealised models from real systems do not matter. As it will be seen below, it is in virtue of integration that one obtains a high degree of physical plausibility.

The fourth quality is comprehensiveness. This is the size of the set of observables over which the robustness and universality argument can be applied. At criticality the arguments apply to all relevant observables and so this quality is again very strong. Although this quality is logically independent of the degree of robustness and universality, a high degree of comprehensiveness is required to have universality arguments that can successfully account for a rich class of phenomena. We will see in the next section that limiting the size of the set of observables over which universality arguments apply also limits the strength of these arguments.

The fifth quality that characterises Wilsonian universality arguments is empirical support. In fact, it is possible to use this framework to calculate explicit values of the critical exponents by linearizing around the fixed points and these values are found to be in good agreement with experimental results – see for example, (Ahlers, Kornblit, and Guggenheim 1975). Thus, the arguments are directly supported by experimental evidence in the relevant range of physically instantiated types. Furthermore, the methods of approximation that underlie the Wilsonian approach are also justified by the fact that they successfully reproduce empirically observed phenomena.¹⁰

The sixth quality, crucial to our analysis, is integration. Integration is the quality whereby the theoretical basis behind the invariance found in the universality arguments and the robustness arguments is mutually consistent. Wilsonian arguments are clearly integrated since, as noted above, identical mathematical properties of the basin of attraction of the fixed point establish robustness and universality. Essentially the same arguments that successfully establish the irrelevance of the micro-details of a given system also allow us to explain the irrelevance of the details that distinguish different systems within the same universality class (c.f. (Batterman 2002; Batterman and Rice 2014)). Wilsonian universality arguments are constituted by the combination of the robustness argument together with general commonality conditions between the types based upon the spatial dimension and the symmetries of the original Hamiltonian. This means that a high degree of robustness will automatically imply a high degree of universality. Furthermore, arguments of this structure will also automatically be highly physically plausible. This is because we can use insensitivity under token-level variations between different possible micro-physics to show insensitivity under i) de-idealisation, which means that the same predictions will also hold for de-idealised realistic models and ii) re-interpretation, which means that we can re-interpret the system type by adding details that characterise par-

¹⁰See (Blum and Joas 2016) for analysis of a historical case study focusing on the interplay between experimental evidence, approximations techniques and ‘emergent entities’ in the context of RG techniques.

ticular models.¹¹ Batterman and Rice (2014) make a similar point, arguing that in the case of universality arguments it is the story that explains why microscopic details are irrelevant for the occurrence of the phenomenon that allows us to understand why idealised minimal models can be used to explain the behaviour of real systems. Batterman and Rice (2014) provide a vivid example in terms of Fisher's model in the explanation of the 1:1 sex ratio in diverse biological systems: there, it is the demonstration that a large class of details of particular systems (e.g. the actual population size) are irrelevant for the 1:1 sex ratio (i.e. robustness) that allows us to delimit the universality class of systems that will display the same behaviour. The delimitation of the universality class then allows us to see which features these heterogeneous systems have in common and, moreover, to understand how such idealised model can be used to explain the 1:1 ratio in real biological populations.

In general terms, integration is clearly a highly desirable feature since it means that the token-level insensitivity established by the robustness arguments is extendable to a wide range of types. When these types include physically realistic systems such integration can be expected to then also lead to physical plausibility since we have then means to demonstrate the insensitivity of effects to de-idealization at type and token level simultaneously. Conversely, unintegrated universality arguments are undesirable since inconsistency between the robustness and universality arguments will imply both that token-level insensitivity is non-universal (i.e. highly type-specific), and that type-level insensitivity is non-robust (i.e. highly token-specific). Such arguments are highly unlikely to be physically plausible or, moreover, useful in providing a general means of dealing with problems due to a breakdown in the separation of scales. Unfortunately, it is precisely a lack of integration that we find in the context of the three sets of universality arguments for Hawking radiation considered in the next section.

4 Universality Arguments for Hawking Radiation

4.1 Arguments from the Equivalence Principle

Our first example is best understood as an independent universality argument for Hawking radiation, rather than an alternative derivation. It is based upon a heuristic argument that runs as follows. Through the equivalence principle, Einstein taught us that gravitation and acceleration are locally indistinguishable. This means that we should expect to be able to translate the local physics experience by a stationary observer in a region outside the event horizon of a black hole spacetime into equivalent physics experienced by a constantly accelerating Rindler observer in a Minkowski spacetime. A translation scheme based upon the equivalence principle would instruct us to identify the Hawking radiation detected by some stationary observer

¹¹It is important to note that this is not equivalent to the claim that one can find fixed-point solutions in a de-idealised model, which is still matter of controversy in the philosophical literature (Batterman and Rice 2014; Palacios 2019; Saatsi and Reutlinger 2018).

in the black hole spacetime with the Unruh radiation detected by a Rindler observer with the same acceleration. In particular, a simple argument – see e.g. (Wallace 2017, pp.23-4) – can be used to numerically identify the Hawking temperature of a black hole with the Unruh temperature of a near-horizon observer red-shifted to infinity.¹² Next, one shows that the Unruh effect is suitably robust under some class of possible ultraviolet modifications (i.e. quantum gravity effects). Using the translation scheme for near-horizon observers one then infers that the Hawking effect should also be robust to ultraviolet modifications. Such reasoning suggests that we were wrong to ever think of Hawking radiation and the Unruh effect as two separate phenomena, just as it would be wrong to distinguish gravitational and inertial mass. Rather, the equivalence principle argument implies the two effects to be instances of single universal phenomenon of Unruh/Hawking radiation connected to acceleration/gravitation. Although suggestive, and to an extent physically insightful, in this section we will isolate the various senses in which such a line of argument is unreliable. We start by analysing the evidence for the robustness of the Unruh effect.

The most physically salient starting point is the ‘detector approach’ to describing Unruh radiation (Unruh 1976; Unruh 1975; Crispino, Higuchi, and Matsas 2008). In this approach, a particle detector is made to follow a constantly accelerating path through Minkowski spacetime. In the simplest case, a scalar field in its vacuum state is coupled to the detector in such a way that the detector will count any sufficiently localized particle that enters the detector. The quantum field theoretic calculation can be done rigorously. In particular, the thermal spectrum can be obtained as the result of a particular branch cut in the relevant integrals over the divergent part of the 2-point functions. Two important factors therefore determine the thermal form of this spectrum: the Lorentz invariance of the procedure, which controls the periodicity of the domain of the analytically continued integrals, and the precise form of the ultraviolet divergences of the 2-point functions (Agullo et al. 2009). The physical mechanism behind the particle production in the Unruh effect can be understood in terms of the physical force responsible for the detector’s sustained acceleration. The ‘energy reservoir’ for the pair production is then seen to arise from whatever source is producing the energy to maintain this force.

Given these insights, it is possible to investigate whether trans-Planckian modifications due to gravitational physics should have a noticeable effect on the spectrum observed by the detector via its response function. In this context, different kinds of ultraviolet modifications have been considered in the literature and all point to a limited degree of robustness of the thermal Unruh spectrum. In (Agullo et al. 2009) the effect of introducing a Lorentz-invariant ultraviolet cutoff for the scalar field modes on the detector response function is considered. It is found that the thermal Unruh spectrum is insensitive to the introduction of this particular cutoff provided it is Lorentz-invariant. Contrastingly, it is also shown that non-Lorentz invariant cutoffs ruin the thermal properties of the spectrum by introducing a time-dependent damping effect on the

¹²The red-shifting here is important because it regularizes the proper acceleration of a stationary observer on the horizon which is formally divergent.

detection rate of Unruh modes. We thus see that the Unruh effect should not be expected to be robust with respect to effects due to violations of Lorentz invariance in the ultraviolet. In more general terms, it can be explicitly shown via non-perturbative Functional Renormalization Group methods that, even in the case of Unruh radiation, the thermality of the spectrum can be ruined by a variety of quantum gravity effects (Alkofer et al. 2016). Such effects are unlikely to improve in the Hawking case, and could even be compounded by non-linear effects during the evaporation process.

Certain features of Unruh radiation can thus be seen to be robust under certain limitations. On such a basis one might plausibly attempt to construct an argument for the universality of Hawking radiation based upon the equivalence principle along the lines of the heuristic argument presented above. Agullo et al. (2009) do this as follows. First, reason that a small detector near the horizon is locally indistinguishable from a Rindler observer in Minkowski space provided the detector is much smaller than the Schwarzschild radius of the black hole. Next, assert that such a detector will have a response rate identical to the one computed in the Unruh case. Finally, use the robustness of the Unruh effect to reason that the response function of the detector *in the black hole spacetime* will be invariant under arbitrary Lorentz-invariant modifications.

There are good reasons to be sceptical regarding this line of reasoning. As we have seen already, one must be cautious about using a local argument concerning the behaviour of a detector near a horizon to make inferences regarding the behaviour of a late-time detector far away from the horizon. While particle detection is itself a local process, the parameters of the Hawking spectrum, such as its temperature, depend functionally on global properties of the spacetime, such as its ADM mass. This is ultimately because the Hawking spectrum is determined in terms of an integrated effect of the field over its entire history. Moreover, as argued in detail by Helfer (2010), the cis-Planckian Hawking modes detected by a stationary detector near the horizon of a black hole can contribute only a negligible proportion of the total Hawking flux detected at late times. Thus, almost all the late-time Hawking fluxes must stem from thermal modes that are trans-Planckian for the near-horizon inertial observers. It is therefore clear that universality arguments for Unruh radiation are completely ill-suited to provide a response to the trans-Planckian problem. They can only show insensitivity of near horizon cis-Planckian thermal spectrum to unknown trans-Planckian physics. By construction, they are silent regarding the properties of (almost all of) the late-time Hawking radiation.

Equivalence principle based universality arguments thus do badly on the qualities of both degree of robustness and physical plausibility. The idealisation that the entire class of observers can be represented by near-horizon detectors limits the argument in terms of both the token-level insensitivity established and applicability to de-idealised token system. In fact, it can be shown that repeating the above argument using *any* observer other than one infinitesimally close to the horizon leads to the wrong answer (Singleton and Wilburn 2011) for the

predicted radiation near i^+ .¹³ This puts into the question the consistency of the entire approach. The score on degree of universality is slightly better since not only are the arguments applicable to Rindler and Schwarzschild spacetimes, they can also plausibly be extended to various other types of black hole spacetime including those representing astrophysical black holes formed via collapse. Contrastingly comprehensiveness is low since the set of observables over which the robustness argument can be applied is only contains one member. Empirical support is completely lacking. Furthermore, and most problematically, the argument is worryingly unintegrated. The theoretical basis for universality (the equivalence principle) in tension with the theoretical basis for robustness (effective field theory arguments applied to the Unruh effect). Plausibly, it is precisely the lack of integration that renders this an argument with rather limited qualities.

4.2 Arguments from Horizon Symmetries

Our next candidate universality argument is best understood as a new derivation of Hawking radiation as a universal phenomena based upon the symmetries of the black hole event horizon. The argument relies upon the cancellation of anomalies of the effective event horizon symmetries to establish robust properties of the resulting Goldstone bosons. These Goldstone bosons are then interpreted as Hawking fluxes which suggests that Hawking radiation is itself robust.

A good starting point is to observe that in general there exists an expansion of the d'Alembertian (in terms of the radial tortoise coordinate) near a fairly general class of horizons within which the leading order term is conformal and the angular contributions can be integrated out. This means that a Klein-Gordon theory reduces to an effective $1 + 1$ conformal field theory near such a horizon (Birmingham et al. 2001; Carlip 2005). These simplifications allow, under fairly general assumptions, for the near-horizon Klein-Gordon modes to be written as representations of a chiral Virasoro algebra whose quantization is well-known. In $1 + 1$ dimensions, there is a conformal anomaly that can be expressed in terms of the topological invariants of the horizon geometry. This quantum mechanically broken conformal symmetry can be seen to lead to the generation of Goldstone bosons that must be present to cancel the anomaly. It is these Goldstone bosons that are interpreted as Hawking fluxes. Two formal observations support this interpretation. First, a state-counting of the Goldstone bosons generated by such an anomaly can be found to exactly reproduce the Bekenstein–Hawking entropy that one would expect for a black hole (Carlip 2005). Second, if one computes the emission rate of the Goldstone bosons from the horizon, the response rate at infinity reproduces a thermal Hawking spectrum (Banerjee et al. 2010; Banerjee and Kulkarni 2008).¹⁴

Arguably, further justification is needed to identify the generation of Goldstone bosons from a broken conformal symmetry with genuine Hawking fluxes. However, given such an

¹³This is effect is most extreme for static observers near i^+ who see no Unruh radiation at all.

¹⁴See also (Iso, Umetsu, and Wilczek 2006a; Iso, Umetsu, and Wilczek 2006b).

identification, the argument from horizon symmetries shows precisely why we should expect Hawking radiation to have a degree of robustness comparable to the Wilsonian approach. In particular, that the anomalies in the near-horizon field theory are related to topological invariants of the horizon geometry means they can be expected to be invariant under a wide range of possible micro-physics. This is because any ultraviolet modifications that preserve the relevant symmetries without disrupting the topological properties of the horizon will produce the same spectrum of Goldstone bosons. This approach thus allows us to demonstrate invariance of the Hawking spectrum under any ultraviolet modifications that are due to quantum gravity effects that preserve the horizon symmetries without disrupting the topological properties of the horizon. Such assumptions are natural in string theory where conformal symmetry of the $1 + 1$ dimensional string worldsheet plays a vital role in the expected ultraviolet finiteness of the theory. However, the assumption that the horizon is smooth and has a definite position – even when probed by arbitrarily high energy modes – requires further justification in the context of quantum gravity. It is not even clear that the notion of horizon itself, which is a globally defined notion that relies on the existence of a time-like singularity, will survive in full theory of quantum gravity.¹⁵ Moreover, it is certainly not clear that physically realistic horizons, such as the event horizons of astrophysical black holes, can indeed be effectively described using Virasoro methods, since there is no proof that the reduction to $1 + 1$ dimensions can adequately encode the physics of collapse. This notwithstanding, given we have justification of the identification of Goldstone bosons as Hawking fluxes, we can take arguments from Horizon symmetries to imply a fairly wide range of token-level insensitivity of Hawking radiation within the idealised system type that they describe. The arguments are thus moderately robust.

A further advantage of these methods is in terms of their comprehensiveness. That is, the size of the set of observables over which the robustness argument can be applied. In particular, it is known that, for CFTs in $1 + 1$, all observables (i.e., n -point functions) can be computed directly from the symmetry considerations of the Virasoro algebra. This means that the arguments apply to all measurable processes rather than a restricted class of observables. With regard to the quality of comprehensiveness the horizon symmetry approach to black holes is comparable to the Wilsonian approach to condensed matter.

To what extent can we think of these arguments for horizon symmetries as universality arguments? Since only quite general properties of the horizon geometry and the Klein-Gordon operator are required to derive the Hawking flux, one might expect that the arguments will be applicable regardless of the type of physical system in question. Clearly this depends upon the extent to which the relevant features of the horizon geometry and the conformal expansion of the Klein-Gordon operator are instantiated in a wide range of physically plausible models. At present the Goldstone bosons and their entropies have only been computed for the a restricted class of horizons where back-reaction and evaporation effects are ignored. As noted above,

¹⁵See (Dougherty and Callender 2019) and (Wallace 2017) for philosophical discussion of the potential issues with global definitions of black hole horizons.

such models are highly idealised. Most significantly, they assume both that the horizon can be treated as a boundary with fixed location even for the highest of trans-Planckian modes and that the equilibrium state of these modes is independent of the details of the non-stationary collapse process itself. It is an interesting open question whether the numerical coincidence between the Goldstone bosons and Hawking modes can be given a plausible physical basis in non-stationary and evaporating black hole models. However, it is certainly not a matter beyond all conjecture; if Hawking radiation really does causally depend upon the creation of the horizon and the violation of time translation invariance, perhaps due to some non-linear quantum back-reaction between the background quantum geometry and the trans-Planckian modes (or some other effect), then the physical basis behind this approach will be undermined. We thus find that both the physical plausibility and degree of universality of these arguments is severely limited. Empirical support is obviously also lacking.

Finally, we find arguments for the universality of Hawking radiation based upon horizon symmetries are not integrated. Unlike in the Wilsonian approach, the robustness arguments are highly type sensitive and thus inconsistent with relevant inter-type invariance arguments. It is natural to diagnose the weakness with regards to degree of universality and physical plausibility, despite high robustness in terms of a lack of integration. Thus although arguments from horizon symmetries do provide a remarkably robust and comprehensive derivation of Hawking radiation, that establishes the effect as originating from very general features of anomaly cancellation, since they are unintegrated, these arguments do not establish the effect as universal and also there are doubts regarding their physical plausibility.

4.3 Arguments from Modified Dispersion Relations

Our final example of a universality argument for Hawking radiation takes the form of a general strategy for modifying derivations of Hawking radiation such that ultra-short distance effects are factored in. The key idea is that quantum gravity corrections to the Hawking spectrum can be modelled in terms of their effects on the propagation of the scalar field. In particular, the corrections are characterised in terms of a set of possible modifications to the dispersion relation of the high-energy Hawking modes. The late-time flux of Hawking modes is then computed with the modified dispersion relations using a straightforward generalisation of Hawking’s original derivation. Provided the modifications to the dispersion relation satisfy a number of criteria the Hawking spectrum can be shown to be insensitive to the modifications.

The original idea behind the modified dispersion relation approach comes from numerical studies of analogue black holes¹⁶ which indicate that one can use a modified dispersion relation to understand the ‘ultraviolet’ breakdown of continuous fluid models due to atomic effects (Jacobson 1991; Jacobson 1993; Unruh 1995). Various generalisations of this approach

¹⁶See (Unruh 1981; Garay et al. 2000; Philbin et al. 2008; Belgiorno et al. 2010; Unruh and Schützhold 2012; Liberati et al. 2012; Nguyen et al. 2015; Jacquet 2018).

to the gravitational case have now been achieved but, for our purposes, the most physically enlightening will prove to be that of [Unruh and Schützhold \(2005\)](#).¹⁷

The universality argument of [Unruh and Schützhold \(2005\)](#) can be reformulated to make the comparison with Wilsonian universality arguments as follows.¹⁸ First, take the family of modified dispersion relations to be parametrized by a single function, F , on momentum space. This function can be interpreted as representing a (non-Lorentz invariant) 3-momentum-dependent mass term for the Klein-Gordon field. In terms of the function F , the modified Klein-Gordon equation then takes the local form:

$$(\square + F(\hat{p}^2)) \phi = 0, \quad (3)$$

where \square is the Klein-Gordon operator in the exterior Schwarzschild spacetime and \hat{p}^2 is some differential representation of the square of the 3-momentum operator in a first quantized Klein-Gordon theory. In momentum space, F can be taken to be some function of the eigenvalue k^2 of the square of the linear momentum operator satisfying four criteria:

- (i) *Analyticity*: $F(s^2)$ has an analytic continuation in terms of some convergent power series expansion in s , for $s \in \mathbb{C}$.
- (ii) *Vanishing in the infrared*: $F(k^2) \rightarrow 0$ when $k \rightarrow 0$ so that the dispersion relation is unmodified for the low energy modes.¹⁹
- (iii) *Sub-luminal*: $F(k^2) \leq k^2$ so that all modes travel slower than the speed of light.
- (iv) *Scaling limit*: $F(k^2) \rightarrow \tilde{F}_\infty k^2$ for $0 < \tilde{F}_\infty < 1$ when $k \rightarrow \infty$ so that the dimensionless mass term, $\tilde{F}(k^2) = \frac{F(k^2)}{k^2}$, flows to a constant, \tilde{F}_∞ in the ultraviolet.

The relevant limits of k above are defined relative to the Planck scale so that the infrared and ultraviolet limits represent the sub- and trans-Planck limits respectively. It is worth noting that, aside from the sub-luminal assumption, the above criteria are remarkably similar to those found in the Wilsonian renormalization framework. In particular, plausibly we can understand $F(k^2)$ as representing a running coupling in a particularly simple truncation of the Klein-Gordon theory space. The reasoning of this approach is to assume that $F(k^2)$ could be obtained from an honest Wilsonian treatment of the coupled Klein-Gordon and quantum-gravity system. But without a concrete proposal for quantum gravity, the specific form of $F(k^2)$ is left free (up to the restrictions mentioned above). Moreover, in the context of the Hawking set-up where the WKB approximation applies to the Hawking modes, the classical (effective) Green's function

¹⁷See also ([Brout et al. 1995](#); [Corley 1998](#); [Himemoto and Tanaka 2000](#); [Barcelo et al. 2009](#); [Coutant et al. 2012](#); [Schützhold and Unruh 2013](#)).

¹⁸Two particular differences in our formulation are that in their analysis condition (ii) is left implicit and condition (iv) is reformulated in a mathematically equivalent way (see its use in Equation (16) of ([Unruh and Schützhold 2005](#))).

¹⁹A rapidity of convergence condition on could also be applied here.

contains most of the information of the full quantum mechanical 2-point function. Thus criterion (iv) formally accomplishes many of the same things as requiring the Klein-Gordon state to be Hadamard or to have a scaling limit.²⁰ Using (3) it is possible to explicitly compute the spectrum of Hawking fluxes by essentially following Hawking’s original procedure. The exponential nature of the red-shift between near-horizon Killing and affine modes drowns out any contributions coming from the power-series expansion of $F(k^2)$, and a thermal spectrum is straightforwardly recovered.

The great strength of the modified dispersion relation approach is its physical plausibility. In fact, we take this style of approach to be the most physically plausible derivation of Hawking radiation available since, like the original Hawking calculation, we incorporate core physical features of astrophysical black holes whilst attempting to address certain aspects of the trans-Planckian problem. Since they are adapted from Hawking’s original treatment of astrophysical black holes, modified dispersion relation approaches are embedded in a physically plausible context for Hawking radiation. Furthermore, as could be expected given its origin, the modified-dispersion approach is readily applicable to a huge range of physical systems and thus also has a high degree of universality. In fact, any system to which a Hawking-style derivation can be applied can be supplemented with a corresponding modified dispersion relation treatment. This includes various types of eternal and astrophysical black holes, and also a wide variety of analogue black hole systems.

With regard to empirical support there is increasing cause for confidence. There are a large number of potential analogue realisations of the Hawking effect compatible with the modified dispersion relation argument (Rousseaux et al. 2008; Philbin et al. 2008; Belgiorno et al. 2010; Unruh and Schützhold 2012; Liberati et al. 2012; Nguyen et al. 2015; Jacquet 2018). Furthermore, in some such cases the physical reliability of the techniques has been evaluated experimentally (Rousseaux et al. 2008; Weinfurter et al. 2011; Steinhauer 2016).²¹ Thus, although the arguments are not supported by experimental evidence for the full range of physically instantiated types, the methods of approximation that underlie the modified dispersion relation approach are (at least in part) justified by the fact that they successfully reproduce empirically observed phenomena in some system types.

What we gain in physical reasonableness and degree of universality it appears, unfortunately, we have lost in degree of robustness. In particular, whilst the modifications considered are consistent with known microphysics of the analogue systems, they are highly restrictive with regard to the class of possible ultraviolet modifications to the physics of black holes. It is reasonable to identify this weakness in terms of a lack of integration. In particular, the same

²⁰For the former, criterion (iv) in equation (16) of (Unruh and Schützhold 2005) guarantees that all divergences of the Green’s function are local. For the latter, the fixed point requirement of $\tilde{F}(k^2)$ is just equivalent to a scaling limit.

²¹For discussion of the problem of making inferences about black holes based upon universality arguments combined with such ‘analogue experiments’ see Dardashti et al. (2017), Dardashti et al. (2019), Thébault (2019). For other examples of analogue experiments combined with universality arguments see Thouless (1989), Prüfer et al. (2018), Erne et al. (2018), Eigen et al. (2018).

features that imply that the modified dispersion relations argument can be applied in a wider range of physical situations, are those that limit its robustness in the specific context of black hole physics. A particular concern is whether sub-luminal modified dispersion relations really are sufficient to model quantum gravity effects. Most significantly, the truncation used in the ‘Klein-Gordon theory space’ is of a very special type. In particular, a k -dependent sub-luminal mass term with a scaling limit in the ultraviolet is only the simplest example of how an effective quantum geometry could affect the propagation of Hawking modes across spacetime. Many quantum gravity proposals exhibit some sort of dimensional reduction or enhancement or non-locality in the ultra-violet where even a smooth local parameter such as k (let alone an analytic expansion in terms of it) may not be defined. The space of all such effects cannot plausibly be modelled in such a limited truncation.²² The arguments also do not do well on comprehensiveness since they only apply to the occupation number of the Hawking modes.

In summary, the modified dispersion relation approach can legitimately be treated as a universality argument since it has the three qualities of robustness, physical plausibility and universality to a non-trivial degree. The arguments also have some empirical support. Modified dispersion relation arguments are, however, highly limited in their comprehensiveness. Moreover, with regard to black holes in particular, the arguments have only a low degree of robustness since there is only a very limited degree to which a modified dispersion relations approach can show the gravitational Hawking effect to be robust. Our diagnose of the cause of this issue is again a lack of integration since the basis behind robustness and universality are not entirely consistent, the degree of token level insensitivity varies to a high degree from type to type.

5 Conclusion

Given that the trans-Planckian problem threatens to undermine the entire semi-classical framework for modelling black holes, it is significant that none of the available universality arguments offer an entirely convincing means of response to it. Universality arguments for Hawking radiation in the context of black hole physics thus do poorly in comparison with the Wilsonian arguments in the context of condensed matter physics on their own terms – that is, over and above more general epistemic worries relating to the quantum description of black hole physics. A natural way forward drawn from the analysis above would be to attempt to combine the horizon symmetry and modified dispersion relation arguments for the universality of Hawking radiation. However, the coherence of such a combined approach is yet to be seen. In particular, the stationarity idealization used in the anomaly cancellation approach is not compatible with a Hawking-style derivation where the (non-stationary) process of the formation of the horizon plays an important role. There is, as yet, no smooth limit in which the origi-

²²This can be seen explicitly in the quantum gravitational analysis of the Unruh effect mentioned above in (Alkofer, D’Odorico, Saueressig, and Versteegen 2016).

nal Hawking derivation can be seen to lead to the context in which the anomaly cancellation arguments are based. A significant challenge is to show that anomaly cancellation arguments can be successfully applied to a more realistic model of a collapsing black hole. In particular, there is no existing universality argument or analogue model that deals explicitly with how quantum fluctuations and back-reaction of the horizon, which should dominate the physics of the trans-Planckian modes, could impact the thermality of the Hawking spectrum. If that could be achieved, then it is plausible that a suitably integrated universality argument for Hawking radiation could be established based upon the combination of the modified dispersion relation and horizon symmetry approaches.

Acknowledgments

This paper has benefited from discussions with a large number of different people over a period of years. We would like to give particular thanks to Vincent Ardourel, Alex Blum, Radin Dardashti, Stephan Hartmann, Maxime Jacquet, Ted Jacobson, Sébastien Rivat, Katie Robertson, Charlotte Werndl, Eric Winsberg and audience members in Cambridge, Munich and Nijmegen. We also appreciate helpful comments from two anonymous referees. Finally, we are deeply indebted to Erik Curiel and David Wallace for extensive discussions and detailed comments on a draft manuscript. Work on this paper was supported by the Arts and Humanities Research Council, UK (Grant Ref. AH/P004415/1).

References

- Adler, R. J., P. Chen, and D. I. Santiago (2001). The generalized uncertainty principle and black hole remnants. *General Relativity and Gravitation* 33(12), 2101–2108.
- Agullo, I., J. Navarro-Salas, G. J. Olmo, and L. Parker (2009). Insensitivity of Hawking radiation to an invariant Planck-scale cutoff. *Phys. Rev. D* 80, 047503.
- Ahlers, G., A. Kornblit, and H. J. Guggenheim (1975). Logarithmic corrections to the Landau specific heat near the Curie temperature of the dipolar Ising ferromagnet LiF_2 . *Physical Review Letters* 34(19), 1227.
- Alkofer, N., G. D’Odorico, F. Saueressig, and F. Versteegen (2016). Quantum Gravity signatures in the Unruh effect. *Phys. Rev. D* 94(10), 104055.
- Almheiri, A., D. Marolf, J. Polchinski, and J. Sully (2013). Black Holes: Complementarity or Firewalls? *JHEP* 02, 062.
- Banerjee, R., S. Gangopadhyay, and S. Kulkarni (2010, Dec). Hawking radiation and near horizon universality of chiral Virasoro algebra. *General Relativity and Gravitation* 42(12), 2865–2871.

- Banerjee, R. and S. Kulkarni (2008, Jan). Hawking radiation and covariant anomalies. *Phys. Rev. D* 77, 024018.
- Barcelo, C., L. J. Garay, and G. Jannes (2009). Sensitivity of Hawking radiation to superluminal dispersion relations. *Phys. Rev. D* 79, 024016.
- Batterman, R. W. (2000). ‘Multiple realizability and universality’. *The British Journal for the Philosophy of Science* 51(1), pp. 115–45.
- Batterman, R. W. (2002). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*. Oxford: Oxford University Press.
- Batterman, R. W. and C. C. Rice (2014). Minimal model explanations. *Philosophy of Science* 81(3), 349–376.
- Belgiorno, F., S. Cacciatori, M. Clerici, V. Gorini, G. Ortenzi, L. Rizzi, E. Rubino, V. Sala, and D. Faccio (2010). Hawking radiation from ultrashort laser pulse filaments. *Physical review letters* 105(20), 203901.
- Birmingham, D., K. S. Gupta, and S. Sen (2001). Near horizon conformal structure of black holes. *Phys. Lett. B* 505, 191–196.
- Blum, A. S. and C. Joas (2016). From dressed electrons to quasiparticles: The emergence of emergent entities in quantum field theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 53, 1–8.
- Boulware, D. G. (1975). Quantum field theory in schwarzschild and rindler spaces. *Physical Review D* 11(6), 1404.
- Brout, R., S. Massar, R. Parentani, and P. Spindel (1995, Oct). Hawking radiation without trans-planckian frequencies. *Phys. Rev. D* 52, 4559–4568.
- Burgess, C., J. Hainge, G. Kaplanek, and M. Rummel (2018). Failure of perturbation theory near horizons: the rindler example. *Journal of High Energy Physics* 2018(10), 122.
- Butterfield, J. (2011). ‘Less is different: emergence and reduction reconciled’. *Foundations of Physics* 41(6), 1065–1135.
- Candelas, P. (1980). Vacuum polarization in schwarzschild spacetime. *Physical Review D* 21(8), 2185.
- Carlip, S. (2005). Horizon constraints and black-hole entropy. *Classical and Quantum Gravity* 22(7), 1303.
- Corley, S. (1998). ‘Computing the spectrum of black hole radiation in the presence of high frequency dispersion: An analytical approach’. *Physical Review D* 57(10), p. 6280.
- Coutant, A., R. Parentani, and S. Finazzi (2012). ‘Black hole radiation with short distance dispersion, an analytical S-matrix approach’. *Physical Review D* 85(2), p. 024021.

- Crispino, L. C. B., A. Higuchi, and G. E. A. Matsas (2008). The unruh effect and its applications. *Reviews of Modern Physics* 80(3), 787–838.
- Curiel, E. (2018, 08). What is a black hole? *Nature Astronomy* (forthcoming).
- Dappiaggi, C., V. Moretti, N. Pinamonti, et al. (2011). Rigorous construction and hadamard property of the unruh state in schwarzschild spacetime. *Advances in Theoretical and Mathematical Physics* 15(2), 355–447.
- Dardashti, R., S. Hartmann, K. Thébault, and 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* (forthcoming).
- Dardashti, R., K. 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.
- DeWitt, B. S. (1975). Quantum field theory in curved spacetime. *Physics Reports* 19(6), 295–357.
- Dey, R., S. Liberati, and D. Pranzetti (2017). The black hole quantum atmosphere. *Physics Letters B* 774, 308–316.
- Dougherty, J. and C. Callender (2019). Black hole thermodynamics: More than an analogy? In A. Ijjas and B. Loewer (Eds.), *Guide to the Philosophy of Cosmology*. Oxford University Press.
- Eigen, C., J. A. Glidden, R. Lopes, E. A. Cornell, R. P. Smith, and Z. Hadzibabic (2018). Universal prethermal dynamics of bose gases quenched to unitarity. *arXiv preprint arXiv:1805.09802*.
- Erne, S., R. Bücker, T. Gasenzer, J. Berges, and J. Schmiedmayer (2018). Universal dynamics in an isolated one-dimensional bose gas far from equilibrium. *Nature* 563(7730), 225.
- Fisher, M. (1998). Renormalization group theory: Its basis and formulation in statistical physics. *Reviews of Modern Physics* 70(2), 653.
- Franklin, A. (2017). On the renormalisation group explanation of universality on the renormalisation group explanation of universality on the renormalization group explanation of universality. *Philosophy of Science* 85(2), 225–248.
- Fraser, J. D. (2018, Jul). Renormalization and the formulation of scientific realism. *Philosophy of Science*.
- Fredenhagen, K. and R. Haag (1990). On the derivation of hawking radiation associated with the formation of a black hole. *Communications in mathematical physics* 127(2), 273–284.

- Garay, L., J. Anglin, J. Cirac, and P. Zoller (2000). ‘Sonic analog of gravitational black holes in Bose-Einstein condensates’. *Physical Review Letters* **85**(22), p. 4643.
- Gibbons, G. (1977). Quantum processing near black holes. In R. Ruffini (Ed.), *Proceedings of the First Marcel Grossmann Meeting on General Relativity*, pp. pp 449–58. North-Holland, Amsterdam.
- Giddings, S. B. (2016). Hawking radiation, the stefan–boltzmann law, and unitarization. *Physics Letters B* **754**, 39–42.
- Goldenfeld, N. (1992). *Lectures on phase transitions and the renormalization group*. CRC Press.
- Harlow, D. (2016). Jerusalem lectures on black holes and quantum information. *Reviews of Modern Physics* **88**(1), 015002.
- Hartle, J. B. and S. W. Hawking (1976). Path-integral derivation of black-hole radiance. *Physical Review D* **13**(8), 2188.
- Hawking, S. and W. Israel (1979). An introduction survey. In S. Hawking and W. Israel (Eds.), *General relativity: an Einstein centenary survey*. Cambridge University Press.
- Hawking, S. W. (1975). ‘Particle creation by black holes’. *Communications in mathematical physics* **43**(3), 199–220.
- Helfer, A. D. (2003). Do black holes radiate? *Reports on Progress in Physics* **66**(6), 943.
- Helfer, A. D. (2010). Comment on ‘Insensitivity of Hawking Radiation to an invariant Planck-scale cutoff’. *Phys. Rev. D* **81**, 108501.
- Himemoto, Y. and T. Tanaka (2000). ‘Generalization of the model of Hawking radiation with modified high frequency dispersion relation’. *Physical Review D* **61**(6), p. 064004.
- Hod, S. (2016). Hawking radiation and the stefan–boltzmann law: The effective radius of the black-hole quantum atmosphere. *Physics Letters B* **757**, 121–124.
- Hollands, S. and R. M. Wald (2015). Quantum fields in curved spacetime. *Physics Reports* **574**, 1–35.
- Iso, S., H. Umetsu, and F. Wilczek (2006a). Anomalies, Hawking radiations and regularity in rotating black holes. *Phys. Rev. D* **74**, 044017.
- Iso, S., H. Umetsu, and F. Wilczek (2006b). Hawking radiation from charged black holes via gauge and gravitational anomalies. *Phys. Rev. Lett.* **96**, 151302.
- Jacobson, T. (1991). ‘Black-hole evaporation and ultrashort distances’. *Physical Review D* **44**(6), 1731.
- Jacobson, T. (1993). ‘Black hole radiation in the presence of a short distance cutoff’. *Physical Review D* **48**(2), 728–41.

- Jacobson, T. (1996). ‘introductory lectures on black hole thermodynamics’.
www.physics.umd.edu/grt/taj/776b/lectures.pdf.
- Jacobson, T. (2005). ‘Introduction to quantum fields in curved spacetime and the Hawking effect’. In *Lectures on Quantum Gravity*, pp. pp. 39–89. New York: Springer.
- Jacquet, M. (2018). *Negative Frequency at the horizon: theoretical study and experimental realisation of analogue gravity physics in dispersive optical media*. Springer.
- Kay, B. S. and R. M. Wald (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.
- Landau, L. (1936). The theory of phase transitions. *Nature* 138(3498), 840–841.
- Leonhardt, U. and T. G. Philbin (2008). ‘The case for artificial black holes’. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 366(1877), pp. 2851–7.
- Liberati, S., A. Prain, and M. Visser (2012). Quantum vacuum radiation in optical glass. *Physical Review D* 85(8), 084014.
- Mainwood, P. (2006). Phase transitions in finite systems. *unpublished MS; (corresponds to Chapter 4 of Mainwood (2006) Available at <http://philsci-archive.pitt.edu/8339/>*.
- Martin, J. and R. H. Brandenberger (2001). Trans-planckian problem of inflationary cosmology. *Physical Review D* 63(12).
- Nguyen, H., D. Gerace, I. Carusotto, D. Sanvitto, E. Galopin, A. Lemaître, I. Sagnes, J. Bloch, and A. Amo (2015). Acoustic black hole in a stationary hydrodynamic flow of microcavity polaritons. *Physical review letters* 114(3), 036402.
- Palacios, P. (2019). Phase transitions: A challenge for intertheoretic reduction? *Philosophy of Science (forthcoming)*.
- Parentani, R. (2010). From vacuum fluctuations across an event horizon to long distance correlations. *Physical Review D* 82(2), 025008.
- Parikh, M. K. and F. Wilczek (2000). Hawking radiation as tunneling. *Physical Review Letters* 85(24), 5042.
- Philbin, T. G., C. Kuklewicz, S. Robertson, S. Hill, F. König, and U. Leonhardt (2008). Fiber-optical analog of the event horizon. *Science* 319(5868), 1367–1370.
- Polchinski, J. (1995). String theory and black hole complementarity. *arXiv preprint hep-th/9507094*.
- Prüfer, M., P. Kunkel, H. Strobel, S. Lannig, D. Linnemann, C.-M. Schmied, J. Berges, T. Gasenzer, and M. K. Oberthaler (2018). Observation of universal dynamics in a spinor bose gas far from equilibrium. *Nature* 563(7730), 217–220.

- Rousseaux, G., C. Mathis, P. Maïssa, T. G. Philbin, and U. Leonhardt (2008). Observation of negative-frequency waves in a water tank: a classical analogue to the Hawking effect? *New Journal of Physics* 10(5), 053015.
- Ruetsche, L. (2011). *Interpreting quantum theories*. Oxford University Press.
- Saatsi, J. and A. Reutlinger (2018). Taking reductionism to the limit: How to rebut the antireductionist argument from infinite limits. *Philosophy of Science* 85(3), 455–482.
- Schützhold, R. and W. G. Unruh (2013). Hawking radiation with dispersion versus breakdown of the WKB approximation. *Phys. Rev. D* 88(12), 124009.
- Sciama, D. W., P. Candelas, and D. Deutsch (1981). Quantum field theory, horizons and thermodynamics. *Advances in Physics* 30(3), 327–366.
- Shech, E. (2018). Infinite idealizations in physics. *Philosophy Compass*, e12514.
- Singleton, D. and S. Wilburn (2011). Hawking radiation, Unruh radiation and the equivalence principle. *Phys. Rev. Lett.* 107, 081102.
- Steinhauer, J. (2016, August). Observation of quantum hawking radiation and its entanglement in an analogue black hole. *Nature Physics Advance Online Publication*.
- Täuber, U. C. (2012). Renormalization group: applications in statistical physics. *Nuclear Physics B-Proceedings Supplements* 228, 7–34.
- Thébault, K. P. Y. (2019, October). What can we learn from analogue experiments? In R. Dardashti, R. Dawid, and K. Thébault (Eds.), *Why Trust a Theory?* Cambridge University Press.
- Thompson, R. and L. Ford (2008). ‘Enhanced black hole horizon fluctuations’. *Physical Review D* 78(2), p. 024014.
- Thouless, M. D. (1989). Some mechanics for the adhesion of thin films. *Thin Solid Films* 181(1-2), 397–406.
- Unruh, W. (1981). ‘Experimental black-hole evaporation?’. *Physical Review Letters* 46(21), 1351–53.
- Unruh, W. (1995). ‘Sonic analogue of black holes and the effects of high frequencies on black hole evaporation’. *Physical Review D* 51(6), p. 2827.
- Unruh, W. and R. Schützhold (2012). Hawking radiation from “phase horizons” in laser filaments? *Physical Review D* 86(6), 064006.
- Unruh, W. G. (1975). Particle Detectors and Black Holes. In *Marcel Grossmann Meeting on the Recent Progress of the Fundamentals of General Relativity Trieste, Italy, July 7-12, 1975*, pp. 527–536.
- Unruh, W. G. (1976). Notes on black-hole evaporation. *Physical Review D* 14(4), 870.

- Unruh, W. G. (1977). Origin of the particles in black-hole evaporation. *Physical Review D* 15(2), 365.
- Unruh, W. G. (2014). Has hawking radiation been measured? *Foundations of Physics* 44(5), 532–545.
- Unruh, W. G. and R. Schützhold (2005). ‘Universality of the Hawking effect’. *Physical Review D* 71(2), 024028.
- Wald, R. M. (1994). *Quantum field theory in curved spacetime and black hole thermodynamics*. University of Chicago Press.
- Wallace, D. (2017, 10). The case for black hole thermodynamics, part i: phenomenological thermodynamics. <https://arxiv.org/abs/1710.02724>.
- Weinfurter, S., E. W. Tedford, M. C. Penrice, W. G. Unruh, and G. A. Lawrence (2011). ‘Measurement of stimulated Hawking emission in an analogue system’. *Physical Review Letters* 106(2), p. 021302.
- Wilson, K. G. (1971). Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture. *Physical review B* 4(9), 3174.
- Wilson, K. G. and J. Kogut (1974). The renormalization group and the ϵ expansion. *Physics Reports* 12(2), 75–199.