

# Black hole entropy from soft hair

Malcolm J. Perry

*Department of Applied Mathematics and Theoretical Physics,  
Centre for Mathematical Sciences,  
University of Cambridge,  
Wilberforce Road,  
Cambridge CB3 0WA,  
UK.*

*\*E-mail: malcolm@damtp.cam.ac.uk*

We start by looking at why we believe that black holes have entropy. According to Boltzmann, the entropy is a measure of the number of microstates of a system. We suggest here that the entropy arises from a holographic conformal field theory on the black hole horizon. Finally, we discuss some of the implications for the information paradox.

*Keywords:* Black holes. Thermodynamics, Holography. Information Paradox.

## 1. Introduction

Black holes were first thought about by John Michell back in 1784.<sup>1</sup> He reasoned that if the escape velocity from an object like a star exceeded the speed of light, then it would give rise to an object that cannot be seen optically but whose gravitational field would betray its existence. These ideas were given substance by Einstein's general theory of relativity and the subsequent discovery of solutions of the field equations that represented stationary black holes.<sup>2–6</sup> One puzzling feature of stationary black holes is that they are completely characterised by just a few parameters; their mass  $M$ , angular momentum  $\mathbf{J}$  and electric charge  $Q$ .<sup>7–11</sup> It is this observation that is the basis for the information paradox.<sup>12</sup>

The thermodynamics of black holes in general relativity has a history starting in 1972. The first relevant discovery, the area theorem, was made by Hawking.<sup>13</sup> He found that the area of a black hole horizon could never decrease provided the null convergence condition  $R_{ab}k^ak^b \geq 0$  holds for every null vector  $k^a$ . Shortly after this, Jacob Bekenstein suggested<sup>14</sup> that the entropy of a black hole must be proportional to the area of its event horizon. His reasoning was based on three observations. The first was that if a Kerr-Newman black hole increased its mass by an amount  $dM$ , then

$$dM = \frac{\kappa dA}{8\pi} + \Phi dQ + \boldsymbol{\Omega} \cdot d\mathbf{J} \quad (1)$$

where  $\kappa$  is the black hole surface gravity,  $A$  the area of the event horizon,  $\Phi$  the electrostatic potential of the black hole and  $\boldsymbol{\Omega}$  its angular velocity.  $\kappa$ ,  $A$ ,  $\Phi$  and  $\boldsymbol{\Omega}$  are all determined in terms of  $M$ ,  $\mathbf{J}$  and  $Q$ . The second piece of evidence was that he reasoned that a black hole must have some kind of internal structure that

resulted from its method of formation. That would give rise to an entropy

$$S = - \sum_n p_n \ln p_n \quad (2)$$

arising from the probability of the occupation of the  $n^{\text{th}}$ -state being  $p_n$ . Finally, he showed that it was necessary for this black hole entropy to be added to the thermodynamic entropy of the rest of the universe in order to have a consistent theory of thermodynamics. This came about because otherwise dropping a box of radiation into a black hole would cause the entropy of the universe to decrease, in contradiction to the second law of thermodynamics.

His ideas were met with a certain amount of scepticism because black holes were thought to have vanishing temperature. Despite that, Bardeen, Carter and Hawking<sup>15</sup> pointed out the similarities between the first law of thermodynamics and (1) and also the second law of thermodynamics and the area theorem.

In 1974, Hawking<sup>16,17</sup> showed that black holes had a temperature  $T_H$  of  $\hbar\kappa/(2\pi)$ . Unlike previous work, his calculation was quantum mechanical in nature. Black holes would emit particles with a thermal spectrum at a temperature given by  $T_H$ . By identifying (1) with the first law of thermodynamics, one can immediately infer that the entropy must be given by  $A/(4\hbar)$ . The area theorem is thereby identified with the second law of thermodynamics. A somewhat different view of entropy was taken in Ref. 18. The idea here was to use the path integral for gravity to derive black hole entropy. Although gravity is unrenormalizable, there is no obstacle to using the path integral to lowest order as the uncontrollable divergences only occur at one loop or beyond. The action for pure gravity is, including the Gibbons-Hawking-York boundary terms<sup>18,19</sup>

$$I[g, h] = \frac{1}{16\pi} \int_{\mathcal{M}} R(g) \sqrt{\|g\|} d^4x + \frac{1}{8\pi} \int_{\partial\mathcal{M}} K \sqrt{\|h\|} d^3x + C[h] \quad (3)$$

where now  $\mathcal{M}$  is the spacetime manifold with metric  $g$  and Ricci scalar  $R(g)$ . The boundary of  $\mathcal{M}$  is  $\partial\mathcal{M}$  with metric  $h$  and second fundamental form  $K$ .  $C[h]$  is any functional of  $h$  and is designed to make the action of flat spacetime vanish. Suppose one wants to find the partition function for a black hole spacetime. Then one wants to compute  $Z = \text{tr}(e^{-\beta\mathcal{H}})$  where  $\beta$  is the inverse temperature and  $\mathcal{H}$  the Hamiltonian. This can be done by realising that  $e^{i\mathcal{H}t}$  is the time evolution operator and so if  $t$  is identified with  $t + i\beta$  then  $Z$  is given by

$$Z = \int D[g] e^{-I[g, h]/\hbar}, \quad (4)$$

where now the integral is over all metrics  $g$  of positive definite signature and that approach flat space at infinity and are periodic in imaginary time  $t$  with period  $\beta = T_H^{-1}$ .

The Schwarzschild metric is

$$ds^2 = -(1 - \frac{2M}{r}) dt^2 + (1 - \frac{2M}{r})^{-1} dr^2 + r^2 d\Omega^2, \quad (5)$$

where  $M$  is the mass of the static black hole with horizon at  $r = 2M$  and  $d\Omega^2$  is the metric on the unit 2-sphere. Taking  $t = i\tau$  so that the geometry is as described above gives the Euclidean metric

$$ds^2 = \left(1 - \frac{2M}{r}\right)d\tau^2 + \left(1 - \frac{2M}{r}\right)^{-1}dr^2 + r^2d\Omega^2 \quad (6)$$

Now,  $r = 2M$  is a conical singularity that is resolved provided that  $\tau$  is identified with period  $8\pi M$ .<sup>20</sup> This periodicity is precisely the same periodicity expected from the Hawking calculation of the black hole temperature. Interpreting the exponential of the action as the partition function, reproduces the black hole entropy. In this calculation, the region  $r < 2M$  has been removed from consideration. Implicitly this means that the internal degrees of freedom have been traced over. One is thereby led to believe that the black hole does have some kind of internal structure that cannot be probed by external observers who just look at the classical geometry. The same kind of reasoning can be applied to the Kerr-Newman metric too.

Black holes evaporate. The black hole uniqueness theorems suggest that the only properties that a stationary black hole has are just the mass, charge and spin. As a consequence, there is a tension with the ideas of quantum mechanics. If a black hole completely disappears, then the final state should be unitarily equivalent to the initial state. Obviously, there are enormous number of ways in which the black hole could form. The black hole, once it has settled down to a more or less equilibrium state, is described by just those three parameters. The Hawking radiation is thermal and characterised by the Hawking temperature. Such a final state consisting of Hawking radiation will not be unitarily related to the initial state that gave rise to the black hole. This is the information paradox. It might be that quantum mechanical information really is lost in gravitational collapse. But then, the whole edifice of quantum mechanics would need to be rethought. The incredible success of quantum mechanics would seem to discourage such a viewpoint. Alternatively, there might be something wrong with the uniqueness theorems. It is this latter possibility that we will investigate here.

In what follows, we will use covariant phase space methods<sup>21–27</sup> to understand the nature of charges in general relativity and the consequences for the physics of black holes. The reason for using the covariant phase space method is to preserve as much as possible of the covariance of theory. Had we picked the more conventional canonical methods, we would be forced to pick a particular time coordinate which would obscure matters. Furthermore, it would be impossible to understand what happens on null surfaces such as the event horizon. Our aim now is to try to understand something about the microscopic origin of black hole entropy. The hope is that this will aid a resolution of the information paradox.

In pure general relativity, one can start with the Einstein-Hilbert action  $I$  given by

$$I = \frac{1}{16\pi} \int_{\mathcal{M}} R(g) \sqrt{\|g\|} d^4x. \quad (7)$$

This action omits the boundary terms, but these are not germane to the discussion that follows. One finds the Einstein equation by performing a variation of the action induced by a variation of the metric  $g_{ab} \rightarrow g_{ab} + h_{ab}$ . This results in the variation  $\delta I$  given by

$$\delta I = \int_{\mathcal{M}} (R_{ab} - \frac{1}{2} R g_{ab}) h^{ab} \sqrt{\|g\|} d^4x + \int_{\partial\mathcal{M}} \theta \quad (8)$$

The three-form on the boundary  $\theta(g, h)$  is known as the presymplectic potential and has components

$$(*\theta)_a = \frac{1}{16\pi} (\nabla_b h_a^b - \nabla_a h) \quad (9)$$

where  $h = h_{ab} g^{ab}$ . In canonical general relativity, the boundary term would be thought of

$$\int_{\partial\mathcal{M}} \sum_i p_i \delta q^i \quad (10)$$

where  $q^i$  are the generalised coordinates,  $p_i$  are the generalised momenta and  $i$  represent the tensor indices of these fields.

The presymplectic density  $\omega(g; h, h')$  is defined by a second variation  $g_{ab} \rightarrow g_{ab} + h'_{ab}$

$$\omega(g; h, h') = \delta\theta(g, h') - \delta'\theta(g, h) \quad (11)$$

Finally, the symplectic form for general relativity is

$$\Omega_\Sigma = \int_\Sigma \omega \quad (12)$$

where  $\Sigma$  is any partial Cauchy surface in the spacetime. In the language of the canonical theory,  $\Omega_\Sigma$  would be

$$\int_\Sigma \sum_i \delta p_i \wedge \delta q^i. \quad (13)$$

One property that  $\omega$  has is that if the background metric  $g_{ab}$  obeys the Einstein equation and both  $h_{ab}$  and  $h'_{ab}$  obey the linearised Einstein equations, then  $\omega$  is closed. Thus  $\Omega_\Sigma$  is constant under variations of  $\Sigma$  as long as the boundaries of  $\Sigma$  are fixed.

In general relativity, the symmetry group is the group of diffeomorphisms. An infinitesimal coordinate transformation is specified by a vector field  $\zeta^a$ . This induces a variation in any tensor field given by the Lie derivative of that field. Thus, for example, the variation of the metric is given by

$$\delta g_{ab} = \mathcal{L}_\zeta g_{ab} = g_{ab} + \nabla_a \zeta_b + \nabla_b \zeta_a. \quad (14)$$

The bulk term in the variation of the action  $\delta I$  is invariant under such a transformation but the boundary term is not. The infinitesimal co-ordinate transformations obey an algebra whose composition law is

$$\mathcal{L}_\zeta \mathcal{L}_\eta - \mathcal{L}_\eta \mathcal{L}_\zeta = \mathcal{L}_{[\zeta, \eta]} \quad (15)$$

where  $\zeta$  and  $\eta$  are two (smooth) vector fields and

$$[\zeta, \eta] = \mathcal{L}_\zeta \eta = -\mathcal{L}_\eta \zeta \quad (16)$$

Suppose now that in  $\Omega$  one makes  $h'_{ab}$  a gauge transformation given by the vector field  $\zeta$ . Then  $\Omega$  can be written as a boundary integral. Explicitly,

$$Q_\zeta = \frac{1}{16\pi} \int_{\partial\Sigma} F_{ab} dS^{ab} \quad (17)$$

with

$$F_{ab} = -2\zeta_{[a}\nabla_{b]}h + 2\zeta_{[a}\nabla^c h_{b]c} - 2\zeta^c\nabla_{[a}h_{b]c} - h\nabla_{[a}\zeta_{b]} + 2h_{c[a}\nabla^c\zeta_{b]}. \quad (18)$$

Let  $\partial\Sigma$  is a closed 2-surface  $S$ , for example the celestial sphere or a black hole event horizon. One would like to interpret  $Q_\zeta$  as the variation in the Noether charge conjugate to  $\zeta$  that is enclosed in the interior of  $S$  as one moves between the metric  $g_{ab}$  and  $g_{ab} + h_{ab}$ . There is a complication with this idea because in such a change, there might be a flux of charge crossing  $S$ . To take account of this possibility, one needs to examine  $Q_\zeta$  and identify such terms and subtract them out. In more mathematical language, one we want  $Q_\zeta$  to be a function of state. As such it must be a 1-form on the infinite-dimensional phase space of the theory. This 1-form needs to be exact so that if one goes along a path  $\Gamma$  between  $g_{ab}$  and  $g_{ab} + h_{ab}$ , then  $Q_\zeta$  is independent of the path  $\Gamma$ , and therefore dependent only the end-points of that path. The definition on  $Q_\zeta$  thus needs to be modified by the addition of a suitable counterterm  $Q_\zeta \rightarrow Q_\zeta + Q_\zeta^{ct}$ . Finding  $Q_\zeta^{ct}$  needs to be done on a case by case basis as has been elegantly explained in detail by Wald and Zoupas.<sup>26</sup>

In the case that  $\zeta$  were a time translation, then  $Q_\zeta$  would be the quasi-local mass enclosed in  $S$ .<sup>28</sup> If it were a spatial translation then the momentum. If  $\zeta$  were a Killing vector, then  $Q_\zeta$  would be the same as the Komar integral.<sup>29</sup> If  $\zeta$  were a supertranslation or super-rotation at null infinity, then  $Q_\zeta$  would be the corresponding supertranslation or super-rotation charge. Equally, one can define charges on the black hole horizon and these are the soft charges or soft black hole hair.<sup>30</sup>

Diffeomorphism invariance of general relativity means that the charges  $Q_\zeta$  lie in some representation of group of coordinate transformations. Thus

$$\delta_\zeta Q_\eta - \delta_\eta Q_\zeta = Q_{[\zeta, \eta]} \quad (19)$$

Were this relation not to hold, general coordinate invariance would be violated, in gross contradiction to our expectations of what should be true in physics. However, what we find is that this relationship does not hold for charges on black hole event horizons. Instead, we find

$$\delta_\zeta Q_\eta - \delta_\eta Q_\zeta = Q_{[\zeta, \eta]} + K(\zeta, \eta) \quad (20)$$

where  $K(\zeta, \eta)$  is a central extension of this algebra.<sup>31</sup> We will now explore a particular example and move on to its interpretation.

We start from the Kerr metric in Boyer-Lindquist coordinates.

$$ds^2 = \rho^2 \left( \frac{dr^2}{\Delta} + d\theta^2 \right) + (r^2 + a^2) \sin^2 \theta d\phi^2 - dt^2 + \frac{2Mr}{\rho^2} (a \sin^2 \theta d\phi - dt)^2 \quad (21)$$

where

$$\Delta = r^2 - 2Mr + a^2 \quad (22)$$

and

$$\rho^2 = r^2 + a^2 \cos^2 \theta. \quad (23)$$

$M$  is the mass of the black hole and  $J = Ma$  is its angular momentum.  $\Delta = 0$  at  $r_{\pm}$  with  $r_+$  being the location of the outer horizon,  $r_-$  the location of the inner horizon and

$$r_{\pm} = M \pm \sqrt{M^2 - a^2} \quad (24)$$

Now we will define “conformal” coordinates<sup>32</sup> and assume that the black hole is not extreme so that  $m^2 > a^2$ .

$$w^+ = \sqrt{\frac{r - r_+}{r - r_-}} e^{2\pi T_R \phi} \quad (25)$$

$$w^- = \sqrt{\frac{r - r_+}{r - r_-}} e^{2\pi T_L \phi - t/2M} \quad (26)$$

$$y = \sqrt{\frac{r_+ - r_-}{r - r_-}} e^{\pi(T_L + T_R)\phi - t/4M} \quad (27)$$

where

$$T_L = \frac{r_+ + r_-}{4\pi a} \quad \text{and} \quad T_R = \frac{r_+ - r_-}{4\pi a} \quad (28)$$

The future outer horizon is  $w^- = 0$  and the past outer horizon is  $w^+ = 0$ . The azimuthal coordinate  $\phi$  is identified with period  $2\pi$  and this induces an identification on  $w^+, w^-$  and  $y$  as

$$w^+ \rightarrow e^{4\pi^2 T_R} w^+, \quad w^- \rightarrow e^{4\pi^2 T_L} w^-, \quad y \rightarrow e^{2\pi^2(T_L + T_R)} y. \quad (29)$$

The line element close to the horizon bifurcation surface  $w^+ = w^- = 0$  is<sup>33</sup>

$$ds^2 = \frac{4\rho_+^2}{y^2} dw^+ dw^- + \frac{16M^2 a^2 \sin^2 \theta}{\rho_+^2 y^2} dy^2 + \rho_+^2 d\theta^2 + O(w^+, w^-) \quad (30)$$

where  $\rho_+^2 = r_+^2 + a^2 \cos^2 \theta$ . If one looks at the  $w^+, w^-, y$ -plane by setting  $\theta$  to be a constant, then this line element is that of  $AdS_3/\Gamma$  with  $\Gamma$  being some discrete group. So close to the horizon bifurcation surface, the geometry of spacetime is some kind of warped product of the line segment  $\theta \in [0, \pi]$  and a deformed portion of three-dimensional anti-de Sitter space. The classic work of Brown and Henneaux<sup>31</sup> shows

that in spacetimes that are asymptotic to anti-de Sitter spacetime, the diffeomorphism algebra has anomalies. One might therefore suspect that something similar happens in the case of the Kerr black hole.

Consider the diffeomorphism given by the vector field  $\zeta_n$

$$\zeta_n = \epsilon_n(w^+) \partial_+ + \frac{1}{2} \epsilon'(w^+) y \partial_y \quad (31)$$

with

$$\epsilon_n(w^+) = 2\pi T_R (w^+)^{(1+\frac{in}{2\pi T_R})} \quad (32)$$

and  $n$  being any integer. It should be noted that under the identifications of either  $w^+ \rightarrow w^+ e^{4\pi^2 T_R}$  or  $y \rightarrow y e^{2\pi^2 (T_L + T_R)}$  that  $\zeta_n$  is invariant. This vector field is well-defined on the future horizon. These vector fields obey the Witt (or centerless Virasoro) algebra with the commutator

$$[\zeta_n, \zeta_m] = i(m-n)\zeta_{n+m} \quad (33)$$

Similarly, one can find a second vector field  $\tilde{\zeta}_n$  given by

$$\tilde{\zeta}_n = \tilde{\epsilon}_n(w^-) \partial_- + \frac{1}{2} \tilde{\epsilon}'(w^-) y \partial_y \quad (34)$$

and  $\tilde{\epsilon}$  being given

$$\tilde{\epsilon}_n(w^-) = 2\pi T_L (w^-)^{(1+\frac{in}{2\pi T_L})} \quad (35)$$

again with  $n$  being any integer. This vector field is well-defined on the past horizon. Again, it is invariant under the identifications  $w^- \rightarrow w^- e^{4\pi^2 T_L}$  or  $y \rightarrow y e^{2\pi^2 (T_L + T_R)}$ . It too obeys the Virasoro algebra

$$[\tilde{\zeta}_n, \tilde{\zeta}_m] = i(m-n)\tilde{\zeta}_{n+m}. \quad (36)$$

Both of these vector fields are well-defined on the bifurcation surface and commute with each other there

$$[\zeta_n, \tilde{\zeta}_m] = 0. \quad (37)$$

These vector fields can be used to generate charges on the bifurcation surface. To do this we need to introduce an appropriate counterterm. This is given by

$$-\frac{1}{8\pi} \int dS^{ab} \nabla_a (\zeta^c h_b^d) N_{cd} \quad (38)$$

where  $N_{ab}$  are the components of the volume form on the normal bundle to the horizon. There is a precisely similar expression for the fields  $\tilde{\zeta}_n$ . One then finds that the charges on the bifurcation surface obey the algebra

$$[Q_n, Q_m] = i(n-m)Q_{n+m} + in^3 J \delta_{n,-m} \quad (39)$$

for the right-handed algebra and

$$[\tilde{Q}_n, \tilde{Q}_m] = i(n-m)\tilde{Q}_{n+m} + in^3 J \delta_{n,-m} \quad (40)$$

for the left-handed algebra. Finally, the left and right algebras commute with each other

$$[Q_n, \tilde{Q}_m] = 0. \quad (41)$$

In both cases, the central terms shown here correspond to the conventionally normalised Virasoro algebra with central charges given by  $c_L = c_R = 12J$ . Thus the diffeomorphism algebra has an anomaly.

We postulate that this anomaly is cancelled by holographic degrees of freedom on the horizon expressed in terms of a two-dimensional conformal field theory. Consider for a moment the expressions for the absorption probabilities for particles incident on a Kerr black hole. Suppose we look at a particle with energy  $\delta E$  and angular momentum parallel to the black hole spin  $\delta J$ . Then we observe that the absorption probability obtains a suggestive factor of

$$|\Gamma(1 + \frac{i\omega_L}{2\pi T_L})|^2 |\Gamma(1 + \frac{i\omega_R}{2\pi T_R})|^2 \quad (42)$$

where

$$\omega_L = \frac{2M^3}{J}\delta E \quad \omega_R = \frac{2M^3}{J}\delta E - \delta J. \quad (43)$$

This is precisely what is to be expected for a conformal field theory where the left-handed degrees of freedom are at a temperature of  $T_L$  and the right-handed degrees of freedom are at a temperature of  $T_R$  and one is asking for the absorption probability for particles of energy  $\omega_L$  in the left-handed sector and energy  $\omega_R$  in the right-handed sector. We take it that there are no coincidences in nature and therefore we really can attribute our observations to the existence of holographic degrees of freedom on the horizon described by a two-dimensional conformal field theory.

A general property of conformal field theories, provided the central charge is sufficiently large, was first described by Cardy.<sup>34</sup> The entropy for a system with central charges  $c_L$  and  $c_R$  for the two sectors at temperatures  $T_L$  and  $T_R$  is given by

$$S = \frac{\pi^2}{3}(c_L T_L + c_R T_R). \quad (44)$$

Plugging in our expressions for  $c_L, c_R, T_L$  and  $T_R$  gives

$$S = \frac{1}{4}A. \quad (45)$$

It is hard to believe that this is a coincidence. It appears therefore we have identified the degrees of freedom responsible for black hole entropy.<sup>33</sup>

Subsequent to the conference, it has been shown that the same methods reproduce the entropy for the Kerr-Newman family of black holes<sup>35</sup> and for uncharged black holes in anti-de Sitter spacetime.<sup>36</sup>



A key question is to ask how this affects our view of the information paradox. We have shown how to account for black hole entropy in terms of a holographic two-dimensional conformal field theory living on the black hole horizon. It is however far from clear that the states of such a theory can record all of the quantum mechanical information that is pertinent to black hole formation from ordinary matter. We are therefore left with a collection of problems that need exploration and solution before there can be any claim of solving the information paradox. We conclude this essay with a summary of outstanding issues. Does the horizon conformal field theory contain a complete description of the black hole formation process? How does the Hawking radiation encode this information so as to preserve unitary time evolution? Why is it that the black hole entropy is independent of the spectrum of elementary particles when the number of ways a black hole can be formed is highly dependent on that spectrum. For example, if there were a million different species of electron, the number of ways a black hole could form would be vastly higher than if there a single type of electron. Nevertheless, the Hawking entropy would be same.

Suppose a particle falls into a black hole. Classically, a co-moving observer sees it pass through the horizon without anything obvious happening. In the case of a Schwarzschild black hole, it will reach the singularity in a finite amount of proper time. The singularity is a boundary of spacetime and so we believe the particle to have disappeared. In the case of rotating black holes, it seems plausible that it will also inevitably reach a singularity as the inner horizon of a Kerr black hole is unstable and is presumed to become singular once any energy-momentum arrives there. However, if the particle is to leave an imprint on the state of the horizon conformal field theory, it appears to have violated the quantum no-cloning theorem. Roughly speaking, the no-cloning theorem says that you cannot duplicate the state of a particle by unitary time evolution. A number of technical assumptions go into this amongst which is a notion of locality, a dubious assumption in the case of gravitation.

Then there are some more challenging issues. What happens to the singularity? It is a classical concept and shows that classical general relativity is an incomplete theory. What happens quantum mechanically? There is no satisfactory answer at present. What are the final stages of black hole evaporation? The picture presented seems to suggest that all symmetries in nature are gauge symmetries and not global symmetries. For example in the standard model, baryon number is a global symmetry, but it is hard to see how this could be encoded in the picture presented here. There is one ambitious theory that predicts that all symmetries are gauge symmetries and that is string theory. Although string theory is successful in resolving the divergence problems of quantum gravity, and potentially geometrizing the spectrum of elementary particles, it is far from being a theory of spacetime. Hopefully, the picture here will provide a guide to the true nature of quantum gravity, but there are immense and exciting challenges to the construction of such a theory. Eventually, we hope that the construction of such a theory will lead to deep insights into the nature of our Universe.

## Acknowledgements

The work presented here was joint with the late Stephen Hawking, Andy Strominger and Sasha Haco. We hope that it is a fitting memory to Stephen who was a dear friend for almost forty years.

We are grateful to Sangmin Choi, Geoffrey Comp'ere, Peter Galison, Monica Guica, Dan Harlow, Roy Kerr, Alex Lupsasca, Juan Maldacena, Alex Maloney, Suvrat Raju and Maria Rodriguez for useful conversations. This paper was supported in part by DOE de-sc0007870, the John Templeton Foundation, the Black Hole Initiative at Harvard, the Radcliffe Institute for Advanced Study and the UK STFC. We would also like to thank the George and Cynthia Mitchell Foundation, whose generous support allowed for over a decade of scientific workshops, spanning Cooks' Branch in Texas and Great Brampton House and Brinsop Court in England, without which this collaboration would not have been possible. The work began at a Mitchell Foundation retreat and was completed at another. We are profoundly grateful to the Foundation and personally to Sheridan Lorenz and George Mitchell for their contribution to the development of fundamental science.

## References

1. J. Michell, "On the Means of Discovering the Distance, Magnitude, of the Fixed Stars, in Consequence of the Diminution of the Velocity of Their Light, in Case Such a Diminution Should be Found to Take Place in any of Them, and Such Other Data Should be Procured from Observations, as Would be Farther Necessary for That Purpose.," *Phil. Trans. Roy. Soc. Lond.* **74** (1784), 35-57. doi:10.1098/rstl.1784.0008
2. K. Schwarzschild, "On the gravitational field of a mass point according to Einstein's theory," *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* (1916), 189-196. [arXiv:physics/9905030 [physics]].
3. H. Reissner, "Über die Eigengravitation des elektrischen Feldes nach der Einsteinschen Theorie". *Annalen der Physik* (1916), 50 (9): 106–120. doi:10.1002/andp.19163550905.
4. G. Nordström, "On the Energy of the Gravitational Field in Einstein's Theory". *Verhandl. Koninkl. Ned. Akad. Wetenschap., Afdel. Natuurk., Amsterdam.* (1918) 26: 1201–1208.
5. R. P. Kerr, "Gravitational field of a spinning mass as an example of algebraically special metrics," *Phys. Rev. Lett.* **11** (1963), 237-238. doi:10.1103/PhysRevLett.11.237
6. E. T. Newman, R. Couch, K. Chinnapared, A. Exton, A. Prakash and R. Torrence, *J. Math. Phys.* **6** (1965), 918-919 doi:10.1063/1.1704351
7. D. C. Robinson, "Uniqueness of the Kerr black hole," *Phys. Rev. Lett.* **34** (1975), 905-906 doi:10.1103/PhysRevLett.34.905
8. P. O. Mazur, "Proof of Uniqueness of the Kerr-Newman Black Hole Solution," *J. Phys. A* **15** (1982), 3173-3180 doi:10.1088/0305-4470/15/10/021

9. B. K. Harrison, K. S. Thorne, M. Wakano, and J. A. Wheeler, "Gravitation Theory and Gravitational Collapse (Univ. of Chicago Press, Chicago, 1965).
10. J. D. Bekenstein, "Nonexistence of baryon number for static black holes," *Phys. Rev. D* **5** (1972), 1239-1246 doi:10.1103/PhysRevD.5.1239
11. J. D. Bekenstein, "Nonexistence of baryon number for black holes. ii," *Phys. Rev. D* **5** (1972), 2403-2412 doi:10.1103/PhysRevD.5.2403
12. S. W. Hawking, "Breakdown of Predictability in Gravitational Collapse," *Phys. Rev. D* **14** (1976), 2460-2473 doi:10.1103/PhysRevD.14.2460
13. S.W. Hawking and J.B. Hartle, "Energy and angular momentum flow into a black hole," *Commun. Math. Phys.* **27** (1972), 283-290 doi:10.1007/BF01645515.
14. J.D. Bekenstein, "Black holes and entropy," *Phys. Rev. D* **7** (1973), 2333-2346 doi:10.1103/PhysRevD.7.2333.
15. J. M. Bardeen, B. Carter and S. Hawking, "The Four laws of black hole mechanics," *Commun. Math. Phys.* **31** (1973), 161-170 doi:10.1007/BF01645742
16. S.W. Hawking, "Particle Creation by Black Holes," *Commun. Math. Phys.* **43** (1975), 199-220 doi:10.1007/BF02345020
17. S.W. Hawking, "Black hole explosions," *Nature* **248** (1974), 30-31 doi:10.1038/248030a0
18. G.W. Gibbons and S.W. Hawking, "Action Integrals and Partition Functions in Quantum Gravity," *Phys. Rev. D* **15** (1977), 2752-2756 doi:10.1103/PhysRevD.15.2752
19. J. W. York, Jr., "Role of conformal three geometry in the dynamics of gravitation," *Phys. Rev. Lett.* **28** (1972), 1082-1085 doi:10.1103/PhysRevLett.28.1082
20. G. W. Gibbons and M. J. Perry, "Black Holes and Thermal Green's Functions," *Proc. Roy. Soc. Lond. A* **A358** (1978), 467-494 doi:10.1098/rspa.1978.0022
21. R. E. Peierls, "The Commutation laws of relativistic field theory," *Proc. Roy. Soc. Lond. A* **A214** (1952), 143-157 doi:10.1098/rspa.1952.0158
22. G. J. Zuckerman, "Action Principles and Global geometry," in "Mathematical Aspects of String Theory," ed S. T. Yau, World Scientific, Singapore, 1986. doi:10.1142/9789812798411.0013
23. C. Crnkovic and E. Witten, "Covariant Description of Canonical Formalism in Geometrical Theories," in eds S. W. Hawking and W. Israel, "Three Hundred Years of Gravitation," Cambridge University Press, 1989.
24. J. Lee and R. M. Wald, "Local symmetries and constraints," *J. Math. Phys.* **31** (1990), 725-743 doi:10.1063/1.528801
25. R. M. Wald, "Black hole entropy is the Noether charge," *Phys. Rev. D* **48** (1993) no.8, 3427-3431 doi:10.1103/PhysRevD.48.R3427 [arXiv:gr-qc/9307038 [gr-qc]].
26. R. M. Wald and A. Zoupas, "A General definition of 'conserved quantities' in general relativity and other theories of gravity," *Phys. Rev. D* **61** (2000), 084027 doi:10.1103/PhysRevD.61.084027 [arXiv:gr-qc/9911095 [gr-qc]].

27. A. Ashtekar and M. Streubel, “Symplectic Geometry of Radiative Modes and Conserved Quantities at Null Infinity,” *Proc. Roy. Soc. Lond. A* **A376** (1981), 585-607 doi:10.1098/rspa.1981.0109
28. J. D. Brown and J. W. York, Jr., “Quasilocal energy and conserved charges derived from the gravitational action,” *Phys. Rev. D* **47** (1993), 1407-1419 doi:10.1103/PhysRevD.47.1407 [arXiv:gr-qc/9209012 [gr-qc]].
29. A. Komar, “Positive-Definite Energy Density and Global Consequences for General Relativity,” *Phys. Rev.* **129** (1963) no.4, 1873 doi:10.1103/PhysRev.129.1873
30. S. W. Hawking, M. J. Perry and A. Strominger, “Soft Hair on Black Holes,” *Phys. Rev. Lett.* **116** (2016) no.23, 231301 doi:10.1103/PhysRevLett.116.231301 [arXiv:1601.00921 [hep-th]].
31. J. D. Brown and M. Henneaux, “Central Charges in the Canonical Realization of Asymptotic Symmetries: An Example from Three-Dimensional Gravity,” *Commun. Math. Phys.* **104** (1986), 207-226 doi:10.1007/BF01211590
32. A. Castro, A. Maloney and A. Strominger, “Hidden Conformal Symmetry of the Kerr Black Hole,” *Phys. Rev. D* **82** (2010), 024008 doi:10.1103/PhysRevD.82.024008
33. S. Haco, S. W. Hawking, M. J. Perry and A. Strominger, “Black Hole Entropy and Soft Hair,” *JHEP* **12** (2018), 098 doi:10.1007/JHEP12(2018)098 [arXiv:1810.01847 [hep-th]].
34. J. L. Cardy, “Operator Content of Two-Dimensional Conformally Invariant Theories,” *Nucl. Phys. B* **270** (1986), 186-204 doi:10.1016/0550-3213(86)90552-3
35. S. Haco, M. J. Perry and A. Strominger, “Kerr-Newman Black Hole Entropy and Soft Hair,” [arXiv:1902.02247 [hep-th]].
36. M. Perry and M. J. Rodriguez, “Central Charges for AdS Black Holes,” [arXiv:2007.03709 [hep-th]].