Second-order phase transitions in black-hole thermodynamics



1980 J. Phys. A: Math. Gen. 13 1113

(http://iopscience.iop.org/0305-4470/13/3/043)

J. Phys. A: Math. Gen. 13 (1980) 1113–1120. Printed in Great Britain

# Second-order phase transitions in black-hole thermodynamics

L M Sokołowski† and P Mazur‡

† Astronomical Observatory of the Jagellonian University, Orla 171, 30–244 Kraków, Poland

‡ Institute of Physics, Jagellonian University, Reymonta 4, 30–059 Kraków, Poland

Received 22 June 1979

**Abstract.** The physical meaning of second-order phase transitions discovered by Davies in black-hole thermodynamics is discussed. It is argued that the phenomenon has nothing to do with phase transitions occurring in rotating self-gravitating fluids in Newtonian theory. We show that the internal state of a black hole remains unaffected after the phase transition. We also argue that the change of sign and infinite discontinuity in the heat capacity are of purely geometric origin, i.e. are determined by the embedding of an event horizon into a black-hole space–time. This result supports the view that temperature cannot be used as a well-behaved fundamental parameter in black-hole physics.

## 1. Introduction

The very remarkable discovery of thermal emission of elementary particles by Schwarzschild black holes has initiated deeper investigations of thermodynamic properties of stationary rotating and charged (i.e. Kerr–Newman) black holes. Those investigations involved conditions for stable equilibrium of non-rotating black holes with a thermal radiation bath (Hawking 1976, Davies 1977, Huţ 1977, Gibbons and Perry 1978, see also review articles by Sciama 1976 and Davies 1978), the fluctuation–dissipation theorem in irreversible thermodynamics (Candelas and Sciama 1977), the black-hole version for the third law of thermodynamics and various specific heats of black holes held in equilibrium (Davies 1977, Hut 1977). It has turned out that the black-hole thermodynamics differs from the theory of normal thermodynamics in a number of details: apart from the unsolved problem of a proper definition of stable equilibrium for Kerr holes (difficulties have arisen with the rotating thermal bath), Hawking (1976) has shown that black holes cannot be described by means of a canonical ensemble (this is closely related to the fact that the black-hole entropy is a global property of the hole, since it cannot be divided up into a number of weakly interacting parts). Davies has found that the Nernst theorem is not fully satisfied for black holes—the entropy has a finite positive limit as the temperature approaches zero. Davies (1977) has also studied black-hole specific heats and has derived a formula for the thermal capacity $C_{J,Q}$ with angular momentum $J$ and charge $Q$ held constant:

$$C_{J,Q} := T\left(\frac{\partial S}{\partial T}\right) = \frac{MTS^3}{\pi J^2 + (\pi/4)Q^4 - T^2S^3} \quad (1)$$

for a generic Kerr–Newman black hole with mass $M$ held in equilibrium (assuming that equilibrium is possible at all for rotating holes immersed in a radiation field). The formula (1) differs slightly from that given in Davies (1977), for we use units with $\hbar = c = G = 1$ and $k = 1$ rather than $k = 1/8\pi$. One sees immediately that $C_{J,Q}$ varies from the negative value $-8\pi M^2$ for a Schwarzschild black hole to zero for the extreme Kerr–Newman hole ($T \to 0$), but it approaches zero through positive values. It means that for certain values of $J$ and $Q$ the denominator in (1) vanishes. In other words, at this critical point $C_{J,Q}$ suffers an infinite discontinuity, after which it changes sign from negative to positive. Putting $J^2 = \alpha M^4$ and $Q^2 = \beta M^2$, Davies has obtained an equation for $\alpha$ and $\beta$ at the critical point:

$$\alpha^2 + 6\alpha + 4\beta - 3 = 0. \tag{2}$$

Introducing for a black hole with angular velocity $\Omega$ and electric potential $\phi$ at its horizon the Gibbs free energy $G := M - TS - \Omega J - \phi Q$ one sees tht it is continuous everywhere and has continuous first derivatives with respect to $T$, $\Omega$ and $\phi$. Second derivatives of $G$, however, are in general discontinuous at the critical point (which is in fact a line rather than a point on the $(\alpha, \beta)$ plane). There is, therefore, a formal similarity of this phenomenon to a second-order phase transition in ordinary thermodynamics. The phase transition occurs at $\alpha = 2\sqrt{3} - 3$ for the Kerr case ($Q = 0$) and at $\beta = \frac{3}{4}$ for the Reissner–Nordström case ($J = 0$). Above the critical point, after the phase transition, the heat capacity becomes positive and it would be easier for a black hole to attain equilibrium with a surrounding radiation thermal bath.

The physical significance of the phase transition, however, has remained unclear. The concept of a phase transition has been introduced into black-hole physics by Gibbons and Perry (1978), who have shown that in an isolated box containing a thermal radiation field at high enough energy densities a black hole would condense out. This black-hole condensation is qualitatively very similar to a process of liquid drop condensation out of a saturated vapour. One has therefore a first-order phase transition in black-hole physics and this process has clear physical meaning. Second-order phase transitions for ordinary physical systems, however, involve much more subtle processes whose physical interpretation requires a sophisticated analysis in each separate case. One can be sure that the phrase 'second-order phase transition for black holes' is merely the name of a physically rather obscure phenomenon and not a description of its essential features. Furthermore, a deeper thermodynamic investigation of the phenomenon is strongly hindered since an underlying statistical-mechanical theory that would describe black hole processes is lacking. By applying a phenomenological approach based on thermodynamic interpretation of the four laws of the black-hole mechanics (i.e. entropy $S \propto$ area of event horizon $A$, temperature $T \propto$ surface gravity $\mathscr{K}$) one would not gain much insight into the physical content of the problem.

Davies (1977, 1978) has suggested that for a Kerr black hole the phenomenon under consideration may be a relativistic black-hole counterpart of a phenomenon which is known for a non-relativistic rotating self-gravitating fluid (Bertin and Radicati 1976). Such a fluid loses its axial symmetry and becomes unstable when the ratio of its rotational energy to its Newtonian gravitational energy exceeds some critical value (about 0·14). The change has the form of a thermodynamic second-order phase transition and in the higher angular-momentum phase the fluid forms a triaxial ellipsoid. Davies' suggestion was therefore: 'This suggests the possibility that in a similar fashion the Kerr black hole becomes unstable at $\alpha^{1/2} \simeq 0{\cdot}68$, and enters a non-axisymmetric dynamic phase in which new thermodynamic degrees of freedom appear (Davies 1977).

In this paper we present a view that the black-hole transitions ought to be explained in a quite different way, without relating them to phase transitions in rotating fluids. First, in our opinion, it is not quite clear what the 'unstable dynamic phase' would in fact mean. It is known that Kerr metrics are stable against small perturbations and one can believe with fairly good confidence that at the critical point there would be no bifurcation and no new classes of solutions of Einstein's field equations would appear. The second, much stronger argument is that the same phase transition occurs for non-rotating, charged black holes and this process has no counterpart for non-relativistic charged bodies. A spherically-symmetric static body can in principle be charged to arbitrarily high values (at least above the black hole limit $Q^2 = M^2$ which is greatly exceeded in nuclear matter) and no second-order phase transition appears. This means that physical processes occurring in material bodies which are sources of gravitational fields representing the asymptotic non-relativistic limits of the Kerr–Newman fields cannot yield much insight into physical processes occurring in black holes themselves. The black-hole phase transition has been discovered by thermodynamic considerations, but taking into account that all the thermodynamic quantities are of geometric origin, one ought to seek a geometric interpretation of the transition. We show that the geometry of the black hole does not suffer any qualitative change at the critical point, which in turn suggests that there are also no significant changes in physical properties of it. Assuming that the phase transition has a purely geometric origin we begin with a discussion of the intrinsic geometry of an event horizon. Consider first the intrinsic geometry of a two-dimensional cross-section of the horizon (i.e. of the surface of a black hole). This problem has been studied in detail by Smarr (1973) who has proven that the metric ${}^{(2)}\mathrm{d}s^2$ of the two-slice and hence its geometry are uniquely determined by two parameters:

$$\eta := (r_+^2 + a^2)^{1/2} \qquad \text{and} \qquad \beta_S := a(r_+^2 + a^2)^{-1/2}, \tag{2}$$

where as usual $r_+ := M + (M^2 - a^2 - Q^2)^{1/2}$ gives the location of the event horizon, and $a := J/M$, instead of three parameters determining the entire four-geometry. For non-rotating black holes ($a = \beta_S = 0$) the two-slice is isometric to a sphere with radius $\eta = r_+$. Thus the geometry of the black-hole surface singles out no value of $Q$ and, in particular, having given a value $\eta$ of the radius one cannot decide whether one measures the geometry of the two-slice for a Schwarzschild hole with mass $\frac{1}{2}\eta$ or for a charged black hole with mass $\frac{2}{3}\eta$ at the critical value of its charge. The geometry of the surface of a charged hole is unaffected by the phase transition. The same conclusion should also hold for rotating black holes, since according to our assumption the phase transitions are due to a common geometric origin. In the Kerr case, however, the situation is slightly more complicated, for the surface of a rotating hole is merely homeomorphic to a sphere. Calculations by Smarr (1973) have revealed that the Gaussian curvature becomes negative near the poles of the spheroid for $J > \frac{1}{2}\sqrt{3}\, M^2$. Detailed investigations by Press have shown that the Kerr solutions remain stable there. Note also that the value of $J$ is irrelevant for the phase transition.

Now one can ask whether the phase transition affects the three-geometry of the entire event horizon. To show that it is not so we apply an approach due to Smarr (1973). As the event horizon is also a Killing horizon we choose such coordinates $t, x^1, x^2, x^3$ that a Killing vector tangent to null generators of $H$ has the form $K = \partial/\partial t$. This restriction is invariant under transformations $x^i \to f^i(x^k)$ where $i, k = 1, 2, 3$. By using one of these transformations the equation for $H$ can be made $x^1=0$. Take $t, x^2, x^3$ as coordinates on $H$, then considering that $K$ is null on $H$ and orthogonal to $H$, it follows that the induced line element on the horizon is given by

$$^{(3)}\mathrm{d}s^2 = g_{AB}\,\mathrm{d}x^A\mathrm{d}x^B \qquad (A, B = 2, 3) \tag{3}$$

Here $g_{AB}$ depends on $x^2$ and $x^3$ only, since $\partial/\partial t$ is the Killing vector and $x^1=0$ on $H$. The surface of a black hole is given by $t=F(x^2, x^3)$ for an appropriate choice of the function $F$. One sees that regardless of the form of $F$, the line elements for the surface of the hole and for the event horizon are identical. The metric on $H$ is, of course, degenerate but it still determines all principal intrinsic geometric properties of the horizon. Once again one cannot determine from the metric tensor the exact values of the mass and charge. We can therefore summarise: the intrinsic metric geometry of the event horizon remains completely unaffected by the phase transition.

The definition of $C_{J,Q}$ involves two quantities which are to be determined on the event horizon, the area of the black hole surface $A$ and the surface gravity $\mathscr{K}$, the latter being an extrinsic quantity for the event horizon geometry. In fact one cannot determine the value of $\mathscr{K}$ from the formula (3) for $^{(3)}\mathrm{d}S^2$ but one needs to know the entire four-metric near $H$, that is one needs a 'rigging' vector for $H$ (Bardeen *et al* 1973). We arrive at a conclusion: the phase transition is due to the geometry of the space–time in the neighbourhood of the horizon, i.e. to the way in which the horizon is embedded into the space–time.

Fortunately we need not study the full line element for the Kerr–Newman black holes but what is in fact needed is to investigate the relation between the surface area of a black hole and its surface gravity. This will be done in detail in the next section.

## 2. The relation between the black-hole temperature and entropy

One of the fundamental aspects of black-hole thermodynamics is the existence of the explicit fundamental relation for the Kerr–Newman family between the energy, entropy, angular momentum and charge:

$$M = [(S/4\pi) + (\pi/S)(J^2 + \tfrac{1}{4}Q^4) + \tfrac{1}{2}Q^2]^{1/2}. \tag{4}$$

The definition of the black hole temperature is based on the first law of black-hole thermodynamics and on the quantum Hawking effect:

$$T := (\partial M/\partial S)_{J,Q} = (8\pi M)^{-1}[1 - (4\pi^2/S^2)(J^2 + \tfrac{1}{4}Q^4)]. \tag{5}$$

By substituting (4) into (5) we obtain $T$ as a function of $S$, $J$ and $Q$:

$$T = (8\pi)^{-1}[(S/4\pi) + (\pi/S)(J^2 + \tfrac{1}{4}Q^4) + \tfrac{1}{2}Q^2]^{-1/2} \cdot [1 - (4\pi^2/S^2)(J^2 + \tfrac{1}{4}Q^4)]. \tag{6}$$

For ordinary thermodynamic systems temperature is one of the most fundamental macroscopic parameters and all thermodynamic functions are expressed in terms of it. In particular, in most applications, when entropy is considered it is used in the form of a function of $T$, $S(T)$. In black-hole thermodynamics, however, one obtains from the basic principles of the theory the temperature as a function of entropy and the 'extensive' parameters $J$ and $Q$ in the form (6). To invert this equation to obtain $S(T, J, Q)$ one would have to solve an algebraic equation of the fifth order in $S$, which turns out to be impossible.

To calculate the heat capacity $C_{J,Q}$ one does not need to know the explicit form of the function $S(T)$. Equation (6) may be transformed into an equation of the form $F(T, S) = 0$. By the implicit function theorem, the unique solution $S(T)$ of the equation exists in a neighbourhood of a point $(T, S)$ and is differentiable iff the derivative $\partial F/\partial S \neq 0$ at this point. This is how $C_{J,Q}$ has been calculated. The critical point at which the phase transition occurs is that at which the theorem is no longer valid, since $\partial F/\partial S = 0$. At this point the equation $F(T, S) = 0$ does not possess a unique local solution with respect to $S$, and the concept of the heat capacity calculated by means of it loses its validity. This suggests that it would not be plausible to view the infinite discontinuity in $C_{J,Q}$ as a real second-order phase transition, basing the entire argument merely on the fact that $S$ and $T$ remain finite and continuous there. To understand better the physical meaning of the occurence of the infinite discontinuity in $C_{J,Q}$ apart from the mathematical one mentioned above, we study the equation (6) in more detail. We are interested in processes in which $J$ and $Q$ are held constant and mass is variable, therefore we begin with an extreme Kerr–Newman black hole of charge $Q$ and angular momentum $J$. Its mass is

$$M_e = \tfrac{1}{2}Q^2 + (J^2 + \tfrac{1}{4}Q^4)^{1/2}. \tag{7}$$

We shall keep $M_e$ constant throughout the paper, fixed at some value; the equation (7) determines a two-surface in the $(M, J, Q)$ space of black-hole states. For the given value of $M_e$, $J$ and $Q$ may vary in some intervals; each value of $Q$ corresponds to a straight line, parallel to the $M$ axis, representing a family of black holes with the same $J$ and $Q$.

For each family of black holes the smallest hole is the extreme one, i.e. with mass $M_e$. Then a hole mass is increased by allowing energy inflow through the horizon, but without changing the angular momentum and charge. We can use $(S, J, Q)$ instead of $(M, J, Q)$ as independent parameters since $S$ is a monotonically increasing function of $M$. Introducing dimensionless parameters

$$\theta := 2^{5/2}\pi M_e T, \qquad \sigma := \frac{S}{2\pi M_e^2}, \qquad q := Q^2/2M_e^2 \tag{8}$$

we replace the relation (6) by

$$\theta = \sigma^{-3/2} \cdot [\sigma^2 - (1-q)^2][\sigma^2 + 2q\sigma + (1-q)^2]^{-1/2}. \tag{9}$$

Here the charge parameter $q$ varies from 0 to 0·5 and the entropy parameter $\sigma$ is greater than $1-q$ to yield a positive value of $\theta$. Thus all black holes are characterised by three independent parameters $M_e$, $q$ and $\sigma$. The function $\theta(\sigma)$ grows from zero for $\sigma = 1-q$ and for very large values of mass vanishes as $\sigma^{-1/2}$; it therefore has a maximum for a value $\sigma_0(q)$ which may be determined from the condition for an extremum:

$$\sigma^4 - \sigma(1-q)^2\sigma^2 - 8(1-q)^2 q\sigma - 3(1-q)^4 = 0. \tag{10}$$

It may be shown analytically that this equation has exactly one root $\sigma_0$ greater than $1-q$.

At the point of maximal temperature a black hole has a mass $M_0(q)$ which is uniquely determined by $M_e$ and $q$:

$$(M_0/M_e)^2 = (2\sigma_0)^{-1}[(\sigma_0 + q)^2 + 1 - 2q]. \tag{11}$$

The point of maximal temperature is certainly the critical point where the phase transition occurs, since $dT/dS = 0$ there. The relation between $J$ and $Q$ and the mass $M_0$ at the critical point may be expressed by means of the dimensionless parameters $\alpha$ and $\beta$:

$$J^2 =: \alpha M_0^4, \qquad Q^2 =: \beta M_0^2$$

then

$$\alpha = (1-2q)\,(M_0/M_e)^{-4}, \qquad \beta = 2q\,(M_0/M_e)^{-2}. \tag{12}$$

The results of computer calculations are given in figures 1 and 2. Figure 1 presents the function $\theta(\sigma)$ for different values of $q$. One sees that the sign of $C_{J,Q}$, as determined by the derivative $d\theta/d\sigma$, is positive for $\sigma < \sigma_0$ and negative above. This behaviour of $C_{J,Q}$ together with its infinite discontinuity is due to the double-valuedness of the $(T, J, Q)$ coordinates—for each set $(T, J, Q)$ there exist two values of $S, M$ and other thermodynamic quantities. It is obvious that the black-hole temperature cannot be used as a fundamental parameter of the system; on the contrary, it should be viewed as a function of the basic parameters $(M, J, Q)$.

The simplest resolution of the problem of the black-hole phase transition would therefore be that the phenomenon is due merely to the fact that the temperature is not a monotonic function of the entropy. In other words, the dynamics of the gravitational field via the Einstein equations establishes the geometry of a stationary, vacuum (except for the electromagnetic field) black hole in the form of Kerr–Newman metrics, and given the metric one can calculate all quantities describing a hole. The dependence of the surface gravity $\mathscr{H}$ on $S$, $J$ and $Q$ is then the result of a direct calculation, needing no comparison with rotating or charged material bodies. This answer may, however, seem unsatisfactory, as one might argue that a full elucidation of the problem should involve

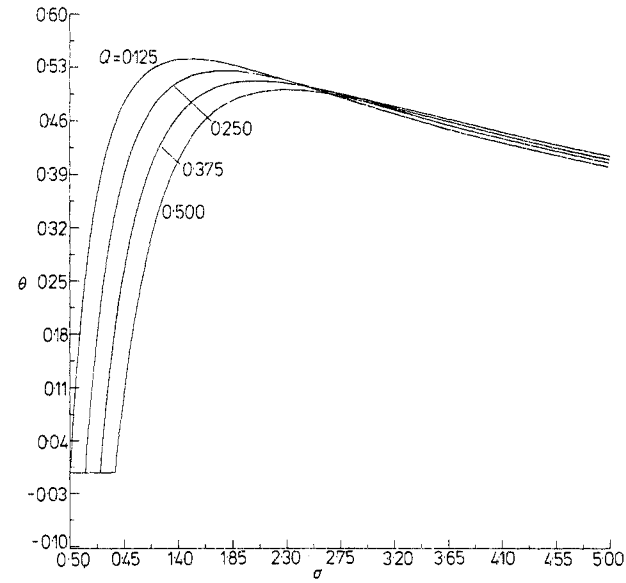


**Figure 1.** The function of $T(S)$ in dimensionless units is drawn for four values of $Q$. For each value of $(T, J, Q)$ there are two values of $S$ and $M$. At the points of maximal temperature $C_{J,Q}$ suffers an infinite discontinuity and changes its sign. Dependence of the temperature on the black-hole mass, $T(M)$ has qualitatively the same character.

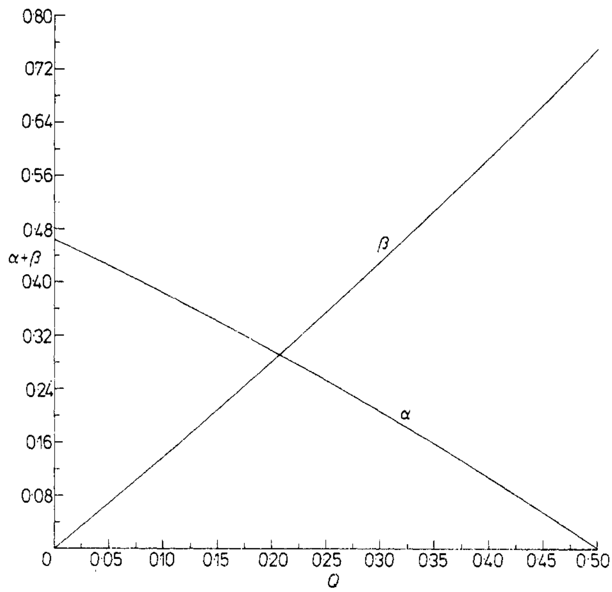


**Figure 2.** The values of $J^2$ and $Q^2$ (again in dimensionless units) are shown at the critical points where the phase transitions occur.

the physical meaning of behaviour of $\mathscr{H}$ for varying masses—why the temperature vanishes for an extreme black hole and in the limit of infinite mass. The latter problem cannot be solved within the Kerr–Newman family of solutions—in this case the explicit form of the metric determines $T$. On the contrary, one should consider a generic stationary black hole surrounded by arbitrary matter and study its thermal properties. This is however very difficult since a generic black hole is described by an unknown number of parameters. We are therefore left with merely heuristic arguments, stating in the case of a Kerr black hole that a reduction of the surface gravity in the process of increasing angular momentum, while keeping the mass constant, is due to an increment of the centrifugal force.

In our opinion, Davies' second-order phase transition for black holes may be explained by studying the manner in which an event horizon is embedded into the Kerr–Newman space–time, and in this sense the phenomenon is said to be of purely geometric origin. However, a full elucidation of the problem is still missing. Qualitatively the black-hole temperature-dependence on mass is the same as in figure 1. As far as we know there are no physical systems, real or theoretical, for which temperature depends on internal energy in this way. Ordinary thermodynamics therefore gives no analogues for black-hole physics and it seems that black holes are unique systems possessing maximal temperature and positive and negative specific heats.

Some of our conclusions are in part implicitly contained in Hut's paper (1977): 'At $Q = \frac{1}{2}\sqrt{3}M$ the heat capacity has an infinite discontinuity. Although this does not affect the internal state of the system as in the case of a phase transition, it is physically important: it indicates a transition from a region where only a microcanonical ensemble is appropriate to a region where a canonical ensemble can also be used to describe the system'. The last part of the statement seems to be rather doubtful, as positiveness of the heat capacity is merely a necessary but not a sufficient condition for a canonical ensemble to exist. Given the statistical weight (i.e. the density of energy states) for a system one can construct a canonical ensemble for it, provided that the integral representing the partition function is convergent. As Hawking (1976) has shown, the partition–function integral for black holes diverges exponentially. His proof holds in the negative specific heat region as well as in the positive one, since in both cases the partition function diverges approximately as $\exp(M^2)$. Thus black holes can be described only by means of a microcanonical ensemble.

## Acknowledgments

We are very grateful to Professors M Demiański and A Staruszkiewicz for very helpful discussions and encouragement. We should also like to thank Dr J Karczmarczuk for performing the computer calculations.

## References

Bardeen J, Carter B and Hawking S W 1973 *Commun. Math. Phys.* **31** 161
Bertin G and Radicati L 1976 *J. Astrophys.* **206** 815
Candelas P and Sciama D W 1977 *Phys. Rev. Lett.* **38** 1372
Davies P C W 1977 *Proc. R. Soc.* A**353** 499
—— 1978 *Rep. Prog. Phys.* **41** 1313
Gibbons G W and Perry M J 1978 *Proc. R. Soc.* A**358** 467
Hawking S W 1976 *Phys. Rev.* D**13** 191
Hut P 1977 *Mon. Not. R. Astr. Soc.* **180** 379
Sciama D W 1976 *Vistas Astr.* **19** 385
Smarr L 1973 *Phys. Rev.* D**7** 289