PHYSICAL REVIEW LETTERS

VOLUME 51 | 11 JULY 1983 | NUMBER 2

Instability of Some Black Holes

Steven Detweiler
Physics Department, University of Florida, Gainesville, Florida 32611

and

Roger Ove
Physics Department, Yale University, New Haven, Connecticut 06520
(Received 21 March 1983)

It has been found that some charged, rotating black holes are unstable to the growth of a small, perturbing, massless scalar field. Additionally, there are strong indications that some uncharged black holes will be unstable to gravitational and electromagnetic perturbations as well.

PACS numbers: 04.30.+x, 97.60.Lf

The stability of black holes has been examined by a number of researchers. Analytical work shows that both charged and uncharged, nonrotating black holes are stable.[1-3] Furthermore, a combination of analytical and numerical work has, until the present, failed to find an instability.[4-7] But, a proof of stability has not yet been found.

In this paper we show in fact that some black holes are unstable. We imagine that there is some, as yet undetermined, critical amount of angular momentum, J, for a black hole of mass M. For angular momenta greater than this an instability will make some initially small perturbation grow in time until the field radiates away enough angular momentum for the black hole to settle down in the stable region again. This picture is conjectural. What we do show here is that there exist some choices of a, the Kerr angular momentum parameter, and Q, the charge, such that the amplitude of an initially small perturbing field will grow exponentially in time. And there are good reasons for believing that the instability will persist even for uncharged black holes.

It is not surprising that this instability has eluded discovery for so long. It appears only for large multipoles and for rapidly rotating black holes. These two conditions make the numerical work quite difficult; and until now there seemed to be no reason for believing that the higher multipoles would be unstable first. However with the advantage of hindsight, this should not be too remarkable. After all, Friedman[8] has shown that general relativity makes *all* rotating stars unstable—the slower the rotation, the longer the growth time and the larger the multipole. In fact he suggested privately some years ago that one should be quite careful in considering the stability of black holes for large l and m.

The instability of a black hole may have important astrophysical consequences. This will depend crucially upon the e-folding time for black holes of reasonable masses and angular momenta. But, for example, an instability may be involved in the energy generation in quasars. And it is quite likely to influence both the amplitude and the wave form of the gravitational waves resulting from any black hole interaction.

Teukolsky[9] first derived the separable wave equation which governs the evolution of a spin-s field in the background geometry of a rotating black hole. One form of the radial part of this equation is

$$\Delta^{-s}\frac{d}{dr}\left(\Delta^{s+1}\frac{dR}{dr}\right)+\left(\frac{K^2+2is(r-M)K}{\Delta}-4is\omega r-\lambda\right)R=0. \tag{1}$$

The quantity $R(r)$ gives the radial dependence of the spin-s field; the time and angular dependence are assumed to be $\exp(-i\omega t+im\varphi)$. The quantity λ is an eigenvalue that comes from the θ-dependent part of the wave equation and is a function of $a\omega$; for $a\omega=0$, $\lambda=l(l+1)-s(s+1)$, where l is a spherical harmonic index. Other quantities in equation (1) are $\Delta\equiv r^2=2Mr+a^2+Q^2$ and $K\equiv(r^2+a^2)\omega-am$. Either Q or s should be set to zero in this equation—for a charged black hole this is the correct equation only for a scalar field; a separable equation for the coupled gravitational and electromagnetic perturbations of a charged, rotating black hole has not yet been found.

As $r\to\infty$ the general asymptotic solution to equation (1) is

$$R\sim(r_+/r)^{2s+1}Z_{\rm in}\exp(-i\omega r_*)+(r_+/r)Z_{\rm out}\exp(i\omega r_*), \tag{2}$$

where r_* is defined by $dr_*/dr=(r^2+a^2)/\Delta$.

The event horizon is at r_+ the greater of the two roots of $\Delta\equiv0$, $r_\pm=M\pm(M^2-a^2-Q^2)^{1/2}$. There we impose the boundary condition that radiation only go into the black hole,

$$R\sim\exp(-ikr_*), \tag{3}$$

where $k\equiv\omega-am/(r_+^2+a^2)=\omega-\omega_{\rm crit}$.

A normal mode is a solution of Eq. (1) which satisfies the boundary condition (3) and in addition has $Z_{\rm in}=0$.[4,10] Thus the normal modes have radiation only going out at infinity and only going down the black hole.

The search for the normal modes constitutes an eigenvalue problem for the complex resonant frequency ω. If $\mathrm{Im}\omega<0$ then the time dependence $\exp(-i\omega t)$ implies that the wave dissipates as it radiates away; this is a *stable* normal mode. However, if $\mathrm{Im}\omega>0$ then the amplitude of the wave grows in time and the mode is *unstable*.

For a black hole with nearly the maximum amount of charge and angular momentum,

$$\sigma\equiv(r_+-r_-)/r_+\ll1, \tag{4}$$

and for a frequency close to the critical frequency of superradiance,

$$\tau\equiv(r_+^2+a^2)k/2r_+\ll1, \tag{5}$$

Eq. (1) is amenable to analytical methods.[7] We use the notation of Teukolsky and Press[11] throughout this paper and, while later we rely heavily upon their results, here we give only a sketchy outline of their methods.

When $x\equiv(r-r_+)/r_+\ll1$ Eq. (1) has a solution which satisfies boundary condition (3) in terms of a hypergeometric function. And when $x\gg\mathrm{Max}(\sigma,\tau)$ the general solution is the sum of two confluent hypergeometric functions. In the region when $\mathrm{Max}(\sigma,\tau)\ll x\ll1$ the different solutions can be matched together: explicit formulas for $Z_{\rm in}$ and $Z_{\rm out}$ result. Specifically

$$Z_{\rm in}=\frac{2\alpha\sigma^{2s+1}(e^{i\pi}/2i\hat\omega\sigma)^{1/2+s+\alpha+2i\omega}\Gamma^2(2\alpha)\Gamma(1+s-4i\tau/\sigma)}{\Gamma(\frac12-s+\alpha-2i\hat\omega)\Gamma(\frac12+s+\alpha-2i\hat\omega)\Gamma(\frac12+\alpha+2i\hat\omega-4i\tau/\sigma)}+(\alpha\to-\alpha), \tag{6}$$

where $\hat\omega\equiv\omega r_+$, $\alpha^2=\lambda+(s+\frac12)^2-4\hat\omega^2$, and $(\alpha\to-\alpha)$ means take the first term with α replaced by $-\alpha$. For a resonant frequency $Z_{\rm in}$ vanishes, which is equivalent to

$$\left(\frac{e^{i\pi/2}}{2\hat\omega\sigma}\right)^{2\alpha}\frac{\Gamma^2(2\alpha)\Gamma(\frac12-s-\alpha-2i\hat\omega)\Gamma(\frac12+s-\alpha-2i\hat\omega)\Gamma(\frac12-\alpha+2i\hat\omega-4i\tau/\sigma)}{\Gamma^2(-2\alpha)\Gamma(\frac12-s+\alpha-2i\hat\omega)\Gamma(\frac12+s+\alpha-2i\hat\omega)\Gamma(\frac12+\alpha+2i\hat\omega-4i\tau/\sigma)}=1. \tag{7}$$

For large values of α this equation usually has no solution because the numerator is much bigger than the denominator. However $\Gamma(-2\alpha)$ has a simple pole when 2α is an integer; near this special case there are some solutions for τ which correspond to resonant frequencies.

To simplify Eq. (7) we make some further assumptions about the relative sizes of different quantities. The real parts of ω and α are assumed to be much larger than their imaginary parts. Then with the assumption that $\tau/\sigma\gg\mathrm{Max}(\alpha,\hat\omega)$ the two Γ functions containing $-4i\tau/\sigma$ are evaluated with Stirling's formula with a branch cut in the complex τ plane along the negative imaginary axis. The reflection formula is used on the remaining Γ functions containing a $-\alpha$ to obtain

$$(e^{i\pi}/8\hat{\omega}\tau)^{2\alpha}4\alpha^2\Gamma^4(2\alpha)\sin^2(2\pi\alpha)$$
$$=|\Gamma(\tfrac{1}{2}-s+\alpha-2i\hat{\omega})\Gamma(\tfrac{1}{2}+s+\alpha-2i\hat{\omega})|^2\sin[\pi(\tfrac{1}{2}-s-\alpha-2i\hat{\omega})]\sin[\pi(\tfrac{1}{2}+s-\alpha-2i\hat{\omega}]. \quad (8)$$

With the additional assumption that $\alpha, \hat{\omega} \gg 1$ Eq. (8) simplifies further to

$$4\sin^2(2\pi\alpha)=\exp(-4\pi i\alpha)\left(\frac{\tau\hat{\omega}(\alpha^2+4\hat{\omega}^2)}{2\alpha^4}\right)^{2\alpha}\exp(4\alpha+4\pi\hat{\omega})\left(\frac{\alpha-2i\hat{\omega}}{\alpha+2i\hat{\omega}}\right)^{-4i\hat{\omega}}. \quad (9)$$

We now make the final and most critical assumption, the justification of which we will give later, that for $\omega=\omega_{\text{crit}}$, $\alpha=j/2+\epsilon$ with j an integer and ϵ a small real number. Thus for frequencies near ω_{crit} a Taylor-series expansion for α about the point $\tau_{j/2}$ at which $\alpha(\tau_{j/2})=j/2$ gives $\alpha=j/2+\alpha'\Delta\tau$, where $\Delta\tau=\tau-\tau_{j/2}$ and $\alpha'=[2r_+^2/(r_+^2+a^2)](d\alpha/d\omega)_{\tau_{j/2}}$. An examination of the angular eigenvalue problem shows that $\alpha'<0$ for $m>0$.

Two solutions of Eq. (9) for τ which are consistent with all of the previous assumptions are

$$\tau=-\epsilon/\alpha'\pm\Delta\tau, \quad (10)$$

where

$$\Delta\tau\approx\frac{1}{4\pi\alpha'}\left(\frac{-2\epsilon\hat{\omega}(j^2+16\hat{\omega}^2)}{\alpha' j^4}\right)^{j/2}\exp(j+2\pi\hat{\omega})\left(\frac{j-4i\hat{\omega}}{j+4i\hat{\omega}}\right)^{-2i\hat{\omega}}, \quad (11)$$

where $\hat{\omega}$ may be assumed to be $\hat{\omega}_{\text{crit}}$ for the evaluation of the right-hand side. For j an even integer, these two values of τ correspond to two normal modes on the real axis with one on either side of that special frequency which has $\alpha=j/2$ precisely. If $\epsilon>0$ these modes are outside the superradiant region; if $\epsilon<0$ then they are inside. Whether these modes are really stable or unstable is determined by higher-order corrections which have been ignored by our myriad assumptions. But the existence of these modes, at least close to the real axis, is certain.

For j an odd integer and $\epsilon>0$ the situation is the same as for j even. But for j odd and $\epsilon<0$, the two modes are just above and just below, in the complex plane, the special frequency which has $\alpha=j/2$ precisely. The mode below is stable, the one above unstable. In addition, both frequencies have a real part less than ω_{crit}, so that the unstable mode is inside the superradiant region as it must be.[4,5]

To demonstrate the existence of unstable modes it remains to show that α is sufficiently close to a half-integer for some choice of parameters. We consider the sequence of black holes with $\sigma=(r_+-r_-)/r_+$ fixed and extremely small. For $a=0$ (a Reissner-Nordstrom black hole close to the limit where $Q=M$), $\alpha^2=(l+\frac{1}{2})^2$ with $l\geq|M|$. Analysis of the angular eigenvalue problem shows that, as long as $\alpha^2>0$, α decreases with increasing $a\omega$ and that, for $\omega=\omega_{\text{crit}}$ and $l=m$, α^2 goes through zero before $a=M$. So along this sequence α passes through half-integral values $l+1$ times. And when it does so, at a value large enough that the approximations leading to Eq. (11) are valid, then a mode is unstable.

A summary of the inequalities which are sufficient for Eq. (11) to be valid is

$$\sigma\ll\hat{\omega}\sigma\ll|\tau|\approx|\epsilon|\ll\hat{\omega}^{-1}\approx\alpha^{-1}\ll 1. \quad (12)$$

In addition, sufficient conditions for the matching of the solutions of Eq. (1) to derive Eq. (6) are that there exist a range of values of x such that $|\hat{\omega}\epsilon|\ll|\hat{\omega}\epsilon|^{1/2}\ll\hat{\omega}x\ll 1$ and that there exist a different range of values of x such that $|\omega\epsilon|\ll\omega x\ll|\omega\epsilon|^{1/2}\ll 1$.[12] It is clear that for sufficiently small σ and sufficiently large α and ω it is possible to find an ϵ such that all of these conditions are satisfied.

We have made estimates of the higher-order corrections to the value of $\Delta\tau$, given in Eq. (11), which are caused by the approximations summarized above. The corrections are always smaller than $\Delta\tau$ by the order of one of the small quantities. Hence, the higher-order corrections will not stabilize the unstable mode.

We have shown that a massless, scalar field in the Kerr-Newman geometry has some unstable modes for some choices of the parameters a and Q near the limit when $M^2-a^2-Q^2\to 0$. And we would be suprised if the instabilities did not persist in the extreme Kerr limit, when $Q=0$ and $a\to M$. After all, Eq. (1) can be considered at least formally when $Q^2<0$ and $a^2>M^2$. Then, along the sequence of black holes with σ^2 fixed, the point where Q^2 vanishes is not the least bit special as far as Eq. (1) is concerned. Furthermore, when α passes through a half-integral value along this sequence, the two modes with j odd move along the real axis and converge toward ω_{crit} until the inequalities (12) are not satisfied. When $\epsilon < 0$ the modes emerge from ω_{crit} with one on either side of the real axis, and the imaginary parts of the frequencies grow as $|\epsilon|^\alpha$ until ϵ is too large for our assumptions to be valid. So when the instability is lost from sight it is moving rapidly away from the real axis, and there is no apparent reason why the mode should become stable again before $Q = 0$. We also expect gravitational and electromagnetic perturbations of the Kerr metric to be unstable, because in the limit of large α and ω the dependence of Eq. (6) on the spin is unimportant.

It remains to be seen whether the instability which we have found can be of any astrophysical significance. An estimate of $\Delta\tau$ from Eq. (11) shows that the time scale for growth of the instability is approximately $(\alpha/\epsilon)^\alpha M$. With the inequalities (12) in mind we might choose α as small as 10 and ϵ as large as 0.01. Even with these choices the growth time is of order $10^{30}M$ —many times the age of the universe for a solar-mass black hole. But, as we pointed out above, it is our analysis that fails for shorter growth times, and it seems quite reasonable to believe that there are instabilities with much smaller growth times. We are proceeding with further numerical work in a search for these instabilities.

This research was supported in part by the National Science Foundation through Grants. No. PHY-8300190 and No. PHY-8204995.

[1]T. Regge and J. A. Wheeler, Phys. Rev. 108, 1063 (1957).

[2]V. C. Vishveshwara, Phys. Rev. D 1, 2870 (1970).

[3]V. Moncrief, Phys. Rev. D 10, 1057 (1974).

[4]W. H. Press and S. A. Teukolsky, Astrophys. J. 185, 649 (1973).

[5]S. Detweiler and J. R. Ipser, Astrophys. J. 185, 675 (1973).

[6]J. M. Stewart, Proc. Roy. Soc. London, Ser. A 344, 65 (1975).

[7]A. A. Starobinsky, Zh. Eksp. Teor. Fiz. 64, 48 (1973) [Sov. Phys. JETP 37, 28 (1973)].

[8]J. L. Friedman, Commun. Math. Phys. 62, 247 (1978).

[9]S. A. Teukolsky, Astrophys. J. 185, 635 (1973).

[10]S. Detweiler, Astrophys. J. 239, 292 (1980).

[11]S. A. Teukolsky and W. H. Press, Astrophys. J. 193, 443 (1974).

[12]To obtain Eq. (6) the asymptotic matching must be done twice, once for each of the two independent solutions of Eq. (1). The regions where the matching is done is different so that we end up with two conditions for the existence of matching regions.