Commun. math. Phys. 63, 243—255 (1978)

Communications in
Mathematical
Physics

Ergosphere Instability*

John L. Friedman

Physics Department, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA

Abstract. We consider stationary asymptotically flat spacetimes having an ergosphere but with no horizon. In the framework of linear perturbation theory such configurations are unstable or marginally unstable to scalar and electromagnetic perturbations.

I. Introduction

Outside the event horizon of any rotating black hole is a region in which no physical object can remain at rest as seen by an inertial observer at infinity: all timelike trajectories rotate with the black hole. Such regions, called ergospheres[1], are also present in models of dense, rotating fluids [1, 2, 3, 4]. Technically an ergosphere is the part of a stationary asymptotically flat spacetime in which the Killing vector that corresponds asymptotically to time translations becomes spacelike. We shall argue here that any configuration having an ergosphere but no horizon will be unstable to scalar and electromagnetic perturbations. One expects that an object which rotates rapidly enough to acquire an ergosphere will radiate its excess angular momentum and spin down until no ergosphere remains (or, perhaps, until another more disruptive instability arises).

Spacetimes with ergospheres are also presumably unstable to gravitational perturbations. In general, however, gravitational waves couple to the source: the linearized field equations include the perturbed matter fields. Thus a stability analysis must specify the nature of the source; and in the case of greatest interest—when the source is a perfect fluid—we show in a companion paper that *all* rotating configurations are unstable (or marginally unstable) to gravitational radiation. Our considerations here will therefore be restricted to nongravitational perturbations, for which the presence of an ergosphere marks the onset of instability along a sequence of rotating equilibrium models.

* Research supported in part by the National Science Foundation under grant MPS 74-17456 with the University of Chicago and grant MPS 74-7456 at the University of Wisconsin-Milwaukee

1 The word is analogous to "atmosphere." Ergospheres are not topological spheres; ergospheres of stars, for example, are toroids

0010-3616/78/0063/0243/$02.60

In the case of scalar and electromagnetic test fields, the ergosphere instability arises in the following way. Associated with the test field's energy-momentum tensor T^{ab} and with the background Killing vector t^a is a canonical energy

$$\mathscr{E}_S = \int_S T_a^b t^a dS_b .$$

Because t^a is a spacelike within an ergosphere, initial data can be chosen on S to make $\mathscr{E}_S$ negative. But, because only positive energy can be radiated at future null infinity, the value of $\mathscr{E}_S$ can only decrease from one asymptotically null hypersurface S to another, say S', in the future of S. Furthermore, we will see that the energy can be negative (in fact nonzero) only when the test field is time dependent. Thus, unless the system can always settle down to a time dependent but nonradiative state, the energy $\mathscr{E}$ will grow without bound. If one assumes sufficient smoothness of the field in a neighborhood of null infinity, an argument based on a timelike uniqueness theorem due to Holmgren [5] rules out the first alternative and implies that the system is strictly unstable.

For axisymmetric spacetimes the instability is associated with non-axisymmetric perturbations, fields having angular dependence $e^{im\phi}$ (where ϕ is the angle about the background symmetry axis). We find that unstable (or marginally unstable) solutions to the test field equations exist for all sufficiently large values of the integer m. When the ergosphere is small, unstable modes have large values of m: along a sequence of models, the instability sets in not through a particular mode, but via the limit as $m \to \infty$ of modes having behavior $e^{im\phi}$.

The question of how rapidly the ergosphere instability is likely to grow is not dealt with here. However Comins and Schutz [4] have recently considered the problem in the case of a scalar field propagating on a background spacetime that approximates a rotating fluid. Using a JWKB method, they find (for reasonably large ergospheres) characteristic growth times long compared to the dynamical timescale but short compared to evolutionary times. Thus the instability is unlikely to play any role in collapse; but it can be used to tighten the upper mass limit on compact objects by ruling out relativistic configurations that rotate rapidly enough to have ergospheres.

In § II and § III we treat scalar and electromagnetic test fields on a background spacetime and carry through the stability argument sketched above. An appendix deals with uniqueness of the timelike initial value problem, applying Holmgren's theorem to wave equations on a curved spacetime.

II. A Test Scalar Field

Consider an asymptotically flat spacetime, M, whose metric g_{ab} admits a Killing vector t^a, $\mathscr{L}_t t_{ab} = 2\nabla_{(a} t_{b)} = 0$. The spacetime is supposed stationary: that is, near infinity t^a is timelike and has asymptotic norm $t^a t_a = -1$. There is to be no horizon, but an ergosphere – a region in which t^a is spacelike – will be present. No further symmetry assumptions need be made, so that if stationary nonaxisymmetric fluids exist in relativity, analogous to the Dedekind ellipsoids of Newtonian theory, the analysis will apply to them as well.

The first aim of this section will be to prove that all such stationary configurations with ergospheres are at least marginally unstable to scalar perturbations, in the sense that there are always perturbations which do not die away at large times; and unless there are time dependent but nonradiative perturbations, such configurations will in fact be strictly unstable, radiating infinite energy in the linearized theory. In fact, when the background geometry is axisymmetric, strict instability can be avoided only if there are time dependent but nonradiative scalar fields having angular behavior $e^{\mathrm{i}m\phi}$ for all sufficiently large integers m, where ϕ is the angle about the symmetry axis. Physically, this would indicate that real perturbations radiate away the angular momentum of the background spacetime until no ergosphere remains.

Denote by S_u a family of Killing related spacelike hypersurfaces, indexed by a scalar u with $t^a \nabla_a u = 1$. With the definition

$$\mu = (-\nabla_a \mathrm{u} \nabla^a \mathrm{u})^{1/2}, \tag{1}$$

the unit normal n_a to S_u is

$$n_a = -\mu^{-1} \nabla_a u. \tag{2}$$

The Killing vector t^a can be written in terms of n_a and an orthogonal vector in the manner

$$t^a = \mu^{-1}(n^a + \alpha k^a), \tag{3}$$

where

$$n^a n_a = -1,$$

$$k^a k_a = 1,$$

and

$$n^a k_a = 0. \tag{4}$$

The projection operator orthogonal to n^a and k^a is

$$j^a{}_b = \delta^a{}_b + 4 n^{[a} k^{c]} n_{[b} k_{c]}. \tag{5}$$

A scalar field on the background spacetime satisfies

$$\nabla^a \nabla_a \psi = 0. \tag{6}$$

Its energy-momentum tensor,

$$T^{ab} = \nabla^a \psi \nabla^b \psi - \tfrac{1}{2} g^{ab} \nabla_c \psi \nabla^c \psi, \tag{7}$$

is divergence-free

$$\nabla_b T^{ab} = 0, \tag{8}$$

and so to the Killing vector t^a corresponds a conserved current

$$J^a = T^{ab} t_b; \tag{9}$$

that is,

$$\nabla_{(a} t_{b)} = 0 \Rightarrow \nabla_a J^a = 0. \tag{10}$$

Consequently, if the field ψ has, say, compact support on a member S_0 of the family of surfaces S_u, then

$$\mathscr{E}_u \equiv \int_{S_u} J^a dS_a = \int_{S_0} J^a dS_a = \mathscr{E}_0, \tag{11}$$

for surfaces S_u near S_0.

Suppose, now, that at large distances the surfaces S_u become null, and consider fields ψ whose support on S_u may extend to null infinity. Let (u, r, θ, ϕ), with $-\infty < u < \infty$, $r > r_0$, $0 \leqq \theta < \pi$, and $0 \leqq \phi < 2\pi$ be a standard null chart for M outside a bounded region. That is, lines of constant u, θ, and ϕ are null geodesics with affine parameter r; lines of constant r, θ, and ϕ are trajectories of the Killing vector t^a; and the metric has the asymptotic behavior given, for example, by Newman and Unti [6],

$$\begin{aligned} &g^{uu} = 0, \quad g^{ur} = -1, \quad g^{u\theta} = 0, \quad g^{u\phi} = 0, \\ &g^{rr} = 1 - \frac{2M}{r} + O(r^{-2}), \quad g^{r\theta} = O(r^{-3}), \quad g^{r\phi} = O(r^{-3}) \\ &g^{\theta\theta} = \frac{1}{r^2} + O(r^{-4}), \quad g^{\theta\phi} = O(r^{-5}), \quad g^{\theta\phi} = \frac{1}{r^2 \sin^2\theta} + O(r^{-4}), \end{aligned} \tag{12}$$

characteristic of a time independent geometry. Let us consider a region $\mathscr{R}$ bounded by the surfaces S_0, S_u, and by an $r = $constant cylinder.

We have

$$0 = \int_{\mathscr{R}} \nabla_a J^a = \int_{\partial\mathscr{R}} J^a dS_a, \tag{13}$$

whence

$$\begin{aligned} \mathscr{E}_u - \mathscr{E}_0 &= -\lim_{r\to\infty} \int_0^u J^r r^2 d\Omega du \\ &= -\lim_{r\to\infty} \int_0^u \dot{\psi}^2 r^2 d\Omega du \end{aligned} \tag{14}$$

where $\dot{\psi} = t^a \nabla_a \psi$ (and where asymptotic regularity, the condition $\psi = \frac{1}{r}\psi_1(u, \theta, \phi) + o(r-1)$, has been assumed). In other words, a radiative solution loses energy between S_0 and S_u and, consequently $\mathscr{E}_u$ is a decreasing of u.

If $\mathscr{E}_0 \geqq 0$ for all initial values $(\psi, \nabla_a \psi)$ on S_0, then the symmetry implies $\mathscr{E}_u \geqq 0$ for all u; by (12), only a finite amount of energy can be radiated and the functional $\mathscr{E}_u$ is bounded by $\mathscr{E}_0$. The scalar field is consequently either strictly stable, $\mathscr{E}_u \to 0$ as $u \to \infty$[2], or at worst marginally unstable, $\mathscr{E}_u$ finite as $u \to \infty$. If, on the other hand, there is some initial data on S_0 for which $\mathscr{E}_0 < 0$, than the field is at best marginally unstable; and, unless it can settle down to a nonradiative state with fixed $\mathscr{E} < \mathscr{E}_0 < 0$, it will radiate infinite energy to null infinity. Furthermore, the energy can be negative only if the field ψ is time dependent, and thus a field for which

2 $\mathscr{E}_u \to 0$ implies that $\nabla_a \psi \to 0$ (i.e. components of $\nabla_a \psi$ along a Killing transported tetrad converge to zero); ψ itself can asymptote to a constant

$\mathscr{E}<0$ will be strictly unstable unless there are time dependent nonradiative states available to it. To see this, one manipulates the expression for $\mathscr{E}$, using the field Eq. (6) and integrating by parts to obtain the following form involving the symplectic product (Klein-Gordon inner product) of the field ψ and its time derivative $\dot{\psi}$:

$$\mathscr{E}=\tfrac{1}{2}\int_S(\dot{\psi}\nabla^a\psi-\psi\nabla^a\dot{\psi})dS_a+\tfrac{1}{2}\int_{\partial S}\psi\nabla^a\psi t^b dS_{ab}. \tag{15}$$

If ψ is a time independent solution of (6), the first term on the right hand side of this equation clearly vanishes; and asymptotic regularity (when $\dot{\psi}=0$) requires that the surface term vanish as well. Thus for time independent fields, $\mathscr{E}=0$.

It is not difficult to show that

Proposition. *$\mathscr{E}<0$ for some $(\psi,\nabla_a\psi)$ on $S\Leftrightarrow$there is an ergosphere.*

We have

$$\begin{aligned}\mathscr{E}&=\int_S T_a{}^b t^a dS_b\\ &=\int_S T_a{}^b\mu^{-1}(n^a+\alpha k^a)n_b dS\,.\end{aligned} \tag{16}$$

Now

$$\begin{aligned}\mu^{-1}T_a^b(n^a+\alpha k^a)n_b=\tfrac{1}{2}\mu^{-1}[&(n^a\nabla_a\psi)^2+2\alpha(n^a\nabla_a\psi)(k^b\nabla_b\psi)+(k^a\nabla_a\psi)^2\\ &+j^{ab}\nabla_a\psi\nabla_b\psi]\,.\end{aligned}$$

But the expression on the right-hand side is positive when $|\alpha|\leqq 1$, whence $\mathscr{E}>0$ when there is no ergosphere.

On the other hand when $|\alpha|>1$ somewhere on S, initial data for which $\mathscr{E}<0$ can be found as follows. Let Ω be an open set in the ergosphere and (t,x,y,z) a chart on Ω chosen so that curves of constant t,y,z have tangent k^a: $k^a\nabla_a f=\partial_x f$. There is some ε with $|\alpha|>1+\varepsilon>0$ on Ω and we can assume $\alpha>0$. Let $\Omega_R<\Omega$ be the ball $r^2\leqq R$ (there is such a ball about p in Ω for small R). Consider a function $\varrho\in C^\infty(\Omega)$ which vanishes outside of a compact subset of Ω, whose value and derivatives are bounded by

$$\|\varrho\|_1\equiv\operatorname*{lub}_\Omega|\varrho|+\operatorname*{lub}_\Omega|k^a\nabla_a\varrho|+\operatorname*{lub}_\Omega(j^{ab}\nabla_a\varrho\nabla_b\varrho)^{1/2}<K\,, \tag{17}$$

and with

$$\varrho=1\quad\text{on}\quad\Omega_R\,. \tag{18}$$

Then the initial data,

$$\begin{aligned}\psi_m&=\varrho\sin m\chi\\ n^a\nabla_a\psi_m&=-k^a\nabla_a\psi_m\\ &=-m\varrho\cos mx+\varrho_{,x}\sin mx\,,\end{aligned} \tag{19}$$

on S gives $\mathscr{E}_S<0$, for large enough m. That is,

$$T_a{}^b t^a n_b=-(\alpha-1)(k^a\nabla_a\psi_m)^2+\tfrac{1}{2}j^{ab}\nabla_a\psi_m\nabla_b\psi_m\,. \tag{20}$$

Denoting the bounds on μ by μ_0 and μ_1, i.e.

$$\mu_0 \leqq \mu(p) \leqq \mu_1, \quad p \in \Omega, \tag{21}$$

we have

$$\begin{aligned}\mathscr{E} = \int_\Omega T_a{}^b t^a n_b dS &\leqq -\varepsilon\mu_1^{-1} \int_{\Omega_R} (k^a \nabla_a \psi_m)^2 dS + \mu_0^{-1} \int_\Omega j^{ab} \nabla_a \psi_m \nabla_b \psi_m dS \\ &\leqq -m^2 \varepsilon \mu_1^{-1} \int_{\Omega_R} \sin^2 mx dS + \mu_0^{-1} K^2 |\Omega|\end{aligned} \tag{22}$$

where $|\Omega| \equiv \int_\Omega \mathrm{dS}$.

As $m \to \infty$, $\int_{\Omega_R} \sin^2 m\chi \to \frac{1}{2}|\Omega_R|$, so for sufficiently large m,

$$\mathscr{E} \leqq -m^2(\tfrac{1}{3}\varepsilon\mu_1^{-1}|\Omega_R|) + K^2 \mu_0^{-1}|\Omega| < 0\,. \tag{23}$$

This concludes our proof of the proposition.

Spacetimes with ergospheres are thus marginally unstable to scalar perturbations, and initial data on S with $\mathscr{E}_S < 0$ must either evolve to a time dependent but nonradiative state, or grow without bound ($|\psi|$ or $\mathscr{E}$ becomes infinite).

The first alternative appears to be ruled out by the following line of argument. Define the domain of dependence $D(T)$ of a timelike hypersurface T to be the locus of all deformations $\hat{T}$ of T with compact support on T and which are themselves timelike (see Appendix II). A uniqueness theorem due to Holmgren [5] implies that when the background spacetime is analytic, any smooth (C^∞) solution ψ to the scalar wave equation is uniquely determined by its "initial data" on T. That is, given ψ and $\nabla_a\psi$ on T, there is at most one solution to $\nabla^a\nabla_a\psi = 0$ on $D(T)$. In the case of a stationary spacetime with ergosphere, the existence of a timelike Killing vector implies that the metric is analytic outside the source and the ergosphere [7]. Moreover, supposing the whole background to be analytic does not restrict the physics that can be described. (Any background metric can be approximated with arbitrary accuracy in, say, a C^n norm by an analytic metric.) So for the remainder of this section we will assume that the background is analytic.

Now suppose that ψ were time dependent inside the ergosphere $\mathbb{E}$. Then $(\psi, \nabla_a\psi)$ cannot be zero on the whole of any distant timelike surface T with $\mathbb{E}$ in its domain of dependence $D(T)$. Moreover, if we take T tangent to the Killing field t^a, the data must be time dependent on T. This is because the Killing translated data on T gives the Killing translated solution on $D(T)$, whence ψ time independent on T implies ψ time independent on $D(T)$. Solutions to the wave equation have the asymptotic behavior

$$\psi \sim \frac{\psi_1(u, \theta, \phi)}{r} + o(r^{-1})\,. \tag{24}$$

The energy radiated to null infinity between u_1 and u_2 is $\int_{u_1}^{u_2} (\partial_u \psi_1)^2 d\Omega du$ and unless $\partial_u \psi_1 \to 0$ as $u \to \infty$, the radiated energy will be infinite.

The possibility remains, however, that although asymptotically the field settles down to a time independent state, its time derivatives remain finitely large in any spatially bounded region. This lacuna can be eliminated if one assumes that the scalar field and the null coordinate components of the metric g^{ab} can be expanded in powers of r^{-1} near future null infinity. That is, suppose that the series

$$\psi = \sum_{n=1}^{\infty} \psi_n(u,\theta,\phi) r^{-n},$$

$$g^{rr} = \sum_{n=1}^{\infty} g_n^{rr}(u,\theta,\phi) r^{-n},$$

$$g^{\theta\theta} = r^{-2} \sum_{n=1}^{\infty} g_n^{\theta\theta}(u,\theta,\phi) r^{-n}, \ldots, \tag{25}$$

converge in a region exterior to some timelike hypersurface T of the form $r = R(u,\theta,\phi)$ and in the future of S_0. Then, following Moret-Bailly and Papapetrou [8], it is easy to show that the scalar wave equation is equivalent to recursion relations of the form

$$(n-1)\partial_u \psi_n + O_1(\theta,\phi,\partial_\theta,\partial_\phi)\psi_{n-1} \pm \ldots + O_{n-1}(\theta,\phi,\partial_\theta,\partial_\phi)\psi_1 = 0, \tag{26}$$

where the O_n are linear operators of second order involving ∂_θ and ∂_ϕ with coefficients depending on the coordinates θ and ϕ. For a nonradiative field, $\partial_u \psi_1$ must vanish; the recursion relations (26) then imply that ψ_n is a polynomial in u of degree less than or equal to $n-1$. Thus either all the ψ_n are of degree zero or else some ψ_n increases without bound. If all ψ_n were degree zero in u, ψ *would* be time independent outside T. By the timelike uniqueness theorem, ψ would then be time independent everywhere to the future of some S_u, which contradicts the assumption that $\mathscr{E} < 0$. If, on the other hand, some ψ_n grew without bound, then ψ would be unbounded as well. Thus, as asserted, any bounded, time-dependent scalar field satisfying (25) must radiate.

There are in principle time-dependent sources with constant amplitude that radiate finite energy (for example, machines that change their shape with a time dependence $Q(u) \sim \sin \log u$ (as in Bardeen and Press [9]), but their time derivatives must become arbitrarily small. In the case at hand, the existence of a finite amplitude solution with arbitrarily small time derivatives would presumably again, by (15), be inconsistent with the fact that the energy is bounded away from 0 by $\mathscr{E} < \mathscr{E}_0 < 0$.

III. Electromagnetic Perturbations

The analogous demonstration that for a test electromagnetic field initial data exists for which $\mathscr{E} = \int_S T_a{}^b t^a dS_b < 0$ if and only if an ergosphere is present is provided in this section. The remaining argument is the same as that for the scalar field, Holmgren's theorem applying also to the free electromagnetic field.

A test electromagnetic field F^{ab} satisfies

$$\nabla_b F^{ab} = 0, \quad \nabla_{[a} F_{bc]} = 0, \tag{27}$$

and has the energy-momentum tensor

$$T^{ab}=F^{ac}F^{b}{}_{c}-\tfrac{1}{4}g^{ab}F_{cd}F^{cd}. \tag{28}$$

Defining the electric and magnetic field associated with the hypersurface S by

$$E^a=F^{ab}n_b, \qquad B^a=\tfrac{1}{2}\varepsilon^{abcd}n_bF_{cd}={}^*F^{ab}n_b, \tag{29}$$

we can characterize an initial data set on S as a pair of vector fields (E^a, B^a) satisfying

$$D_aE^a=0 \quad \text{and} \quad D_aB^a=0, \tag{30}$$

where the operator D_a is the covariant derivative on S. Equivalently, one could specify vector fields A^a and $\hat{A}^a$ on S with

$$A^an_a=0, \qquad \hat{A}^an_a=0\ ; \tag{31}$$

then

$$E^a=\varepsilon^{abcd}n_b\nabla_cA_d$$

and

$$B^a=\varepsilon^{abcd}n_b\nabla_c\hat{A}_d \tag{32}$$

will satisfy equation (39). We will first show that the electromagnetic energy $\mathscr{E}$ can be nonzero only if the field is time dependent[3], and will then turn to the more complicated demonstration that $E_S<0$ for some initial data on a hypersurface S if and only if the background spacetime has an ergosphere.

By defining fields $\tilde{E}_a=F_{ab}t^b$ and $\tilde{B}_a={}^*F_{ab}t^b$, one writes the energy,

$$\mathscr{E}=\int_S T_a{}^bt^adS_b \tag{33}$$

in the form

$$\mathscr{E}=\tfrac{1}{2}\int_S(\tilde{E}_aE^a+\tilde{B}_aB^a)dS. \tag{34}$$

From the second of Maxwell's equations (27) and the fact that t^a is a Killing vector follow the relations

$$\nabla_a\tilde{E}_b-\nabla_b\tilde{E}_a=\mathscr{L}_tF_{ab} \tag{35}$$

$$\nabla_a\tilde{B}_b-\nabla_b\tilde{B}_a=\mathscr{L}_t{}^*F_{ab}. \tag{36}$$

Then, using (32) to express E^a and B^a in (34) in terms of A_a and $\hat{A}_a$, we obtain

$$\mathscr{E}=\tfrac{3}{2}\int_S(A_{[a}\mathscr{L}_tF_{bc]}+\hat{A}_{[a}\mathscr{L}_t{}^*F_{bc]})dS^{abc}+\int_{\partial S}(A_{[a}E_{b]}+\hat{A}_{[a}B_{b]})dS^{ab}. \tag{37}$$

3 Alternative versions of the demonstration below are apparently known. I am indebted to R. Geroch for the one given here

When F_{ab} is time independent, asymptotic regularity requires that the surface term at infinity vanish, and since the first term in the above expression for $\mathscr{E}$ is manifestly zero, the energy $\mathscr{E}$ vanishes as well. In analogy with Eq. (15) for the scalar field, this equation requires finitely large time derivatives in order that the energy $\mathscr{E}$ be bounded away from zero.

The stability of a test electromagnetic field then depends on whether the quantity $\mathscr{E}_S$ is positive for all initial data sets on S. Using Eq. (3) and the relation

$$F^{ab} = 2n^{[a}E^{b]} + \varepsilon^{abcd}B_c n_d \tag{38}$$

we can write the integrand in (33) in terms of E^a and B^a:

$$T_a{}^b t^a n_b = \mu^{-1}[\tfrac{1}{2}(E^aE_a + B^aB_a) + \alpha\varepsilon^{ab}E_aB_b] \tag{39}$$

where $\varepsilon^{ab} = \mathrm{e}^{abcd}\mathrm{K}_c\mathrm{n}_d$ is the antisymmetric tensor in the subspace orthogonal to K^a and n^a. When there is no ergosphere, the vectors t^a and n^a are both timelike and the integrand is itself positive: this is the dominant energy condition.

Explicitly,

$$|\varepsilon^{ab}E_aB_b| \leqq (E_aE^a)^{1/2}(B_bB^b)^{1/2} \equiv EB \tag{40}$$

whence

$$T_a{}^b t^a n_b \geqq \tfrac{1}{2}[E^2 + B^2 - 2\alpha EB] > 0 \quad \text{when} \quad \alpha \leqq 1\,. \tag{41}$$

Within an ergosphere, however, the integrand (39) can be negative, and we will find initial data for which the integral $\mathscr{E}$ is negative as well. Consider as in II, an open set Ω in $\mathbb{E}$ and a chart with origin at some point $p \in \mathbb{E}$. The chart is to be spatially geodesic at p so that (writing concrete indices i, j, k, to refer to components in the chart (t, x, y, z)),

$$|g_{ij} - \eta_{ij}| < Kr^2 \tag{42}$$

for some constant K, where $\eta_{ij} = \mathrm{diag}(-1, 1, 1, 1)$; by aligning the coordinate axes at p, we can require

$$|n_i - \delta_i^t| < Kr^2 \quad \text{and} \quad |k_i - \delta_i^x| < Kr^2 \tag{43}$$

(i.e., at p, $n^i\partial_i = \partial_t$, $K^i\partial_i = \partial_x$), redefining K if necessary in order that (27) and (28) hold for a single constant K. By Eq. (3), $t^i = \mu^{-1}(n^i + \alpha k^i)$. Because $\bar{\Omega}$ is compact in $\mathbb{E}$, $\alpha > 1 + \delta$ on Ω. Similarly, as in II, μ is bounded on Ω by

$$O < \mu_0 < \mu(p) < \mu, \qquad p \in \Omega\,.$$

Writing

$$\tau = \varepsilon^{txyz}n_t\,,$$

we have by (42) that

$$|\tau - 1| < Kr^2\,. \tag{44}$$

Again, as in treating the scalar field, we will use functions of the form $\psi_m = \varrho \sin m\chi$, where ϱ vanishes outside Ω and satisfies Eqs. (17) and (18).

Here we will take $\varrho_R=\varrho\left(\frac{r}{R}\right)$, where $\varrho(s)$ is a smooth function, vanishing for $s>1$ and satisfying

$$\varrho(s)=1, \qquad s<\tfrac{1}{2}, \tag{45}$$

$$\|\varrho\|_1=\operatorname{lub}|\varrho|+\operatorname{lub}|\varrho'|<K. \tag{46}$$

It will also be convenient to introduce the following shorthand. The letter Γ will represent functions bounded on $\bar{\Omega}$ by a constant independent of the integer m. Only a finite number of such functions will be considered and so we can assume that they are all bounded by a single constant K. Then

$$\psi_m=\Gamma, \quad \partial_y\psi_m=\Gamma, \quad \partial_z\psi_m=\Gamma, \quad \partial_x\psi_m=m\Gamma \tag{47}$$

and

$$\tau=1+\Gamma r^2. \tag{48}$$

(In each occurence of the letter Γ it represents a different function bounded by K).
Consider now initial data of the form

$$\begin{aligned} &A_x=\psi_m, \qquad A_x=A_z=A_t=0,\\ &\hat{A}_z=-\psi_m, \quad \hat{A}_x=\hat{A}_y=\hat{A}_t=0, \end{aligned} \tag{49}$$

with

$$\psi_m=\varrho\sin m\chi. \tag{50}$$

We have

$$\begin{aligned} E^x&=\varepsilon^{xyzt}n_t(\partial_yA_z-\partial_zA_y)\\ &=-\tau\varrho_{,z}\sin m\chi\\ &=\Gamma,\\ E^y&=0,\\ E^z&=m\tau\varrho\cos m\chi+\Gamma, \end{aligned} \tag{51}$$

and

$$\begin{aligned} B^x&=\Gamma\\ B^y&=m\tau\varrho\cos m\chi+\Gamma\\ B^z&=0. \end{aligned} \tag{52}$$

Then, from (39)

$$\begin{aligned} T_a{}^bt^an_b&=\tfrac{1}{2}\mu^{-1}[g_{zz}(E^z)^2+g_{yy}(B^y)^2+2\alpha E^zB^y]+m\Gamma+\Gamma\\ &=\tfrac{1}{2}\mu^{-1}(g_{yy}+g_{zz}-2\alpha\varepsilon_{yz})m^2\tau^2\varrho^2\cos^2 m\chi+m\Gamma+\Gamma. \end{aligned} \tag{53}$$

Now Eqs. (43), (44), and the definition of the tensor ε^{ab} imply

$$g_{yy}+g_{zz}-2\alpha\varepsilon_{yz}<2-2\alpha+\Gamma r^2 \tag{54}$$

it follows that by making Ω sufficiently small, we can require that on Ω,

$$g_{yy}+g_{zz}-2\varepsilon_{yz}<-\delta \tag{55}$$

for some positive number δ. Then

$$\begin{aligned}\mathscr{E}_S &= \int_{\Omega} T_a{}^b t^a n_b dS \\ &\leqq -\tfrac{1}{2}\delta \int_{\Omega} m^2\mu^{-1}\tau^2\varrho^2\cos^2 m\chi dS + mC_1 + C_2 \\ &\leqq -\frac{m^2}{2}\mu_1^{-1}\delta \int_{\Omega_R} \tau^2\cos^2 m\chi + mC_1 + C_2\end{aligned} \tag{56}$$

where C_1 and C_2 are constants.

Again, by restricting the size of Ω we can make $\tau-1$ small; and, as $m\to\infty$,

$$\int_{\Omega_R}\cos^2 m\chi dS \to \tfrac{1}{2}\int_{\Omega_R} dS \equiv \tfrac{1}{2}|\Omega_R|. \tag{57}$$

Thus for sufficiently large m,

$$\mathscr{E}_S \leqq -m^2(\tfrac{1}{4}\mu_1^{-1}|\Omega_R|)\delta + mC_1 + C_2 < 0 \tag{58}$$

as was to be proved.

In the case of an axisymmetric background with an axial Killing vector ϕ^a, one considers data with angular dependence $e^{\mathrm{i}m\phi}$, where ϕ is an angular coordinate about the axis of symmetry chosen in such a way that $\phi^a\nabla_a\phi=1$; in other words, $(E^a, B^a)=\mathrm{Re}(\hat{E}^a, \hat{B}^a)$, where

$$\mathscr{L}_\phi \hat{E}^a = \mathrm{i}m\hat{E}^a, \qquad \mathscr{L}_\phi \hat{B}^a = \mathrm{i}m\hat{B}^a. \tag{59}$$

Then for all integers m greater than some m_0 there is initial data with angular dependence $e^{\mathrm{i}m\phi}$ for which $\mathscr{E}_S>0$. Because the time evolution preserves the ϕ-dependence of the perturbation, each such initial data set gives rise to an independent unstable or marginally stable perturbation with angular dependence $e^{\mathrm{i}m\phi}$. The magnitude of m_0 depends on the detailed configuration, and crucially upon the size of the ergosphere. In particular, suppose $Q(J)$ represents a continuous sequence of equilibria, parameterized, say, by increasing angular momentum J; and suppose that for $J>J_0$ there is an ergosphere that shrinks to a point as $J\to J_0$. Then $m_0\to\infty$ as $J\to J_0{}^+$ and the instability can be said to set in as a limit $m\to\infty$ of perturbations with angular dependence $e^{\mathrm{i}m\phi}$.

Finally, as in § II one expects on the basis of Holmgren's theorem that initial perturbations having $\mathscr{E}_S<0$ will grow without bound and therefore that configurations with ergospheres will be strictly unstable. The expectation relies, however, on the fact that by Eq. (37), if $\mathscr{E}<0$ the perturbation must be time dependent and on the assumption that time dependent perturbations will be radiative; and even for the scalar wave equation, there is no formal demonstration that all time dependent solutions on a stationary background (with or without ergosphere) radiate energy to null infinity.

A related result for quantum fields on a background spacetime with ergosphere obtained by Ashtekhar and Magnon [10] should be mentioned. In constructing a

Hilbert space from solutions to the scalar wave equation, one introduces a complex structure analogous to that obtained in flat space by the decomposition of a real solution into its positive and negative frequency parts. Ashtekhar and Magnon show that any definition of complex structure for which the Klein-Gordon inner product is positive definite must be time dependent; that is, the complex structure cannot be Lie derived by the asymptotically timelike Killing field. Consequently any spacetime with ergosphere is unstable to particle creation, and so is quantum mechanically as well as classically unstable. (For astrophysical objects the particle creation would be negligible.)

Acknowledgements. I want to thank James Ipser for suggesting the problem considered here, Robert Geroch for helpful conversations, and Bernard Schutz for comments on a previous version of the manuscript.

Appendix

We establish here a uniqueness theorem for C^∞ solutions to linear wave equations on analytic spacetimes of the form

$$\nabla_m \nabla^m \psi^{a\ldots b} = 0 \tag{A1}$$

where $\psi^{a\ldots b}$ is an n-index tensor (see also [11]). As stated in § II, one can define a domain of dependence for timelike surfaces as follows. Let $T \subset M$ be a timelike hypersurface and let $\tau \subset T$ be an analytic submanifold of T with compact closure. Consider the set of all timelike surfaces $\hat{\tau}$ which can be obtained from τ by an analytic deformation that leaves $\delta\tau$ fixed. The union of all such $\hat{\tau}$ for all compact $\tau \subset T$ is the domain of timelike dependence of T, written $D(T)$.

We assume that M is an analytic manifold and that the metric g_{ab} is an analytic tensor field on M. The uniqueness theorem is

Proposition. *If $\lambda^{a\ldots b}$ and $\mu^{a\ldots b}$ are two C^∞ n-index tensor fields on T, there is at most one C^∞ tensor field $\psi^{a\ldots b}$ on $D(T)$ satisfying*

$$\nabla_m \nabla^m \psi^{a\ldots b} = 0$$

$$\psi^{a\ldots b}|_T = \lambda^{a\ldots b}$$

and

$$n^m \nabla_m \psi^{a\ldots b}|_T = \mu^{a\ldots b}, \tag{A2}$$

where n^a is the unit normal to T.

Proof. By the linearity of (A2), it suffices to prove that $\lambda^{a\ldots b}=0$ and $\mu^{a\ldots b}=0$ imply $\psi^{a\ldots b}=0$ on $D(T)$. The proof is based on a theorem due to Holmgren [5], a version of which can be stated in the following manner. Let $L(u)=0$ be a hyperbolic system of r linear analytic partial differential equations of order m in r functions u_j of k variables. Then if for all j, u_j and its derivatives of order less than or equal to $m-1$ vanish on a noncharacteristic manifold and if u_j is C^∞, u_j must vanish in a neighborhood of the manifold.[4]

4 The proof given by Holmgren is for single hyperbolic equations, but its extension to hyperbolic systems is straightforward

There is a neighborhood N of any point p of T and a chart on N for which the Eq. (A1) satisfies the conditions of Holmgren's theorem, and so if $\psi^{a\dots b}$ and $n^m \nabla_m \psi^{a\dots b}$ vanish on T, $\psi^{a\dots b}$ must vanish in some neighborhood of each point p of T.

Now let τ be a submanifold of T with compact closure. Suppose that for each $S \in [0, 1]$, $\chi_S : \tau \to M$ is a diffeomorphism of τ to $\tau_S = \chi_S(\tau)$, that τ_S is timelike and that $\partial \tau_S = \partial \tau$. Further, suppose that the map $[0, 1] \to M$, $S \to \chi_S(p)$ is continuous for all $p \in \tau$; in other words that χ_S is a deformation of τ. We want to show that $\psi^{a\dots b}$ vanishes on all surfaces τ_S. Let S_0 be the greatest lower bound of all $S \in [0, 1]$ for which $\psi^{a\dots b}|_{\tau_S}$ is not identically zero. By continuity, the field $\psi^{a\dots b}$ and all its derivatives vanish on τ_{S_0}, and since τ_{S_0} is timelike, Holmgren's theorem implies that $\psi^{a\dots b}$ vanishes in a neighborhood of each point of τ_{S_0}. But $\bar{\tau}_{S_0}$ is compact, and so is covered by a finite collection of open sets O_α on each of which $\psi^{a\dots b}$ vanishes. Then if $S_0 \neq 1$, $\tau_S \subset \bigcup O_\alpha$ for a finite range of $S > S_0$, say for $S_0 < S < S_1$. This means that $\psi^{a\dots b}$ vanishes on τ_S for all $S < S_1$, contradicting the assumption that S_0 was the greatest lower bound for surfaces on which τ_S vanished identically. Whence $S_0 = 1$, and we conclude that $\psi^{a\dots b}$ vanishes on $D(T)$.

References

1. Butterworth, E.M., Ipser, J.R.: Ap. J. **204**, 200 (1975)
2. Wilson, J.R.: Ap. J. **176**, 195 (1972)
3. Bardeen, J.M., Wagoner, R.V.: Ap. J. **167**, 359 (1971)
4. Comins, N., Schutz, B.F.: (to be published)
5. Holmgren, E.: Vetenskaps-Akad. Fohr. **58**, 91 (1901)
6. Newman, E.T., Unti, T.W.J.: J. Math. Phys. **3**, 891 (1962)
7. Müller zum Hagen, H.: Camb. Phil. Soc. **68**, 199 (1970)
8. Moret-Bailly, J., Papapetrou, A.: Ann. Inst. H. Poinc. **6**, 205 (1967)
9. Bardeen, J.M., Press, W.H.: J. Math. Phys. **14**, 7 (1972)
10. Ashtekar, A., Magnon-Ashtekar, A.: C. R. Acad. Sci. Paris **281** A, 875 (1975)
11. Friedlander, F.G.: Proc. Roy. Soc. (Lond.) A **269**, 53 (1962)

Communicated by R. Geroch

Received June 8, 1978