

A New Proof of the Positive Energy Theorem[★]

Edward Witten

Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544, USA

Abstract. A new proof is given of the positive energy theorem of classical general relativity. Also, a new proof is given that there are no asymptotically Euclidean gravitational instantons. (These theorems have been proved previously, by a different method, by Schoen and Yau.) The relevance of these results to the stability of Minkowski space is discussed.

I. Introduction

In most classical field theories that play a role in physics, the total energy is the integral of a positive definite energy density T_{00} . This positivity of the energy is usually responsible for ensuring the stability of the ground state.

In gravity, the situation is very different. Even in the weak field case, there is no satisfactory way to define the energy density of the gravitational field. An energy momentum pseudotensor can be defined [1], but it is not a true tensor and is not positive definite. The positivity of the energy in general relativity and the stability of Minkowski space as the ground state are therefore far from obvious.

Although there is no satisfactory way to define the local energy density when gravity is present, one can define the total energy of a gravitating system [2]. The total energy (and momentum and angular momentum) of a gravitating system can be defined in terms of the asymptotic behavior, at large distances, of the gravitational field. However, it is far from obvious that the total energy so defined is always positive.

It is an old conjecture that this total energy is in fact always strictly positive, except for flat Minkowski space, which has zero energy. This matter has been studied by a variety of means.

The energy of a class of gravitational waves was studied by Weber and Wheeler [3]. Positivity of the energy for gravitating systems of special classes was demonstrated by Araki, by Brill, and by Arnowitt, Deser, and Misner [4]. The

[★] Research partially supported by NSF Grant PHY78-01221

paper by Brill gives references to earlier unpublished work by Bondi, Bonnor, Weber, and Wheeler.

Brill and Deser [5] showed that among spaces that are topologically Minkowskian, Minkowski space is the unique stationary point of the energy; they also showed that it is locally a minimum. Brill et al. [6] gave the variational argument in a canonical form.

The positive energy problem with spherically symmetric initial data was discussed by Leibovitz and Israel, and by Misner [7]. Geroch [8] gave a simple argument for spaces with Minkowski topology admitting a maximal hypersurface and also reviewed the status of the problem. O'Murchadha and York [9] analyzed time symmetric initial data and spaces with maximal hypersurfaces. Jang [10] proved that spaces with a flat initial hypersurface have positive energy. Leite [11] proved the positive energy theorem for spaces whose initial value surface can be isometrically embedded in R^4 .

Choquet-Bruhat and Cantor, Fischer, Marsden, and O'Murchadha [12] proved the existence, under appropriate conditions, of maximal hypersurfaces. Choquet-Bruhat and Marsden [13] proved positive energy for any space which is in a sufficiently small neighborhood (in the sense of functional analysis) of Minkowski space.

Deser and Teitelboim [14] and Grisaru [15] pointed out that, formally, in supergravity theory the total energy operator is a sum of squares and therefore positive. Deser and Teitelboim showed that global supersymmetry charges can really be defined in supergravity. Grisaru suggested that it might be possible to give a rigorous, purely classical proof of the classical positive energy theorem by taking the limit as $\hbar \rightarrow 0$ of the supergravity argument.

Finally, Schoen and Yau [16] used a geometrical method to prove the positive energy theorem for the key case of a space with a maximal spacelike slice. Using an auxiliary equation introduced by Jang [17] who also made considerable progress in generalizing the approach of [8], Schoen and Yau have generalized their proof to a general proof of the positive energy theorem [18], thus finally resolving this long-standing problem. They have also applied their method [19] to prove the positive action conjecture [20].

In this paper, new and simple proofs of the positive energy theorem, and of another theorem which is also relevant to the stability of Minkowski space, will be presented.

Related to the question of whether the energy is always positive is the question of whether Minkowski space is stable against semiclassical decay processes. In the last few years it has been learned [21] that the decay of unstable vacuum states in quantum field theory can be systematically studied on the basis of "bounce" solutions of the classical Euclidean equations of motion. Might Minkowski space itself be unstable against a semiclassical decay process in general relativity?

In Sect. II of this paper a simple proof is given that such a semiclassical decay of Minkowski space does not occur, at least in pure gravity. In Sect. III a new proof is presented of the positive energy theorem, which states that in classical general relativity, Minkowski space is the unique space of lowest energy. This is a more far-reaching indication of the stability of Minkowski space than the absence of a semiclassical decay mechanism, because it shows that irrespective of the

mechanism, there is no state to which it is energetically possible for Minkowski space to decay. The proofs presented here were found in the course of an attempt to take the limit as $\hbar \rightarrow 0$ of the formal argument by Deser and Teitelboim and by Grisaru that in supergravity the total energy is a sum of squares.

II. Semiclassical Stability of Minkowski Space

Before considering the positive energy theorem, I will first prove a related but simpler theorem, whose proof involves fewer technicalities. This theorem has been proved in a different way by Schoen and Yau [19].

If the positive energy theorem were false and a state of negative energy existed in general relativity, Minkowski space would presumably be unstable and would decay into the negative energy state. How would this decay take place? Since Minkowski space is known to be stable against small fluctuations (the energy of linearized gravitational waves is positive!), the hypothetical decay of Minkowski space would presumably occur via barrier penetration, or quantum mechanical tunneling through a barrier.

Actually, it has been shown [22] by Perry and by Gross, Perry and Yaffe that at non-zero temperature quantum gravity is unstable against such a process. Also, it will be shown in a separate paper [23] that the ground state of the Kaluza-Klein unified theory of gravitation and electromagnetism is unstable and decays by barrier penetration. So it is not idle to ask whether ordinary Minkowski space could have a semiclassical instability.

In the last few years it has been understood [21] how to analyze in field theory the decay by barrier penetration of an unstable ground state. One looks for instanton-like "bounce" solutions of the classical Euclidean field equations in which at large distances the fields approach their values in the unstable vacuum state. Instability of that vacuum state shows up in the form of negative action modes for small fluctuations around the instanton; inclusion of these modes in a functional determinant gives an imaginary part to the energy of the false vacuum.

To investigate by these means the stability of Minkowski space, one should look for a Euclidean metric (signature + + + +) solution of the Einstein equations which at large distances asymptotically approaches Euclidean space. Here only the case of pure gravity, without matter fields, will be considered, so we should study the source-free Einstein equations

$$R_{\mu\nu} = 0 \tag{1}$$

with the boundary condition that outside a compact region one can introduce coordinates x^i in which the metric g_{ij} is asymptotically Euclidean,

$$g_{ij} = \delta_{ij} + a_{ij}, \quad a_{ij} \rightarrow 0 \quad \text{as} \quad |x| \rightarrow \infty. \tag{2}$$

The energy of Minkowski space would get an imaginary contribution, indicating an instability, if there existed a metric satisfying (1) and (2) and such that for small fluctuations about this metric there were negative action modes. [A non-flat space satisfying (1) and (2) would probably have interesting consequences even if negative action modes did not exist.] However, it will now be shown that the only space that satisfies (1) and (2) is flat Euclidean space.

At very large distances, the quantity a_{ij} defined in (2) is extremely small; it becomes asymptotically a solution of the linearized Einstein equations. In four dimensions, every solution of the linearized Einstein equations that vanishes at infinity vanishes asymptotically at least as fast as $1/r^4$ (with $r=|x|$) if the coordinates x^i are chosen to be, for instance, harmonic coordinates (asymptotically). This fact may require some explanation. In four dimensions a monopole field (a solution of the massless spin zero equation $\Delta^2\phi=0$) vanishes at least as fast as $1/r^2$. A dipole field is the derivative of a monopole field and vanishes at least as $1/r^3$. A quadrupole field such as the gravitational field is the derivative of a dipole field, and vanishes at least as fast as $1/r^4$.

We therefore can assume for the metric the asymptotic behavior

$$g_{ij} = \delta_{ij} + O(1/r^4) \tag{3}$$

so that the affine connection has asymptotic behavior

$$\Gamma^i_{jk} \sim 1/r^5. \tag{4}$$

I will now prove that an asymptotically Euclidean space with $R_{\mu\nu}=0$ must be flat by constructing coordinates in which the metric tensor is identically δ_{ij} .

The ordinary Cartesian coordinates of flat Euclidean space are harmonic functions. This simply means that the ordinary coordinates $t, x, y,$ and $z,$ regarded as scalar functions, satisfy Laplace's equation. For instance,

$$\left(\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)t=0 \tag{5}$$

and likewise for $x, y,$ and $z.$ Of course, Laplace's equation can be written in a generally covariant form. If instead of the Cartesian coordinates $t, x, y,$ and z one describes flat space by a general curvilinear coordinate system, the functions $t, x, y,$ and z still satisfy the covariant form of Laplace's equation. This means, for instance, that $D^2t=0,$ where $D^2=g^{\mu\nu}D_\mu D_\nu$ is the Laplacian defined in the curvilinear coordinate system.

In our problem, we are given a space with $R_{\mu\nu}=0$ and a coordinate system x^i (defined at least outside a compact region) in which the metric satisfies condition (3). We wish to prove that this space is flat Euclidean space in disguise and that the x^i are simply curvilinear coordinates for flat space. To prove this we will construct Cartesian coordinates ϕ^i in which the metric will be identically $\delta_{ij}.$

If such coordinates ϕ^i exist they certainly satisfy Laplace's equation since we have already noted that Cartesian coordinates satisfy this equation. Therefore

$$D^2\phi=0, \tag{6}$$

where ϕ may be any of the $\phi^i.$ We are thus lead to study the solutions of Eq. (6).

Laplace's equation has no nonzero solutions which vanish at infinity because the operator $-D^2$ is positive. If $D^2\phi=0$ then

$$\int dx\phi(-D^2\phi)=0. \tag{7}$$

(The measure $d^4x\sqrt{g}$ will often be denoted just as $dx.$)

Integrating by parts, and discarding the surface term because ϕ vanishes at large distances, we find

$$\int d^4x \sqrt{g} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi = 0. \tag{8}$$

(Since any harmonic function that vanishes at infinity vanishes at least as $1/r^2$, we may, in fact, discard the surface term.) Since the integrand is positive definite, this implies that $\partial_\mu \phi = 0$ everywhere, so ϕ is a constant; the constant vanishes because ϕ vanishes at infinity.

Although $D^2\phi = 0$ has no nonzero solutions which vanish at infinity, there certainly exist solutions if ϕ is not required to vanish at infinity. In fact, we will see that, if a^i is any constant vector, there is always a unique solution ϕ of $D^2\phi = 0$ such that

$$\phi = a \cdot x + O(1/r^3) \quad \text{as } |x| \rightarrow \infty. \tag{9}$$

To prove that ϕ exists, consider first a convenient trial function ϕ_1 . The trial function ϕ_1 should be linear in a^i and should equal $a \cdot x$ identically for large enough x . It would be adequate to define $\phi_1 = a \cdot x$ everywhere if this formula made sense. It may not make sense because, depending on the topology of the manifold, the coordinates x^i may not be defined everywhere, but only outside a compact set. To allow for this, ϕ_1 may be defined as follows. Suppose that the x^i are defined for $|x| > R_0$ and let R be some number greater than R_0 . Then let $\phi_1 = a \cdot x f(|x|)$ where f is any smooth function which is identically one for $|x| > R$ and identically zero for $|x| < R_0$.

We now write the desired harmonic function ϕ as $\phi = \phi_1 + \phi_2$, where ϕ_2 must satisfy

$$D^2\phi_2 = -D^2\phi_1. \tag{10}$$

Formally, this can be solved by

$$\phi_2(x) = - \int dy G(x, y) D^2\phi_1(y), \tag{11}$$

where $G(x, y)$ is the Green's function of the Laplacian operator with boundary conditions that $G(x, y) \rightarrow 0$ as $|x| \rightarrow \infty$. (This Green's function exists because we have already seen that the Laplacian has no zero modes.) Equation (11) makes sense and satisfies (10) provided that the integral converges.

To see that the integral does converge, note that with $\phi_1(y) = a \cdot y$ for $|y| > R$ and the asymptotic behavior of the metric given by Eqs. (3) and (4), $D^2\phi_1(y)$ is of order $1/|y|^5$ for large $|y|$. This ensures the convergence of the integral.

What is the asymptotic behavior of the function $\phi_2(x)$ defined in (11)? Since in an asymptotically flat four dimensional manifold $G(x, y) \sim -1/(2\pi^2|x|^2)$ for large x , independent of y , the large x behavior is

$$\phi_2(x) = 1/(2\pi^2|x|^2) \int dy D^2\phi_1(y) + O(1/|x|^3) \tag{12}$$

provided that the integral converges. Actually, the integral not only converges but vanishes, since the integrand is a total divergence (and in view of the asymptotic behavior of ϕ_1 and of the metric, there is no surface term). So $\phi_2(x)$ is of order $1/|x|^3$ for large $|x|$.

The net effect of this is that

$$\phi(x) = a \cdot x f(|x|) - \int dy G(x, y) a \cdot y f(|y|) \tag{13}$$

is a harmonic function which satisfies the boundary condition $\phi(x) = a \cdot x + O(1/r^3)$.

The next step is to prove that $K_\nu = \partial_\nu \phi$ is covariantly constant, $D_\mu K_\nu = 0$. This will be accomplished by proving that $\int dx (D_\mu K_\nu)^2$ vanishes. Since the integrand is nowhere negative, the integral can vanish only if the integrand vanishes identically.

In general

$$\int dx (D_\mu K_\nu)^2 = \frac{1}{2} \int dx (D_\mu K_\nu - D_\nu K_\mu)^2 + \int dx D_\mu K_\nu D_\nu K_\mu. \tag{14}$$

In the case at hand the first term on the right-hand side vanishes because $K_\nu = \partial_\nu \phi$ is a gradient. After integrating twice by parts in the second term of Eq. (14), we find

$$\begin{aligned} \int dx (D_\mu K_\nu)^2 &= \int dx (D_\mu K^\mu)^2 - \int dx K^\nu [D_\mu, D_\nu] K^\mu \\ &\quad + \int dx \partial_\mu (K^\nu D_\nu K^\mu - K^\mu D_\nu K^\nu) \\ &= \int dx (D_\mu K^\mu)^2 - \int dx K^\nu K^\mu R_{\mu\nu} \\ &\quad + \int dx \partial_\mu (K^\nu D_\nu K^\mu - K^\mu D_\nu K^\nu). \end{aligned} \tag{15}$$

On the right-hand side of (15), the first term vanishes because $D_\mu K^\mu = D^2 \phi = 0$. The second term vanishes because we are given that $R_{\mu\nu} = 0$. And the last term vanishes because the integral converges rapidly enough that there is no surface contribution [In view of Eqs. (3), (4), and (9) concerning the asymptotic behavior of ϕ and of the metric, $K_\nu = \partial_\nu \phi$ is of order one for large $|x|$, and the covariant derivative of K is at most of order $1/r^5$.] So $\int (D_\mu K_\nu)^2 = 0$, and $D_\mu K_\nu$ must vanish identically.

At this point we have constructed a covariantly constant vector field K_μ . But because a constant vector a^i entered the construction, and four linearly independent choices of a^i are possible, we actually have four linearly independent vector fields K which are each covariantly constant. A four dimensional manifold with four independent covariantly constant vector fields must be flat, so our space is flat.

Actually this construction yields explicitly a Cartesian coordinate system. The function ϕ of Eq. (13) is obviously of the form $\phi = a^i \phi_i$, with

$$\phi_i = x_i f(|x|) - \int dy G(x, y) y_i f(|y|). \tag{16}$$

The four functions ϕ_i are the desired Cartesian coordinates. In fact, we have shown that in any coordinate system $D_\mu D_\nu \phi = 0$ (since $D_\mu D_\nu \phi$ is the same as $D_\mu K_\nu$). If the ϕ_i are chosen as coordinates, we have

$$0 = D_i D_j \phi = \frac{\partial}{\partial \phi^i} \frac{\partial}{\partial \phi^j} \phi - \Gamma_{ij}^k \frac{\partial}{\partial \phi^k} \phi \tag{17}$$

with $\phi = a^i \phi_i$, $\frac{\partial}{\partial \phi^k} \phi = a_k$ and $\frac{\partial}{\partial \phi^i} \frac{\partial}{\partial \phi^j} \phi = 0$.

So

$$0 = -\Gamma_{ij}^k(\phi_l) a_k. \tag{18}$$

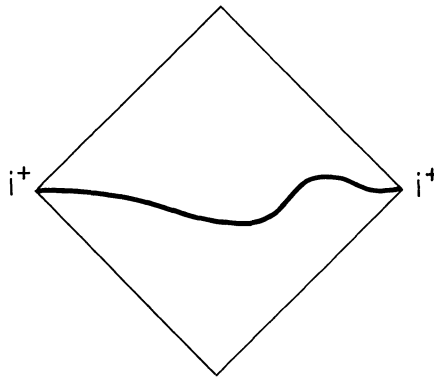


Fig. 1. An asymptotically Euclidean initial value hypersurface

Since the choice of a_k was arbitrary, the affine connection Γ_{ij}^k vanishes everywhere. Hence the metric tensor is a constant in terms of the coordinates ϕ_i . The constant is δ_{ij} since asymptotically [see Eq. (16)] $\phi_i = x_i$. So the ϕ_i are indeed Cartesian coordinates.

While the argument has been presented for four dimensions, it is valid in any number of dimensions. Only the power counting depends on the number of dimensions. The relevant modifications are that Eq. (3) becomes $g_{ij} = \delta_{ij} + O(1/r^n)$ and Eq. (9) becomes $\phi = a \cdot x + O(1/r^{n-1})$.

III. The Positive Energy Theorem

In this section, a new proof of the positive energy theorem will be described. The problem which must be addressed is the following. One is given a space-time which satisfies Einstein's equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi GT_{\mu\nu}. \tag{19}$$

The only requirement on the energy momentum tensor $T_{\mu\nu}$ is that the local energy density T_{00} is positive (or zero) at each point in space-time and in each local Lorentz frame.

It is assumed, moreover, that in this space-time there exists (Fig. 1) a space-like hypersurface (which can be regarded as the initial value surface) that is asymptotically Euclidean. More specifically, we suppose that in the vicinity of this space-like hypersurface the metric behaves at spatial infinity as

$$g_{\mu\nu} = \eta_{\mu\nu} + O(1/r) \tag{20}$$

$$\frac{\partial}{\partial x^k} g_{\mu\nu} = O(1/r^2),$$

where $\eta_{\mu\nu}$ is the flat space metric (signature $- + + +$). [The second condition in (20) is needed so that the energy integral defined below should converge.] We will make no assumption about the topology of the initial value surface.

The total energy of this system is defined as a surface integral over the asymptotic behavior of the gravitational field,

$$E = \frac{1}{16\pi} \int d^2S^j \left(\frac{\partial}{\partial x^k} g_{jk} - \frac{\partial}{\partial x^j} g_{kk} \right), \quad (21)$$

where the integral is evaluated over a bounding surface in the asymptotically flat region of the initial value surface. The problem is to prove that this total energy E is always positive or zero, and zero only for flat Minkowski space.

There certainly are spaces for which the flux integral (21) is negative and which satisfy Einstein's equations at large distances. For instance, for the Schwarzschild space

$$ds^2 = -dt^2 \left(1 - \frac{2GM}{r} \right) + \frac{dr^2}{1 - 2GM/r} + r^2 d\Omega^2 \quad (22)$$

the integral (21) is easily evaluated, and one finds $E = M$. Whether M is positive or negative, the Schwarzschild space satisfies the vacuum Einstein equations everywhere except at the singular point $r = 0$. In the negative M case, every asymptotically flat spacelike surface passes through this singular point, where the equations are not satisfied. A special case of the positive energy theorem is that the negative M Schwarzschild solution cannot be matched onto an interior solution with $T_{\mu\nu}$ satisfying the positivity condition. (In contrast, in the positive M Schwarzschild space there are asymptotically flat spacelike surfaces that do not pass through the singularity, and of course it is also possible to match the positive M Schwarzschild solution with an allowable interior solution that has no singularity.)

In the special case in which matter is absent and $T_{\mu\nu}$ is zero everywhere, the positive energy theorem states that the total energy of the pure gravitational field is always positive (or zero for flat space). When matter is present, the positive energy theorem states that the total energy of a gravitating system, including the energy of the matter and also the energy of the field, is always positive, if the matter contribution is positive ($T_{00} \geq 0$ everywhere and in each frame).

The energy defined as in (20) is always conserved because, since no signal can travel faster than light, no physical process can change the asymptotic behavior at spacelike infinity of the gravitational field. Semiclassical barrier penetration, the process which was discussed in the last section, conserves the total energy, and presumably this would also be true for any hypothetical alternative process by which Minkowski space might decay. Therefore, the positive energy theorem, according to which Minkowski space is the unique space of minimum energy, ensures the stability of Minkowski space.

Let us now turn to the proof of the positive energy theorem. The proof will involve a consideration of solutions of the Dirac equation

$$i\mathcal{D}\varepsilon = 0 \quad (23)$$

on the initial value hypersurface. The need to consider spinors is perhaps surprising. Spinors have been used in the past to prove various results in differential geometry, including some that are at least roughly related to the positive energy theorem [24]. Their use in this paper was suggested by considerations involving supersymmetry, which will be discussed later.

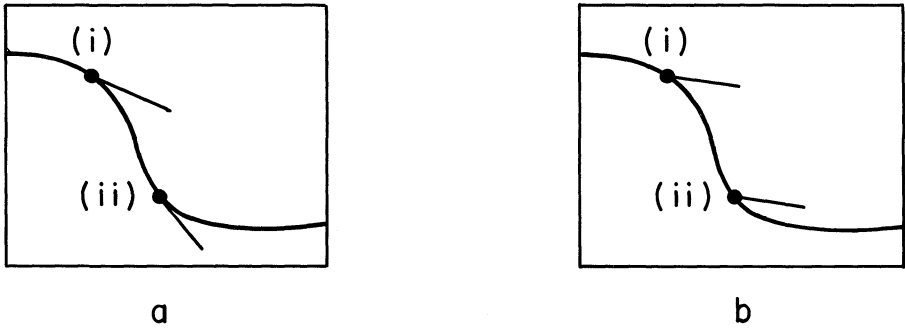


Fig. 2a and b. Parallel transport within a hypersurface of a Riemannian manifold may be defined relative to the intrinsic geometry of the hypersurface, as shown in **a**, or relative to the geometry of the full manifold, as shown in **b**. In the first case a tangent sector to the hypersurface remains tangent after parallel transport from (i) to (ii); in the second case it does not

The Dirac operator that we will consider is not the four dimensional Dirac operator but rather is a three dimensional Dirac operator defined on the initial value surface. Thus, by “ \mathcal{D} ” is meant $\sum_{i=1}^3 \gamma^i D_i$, where the sum runs only over the three directions tangent to the initial value surface, and the spinor ε that satisfies Eq. (23) is defined only on this surface. (The γ matrices satisfy $\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\eta_{\mu\nu}$.)

To fully explain what Eq. (23) means, it is still necessary to define the covariant derivative D_i that appears in the formula $\mathcal{D} = \sum_i \gamma^i D_i$. There are two obvious possibilities. One could define D_i as the covariant derivative with respect to the intrinsic geometry of the three dimensional initial value surface. Or one can use the usual covariant derivative D_μ of the full four dimensional geometry.

The difference between the two possibilities is illustrated in Fig. 2. Parallel transport within a hypersurface of a Riemannian manifold can be defined using the intrinsic geometry of the hypersurface, in which case a vector tangent to the hypersurface will remain tangent (Fig. 2a), or using the full four dimensional geometry, in which case a tangent vector may not remain tangent (Fig. 2b).

To prove the positive energy theorem, we will make the second choice, so the covariant derivative in Eq. (23) is the covariant derivative of the four dimensional space-time although we will always use it only to differentiate within the three dimensional initial value surface.

The next step is to prove that the Dirac equation (23) has no solution ε that vanishes at large distances. In fact, if $i\mathcal{D}\varepsilon = 0$, then

$$0 = (i\mathcal{D})^2 \varepsilon = - \sum_i D^i D_i \varepsilon - \frac{1}{4} \sum_{i,j} [\gamma^i, \gamma^j] [D_i, D_j] \varepsilon, \tag{24}$$

where summations over Latin letters i and j run over the three directions tangent to the initial value surface.

Now we recall that for spinors

$$[D_i, D_j] \varepsilon = \frac{1}{8} \sum_{\alpha\beta} R_{i\alpha\beta} [\gamma^\alpha, \gamma^\beta] \varepsilon \tag{25}$$

where the sum over Greek letters α and β ranges over all four directions in the space-time, normal or tangent to the initial value surface, because D_i is the covariant derivative of the four-manifold. So

$$0 = - \sum_i D^i D_i \varepsilon - \frac{1}{32} \sum_{i,j} \sum_{\alpha,\beta} [\gamma^i, \gamma^j] [\gamma^\alpha, \gamma^\beta] R_{ij\alpha\beta} \varepsilon. \tag{26}$$

Now we use the Dirac algebra identity

$$\begin{aligned} [\gamma^\mu, \gamma^\nu] [\gamma^\alpha, \gamma^\beta] &= 4\varepsilon^{\mu\nu\alpha\beta} \gamma_5 - 4(g^{\mu\alpha} g^{\nu\beta} - g^{\mu\beta} g^{\nu\alpha}) \\ &\quad + 2(-g^{\mu\alpha} [\gamma^\nu, \gamma^\beta] + g^{\mu\beta} [\gamma^\nu, \gamma^\alpha] \\ &\quad + g^{\nu\alpha} [\gamma^\mu, \gamma^\beta] - g^{\nu\beta} [\gamma^\mu, \gamma^\alpha]). \end{aligned} \tag{27}$$

The term involving $\varepsilon^{\mu\nu\alpha\beta}$ does not contribute in (26) because $\varepsilon^{\mu\nu\alpha\beta} R_{\mu\nu\alpha\beta} = 0$. The term $-4(g^{\mu\alpha} g^{\nu\beta} - g^{\mu\beta} g^{\nu\alpha})$, when inserted in (26), gives

$$\frac{1}{4} \sum_{i,j} R_{ij}^i \varepsilon \tag{28}$$

and the last term in (27) gives (after some use of the algebraic symmetries of the Riemann tensor)

$$-\frac{1}{2} \sum_{j,k} R_{0kj}^k \gamma^0 \gamma^j \varepsilon, \tag{29}$$

where the index “0” refers to a unit vector normal to the hypersurface.

But the expressions (28) and (29) are related to the energy-momentum tensor by Einstein’s equations! By Einstein’s equations

$$R_{00} - \frac{1}{2} g_{00} R = 8\pi G T_{00}. \tag{30}$$

Since the “0” direction is orthonormal to the hypersurface, $g_{00} = -1$. Also $R_{00} = R_{0i0}^i$ and $R = R_{\mu\nu}^{\mu\nu} = 2R_{0i}^{0i} + R_{ij}^{ij}$. So altogether $R_{00} - \frac{1}{2} g_{00} R = \frac{1}{2} R_{ij}^{ij}$, and hence

$$R_{ij}^{ij} = 16\pi G T_{00}. \tag{31}$$

Likewise

$$R_{0j} - \frac{1}{2} g_{0j} R = 8\pi G T_{0j}. \tag{32}$$

But $g_{0j} = 0$, because the j direction is tangent and the 0 direction is normal to the initial value surface. Since $R_{0j} = R_{0kj}^k$, we have

$$R_{0kj}^k = 8\pi G T_{0j}. \tag{33}$$

Combining these equations, we learn that any solution of $\mathcal{D}\varepsilon = 0$ also satisfies

$$- \sum_i D^i D_i \varepsilon + 4\pi G \left(T_{00} + \sum_j T_{0j} \gamma^0 \gamma^j \right) \varepsilon = 0. \tag{34}$$

We can now verify the claim that no solution of $\mathcal{D}\varepsilon = 0$ vanishes at infinity. In fact, after multiplying by ε^* and integrating over the three surface on which ε is defined, we obtain

$$\int d^3x \sqrt{g} \varepsilon^* \left(- \sum_i D_i D^i \varepsilon + 4\pi G \left(T_{00} + \sum_j T_{0j} \gamma^0 \gamma^j \right) \varepsilon \right) = 0. \tag{35}$$

We now integrate by parts. There is no surface term because on an asymptotically flat three dimensional hypersurface any solution of the Dirac equation that vanishes at infinity vanishes at least as fast as $1/r^2$. So

$$\int d^3x \sqrt{g}(D_i \varepsilon^* D_i \varepsilon) + 4\pi G \int d^3x \sqrt{g} \varepsilon^* \left[T_{00} + \sum_j T_{0j} \gamma^0 \gamma^j \right] \varepsilon = 0. \tag{36}$$

The crucial point is now that the matrix $T_{00} + \sum_j T_{0j} \gamma^0 \gamma^j$ is a positive (semi-definite) matrix. The positivity condition on the energy momentum tensor requires T_{00} to be equal to or greater than the magnitude of the vector T_{0j} of momentum flux; otherwise T_{00} would be negative in some local inertial frame. The (hermitian) matrix $\sum_j T_{0j} \gamma^0 \gamma^j$ has eigenvalues equal to plus or minus the magnitude of the momentum flux, so the eigenvalues of $T_{00} + \sum_j T_{0j} \gamma^0 \gamma^j$ are positive or zero.

Since the second term in (36) is known to be non-negative, (36) can vanish only if $D_i \varepsilon = 0$. But if $D_i \varepsilon = 0$, and ε is not identically zero, ε does not vanish at infinity. So no nonzero solution of $\mathcal{D}\varepsilon = 0$ can vanish at infinity.

We will prove the positive mass theorem by studying spinors that satisfy the Dirac equation and do not vanish at infinity. We will see, in fact, that there always exists a solution of $\mathcal{D}\varepsilon = 0$ satisfying

$$\varepsilon \xrightarrow{r \rightarrow \infty} \varepsilon_0 + O(1/r) \tag{37}$$

for any constant ε_0 .

To prove that ε always exists, we proceed in a fashion similar to the argument in Sect. II. We first write $\varepsilon = \varepsilon_1 + \varepsilon_2$ where ε_1 is a trial function of the form $\varepsilon_1 = \varepsilon_0 + O(1/r)$. Then we will define ε_2 in terms of ε_1 , using the Green's function of the Dirac operator.

A broad class of trial functions satisfying $\varepsilon_1 = \varepsilon_0 + O(1/r)$ could be used to prove that ε exists. However, because we will need more accurate information about the asymptotic behavior of ε than is indicated in (37), we must choose ε_1 more carefully.

We wish to choose $\varepsilon_1(x)$ so that $\mathcal{D}\varepsilon_1 = O(1/r^3)$; we wish to show that a spinor $\varepsilon_1(x)$ satisfying this condition and approaching the constant ε_0 at infinity exists and is uniquely determined up to terms of order $1/r^2$. In fact, let us write $\varepsilon_1(x) = \varepsilon_0 + (1/r)\tilde{\varepsilon}(\theta, \phi) + O(1/r^2)$ where $\tilde{\varepsilon}$ is a function only of the polar angles θ and ϕ (which are well defined at large distances on an asymptotically Euclidean three surface) and not of r . We wish to show that $\tilde{\varepsilon}(\theta, \phi)$ can be chosen so that $\mathcal{D}\varepsilon_1 = O(1/r^3)$, and, moreover, we wish to show that this choice is unique.

The Dirac operator \mathcal{D} can be written as $\gamma^i(\partial_i + \Gamma_i)$ where ∂_i is the ordinary derivative and Γ_i involves the spin connection. For the class of metrics we are considering [Eq. (20)], the spin connection Γ_i is of order $1/r^2$. Since ε_0 is a constant (in asymptotically flat coordinates that we are using), $\mathcal{D}\varepsilon_0 = \gamma^i \Gamma_i \varepsilon_0$ and is of order $1/r^2$. Let us write $\gamma^i \Gamma_i \varepsilon_0 = (1/r^2)A(\theta, \phi) + O(1/r^3)$; that is,

$$A(\theta, \phi) = \lim_{r \rightarrow \infty} r^2 \gamma^i \Gamma_i(r, \theta, \phi) \varepsilon_0.$$

On the other hand,

$$\mathcal{D}((1/r)\tilde{\varepsilon}(\theta, \phi)) = \gamma^i \partial_i((1/r)\tilde{\varepsilon}(\theta, \phi)) + O(1/r^3),$$

so in calculating terms of order $1/r^2$ we may replace $\mathcal{D}((1/r)\tilde{\varepsilon})$ by $\tilde{\mathcal{D}}((1/r)\tilde{\varepsilon})$, where $\tilde{\mathcal{D}} = \sum_i \gamma^i \partial_i$ is the flat space Dirac operator.

We can therefore guarantee that $\mathcal{D}(\varepsilon_0 + (1/r)\tilde{\varepsilon}) = O(1/r^3)$ if

$$\tilde{\mathcal{D}}((1/r)\tilde{\varepsilon}(\theta, \phi)) = -\frac{1}{r^2} A(\theta, \phi). \tag{38}$$

To see that the solution exists and is unique, we write the free Dirac operator $\tilde{\mathcal{D}}$ as $\gamma^r \frac{\partial}{\partial r} + \frac{1}{r} \tilde{\mathcal{D}}^T$, where $\tilde{\mathcal{D}}^T$ is the angular or transverse part of the free Dirac operator, which acts on spinor functions defined on a sphere. In terms of $\tilde{\mathcal{D}}^T$, the equation for $\tilde{\varepsilon}$ is

$$\left(-\frac{\gamma^r}{r^2} \tilde{\varepsilon}(\theta, \phi) + 1/r^2 \tilde{\mathcal{D}}^T \tilde{\varepsilon}(\theta, \phi) \right) = -1/r^2 A(\theta, \phi) \tag{39}$$

or [since $(\gamma^r)^2 = 1$]

$$(1 - \gamma^r \tilde{\mathcal{D}}^T) \tilde{\varepsilon}(\theta, \phi) = \gamma^r A(\theta, \phi). \tag{40}$$

The solution of Eq. (40) exists, and is unique, provided that the operator $(1 - \gamma^r \tilde{\mathcal{D}}^T)$ is invertible. It is invertible, because it has no zero eigenvalues. In fact, if there were a zero eigenvalue $\alpha(\theta, \phi)$, satisfying $(1 - \gamma^r \tilde{\mathcal{D}}^T)\alpha(\theta, \phi) = 0$, then $(1/r)\alpha(\theta, \phi)$ would be a solution of the free Dirac equation, vanishing for large r as $1/r$. But in three dimensions every solution of the Dirac equation that vanishes at infinity vanishes at least as fast as $1/r^2$.

Since $(1 - \gamma^r \tilde{\mathcal{D}}^T)$ is invertible, $\tilde{\varepsilon}$ satisfying (40) exists and is unique; the solution can be written formally as

$$\tilde{\varepsilon}(\theta, \phi) = \frac{1}{1 - \gamma^r \tilde{\mathcal{D}}^T} \gamma^r A(\theta, \phi). \tag{41}$$

For future use, let us write the equation for $\tilde{\varepsilon}$ in this form:

$$\left(1 - \gamma^r \frac{1}{r} \tilde{\mathcal{D}} \right) \tilde{\varepsilon}(\theta, \phi) = r^2 \gamma^r \gamma^i \Gamma^i \varepsilon_0 + O\left(\frac{1}{r}\right). \tag{42}$$

Note that, acting on $\tilde{\varepsilon}$, $\tilde{\mathcal{D}}^T$ can be replaced by $\frac{1}{r} \tilde{\mathcal{D}}$ because $\tilde{\varepsilon}$ is independent of r .

Let us now return from this technical interlude to the main thread of the argument. We have found ε_1 which approaches ε_0 at infinity and with $\mathcal{D}\varepsilon_1 = O(1/r^3)$. Now we wish to find a spinor $\varepsilon(x)$ approaching ε_0 at infinity and with $\mathcal{D}\varepsilon = 0$. As in Sect. II, we write $\varepsilon = \varepsilon_1 + \varepsilon_2$ and attempt to satisfy $\mathcal{D}\varepsilon_2 = -\mathcal{D}\varepsilon_1$, with ε_2 vanishing at large distances. This can be formally solved by

$$\varepsilon_2(x) = -\int d^3 y S(x, y) \mathcal{D}\varepsilon_1(y), \tag{43}$$

where $S(x, y)$ is the Green's function of the Dirac operator \mathcal{D} , with boundary conditions that $S(x, y)$ vanishes for large x or y .¹ The expression (43) makes sense and defines a spinor ε_2 that satisfies $\mathcal{D}\varepsilon_2 = -\mathcal{D}\varepsilon_1$ provided that the integral over y converges. This three dimensional integral does, in fact, converge, because the Green's function $S(x, y)$ vanishes for large y as $1/y^2$, and we have chosen ε so that $\mathcal{D}\varepsilon_1(y) \sim 1/y^3$.

What is the asymptotic behavior of $\varepsilon_2(x)$? We are interested in verifying that $\varepsilon_2(x)$ vanishes faster than $1/r$ for large r . In fact, on any asymptotically flat three manifold, $S(x, y)$ behaves for large x as $\frac{1}{4\pi r^2} \gamma \cdot \hat{x} + O(1/r^3)$ (where x is a unit vector the x direction). So

$$\varepsilon_2(x) = -\frac{1}{4\pi r^2} \int dy \gamma \cdot \hat{x} \mathcal{D}\varepsilon_1(y) + O(1/r^3) \tag{44}$$

provided that the integral converges. In fact, the integral may diverge, but only logarithmically, since we have $\mathcal{D}\varepsilon_1(y) \sim 1/y^3$. If the integral in (44) is logarithmically divergent, $\varepsilon_2(x)$ behaves for large r not as $1/r^2$ but as $(1/r^2) \ln r$. In any case, $\varepsilon_2(x)$ vanishes for large x faster than $1/r$.

[In proving that $\varepsilon_2(x)$ vanishes faster than $1/r$, we used the fact that $S(x, y)$ vanishes for large x as $1/r^2$. This is so because any solution of the free Dirac equation which vanishes at infinity vanishes at least as fast as $1/r^2$. The latter fact was used in establishing the uniqueness of the solution of (40). Actually, the fact that $\varepsilon_2(x)$ vanishes faster than $1/r$ follows from our previous uniqueness argument, for if ε_2 vanished only as $1/r$, then $\varepsilon_1 + \varepsilon_2$ would be a counterexample to the claim that $\varepsilon_1(x)$ is uniquely determined, up to terms vanishing faster than $1/r$, by the conditions we placed on it.]

We are now almost ready to prove the positive mass theorem. We may repeat the entire analysis leading to Eq. (34), which used the fact that $\mathcal{D}\varepsilon = 0$ but did not depend upon the boundary conditions satisfied by ε . So we have again

$$-\sum_i D_i D^i \varepsilon + 4\pi G \left(T_{00} + \sum_j T_{0j} \gamma^0 \gamma^j \right) \varepsilon = 0. \tag{45}$$

We may once again multiply by ε^* . We can again integrate over x . And we can again integrate by parts. The only difference from the derivation of (36) is that now, because ε does not vanish at infinity, we pick up a surface term. We therefore obtain

$$\int d^3x x \partial_k (\varepsilon^* D^k \varepsilon) = \int d^3x D_i \varepsilon^* D_i \varepsilon + \int d^3x \varepsilon^* \left(T_{00} + \sum_j T_{0j} \gamma^0 \gamma^j \right) \varepsilon. \tag{46}$$

Since the right-hand side is positive, we have

$$S \geq 0, \tag{47}$$

1 We have shown that the Dirac operator has no zero eigenvalue. Using this fact, we presume that standard methods can be used to yield existence of $S(x, y)$ and local regularity of $S(x, y)$ for $x \neq y$. The existence of the Dirac Green's function would also follow from the existence of the Green's function of the positive definite, hermitian operator $(i\mathcal{D})^2$ that appears in Eq. (34), because the Green's function of the Dirac operator is $i\mathcal{D}$ times the Green's function of this second order operator

where

$$S = \int d^3x \partial^k (\varepsilon^* D_k \varepsilon) = \int d^2S^k \varepsilon^* D_k \varepsilon. \tag{48}$$

To evaluate the surface integral (48), it is enough to calculate $D_k \varepsilon$ up to and including terms of order $1/r^2$. For this purpose it is sufficient to calculate ε up to and including terms of order $1/r$. We have seen [in connection with Eqs. (41) and (42)] that to calculate ε up to and including terms of order $1/r$ it is sufficient to know the terms of order $1/r$ in the metric tensor. Therefore, S depends only on this leading large distance correction to the flat space metric. In fact, an explicit formula will be presented below.

S is, moreover, obviously an invariant of the initial value three surface. The only invariants that can be formed from the $1/r$ term in the metric tensor are the total energy and the total momentum. Therefore, it must be possible to identify S in terms of the energy and momentum of the system. This will now be demonstrated.

For a general weak gravitational field, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, one can introduce the orthonormal basis $e_\mu^i = \delta_\mu^i + \frac{1}{2} h_\mu^i$. Relative to this basis, the linearized form of the spin connection matrices is

$$\Gamma_\lambda = \frac{1}{16} \sum_{\alpha\beta} (\partial_\beta h_{\alpha\lambda} - \partial_\alpha h_{\beta\lambda}) [\gamma^\alpha, \gamma^\beta]. \tag{49}$$

In an asymptotically flat space, the gravitational field is weak at large distances, and we may use the linearized form (49) in evaluating S .

Since the calculations are somewhat intricate in general, let us first consider the special case of an object “at rest”, whose field is expected to be asymptotically of the Schwarzschild form. This means that the terms of order $1/r$ are

$$\begin{aligned} h_{ij} &= \frac{2GM}{r} \delta_{ij} \\ h_{0i} &= 0 \\ h_{00} &= \frac{2GM}{r}. \end{aligned} \tag{50}$$

Again, i and j refer to the three directions tangent to the initial value surface and 0 to the normal direction. We wish to prove that $M \geq 0$ for any space of asymptotic form (50) if the positivity condition on $T_{\mu\nu}$ is satisfied everywhere.

It is straightforward to calculate from (49) that the asymptotic form of Γ_i is

$$\Gamma_i = -\frac{1}{4} \frac{GM}{r^3} [\gamma_i, \boldsymbol{\gamma} \cdot \mathbf{x}] + O(1/r^3), \tag{51}$$

where $\boldsymbol{\gamma} \cdot \mathbf{x} = \sum_k \gamma^k x^k$. The asymptotic form of the Dirac operator is

$$\mathcal{D} = \not{\partial} + \gamma^i \Gamma_i = \not{\partial} - \frac{GM}{r^3} \boldsymbol{\gamma} \cdot \mathbf{x}. \tag{52}$$

The equation $\mathcal{D}\varepsilon = 0$ with the boundary condition $\varepsilon \rightarrow \varepsilon_0$ at infinity is satisfied by

$$\varepsilon = \left(1 - \frac{GM}{r}\right) \varepsilon_0 + O(1/r^2). \tag{53}$$

(This could also have been derived by a conformal transformation using the fact that the Schwarzschild three geometry is conformally flat.)

We can now readily evaluate

$$S = \int dS^k \varepsilon^* D_k \varepsilon. \quad (54)$$

If the integration surface is taken to be a large sphere, then [since Γ_r vanishes by Eq. (51)], equation (54) is just

$$S = \int d\Omega r^2 \varepsilon^* D_r \varepsilon = \int d\Omega r^2 \varepsilon_0^* \frac{\partial}{\partial r} \left(1 - \frac{GM}{r} \right) \varepsilon_0^* = 4\pi GM \varepsilon_0^* \varepsilon_0. \quad (55)$$

So S is a positive constant times M , and positivity of S is positivity of the mass. This proves that an object whose field is asymptotically Schwarzschild must have positive total energy.

Let us now turn to the general case in which it is assumed only that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, with $h_{\mu\nu} \sim 1/r$ and $\partial_\alpha h_{\mu\nu} \sim 1/r^2$ at spacelike infinity. We will see that the generalization of (55) is

$$S = 4\pi G \left(E \varepsilon_0^* \varepsilon_0 + \sum_k P_k \varepsilon_0^* \gamma^0 \gamma^k \varepsilon_0 \right), \quad (56)$$

where E is the total energy of the system and P_k are the components of the total momentum. Regardless of the value of P_k , it is always possible to choose ε_0 as an eigenvector of the matrix $\gamma^0 \gamma \cdot \mathbf{P}$ in order that $\sum_k P_k \varepsilon_0^* \gamma^0 \gamma^k \varepsilon_0 = -|P| \varepsilon_0^* \varepsilon_0$, where $|P|$ is the magnitude of the momentum. With such a choice of ε_0 , $S = 4\pi G \varepsilon_0^* \varepsilon_0 (E - |P|)$, and positivity of S means that

$$E \geq |P|. \quad (57)$$

This is, of course, the necessary and sufficient condition for the total energy to be positive in every asymptotic Lorentz frame.

It remains to derive Eq. (56). We have

$$S = \int d\Omega r^2 \varepsilon^* D_r \varepsilon. \quad (58)$$

We must evaluate the term of order $1/r^2$ in $D_r \varepsilon$, which is the only term that contributes.

With $\varepsilon = \varepsilon_0 + \tilde{\varepsilon}(\theta, \phi)/r$ and $\Gamma_r \sim 1/r^2$, we have

$$D_r \varepsilon = \Gamma_r \varepsilon_0 - 1/r^2 \tilde{\varepsilon}(\theta, \phi) + O(1/r^3) \quad (59)$$

so

$$S = \int d\Omega r^2 \varepsilon_0^* \Gamma_r \varepsilon_0 - \int d\Omega \varepsilon_0^* \tilde{\varepsilon}(\theta, \phi). \quad (60)$$

Let us now simplify the second term in Eq. (60) to eliminate $\tilde{\varepsilon}$ in favor of ε_0 .

(In what follows, surface integrals will often be rewritten as volume integrals of a total divergence to make possible simplifications. The integrals will always depend only on the large r behavior of quantities such as $\tilde{\varepsilon}$, which have in fact been defined only asymptotically.)

Using Eq. (41), we have

$$\int d\Omega \varepsilon_0^* \tilde{\varepsilon}(\theta, \phi) = \int d\Omega \varepsilon_0^* (\gamma^r/r) \partial \tilde{\varepsilon} + \int d\Omega r^2 \varepsilon_0^* \gamma^r \gamma^i \Gamma^i \varepsilon_0. \quad (61)$$

It is possible, but unfortunately somewhat tedious, to show that the first term on the right side of (61) vanishes. We have

$$\begin{aligned} \int d\Omega e_0^* \gamma^r / r \partial \tilde{\epsilon}(\theta, \phi) &= \int d^3 x \partial / \partial x^k [e_0^* (\gamma^k / r) \partial \tilde{\epsilon}] \\ &= \int d^3 x \partial / \partial x^k [e_0^* (\gamma^k \gamma^l / r) \partial / \partial x^l \tilde{\epsilon}] \\ &= \int d^3 x \partial / \partial x^k \partial / \partial x^l [e_0^* (\gamma^k \gamma^l / r) \tilde{\epsilon}] \\ &\quad + \int d^3 x \partial / \partial x^k [e_0^* \gamma^k (\gamma \cdot \mathbf{x} / r^3) \tilde{\epsilon}]. \end{aligned} \tag{62}$$

The first term in the last expression is

$$\begin{aligned} \int d^3 x \partial / \partial x^k \partial / \partial x^l [e_0^* \gamma^k \gamma^l / r \tilde{\epsilon}] &= \int d^3 x \partial / \partial x^k \partial / \partial x^k [e_0^* (1/r) \tilde{\epsilon}] \\ &= \int dS^k \partial / \partial x^k [e_0^* (1/r) \tilde{\epsilon}] \\ &= \int d\Omega r^2 \partial / \partial r [e_0^* (1/r) \tilde{\epsilon}(\theta, \phi)] \\ &= - \int d\Omega e_0^* \tilde{\epsilon}(\theta, \phi). \end{aligned} \tag{63}$$

The last term in (62) is

$$\begin{aligned} \int d^3 x \partial / \partial x^k [e_0^* \gamma^k \gamma \cdot \mathbf{x} / r^3] \tilde{\epsilon} &= \int d\Omega r^2 [e_0^* \gamma^r \gamma \cdot \mathbf{x} / r^3] \tilde{\epsilon} \\ &= \int d\Omega e_0^* \tilde{\epsilon}(\theta, \phi) \end{aligned} \tag{64}$$

(because $\gamma \cdot \mathbf{x} = r\gamma^r$ and $\gamma^r \gamma \cdot \mathbf{x} = r$). So after all, since (63) and (64) cancel, the first term on the right side of (61) indeed vanishes.

We may now use (61) to simplify Eq. (60), and we obtain the desired formula

$$S = \int dS^k e_0^* (\Gamma_k - \gamma_k \gamma^i \Gamma_i) \epsilon_0, \tag{65}$$

which expresses S in terms of the arbitrary spinor ϵ_0 and the asymptotic behavior of Γ only.

Making use of (49), it is straightforward to calculate that

$$\begin{aligned} \Gamma_k - \gamma_k \gamma^i \Gamma_i &= \frac{1}{4} (\partial_i h_{ki} - \partial_k h_{ii}) + \frac{1}{8} (\partial_\beta h_{ki}) [\gamma_i, \gamma_\beta] \\ &\quad - \frac{1}{8} \partial_k h_{\beta i} [\gamma^i, \gamma^\beta] - \frac{1}{8} \partial_\beta h_{ii} [\gamma_k, \gamma_\beta] + \frac{1}{8} \partial_i h_{\beta i} [\gamma_k, \gamma_\beta], \end{aligned} \tag{66}$$

where a repeated Latin letter “ i ” refers to a summation over the three spatial directions, and a repeated Greek index “ β ” refers to a sum over all four space-time dimensions.

It is possible to see that the terms in (66) with $\beta \neq 0$ do not contribute to the surface integral which defines S . This follows from the identities

$$\int dS^k \partial_j h_{ii} [\gamma_j, \gamma_k] = \int d^3 x \partial_j \partial_k h_{ii} [\gamma_j, \gamma_k] = 0, \tag{67}$$

$$\begin{aligned} \int dS^k [\partial_j h_{ki} [\gamma_i, \gamma_j] + \partial_i h_{ji} [\gamma_k, \gamma_j]] \\ = \int d^3 x [\partial_j \partial_k h_{ki} [\gamma_i, \gamma_j] + \partial_k \partial_i h_{ji} [\gamma_k, \gamma_j]] = 0, \end{aligned} \tag{68}$$

and from the symmetry of $h_{\mu\nu}$, which ensures $[\gamma^i, \gamma^j] h_{ij} = 0$. [The fact that (68) vanishes is seen by cyclic redefinition of indices $k \rightarrow i, i \rightarrow j, j \rightarrow k$ in the second integral on the right-hand side; the two terms then cancel because of the symmetry of h_{jk} .]

We may therefore set $\beta=0$ in (66) to obtain a formula for S :

$$S = \frac{1}{4}\epsilon_0^*\epsilon_0 \int dS^j(\partial_i h_{ji} - \partial_j h_{ii}) + \frac{1}{4} \sum_k \epsilon_0^* \gamma^0 \gamma^k \epsilon_0 \cdot \int dS^j(\partial_j h_{0k} - \partial_0 h_{jk} + \delta_{jk} \partial_0 h_{ii} - \delta_{jk} \partial_i h_{0i}). \tag{69}$$

Comparing with the standard formulas for the total energy and momentum of a gravitating system

$$E = \frac{1}{16\pi G} \int dS^j(\partial_i h_{ji} - \partial_j h_{ii}) \tag{70}$$

$$P_k = \frac{1}{16\pi G} \int dS^j(\partial_j h_{0k} - \partial_0 h_{jk} + \delta_{jk} \partial_0 h_{ii} - \delta_{jk} \partial_i h_{0i})$$

we arrive at the promised result

$$S = 4\pi G \left(\epsilon_0^* \epsilon_0 E + \sum_k \epsilon_0^* \gamma^0 \gamma^k \epsilon_0 P_k \right). \tag{71}$$

As was noted above, by choosing ϵ_0 to be an eigenvector of the hermitian matrix $\gamma^0 \boldsymbol{\gamma} \cdot \mathbf{P}$ with eigenvalue $-|P|$, one immediately concludes

$$E \geq |P| \tag{72}$$

because S is known to be positive or zero.

To complete the proof of the positive mass theorem, we wish to prove that while the energy is positive or zero for every system, only flat, empty Minkowski space has zero energy.

If $E=0$, then by (72) $|P|=0$, and so by (71) $S=0$ for any choice of ϵ_0 . From (46), we see that for S to vanish, we must have $D_i \epsilon = 0$ throughout the initial value surface. In particular, being covariantly constant, ϵ nowhere vanishes on this surface. But from $D_i \epsilon = 0$ follows $[D_i, D_j] \epsilon = 0$ or

$$R_{ij\alpha\beta}[\gamma^\alpha, \gamma^\beta] \epsilon = 0. \tag{73}$$

The existence of a single nonzero ϵ satisfying (73) does not imply the vanishing of the $R_{ij\alpha\beta}$. But because an arbitrary ϵ_0 entered the construction, and S vanishes for any ϵ_0 , we have not just a single covariantly constant spinor, but a basis of nonzero covariantly constant spinor fields. The existence of such a basis implies that indeed

$$R_{ij\alpha\beta} = 0 \tag{74}$$

throughout the initial value hypersurface.

We have not yet proved that the space-time is flat because (74) has been shown to hold only on the initial value surface, and moreover (74) refers not to the full $R_{\mu\nu\alpha\beta}$ but only to the components for which $\mu=i$ and $\nu=j$ are tangent to this surface. But now we may make use of the “many fingered nature of time” in general relativity. Let us deform the initial value surface locally without changing its form at infinity (Fig. 3). Since the behavior at infinity has not been changed, E and \mathbf{P} are unchanged and S is still zero. Hence $R_{ij\alpha\beta} = 0$ also on the new surface. In order for $R_{ij\alpha\beta}$ to vanish on every space-like surface into which the initial value

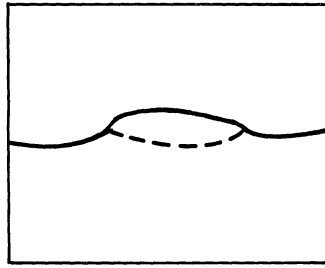


Fig. 3. A small deformation of the initial value hypersurface

surface can be deformed, we must in fact have $R_{\mu\nu\alpha\beta} = 0$ everywhere. So the space is flat.

The space is also empty because with $R_{\mu\nu\alpha\beta} = 0$, Einstein's equations $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi GT_{\mu\nu}$ immediately imply that $T_{\mu\nu} = 0$. This completes the proof that only Minkowski space has zero energy.

It is actually possible to go further and use these methods to classify the possible asymptotically flat spaces for which $E = |P|$ (but $E \neq 0$). If $E = |P|$, then by (71), S vanishes if ε_0 is chosen so that $\gamma^0 \boldsymbol{\gamma} \cdot \mathbf{P} \varepsilon_0 = -|P| \varepsilon_0$. This equation has a unique solution for each possible chirality of ε_0 ; the positive and negative chirality solutions are complex conjugates of one another.

Choosing ε_0 properly, we obtain unique covariantly constant spinors ε and ε^* of positive and negative chirality. From them we may construct a covariantly constant vector $n^\mu = \bar{\varepsilon} \gamma^\mu \varepsilon$. It is a property of the Lorentz group that a vector constructed in this way is always light-like, $n^\mu n_\mu = 0$. The existence of a covariantly constant light-like vector means that our system is propagating at the speed of light, as expected for a system of $E = |P|$. The vector n^μ points in the direction of propagation of the system.

From the spinor ε we may also construct a covariantly constant antisymmetric tensor $K^{\mu\nu} = \varepsilon^T C \sigma^{\mu\nu} \varepsilon$ (C is the charge conjugation matrix). Ehlers and Kundt [25] have classified the spaces admitting such covariantly constant tensors. Any such space is a so-called plane fronted (pp) wave. The line element can always be put in the form $dS^2 = d\sigma d\tau + H(x, y, \tau) d\tau^2 - dx^2 - dy^2$. The boundary condition of asymptotic flatness is that H must vanish for large $|\tau|$ and may grow for large x and y at most like $\ln|x^2 + y^2|$. [The logarithmic behavior looks dangerous, but can be eliminated by a coordinate transformation $\sigma' = \sigma + F(\tau, x, y)$ where $\partial F / \partial \tau = H$.]

As shown by Ehlers and Kundt, the vacuum Einstein equations require $\frac{d^2 H}{dx^2} + \frac{d^2 H}{dy^2} = 0$. There are no non-singular solutions that satisfy the boundary condition just stated except $H = H(\tau)$ (H a function of τ only), but this is Minkowski space in disguise [let $t + z = \tau$, $t - z = \sigma + F(\tau)$ with $F' = H$]. So the only asymptotically flat solution of the vacuum Einstein equations with $E = |P|$ is Minkowski space.

In the presence of matter, $\frac{d^2 H}{dx^2} + \frac{d^2 H}{dy^2}$ no longer vanishes; instead it is proportional to the matter density. From formulas of Ehlers and Kundt it can be

seen that there also are no nontrivial asymptotically flat solutions of the Einstein-Maxwell equations with $E=|P|$. However, it is possible that such states may exist in some other field theories. In fact, theories of the Kaluza-Klein type might be good candidates.

Let us briefly consider some generalizations. Although stated above in a 3+1 dimensional language, the proof in fact applies in any number of dimensions provided that the topology of the initial value surface is such that spinors can be defined on this surface. On a three dimensional surface spinors can always be defined and they also can always be defined on any space-time in which an initial value surface can be introduced [26]. In a world of more than three space dimensions, it may be impossible for topological reasons to consistently define spinors throughout the space. Therefore, the proof given here, which depends on the use of spinors, does not fully generalize to the case of more than three space or four space-time dimensions.

One may also consider the case of a space-time in which the initial value surface connects between several different asymptotically flat worlds. (The most famous example is the analytically continued Schwarzschild solution.) In this case observers in each asymptotically flat world would, in general, observe a different mass. One can prove that each of these masses is positive by introducing solutions of the Dirac equation which are asymptotically constant in one world, and asymptotically vanish in the other worlds.

IV. Connection with Supergravity

In this section a few speculative remarks will be made about the not altogether clear relation between the previous argument and supergravity.

Formally, in supergravity theory the Hamiltonian is the sum of squares of the hermitian supersymmetry charges Q_α ,

$$H = \frac{1}{\hbar} \sum_{\alpha} Q_{\alpha}^2. \quad (75)$$

As noted by Deser and Teitelboim and by Grisaru, this formally proves the positivity of the energy in supergravity. However, it is not clear, because of the ultraviolet problems, that supergravity really makes sense as a quantum field theory. And supergravity, like other theories with fermions (including ordinary quantum electrodynamics), surely does not make sense as a classical field theory. Consequently, the precise meaning of (75) is not obvious.

It was suggested by Grisaru that it might be possible to obtain a well-defined statement from (75) by taking the classical limit, $\hbar \rightarrow 0$. The fermions, being anticommuting fields whose square is of order \hbar , vanish as $\hbar \rightarrow 0$. The boson theory becomes classical general relativity as $\hbar \rightarrow 0$. Notice that the Q_{α} vanish as $\hbar \rightarrow 0$ because they are proportional to the fermi fields. But the explicit factor of $1/\hbar$ in (75) ensures that the right-hand side of (75) does not vanish as $\hbar \rightarrow 0$, since it is equal to the energy H , which presumably does not vanish as $\hbar \rightarrow 0$. This suggests that it might make sense to take the limit of (75) as $\hbar \rightarrow 0$, and that one might obtain in this way a proof in terms of well-defined classical objects of the positivity of the

energy in classical general relativity. It seems likely that the argument in Sect. III of this paper can be understood as the limit as $\hbar \rightarrow 0$ of Eq. (75).

What is meant by the global supersymmetry charges Q ? In the absence of gravity a global supersymmetry charge is

$$Q = \int d^3x \bar{\varepsilon}_\alpha S_{0\alpha}, \quad (76)$$

where ε_α is a c -number constant spinor. But in curved space time there are no (covariantly) constant spinors. As discussed by Deser and Teitelboim, in supergravity the global supersymmetry charges are defined as in (76), but with ε_α being *asymptotically* covariantly constant. This suggests that the classical object which should be used in trying to construct a classical limit of Eq. (75) is the asymptotically constant spinor ε .

The choice of ε in (76) is arbitrary, as long as ε is constant at infinity. Because of local supersymmetry, any choices of ε that are equal at infinity give the same integral for Q . One would like to impose some condition on ε to facilitate its use in proving positive energy. The choice of ε is similar to a choice of gauge in supergravity. Since the condition $D_i \varepsilon = 0$ is too strong (no solution exists) it is natural to try the weaker condition $\gamma^i D_i \varepsilon = 0$. This condition on ε probably corresponds in supergravity to a gauge condition $\gamma^i \psi_i = 0$, since $\delta(\gamma^i \psi_i) = \gamma^i \delta \psi_i = \gamma^i D_i \varepsilon$. Requiring $\mathcal{D} \varepsilon = 0$ on the initial value surface leads to a unique choice of ε and to the argument in Sect. III.

Readers conversant with supergravity may notice some analogies between manipulations in Sect. III and manipulations in supergravity theory. For instance, the manipulations leading from (26) to (34) are similar to those which enter in proving the gauge invariance and consistency of supergravity. Also, the final formula

$$S = \varepsilon_0^* (E + \gamma_0 \boldsymbol{\gamma} \cdot \mathbf{P}) \varepsilon_0 = \bar{\varepsilon}_0 \gamma^\mu P_\mu \varepsilon_0$$

is reminiscent of supersymmetry and strongly suggests that S in the limit as $\hbar \rightarrow 0$ of $\text{Tr} \{ \bar{\varepsilon}_0 Q, \bar{Q} \varepsilon_0 \}$.

V. Discussion

As has been noted, the main importance of the positive energy theorem is that it is related to the stability of Minkowski space as the ground state of general relativity. The proof of the positive energy theorem in this paper suggests a connection between the classical and semiclassical stability of general relativity and the existence of supergravity. This may add to the attractiveness of supergravity. One may even wonder whether in the absence of local supersymmetry, the stability of general relativity could survive in a full quantum theory.

It is unfortunate that the proof of the positive energy theorem in Sect. III does not fully generalize to more than four space-time dimensions (because spinors, which were used extensively, do not always exist). In connection with theories of the Kaluza-Klein type, one would like to know whether in more than 3+1 dimensions Minkowski space is stable, or whether it undergoes instead "spontaneous compactification" [27]. I believe that it is an important question to know whether the positive energy theorem is true in any number of dimensions.

The positive action conjecture in d dimensions is a special case of the positive energy conjecture in $d+1$ dimensions [20]. The argument of Sect. III proves the positive action conjecture in any number of dimensions for manifolds on which spinors can be defined. For a proof of this theorem in four dimensions that does not depend on spinors, see [19].

The absence of asymptotically Euclidean solutions of $R_{\mu\nu}=0$, for which a new proof was given in Sect. II, is actually a special case of the positive action conjecture. If $R_{\mu\nu}=0$, then the Hilbert action $-\int dx \sqrt{g}R$ obviously vanishes. And the surface term in the action also vanishes; Einstein's equations force an approach to flatness too rapid to permit surface contributions. So the action of an asymptotically Euclidean solution of $R_{\mu\nu}=0$ vanishes, and the space must be flat if the positive action theorem holds. For manifolds on which spinors can be defined, the result of Sect. II is thus a corollary of the result of Sect. III.

References

1. Trautman, A.: Conservation laws in general relativity. In: Gravitation: an introduction to current research, Witten, L. (ed.). New York: Wiley 1962;
Misner, C., Thorne, K.S., Wheeler, J.A.: Gravitation. San Francisco: W. H. Freeman 1973;
Weinberg, S.: Gravitation and cosmology. New York: Wiley 1972
2. Arnowitt, R., Deser, S., Misner, C.W.: Phys. Rev. **117**, 1595 (1960); **118**, 1100 (1960); **122**, 997 (1961); The dynamics of general relativity. In: Witten, L. op.cit
3. Weber, J., Wheeler, J.A.: Rev. Mod. Phys. **29**, 509 (1957)
4. Araki, H.: Ann. Phys. (N.Y.) **7**, 456 (1959);
Brill, D.: Ann. Phys. (N.Y.) **7**, 466 (1959);
Arnowitt, R., Deser, S., Misner, C.: Ann. Phys. (N.Y.) **11**, 116 (1960)
5. Brill, D., Deser, S.: Ann. Phys. (N.Y.) **50**, 548 (1968)
6. Brill, D., Deser, S., Faddeev, L.: Phys. Lett. **26 A**, 538 (1968)
7. Leibovitz, C., Israel, W.: Phys. Rev. **D1**, 3226 (1970);
Misner, C.: In: astrophysics and general relativity, Chretien, M., Deser, S., Goldstein, J. (ed.). New York: Gordon and Breach 1971
8. Geroch, R.: Ann. N.Y. Acad. Sci. **224**, 108 (1973); Proc. Symp. Pure Math. **27**, 401 (1975)
9. O'Murchadha, N., York, J.W., Jr.: Phys. Rev. **D10**, 2345 (1974)
10. Jang, P.S.: J. Math. Phys. **17**, 141 (1976)
11. Leite, M.: Ph.D. Dissertation, MIT (1977) (unpublished)
12. Choquet-Bruhat, Y.: C.R. Acad. Sci. **280**, 169 (1975);
Cantor, M., Fischer, A., Marsden, J., O'Murchadha, N., York, J.W., Jr.: Commun. Math. Phys. **49**, 187 (1976)
13. Choquet-Bruhat, Y., Marsden, J.: C.R. Acad. Sci. **282**, 609 (1976); Commun. Math. Phys. **51**, 283 (1976)
14. Deser, S., Teitelboim, C.: Phys. Rev. Lett. **39**, 249 (1977)
15. Grisar, M.: Phys. Lett. **73 B**, 207 (1978)
16. Schoen, P., Yau, S.T.: Commun. Math. Phys. **65**, 45 (1979)
17. Jang, P.S.: J. Math. Phys. **19**, 1155 (1978)
18. Schoen, P., Yau, S.T.: Phys. Rev. Lett. **43**, 1457 (1979); Commun. Math. Phys. (to appear)
19. Schoen, P., Yau, S.T.: Phys. Rev. Lett. **42**, 547 (1979)
20. Gibbons, G.W., Hawking, S.W., Perry, M.J.: Nucl. Phys. **B138**, 141 (1978);
Hawking, S.W.: Phys. Rev. **D18**, 1747 (1978);
Page, D.N.: Phys. Rev. **D18**, 2733 (1978);
Gibbons, G.W., Pope, C.N.: Commun. Math. Phys. **66**, 267 (1979)
21. Coleman, S.: Phys. Rev. **D15**, 2929 (1977);
Coleman, S., Callan, C.G., Jr.: Phys. Rev. **D16**, 1762 (1977)

22. Perry, M.: Princeton Preprint, to appear in 1980 Nuffield workshop proc.; Gross, D.J., Perry, M., Yaffe, L.: Princeton Preprint (to appear)
23. Witten, E.: Princeton Preprint (to appear)
24. Lichnerowicz, A.: C.R. Acad. Sci. (Paris) A-B, **257**, 7 (1963); Hitchin, N.: Adv. Math. **14**, 1 (1974)
25. Ehlers, J., Kundt, W.: In Witten, L.: op.cit.
26. Geroch, R.: J. Math. Phys. **11**, 343 (1970)
27. Cremmer, E., Scherk, J.: Nucl. Phys. B **103**, 399 (1976)

Communicated by S. T. Yau

Received November 26, 1980; in revised form January 21, 1981

Note added in proof. A technical error in the calculation, which cancels out in the final formulas, has been pointed out by J. Nestor (preprint, Univ. of Alberta) and by M. Perry (private communication). In squaring the Dirac operator Eq. (24) an extra term appears proportional to the second fundamental form of the initial value surface. It arises as follows: If t^α is the unit normal to the initial value surface, then our Dirac operator is $(g^{\alpha\beta} - t^\alpha t^\beta)\gamma_\alpha D_\beta$. In squaring it, one obtains a term proportional to the derivative of t^α , which was omitted in Eq. (24).

In the integration by parts leading to Eq. (36), a similar term, proportional to the derivative of t^α , should appear; it precisely cancels the term just mentioned. This extra term appears because of the unusual covariant derivative for spinors and the fact that the integrand in (36) is a scalar in the three dimensional sense but is not a four dimensional Lorentz scalar.

Neither of these delicacies occurs in the often considered case of time symmetric initial conditions. I wish to thank M. Perry for discussions.