JERRY L. KAZDAN
Positive energy in general relativity

Astérisque, tome 92-93 (1982), Séminaire Bourbaki, exp. n° 593, p. 315-330

Most classical field theories in physics express energy as a sum of squares. Hence the energy is zero only if the field is zero. For gravitational fields, one is given a space-time \((N,\gamma)\) with a metric \(\gamma\) of signature \((-+++\)) . This metric satisfies Einstein's equations

\[
R_{\mu\nu} - \frac{1}{2} R \gamma_{\mu\nu} = 8\pi G T_{\mu\nu},
\]

where \(R_{\mu\nu}\) is the Ricci curvature tensor and \(R\) is the scalar curvature of \((N,\gamma)\), while \(T_{\mu\nu}\) is the energy-momentum tensor and \(G\) is a physical constant. One can think of \(T_{\mu\nu}\) as describing the physical distribution of matter. (We use the convention that Greek indices run from 0 to 3, Latin ones go from 1 to 3, and we sum on repeated indices).

Physical considerations \([G,p.408]\) and \([H-E, \S 4.3]\) suggest that in any orthonormal frame field \(\{e_0, e_1, e_2, e_3\}\), with \(e_0\) a time-like vector, then

\[
T^{00} \geq |T^{\mu\nu}| \quad \text{and} \quad T^{00} \geq (-T_{0i} T^{0i})^{1/2},
\]

so \(T^{00}\) dominates the other components of \(T^{\mu\nu}\) and the vector \(T^{0\nu}\) is non-spacelike. The (local) inequalities (A.2) are called the dominant energy condition.

In the special case where there is a space-like hypersurface \(M\) (i.e. a choice of space at one instant of time) which is asymptotically Euclidean, physical considerations have also suggested a definition of the

* Supported in part by an N.S.F. Grant.
total energy of the system. To be precise, we are assuming there is an oriented spacelike hypersurface \((M, g) \subset (N, \gamma)\) with induced metric \(g_{ij}\) and second fundamental form \(h_{ij}\). \((M, g)\) is assumed to be asymptotically Euclidean in the sense that for some compact set \(K \subset M\), the remainder, \(M - K\), consists of a finite number of subsets \(M_1, \ldots, M_n\) (called the ends of \(M\)), each of which is diffeomorphic to the exterior of a ball in \(\mathbb{R}^3\) (see figure). Moreover, under this diffeomorphism, in standard coordinates on \(\mathbb{R}^3\), \(g\) and \(h\) have the asymptotic behavior

\[
    g_{ij} = \delta_{ij} + p_{ij},
\]

where

\[
    p_{ij} = O(1/r) \quad \text{and} \quad h_{ij} = O(1/r^2)
\]

(actually, we need \(\partial^\alpha p = O(r^{-1-|\alpha|})\) for \(|\alpha| \leq 2\) and \(\partial^\beta h = O(r^{-2-|\beta|})\) for \(|\beta| \leq 1\)).
Then the total energy and total momentum of $M_\alpha$ are defined by the limiting expressions \[A-D-M,1,2\]

\[
E_\alpha = \lim_{R \to \infty} \frac{1}{16\pi G} \int_{S_{R,\alpha}} (\delta_i g_{ij} - \delta_j g_{ij}) dS^i
\]

\[
P_{aj} = \lim_{R \to \infty} \frac{1}{8\pi G} \int_{S_{R,\alpha}} (h_{ij} - \delta_i h_{j}) dS^i.
\]

One can think of $E_\alpha$ and $P_{aj}$, $j = 1, 2, 3$, as the four components of an energy-momentum vector associated to the $\alpha$ th end. The goal is to prove the following.

**Positive Energy Theorem.** Let $(M, g) \subset (N, \gamma)$ be an asymptotically Euclidean space-like hypersurface, where the metric $\gamma$ satisfies Einstein's equations (A.1) and the dominant energy condition (A.2). Then $E_\alpha \geq |P_{a}|$ for each end $M_\alpha$. Moreover, if $E_\alpha = 0$ for some $\alpha$, then $M$ is flat and has only one end.

Because the integrals (A.4) - (A.5) are evaluated at spatial infinity, and since no signal can travel faster than light, the value of the integrals (A.4) and (A.5) are conserved in time, and hence, so is the energy inequality $E \geq |P|$.

The first complete proof of this result was by Schoen-Yau [S-Y, 2,3,4] (see these papers as well as [G], [W] for a discussion of earlier work by others), who used minimal surfaces in a manner similar to the traditional use of geodesics. Subsequently, E. Witten [W] found a different proof using harmonic spinors (some mathematical considerations were clarified in [P-T] and, independently, in [C]).
The Schoen-Yau proof proceeds by contradiction, while Witten exhibits \( E - |P| \) explicitly as a sum of squares. Both proofs, and indeed the whole problem, are closely related to similar ones concerning the existence of positive scalar curvature metrics on Riemannian manifolds, for instance, on the torus \( T^3 \). Mathematicians became aware of both problems, and of their close relationship, from the lecture of Geroch [G] and Kazdan [K-W] at a 1973 symposium on differential geometry. Schoen-Yau [S-Y,1], using minimal surfaces, then resolved some of the questions concerning scalar curvature. Their work on the positive energy theorem is a direct outgrowth of that work. Independently, Gromov-Lawson [G-L,1,2], using harmonic spinors obtained other results on positive scalar curvature (see also the important earlier work of Lichnerowicz [Li] and Hitchin [H], as well as the survey [BB]).

B. **First Proof (Schoen-Yau)**

1. **The idea.** Geodesics minimizing arc length are a standard tool in geometry. If \( \gamma : \mathbb{R} \to M \) is a curve and \( L(\gamma) \) the arc length functional, then the first variation, \( \delta L(\gamma) = 0 \), gives the equation of geodesics, which are, by definition, the critical points of \( L \). If \( \gamma \) actually minimizes arc length, then the second variation yields the inequality

\[
0 \leq \delta^2_L(\gamma) .
\]

This inequality yields the Jacobi equation, and has been a rich source of seeing the effect of curvature on geodesics, and hence on many geometric phenomena.

The two dimensional analogue of a geodesic is a minimal surface. In this case one seeks the critical points of the surface area functional, \( A(S) \). Once one knows the existence of minimal surfaces (these existence proofs are
not trivial), it is natural to use the second variation, analogously to (B.1), for a surface $S$ that is actually a minimum, not just a critical point. For $S \hookrightarrow M$ with $\dim M = 3$ this second variation inequality (see [L, §9]) asserts that for all functions $f \in C^2(S)$ with compact support

$$0 \leq \delta^2A(S) = \int_S \left[ |\nabla_S f|^2 - (\text{Ric}_M(v) + \|b\|^2)f^2 \right] dx$$

$$= -\int_S \left[ \Delta_S f + (\text{Ric}_M(v) + ||b||^2)f \right] f dx ,$$

where $dx$ is the element of area on $S$, while $b = (b_{ij})$, $i,j = 1,2$ is the second fundamental form of the embedding $S \hookrightarrow M$ with $\|b\|^2 = \sum_{i,j} b_{ij}^2$, and $\text{Ric}_M(v)$ is the Ricci curvature of $M$ in the direction $v$ normal to $S$. (A minimal surface $S$ is called stable if $\delta^2A(S) \geq 0$ for all variations with compact support).

Now $S$ being minimal means that $0 = \text{trace } b = b_{11} + b_{22}$, so by standard formulas one has

$$\text{Ric}_M(v) + \|b\|^2 = \frac{1}{2}(R_M + \|b\|^2) - K_S ,$$

where $R_M$ is the scalar curvature of $M$ and $K_S$ is the Gaussian curvature of $S$. The stability inequality (B.2) may thus be rewritten as

$$0 \leq \int_S \left[ |\nabla_S f|^2 - (\frac{1}{2}R_M + \frac{1}{2}\|b\|^2 - K_S)f^2 \right] dx$$

for all functions $f \in C^2(S)$ with compact support.

If $S$ is compact, we may let $f = 1$ to conclude, by Gauss-Bonnet, that
This proves half of the following

\textbf{Theorem B.5} \cite{S-Y,1} \textit{On the torus, }$M = T^3$ \textit{there is no Riemannian metric }$g$ \textit{having positive scalar curvature, }$R_M$ \textit{. Moreover, if }$R_M \geq 0,$ \textit{then }$g$ \textit{is flat and }$R_M = 0.$

\textbf{Proof} \textit{As is proved in [S-Y,1] (see also [Sa-U,1,2]) given }$(T^3, g)$ \textit{one can find a compact stable minimal torus, }$T^2.$ \textit{Since }$\chi(T^2) = 0,$ \textit{by using the stability inequality (B.4) we conclude that the scalar curvature }$R$ \textit{of }$(T^3, g)$ \textit{must be negative somewhere - unless it is identically zero. However, if }$R \equiv 0$ \textit{but }$\text{Ric} \not\equiv 0,$ \textit{then ([Bo], see also [K-W]) there is a new metric }$g_1$ \textit{with positive scalar curvature, a contradiction. Consequently }$\text{Ric} \equiv 0.$ \textit{But on a three manifold this implies the sectional curvature is everywhere zero. Thus, }$g$ \textit{is flat. Q.E.D.}

This theorem has been improved by Schoen-Yau, and Gromov-Lawson to give quite detailed information on which compact manifolds admit metrics of positive scalar curvature. In particular, for all }$n$ \textit{the torus }$T^n$ \textit{has no metric of non-negative curvature except the flat metric (see \cite{G-L,1}). Gromov-Lawson \cite{G-L,3} have recently found much more information on complete (non-compact) manifolds having metrics with positive scalar curvature.}

One easy corollary of Theorem B.5 is the following, which, as we shall see shortly, is a special case of the positive energy theorem.

\textbf{Corollary B.6} \textit{Let }$g$ \textit{be a metric on }$\mathbb{R}^3$ \textit{with the properties}

\begin{enumerate}
\item[(i)] $g$ \textit{is standard metric }$\delta_{ij}$ \textit{outside a compact set},
\item[(ii)] the scalar curvature of }$g$ \textit{is non-negative. Then }$g$ \textit{is the standard metric everywhere.}
**POSITIVE ENERGY**

**Proof** Place the compact set in a large cube. Continue this cube periodically to obtain a torus to which the previous result applies. Q.E.D.

To relate this corollary to the positive energy theorem, we restate the dominant energy condition (A.2) in terms of the intrinsic geometry of the space-like hypersurface \((M, g) \hookrightarrow (N, \gamma)\) with second fundamental form \(h_{ij}\) and scalar curvature \(R_M\) (see [G,p.408]). Let

\[\mu = \frac{1}{2}[R_M - \|h\|^2 + \langle h^\ell, i \rangle^2]\]

and

\[J^i = \nabla_j [h^{ij} - h^\ell, i g^{ij}]\ .\]

Then the dominant energy condition (A.2) is replaced by

(B.7) \[\mu \geq |J^i J_i|^{\frac{1}{2}}\ .\]

Note that in the special case where \(h^\ell = 0\), the condition (B.7) implies that \(R_M \geq 0\), while assumption (i) of Corollary B.6 may be viewed as a strong version of the asymptotic Euclidean condition. Consequently, the corollary is indeed a special case of the positive energy theorem. Next, we relax the stringent assumption (i) of Corollary B.6.

2. **A special case.** To extend the above proof to cover the situation of interest in general relativity, one must replace the portions where compactness of \(M\) was used by corresponding statements for non-compact but asymptotically flat manifolds. The compactness entered in two places; (i) constructing the stable minimal submanifold, and (ii) via the Gauss-Bonnet theorem.

321
Schoen-Yau first treat a special case, making the assumption - just as above - that \( \Sigma h^k_k = 0 \). To repeat, in this situation the dominant energy condition (B.7) then implies that the scalar curvature \( \Sigma R_m \geq 0 \). The positive energy theorem is then a consequence of

**Theorem (B.8)** Let \((M,g)\) be an asymptotically flat oriented three manifold. If \( \Sigma R_m \geq 0 \), then \( \Sigma E_m \geq \Sigma |P_m| \) for each end \( M_\alpha \). Moreover, if \( E_\alpha = 0 \) for some end \( M_\alpha \), then \((M,g)\) is isometric to \( \mathbb{R}^3 \) with its standard flat metric.

Reasoning by contradiction, Schoen-Yau reduce to the situation where the metric satisfies

\[
g_{ij} = \left( 1 + \frac{2m_\alpha}{r} \right) \delta_{ij} + \frac{1}{r^2}
\]

on each end \( M_\alpha \), with \( P_m = 0 \) (see [S-Y,3]), with \( R_m \geq 0 \) and \( R_m > 0 \) outside of a compact set. By a computation, \( E_\alpha = m_\alpha \Sigma^{-1} \).

Again by contradiction, assume \( m_\alpha < 0 \) for some end \( M_\alpha \). Using this assumption, one proves the existence of an appropriate area minimizing surface \( S \) in \( M \). The area minimizing property gives the stability inequality (B.2). A geometric argument, which substitutes for the Gauss-Bonnet theorem in the proof of Theorem B.5 gives a contradiction, thus establishing that \( m_\alpha \geq 0 \) (and hence \( E_\alpha \geq 0 \)) for each end \( M_\alpha \).

To complete the proof, one must show that if \( R_m \geq 0 \) and some \( m_\alpha = 0 \) then \((M,g)\) is just \( \mathbb{R}^3 \) with its standard flat metric. Because \( \dim M = 3 \) and \( M \) is asymptotically flat, it is enough to show that \( \text{Ric}_M(g) \equiv 0 \). One first shows that, under these assumptions, \( R_m \equiv 0 \) (if not, one can find a conformal asymptotically flat metric with
zero scalar curvature and negative total energy, which contradicts the first part of this proof. Then by a perturbation analysis (as in [K-W, Lemma 5.2], see also [K]) one shows that if $\text{Ric}_M(g) \not\equiv 0$ the new metric $g_t = g - t \text{Ric}_M(g)$ has scalar curvature $R_t \geq 0$ for all small $t > 0$, but has negative total energy. This again contradicts the first part of the proof. Thus $\text{Ric}_M(g) = 0$.

3. The general case. Up to now, we have made the restrictive assumption that the second fundamental form $h$ of $M$ satisfies $h_{ij} = 0$. The general case of the theorem is proved by deforming both the metric $g$ and the embedding to the situation of the special case discussed above. It involves a difficult and technical existence proof of the existence of a solution $f$ to the mean curvature equation (first proposed by P.S. Jang)

$$\sum_{i,j} g_{ij} \frac{D_i D_j f}{(1 + |Df|^2)^k} = \sum_{i,j} g_{ij} h_{ij},$$

where $D_i$ is the covariant derivative on $(M, g)$, and $g^{-1}_{ij}$ is the inverse of $g_{ij} = g_{ij} + f_{i} x_j^j$. Since any adequate summary would be too long, we just refer the reader to the paper [S-Y, 4].
C. Second Proof (Witten)

1. The idea. Witten's procedure is similar to that used to prove most "vanishing theorems" in geometry. To describe the procedure, let $(M,g)$ be a Riemannian manifold, possibly with boundary, and let $L$ be a differential operator of the form

\[ Lu = V^*Vu + Qu , \]

where $u$ is a section of some bundle with covariant derivative $V$, having $V^*$ as its formal adjoint, and $Q$ is a self-adjoint endomorphism of the bundle. If $Lu = 0$, then, taking the inner product of (C.1) with $u$ and integrating over $M$, we have

\[ \int_M (|Vu|^2 + (u,Qu)) = \int_{\partial M} \text{(something)} , \]

where the boundary integral results from the integration by parts.

In practice, the operator $L$ is a natural operator - such as the Hodge Laplacian on differential forms, or the square of the Dirac operator on spinors - and $Q$ is expressed in terms of the curvature of $(M,g)$. Equation (C.1) is often called a "Weitzenböck formula". If one assumes that $Q > 0$ and $M$ is compact without boundary, then (C.2) implies that $u = 0$, that is, $\ker L = 0$. If $M$ is not compact or has a boundary, then one must impose either growth conditions or boundary conditions on $u$ to control the boundary integral in (C.2).

To be brief, for this proof of the positive energy theorem, the operator $L$ will be $D^2$, where $D$ is the Dirac operator on spinors, the endomorphism $Q$ will be positive because of the dominant energy condition, while, by choosing the spinor $u$ appropriately, the boundary
integral will exactly be the integrals (A.4) - (A.5) for the difference $E - |P|$. Thus, the identity (C.2) will be the desired expression of $E - |P|$ as a sum of squares.

With hindsight, one should anticipate that Clifford algebras (and spinors) play a significant role in Riemannian differential geometry, perhaps even more so than the exterior algebra of differential forms. The reason is that the exterior algebra does not utilize the metric, while the inner product itself is directly used as the basic quadratic form in the construction of the Clifford algebra. Thus, the Clifford algebra, with its differential operators, inherently embodies more information concerning the inner product and metric than the exterior algebra does.

2. Some details. We will consider the Dirac operator $D$ acting on the spinor fields, $S$, of the space-like hypersurface $M$. To define $D$, Witten does not use the intrinsic spinor covariant derivative on $(M, g)$, but rather uses the full space-time covariant derivative on $(N, \gamma)$ restricted to $M$. Thus if $\{e_u\}$ is an adapted orthonormal frame field for $N$, with $e_0$ normal to $M$ and $e_1, e_2, e_3$ tangent to $M$, then, in terms of the corresponding coframe $\{e^i\}$, the Dirac operator is

$$D u = \sum_{i=1}^{3} e^i \cdot \nabla_i u$$

where $\cdot$ is Clifford multiplication and $u \in \Gamma(S)$.

Step 1 (Weitzenböck formula) One computes $D^2$ and the formal adjoint operators $D^*$ and $\nabla^*$. It turns out that $D = D^*$ and $\nabla^* = - \nabla + \text{(correction)}$. These give
(C.3) \[ \mathcal{D}^2 = \nabla^* \nabla + Q, \]

where

\[ Q = \frac{1}{4}(R^2 + 2R_{00} + 2R_{01}^0 e^0 \cdot e^1) \in \text{End}(S). \]

The key observation relating \( Q \) to the physical assumptions is that by using Einstein's equations (A.1) we may rewrite \( Q \) as

\[ Q = 4\pi G(T_{00} + T_{01}^0 e^0 \cdot e^1) \]

\[ \geq 4\pi G[T_{00} - (-T_{01}^0) \frac{1}{2}] . \]

Thus the dominant energy condition (A.2) gives \( Q \geq 0 \). One can not help but be impressed at how well this fits \( \mathcal{D}^2 \) in (C.3). (In contrast, if one uses the intrinsic covariant derivative on \( M \), then, as Lichnerowicz first observed [Li], \( Q = \frac{1}{4} \) (scalar curvature of \( M \)), which is less helpful here - but obviously useful in discussing scalar curvature).

**Step 2.** Next one must solve the Dirac equation \( D_u = 0 \) on \( M \) in such a way that one has control of the asymptotic behavior of \( u \) on each end \( M^a \), in order to evaluate the right side of (C.2). So far, there is no adequate theory of linear elliptic operators on complete, noncompact manifolds. Fortunately, there is a good theory in the special case of \( \mathbb{R}^n \), with its flat metric, as long as the given elliptic operator is asymptotic to one with constant coefficients. This work was first carried out by Nirenberg-Walker [N-W], and developed further in [Ca], [CS-CB], [CB-C], and [P-T].

In our application, the asymptotic Euclidean assumption on \( (M,g) \) enables us to use the above theory to show that given any constant spinor field \( u_0 \) (it may be a different constant in each end,
\( u_0 = u_{0\alpha} \text{ in } M_\alpha \), there is a unique spinor field \( u \) such that

\[
\mathcal{D}u = 0 \quad \text{and} \quad u = u_0 + o\left(\frac{1}{r^{1-\varepsilon}}\right),
\]

that is,

\[
u = u_{0\alpha} + o\left(\frac{1}{r^{1-\varepsilon}}\right) \text{ in } M_\alpha .
\]

**Step 3.** For this solution \( u \) of (C.5), compute the boundary integral in (C.2). Of course, now the boundary integrals are the limits of boundary integrals taken over large spheres at each end, \( M_\alpha \), just as in (A.4) - (A.5). The result is

\[
\frac{1}{4\pi G} \int_M [\|\nabla u\|^2 + (u,Qu)] = \sum_{\alpha=1}^n [E_\alpha|u_{0\alpha}|^2 + \langle u_{0\alpha},P_{\alpha}u_{0\alpha}\rangle],
\]

where, for any constant spinor \( v \), we let

\[
P_{\alpha}v = \sum_{j=1}^3 p_{\alpha j}dx^j \cdot v,
\]

with \( \{dx^\mu\} \) the standard basis for \( \mathbb{T}^*(\mathbb{R}^3,1) \). Note that \( P_{\alpha} \in \text{End}(\mathcal{S}) \) is self-adjoint with \( P^2_{\alpha} = |P_{\alpha}|^2 \text{Id} \) (here \( |P_{\alpha}|^2 = p_{\alpha 1}^2 + p_{\alpha 2}^2 + p_{\alpha 3}^2 \)). Thus, the eigenvalues of \( P_{\alpha} \) are \( \pm |P_{\alpha}| \) so for each end \( M_\alpha \) there is a constant spinor \( u_{0\alpha} \) of length one so that \( P_{\alpha}u_{0\alpha} = -|P_{\alpha}|u_{0\alpha} \) (no sum on \( \alpha \) here). With this choice of \( u_0 \), equation (C.6) and the positivity of \( Q \), make obvious the desired positivity:

\[
\frac{1}{4\pi G} \int_M [\|\nabla u\|^2 + (u,Qu)] = \sum_{\alpha=1}^n [E_\alpha|u_{0\alpha}|^2 + \langle u_{0\alpha},P_{\alpha}u_{0\alpha}\rangle] \geq 0 .
\]

**Step 4.** It remains to show that if some \( E_\alpha = 0 \), then there is only one end and the metric on \( M \) is flat. As in [P-T], the key observations are that (i) if a spinor field \( u \) satisfies \( \nabla u = 0 \) in \( M \),
and \( u(x) + 0 \) as \( |x| \to \infty \) in one end, then \( u \equiv 0 \), and the related fact that (ii) if \( \{u_i\} \) are smooth spinor fields with \( \nabla u_i = 0 \) and with the \( \{u_i\} \) linearly independent in one end, then they are linearly independent everywhere.

References


POSITIVE ENERGY


