# PHYSICAL REVIEW D

## PARTICLES AND FIELDS

THIRD SERIES, VOLUME 27, NUMBER 12 15 JUNE 1983


# Witten's expression for gravitational energy


Gary T. Horowitz and Andrew Strominger

*Institute for Advanced Study, Princeton, New Jersey 08540*



(Received 25 October 1982)



The origin and physical interpretation of Witten's integral expression for gravitational energy is examined. It is shown that this expression (including nonzero $T_{\mu\nu}$) arises naturally from a Hamiltonian treatment of "classical supergravity," without consideration of the quantum theory. In addition, this expression is evaluated explicitly in several examples. It is found that the individual terms do not have a simple physical interpretation. New integral expressions are then introduced in which the gravitational energy divides naturally into a conformally invariant and nonconformally invariant contribution. It is argued that in certain circumstances these contributions can be interpreted as the gravitational radiation energy and total binding energy for strongly gravitating systems.


## I. INTRODUCTION

Perhaps one of the most important advances in general relativity in recent years is Witten's discovery of an integral expression for gravitational energy.[1] Since there does not exist a local gravitational energy density, the total energy of an isolated system is defined only in terms of the asymptotic behavior of the fields. It was generally believed that this energy could not be expressed as a simple integral over a spacelike three-surface of a positive-definite quantity. However, Witten's expression is precisely of this form.

An immediate consequence of this expression is that the total energy at spatial infinity— the Arnowitt-Deser-Misner (ADM) energy[2] $E$—must be positive. (This result was first proved by Schoen and Yau.[3]) The simplicity of Witten's expression has encouraged its application to a number of related problems. A minor modification leads to a proof of the positivity of the total energy at null infinity[4]—the Bondi energy—which represents the energy remaining after radiation has been emitted from the system. Further modifications have been used to prove the positivity of $E^2-Q^2$ where $Q$ is the total electric charge of the spacetime,[5] and the positivity of a quantity that can be interpreted as the total mass in asymptotically anti–de Sitter spacetimes.[6,7] All of these results have been extended to the case where the spacetimes contain black holes.[7] Witten's approach has also been used to prove the positivity of energy in several classical supersymmetric theories.[8] Finally, one can obtain an exact analog of the quadrupole formula for general relativity.[9]

Witten's original expression can be obtained as follows. Let $M,g_{\mu\nu}$ be an asymptotically flat spacetime satisfying the dominant energy condition: $T_{\mu\nu}t^{\mu}\hat{t}^{\nu}\geq 0$ for all future-directed $t^{\mu}$, $\hat{t}^{\nu}$. Let $\Sigma$ be an asymptotically flat three-surface with induced metric $q_{ab}$, and let $D_a$ be the projection of the four-dimensional covariant derivative into $\Sigma$. It has been shown[10,11] that there exists a unique spinor $\alpha$ which satisfies the "spatial Dirac equation"

$$\not{D}\alpha=0 \tag{1.1}$$

and approaches a constant spinor $\alpha^0$ at infinity. Let $\xi^{\mu}=\bar{\alpha}\gamma^{\mu}\alpha$ be the Dirac current of $\alpha$, and choose $\alpha^0$ such that $\xi^{0\mu}$ equals the unit normal to $\Sigma$, $t^{\mu}$.[12] Then by taking a second derivative of (1.1), commuting derivatives, multiplying by $\alpha^{\dagger}$, and integrating over $\Sigma$, one finds[1,13]

$$E=\int_{\Sigma}[T_{\mu\nu}\xi^{\mu}t^{\nu}+2(D_m\alpha)^{\dagger}(D^m\alpha)]d\Sigma\ . \tag{1.2}$$

This is Witten's expression relating the total energy to a positive integral over a spacelike three-surface.

The purpose of this paper is to investigate the physical interpretation of (1.2) in general relativity and its origin in the theory of supergravity. Why is there such a simple expression for gravitational energy in general relativity? Why does it involve spinors? The use of spinors was originally motivated by consideration of the quantum theory of supergravity. We will show, however, that the supersymmetric origin of Witten's expression can be understood in terms of classical supergravity. This theory is an extension of general relativity which admits an enlarged group of symmetry transformations. As we show in Sec. II, these symmetries enable one to see relations in pure general relativity that might have otherwise remained obscure. In particular, diffeomorphisms can be expressed as the square of supersymmetry transformations. In the Hamiltonian formulation, one obtains a classical equation relating the generator of diffeomorphisms to the square of the generator of supersymmetry transformations. For the case that this diffeomorphism is a global time translation, one obtains Witten's expression for the total energy.

Witten's expression is, of course, valid for spacetimes in which the stress energy is nonzero. Such spacetimes are not described by classical supergravity, since the addition of the source term to the action necessary to produce the classical equation $G_{\mu\nu}=T_{\mu\nu}$ violates supersymmetry. In Sec. II we show, however, that by coupling supermatter one can derive Witten's expression for a spacelike surface with nonzero $T^{0\mu}$.

What is the physical interpretation of Witten's expression from the standpoint of classical general relativity? Equation (1.2) expresses the total energy as the integral of two terms—one which involves the stress energy tensor directly and one which does not. At first sight, it therefore seems natural to interpret these two terms as the matter and gravitational contributions to the energy.

To see whether this interpretation is justified, we require more information about the behavior of the spinor $\alpha$ satisfying (1.1). The matter term in (1.2) is just $T_{\mu\nu}\xi^\mu t^\nu$. The qualitative behavior of the vector $\xi^\mu$ can be found from a simple argument due to Reula[11]: Consider the following functional of spinors $\psi$ which satisfy $\bar\psi\gamma^\mu\psi\rightarrow t^\mu$ asymptotically:

$$\mathscr{E}[\psi]=\int_\Sigma [T_{\mu\nu}(\bar\psi\gamma^\mu\psi)t^\nu+2(D_m\psi)^\dagger(D^m\psi)]d\Sigma\ . \tag{1.3}$$

The requirement that $\psi$ minimize this functional is equivalent to having $\psi$ satisfy (1.1). So the minimum of (1.3) is precisely Witten's expression. Since $T_{\mu\nu}$ satisfies the dominant energy condition, the minimum of $\mathscr{E}$ is obtained by decreasing the norm of $\xi^\mu=\bar\psi\gamma^\mu\psi$ in regions where $T_{\mu\nu}\neq 0$. Thus, the qualitative behavior of $\xi^\mu$ agrees with one's intuition about a "gravitational red-shift factor," and hence supports the interpretation of $T_{\mu\nu}t^\mu\xi^\nu$ as the matter contribution to the total energy.

Can one obtain more quantitative information about both the direction and magnitude of $\xi^\mu$? With respect to the direction of $\xi^\mu$, one can determine when $\xi^\mu$ will be orthogonal to $\Sigma$. We will show in Sec. III that *$\xi^\mu$ is everywhere orthogonal to $\Sigma$ if and only if $\mathscr{E}$ is a maximal surface,* i.e., the trace of the extrinsic curvature of $\Sigma$, $\pi=\pi_m{}^m$, vanishes. This result removes a slight difference between the positive-energy theorem proved by Schoen and Yau[3] and the one proved by Witten.[1] Although both theorems require the dominant energy condition on $T_{\mu\nu}$ to prove $E\geq 0$ for an arbitrary slice $\Sigma$, Schoen and Yau show that if $\pi=0$ then it suffices to impose the weak energy condition: $T_{\mu\nu}t^\mu t^\nu\geq 0$. The fact that $\xi^\mu$ is orthogonal to a maximal slice $\Sigma$ shows that in Witten's proof as well one can weaken the energy condition required on a maximal slice.

With respect to the magnitude of $\xi^\mu$, it is interesting to evaluate $\xi^\mu$ on a static slice $\Sigma$ in a static spacetime. There are at least two reasons, one physical and one mathematical, for suspecting that $\xi^\mu$ might agree with the timelike Killing field $\eta^\mu$ evaluated on $\Sigma$. The physical reason is that both vectors can be interpreted as a measure of the gravitational red-shift. The mathematical reason is that we know from the above results that $\xi^\mu$ and $\eta^\mu$ are parallel, both have norm one at infinity and decreasing norm in the interior, and both are obtained in a natural way from the geometry. Nevertheless, by examining specific examples, we find in Sec. III that these vectors are not equal. *For a static slice in a (nonflat) static spacetime, $\xi^\mu$ does not agree with the timelike Killing field.* Since the norm of the Killing field $\eta^\mu$ is the red-shift factor only for test fields, there is perhaps little reason to require this function to also be the red-shift factor for finite matter. In order to see whether the norm of $\xi^\mu$ yields a reasonable red-shift factor for finite matter, we must turn to the gravitational term in Witten's expression.

Intuitively, one expects two different types of contributions to gravitational energy. There is presumably a positive contribution from gravitational radiation and a negative contribution from gravitational binding energy. Yet in Witten's expression, there is only one positive term other than the matter. Where is the negative binding energy? To answer this question, it is instructive to consider Newtonian gravity.

The total energy of a Newtonian system is

$$E_N=\int(\mu-\partial_m\phi\partial^m\phi)d^3x\ , \tag{1.4}$$

where we have included the rest-mass energy $\mu$. The gravitational energy is clearly negative. However, there are different ways of distributing the energy between the matter and gravitational terms. Since

$$\partial^2\phi=\tfrac{1}{2}\mu\ , \tag{1.5}$$

one can integrate by parts and reexpress (1.4) in the form

$$E_n=\int[\mu(1+\tfrac{1}{2}n\phi)+(n-1)\partial_m\phi\partial^m\phi]d^3x\ , \tag{1.6}$$

where $n$ is any real number. For $n>1$, the gravitational contribution to the energy is positive. Roughly speaking, as $n$ increases, one increases the "red-shift" of the matter (recall $\phi<0$) and therefore must compensate by adding a positive contribution from the gravitational field. We will show in Sec. III that the Newtonian limit of Witten's expression is precisely (1.6) with $n=4$. This helps to explain why there is no negative-binding-energy term in Witten's expression. It also gives some understanding of why $\xi^\mu$ does not agree with the timelike Killing field $\eta^\mu$ in static spacetimes. In the Newtonian limit, the red-shift associated with the Killing field corresponds to $n=2$.

One can view Witten's expression as the analog of (1.6) with $n=4$ for an arbitrary spacetime. It is thus somewhat misleading to interpret the individual terms in Witten's expression as the "matter contribution" and "gravitational contribution" to the total energy. What one would like is the analog of (1.6) with $n=0$ for an arbitrary spacetime since this expression is closer to one's physical intuition. Does such an expression exist? We will show in Sec. IV that modulo a technical conjecture about the spinors $\alpha$ the answer is yes. In fact, for any $n$, there exists an integral expression for the total energy which reduces to (1.6) in the Newtonian limit. If $\pi=0$, then the gravitational contribution to the energy divides naturally into a conformally invariant $\mu_{\rm I}$ and nonconformally invariant $\mu_{\rm NI}$ part. We show that in the limit of a weak gravitational wave, $\mu_{\rm NI}$ vanishes and $\mu_{\rm I}$ reduces to the standard expression for the energy density of a spin-two field. This supports the interpretation in strongly curved spacetimes of $\int\mu_{\rm I}d\Sigma$ as the energy in gravitational radiation and $\int\mu_{\rm NI}d\Sigma$ as the total binding energy of the system.

The outline of this paper is as follows. In Sec. II we discuss the Hamiltonian formulation of classical supergravity and show how Witten's expression arises naturally in this framework. The case when $T_{\mu\nu}=0$ is discussed first and the argument is then generalized to allow nonzero $T_{\mu\nu}$. Section III begins with a proof that $\xi^\mu$ is orthogonal to $\Sigma$ if and only if $\pi=0$. We then show, by means of examples, that on a static slice in a static spacetime $\xi^\mu$ does not agree with the timelike Killing field. Finally, we discuss the generalization of (1.6) to an arbitrary conformally flat initial-data set. The generalization to an arbitrary initial-data set is described in Sec. IV. The effect of conformal transformations on the individual terms in this expression is examined and their physical interpretation is discussed.

To make contact with supergravity literature, we use four-dimensional Dirac spinors in Sec. II. However, in Secs. III and IV, it is convenient to work entirely with three-dimensional SU(2) spinors because the three-dimensional analog of (1.1) is conformally invariant.

## II. SUPERGRAVITY AND WITTEN'S EXPRESSION

It was observed by Teitelboim[14] that the quantum supergravity Hamiltonian $\mathscr{H}$ may be written

$$\mathscr{H}=\frac{1}{\hbar}\mathrm{tr}Q^2\ , \tag{2.1}$$

where $Q$ is the supercharge. Deser and Teitelboim[15] then argued that this implies that quantum supergravity has positive energy. Grisaru[16] commented that the positive-energy theorem in classical relativity follows by taking the $\hbar\to0$ limit. Inspired by these arguments, Witten[1] gave a rigorous and elegant proof.

Supergravity has already proven to be a very powerful tool for investigating the structure of classical relativity. Consideration of the *quantum* theory, and the subsequent $\hbar\to0$ limit, however, is an unnecessarily cumbersome procedure. Instead, one may work entirely within the context of "classical supergravity." This is a classical field theory in which the fields take values in a Grassmann algebra.[17] In a Hamiltonian formulation, equations of motion are generated by Dirac brackets.[18] The classical analog of (2.1) may be directly derived.[19] At the end of the calculation, physically relevant relations are obtained by setting all anticommuting fields to zero. These fields therefore have no physical interpretation—they are solely a calculational device.

This approach avoids many of the difficulties of quantum supergravity. Classical supergravity is a far simpler theory. Much of the voluminous literature on supergravity is concerned with the nonclosure of the supersymmetry algebra. The nonclosure terms vanish, however, when the field equations are satisfied—so this is less troublesome in the classical theory.

We view the supersymmetric extension of general relativity to classical supergravity as much like the extension of real functions to the complex plane. Many properties of real functions (e.g., certain integrals) are most easily discovered by considering their analytic extensions. Similarly, the supersymmetric extension of general relativity allows a greatly expanded group of symmetry transformations. These provide a powerful tool to derive properties and understand relations which might otherwise remain obscure.

Witten's integral expression for gravitational energy, although mysterious from other viewpoints, arises naturally in the context of classical supergravity. In this section, we will derive this expression using the Dirac Hamiltonian formulation of classical supergravity. Many of the details of Hamiltonian supergravity are (fortunately) irrelevant for the purpose of deriving (1.2), and will be omitted. For more comprehensive discussions of Hamiltonian supergravity, the reader is referred to the literature.[14,20] We begin with an outline of our argument and then proceed with a more detailed discussion.

Consider the action for $N=1$ supergravity[21]:

$$S=-\tfrac{1}{2}\int d^4x(eR+i\epsilon^{\mu\nu\rho\sigma}\bar{\psi}_\mu\gamma_5\gamma_\nu\nabla_\rho\psi_\sigma)\ , \qquad (2.2)$$

where $e$ is the determinant of the vierbein $e_{a\mu}$, $\nabla_\rho$ is a covariant derivative with torsion,[21] and $\bar{\psi}_\mu$ and $\psi_\nu$ are anticommuting spinor vector fields. $S$ is an even element of a Grassmann algebra. The field equations are

$$\epsilon^{\mu\nu\rho\sigma}\gamma_5\gamma_\nu\nabla_\rho\psi_\sigma=0\ , \qquad (2.3a)$$

$$eG^{a\nu}=\frac{i}{2}\bar{\psi}_\lambda\gamma_5\gamma^a\epsilon^{\lambda\nu\rho\sigma}\nabla_\rho\psi_\sigma\ . \qquad (2.3b)$$

For $\psi_\mu=0$, these reduce to the vacuum Einstein equation.

The action (2.2) is invariant under local supersymmetry transformations [parametrized by a spinor $\epsilon(x)$] up to a total derivative:

$$\delta_\epsilon S=\int d^4x\ \nabla_\mu\theta^\mu_\epsilon\ , \qquad (2.4)$$

where "$\delta_\epsilon$" generates an infinitesimal local supersymmetry transformation. The conserved supercurrent associated with this symmetry can be obtained by the standard Noether construction:

$$J^\mu_\epsilon=\frac{\delta S}{\delta\nabla_\mu\Phi}\delta_\epsilon\Phi-\theta^\mu_\epsilon\ . \qquad (2.5)$$

The conserved Noether charge is

$$Q\cdot\epsilon=\int_\Sigma d\Sigma_\mu J^\mu_\epsilon\ . \qquad (2.6)$$

$\Phi$ represents all the fields (indices suppressed) and $d\Sigma_\mu$ is the volume element on the hypersurface $\Sigma$ on which $Q\cdot\epsilon$ is defined.

In analogy to general relativity[22] one can view transformations $\delta_\epsilon$ in which $\epsilon\to0$ asymptotically as "pure gauge" transformations, and those in which $\epsilon\to$constant as "physical supersymmetry" transformations.[23] To cast the theory in Hamiltonian form, we constrain this gauge freedom. After this is done, $\epsilon$ is uniquely determined by its value at infinity. We then adopt the Hamiltonian formalism for constrained systems developed by Dirac[19] and extended to anticommuting fields by Senjanovic, Casalbuoni, and others.[14,20,24] In this formalism, $Q$ generates physical supersymmetry transformations via its Dirac bracket. As usual, time translations are generated by the Hamiltonian:

$$\dot{\Phi}(x)=\{\mathscr{H},\Phi(x)\}_D\ , \qquad (2.7)$$

where $\{,\}_D$ denotes the Dirac bracket. However, one finds that time translations are also generated by the square of a supersymmetry transformation:

$$\begin{aligned}\dot{\Phi}(x)&=\mathrm{tr}\{Q,\{Q,\Phi(x)\}_D\}_D\\&=\frac{\mathrm{tr}}{2}\{\{Q,Q\}_D,\Phi(x)\}_D\ .\end{aligned} \qquad (2.8)$$

The second equality follows from an extra minus sign in the definition of the Dirac bracket for two anticommuting fields.[24] Equations (2.7) and (2.8) imply

$$\tfrac{1}{2}\mathrm{tr}\{Q,Q\}_D=\mathscr{H}\ . \qquad (2.9)$$

Explicit evaluation of this expression yields Witten's expression for vacuum spacetimes.

To obtain Witten's expression for spacetimes with matter, one repeats the above analysis with the addition of supersymmetric matter to the action (2.2).

A nice feature of the above construction is that, if (2.9) is evaluated for $\psi_\mu=0$ configurations, *the result is independent of the canonical decomposition of the metric.* Thus, although in principle the metric must be put into canonical form in order to define the Dirac brackets, in practice it is unnecessary to do so explicitly. This is because the terms in (2.9) proportional to Dirac brackets of the metric are also proportional to $\psi_\mu$.

We now proceed with details of the derivation. The local supersymmetry transformations are

$$\begin{aligned}&\delta_\epsilon e^m{}_\mu(x)=i\bar{\epsilon}(x)\gamma^m\psi_\mu(x)\ ,\\&\delta_\epsilon\psi_\lambda(x)=2\nabla_\lambda\epsilon(x)\ ,\end{aligned} \qquad (2.10)$$

where $\epsilon(x)$ is an arbitrary anticommuting spinor. The action density $\mathscr{L}(x)$ changes by a total derivative under this transformation:

$$\begin{aligned}&\delta_\epsilon\mathscr{L}(x)=\nabla_\mu\theta^\mu_\epsilon(x)\ ,\\&\theta^\mu_\epsilon(x)=i\epsilon^{\mu\nu\rho\sigma}\bar{\psi}_\nu(x)\gamma_5\gamma_\rho\nabla_\sigma\epsilon(x)\ .\end{aligned} \qquad (2.11)$$

It follows directly from (2.10) that $\delta_\epsilon$ is the "square root" of a coordinate transformation.

Equations (2.10), along with the spin-$\frac{3}{2}$ field equation (2.3a) imply[21]

$$[\delta_\epsilon,\delta_{\epsilon'}]=\delta_G(K^\mu)+\delta_L(K^\mu\omega_{\mu mn})+\delta_{-K^\mu\psi_\mu}\ , \qquad (2.12)$$

where

$$K^\mu(x)=2i\bar{\epsilon}'(x)\gamma^\mu\epsilon(x)\ .$$

This relation holds, after use of (3.3a) for $[\delta_\epsilon,\delta_{\epsilon'}]$ acting on any function of the fields. $\delta_G(K^\mu)$ is the generator of transformations $x^\mu \rightarrow x^\mu+K^\mu$, $\delta_L$ rotates the tetrad frames, and $\delta_{-K^\mu\psi_\mu}$ is again a supersymmetry transformation. Notice that when $[\delta_\epsilon,\delta_{\epsilon'}]$ acts on a $\psi_\mu=0$ configuration, the resulting configuration again has $\psi_\mu=0$. This is the main reason why supersymmetry is relevant to classical relativity: Diffeomorphisms of the spacetime may be represented as repeated supersymmetry transformations.

The conserved Noether current derived from supersymmetry is

$$J^\mu_\epsilon=-2i\epsilon^{\mu\nu\rho\sigma}\bar{\psi}_\nu\gamma_5\gamma_\rho\nabla_\sigma\epsilon+\cdots\ , \qquad (2.13)$$

where the dots indicate torsion terms that vanish in the final expression for $E$ when the field equations are imposed and $\psi_\mu$ is set to zero. Of course, one cannot in general impose field equations before Dirac brackets are computed. In this case, however, we save some effort by noting that these terms cannot contribute to the final expression. The "$+\cdots$" will hereafter be implicit in all expressions. The conserved charge is

$$Q\cdot\epsilon=\int_\Sigma d\Sigma_\mu J^\mu_\epsilon\ . \qquad (2.14)$$

In order to show that $Q\cdot\epsilon$ generates supersymmetry transformations via its Dirac bracket, we must perform a Hamiltonian decomposition. Ultimately, however, we are only interested in those terms which do not vanish when $\psi_\mu=0$. Since $Q\cdot\epsilon$ and $J^\mu_\epsilon$ are linear in $\psi_\mu$, expressions such as (2.9) involve only Dirac brackets of $\psi_\mu$ and not of $e^a_\mu$. We are thus spared explicit consideration of the Hamiltonian decomposition of the vierbein.

The Hamiltonian treatment of the spin-$\frac{3}{2}$ field cannot be avoided. For a full discussion the reader is referred to the literature.[14,20] Because of the gauge invariance of $S$, not all of the fields represent physical degrees of freedom. One begins by computing the constraints necessary to eliminate the unphysical degrees of freedom. There are 8 first-class constraints and 12 second-class constraints. For each first-class constraint a gauge condition must be chosen. We choose[25]

$$\left.\begin{aligned}&\psi_0=0\ ,\\&\gamma^i\psi_i=0\ ,\end{aligned}\right.\ \text{on }\Sigma \qquad (2.15)$$

where $\Sigma$ is an asymptotically flat spacelike slice on which $\mathscr{H}$ will be defined. $i=1,2,3$ is a spatial index in the hypersurface $\Sigma$ and $\psi_0$ is the component normal to $\Sigma$. The Dirac bracket of $\psi_\mu$ and its conjugate momentum

$$\pi^\mu=-\frac{i}{2}\epsilon^{0\mu\alpha\beta}\bar{\psi}_\alpha\gamma_5\gamma_\beta \qquad (2.16)$$

may now be computed. The key idea of the Dirac formalism is to have nonzero brackets only between the "true degrees of freedom." This implies that the constraint equations or gauge conditions can be imposed at any stage of the calculation, i.e.,

$$\{\chi_i,\Phi\}_D=0\ , \qquad (2.17)$$

where $\chi_i$ is any constraint or gauge condition and $\Phi$ is any function of the fields and their momenta. The Dirac bracket $\{,\}_D$ is defined by

$$\{\Phi_A,\Phi_B\}_D=\{\Phi_A,\Phi_B\}_P-\{\Phi_A,\chi_i\}_P[\{\chi_i,\chi_j\}_P]^{-1}\{\chi_j,\Phi_B\}_P\ , \qquad (2.18)$$

where $\{\Phi_A,\Phi_B\}_P$ is the ordinary Poisson bracket defined with an extra minus sign for anticommuting fields. For our gauge choice (2.15), the relevant Dirac bracket is[26]

$$\{\psi_i(x),\pi_j(y)\}_D=\tfrac{1}{2}[g_{ij}\delta^3(x,y)-\tfrac{1}{2}(D_iD_j+D_jD_i)S(x,y)-\tfrac{1}{2}\epsilon_{ijk}\gamma_0\gamma_5(D^k\not{D}+\not{D}D^k)S(x,y)]\ , \qquad (2.19)$$

where

$$D^kD_kS(x,y)=\delta^3(x,y)\ ,$$

$$\not{D}=\gamma^kD_k\ ,$$

and $D_i$ is the projection of $\nabla_\mu$ into $\Sigma$. There could also be curvature terms in (2.19). Their coefficients have not been computed, however, as one can show that they cannot contribute to the final expression.

Physical supersymmetry transformations may now be generated via Dirac brackets with the Noether charge $Q\cdot\epsilon$. The gauge condition $\gamma^k\psi_k=0$ requires that $\epsilon$ be an asymptotically constant solution of

$$\gamma^kD_k\epsilon=0\ . \qquad (2.20)$$

There are four linearly independent such solutions $\epsilon_N(x)$ $(N=1,2,3,4,)$ with nonzero asymptotic values $\lim_{x\rightarrow\infty}\epsilon_N(x)=\epsilon^0_N$. These four solutions determine the four supercharges $Q_N\equiv Q_{\epsilon_N}$ (Ref. 27):

$$Q_N=\int d^3\Sigma[2\pi^k D_k\epsilon_N+i(D^k\epsilon_N)^\dagger\psi_k]$$
$$=\int d^2S^k(2\pi_k\epsilon_N+i\epsilon_N^\dagger\psi_k)\ , \qquad (2.21)$$

where the constraints and gauge conditions must be used to obtain this expression from (2.13) and (2.14).

After a short calculation, the Dirac brackets of the supercharges are found to be

$$\{Q_N,Q_M\}_D=4i\int d^3\Sigma(D^k\epsilon_N)^\dagger D_k\epsilon_M\ . \qquad (2.22)$$

Comparing with (2.9) and (2.12), however, we realize this quantity also generates a global translation whose asymptotic value is $K^\mu=2i\bar\epsilon_M^0\gamma^\mu\epsilon_N^0$. Therefore

$$\{Q_N,Q_M\}_D=2i\bar\epsilon_M^0\gamma^\mu\epsilon_N^0P_\mu\ , \qquad (2.23)$$

where $P_\mu$ is the ADM four-momentum.

To extract an explicit formula for the ADM energy, we must factor out the anticommuting variables from (2.22) and (2.23). Define

$$\epsilon_N(x)=\theta_N\alpha_N(x)\ , \qquad (2.24)$$

where the $\theta_N$ are a basis for the Grassmann algebra, and $\alpha_N$ are commuting spinors obeying

$$\not D\alpha_N(x)=0\ ,$$
$$\lim_{x\to\infty}\alpha_N(x)=\alpha_N^0\ . \qquad (2.25)$$

Integrating[28] (2.22) and (2.23) over $\theta_N$ and $\theta_M$, we find

$$-2\int d^3\Sigma[(D^k\alpha_N)^\dagger D_k\alpha_M]=(\bar\alpha_N^0\gamma^\mu\alpha_M^0)P_\mu\ . \qquad (2.26)$$

Since (2.26) holds for all $\alpha_A$ satisfying (2.25), it follows that

$$E=2\int d^3\Sigma(D^k\alpha)^\dagger D_k\alpha\ , \qquad (2.27)$$

where $\alpha$ is any spinor satisfying (2.25) and has $\bar\alpha^0\gamma^\mu\alpha^0$ equal to the timelike normal to $\Sigma$ at infinity. When $\psi_\mu$ is set to zero, $D_i$ becomes the projection into $\Sigma$ of the standard (torsion-free) derivative operator. Equation (2.27) is then precisely Witten's expression for vacuum spacetimes.

One would like to extend this analysis to obtain the full Witten expression (1.2) for metrics obeying $G_{\mu\nu}=T_{\mu\nu}$. The addition of an arbitrary source term to the action, however, spoils supersymmetry and the connection between the energy and the squared supercharge. One is therefore led to consider supermatter as a source for the gravitational field. The literature contains many constructions of supersymmetric actions that incorporate supergravity and supermatter multiplets.

This procedure, however, will only lead to Witten's expression for configurations that are solutions of all (including the supermatter) field equations. It is unlikely that an arbitrary $T_{\mu\nu}$ corresponds to such a solution.

This is not, however, a serious difficulty because Witten's expression involves only the space-time and time-time components $T^{0\mu}$ of the stress energy at one moment of time, and not the space-space components $T^{ij}$. Therefore, as far as positive-energy theorems are concerned, it is completely general to consider supermatter for which the initial-value formulation on a spacelike hypersurface allows arbitrary specification of $T^{0\mu}$.

As a simple example we consider the derivation of Witten's expression for supersymmetric Einstein-Maxwell fields. The supermatter multiplet consists of a photon field $A_\mu$ and a neutrino field $\lambda$. $\gamma^k\psi_k=0$ remains an appropriate gauge condition and still implies $\not D\epsilon=0$. The volume integral expression for the supercharge acquires an extra contribution linear in the neutrino fields:

$$Q_N=\int d^3\Sigma(2\pi^kD_k\epsilon_N+iD^k\epsilon_N^\dagger\psi_k-\tfrac{1}{2}\bar\epsilon_N\sigma_{\mu\nu}F^{\mu\nu}\gamma^0\lambda)\ , \qquad (2.28)$$

where $F^{\mu\nu}$ is the Maxwell tensor. The constraint $D_k\psi^k=0$ becomes

$$D_k\psi^k=\frac{i}{4}\sigma_{\mu\nu}F^{\mu\nu}\gamma^0\lambda \qquad (2.29)$$

so the surface integral form for $Q_N$ remains unchanged:

$$Q_N=\int d^2S^k(2\pi_k\epsilon_N+i\epsilon_N^\dagger\psi_k)\ . \qquad (2.30)$$

When computing the Dirac bracket of the supercharge one finds (when using the volume integral form for $Q_N$) an extra contribution from the term linear in the neutrino field. This term is proportional to $\bar\epsilon\gamma^\mu\epsilon T_{0\mu}$, and one may thus obtain the full Witten expression (1.2).

To demonstrate that one may arbitrarily specify $T_{0\mu}$ in this theory, one must prove existence of solutions to the equations $T_{00}=\frac{1}{2}(\vec{E}^2+\vec{B}^2)$, $T_{0k}=\frac{1}{2}(\vec{E}\times\vec{B})_k$. Although this is four equations for four unknowns, the massless nature of $\vec{E}$ and $\vec{B}$ makes an existence proof difficult. This can be circumvented by considering more complicated supermatter containing, e.g., massive vector bosons.

## III. EXAMPLES

In order to gain a better physical understanding of the individual terms in Witten's expression, we now evaluate this expression for several spacetimes. It is convenient to work just with the initial data for the spacetime, which consists of a noncompact three-manifold $\Sigma$, an asymptotically flat positive-definite metric $q_{ab}$ on $\Sigma$, and a symmetric tensor field $\pi_{ab}$ representing the extrinsic curvature.

Witten's expression can be written in terms of intrinsic three-dimensional SU(2) spinors on the initial-data set.[29] Let $(\Sigma, q_{ab}, \pi_{ab})$ be an initial-data set satisfying the dominant energy condition

$$\mu \geq (J_a J^a)^{1/2} , \tag{3.1}$$

where $\mu$ and $J^a$ are the energy and momentum densities of the matter which are related to $q_{ab}$ and $\pi_{ab}$ by the constraint equations.[30] Let $D_{AB}$ denote the (torsion-free) covariant derivative compatible with $q_{ab}$ and define a new derivative operator by

$$\mathscr{D}_{AB}\lambda_C = D_{AB}\lambda_C + \frac{i}{\sqrt{2}}\pi_{ABC}{}^{D}\lambda_D , \tag{3.2}$$

where $\pi_{ABCD} = \pi_{(AB)(CD)}$ is the spinor representation of $\pi_{ab}$. Let $\beta^B$ be a solution to

$$\mathscr{D}_{AB}\beta^B = D_{AB}\beta^B + \frac{i\pi}{2\sqrt{2}}\beta_A = 0 , \tag{3.3}$$

which is the three-dimensional version of (1.1). Then one can show

$$D_m(\beta^{\dagger C}\mathscr{D}^m\beta_C) = \tfrac{1}{2}(\mu\epsilon_{MN} - i\sqrt{2}J_{MN})\beta^M\beta^{\dagger N} + (\mathscr{D}^{AB}\beta^C)^\dagger(\mathscr{D}_{AB}\beta_C) . \tag{3.4}$$

Integrating over $\Sigma$ thus yields

$$2\int_S(\beta^{\dagger C}\mathscr{D}_a\beta_C)dS^a = \int_\Sigma[(\mu\epsilon_{MN} - i\sqrt{2}J_{MN})\beta^M\beta^{\dagger N} + 2(\mathscr{D}^{AB}\beta^C)^\dagger(\mathscr{D}_{AB}\beta_C)]d\Sigma , \tag{3.5}$$

where $S$ is an asymptotic two sphere on $\Sigma$. If $\beta^A$ asymptotically approaches a constant spinor $\beta^{0A}$, then the surface integral is related to the total ADM energy $E$ and momentum $P^a$ of the initial-data set by

$$2\int_S(\beta^{\dagger C}\mathscr{D}_a\beta_C)dS^a = (E\epsilon_{MN} - i\sqrt{2}P_{MN})\beta^{0M}\beta^{0\dagger N} . \tag{3.6}$$

Notice that (3.6) involves the total three-momentum $P^a$ whereas (1.2) does not. To obtain the analog of (1.2) in terms of three-dimensional spinors one must use two solutions of (3.3). Let $\beta^{0A}$ be any constant spinor normalized so that $\beta^{0\dagger A}\beta^0_A = \frac{1}{2}$. Let $\beta^{1A}$ be the solution to (3.3) which asymptotically approaches $\beta^{0A}$ and $\beta^{2A}$ be the solution which approaches $\beta^{0\dagger A}$. We define

$$f = \beta^{j\dagger A}\beta^j_A , \tag{3.7}$$

$$v^{AB} = -i\beta^{j\dagger(A}\beta^{jB)} , \tag{3.8}$$

where here and throughout a sum over the repeated index $j = 1,2$ is assumed. The function $f$ and real vector $v^a$ are, in fact, independent of the choice of $\beta^{0A}$.[31] Using the fact that $\beta^{0\dagger\dagger A} = -\beta^{0A}$, we evaluate (3.5) for each $\beta^{jA}$ and add the resulting expressions to obtain

$$E = \int_S D_a f dS^a = \int[(f\mu + v^a J_a) + 2(\mathscr{D}^{AB}\beta^{jC})^\dagger(\mathscr{D}_{AB}\beta^j_C)]d\Sigma . \tag{3.9}$$

This is the expression we wish to evaluate in specific examples. However, before we begin, we prove a general result, mentioned in the Introduction, about the behavior of $v^a$.

We now show that $v^a$ vanishes if and only if $\Sigma$ is a maximal slice. Suppose $\pi = 0$. Then

$$\mathscr{D}_{AB}\beta^{1\dagger B} = (\mathscr{D}_{AB}\beta^{1B})^\dagger = 0 . \tag{3.10}$$

Thus $\beta^{1\dagger A}$ is a solution to (3.3) which asymptotically approaches $\beta^{0\dagger A}$. Since this equation has unique solutions, $\beta^{2A}$ must be precisely $\beta^{1\dagger A}$. Hence $v^a = 0$. To show the converse, we use the fact that

$$0 = \mathscr{D}_{AB}v^{AB} = D_a v^a + \frac{1}{\sqrt{2}}\pi f . \tag{3.11}$$

Therefore, if $v^a = 0$, then $\pi = 0$. This completes the proof. In terms of the four-dimensional spinor $\alpha$ and four vector $\xi^\mu = \bar{\alpha}\gamma^\mu\alpha$, the requirement that $v^a = 0$ is equivalent to requiring that $\xi^\mu$ be orthogonal to $\Sigma$. Thus we have proved the result stated in the Introduction.[32]

We now begin our discussion of examples by evaluating Witten's expression on a static slice in the Schwarzschild spacetime. This example is of interest because one can compare the function $f$ [which looks like a gravitational red-shift factor in (3.9)] with the norm of the timelike Killing field (which *is* the red-shift factor for test fields). Fortunately, this example is easy to evaluate since one can obtain the solutions to (3.3) for this initial data by a conformal transformation.

Equation (3.3) is conformally invariant: If $\hat{q}_{ab} = \varphi^4 q_{ab}$, $\hat{\pi}_{ab} = \varphi^2\pi_{ab}$, and $\hat{\beta}^A = \varphi^{-3}\beta^A$, then $\hat{\mathscr{D}}_{AB}\hat{\beta}^B = \varphi^{-3}\mathscr{D}_{AB}\beta^B$. Therefore, if $\beta^A$ is a solution for $(q_{ab}, \pi_{ab})$, $\hat{\beta}^A$ will be a solution for $(\hat{q}_{ab}, \hat{\pi}_{ab})$.

As is well known, a static slice in the Schwarzschild geometry is conformally flat, with the conformal factor given by

$$\varphi = 1 + \frac{M}{2\rho} . \tag{3.12}$$

The coordinate $\rho$ is a radial coordinate in the flat three-space which is related to the standard Schwarzschild coordinate $r$ by

$$r=\rho\left[1+\frac{M}{2\rho}\right]^2 . \tag{3.13}$$

The constant $M$ is related to the total ADM energy $E$ on a static slice by $E=8\pi M$.

Since the only asymptotically constant solution to (3.3) when $q_{ab}$ is flat and $\pi_{ab}$ vanishes are the constant spinors $\beta^{0A}$, $\hat{\beta}^A=\varphi^{-3}\beta^{0A}$, and $\hat{\beta}^{\dagger A}=\varphi^{-3}\beta^{0\dagger A}$ are the solutions to (3.3) in Schwarzschild which asymptotically approach $\beta^{0A}$ and $\beta^{0\dagger A}$, respectively. Thus the function $f$ is simply given by

$$f=2\hat{\epsilon}_{AB}\hat{\beta}^A\hat{\beta}^{\dagger B}=\varphi^{-4} . \tag{3.14}$$

By inverting (3.13), we can express $\varphi$ in terms of $r$ to obtain

$$f=\tfrac{1}{4}\left[1-\frac{M}{r}+\left[1-\frac{2M}{r}\right]^{1/2}\right]^2 . \tag{3.15}$$

This is certainly *not* equal to the norm of the time-like Killing field:

$$(-\lambda)^{1/2}=\left[1-\frac{2M}{r}\right]^{1/2} . \tag{3.16}$$

In terms of four-dimensional quantities, this statement says that the vector $\xi^\mu$ does not agree with the timelike Killing field $\eta^\mu$ on a static slice. These two vectors are parallel to each other, but have different magnitudes. From (3.15) and (3.16) one sees that the norm of $\xi^\mu$ is *less* than $\eta^\mu$ for large $r$, but *greater* than $\eta^\mu$ for $r$ near $2M$.

This difference between $\xi^\mu$ and $\eta^\mu$ in Schwarzschild can be understood as follows. Recall a static slice in the maximally extended Schwarzschild solutions two asymptotic regions. It is clear from (3.13) that the coordinates $\rho>0$, $\theta,\varphi$ cover both asymptotic regions. Thus $\beta^A=\varphi^{-3}\beta^{0A}$ is a solution on the entire maximally extended slice. In one asymptotic region ($\rho\to\infty$) $\beta^A$ approaches the constant spinor $\beta^{0A}$, while in the other ($\rho\to 0$) $\beta^A$ vanishes. Since the Killing field $\eta^\mu$ vanishes at the $r=2M$ throat and changes direction on the other side, while the vector $\xi^\mu$ is future directed everywhere, it is clear that they cannot be equal.

The maximally extended Schwarzschild solution is perhaps not of direct physical interest because of the presence of a white hole. It is also inappropriate for discussing the red-shift of matter because all the matter is concentrated in a spacetime singularity. Therefore, as our second example, we consider a static slice in the geometry of a static star. For simplicity, we assume the matter density $\mu$ is constant inside the star so the star is spherically symmetric.[33] Thus the exterior metric is again Schwarzschild, and the interior metric is part of a time symmetric slice in a closed Friedmann universe. The matching condition requires the boundary of the star to have radius $R$ given by

$$M=\tfrac{4}{3}\pi R^3\mu_0 . \tag{3.17}$$

The interior metric is also conformally flat with conformal factor

$$\varphi_{in}(r)=\left[1+\left[1-\frac{2Mr^2}{R^3}\right]^{1/2}\right]^{1/2}\Big/K^{1/2} , \tag{3.18}$$

where the constant $K$ is given by

$$K=\left[1-\frac{3M}{2R}\right]+\left[1-\frac{M}{2R}\right]\left[1-\frac{2M}{R}\right]^{1/2} . \tag{3.19}$$

Since the entire three-geometry of the static star is conformally flat, the asymptotically constant solutions to (3.3) are again $\beta^A=\varphi^{-3}\beta^{0A}$. The norm of the vector $\xi^\mu$ in the exterior region is thus identical to the maximally extended Schwarzschild case, and hence does not agree with the Killing field. In the interior region, we have

$$f=K^2\left[1+\left[1-\frac{2Mr^2}{R^3}\right]^{1/2}\right]^{-2} \tag{3.20}$$

which again does not agree with the norm of the Killing field $\eta^\mu$:

$$(-\lambda)^{1/2}=\frac{3}{2}\left[1-\frac{2M}{R}\right]^{1/2}-\frac{1}{2}\left[1-\frac{2Mr^2}{R^3}\right]^{1/2} . \tag{3.21}$$

Notice that both norms increase with increasing radius as one would expect from the fact that as one encloses more matter, one increases the red-shift. However, they increase at different rates.

Finally, we consider a general conformally flat initial-data set: $q_{ab}=\varphi^4\overset{0}{q}_{ab}$, $\pi_{ab}=0$ where $\overset{0}{q}_{ab}$ is a flat metric on $\Sigma$. Physically, this represents initial data with no gravitational radiation. In addition to the two previous examples, this class includes the well-known initial data for a momentarily static configuration of black holes.[34] By conformal invariance, the solutions to $\mathscr{D}_{AB}\beta^B=0$ are $\beta^{1B}=\varphi^{-3}\beta^{0B}$, $\beta^{2B}=\varphi^{-3}\beta^{0\dagger B}$. Therefore

$$f=\varphi^{-4} . \tag{3.22}$$

Witten's expression (3.9) is easily evaluated for this initial data as follows. The surface integral in (3.9) becomes

$$E=\int_S D^a\varphi^{-4}dS_a . \tag{3.23}$$

Reexpressing the right-hand side in terms of an integral in the conformally related flat space yields

$$E=-4\int_S \varphi^{-3}\partial^a\varphi dS_a \ , \tag{3.24}$$

where the index of the flat derivative $\partial_a$ is raised with $q^0_{ab}$ and the index of $D_a$ is raised with $q_{ab}$. We can now take the flat space divergence of the right-hand side to obtain

$$E=\int(\varphi^2\mu+12\varphi^{-4}\partial_a\varphi\partial^a\varphi)d^3x \ , \tag{3.25}$$

where we have used the fact that

$$\partial^2\varphi=-\tfrac{1}{4}\varphi^5\mu \ , \tag{3.26}$$

which follows from the scalar constraint equation. Finally, we express the right-hand side back in terms of the physical metric $q_{ab}$ and volume element $d\Sigma$:

$$E=\int(\varphi^{-4}\mu+12\varphi^{-6}D_a\varphi D^a\varphi)d\Sigma \ . \tag{3.27}$$

The Newtonian limit of Witten's expression is just a special case of (3.27). Define the Newtonian potential $\phi$ by

$$\varphi=1-\frac{1}{2c^2}\phi \ . \tag{3.28}$$

Then putting the factors of $c^2$ back into (3.26) and (3.27) and taking the limit as $c\to\infty$, we obtain[35]

$$\partial^2\phi=\tfrac{1}{2}\mu \ , \tag{3.29}$$

$$E=\int[\mu(c^2+2\phi)+3\partial_a\phi\partial^a\phi]d^3x \ . \tag{3.30}$$

Comparing (3.30) with the general formula for Newtonian energy (1.6) discussed in the Introduction, we see that Witten's expression corresponds to precisely this expression with $n=4$.

Can one extend the Newtonian formula (1.6) for arbitrary $n$ to a general conformally flat initial-data set? If there is only one asymptotic region, then the answer is yes. (For more than one asymptotic region the extra boundary integrals diverge unless $n\geq 1$.) Since $\varphi\to 1$ asymptotically we can insert a factor of $\varphi^{4-n}$ into the surface integral (3.23). Taking the divergence as before, we now obtain

$$E=\int[\varphi^{-n}\mu+4(n-1)\varphi^{-(n+2)}D_a\varphi D^a\varphi]d\Sigma \ . \tag{3.31}$$

This is our desired result. When $n=4$, this expression reduces to (3.27), and in the Newtonian limit, this expression (for arbitrary $n$) reduces to (1.6). Notice that the integrand is positive definite for all $n\geq 1$. In the first proof of the positivity of energy for this class of initial-data sets, Arnowitt, Deser, and Misner[36] use (3.31) with $n=6$.

## IV. NEW EXPRESSIONS FOR GRAVITATIONAL ENERGY

In this section, we derive a one-parameter family of integral expressions for gravitational energy. For the case of conformally flat initial data, these expressions reduce to Eq. (3.31) obtained in the previous section. The key idea is to use the norm of the solutions to (3.3) as a conformal factor and follow the construction used for conformally flat initial data.

Let $(\Sigma,q_{ab},\pi_{ab})$ be an initial-data set satisfying the dominant energy condition with one asymptotically flat region. Define $\beta^{jA}$, $j=1,2$ as before to be solutions to $\mathscr{D}_{AB}\beta^{jA}=0$ which asymptotically approach $\beta^{0A}$ and $\beta^{0\dagger A}$, where $\beta^{0\dagger A}\beta^0_A=\frac{1}{2}$. Define a real nonnegative function $\omega$ by

$$\omega=f^{1/4}=(\beta^{j\dagger A}\beta^j_A)^{1/4} \ . \tag{4.1}$$

There is some evidence for the conjecture that $\omega$ is always strictly positive.[37] We shall assume that this is the case in what follows. The function $\omega^{-1}$ is then a natural generalization of $\varphi$ for nonconformally flat initial data.

We begin with (3.4) which is satisfied for each $\beta^{jA}$. Multiplying this equation by $\omega^{n-4}$ and integrating over $\Sigma$ yields

$$\begin{aligned} E&=2\int_S \omega^{n-4}\beta^{j\dagger C}\mathscr{D}_a\beta^j_C dS^a \\ &=\int_\Sigma[\omega^n\mu+\omega^{n-4}v^aJ_a \\ &\qquad +2\omega^{n-4}(\mathscr{D}^{AB}\beta^{jC})^\dagger(\mathscr{D}_{AB}\beta^j_C) \\ &\qquad +2(n-4)\omega^{n-5}(D^a\omega)\beta^{j\dagger C}\mathscr{D}_a\beta^j_C]d\Sigma \ , \end{aligned} \tag{4.2}$$

where the first equality uses the fact that $\omega\to 1$ asymptotically. The last term in this equation can have either sign. However, when $\pi=0$, this is no longer the case. As we showed in the last section, $\pi=0$ implies that $\beta^{2A}=\beta^{1\dagger A}$. Therefore (4.2) simplifies to

$$\begin{aligned} E=\int_\Sigma[&\omega^n\mu+2\omega^{n-4}(\mathscr{D}^{AB}\beta^{jC})^\dagger(\mathscr{D}_{AB}\beta^j_C) \\ &+4(n-4)\omega^{n-2}D_a\omega D^a\omega]d\Sigma \ . \end{aligned} \tag{4.3}$$

The first two terms are positive and the last term is a positive function times $(n-4)$. Thus, this expression is positive definite for all $n\geq 4$ and reduces to Witten's expression when $n=4$. If $\pi\neq 0$, then the only expression which remains positive definite is the one with $n=4$.[38] However, the virtue of considering initial data with $\pi=0$ is that, as we now show, one can, in fact, rewrite (4.3) so that it is manifestly positive for all $n\geq 1$, and thus is in close analogy to the Newtonian expression (1.6) and the conformally flat expression (3.31).

Consider the conformally related initial-data set $\tilde{q}_{ab}=\omega^4 q_{ab}$, $\tilde{\pi}_{ab}=\omega^2\pi_{ab}$. Define $\tilde{\beta}^{jA}=\omega^{-3}\beta^{jA}$. Then $\tilde{\beta}^{jA}$ satisfies (3.3) for this initial data

$$\tilde{\mathscr{D}}_{AB}\tilde{\beta}^{jA}=0 \ . \tag{4.4}$$

Notice that for conformally flat initial data $q_{ab}=\varphi^4 q^0_{ab}$ and $\pi_{ab}=0$, we have $\omega=\varphi^{-1}$ so that $\tilde{q}_{ab}$ is flat and $\tilde{\beta}^{jA}$ are constant spinors. Using $\tilde{\epsilon}_{AB}=\omega^2\epsilon_{AB}$ to raise and lower indices of tilded objects, we find

$$(\tilde{\mathscr{D}}^{AB}\tilde{\beta}^{jC})^\dagger(\tilde{\mathscr{D}}_{AB}\tilde{\beta}^j_C) = \omega^{-8}(\mathscr{D}^{AB}\beta^{jC})^\dagger(\mathscr{D}_{AB}\beta^j_C)-6\omega^{-6}D_a\omega D^a\omega \ . \tag{4.5}$$

Substituting into (4.3), we obtain our final expression

$$E=\int_\Sigma [\omega^n\mu+2\omega^{n+4}(\tilde{\mathscr{D}}^{AB}\tilde{\beta}^{jC})^\dagger(\tilde{\mathscr{D}}_{AB}\tilde{\beta}^j_C) +4(n-1)\omega^{n-2}D_a\omega D^a\omega]d\Sigma \ . \tag{4.6}$$

There is a striking similarity between this expression and the one for conformally flat initial data (3.31). The first and last terms are identical under the substitution $\omega=\varphi^{-1}$. The only changes when the initial data is not conformally flat is the middle term—a new positive contribution to the energy.

*This new term is in fact conformally invariant:* If $\hat{q}_{ab}=\Omega^4 q_{ab}$, $\hat{\pi}_{ab}=\Omega^2\pi_{ab}$ where $\Omega$ is any positive function that approaches one asymptotically, then $\hat{\omega}=\Omega^{-1}\omega$ so that $\tilde{q}_{ab},\tilde{\pi}_{ab}$ are unchanged. This implies that the middle term in (4.6) just scales by an overall power of $\Omega$.

We now define

$$\mu_M=\mu \ , \tag{4.7}$$

$$\mu_{\rm I}=2\omega^4(\tilde{\mathscr{D}}^{AB}\tilde{\beta}^{jC})^\dagger(\tilde{\mathscr{D}}_{AB}\tilde{\beta}^j_C) \ , \tag{4.8}$$

$$\mu_{\rm NI}=4\omega^{-2}D_a\omega D^a\omega \ , \tag{4.9}$$

so that (4.6) becomes[39]

$$E=\int \omega^n[\mu_M+\mu_{\rm I}+(n-1)\mu_{\rm NI}]d\Sigma \ . \tag{4.10}$$

The subscripts I and NI refer to "invariant" and "noninvariant" under conformal transformations. Roughly speaking, the direction of the spinors $\beta^{jA}$ determine the conformally invariant part of the gravitational energy $\mu_{\rm I}$, while their norm determines the noninvariant part $\mu_{\rm NI}$. One can easily show that $\mu_{\rm I}\geq 0$ with $\mu_{\rm I}=0$ if and only if $q_{ab}$ is conformally flat and $\pi_{ab}=0$. Similarly, $\mu_{\rm NI}\geq 0$ with $\mu_{\rm NI}=0$ if and only if $q_{ab}$ is flat and $\pi_{ab}=0$. We have thus obtained a covariant decomposition of gravitational energy into a conformally invariant and nonconformally invariant part on a maximal slice.

What is the physical interpretation of this decomposition? We have already seen that for conformally flat initial data, $\mu_{\rm I}=0$ and $\mu_{\rm NI}$ can be viewed as the gravitational binding energy. Since conformally flat initial data represent data with no gravitational waves, it seems natural to interpret $\mu_{\rm I}$ as the energy in gravitational radiation. For vacuum spacetimes, this interpretation is strongly supported by the following fact: $\mu_{\rm I}$ depends only on the conformal metric and transverse trace-free parts of $\pi_{ab}$. Recall that York and co-workers[40] have argued that precisely these two fields represent the unconstrained initial data for gravity. Indeed, it has been shown[41] that given any $q_{ab}$ with non-negative scalar curvature, and any $\pi_{ab}$ which vanishes sufficiently fast asymptotically, then one can obtain initial data $(\hat{q}_{ab},\hat{\pi}_{ab})$ for the vacuum field equation where $\hat{q}_{ab}$ is conformally related to $q_{ab}$ and $\hat{\pi}_{ab}$ is (conformally related to) the transverse trace-free part of $\pi_{ab}$. The fact that $\mu_{\rm I}$ depends only on the freely specifiable initial data strongly supports its interpretation as gravitational radiation energy. Unfortunately, for nonvacuum spacetimes the interpretation of $\mu_{\rm I}$ is less clear for two reasons. First, $\mu_{\rm I}$ now depends on the longitudinal part of $\pi_{ab}$, as well as the transverse trace-free part. Second, there exist spacetimes that have no radiation but have $\mu_{\rm I}\neq 0$ (e.g., static spacetimes with nonconformally flat static slices). Thus one has a simple physical interpretation when one of the three terms $\mu_M$, $\mu_{\rm I}$, or $\mu_{\rm NI}$ vanishes.

The overall factor of $\omega^n$ in (4.10) can be viewed as a red-shift factor. Notice that gravitational energy as well as matter energy is red-shifted in this expression. As one increases $n$, one increases the effective red-shift, but compensates by adding a greater multiple of the binding energy. By analogy with the standard expression for energy in Newtonian theory (1.6), the value of $n$ which is closest to physical intuition is perhaps $n=0$. In this case, there is no explicit red-shift. The negative gravitational potential energy is simply reflected in the negative binding energy.

To summarize, given an initial-data set that satisfies the dominant energy condition, there exists a one parameter family of integral expressions for the total energy (4.2). [For certain values of the parameter, one must assume that there is only one asymptotic region, and that solutions to (3.3) do not vanish.] If $\pi^m_m=0$, then these expressions take a particularly simple form (4.10) in which the gravitational contribution is divided into a conformally invariant $\mu_{\rm I}$ and nonconformally invariant $\mu_{\rm NI}$ part. If, in addition, the matter energy density $\mu_M$ vanishes, then $\int\mu_{\rm I}d\Sigma$ can be interpreted as the energy in gravitational radiation and $\int\mu_{\rm NI}d\Sigma$ can be interpreted as the binding energy. Thus one obtains a covariant decomposition of the total energy into radiation and binding-energy contributions.

We now discuss the vacuum case in more detail and obtain further support for the interpretation of

$\mu_{\rm I}$ and $\mu_{\rm NI}$. For strong gravitational waves, the binding energy $\mu_{\rm NI}$ is nonzero—as one might expect—since these spacetimes have nonzero total energy. However, the positive-energy theorem shows that the binding energy can never exceed the radiation energy. For weak gravitational waves, we now show that $\mu_{\rm NI}$ vanishes and $\mu_{\rm I}$ reduces to the standard energy density for a spin-two field. (For a further discussion of this result and its relation to a new canonical approach to general relativity, see Ref. 42.)

Let $q^0_{ab}$ be flat and $\pi^0_{ab}$ be zero. A weak gravitational wave is described by a perturbed initial-data set $q_{ab}=q^0_{ab}+\epsilon\gamma_{ab}$, $\pi_{ab}=\frac{1}{2}\epsilon\dot{\gamma}_{ab}$ where $\gamma_{ab}$ and $\dot{\gamma}_{ab}$ are two independent transverse trace-free tensors. The only asymptotically constant solution to $\mathscr{D}_{AB}\beta^B=0$ for the initial data $(q^0_{ab},\pi^0_{ab})$ are the constant spinors $\beta^{0B}$. Let[43] $\beta^B=\beta^{0B}+\epsilon\gamma^B$ where $\gamma^B\to 0$ asymptotically. Then to first order, $\mathscr{D}_{AB}\beta^B=0$ implies

$$\mathscr{D}^0_{AB}\gamma^B=-\mathscr{D}^1_{AB}\beta^{0B}\ , \tag{4.11}$$

where $\mathscr{D}^0_{AB}$ is the derivative operator associated with $(q^0_{ab},\pi^0_{ab})$ and $\mathscr{D}^1_{AB}$ is the first-order change in this operator. But the right-hand side of (4.11) vanishes (essentially because one cannot construct a nonzero vector from $\gamma_{ab}$ and $\dot{\gamma}_{ab}$). Since $\mathscr{D}^0_{AB}$ has vanishing kernel, this implies that $\gamma^B\equiv 0$. Therefore, for a weak gravitational wave, $\beta^A$ is unchanged to first order. Consequently, $\omega=(\beta^{j\dagger A}\beta^j_A)^{1/4}$ has no first-order correction. Since $\omega$ remains constant, $\mu_{\rm NI}=0$ for a weak gravitational wave. To lowest order in $\epsilon$, $\mu_{\rm I}$ is given by

$$\mu_{\rm I}=2[(\mathscr{D}^{1AB}\beta^{0C})^\dagger(\mathscr{D}^1_{AB}\beta^0_C)+(\mathscr{D}^{1AB}\beta^{0\dagger C})^\dagger(\mathscr{D}^1_{AB}\beta^{0\dagger}_C)]\ . \tag{4.12}$$

Using the fact that

$$\mathscr{D}^1_{AB}\beta^0_C=\tfrac{1}{2}(D^{MD}\gamma_{ABCD})\beta^0_M+\frac{i}{2\sqrt{2}}\dot{\gamma}_{ABC}{}^D\beta^0_D \tag{4.13}$$

we obtain

$$\mu_{\rm I}=\tfrac{1}{8}[2(D^{MD}\gamma^{ABC}{}_M)(D^N_D\gamma_{ABCN})+\dot{\gamma}^{ABCD}\dot{\gamma}_{ABCD}]\ . \tag{4.14}$$

Since

$$\sqrt{2}D^M_D\gamma_{ABCM}=\epsilon^{mn}_c D_m\gamma_{na} \tag{4.15}$$

we recognize (4.14) as the standard expression for the energy density of a spin-two field.

We have concentrated so far on expression (4.10) with $n=0$. However, other values of $n$ may well be of interest. For example, consider (4.10) with $n=1$. Then $\mu_{\rm NI}$ does not contribute, so for a static spherically symmetric star, the integrand vanishes identically outside the star. This expression may thus prove useful in establishing the inequality between $E$ and the area of the outermost trapped surface conjectured by Penrose.[44]

## ACKNOWLEDGMENTS

It is a pleasure to thank A. Ashtekar, D. Boulware, and B. Ovrut for interesting discussions. G. T. H. is an Albert Einstein Fellow and the work of A. S. was supported in part by the Department of Energy under Grant No. DE-AC02-76ER02220.

[1]E. Witten, Commun. Math. Phys. 80, 381 (1981). Some points of rigor are clarified in O. Reula, J. Math. Phys. 23, 810 (1982) and T. Parker and C. H. Taubes, Commun. Math. Phys. 84, 223 (1982).

[2]R. Arnowitt, S. Deser, and C. W. Misner, Phys. Rev. 117, 1595 (1960); 118, 1100 (1960); 122, 997 (1961).

[3]R. Schoen and S. T. Yau, Commun. Math. Phys. 65, 45 (1979); 79, 231 (1981).

[4]G. T. Horowitz and M. J. Perry, Phys. Rev. Lett. 48, 371 (1982); M. Ludvigsen and J. A. G. Vickers, J. Phys. A 15, L67 (1982). For another proof of this result, see R. Schoen and S. T. Yau, Phys. Rev. Lett. 48, 369 (1982).

[5]G. W. Gibbons and C. M. Hull, Phys. Lett. 109B, 190 (1982). M. Ludvigsen and J. A. G. Vickers, University of Canterbury report, 1982 (unpublished). This result holds only when the local charge density is less than the local energy density at each point.

[6]L. Abbott and S. Deser, Nucl. Phys. B195, 76 (1982).

[7]G. W. Gibbons, S. W. Hawking, G. T. Horowitz, and M. J. Perry, Commun. Math. Phys. (to be published).

[8]G. W. Gibbons, C. M. Hull, and N. Warner, Department of Applied Mathematics and Theoretical Physics report (unpublished).

[9]G. T. Horowitz and A. Strominger (unpublished).

[10]T. Parker and C. H. Taubes, cited in Ref. 1.

[11]O. Reula, cited in Ref. 1.

[12]For other choices of $\alpha^0$ one obtains expressions for other components of the total four-momentum. However, in this paper, we will focus on just the total energy.

[13]We use metric signature $(-+++)$ and units $c=8\pi G=1$. The curvature tensors are defined by $2\nabla_{[a}\nabla_{b]}v_c=R_{abc}{}^d v_d$, $R_{ab}=R_{amb}{}^m$.

[14]C. Teitelboim, Phys. Lett. 69B, 240 (1977).

[15]S. Deser and C. Teitelboim, Phys. Rev. Lett. 39, 249 (1977).

[16]M. Grisaru, Phys. Lett. 73B, 207 (1978).

[17]Although it may seem unnatural to use anticommuting fields to obtain a result involving ordinary spinors in general relativity, from the point of view of this section it appears to be essential.

[18]P.A.M. Dirac, Can. J. Math. 2, 129 (1950).

[19]After completion of this work, we were informed of related analyses by S. Deser, following paper, Phys. Rev. D 27, 2805 (1983) and C. M. Hull, Cambridge report (unpublished).

[20]M. Pilati, Nucl. Phys. B132, 138 (1978); G. Senjanovic, Phys. Rev. D 16, 307 (1977); S. Deser, B. H. Kay, and K. S. Stelle, *ibid.* 16, 2448 (1977).

[21]See the review article, P. van Nieuwenhuizen, Phys. Rep. 68, 191 (1981); our notation follows S. Deser and B. Zumino, Phys. Lett. 62B, 335 (1976).

[22]See, e.g., A. Ashtekar and G. T. Horowitz, Phys. Rev. D 26, 3342 (1982).

[23]The assumption of finite superchange requires $\epsilon$ to approach a constant asymptotically. See Teitelboim, Ref. 14.

[24]R. Casalbuoni, Nuovo Cimento 33A, 115 (1976); 33A, 389 (1976); A. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Systems* (Academia Nazionale dei Lincei, Rome, 1976).

[25]A. Das and D. Z. Freedman, Nucl. Phys. B114, 271 (1976).

[26]Other Dirac brackets may easily be obtained from the identity $\pi^k = -(i/2)\psi^{\dagger k}$ which is a result of the constraints and gauge conditions.

[27]It may seem surprising that a surface integral can generate transformations in the interior of the spacetime. This is due to the nonlocal nature of the Dirac bracket. For further discussion of this point see Teitelboim, Ref. 14.

[28]F. A. Berezin, *Method of Second Quantization* (Academic, New York, 1966).

[29]This approach was used in Ref. 11. However, since we consider positive-definite three-metrics, our formula differs from those of Ref. 11 by various factors of $i$. In particular, real vectors are represented by symmetric spinors $v^{AB}$ satisfying $v^{\dagger AB} = v^{AB}$, so for any spinor $\lambda^A$, $i\lambda^{\dagger(A}\lambda^{B)}$ is a real vector.

[30]Of course, in terms of the four-dimensional stress energy tensor $T_{\mu\nu}$, $(\mu, J^a)$ are the time and space components of the vector $T_{\mu\nu}t^\nu$.

[31]Since $\beta^{0A}$ and $\beta^{0\dagger A}$ form a basis, any other spinor $\gamma^{0A}$ must be a linear combination. Hence the solutions $\gamma^{jA}$ are related to $\beta^{jA}$ by a constant linear transformation which satisfies $\gamma^{j\dagger A}\gamma^{jB} = \beta^{j\dagger A}\beta^{jB}$.

[32]It follows immediately that if one defines $g$ and $w^a$ analogous to $f$ and $v^a$ by taking the *difference* between $j=1$ and $j=2$ in (3.7) and (3.8), then $g=0$ if $\pi=0$. In four-dimensional language, this says that if $\xi^\mu$ is constructed from solutions to (1.1) and is asymptotically tangent to $\Sigma$, then it will remain tangent everywhere provided $\pi=0$.

[33]L. Lindblom, J. Math. Phys. 21, 1455 (1980).

[34]C. W. Misner and J. A. Wheeler, Ann. Phys. (N.Y.) 2, 525 (1957).

[35]We have used the fact that the total mass $\int \mu d\Sigma$ is held fixed when taking the Newtonian limit.

[36]R. Arnowitt, S. Deser, and C. W. Misner, Ann. Phys. (N.Y.) 11, 116 (1960).

[37]For conformally flat initial data, $\omega = \phi^{-1}$ and hence is positive everywhere. Furthermore, the analogous result for scalar fields is a direct consequence of the Hopf maximum principle. See, e.g., C. Morrey, *Multiple Integrals in the Calculus of Variations* (Springer, New York, 1966), p. 61.

[38]If one changes Eq. (3.3) then there are other expressions which are positive even when $\pi \neq 0$. See Horowitz and Perry, Ref. 4.

[39]Given the form of (4.10), the question naturally arises of whether one can sensibly interpret the integrand (for some value of $n$) as an energy *density* for a gravitating system. We investigated this question in some detail for the case $n=4$ (Witten's expression). Our conclusion is essentially negative.

The difficulty lies in the extreme nonlocality of such an energy density. One knows that there is no entirely local covariant energy density. But on physical grounds, one does not necessarily require an energy density to be entirely local. For example, in an asymptotically flat spacetime one might want to incorporate a red-shift factor defined relative to spatial infinity. Such an energy density might be useful, for example, in thermodynamics.

It does not seem possible, however, to give such a physical interpretation to the type of nonlocality embodied in the integrand of Eq. (1.2). Consider this expression on a three-slice describing an arbitrary solution of Einstein's equation. Now perturb slightly the stress energy in a small compact region of the slice. One may compute the effect of this perturbation on expression (1.2) in a power series. It is easily seen that the effect of this perturbation is spread out over the entire three-slice, and is not localized near the perturbation. Similarly, one may leave the spacetime intact and perturb slightly in a compact region the three-slice on which the elliptic equation (1.1) is to be solved. The effects of such a perturbation are again spread over the entire three-slice. Although we specifically looked at the case $n=4$, it seems likely that similar results will hold for any value of $n$.

[40]N. O'Murchadha and J. W. York, Phys. Rev. D 10, 428 (1974); J. W. York, in *Sources of Gravitational Radiation*, edited by L. Smarr (Cambridge University Press, London, 1979).

[41]D. Christodoulou and N. O'Murchadha, Commun. Math. Phys. 80, 271 (1981).

[42]A. Ashtekar and G. T. Horowitz (unpublished).

[43]The perturbation of spinors is most easily understood by explicitly introducing the isomorphism $\sigma_a^{BC}$ from spin space to the tangent space of $\Sigma$. Then $\sigma$ depends on $q_{ab}$ but individual spinors $\beta^A$ do not. One finds that the perturbation of (3.3) does not involve the change in $\sigma$ because $\beta^{0A}$ is constant.

[44]R. Penrose, in *Seminar on Differential Geometry*, edited by S.-T. Yau (Princeton University Press, Princeton, N.J., 1982).