63

Maximal Hypersurfaces and Positivity of Mass (*).

Y. CHOQUET-BRUHAT

Department de Mécanique, Université de Paris VI - Paris

A. E. FISCHER

Department of Mathematics, University of California - Santa Cruz, Cal.

J. E. MARSDEN

Department of Mathematics, University of California - Berkeley, Cal.

Introduction.

This paper discusses some recent results on the related problems of positivity of mass and the existence of maximal spacelike hypersurfaces Σ for asymptotically flat space-times. If k_Σ denotes the second fundamental form of Σ, Σ is called *maximal* if $\operatorname{tr} k_\Sigma = 0$, has *constant mean extrinsic curvature* if $\operatorname{tr} k_\Sigma = \text{const}$ on Σ, and is a *moment of time symmetry* if $k_\Sigma = 0$.

We will give a certain amount of background material in order to keep the exposition self-contained. However, we do assume familiarity with the ADM formalism. The notation and basic results we need are summarized in [1].

We define an asymptotically flat space-time to be a Lorentz metric on $\boldsymbol{R}^4$ which, in the Euclidean co-ordinates on $\boldsymbol{R}^4$, satisfies the asymptotic conditions

$$g_{\mu\nu} = \eta_{\mu\nu} + O\left(\frac{1}{r}\right), \qquad g_{\mu\nu,\alpha} = O\left(\frac{1}{r^2}\right),$$

as $r \to \infty$ on $t = \text{const}$ hypersurfaces. (See sect. 1 for the more technical definition in terms of function spaces.) Here η denotes the standard Minkowski metric $(\operatorname{diag}(-1, 1, 1, 1))$ on $\boldsymbol{R}^4$. Asymptotically flat space-times are often referred to as *isolated systems*.

(*) Partially supported by the National Science Foundation (U.S.A.) and the National Research Council (U. K.).

By an *asymptotically flat hypersurface* $\Sigma \subset \boldsymbol{R}^4$ of an asymptotically flat space-time $(\boldsymbol{R}^4, {}^{(4)}g)$ we mean a spacelike hypersurface $\Sigma = i(\boldsymbol{R}^3)$, where

$$i : \boldsymbol{R}^3 \to \boldsymbol{R}^4$$

is a spacelike embedding and such that the induced metric $g_\Sigma = i^*({}^{(4)}g)$ (with Σ identified to $\boldsymbol{R}^3$) and second fundamental form satisfy (in the Euclidean co-ordinates on $\boldsymbol{R}^3$) the asymptotic conditions

$$g_{ij} = \gamma_{ij} + O\left(\frac{1}{r}\right), \qquad g_{ij,k} = O\left(\frac{1}{r^2}\right), \qquad k_{ij} = O\left(\frac{1}{r^2}\right), \qquad \text{as } r \to \infty,$$

where $\gamma = \operatorname{diag}(1, 1, 1)$ denotes the Euclidean metric on $\boldsymbol{R}^3$.

For such a hypersurface, the mass (energy) is given by the surface integral

$$m(g_\Sigma) = \frac{1}{16\pi} \oint_\infty (g_{ij,j} - g_{jj,i})\, \mathrm{d}S_i$$

evaluated at infinity in the Euclidean co-ordinates in $\boldsymbol{R}^3$. This formula, first derived by EINSTEIN [2] and KLEIN [3], has been rederived and discussed by many authors, such as VON FREUD [4], PAPAPETROU [5], GUPTA [6], GOLDBERG [7], ARNOWITT, DESER and MISNER [8] and TRAUTMAN [9]. Below in sect. **5** we shall review the textbook derivation.

Roughly speaking, m is an average of the $1/r$ part of ${}^{(4)}g$ at spatial infinity, and represents the total mass-energy of both the gravitational and all nongravitational fields present. For the Schwarzschild and Kerr metrics one recovers the usual mass parameter.

The most satisfactory interpretation of the mass function, but also the most subtle, views m as the generator of time translations in the Hamiltonian formulation. This view, initiated by DIRAC [10], and developed by DEWITT [11], has recently been revived by the important contribution of REGGE and TEITELBOIM [12], who show that this surface integral representing the mass *must* be taken as part of the Hamiltonian if the Hamiltonian is to generate the dynamical equations (see proposition 7.1). In sect. **10**, we outline a possible symplectic version of their results by showing that time translations induce a group of symplectic transformations on the space of gravitational degrees of freedom. On this space the mass alone (not supplemented by the ADM Hamiltonian) is the correct Hamiltonian function. This viewpoint is based on a suggestion of WALKER, who together with ASHTEKAR takes a symplectic viewpoint regarding the Bondi-Metzner-Sachs group and the Bondi-Sachs mass, which, roughly speaking, is the mass evaluated on lightlike hypersurfaces at infinity and is the part of the mass associated with gravitational radiation (see [13]).

Many physicists have argued that for nonflat space-times $m > 0$ on the

grounds that gravity is an attractive force. (See, for example, [14], p. 413.) In understanding this, we must keep in mind that m includes not only masses of source, but of the purely gravitational energy as well. The fact that m is supposed to vanish only for flat space is one version of Mach's principle (see, for example, [15]).

The key difficulty in establishing positivity of m occurs already for an empty, but nonflat space-time. That there can exist pure gravitational waves was emphasized by TAUB [16], who pointed out some key differences with the electromagnetic case.

The first proofs of positivity of mass in some important cases were due to ARAKI [17] and BRILL [18]. ARAKI proved the positivity of the second variation of the Schwarzschild mass of a certain class of time-symmetric solutions constructed by conformal methods. BRILL proved positivity for time-symmetric and axial symmetric empty space-times. ARNOWITT, DESER and MISNER [19] showed positivity in case one can find a hypersurface which is maximal, isotropic and asymptotically flat.

An attempt at the general case, which arose out of work of BERGMAN [20] and MOLLER [21], was made by KOMAR [15]. Although the method was unsuccessful [22, 23], many of the ideas helped the later development.

After a lull of several years, the important paper of Brill and Deser [24] appeared and seemed to prove positivity of m once and for all, in general.

The method of Brill and Deser is to show that the mass function has only one critical point, namely at flat space and that the second variation is strictly positive there. The proof is, however, incomplete for four reasons. It is worth detailing these reasons, as some are subtle. Of course, BRILL and DESER were well aware that there were serious mathematical problems.

First of all, they assumed the existence of maximal slices which was then open to question. It is worth recalling why they believed this difficulty could be avoided:

« The existence of at least one minimal $\operatorname{tr} \pi = 0$ hypersurface can be thought of as part of our definition of nonpathological space-time. The value of $\operatorname{tr} \pi$ corresponds to a choice of time co-ordinate. If one examines the transformation needed to reach $\operatorname{tr} \pi = 0$ from arbitrary π one is led to a Poisson-like equation for the co-ordinate function ».

The heart of the detailed proof (see sect. 2) uses precisely this idea.

Secondly, the topology in which the second variation of m is positive definite is not the same as the topology needed on the initial data to properly capture the $1/r$ behaviour at infinity. They were also aware of this problem:

« We note here that in drawing the conclusions from our variational results, we are assuming that the usual extremum theorems for functions of a finite number of variables hold also for our functional, as sufficiently relevant mathematical theorems are not yet available. »

Actually the situation is technically more complex than this quote indicates. When the function space topology and the topology associated with the second variation do not match up, counter-examples are possible. The matter is serious, for positivity of the mass depends on delicate asymptotic behaviour. The work of Tromba [25] is indicative of the mathematical complexity.

Thirdly, the gauge problem was dealt with only on an infinitesimal level. To really establish the result one must show that the space of Cauchy data with the gauges divided out forms a smooth infinite-dimensional manifold. In view of the well-known singularities in superspace [26, 27], this problem requires careful attention.

Fourthly, the global assertion that a C^∞ real-valued function with a single nondegenerate local minimum has that point as a global minimum is not true. For example, we can easily construct such a function h on $\mathbf{R}^2$ as follows. Let D be the open unit disc in the plane. Let f be a function such that $f(0, 0) = 0$, $(0, 0)$ is a nondegenerate minimum, there is a finite number of other critical points and f takes negative values. Cut out these other critical points by drawing nonoverlapping arcs from the boundary of the disc in a manner that f still takes negative values on the region D' remaining. Let $\varphi: \mathbf{R}^2 \to D'$ be a diffeomorphism and let $h = f \circ \varphi$. Specific example:

$$h(x, y) = 2(x+1)^3 + 3\exp[2y] - 6(x+1)\exp[y] + 1 .$$

These sorts of difficulties have been noted by Geroch [28]. In addition, because of the possible nonexistence of maximal slices, O'Murchadha and York [29] speculated on the possibility of negative mass. An important result obtained by them is that, if one can find a maximal hypersurface, then there is an other vacuum time-symmetric space-time with smaller or equal mass, thereby reducing the positivity problem to the time-symmetric vacuum case, *i.e.* $k = 0$ on an asymptotically flat hypersurface Σ.

Difficulties with positivity of the Bondi-Sachs mass similar to those of the ADM mass were pointed out by Robinson and Winicour [30]. It should be possible, however, to make use of $m \geqslant 0$ to help establish $m_{\text{Bondi-Sachs}} \geqslant 0$. Indeed, as Walker has suggested, it would appear that the symplectic structure on the space of gravitational degrees of freedom in closely related to the symplectic structure defined on the gravitational-radiation fields at future null infinity by Ashtekar. The fact tha the ADM mass and the Bondi-Sachs mass are associated with time translations leaving these structures invariant lends some weight to the conjecture that the ADM mass is the past limit, before any radiation has been emitted, of the Bondi-Sachs mass. See sect. **10** for further comments.

Returning to the difficulties with the Brill, Deser proof mentioned above, the problems as they left them can be divided into two parts:

i) Local problem: is the mass positive near flat space?

ii) Global problem: is the mass positive for a space-time that can be connected to flat space?

None of the previous papers had answered either question rigorously. Here we shall give the proof, following CHOQUET-BRUHAT and MARSDEN [31], that the answer to i) is « yes ».

As the above discussion has shown, there is a link between the problem of positivity of mass and the existence of maximal hypersurfaces. However, maximal hypersurfaces are of interest in their own right. For example, in [18], the « moment » of time symmetry (which is in particular maximal) indicates a division between implosion and explosion of the gravitational waves. For closed universes, maximal hypersurfaces may signal a division between expansion and collapse. Besides this, maximal (or constant trace) hypersurfaces have been important in the study of the constraint equations, for these equations partly decouple in such circumstances. This decoupling occurs in both the conformal approach to these equations [29, 32-34 and references therein] and in the direct approach [27, 35]. Maximal hypersurfaces have also proved to be important in numerical computation of space-times; see, *e.g.*, [36].

The first existence and uniqueness theorem for maximal hypersurfaces is due to CHOQUET-BRUHAT [37-39] and CANTOR *et al.* [40]. This theorem is local in nature and will be presented in sect. **2**. Global uniqueness is proved in BRILL and FLAHERTY [41]; see GODDARD [42] and TIPLER and MARSDEN [43] for generalizations. There is still a number of difficulties preventing a proof of global existence. In particular, the results on existence in AVEZ [44] contain an error (this error is propagated in [45]). These difficulties are also discussed in sect. **2**.

We continue our discussion of the history of the mass problem. The local problem being solved, what can be said about the global problem? Although several additional papers of interest were published (such as [46-48]), the next significant attempt was that using supergravity. This attempt was initiated by DESER and TEITELBOIM [49] for quantum gravity. The classical limit $\hbar \to 0$ was taken by GRISARU [50], which indicates that $m \geqslant 0$. However, as it stands, the argument has even more technical objections than those of the original BRILL and DESER paper discussed above [24]. Neverthless, as GRISARU has suggested, it may be possible to give a purely classical rigorous proof once the precise mathematical nature of classical supergravity is understood. (According to ZUMINO and SINGER, one has to study the frame bundle of a Dirac spin bundle over space-time with a subtle graded algebra structure added ... cf. articles by STERNBERG, KOSTANT and ZUMINO in [51]).

The culmination of the mass problem occurs with the paper of SCHOEN and YAU [52]. They finally prove that $m \geqslant 0$ for any asymptotically flat space-time with a maximal slice, with $m = 0$ only for flat space. Their proof involves an ingenious use of the plateau problem and some estimates obtained

from the Gauss-Bonnet formula. Their proof was inspired by a question in pure geometry posed by GEROCH [28].

This latest history, namely supergravity and the method of Schoen and Yau, will not be discussed further in this paper, because of timing and lack of space.

We wish to thank J. ARMS, M. CANTOR, J. EHLERS, G. GIBBONS, R. HANSEN, S. HAWKING, V. MONCRIEF, R. SACHS, B. SCHMIDT, R. SCHOEN, A. TAUB, A. TROMBA, S. T. YAU, M. WALKER, A. WEINSTEIN and J. W. YORK for their helpful comments, and the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, for their hospitality.

1. – Weighted Sobolev and Hölder spaces.

For compact manifolds, the ordinary Sobolev $W^{s,p}$ (or Hölder, $C^{k+\alpha}$) spaces serve as adequate function spaces. The $W^{s,p}$ spaces, with $H^s = W^{s,2}$ were described in [1]. However, for the noncompact case it is essential to modify these spaces in order to properly capture the $1/r$ behaviour needed at infinity.

Hölder spaces with asymptotic conditions have been described and used by CHOQUET-BRUHAT and DESER [53]. We define, for an integer k, $0 < \alpha < 1$, and $a > 0$,

$$C_a^{k+\alpha}(\boldsymbol{R}^n) = \left\{ u \in C^k(\boldsymbol{R}^n) \,|\, \|u\|_{k+\alpha,a} = \sup_{x,y \in \boldsymbol{R}^n} \frac{\|D^k u(x) - D^k u(y)\|}{\|x-y\|^\alpha} + \right.$$
$$\left. + \sum_{\lambda=0}^{k} \sup_{x \in \boldsymbol{R}^n} \|x\|^{n-3+\lambda+a} \|D^\lambda u(x)\| < \infty \right\}$$

and

$$\tilde{C}_a^{2+\alpha} = C_a^{1+\alpha} \cap \{u \in C^2 | \Delta u \in C_{a+2}^{0+\alpha}\} .$$

These spaces do capture the required behaviour in which the elliptic estimates are possible. For example, if $n = 3$ and $0 < a < 1$, then $u \in C_a^{k+\alpha}$ implies

$$u = O\left(\frac{1}{r^a}\right), \qquad \mathrm{D}u = O\left(\frac{1}{r^{1+a}}\right),$$

which corresponds to $g - \gamma = O(1/r)$. *Moreover* [53],

$$\Delta : \tilde{C}_a^{2+\alpha} \to C_{a+2}^{0+\alpha}$$

is an isomorphism. It is the latter fact that enables one to use the implicit function theorem.

Thus we can say that a metric $g_{\mu\nu}$ is asymptotically flat when

$$g_{\mu\nu} - \eta_{\mu\nu} \in C_a^{k+a}, \qquad 0 < a < 1 .$$

To motivate the necessity of these sorts of spaces, recall that the solution of Poisson's equation $\nabla^2\varphi = \varrho$ in $\boldsymbol{R}^3$ is

$$\varphi(x) = -\frac{1}{4\pi}\int_{\boldsymbol{R}^3} \frac{\varrho(y)}{\|x-y\|}\,\mathrm{d}y$$

and that, if $|Q| = \int|\varrho(y)|\,\mathrm{d}y < \infty$, then $\varphi = O(1/r)$ as $r \to \infty$. However, a faster fall-off, $\varphi = O(1/r^2)$, is not possible unless $Q = \int\varrho(y)\,\mathrm{d}y = 0$. The above spaces have relaxed $1/r$ to $1/r^a$ and integrability of Q, so Δ will be an isomorphism.

As an alternative to C_a^{k+a} which we shall follow, one can also use weighted Sobolev spaces $M^p_{s,\delta}$, introduced by NIRENBERG and WALKER [54] and CANTOR [55]. They are defined as follows:

Let C_0^∞ denote the C^∞ functions with compact support from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$. For $f \in C_0^\infty$, we define the $M^p_{s,\delta}$ norm of f by

$$\|f\|_{p,s,\delta} = \sum_{0\leqslant\alpha\leqslant s} \|\sigma^{\alpha+\delta} D^\alpha f\|_{L_p} .$$

Here $1 < p < \infty$, $\delta \in \boldsymbol{R}$, s is a nonnegative integer and $\sigma(x) = (1 + \|x\|^2)^{\frac{1}{2}}$. $M^p_{s,\delta}(\boldsymbol{R}^n, \boldsymbol{R}^m)$ is the completion of C_0^∞ in this norm. As with the Sobolev spaces, we shall usually abbreviate the notation to $M^p_{s,\delta}$.

Analogous to the Sobolev properties we have:

1) *Embedding in C^k*: Sobolev spaces may be defined on all of $\boldsymbol{R}^n$, and the Sobolev embedding still holds in a weakened form: the embedding is continuous but not compact. For $\delta \geqslant 0$, the $M^p_{s,\delta}$ norm is stronger than the $W^{s,p}$ norm, so we have the continuous embedding (*),

$$M^p_{s,\delta}(\boldsymbol{R}^n, \boldsymbol{R}^m) \to W^{s,p}(\boldsymbol{R}^n, \boldsymbol{R}^m) \to C^k(\boldsymbol{R}^n, \boldsymbol{R}^m) .$$

2) *Multiplication*: If $p > 1$, $s > n/p$, $0 \leqslant l \leqslant s$ and $\delta > 0$, then pointwise multiplication $M^p_{s,\delta} \times M^p_{s-l,\delta+l} \to M^p_{s-l,\delta+l}$ is continuous and hence C^∞. (There is a number of results of this type.)

3) *Composition*: If $p > 1$, $s > n/p + 1$, $\delta \in \boldsymbol{R}$, and f is a diffeomorphism such that f and f^{-1} both belong to $M^p_{s,\delta}$, then the compositions $M^p_{s,\delta} \to M^p_{s,\delta}$,

(*) The noncompactness of the embedding in the $M^p_{s,\delta}$ case was pointed out in a private communication by S. AGMON (1975).

$g \mapsto f \circ g$, $g \mapsto g \circ f$ are continuous (the former is C^l if f and f^{-1} are $M^p_{s+l,\delta-l}$ and the latter is C^∞, being linear).

For these spaces the following have been proven:

1.1. *Theorem* (NIRENBERG-WALKER-CANTOR). – *If*

$$p > \frac{n}{n-2}, \qquad \frac{1}{p} + \frac{1}{p'} = 1$$

and $-n/p < \delta < -2 + n/p'$, *then*

$$\Delta : M^p_{s,\delta} \to M^p_{s-2,\delta+2}$$

is an isomorphism.

CANTOR has generalized this to operators with nonconstant coefficients:

1.2. *Theorem* (CANTOR [56]). – *Let* $n > k$ *and* $A_\infty = \sum_{|\alpha|=k} \bar{a}_\alpha D^\alpha$ *be an elliptic homogeneous operator with constant coefficients for systems, on* $\mathbf{R}^n$. *Suppose we have an elliptic operator*

$$A = \sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha$$

on $\mathbf{R}^n$ *such that for* $s \geq k$, $a_\alpha \in C^{s-k}$ *and for each* γ, $0 \leq |\gamma| \leq s-k$,

$$\sup |D^\gamma(a_\alpha(x)) \cdot \sigma(x)^{m-|\alpha|+|\gamma|}| < \infty \qquad \textit{for } |\alpha| < k$$

and

$$\limsup_{|x| \to \infty} |D^\alpha(a_\alpha(x) - \bar{a}_\alpha)\sigma^\gamma| < \varepsilon \qquad \textit{for } |\alpha| = k.$$

Then, if

$$p > \frac{n}{n-k}, \qquad 0 < \delta < -k + \frac{n(p-1)}{p},$$

and ε *is sufficiently small,*

$$A : M^p_{s,\delta} \to M^p_{s-k,\delta+k}$$

has closed range and finite-dimensional kernel.

For a second-order operator, such as the Laplacian on $\mathbf{R}^3$, these theorems require $p > 3$ and $0 < \delta < (p-3)/p$. In this case $M^p_{s,\delta}$ includes the functions that satisfy $f = O(1/r)$ and $Df = O(1/r^2)$ at infinity; in fact, it is easy to see that $f \in M^p_{s,\delta}$ is, on an intuitive level, another way to say f is asymptotically like $1/r$, Df is asymptotically like $1/r^2, \ldots, D^s f = O(1/r^{s+1})$. Note that such an f is not in L_2, but Df is in L_2.

A partial Fredholm alternative can often be obtained from the following:

1.3. *Theorem* (CANTOR [56]). – *Let the hypotheses of the above theorem hold. If $A_\infty + t(A - A_\infty)$, $0 \leqslant t \leqslant 1$, is an injection, then A is an isomorphism.* (*The line $A_\infty + t(A - A_\infty)$ may be replaced by any C^0 curve joining A to A_∞.*)

The following three lemmas illustrate the sort of estimates one makes in these spaces and establish some notation.

1.4. *Lemma.* – *If $\sigma f \in L_p(\mathbf{R}^3)$, then $f \in L_r$ provided*

$$p \geqslant r \geqslant 2p/(p+3)\,.$$

Proof. By Hölder's inequality,

$$\int |f|^r \,\mathrm{d}x \leqslant \left(\int |\sigma f|^p \,\mathrm{d}x\right)^{1/q} \left(\int \sigma^{-rq'} \,\mathrm{d}x\right)^{1/q'},$$

where $q = p/r$, $1/q + 1/q' = 1$. The last integral is finite if $rq' = rp/(p-r) > 3$, *i.e.* if $r > 2p/(p+3)$. □

1.5. *Lemma.* – *If $n = 3$, $2 \leqslant p < 6$, $\delta = 0$ and $s \geqslant 2$, then $f \in M^p_{s,\delta}$ implies $f \in L_6$ and $\mathrm{D}f \in L_2$. In fact, if we write $\|f\|_E^2 = \|\mathrm{D}f\|_{L_2}^2$ (the « energy norm »), we have inequalities*

$$\|f\|_{L^6} \leqslant \mathrm{const} \cdot \|f\|_E \leqslant \mathrm{const} \cdot \|f\|_{p,s,\delta}\,.$$

Proof. If f is in some L_q space, we have the inequality $\|f\|_{L^6} \leqslant C\|\mathrm{D}f\|_{L^2}$ (see [57]). Thus we need only show $\mathrm{D}f \in L^2$. However, $\sigma \mathrm{D}f \in L_p$, so, by the previous lemma, $\mathrm{D}f \in L_r$, $p \geqslant r > 2p/(p+3)$. For $p < 6$, $3p/(p+3) < 2$, so $r = 2$ can be chosen. □

1.6. *Lemma.* – *Let n, p, δ, s be as in lemma 1.5. If $h \in M^p_{s,\delta}$ and $g \in M^p_{s-2,\delta+2}$, then $hg \in L_1$.*

Remark. The intuitive reason is that h falls off like $1/r$, g like $1/r^3$, so hg falls off like $1/r^4$ which leaves room for integrability ($1/r^{3+\varepsilon}$ is integrable at infinity).

Proof. Again the argument is by Hölder's inequality. To show $hg \in L_1$, we need to show that $g \in L_{p'}$, where $p' = p/(p-1)$. However,

$$\int |g|^{p'} \,\mathrm{d}x = \int |\sigma^2 g|^{p'} \cdot \sigma^{-2p'} \,\mathrm{d}x \leqslant \left(\int |\sigma^2 g|^p \,\mathrm{d}x\right)^{p'/p} \left(\int \sigma^{-2p'q'} \,\mathrm{d}x\right)^{1/q'},$$

where

$$q = p/p'\,, \qquad q' = \frac{q}{q-1} = \frac{p}{p-p'}\,.$$

Thus

$$2p'q' = \frac{2pp'}{p-p'} = \frac{2}{1-2} > 3$$

if $p < 6$. Hence $\sigma^{-2p'q'}$ is integrable, so $g \in L_{p'}$. $\square$

These lemmas will be used later for establishing important differences between the mass in Newtonian theory and the mass in general relativity. First of all, a Newtonian potential φ, everywhere regular, zero at infinity and satisfying $\nabla^2\varphi = 0$ is identically zero, *e.g.*, by the maximum principle. In general relativity, however, one can have nonflat, asymptotically flat, vacuum solutions.

Secondly, if $\varphi \in M^p_{s,\delta}$ for p, s, δ as above, $\varrho = \nabla^2\varphi \in M^p_{s-2,\delta+2}$ need not be integrable, so the total mass might be infinite. For example, if $\varrho = 1/r^3$ outside a bounded set, but is otherwise C^∞, $\varrho \in M^p_{s-2,\delta+2}$ for any s, so $\varphi = \Delta^{-1}\varrho \in M^p_{s,\delta}$. Thus φ will only fall off at a rate $1/r^\alpha$, $0 < \alpha < 1$. In general relativity, if the metric differs from the Minkowski metric by something in $M^p_{s,\delta}$, then the vacuum will have finite energy. The above lemmas will be used later to prove this. However, sources which are $M^p_{s-2,\delta+2}$ need not have finite total energy, in either the Newtonian or relativistic case. The above lemmas will be important for probing this later.

We can introduce notation for spaces of metrics like $\mathscr{M}^{s,p}$ which is used in the compact case (see [1]) as follows:

Let γ denote the standard Euclidean metric on $\boldsymbol{R}^3$ and let $\mathscr{M}^p_{s,\delta}$ denote the set of Riemannian (positive definite) metrics g such that $g - \gamma \in M^p_{s,\delta}$. Then $\mathscr{M}^p_{s,\delta}$ is an open cone in $S^p_{s,\delta} + \{\gamma\}$, where $S^p_{s,\delta}$ denotes the Banach space of 2-covariant symmetric tensors of class $M^p_{s,\delta}$. Thus, the tangent space to $\mathscr{M}^p_{s,\delta}$ at g is $T_g\mathscr{M}^p_{s,\delta} = S^p_{s,\delta}$. (Here we assume $s > n/p$ and $\delta \geq 0$.)

Let $I\colon \boldsymbol{R}^3 \to \boldsymbol{R}^3$ be the identity map and let $\mathscr{D}^p_{s,\delta}$ denote the diffeomorphisms η of $\boldsymbol{R}^3$ such that $\eta - I$ and $\eta^{-1} - I$ are of class $M^p_{s,\delta}$. If $s > n/p + 1$, $\delta \geq 0$, then $\mathscr{D}^p_{s,\delta}$ is a smooth manifold and a topological group. From the properties listed above, we find that $\mathscr{D}^p_{s+1,\delta-1}$ acts continuously on $\mathscr{M}^p_{s,\delta}$ by pull-back. One can show that the orbit of γ under $\mathscr{D}^p_{s+1,\delta}$ is a submanifold of $\mathscr{M}^p_{s,\delta}$. However, for $1/r$ fall-off at ∞, $\mathscr{D}^p_{s+1,\delta-1}$ is more appropriate, as is seen by examining the definition of pull-back $\eta^* g$ for $\eta \in \mathscr{D}^p_{s+1,\delta-1}$ and $g \in \mathscr{M}^p_{s,\delta}$.

It is convenient to enlarge $\mathscr{D}^p_{s+1,\delta-1}$ somewhat. Let $\overline{\mathscr{D}}^p_{s+1,\delta-1}$ denote those diffeomorphisms η of $\boldsymbol{R}^3$ such that $D\eta$ and $D(\eta^{-1})$ are of class $M^p_{s,\delta}$. Again, $\overline{\mathscr{D}}^p_{s+1,\delta-1}$ is a topological group and a smooth manifold if $s > n/p$ and $\delta \geq 0$.

1.7. *Lemma.* – $\overline{\mathscr{D}}^p_{s+1,\delta-1}$ *acts continuously on* $\mathscr{M}^p_{s,\delta}$ *by pull-back*: $(\eta, g) \mapsto \eta^* g$.

This uses the multiplication property 2) above and the fact that composition

$$M^p_{s,\delta} \times \overline{\mathscr{D}}^p_{s,\delta-1} \to M^p_{s,\delta}, \qquad (f, \eta) \mapsto f \circ \eta$$

is continuous.

Let $\mathcal{C}_\gamma$ be the orbit of γ under $\overline{\mathcal{G}}^p_{s+1,\delta-1}$. This is not a smooth submanifold of $\mathcal{M}^p_{s,\delta}$. If it were, its tangent space at γ would be the set of Lie derivatives $L_X\gamma$. The trouble is that the canonical decomposition $h = \mathring{h} + L_X\gamma$ need not hold if h only falls off like $1/r$.

What is true is that the orbit $\hat{\mathcal{C}}_\gamma$ of $\gamma \in \mathcal{M}^p_{s,\delta+1}$ under $\overline{\mathcal{G}}^p_{s+1,\delta}$ is a smooth manifold with tangent space the set of Lie derivatives $L_X\gamma$, $X \in M^p_{s+1,\delta}$. We shall need this fact in sect. **10** (here $\eta^*\gamma$ differs from γ like $1/r^2$ at ∞). The crucial fact needed to prove this is the following splitting lemma. The rest of the argument can then proceed as in [58].

1.8. *Lemma.* – *Suppose $h \in S^p_{s,\delta+1}$ and $g \in M^p_{s+1,\delta}$. Then h splits uniquely as*

$$h = \mathring{h} + L_X g\,,$$

where $\delta_g \mathring{h} = 0$ and X is a vector field of class $M^p_{s+1,\delta}$.

Proof. Consider the curve of metrics $g_t = tg + (1-t)\gamma$ joining g to γ. Thus we get a curve of operators $A_t(X) = \delta_{g_t}(L_X g_t)$ joining $A_g(X) = \delta_g(L_X)$ to $A_\gamma(X) = \delta_\gamma(L_X\gamma)$. The latter is a homogeneous second-order elliptic operator with constant coefficients. To show A_g is an isomorphism, we thus need only show A_t is an injection by theorem 1.3. We can write g for g_t.

Thus assume $A_g(X) = 0$. Then

$$0 = \int_{R^n} \delta_g(L_X g)\cdot X\mu(g) = \tfrac{1}{2}\int_{R^n} L_X g\cdot L_X g\mu(g)\,,$$

where $\mu(g)$ is the volume element associated with g.

The integration by parts is justified since $L_X g \in M^p_{s,\delta+1}$ and $X \in M^p_{s+1,\delta}$. Thus $L_X g = 0$. We now need to prove from this that $X = 0$. We sketch two proofs.

Suppose first that g equals γ outside a ball. Then $L_X g = 0$ implies that outside the ball, X is a Killing field for γ and, since none of these is in $M^p_{s+1,\delta}$, $X = 0$ outside the ball. Since any isometry which fixes a point and a frame is the identity, $X = 0$ identically. Thus, for such metrics g, A_g is an isomorphism.

For any $g \in \mathcal{M}^p_{s,\delta}$, let g_R be a metric equalling g on a ball of radius R and equalling γ outside a ball of radius $R+1$. Specifically, let $g_R = \varphi g + (1-\varphi)\gamma$, where φ is a C^∞ function, $0 \leqslant \varphi \leqslant 1$, which is 0 outside a ball of radius $R+1$ and is 1 inside a ball of radius R. Then write

$$A_g = A_{g_R} + B_R\,,$$

where $B_R(X) = \delta_{(g-g_R)} L_X g + \delta_{g_R} L_X(g - g_R)$. The norm of $A^{-1}_{g_R}$ is uniformly bounded below, by, say, ε, as $R \to \infty$. This requires the observation that the norm of the inverse depends only on the modulus of ellipticity and the norms of the coefficients in theorem 1.2. (Details are given in [56].) Secondly,

$$\|B_R\|_{M^p_{s+1,\delta}, M^p_{s,\delta+1}} \to 0 \qquad \text{as } R \to \infty\,,$$

since $g - g_R$ occurs differentiated. Thus if we choose R large enough so that

$$B_R \|_{M^p_{s+1,\delta}, M^p_{s,\delta+1}} < \varepsilon^{-1} ,$$

A_g will be invertible.

The splitting in the lemma now follows by choosing $X = A_g^{-1}(\delta_g h)$ and letting $\check{h} = h - L_X g$.

Here is a second proof that, if $g \in \mathscr{M}^p_{s,\delta}$, $X \in M^p_{s,\delta}$ and $L_X g = 0$, then $X = 0$. First of all, integration by parts shows that

$$\int_{R^3} L_X g \cdot L_X g = 2 \int_{R^3} (\nabla_i X_j)^2 ,$$

so

$$\nabla X = 0 .$$

Secondly, from $\nabla X = 0$, on each line $x^i = a^i t$, X satisfies the equation $\mathrm{d}X_j/\mathrm{d}t = a^i \Gamma^l_{ij} X_l$. But from $g \in \mathscr{M}^p_{s,\delta}$, $s \geqslant 2$, $p > 3$, $\delta \geqslant 0$, $\Gamma^l_{ij}(a^k t)$ is continuous in t, and integrable for almost all $\{a^k\}$. For such an $\{a^k\}$, the equation only has the zero solution vanishing for $t = -\infty$ and $t = +\infty$. (This argument only requires $g \in \mathscr{M}^p_{s,\delta}$ and $\lim_{|x| \to \infty} X(x) = 0$.) $\square$

Some additional notation will prove useful. We shall say that a Lorentz metric ${}^{(4)}g$ on $\boldsymbol{R}^4$ is *asymptotically flat* if, on every $t = \text{const}$ hypersurface, ${}^{(4)}g - \eta$ is of class $M^p_{s,\delta}$ and $(\partial/\partial t)^{(4)}g$ is of class $M^p_{s-1,\delta+1}$. The set of all such metrics is denoted $\mathscr{L}^p_{s,\delta}$. (As above, η is the standard Minkowski metric.)

An embedding $i : \boldsymbol{R}^3 \to \boldsymbol{R}^4$ is called *asymptotically flat* if $i - i_t$ is of class $M^p_{s+1,\delta-1}$, where i_t is the standard $t = \text{const}$ embedding. For such, we have the induced metric $g_{\underline{x}} = i^*({}^{(4)}g) \in \mathscr{M}^p_{s,\delta}(\boldsymbol{R}^3)$ and second fundamental form $k_{\underline{x}} \in S^p_{s-1,\delta+1}(\boldsymbol{R}^3)$. Thus $(i, g_{\underline{x}}, k_{\underline{x}})$ will be called an *asymptotically flat initial data set*.

These concepts correspond to the asymptotic relations required for isolated systems, as described in the introduction.

As for co-ordinate transformations, we have already described $\mathscr{D}^p_{s,\delta}(\boldsymbol{R}^3)$. If we write $F \in \mathscr{D}^p_{s,\delta}(\boldsymbol{R}^4)$ or speak of a diffeomorphism asymptotic to the identity, we shall mean that $i = F \circ i_t$ is asymptotically flat for each of the standard $t = \text{const}$ embeddings i_t.

Spaces with asymptotic conditions like $M^p_{s,\delta}$ can also be studied by a compactification method; see [59].

2. – The existence of maximal hypersurfaces.

In this section we prove that for asymptotically flat space-times near Minkowski space there exists a maximal ($\operatorname{tr} k = 0$) spacelike hypersurface dif-

feomorphism to $\mathbf{R}^3$ which is asymptotically flat. This is the « local » existence-of-maximal-hypersurface theorem. We shall also give the outline of a possible extension of this result to the « global » case.

The existence of maximal hypersurfaces has important consequence for the set of constraint equations, as will be discussed in sect. 3. On such a hypersurface the system of constraint equations can be split into a linear system and a single nonlinear equation, following the conformal techniques initiated, by LICHNEROWICZ [32] and developed by CHOQUET-BRUHAT [33, 60], YORK [61] and O'MURCHADHA and YORK [34, 62]. In sect. 4 we shall give some geometric applications and, in later sections, maximal hypersurfaces will be used in the mass problem.

The results of this section are based on [37-40]. (See also [43, 63].)

We begin by defining two basic energy conditions on space-times $(V_4, {}^{(4)}g)$. Let $\mathrm{Ric}\,({}^{(4)}g)$ denote the Ricci curvature of ${}^{(4)}g$, and $\mathrm{Ein}\,({}^{(4)}g)$ the Einstein tensor. Then ${}^{(4)}g$ satisfies the *weak energy condition* (respectively, the *strong energy condition*) if, for every $x \in V_4$ and timelike vector ${}^{(4)}t_x \in T_x V_4$,

$$\mathrm{Ein}\,({}^{(4)}g)({}^{(4)}t_x, {}^{(4)}t_x) \geqslant 0$$

(respectively, $\mathrm{Ric}\,({}^{(4)}g)({}^{(4)}t_x, {}^{(4)}t_x) \geqslant 0$).

In these cases we shall write $\mathrm{Ein}\,({}^{(4)}g) \geqslant 0$ and $\mathrm{Ric}\,({}^{(4)}g) \geqslant 0$, respectively. A discussion of these energy conditions is given in [64]. (The notation is different if there is a cosmological constant.)

If Σ_0 is a spacelike hypersurface of V_4 with (forward pointing) unit normal ${}^{(4)}Z_{\Sigma_0}$, then we let

$${}^{(4)}G_{\perp\perp} = {}^{(4)}G_{\perp\perp}(\Sigma_0) = \mathrm{Ein}\,({}^{(4)}g_0)\cdot({}^{(4)}Z_{\Sigma_0}, {}^{(4)}Z_{\Sigma_0}) = G_{\alpha\beta}Z^\alpha Z^\beta$$

and

$${}^{(4)}R_{\perp\perp} = {}^{(4)}R_{\perp\perp}(\Sigma_0) = \mathrm{Ric}\,({}^{(4)}g)\cdot({}^{(4)}Z_{\Sigma_0}, {}^{(4)}Z_{\Sigma_0}) = R_{\alpha\beta}Z^\alpha Z^\beta\,,$$

the « perpendicular-perpendicular » projections of the Einstein and the Ricci tensor, respectively.

Many of the results below do not depend on ${}^{(4)}g$ being a solution of the field equations ${}^{(4)}G_{\alpha\beta} = 8\pi T_{\alpha\beta}$ but require only that ${}^{(4)}g$ satisfy either the strong or weak or both energy conditions. Thus to keep the results as general as possible, we consider space-times which satisfy the field equations only when necessary.

Secondly, let us recall the derivation of the $\mathrm{tr}\,k$ evolution equation. The following formula has been obtained in the absence of a shift by LICHNEROWICZ [65] and in the general case by FOURES-(CHOQUET) BRUHAT [60]. See also RAYCHAUDHURI [66].

2.1. *Proposition.* – *Let* $(V_4, {}^{(4)}g)$ *be a space-time and* Σ_λ *a* 1-*parameter family of spacelike hypersurfaces with induced metric* g_λ *and second fundamental form* k_λ. *Let* (N_λ, X_λ) *be the lapse and shift for the embeddings. Then*

$$\frac{\partial}{\partial\lambda}(\operatorname{tr} k) = (k\cdot k + {}^{(4)}R_{\perp\perp} + \Delta)N - X\cdot \mathrm{d}(\operatorname{tr} k)\,,$$

where $\Delta N = -g^{ab}N_{|a|b}$ *is the Laplacian of* N *formed from* g *and* ${}^{(4)}R_{\perp\perp} = \operatorname{Ric}({}^{(4)}g)({}^{(4)}Z_{\Sigma_\lambda}, {}^{(4)}Z_{\Sigma_\lambda})$.

Remark. The evolution equation for $\operatorname{tr} k$ is true for any space-time, independent of the field equations.

Proof. The proof is by taking the various projections of the Riemann curvature tensor. A good reference for these projection formulae is [67]. From these we have (using the notation from [1])

$$\frac{\partial}{\partial\lambda}k = N({}^{(4)}R_{\perp a\perp b} - k\times k) - \operatorname{Hess} N - L_X k\,.$$

Thus

$$\frac{\partial}{\partial\lambda}(\operatorname{tr} k) = -\left(\frac{\partial g}{\partial\lambda}\right)^{\#}\cdot k + g^{-1}\cdot\left(\frac{\partial k}{\partial\lambda}\right) = (2Nk + L_X g)\cdot k + N({}^{(4)}R_{\perp\perp} - k\cdot k) + \\ + \Delta N - \operatorname{tr}(L_X k) = (k\cdot k + {}^{(4)}R_{\perp\perp} + \Delta)N - X\cdot \mathrm{d}\operatorname{tr} k\,,$$

where we have used

$$\frac{\partial g}{\partial\lambda} = -2Nk - L_X g\,, \qquad \Delta N = -\operatorname{tr}\operatorname{Hess} N\,,$$

and $\operatorname{tr}(L_X k) = X\cdot \mathrm{d}\operatorname{tr} k + k\cdot L_X g$. □

Now we are ready to begin the proof of the existence of maximal hypersurfaces. Let $\Sigma_0 = i_0(M)$ be a Cauchy hypersurface of V_4 with induced metric g_0 and second fundamental form k_0. As usual the metrics and other functions used in the argument must belong to appropriate $W^{s,p}$ spaces in the compact case and must be asymptotically flat or zero in the appropriate $M^p_{s,\delta}$ sense in the noncompact case.

Let V_4 be diffeomorphic to $\boldsymbol{R}\times M$ (in the compact case) or $\boldsymbol{R}^4$ (in the noncompact case). In the noncompact case all our metrics ${}^{(4)}g$ will be asymptotically flat, *i.e.* of class $\mathscr{L}^p_{s,\delta}$, and be joinable to η by a curve in the same class.

We shall need to consider embeddings of a certain class, as described in sect. 1. Thus, in the compact case, we can let $\operatorname{Emb}^{s,p}(M, V_4, {}^{(4)}g)$ denote the spacelike embeddings of M to V_4 which are of class $W^{s,p}$ and, in the noncompact case, $\operatorname{Emb}^p_{s,\delta}(\boldsymbol{R}^3, V_4, {}^{(4)}g, i_0)$ denote embeddings i of $\boldsymbol{R}^3$ to V_4 which

differ from a given embedding i_0 by a function of class $M^p_{s,\delta}$. $\mathrm{Emb}^p_{s,\delta}(\mathbf{R}^3, V_4, {}^{(4)}g)$ refers to the case where i_0 is the standard $t=0$ embedding. We shall consider primarily the noncompact case, but shall discuss differences with the compact case at the appropriate point.

As usual, given a metric ${}^{(4)}g$ and an embedding i, we let $\Sigma = i(\mathbf{R}^3)$ and let (g_Σ, k_Σ) denote the associated metric and second fundamental form. If ${}^{(4)}g \in \mathscr{L}^p_{s,\delta}$ and $i \in \mathrm{Emb}^p_{s,\delta}$, then k_Σ will be of class $M^p_{s-1,\delta+1}$ $\big($*i.e.* $O(1/r^2)\big)$, but $\mathrm{tr}\, k_\Sigma$ will be of class $M^p_{s-1,\delta+2}$ $\big($*i.e.* $O(1/r^3)\big)$. This extra fall-off of $\mathrm{tr}\, k_\Sigma$ is seen by writing out $\mathrm{tr}\, k_\Sigma$ in Gaussian normal co-ordinates about Σ; see [37] for the formula.

Fix ${}^{(4)}g$ at ${}^{(4)}g_0$ for the moment, and define

$$P: \mathrm{Emb}^p_{s,\delta}(\mathbf{R}^3, \mathbf{R}^1, {}^{(4)}g) \to M^p_{s-2,\delta+2}(\mathbf{R}^3)$$

by

$$P(i) = \mathrm{tr}\,(k_\Sigma)$$

(we drop into $s-2$ so P will be C^1-differentiable; see sect. 1).

The tangent space of the space of embeddings at i_0 is $\mathscr{X}^p_{s,\delta}(i_0)$, the vector fields ${}^{(4)}X$ on $\mathbf{R}^3$ to V_4 which cover the map i_0 $\big($*i.e.* ${}^{(4)}X(x)$ lies in the tangent space to V_4 at $i_0(x)\big)$. Thus the derivative of P maps as follows

$$\mathrm{D}P(i_0): \mathscr{X}^p_{s,\delta} \to M^p_{s-2,\delta+2}\,.$$

2.2. *Proposition.* – *The derivative of P is given by*

$$\mathrm{D}P(i_0)\cdot{}^{(4)}X = (k_0\cdot k_0 + {}^{(4)}R_{\perp\perp} + \Delta)N - X\cdot \mathrm{d}(\mathrm{tr}\, k_0)\,,$$

where N, X are the lapse and shift functions of ${}^{(4)}X$, k_0 is the second fundamental form on Σ_0 and ${}^{(4)}R_{\perp\perp}$ is the perpendicular-perpendicular projection of $R_{\mu\nu}$ on Σ_0.

If $k_0\cdot k_0 + {}^{(4)}R_{\perp\perp} \geqslant 0$ and ${}^{(4)}g_0$ may be joined to η by a curve of metrics satisfying this condition, then $\mathrm{D}P(i_0)$ is surjective and its kernel splits.

Proof. If $i(\lambda)$ is a curve of embeddings through i_0 tangent to ${}^{(4)}X$, then, by the chain rule,

$$\mathrm{D}P(i_0)\cdot{}^{(4)}X = \frac{\mathrm{d}}{\mathrm{d}\lambda}\,\mathrm{tr}\, k_\lambda\Big|_{\lambda=0}\,.$$

The expression for $\mathrm{D}P(i_0)$ now follows from proposition 2.1. □

From the elliptic theory in sect. 1, $A(N) = (k_0\cdot k_0 + {}^{(4)}R_{\perp\perp} + \Delta)N$ is an isomorphism from $M^p_{s,\delta}$ to $M^p_{s-2,\delta+2}$. Indeed, since $k_0\cdot k_0 + {}^{(4)}R_{\perp\perp} \geqslant 0$, A is injective and may be joined to Δ_γ by a curve of operators which are injective, by hypothesis. Therefore, by theorem 1.2, A is an isomorphism. Thus, $\mathrm{D}P(i_0)$ is onto, taking $X=0$ and variable N.

The kernel of $DP(i_0)$ splits as follows. Given X, let $N(X)$ be defined by $N(X) = A^{-1}(X \cdot \mathrm{d}\,\mathrm{tr}\,k_0)$ (*e.g.*, if $\mathrm{tr}\,k_0 = 0$, $N(X) = 0$) and decompose

$$(N, X) = \big(N(X), X\big) + \big(N - N(X), 0\big)\,.$$

Remarks.

1) In the noncompact case, $k_0 = 0$, ${}^{(4)}R_{\perp\perp} = 0$ is allowed, for Δ is an isomorphism. In the compact case this is exceptional and instead we map $\mathrm{Emb}^{s,p}(M, V_4, {}^{(4)}g_0)$ to $\mathscr{W}_0^{s,p}(M)$, the $W^{s,p}$ volume elements on M with zero integral, by

$$\hat{P}: \mathrm{Emb}^{s,p} \to \mathscr{W}_0^{s-2,p}\,,$$

$$\hat{P}(i) = \left\{(\mathrm{tr}\,k_\Sigma) - \frac{1}{\mathrm{vol}\,(M)} \int_M (\mathrm{tr}\,k_\Sigma)\mu(g)\right\}\mu(g)\,.$$

The derivative of P at i_0 is just ${}^{(4)}X \mapsto \Delta N$ if $k_0 = 0$, ${}^{(4)}R_{\perp\perp} = 0$, which is surjective from $W^{s,p}$ to $\mathscr{W}_0^{s-2,p}$, with kernel the constants.

2) Note that, if $\mathrm{tr}\,k_0 = 0$ or constant, then X does not appear in $DP(i_0)$. This corresponds to the obvious fact that, if $f: \mathbf{R}^3 \to \mathbf{R}^3$ (or $M \to M$) is a diffeomorphism asymptotic to the identity, then $i = i_0 \circ f$ is another embedding onto the same hypersurface (*i.e.* i and i_0 have the same range) and with $\mathrm{tr}\,k_0$ still 0 (or constant). Thus we cannot expect maximal hypersurfaces to be uniquely embedded, but we might expect the surface itself to be unique.

If we regard the variables in proposition 2.2 to be functions N, diffeomorphisms f of $\mathbf{R}^3$ and four metrics ${}^{(4)}g$, then proposition 2.2 asserts that the partial derivative of P with respect to N is an isomorphism. Thus the implicit function theorem gives the following main local existence theorem for maximal hypersurfaces.

2.3. *Theorem.* – *Let the hypotheses of proposition 2.2 hold.* i) *For* ${}^{(4)}g$ *sufficiently close to* ${}^{(4)}g_0$ *in* $\mathscr{L}^p_{s,\delta}$ *there exists a spacelike embedding* i *asymptotic to* i_0 *on which* $\mathrm{tr}\,k_\Sigma = \mathrm{tr}\,k_0$ ($\mathrm{tr}\,k_\Sigma$ *is relative to* ${}^{(4)}g$ *and* $\mathrm{tr}\,k_0$ *relative to* ${}^{(4)}g_0$). ii) *If* $\mathrm{tr}\,k_0 = 0$, *then* i *is unique up to a diffeomorphism* f *of* $\mathbf{R}^3$ *asymptotic to the identity.*

For M *compact,* $k_0 \cdot k_0 + {}^{(4)}R_{\perp\perp} \geqslant 0$, *and not identically zero, and for* ${}^{(4)}g$ *sufficiently close to* ${}^{(4)}g_0$, *there exists an embedding* i *close to* i_0 *for which* $\mathrm{tr}\,k_\Sigma = \mathrm{tr}\,k_0$; i *is unique up to diffeomorphism if* $\mathrm{tr}\,k_0 = \mathrm{const}$. *If*

$$k_0 \cdot k_0 + {}^{(4)}R_{\perp\perp} = 0\,, \qquad \mathrm{tr}\,k_0 = \mathrm{const}\,,$$

and ${}^{(4)}g$ *is close to* ${}^{(4)}g_0$, *there is an embedding* i *on which* $\mathrm{tr}\,k_\Sigma = \mathrm{const}$, *a possibly different constant.*

As a corollary, note that, if ${}^{(4)}g$ is near Minkowski space, then ${}^{(4)}g$ admits a maximal hypersurface Σ. This surface is unique when it is specified to which $t = \text{const}$ standard hypersurface it should be asymptotic.

Once one has found one hypersurface on which $\operatorname{tr} k = 0$ (or constant), one can find a whole family of them in the noncompact case. Indeed, in proposition 2.1, set $X = 0$ and let N be unknown. Then, if we let N be the function of (g, k) defined by

$$(k \cdot k + {}^{(4)}R_{\perp\perp} + \Delta) N = 0 , \qquad N - 1 \in M^p_{s,\delta},$$

which is possible by ellipticity (see theorem 1.3), the hypersurfaces defined by the lapse N and shift $X = 0$ will have $(\mathrm{d}/\mathrm{d}\lambda) \operatorname{tr} k = 0$. Since $\operatorname{tr} k$ starts off 0, it will remain so. In the compact case one can produce a family of hypersurfaces of constant mean curvature by solving $(\mathrm{d}/\mathrm{d}\lambda) \operatorname{tr} k = C(\lambda)$ for an appropriately chosen constant $C(\lambda)$ depending on λ.

Thus, we have proven

2.4. *Corollary.* – *If ${}^{(4)}g$ is as in theorem 2.3, then a whole neighbourhood of Σ can be written as the union of spacelike hypersurfaces on which the second fundamental form has zero (or constant in the compact cases) trace.*

Now we turn to the question of finding maximal hypersurfaces for ${}^{(4)}g$ when ${}^{(4)}g$ is far away from Minkowski spaces, but connected to it by a curve of space-times.

A possible procedure to deal with this case is based on the global inverse function theorem which we recall in the following. (See, *e.g.*, [68] for a proof.)

2.5. *Lemma.* – *Let $\mathbf{E}$ and $\mathbf{F}$ be Banach spaces and $f: \mathbf{E} \to \mathbf{F}$ a C^1 mapping. Assume $\mathrm{D}f(x)$ is an isomorphism for each $x \in \mathbf{E}$ and either f is proper or $\|\mathrm{D}f(x)\| \geq \varepsilon$ for some $\varepsilon > 0$. Then $f: \mathbf{E} \to \mathbf{F}$ is a diffeomorphism of $\mathbf{E}$ onto $\mathbf{F}$.*

We now give a plausible conjecture with possible indications of a proof.

2.6. *Conjecture.* – *Let ${}^{(4)}g_0 \in \mathscr{L}^p_{s,\delta}(\mathbf{R}^4)$ satisfy the strong energy condition* $\operatorname{Ric}({}^{(4)}g_0) > 0$, *and suppose there exists a curve of space-times ${}^{(4)}g(\varrho) \in \mathscr{L}^p_{s,\delta}$, $0 \leq \varrho \leq 1$, such that ${}^{(4)}g(0) = {}^{(4)}g_0$, ${}^{(4)}g(1) = \eta =$ Minkowski metric on $\mathbf{R}^4$, and such that, for each ϱ,* $\operatorname{Ric}({}^{(4)}g(\varrho)) \geq 0$.

Then there exists an asymptotically flat maximal hypersurface for ${}^{(4)}g_0$ (and in fact a foliation of space-time by them).

Indications for a proof. Let E be the component of $\operatorname{Ric} \geq 0$ space-times containing ${}^{(4)}g_0$ and for $\varepsilon > 0$ let S_ε be those ${}^{(4)}g \in E$ which have a maximal slice on which

$$g_x(\xi, \xi) \geq \varepsilon \gamma(\xi, \xi) \qquad \text{for all vectors } \xi$$

($\gamma =$ the Euclidean metric on $\mathbf{R}^3$).

By the preceding theorem, $S = \bigcup_{\varepsilon>0} S_\varepsilon$ is open. To show that $S = E$ we have to show that S is closed as well, that is that, if ${}^{(4)}g_n \to {}^{(4)}g$, with ${}^{(4)}g_n \in S$, then ${}^{(4)}g \in S$. Now we note that, if ${}^{(4)}g_n \in S_\varepsilon$ for all n, then the inverse of $D_y P$ is uniformly bounded below, independent of n, so the argument of theorem 2.3 ought to enable us to prove that ${}^{(4)}g$ has a maximal slice. Thus the problem then rests on the possibility of preventing the maximal slices in the metric ${}^{(4)}g_n$ from turning null-like as $n \to \infty$, *i.e.* of arranging ${}^{(4)}g_n$ to lie in a fixed S_ε. It is plausible that this can be arranged by making suitable hypotheses on the space-time singularities. (*Note.* A previous study of Avez [44] for the existence of maximal slices in periodic space-times with compact spacelike sections encountered the same difficulty. The results of CHENG and YAU [69] are crucial in overcoming this difficulty; see [43].)

3. – Applications to linearization stability.

Linearization stability for the vacuum Einstein equations in the case of compact hypersurfaces was discussed in the previous paper. Here we shall discuss the time-symmetric asymptotically flat case. Near flat space, linearization stability by using weighted Hölder spaces and conformal methods was proved by CHOQUET-BRUHAT and DESER [53] with special asymptotic conditions and in weighted Sobolev spaces, by using a direct approach, by FISCHER and MARSDEN [27, 35, 70]. Here we prove linearization stability for the general time-symmetric asymptotically flat case following CHOQUET-BRUHAT, FISCHER and MARSDEN [71].

As in the compact case, we shall consider the map

$$\Phi : \mathscr{M}^p_{s,\delta} \times S^p_{s-1,\delta+1} \to M^p_{s-2,\delta+2} \times A^p_{s-2,\delta+2} ,$$
$$(g, \pi) \mapsto \Phi(g, \pi) = (\mathscr{H}(g, \pi), \mathscr{J}(g, \pi)) .$$

(Here we use notation from [1], but have dropped the superscript 2 on S and 1 on A, denoting 2-tensors and 1-forms, respectively, and the subscript d on M and A, denoting densities.) The associated constraint space is then written

$$\mathscr{C}^p_{s,\delta} = (\mathscr{C}_{\mathscr{H}})^p_{s,\delta} \cap (\mathscr{C}_\delta)^p_{s-1,\delta+1} = \Phi^{-1}(0) =$$
$$= \{(g, \pi) \in \mathscr{M}^p_{s,\delta}(\mathbf{R}^3) \times S^p_{s-1,\delta+1}(\mathbf{R}^3) | \mathscr{H}(g, \pi) = 0, \delta_g \pi = 0\} .$$

We shall always work with space-times that admit hypersurfaces of constant or zero mean curvature.

There are two technical points requiring attention when studying the asymptotically flat case. The first is that the evolution equations take place in H^s spaces, not in $M^p_{s,\delta}$ spaces. This is apparently not serious however, for the main

difference is that $M^p_{s,\delta}$ allows a variable M/r term, while this is a « constant of the motion » in the evolution of a space-time from Cauchy data.

The second difficulty is more serious. It is that the splitting theorems for differential operators do not hold in general and great care is needed in their application. Thus, to show that $D\Phi(g, \pi)$ is surjective and that its kernel splits, we must proceed directly, rather than appealing to injectivity of the adjoint $D\Phi(g, \pi)^*$. Thus in the following theorem we must use special arguments to show that $D\mathscr{H}(g, \pi)$, $D\mathscr{J}(g, \pi)$ and $D\Phi(g, \pi)$ are surjective, and that their kernels split.

Using this direct approach, we prove our results in case $\pi = 0$, that is on the time-symmetric initial data set; for the case $\operatorname{tr} \pi$ a constant, see CHOQUET-BRUHAT, FISCHER and MARSDEN [71]. For Friedman models, see D'EATH [72].

In the following theorem, one may use either weighted Sobolev or Hölder spaces.

3.1. *Theorem.* – *Let $(g, \pi) \in (\mathscr{C}_{\mathscr{H}})^p_{s,\delta}$. Then $\mathscr{H}(g, \pi) = 0$ is linearization stable at (g, π) and $(\mathscr{C}_{\mathscr{H}})^p_{s,\delta}$ is a smooth manifold in a neighbourhood of (g, π).*

If $(g, \pi) \in (\mathscr{C}_\delta)^p_{s-1,\delta+1}$, then $\mathscr{J}(g, \pi) = 0$ is linearization stable at (g, π) and $(\mathscr{C}_\delta)^p_{s-1,\delta+1}$ is a smooth manifold in a neighbourhood of (g, π).

Let

$$\mathscr{C}^p_{s,\delta} = (\mathscr{C}_{\mathscr{H}})^p_{s,\delta} \cap (\mathscr{C}_\delta)^p_{s-1,\delta+1} ,$$

and let

$$\mathring{\mathscr{C}}^p_{s,\delta} = \{(g, \pi) \in \mathscr{M}^p_{s,\delta} \times S^p_{s-1,\delta-1} | R(g) = 0 \text{ and } \pi = 0\} ,$$

the set of initial data for time-symmetric space-times. Then the equation $\Phi(g, \pi) = 0$ is linearization stable at $(g, 0) \in \mathring{\mathscr{C}}^p_{s,\delta}$, and $\mathring{\mathscr{C}}^p_{s,\delta}$ is a smooth manifold in a neighbourhood of $(g, 0) \in \mathscr{C}^p_{s,\delta}$.

Proof. As in the compact case, it is enough to show that $D\mathscr{H}(g, \pi)$, $D\mathscr{J}(g, \pi)$ and $D\Phi(g, 0)$ are surjective and that their kernels split. The linearization stability results and manifold assertions for these maps will then follow from the implicit function theorem for the spaces $M^p_{s,\delta}$.

Let (h, ω) be an infinitesimal deformation of (g, π). We shall consider special (h, ω) to prove that $D\mathscr{H}(g, \pi)$, $D\mathscr{J}(g, \pi)$, $D\Phi(g, 0)$ are surjective. First let $h = fg$, $f \in M^p_{s,\delta}(\mathbf{R}^3)$. Such an f represents a pointwise conformal deformation of g. Then, using the general formulae from [1], we have

$$D\mathscr{H}(g, \pi) \cdot (fg, \omega) = \tfrac{1}{2}\mathscr{H}(g, \pi) f - 2(\Delta f)\mu(g) + 2\big((\pi')^\flat - \tfrac{1}{2}(\operatorname{tr} \pi')g\big) \cdot \omega .$$

Since $\mathscr{H}(g, \pi) = 0$, we get

$$D\mathscr{H}(g, \pi) \cdot (fg, 0) = -2(\Delta_g f)\mu(g) .$$

Since $\Delta_g : M^p_{s,\delta}(\mathbf{R}^3) \to M^p_{s-2,\delta+2}(\mathbf{R}^3)$ is an isomorphism (unlike the compact case), the map

$$f \mapsto \mathrm{D}\mathscr{H}(g, \pi)\cdot(fg, 0) = -2(\Delta_g f)\mu(g)$$

is an isomorphism. Hence $\mathrm{D}\mathscr{H}(g, \pi)$ is surjective.

We now show that we can split $S^p_{s,\delta} \times S^p_{s-1,\delta+1}$ as a topological sum $E_1 \oplus E_2$, where $E_1 = \ker \mathrm{D}\mathscr{H}(g, \pi)$, and E_2 is a closed complement. Indeed for $(h, \omega) \in \in S^p_{s,\delta} \times S^p_{s-1,\delta+1}$, let f be such that $\mathrm{D}\mathscr{H}(g, \pi)\cdot(fg, 0) = \mathrm{D}\mathscr{H}(g, \pi)\cdot(h, \omega)$. As above, there is a unique solution f. Then split (h, ω) as

$$(h, \omega) = (h - fg, \omega) + (fg, 0)\,.$$

The set of such $(fg, 0)$ forms the closed space E_2.

For the equation $\mathscr{J}(g, \pi) = 0$, we proceed similarly. If we take $\mathscr{J}(g, \pi) = 0$ and $h = 0$,

$$\mathrm{D}\mathscr{J}(g, \pi)\cdot(0, \omega) = -2\delta_g\omega\,,$$

which is surjective on deformations $\omega = (L_X g)^\sharp \mu(g)$, where $X \in \mathscr{X}^p_{s,\delta}(\mathbf{R}^3)$, since

$$\delta_g \circ \alpha_g : \mathscr{X}^p_{s,\delta}(\mathbf{R}^3) \to \mathscr{X}^p_{s-2,\delta+2}(\mathbf{R}^3)\,,$$
$$X \mapsto \delta_g(L_X g)\,,$$

is an isomorphism as we saw in sect. 1. (In the compact case $\ker \delta_g \circ \alpha_g =$ Killing vector fields on M.)

Again, from surjectivity of $\mathrm{D}\mathscr{J}(g, \pi)$, we can split $S^p_{s,\delta} \times S^p_{s-1,\delta+2}$ topologically as

$$S^p_{s,\delta} \times S^p_{s-1,\delta+1} = \ker \mathrm{D}\mathscr{J}(g, \pi) \oplus E_2\,,$$

where $E_2 = \{(0, (L_X g)^\sharp \mu(g) \mid X \in \mathscr{X}^p_{s,\delta}(\mathbf{R}^3)$ and X solves $-2\,\delta_g(L_X g) = \mathrm{D}\mathscr{J}(g, \pi)\cdot \cdot(h, \omega)$ for some $(h, \omega)\}$.

Now we consider the equation $\Phi(g, \pi) = 0$. For $h = fg$,

$$\begin{aligned}\mathrm{D}\Phi(g, \pi)(fg, \omega) &= \Big(\tfrac{1}{2}\mathscr{H}(g, \pi)f - 2(\Delta f)\mu(g) + \\ &\quad + 2\big((\pi')^\flat - \tfrac{1}{2}(\operatorname{tr}\pi')g\big)\cdot\omega,\ -2\big(\delta_g\omega + f\,\delta_g\pi + \tfrac{1}{2}(\operatorname{tr}\pi)(\operatorname{grad} f) - \pi\cdot \mathrm{d}f\big)\Big) = \\ &= \Big(-2(\Delta f)\mu(g) + 2\big((\pi')^\flat - \tfrac{1}{2}(\operatorname{tr}\pi')g\big)\cdot\omega,\ -2\,\delta_g\omega - (\operatorname{tr}\pi)\operatorname{grad} f + 2\pi\cdot \mathrm{d}f\Big)\,,\end{aligned}$$

where the second equality uses the fact that (g, π) satisfies the constraints. The problem here is to show that $\mathrm{D}\Phi(g, \pi)\cdot(fg, \omega)$ is surjective. This difficulty is related to the use of the $\operatorname{tr}\pi' = \text{const}$ condition to get surjectivity in the compact case (see, in [1], theorem 3.5). Our present arguments assume $\pi = 0$.

Then

$$D\Phi(g, 0)\cdot(fg, \omega) = \big(-2\,(\Delta f)\mu(g), -2\,\delta_g\omega\big)\,,$$

which we have just seen is surjective onto $M^p_{s-2,\delta+2}\times\Lambda^p_{s-2,\delta+2}$ with ω's of the form $(L_X g)^\sharp\mu(g)$, $X\in\mathscr{X}^p_{s,\delta}$. As in the previous cases, the spaces split:

$$S^p_{s,\delta}\times S^p_{s-1,\delta+1} = \ker D\Phi(g, 0)\oplus E_2\,.$$

This splitting is obtained as before: namely $(h, \omega)\in S^p_{s,\delta}\times S^p_{s-1,\delta+1}$ can be written as

$$(h, \omega) = \big(h - fg, \omega - (L_X g)^\sharp\mu(g)\big) + \big(fg, (L_X g)^\sharp\mu(g)\big)\,,$$

where (f, X) satisfy

$$D\Phi(g, 0)\cdot(h, \omega) = D\Phi(g, 0)\cdot\big(fg, (L_X g)^\sharp\mu(g)\big) = \big(-2(\Delta f)\mu(g), -2\,\delta_g(L_X g)^\sharp\mu(g)\big). \quad \square$$

Remarks. Let

$$\mathring{\mathscr{M}}^p_{s,\delta} = \{g\in\mathscr{M}^p_{s,\delta}\,|\,R(g) = 0\}\,.$$

Then

$$\mathring{\mathscr{C}}^p_{s,\delta} = \mathring{\mathscr{M}}^p_{s,\delta}\times\{0\}\,.$$

If we use the same method of proof as for $(\mathscr{C}_{\mathscr{X}})^p_{s,\delta}$, it follows that $\mathring{\mathscr{M}}^p_{s,\delta}$ is a smooth manifold, since, for $R(g) = 0$, $DR(g)\cdot fg = 2\,\Delta_g f$ is surjective. Hence one can use this argument to show that $\mathring{\mathscr{C}}^p_{s,\delta}$, the time-symmetric initial Cauchy data, is a smooth manifold. Note, however, that this procedure is *not* sufficient to conclude that $\Phi(g, \pi) = 0$ is linearization stable at such a $(g, 0)\in\mathring{\mathscr{C}}^p_{s,\delta}$, since one must set $\pi = 0$ after one takes the variation of the divergence constraint. This is the reason the full argument of theorem 3.1 is needed. Here is an alternative, slightly more flexible argument that the time-symmetric initial data sets are a smooth submanifold of $\mathscr{M}^p_{s,\delta}$.

Proof. We consider the subset $R(g) = 0$ of $\mathscr{M}^p_{s,\delta}$. The mapping $g\mapsto R(g)$ is C^∞. Its derivative at any point g is

$$DR(g)\cdot h = \Delta(\operatorname{tr} h) + \delta\,\delta h - \operatorname{Ric}(g)\cdot h\,.$$

We show that it is surjective and its kernel splits. Indeed

1) Take $h = k + \frac{1}{3}g\tau$, with τ a scalar function in $M^p_{s,\delta}$, and $k\in S^p_{s,\delta}$ given. Write

$$DR(g)\cdot h = \tfrac{2}{3}\Delta_g\tau + \Delta_g k + \delta\,\delta k - \operatorname{Ric}(g)\cdot k\,.$$

The mapping $\tau\mapsto DR(g).\,h$ is then surjective by the invertibility of Δ_g.

2) We show that we can split $h \in S^p_{s,\delta}$ as a topological sum

$$h = h_1 \oplus h_2$$

with $h_1 \in \ker DR(g)$ and $h_2 \in E_2$, where E_2 is the closed subspace of $S^p_{s,\delta}$ consisting of tensors of the form λg, with a scalar function chosen so that

$$h - \frac{\lambda}{3} g \in \ker DR(g) ,$$

which is always possible, in a unique way, by what we have just seen. □

In 3.1, the choice $\omega^\flat = (L_X g)\mu(g) - \pi\tau$ allows one to treat the case $\operatorname{tr}\pi =$ constant. See [71].

Using the method of proof of theorem 5.4 of [1] and theorem 3.1 above, we have the following linearization stability theorem for time-symmetric *asymptotically flat* empty-space solutions to Einstein's equations.

3.2. *Theorem.* – Let $(\mathbf{R}^4, {}^{(4)}g_0)$, ${}^{(4)}g_0 \in \mathscr{L}^p_{s,\delta}$, *be an asymptotically flat Einstein-flat space-time,* $\operatorname{Ein}({}^{(4)}g_0) = 0$.

Let $i_0 \colon \mathbf{R}^3 \to \mathbf{R}^4$ *be an asymptotically flat spacelike embedding of* $\mathbf{R}^3$ *into* $\mathbf{R}^4$ *such that* $\Sigma_0 = i_0(\mathbf{R}^3)$ *is a moment of time symmetry,* $k_{\Sigma_0} = 0$ *(and* $g_{\Sigma_0} \in \mathscr{M}^p_{s,\delta}(\mathbf{R}^3)$*). Then the Einstein empty-space field equations*

$$\operatorname{Ein}({}^{(4)}g) = 0$$

are linearization stable at ${}^{(4)}g_0$ *on* $\mathbf{R}^4$, *regarded as the maximal Cauchy development of* $(g_{\Sigma_0}, 0)$. *In particular, flat space is a point of linearization stability.*

Remarks.

1) For linearization stability in the asymptotically flat case we restrict ourselves to spatially asymptotically zero deformations ${}^{(4)}h \in S^p_{s,\delta}(\mathbf{R}^4)$, and then we find a curve of asymptotically flat exact solutions in $\mathscr{L}^p_{s,\delta}$.

2) The space-time condition corresponding to the conditions here is the absence of a timelike Killing field. Compare sect. 5 of [1] and theorem 4.1 below.

3) In the asymptotically flat cases, an extra argument is needed because the Cauchy data lie in $M^p_{s,\delta}$ spaces, whereas the evolution occurs in H^s spaces. However, since the Cauchy data (g, π) are in $\mathscr{M}^p_{s,\delta} \times S^p_{s-1,\delta+1}$ they are in $(W^{s,p}, W^{s-1,p})$ (ordinary Sobolev spaces) on each compact set of $\Sigma_0 \approx \mathbf{R}^3$. On the other hand, we know that on a compact set of $\mathbf{R}^3$, $H^s \supset W^{s,p}$ if $p > 2$. The evolution of the Cauchy data is governed by hyperbolic equations and the corresponding existence theorem is in fact valid in H^{loc}_s (function H^s on

each compact set). Thus, one gets linearization stability in this sense. (The problem of finding a suitable function space which captures the asymptotic behaviour *and* which is preserved by the evolution equation remains open, however.)

4) For a study of existence of solution to the constraint equations on a noncompact manifold using conformal methods, see [17, 29, 73].

4. – Geometric applications of maximal hypersurfaces.

In this section we shall use the evolution equation for $\operatorname{tr} k$, and also the expression of $\operatorname{tr} k$ itself as an elliptic differential operator, to derive some theorems concerning maximal (or constant $\operatorname{tr} k$) hypersurfaces.

We first give an old result of Lichnerowicz [65] concerning stationary space-times, which is related to the Komar expression for the mass of a stationary space-time [15]. We stress, however, that the results of Lichnerowicz do not assume the existence of maximal hypersurfaces.

A space-time $(V_4, {}^{(4)}g)$ is said to be *stationary* if it admits a one-parameter isometry group, acting effectively on V_4 with timelike trajectories. It is supposed moreover that there is a diffeomorphism which maps V_4 onto $\Sigma \times \boldsymbol{R}$, where Σ is a differentiable 3-manifold and where the timelike trajectories are the pull-back of the lines $\{x\} \times \boldsymbol{R}$.

A stationary space-time is said to be *static* if the trajectories of its 1-parameter isometry group (equivalently of its Killing vector field ξ) are hypersurface orthogonal.

4.1. *Theorem* (LICHNEROWICZ).

i) *A stationary space-time, vacuum or satisfying the strong energy condition, and with Σ compact, is static. If it is vacuum or if ${}^{(4)}R_{\perp i \perp j} = 0$ implies ${}^{(4)}R_{ij} = 0$ (mixed energy condition), then it is flat.*

ii) *A stationary vacuum space-time with Σ complete and $|\xi|^2 = -\xi_\alpha \xi^\alpha$ tending uniformly to 1 from below at infinity is flat.*

Proof. Suppose first that the space-time is static. Choose the slicing (lapse N, shift X) of the space-time to be associated with the Killing field, *i.e.* $X = 0$. In such a slicing $k = 0$ and g is independent of λ. Thus the evolution equation for $\operatorname{tr} k$ (see proposition 2.1)

$$\frac{\partial}{\partial \lambda}(\operatorname{tr} k) = (Nk \cdot k + \Delta N) - X \cdot \mathrm{d}(\operatorname{tr} k) + N\, {}^{(4)}R_{\perp\perp}$$

reduces to

$$\Delta N + N\, {}^{(4)}R_{\perp\perp} = 0;$$

if the spacelike slices Σ are compact, this equation implies ${}^{(4)}R_{\perp\perp} = 0$ and $N = \text{const}$; the evolution equation for k then gives $N\,{}^{(4)}R_{\perp a \perp b} = 0$. Therefore, since $N \neq 0$ and $k = 0$, the Gauss equation and the mixed energy condition gives $\mathrm{Ric}\,(g) = 0$, so g is flat. The whole space-time is thus flat.

When the spacelike sections are noncompact, the equation for N has a unique positive bounded solution N if ${}^{(4)}R_{\perp\perp} \geqslant 0$ and appropriately approaches 1 at infinity. Only in the case ${}^{(4)}R_{\perp\perp} = 0$ can we deduce from the equation that $N = \text{const}$, and the flatness of the space-time follows as before.

Next we consider the stationary case. The equation of evolution for $\mathrm{tr}\,k$ still gives the result in the case Σ is compact under the added hypothesis that $\mathrm{tr}\,k = 0$, but even with this supplementary assumption it will not lead to a conclusion in the noncompact case.

The hint, in the stationary case, is not to use the evolution equation for $\mathrm{tr}\,k$ but another equation obtained by taking as 4-vector off the 3-manifold not its normal but the Killing field (and here Σ should be considered as the quotient of V_4 by the isometry group rather than as an embedded hypersurface). A straightforward computation, using for instance orthonormal frames, gives

$$\Delta U - Uh\cdot h + U\,{}^{(4)}\mathrm{Ric}\,(\xi, \xi) = 0\,,$$

where $U^2 = -\xi_\lambda \xi^\lambda$ is the square length of the Killing field, $h\cdot h$ is the square of the vorticity tensor associated with its shift and Δ the Laplace operator ($\Delta = -\nabla\cdot\nabla$) in the positive quotient metric on Σ. In the case of compact Σ we deduce from this equation that, in the vacuum case, $U = \text{const}$ and $h = 0$.

The space-time is therefore static, and the above conclusions apply.

If Σ is noncompact the conclusion follows from the maximum principle, if U is to be a positive function attaining its minimum on Σ (and not at infinity).

To complete the proof, we note the further identity for stationary space-times:

$$UR_{\lambda\mu}\xi^\lambda n^\mu = -\nabla\cdot(\mathrm{grad}\,U + X\cdot n)\,,$$

which implies $R_{\lambda\mu} = 0$ if Σ is compact and if $R_{\lambda\mu}\xi^\lambda n^\mu$ vanishes only for vacuum space-times (mixed energy condition). $\square$

If the space-time is asymptotically flat and if U has the form

$$U = 1 - \frac{2M}{r} + O\left(\frac{1}{r^2}\right),$$

the above identity implies

$$\int_\Sigma UR_{\lambda\mu}\xi^\lambda h^\mu \mu(g) = 8\pi M$$

and, therefore, $M \geqslant 0$ if the sources satisfy $R_{\lambda\mu}\xi^\lambda h^\mu \geqslant 0$. Such an inequality is satisfied by the perfect fluids. In the static case it is a consequence of the strong energy condition. If we restrict our attention to space-times which are asymptotically spherically symmetric, *i.e.* have a Schwarzschild-type asymptotic behaviour, the mass we have just computed (called the Komar mass) is identical with the usual, or ADM mass. However, in general, they are different and the above proof of positivity of the (Komar) mass does not help much with the proof of positivity of the usual mass. See [23] for further details.

A straightforward consequence of the evolution equation for $\operatorname{tr} k$ is the following theorem due to KOMAR [15], which does not assume stationarity:

4.2. *Theorem.* – *Suppose* $(V_4, {}^{(4)}g)$ *is a space-time satisfying the mixed energy condition and which admits a slicing by compact maximal hypersurfaces. Then the spacetime is flat.*

Proof. If $\operatorname{tr} k = 0$ for each λ of the slicing, we have on each slice

$$\Delta N + N(k\cdot k + {}^{(4)}R_{\perp\perp}) = 0 .$$

Therefore, if the slices are compact,

$$k = 0 , \quad {}^{(4)}R_{\perp\perp} = 0 \quad \text{and} \quad N = \text{constant} \quad \text{on each slice} .$$

The evolution equation for k and the mixed energy condition then implies

$$\operatorname{Ric}(g) = 0 , \qquad \textit{i.e. } g \text{ is flat} .$$

The space-time being empty, with Cauchy data ($k = 0$, g flat) is flat. □

We note that the same proof shows that the theorem is still true if we replace « maximal » by « with a given constant $\operatorname{tr} k$ ».

It is also implicit in Komar's work that maximal hypersurfaces must be isolated or be moments of time symmetry. The following holds:

4.3. *Proposition.* – *Let* $(V_4, {}^{(4)}g)$ *satisfy the strong energy condition and have a compact Cauchy surface* Σ *on which* $\operatorname{tr} k = 0$. *Then no other nearby hypersurface can have* $\operatorname{tr} k = 0$ *unless* Σ *is a moment of time symmetry.*

Proof. Again, examine the evolution equation for $\operatorname{tr} k$

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}(\operatorname{tr} k) = Nk\cdot k + \Delta N + N\,{}^{(4)}R_{\perp\perp} - X\cdot \mathrm{d}(\operatorname{tr} k)$$

at $\lambda = 0$ in any curve of spacelike embeddings, $N \neq 0$. This equation implies

that either

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\int \operatorname{tr} k\,\mu(g)\Big|_{\lambda=0} = \int \left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\operatorname{tr} k\right)\mu(g)\Big|_{\lambda=0} > 0$$

or $k = 0$ (and $N = \text{const}$, ${}^{(4)}R_{\perp\perp} = 0$, as above). Thus, if $k \neq 0$ on any (nearby) hypersurface, $\int_{\Sigma} \operatorname{tr} k\mu(g)$ must be greater, so the result follows. □

From these ideas we can also see where the term « maximal » (or « minimal » depending on conventions) arises. Indeed, in a slicing,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}(\operatorname{Vol}\Sigma_\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{\Sigma_\lambda}\mu(g) = \int_{\Sigma_\lambda}\frac{1}{2}\operatorname{tr}\left(\frac{\partial g}{\partial\lambda}\right)\mu(g) = -\int_{\Sigma_\lambda} N(\operatorname{tr} k)\mu(g)\,,$$

since $\mathrm{d}g/\mathrm{d}\lambda = -2Nk - L_X g$ and Σ_λ is compact. Thus at Σ_0, if $\operatorname{tr} k = 0$, the function $\operatorname{Vol}(\Sigma_\lambda)$ has a critical point. Also,

$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}(\operatorname{Vol}\Sigma_\lambda)|_{\lambda=0} = -\int_{\Sigma_0} N\frac{\partial}{\partial\lambda}(\operatorname{tr} k)\mu(g) =$$

$$= -\int N^2 k\cdot k\mu(g) - \int(\nabla N)^2\mu(g) - \int N^2 R_{\perp\perp}\mu(g) < 0$$

if k and N are not zero.

This argument gives the following variation of this type of result.

4.4 *Proposition.* – *Suppose* $(V_4, {}^{(4)}g)$ *is a space-time with* $\operatorname{Ric}{}^{(4)}g \geqslant 0$ *and* Σ_0 *a compact Cauchy surface, with* $\operatorname{tr} k = 0$. *Either every compact Cauchy surface* Σ *near* Σ_0 *has volume* $\operatorname{Vol}\Sigma < \operatorname{Vol}\Sigma_0$ *or else* Σ_0 *is a moment of time symmetry.*

Remark. In [70], p. 527, there is an isolation result related to this one. Namely any perturbation of a flat space-time (with g flat, $k = 0$ on Σ_0) which satisfies the weak energy condition and which has a maximal hypersurface is flat. In fact, any perturbation of a flat space-time which preserves a timelike Killing field is flat (cf. theorem 4.1 above and the equation $R(g) = k\cdot k - (\operatorname{tr} k)^2 + 2\,{}^{(4)}G_{\perp\perp}$).

The results of Komar may be considered as a weaker version of a uniqueness (or isolation) theorem obtained recently by CHOQUET-BRUHAT [37-39] and BRILL and FLAHERTY [41] for compact spacelike hypersurfaces with $\operatorname{tr} k = a$, a given constant. While BRILL and FLAHERTY base their argument on a fourth-order variational argument (suggested by work of FRANKEL in 1961), CHOQUET-BRUHAT uses a maximum principle argument. The result is as follows.

4.5. *Theorem.* – *Suppose that* $(V_4, {}^{(4)}g)$ *is a nowhere source-free space-time* (*i.e.* ${}^{(4)}\operatorname{Ric}(u, u) > 0$ *for every timelike vector field* u). *Suppose that* Σ_0 *is a compact spacelike hypersurface with* $\operatorname{tr} k = a$ *a given constant. Then there exists no other such hypersurface in a neighbourhood of* Σ_0.

Proof. Consider a Gaussian normal co-ordinate neighbourhood U about Σ_0. There is a one-to-one correspondence between hypersurfaces Σ_φ near Σ_0 and functions φ on Σ_0; in fact Σ_φ can be taken as the graph of φ. Suppose that $P(\varphi) = a$, where $P(\varphi)$ is the trace of the second fundamental form of Σ_φ. Now $P(\varphi)$ is a second-order (nonlinear) elliptic operator on φ. At a critical point $\bar{x}$ of φ where $\nabla\varphi = 0$, this operator reduces to

$$P(\varphi)(\bar{x}) = (\Delta\varphi + \operatorname{tr} k_{\bar{\lambda}})(\bar{x}) = a\,,$$

where $\bar{\lambda} = \varphi(\bar{x})$ and k_λ is the trace of the second fundamental form of the $\bar{\lambda} = t$ slice in the Gaussian co-ordinates. If $R_{\perp\perp} > 0$, $\operatorname{tr} k_\lambda$ is a strictly increasing function of λ, as above. Thus φ cannot have a positive maximum $\bar{\lambda}$ nor a negative minimum; thus φ must be identically sero if Σ_0 is compact. □

The arguments of BRILL and FLAHERTY [41] also show *global* uniqueness.

An analogous argument proves that a Robertson-Walker space-time, *i.e.* $V_4 = M \times \boldsymbol{R}$, ${}^{(4)}g = -\,\mathrm{d}t^2 + f(t)\,\mathrm{d}s^2$, admits a maximal spacelike hypersurface if and only if it admits a totally geodesic submanifold (*i.e.* a moment of time symmetry).

A different method gives the following theorem [39] which includes the case of flat space-times:

4.6. *Theorem.* – *Let Σ_0 be a maximal spacelike hypersurface of a space-time $(V_4, {}^{(4)}g)$ with constant Riemannian curvature $c \leqslant 0$. Then Σ_0 is totally geodesic (i.e. a moment of time symmetry; $k = 0$) in the following cases:*

1) *Σ_0 is compact,*

2) *Σ_0 is complete and $k \cdot k$ tends to zero at infinity.*

Proof. A computation using standard identities in Riemannian geometry shows that for a maximal hypersurface in a space of constant curvature

$$\Delta(k\cdot k) + \nabla k \cdot \nabla k + k\cdot k(k\cdot k - c) = 0\,,$$

which gives the result by the maximum principle. □

In the case $c = 0$, $V_4 = \boldsymbol{R}^4$ and ${}^{(4)}g = \eta$ (Minkowski space-time) the above theorem is a weaker version of the Bernstein-Calabi theorem (which makes no restriction on $k\cdot k$ at infinity, but has a more involved proof); see CHENG and YAU [69].

It is possible to give a stronger version of theorem 4.5 which also includes the case of flat space. Namely theorem 4.5 remains valid if ${}^{(4)}\mathrm{Ric}\,(u, u) > 0$ is replaced by either the mixed energy condition or a generic condition (see [64]). For the proof see TIPLER and MARSDEN [43].

5. – Derivation of the mass formula.

This section reviews the textbook derivation of the formula for the mass and total momentum of an asymptotically flat space-time. The derivation here is to be compared with the more intrinsic description of mass as the generator of time translations, given in sect. **10**. Here we shall follow the description of Misner, Thorne and Wheeler [74], Chapt. 19 and 20, although we shall use the linearized Einstein tensor to give a more compact presentation. (The book of Weinberg [75] is also a useful reference.)

To define energy and angular momentum the space-time must be asymptotically flat. Co-ordinates can then be chosen so that asymptotically they are the standard co-ordinates on $\boldsymbol{R}^4$; in these co-ordinates, ${}^{(4)}g$ will be asymptotic to the Minkowski metric η.

For an arbitrary *isolated* gravitating system described by a stress-energy tensor $T_{\mu\nu}$, let ${}^{(4)}g$ be an asymptotically flat solution to the full nonlinear field equations

$$G_{\mu\nu} = 8\pi T_{\mu\nu}\,,$$

where ${}^{(4)}g$ is taken in asymptotically Minkowskian co-ordinates. In such co-ordinates, the gravitational field for the full nonlinear equations with arbitrary isolated strong sources has the same $1/r$ fall-off as in the linearized theory for weak sources. Thus the flux integrals from linearized theory can be used to calculate the total linear momentum P^μ and angular momentum $J^{\mu\nu}$ for any isolated source, weak or strong, in full general relativity. These flux integrals, in either the full or linearized theory, represent the total linear and angular momentum of *both* the sources and the gravitational fields.

Since ${}^{(4)}g$ is asymptotically Minkowskian, we can compare ${}^{(4)}g$ with the Minkowski metric η on $\boldsymbol{R}^4$ by writing

$${}^{(4)}g = \eta + {}^{(4)}h\,.$$

Expanding Einstein's equations in a Taylor series about η, we have

$$\begin{aligned} 8\pi T_{\mu\nu} &= \mathrm{Ein}\,({}^{(4)}g) = \mathrm{Ein}\,(\eta + {}^{(4)}h) = \mathrm{Ein}\,(\eta) + \mathrm{D}\,\mathrm{Ein}\,(\eta)\cdot{}^{(4)}h + \\ &\qquad + \text{nonlinear correction terms} = \\ &= 0 + \mathrm{D}\,\mathrm{Ein}\,(\eta)\cdot{}^{(1)}h - [\mathrm{D}\,\mathrm{Ein}\,(\eta)\cdot{}^{(1)}h - \mathrm{Ein}\,({}^{(1)}g)] = 0 + 8\pi T^{\mathrm{eff}}_{\mu\nu} - 8\pi t_{\mu\nu}\,, \end{aligned}$$

where $T^{\mathrm{eff}}_{\mu\nu}$ and $t_{\mu\nu}$ are defined by

$$\begin{aligned} 8\pi T^{\mathrm{eff}}_{\mu\nu} &= \mathrm{D}\,\mathrm{Ein}\,(\eta)\cdot{}^{(1)}h\,, \\ 8\pi t_{\mu\nu} &= \mathrm{Ein}\,({}^{(4)}g) - \mathrm{D}\,\mathrm{Ein}\,(\eta)\cdot{}^{(1)}h\,, \end{aligned}$$

so that

$$T^{\text{eff}}_{\mu\nu} = T_{\mu\nu} + t_{\mu\nu} .$$

Thus $8\pi T^{\text{eff}}_{\mu\nu}$ is the first-order approximation of Ein $({}^{(4)}g)$, expressed in terms of the linearized Einstein tensor around Minkowski space and ${}^{(4)}h = {}^{(4)}g - \eta$. Note that

$$\mathrm{D\,Ein}\,(\eta)\cdot{}^{(4)}h = \mathrm{D\,Ein}\,(\eta)\cdot({}^{(4)}g - \eta) = \mathrm{D\,Ein}\,(\eta)\cdot{}^{(4)}g .$$

The nonlinear correction terms, defined by $8\pi t_{\mu\nu}$, are a « stress-energy pseudotensor for the gravitational field », as quoted in [74], p. 465. Also $T^{\text{eff}}_{\mu\nu} = T_{\mu\nu} + t_{\mu\nu}$ represents the total stress-energy tensor for both the isolated gravitating sources and the gravitational fields.

By the linearized contracted Bianchi identities (theorem 4.2 of [1]) and since $\mathrm{Ein}\,(\eta) = 0$, $T^{\text{eff}}_{\mu\nu}$ automatically has zero divergence in the Minkowski metric

$$(T^{\text{eff}})_{\mu}{}^{\nu}{}_{,\nu} = (T_{\mu}{}^{\nu} + t_{\mu}{}^{\nu})_{,\nu} = 0 ,$$

where indices are raised with respect to $\eta^{\mu\nu}$. These equations are the flat-space equivalent of $T_{\mu}{}^{\nu}{}_{;\nu} = 0$.

Also from sect. 4 of [1], we know that

$$\mathrm{D\,Ein}\,(\eta)\cdot{}^{(4)}h = \tfrac{1}{2}(-\eta^{\mu\nu}\,{}^{(4)}\bar{h}_{\lambda\beta,\mu,\nu} + {}^{(4)}\bar{h}_{\lambda}{}^{\mu}{}_{,\mu,\beta} + {}^{(4)}\bar{h}_{\beta}{}^{\mu}{}_{,\mu,\lambda} - \eta^{\mu\nu}\,{}^{(4)}h_{\mu\nu,\lambda,\beta}) ,$$

where again all indices are raised with respect to $\eta^{\mu\nu}$; from this expression we can see directly that for all ${}^{(4)}h$

$$\delta_{\eta}(\mathrm{D\,Ein}\,(\eta)\cdot{}^{(4)}h) = 0 .$$

Remarkably, $\mathrm{D\,Ein}\,(\eta)\cdot{}^{(4)}h$ also has a potential, namely

$$(\mathrm{D\,Ein}\,(\eta)\cdot{}^{(4)}h)^{\mu\nu} = \tfrac{1}{2}Q^{\mu\nu\beta}{}_{,\beta} ,$$

where $Q^{\mu\nu\beta}$, antisymmetric in $\nu\beta$, is defined by

$$Q^{\mu\nu\beta} = -\eta^{\lambda\beta}\,{}^{(4)}\bar{h}^{\mu\nu}{}_{,\lambda} + \eta^{\mu\beta}\,{}^{(4)}\bar{h}^{\lambda\nu}{}_{,\lambda} + \eta^{\lambda\nu}\,{}^{(4)}\bar{h}^{\mu\beta}{}_{,\lambda} - \eta^{\mu\nu}\,{}^{(4)}\bar{h}^{\lambda\beta}{}_{,\lambda} .$$

Since the $1/r$ fall-off of full general relativity for the gravitational field of arbitrarily strong isolated sources is the same as for the linearized theory with weak sources, *asymptotically* the linearized theory around η and the full general theory are « the same ». Hence the flux integrals from linearized theory can be used to calculate the total linear and angular momentum of both the isolated

gravitating sources and the gravitational fields. Thus, we define P^μ by

$$P^\mu = \frac{1}{16\pi}\oint_S Q^{\mu 0 j}\,\mathrm{d}S_j = \frac{1}{16\pi}\int_{R^3} Q^{\mu 0 j}{}_{,j}\,\mathrm{d}^3x = \frac{1}{16\pi}\int_{R^3} Q^{\mu 0 \beta}{}_{,\beta}\,\mathrm{d}^3x = \int_{R^3} T^{\mu 0}_{\mathrm{eff}}\,\mathrm{d}^3x$$

and similarly

$$J^{\mu\nu} = \int_{R^3} (x^\mu T^{\nu 0}_{\mathrm{eff}} - x^\nu T^{\mu 0}_{\mathrm{eff}})\,\mathrm{d}^3x\,.$$

An examination of $t_{\mu\nu}$ shows that it is $O(1/r^4)$, so is integrable; thus, P^μ is well defined for finite sources. $J^{\mu\nu}$ is a surface integral of something $O(1/r^2)$, and so it is well defined as well [74, ex. 20.2].

These formulae are valid in full general relativity for an isolated source, *provided* the closed surface of integration S is taken in the asymptotically flat region surrounding the source and the space-time metric ${}^{(4)}g$ is taken asymptotically in Minkowskian co-ordinates. P^μ and $J^{\mu\nu}$ are then tensors in the asymptotically flat region surrounding the sources, *i.e.*, under a co-ordinate transformation asymptotic to the identity at infinity, P^μ is invariant, and, under a Lorentz transformation, P^μ transforms as a 4-vector (see [75], p. 169). Similarly $J^{\mu\nu}$ transforms as an antisymmetric 2-tensor.

Moreover, if the evolution equations are taken with lapse asymptotically one and shift asymptotically zero, the momentum P^μ is conserved on $t =$ const hypersurfaces (in particular the mass is conserved). This fact follows formally from $T^{\mu\nu}_{\mathrm{eff},\nu} = 0$, but a rigorous proof requires a careful examination of the Cauchy problem.

The expression for P^0, which defines the total mass-energy of the isolated gravitating system and the gravitational fields, and referred to as the mass of the system, is given by

$$m = P^0 = \frac{1}{16\pi}\oint_S (h_{ij,j} - h_{jj,i})\,\mathrm{d}S_i = \frac{1}{16\pi}\oint_S (g_{ij,j} - g_{jj,i})\,\mathrm{d}S_i =$$
$$= \frac{1}{16\pi}\int_{R^3} (g_{ij,j,i} - g_{jj,i,i})\,\mathrm{d}^3x\,.$$

Note that m involves only the metric coefficients of the induced metric on the $t =$ const hypersurface and not the second fundamental form of this hypersurface.

Interestingly, the integrand in the volume expression for m is just

$$\mathrm{D}R(\gamma)\cdot h = \mathrm{D}R(\gamma)\cdot g = \Delta_\gamma \mathrm{tr}_\gamma g + \delta_\gamma\delta_\gamma g = -(\gamma^{ik}g_{ki})_{,i,i} + g_{ii,j,j}\,,$$

where γ is the Euclidean metric on R^3 and $h = g - \gamma$. Thus we can write the

mass formula as

$$m = \frac{1}{16\pi}\int_{R^3} \mathrm{D}R(\gamma)\cdot g\,\mathrm{d}^3x\,.$$

Note that, if the volume integrals converge absolutely, it proves via Gauss' theorem that the surface integral expressions have unambiguous meaning; the integrals over spheres of radius R will converge as $R\to\infty$.

In order to see that m defines a smooth function on $\mathscr{M}^p_{s,\delta}$, it is necessary to rewrite the formula for m as a quadratic expression in Christoffel symbols of g. This will be shown in detail in sect. 7.

6. – Time-symmetric initial data sets as minima of mass.

In this section we describe an important contribution of O'Murchadha and York [29]. They show that the time-symmetric vacuum initial data sets ($\pi=0$) are minima of the mass function among $\operatorname{tr}\pi=0$ initial data sets. Since we have seen that every space-time (near flat space) has such a maximal hypersurface, this will reduce the question of positivity of mass to the time-symmetric empty-space case.

Their first result compares the mass of conformally related 3-metrics on $\mathbf{R}^3$.

6.1. *Lemma.* – *Let g be asymptotically flat on $\mathbf{R}^3$, i.e. $g\in\mathscr{M}^p_{s,\delta}$. Let $\varphi>0$ and let φ be asymptotically 1, $\varphi-1\in M^p_{s,\delta}$. Let $\bar g=\varphi^4 g$ (i.e. $\bar g$ is conformally related to g). Then $\bar g$ is also asymptotically flat (i.e. $\bar g\in\mathscr{M}^p_{s,\delta}$) and*

$$16\pi m(\bar g) = 16\pi m(g) - 8\oint_\infty (\operatorname{grad}\varphi)^i\,\mathrm{d}S_i = 16\pi m(g) + 8\int_{R^3}\Delta_g\varphi\,\mu(g)$$

(grad φ *and the surface integral are with respect to either g or γ on* $\mathbf{R}^3$).

Proof. That $\bar g=\varphi^4 g\in\mathscr{M}^p_{s,\delta}$ follows from the fact that, if $\varphi-1\in M^p_{s,\delta}$, then $\varphi^4-1\in M^p_{s,\delta}$. Hence, if $g-\gamma\in S^p_{s,\delta}$, then $\bar g-\gamma\in S^p_{s,\delta}$. Thus $m(\bar g)$ is defined. Moreover, $\mathrm{d}S_i=\mathrm{d}\bar S_i$ is a conformal invariant. Hence the proposition then follows by this computation:

$$\begin{aligned} m(\bar g) &= \frac{1}{16\pi}\oint_\infty \bar g^{ab}\bar g^{ij}(\bar g_{ai,j}-g_{ij,a})\sqrt{\det\bar g}\,\mathrm{d}\bar S_b = \\ &= \frac{1}{16\pi}\oint_\infty \varphi^{-8} g^{ab}g^{ij}[(\varphi^4 g_{ai})_{,j}-(\varphi^4 g_{ij})_{,a}]\varphi^6\sqrt{\det g}\,\mathrm{d}S_b = \\ &= \frac{1}{16\pi}\oint_\infty \varphi^2 g^{ab}g^{ij}[g_{ai,j}-g_{ij,a}]\sqrt{\det g}\,\mathrm{d}S_b + \end{aligned}$$

$$+\frac{4}{16\pi}\oint_{\infty} \varphi^{-2} g^{ab} g^{ij} [\varphi_{,j} g_{ai} - \varphi_{,a} g_{ij}] \sqrt{\det g}\, dS_b =$$

$$= m(g) + \frac{4}{16\pi}\oint_{\infty} \varphi^{-2}(1-3)(\operatorname{grad}\varphi)^b\, dS_b = m(g) - \frac{8}{16\pi}\oint_{\infty} (\operatorname{grad}\varphi)^b\, dS_b\,,$$

where the last 2 steps have used the fact that $\varphi \in M^p_{s,\delta}$. □

Secondly, O'Murchadha and York prove the following

6.2. *Proposition.* – *Let g be an asymptotically flat metric on $\mathbf{R}^3$, $g - \gamma \in S^p_{s,\delta}$, such that $R(g) \geqslant 0$. Then there exists a unique $\varphi > 0$, $\varphi - 1 \in M^p_{s,\delta}$, such that the pointwise conformally related asymptotically flat metric $\bar{g} = \varphi^4 g$, $\bar{g} - \gamma \in S^p_{s,\delta}$, satisfies*

$$R(\bar{g}) = 0\,.$$

Moreover,

$$m(\bar{g}) \leqslant m(g)\,.$$

If g is in a $S^p_{s,\delta}$ neighbourhood of flat space, then so is $\bar{g}$.

Proof. From Lichnerowicz' famous formula [32], if $\bar{g} = \varphi^4 g$, $\varphi > 0$, then

$$R(\bar{g}) = 8\varphi^{-5}(\Delta_g \varphi + \varphi R(g))\,.$$

Since $R(g) \geqslant 0$, by the elliptic theory from sect. 1 and the maximum principle, the system

$$\Delta_g \varphi + R(g)\varphi = 0$$

has a unique positive solution φ such that $\varphi - 1 \in M^p_{s,\delta}$. For this solution φ, $R(\varphi^4 g) = R(\bar{g}) = 0$. Hence $\Delta_g \varphi = -R(g)\varphi$, so from lemma 6.1

$$m(\bar{g}) = m(g) + \frac{1}{2\pi}\int_{R^3} \Delta_g \varphi\, \mu(g) = m(g) - \frac{1}{2\pi}\int_{R^3} \varphi R(g) \mu(g)\,.$$

Since $\varphi > 0$ and $R(g) \geqslant 0$, $m(\bar{g}) \leqslant m(g)$. The integral $= 0$ if and only if $R(g) = 0$ if and only if $\varphi = 1$ if and only if $\bar{g} = g$.

If g is in a $S^p_{s,\delta}$ neighbourhood of flat space, $R(g)$ is in a $M^p_{s-2,\delta+2}$ neighbourhood of 0, and so, by regularity of the Laplacian and continuous dependence of solutions on coefficients, φ is in a $M^p_{s,\delta}$ neighbourhood of 1. Hence $\bar{g} = \varphi^4 g$ is in a $S^p_{s,\delta}$ neighbourhood of flat space also. □

Note that in 6.2, the space-time plays no role.

The following main result correlates the previous discussions.

6.3. *Theorem* (O'MURCHADHA and YORK [29]). – *Let* $(V_4, {}^{(4)}g)$ *be an asymptotically flat space-time satisfying the weak energy condition* $\mathrm{Ein}\,({}^{(4)}g) \geqslant 0$. *Let* $\Sigma = i(\mathbf{R}^3)$ *be a maximal asymptotically flat spacelike hypersurface. Then there exists a unique asymptotically flat conformally related metric* $\bar{g} = \varphi^4 g$ *on* Σ *such that* $R(\bar{g}) = 0$ *and* $m(\bar{g}) \leqslant m(g)$. *Furthermore,* $m(\bar{g}) = m(g)$ *if and only if* ${}^{(4)}G_{\perp\perp} = 0$ *and* $k_{\Sigma} = 0$.

Remark. Thus $(\bar{g}, 0)$ is a solution to the empty-space time-symmetric constraint equations with $m(\bar{g}) \leqslant m(g)$. Thus, this theorem proves that the mass of maximal Cauchy data in space-times that satisfy the weak energy condition can always be bounded below by the mass of a conformally related metric $\bar{g}$, such that $(\bar{g}, 0)$ satisfies the empty-space time-symmetric constraint equations. It will be important for us later to observe that $\bar{g}$ is in fact near flat space if g is. This requires an analysis of the continuity of the solutions of the above differential equations for φ as a function of the metric g.

Proof. The following is an identity for any spacelike hypersurface in any space-time $(V_4, {}^{(4)}g)$:

$$R(g) = k \cdot k - (\operatorname{tr} k)^2 + 2\,{}^{(4)}G_{\perp\perp}\,,$$

where (g, k) are the metric and second fundamental form on Σ and

$${}^{(4)}G_{\perp\perp} = {}^{(4)}G_{\perp\perp}(\Sigma) = \mathrm{Ein}\,({}^{(4)}g) \cdot ({}^{(4)}Z_{\Sigma}, {}^{(4)}Z_{\Sigma})\,.$$

Thus, if Σ is maximal, $\operatorname{tr} k = 0$, and the space-time satisfies the weak energy condition ${}^{(4)}G_{\perp\perp} \geqslant 0$, then

$$R(g) = k \cdot k + 2\,{}^{(4)}G_{\perp\perp} \geqslant 0\,.$$

The result then follows from proposition 6.2. □

Basically, theorem 6.2 says that, if we add matter to an empty-space time-symmetric asymptotically flat solution of the Einstein equations such that the weak energy condition is satisfied, or if we add momentum to the time-symmetric Cauchy data such that $\operatorname{tr} \pi = 0$, then the mass is increased.

The fact that this theorem reduces the (local) positivity-of-mass question for asymptotically flat space-times near flat space that satisfy the weak and strong energy conditions to the case of Einstein-flat asymptotically flat time-symmetric space-times near flat space will be exploited in subsequent sections. See [47] for further discussion.

7. – Positivity of mass for space-times satisfying the weak and strong energy conditions.

We now come to the question of local positivity of mass. Again, local here means in a $M^p_{s,\delta}$ neighbourhood of Minkowski space. We first prove positivity for an asymptotically flat time-symmetric empty-space solution to Einstein's equations, and then for a general asymptotically flat space-time that satisfies the weak and strong energy conditions.

Before proceeding to prove the positivity of mass we mention the important work of Regge and Teitelboim [12]. Their work is relevant here because they show that for asymptotically flat space-times the Hamiltonian which actually generates the field equations in dynamical form must contain the mass. The mass arises in their Hamiltonian because, in taking the variational derivative of the scalar curvature, the second derivatives of the metric coefficient give rise to a surface term which does not go to zero at infinity, and in fact is just the negative of the mass. The mass is then added to compensate for this surface term (see proposition 7.1).

Thus, they show that the Hamiltonian which generates the dynamical equations for empty-space asymptotically flat space-time is

$$G_{\mathrm{RT}}(g,\pi) = 16\pi m(g) + \int_{R^3} (N\mathscr{H}(g,\pi) + X\cdot\mathscr{J}(g,\pi)) = 16\pi m(g) + G_{\mathrm{ADM}}(g,\pi)\,,$$

where $G_{\mathrm{ADM}} = \int (N\mathscr{H} + X\mathscr{J})$ is the generator of the dynamical equations for space-times with compact Cauchy hypersurfaces; see sect. **2** of [1] for the other notation.

In the time-symmetric case, the relevant generator reduces to

$$16\pi\overline{m}(g) \equiv G_{\mathrm{RT}}(g,0) = 16\pi m(g) - \int_{R^3} NR(g)\mu(g)$$

for a given lapse $N > 0$, and N asymptotically 1, $N - 1 \in M^p_{s,\delta}(\mathbf{R}^3)$, so that $\overline{m}$ is taken as a map

$$\overline{m} : \mathscr{M}^p_{s,\delta} \to \mathbf{R}\,.$$

The importance of $\overline{m}(g)$ is that it does not contain second derivatives of the metric coefficients. Indeed, from the volume integral for the mass,

$$16\pi m(g) = \int_{R^3} (Ng^{ij}g^{kl}(g_{ik,j} - g_{ij,k})\sqrt{\det g})_{,l}\, \mathrm{d}^3x\,,$$

we see that the second-derivative terms of the metric tensor that occur in the

formula for $16\pi m(g)$ are exactly the second-derivative terms that occur in the scalar curvature

$$R(g) = g^{ij}g^{kl}(g_{ik,j,l} - g_{ij,k,l}) + \Gamma^i_{jl}\Gamma^j_{kl}g^{kl} - \Gamma^i_{jk}\Gamma^j_{il}g^{kl}\,.$$

Hence $\overline{m}(g)$ contains only first derivatives of g and these occur as products. These product terms are in L_1 by the lemmas from sect. 1. It follows that $\overline{m}(g)$ is a C^∞ function of g. Also, since m and $\overline{m}$ agree on $\mathring{\mathscr{M}}^p_{\varepsilon,\delta}$, m restricted to $\mathring{\mathscr{M}}^p_{\varepsilon,\delta}$ is a C^∞ function of g as well.

Now we compute the derivative (with respect to g) of $\overline{m}(g)$. This is the crucial computation in the Regge-Teitelboim analysis. The asymptotic conditions that we use are recalled as follows:

$$\begin{aligned}
g_{ij} &= \delta_{ij} + O\left(\frac{1}{r}\right),\\
g_{ij,k} &= O\left(\frac{1}{r^2}\right),\\
g_{ij,k,l} &= O\left(\frac{1}{r^3}\right),\\
h_{ij} &= O\left(\frac{1}{r}\right),\\
h_{ij,k} &= O\left(\frac{1}{r^2}\right),\\
N &= 1 + O\left(\frac{1}{r}\right),\\
N_{,i} &= O\left(\frac{1}{r^2}\right),\\
N_{,i,j} &= O\left(\frac{1}{r^3}\right),
\end{aligned}
\qquad \text{as } r \to \infty\,.$$

One can use the lemmas of sect. 1 to verify our assertions, or, to see the results quickly on a first reading, one can use the following two elementary facts:

1) A volume integral $\int_{R^3} f\mu(g)$ is convergent if $f = O(1/r^{3+\varepsilon})$, $\varepsilon > 0$ (and may or may not be convergent if $f = O(1/r^3)$).

2) A volume integral which is a divergence

$$\int_{R^3} \operatorname{div} X\,\mu(g) = \oint_\infty X^k \cdot \mathrm{d}S_k$$

converges if $X = O(1/r^2)$, and converges to zero if $X = O(1/r^{2+\varepsilon})$, $\varepsilon > 0$.

7.1. *Proposition.* – *Let* $\overline{m}: \mathcal{M}^p_{s,\delta} \to \mathbf{R}$ *be defined by*

$$16\pi\overline{m}(g) = 16\pi m(g) - \int_{R^3} NR(g)\mu(g) .$$

Then, for $h \in S^p_{s,\delta}$,

$$16\pi\, \mathrm{d}\overline{m}(g)\cdot h = \int_{R^3} N \operatorname{Ein}(g)\cdot h\mu(g) - \int_{R^3} (\operatorname{Hess} N + g\Delta N)\cdot h\mu(g) .$$

Proof. First we compute the derivative of $-\int NR(g)\mu(g)$. From p. 387 of [1],

$$\mathrm{d}\Big(\int_{R^3} NR(g)\mu(g)\Big)\cdot h = -\int_{R^3} N\,\mathrm{D}R(g)\cdot h\mu(g) - \int_{R^3} NR(g)\tfrac{1}{2}(\operatorname{tr} h)\mu(g) =$$
$$= -\int_{R^3} N\big(\delta\,\delta h + \Delta(\operatorname{tr} h)\big)\mu(g) + \int_{R^3} N \operatorname{Ein}(g)\cdot h\mu(g) .$$

Now $\operatorname{Ein}(g) = O(1/r^3)$, $h = O(1/r)$, $N = 1 + O(1/r)$, so $N \operatorname{Ein}(g)\cdot h = O(1/r^4)$, and hence the last integral converges. The first integral, however, involving second derivatives of h like $\delta\,\delta h = O(1/r^3)$, may or may not converge. If $N = \text{const}$, then this integral is a divergence and so does converge.

To evaluate this first integral, and to see if it does converge, use the identity

$$(7.1) \qquad -N(\delta\,\delta h + \Delta \operatorname{tr} h) = -(\operatorname{Hess} N + g\Delta N)\cdot h -$$
$$- \delta\big(N\,\mathrm{d}(\operatorname{tr} h) - (\operatorname{tr} h)\,\mathrm{d}N + h\cdot\mathrm{d}N + N\,\delta h\big)$$

to get

$$-\int_{R^3} N(\delta\,\delta h + \Delta \operatorname{tr} h)\mu(g) = -\int_{R^3} (\operatorname{Hess} N + g\Delta N)\cdot h\mu(g) -$$
$$-\int_{R^3} \delta\big(N(\delta h + \mathrm{d}\operatorname{tr} h)\big)\mu(g) - \int_{R^3} \delta\big(h\cdot\mathrm{d}N - (\operatorname{tr} h)\,\mathrm{d}N\big)\mu(g) .$$

Consider first the last integral. Since $h = O(1/r)$, $\operatorname{tr} h = O(1/r)$, $\mathrm{d}N = O(1/r^2)$, the integrand is a divergence of something $O(1/r^3)$. Hence converting it to a surface integral shows that it is zero. The next to last integral is the divergence of something of the order $\delta h = O(1/r^2)$, since $N = 1 + O(1/r)$. Hence the next to last integral is finite, and in fact is exactly $-16\pi\,\mathrm{d}m(g)\cdot h$, as we shall see.

The first integral, involving second derivatives of N, like $\operatorname{Hess} N = O(1/r^3)$, contracted with $h = O(1/r)$, is $O(1/r^4)$ at infinity and hence is a convergent integral.

Thus $-\int_{R^3} N(\delta\,\delta h + \Delta \operatorname{tr} h)\mu(g)$ is convergent.

Althogether, we have

$$\mathrm{d}\Big(-\int_{\hat{R}^3} NR(g)\mu(g)\Big)\cdot h = \int_{\hat{R}^3} N\,\mathrm{Ein}\,(g)\cdot h\mu(g) - \\ -\int_{\hat{R}^3}(\mathrm{Hess}\,N + g\Delta N)\cdot h\mu(g) - \int_{\hat{R}^3}\delta\big(N(\delta h + \mathrm{d}\,\mathrm{tr}\,h)\big)\mu(g)\,.$$

To evaluate $\mathrm{d}m(g)\cdot h$, we use the volume expression for $m(g)$ to get

$$16\pi\,\mathrm{d}m(g)\cdot h = -\int_{\hat{R}^3}\Big(N\big(h^{ij}g^{kl}(g_{ik,j}-g_{ij,k}) + g^{ij}h^{kl}(g_{ik,j}-g_{ij,k})\big)\sqrt{\det g}\Big)_{,l}\,\mathrm{d}^3x + \\ + \int_{\hat{R}^3}\big(Ng^{ij}g^{kl}(h_{ik,j}-h_{ij,k})\sqrt{\det g}\big)_{,l}\,\mathrm{d}^3x + \\ + \int_{\hat{R}^3}\Big(Ng^{ij}g^{kl}(g_{ik,j}-g_{ij,k})\frac{1}{2}(\mathrm{tr}\,h)\sqrt{\det g}\Big)_{,l}\,\mathrm{d}^3x\,.$$

Now the terms involving derivatives of g, $g_{ik,j} = O(1/r^2)$ are all multiplied by $h = O(1/r)$. Hence these terms are the divergence of something $O(1/r^3)$ and hence integrate to zero. Similarly, we can complete $h_{ik,j} - h_{ij,k}$ to $h_{ik|j} - h_{ij|k}$ by products of terms like $h_{ab}\Gamma^c_{de} = O(1/r^3)$. Since these terms occur as divergences, they integrate to zero. Hence

$$16\pi\,\mathrm{d}m(g)\cdot h = \int_{\hat{R}^3}\big(Ng^{ij}g^{kl}(h_{ik,j}-h_{ij,k})\sqrt{\det g}\big)_{,l}\,\mathrm{d}^3x = \\ = \int_{\hat{R}^3}\big(Ng^{ij}g^{kl}(h_{ik|j}-h_{ij|k})\sqrt{\det g}\big)_{,l}\,\mathrm{d}^3x = \\ = \int_{\hat{R}^3}\Big(N\big(h^{il}{}_{|l} - g^{ik}(\mathrm{tr}\,h)_{|k}\big)\sqrt{\det g}\Big)_{|i}\,\mathrm{d}^3x = \int_{\hat{R}^3}\delta\big(N(\delta h + \mathrm{d}(\mathrm{tr}\,h)\big)\mu(g)\,,$$

thereby identifying $16\pi\,\mathrm{d}m(g)\cdot h$ with exactly the negative of the «extra term» that occurs in the variation of $-\int_{\hat{R}^3} NR(g)\mu(g)$ in the noncompact (asymptotically flat) case. □

Remarks.

1) Perhaps the main point of this proposition can be summarized schematically as follows:

$$(\text{variation of } g)\cdot(\mathrm{D}g) = (h)\,\mathrm{D}g = O(1/r^3)\,,$$

and so the integral of its divergence is zero. But

$$(g)\cdot(\text{variation of } \mathrm{D}g) = (g)\cdot(\mathrm{D}h) = O(1/r^2)\,,$$

and so the integral of its divergence is finite, and is the mass integral. This difference arises in the behaviour of these integrals because $g_{ij} = \delta_{ij} + O(1/r)$, and these constants are annihilated in $\mathrm{D}g = O(1/r^2)$. This is the reason why the terms appearing in $\mathrm{d}m(g)\cdot h$ behave so differently. If $g = O(1/r)$, then there would be no difference, all of these terms would be zero. Similarly,

$$\int_{R^3} \delta(h\cdot \mathrm{d}N - (\operatorname{tr} h)\,\mathrm{d}N)\mu(g) = 0\,,$$

whereas

$$\int_{R^3} \delta(N(\delta h + \mathrm{d}\operatorname{tr} h))\mu(g) \neq 0\,.$$

2) The variation of the combinations of g's that occur in the integrand of $m(g)$, *viz.*

$$g^{ij}g^{kl}(g_{ik,j} - g_{ij,k})\sqrt{\det g}\,,$$

is not a tensor. Note that the variation of an expression involving first derivatives of g may or may not be a tensor, *e.g.* $\mathrm{D}(\Gamma^i_{jk})\cdot h$ is a tensor, whereas $\mathrm{D}(g^{jk}\Gamma^i_{jk})\cdot h$ is not a tensor. Since the g's that occur cannot be written as an expression involving Christoffel symbols alone (the criterion for the variation to be a tensor), the variation of the above expresion is not a tensor. What we have shown however is that it differs by a tensor by terms that integrate to zero.

3) The identity 7.1 is analogous to the identity used in potential theory

$$\psi_1\Delta\psi_2 - \psi_2\Delta\psi_1 = \operatorname{div}(\psi_2\nabla\psi_1 - \psi_1\nabla\psi_2)\,.$$

As in proposition 7.1, the difference of an operator and its formal adjoint is the divergence of an antisymmetric bilinear form.

The above proposition contains a key part of the Regge-Teitelboim analysis. In taking the variation of the ADM Hamiltonian, the second-derivative terms of the scalar curvature gave rise to two surface integrals. One of these vanishes, but the other is $16\pi\,\mathrm{d}m(g)\cdot h$, which does not vanish. Thus the ADM Hamiltonian does not generate the evolution equations. To get local differential equations from Hamiltonian equations, REGGE and TEITELBOIM thus conclude that the ADM Hamiltonian has to be supplemented by $16\pi m(g)$.

Now we return to the positivity-of-mass question and restrict ourselves to $N = 1$, so that

$$16\,\mathrm{d}\overline{m}(g)\cdot h = \int_{R^3} \mathrm{Ein}\,(g)\cdot h\mu(g)\,.$$

Thus $g \in \mathscr{M}^p_{s,\delta}$ is a critical point of $\overline{m}$ if and only if $\mathrm{Ein}\,(g) = 0$ if and only if g is flat.

The same is true of $\overline{m}$ restricted to $\mathring{\mathscr{M}}^p_{s,\delta}$. However, because we are restricting ourselves to a submanifold, we need a further argument.

7.2. *Proposition.* – *Let* $m : \mathring{\mathscr{M}}^p_{s,\delta} \to \mathbf{R}$ *be* $\overline{m}$ *restricted to* $\mathring{\mathscr{M}}^p_{s,\delta}$. *Then* $g \in \mathring{\mathscr{M}}^p_{s,\delta}$ *is a critical point of* m *if and only if* $\mathrm{Ric}\,(g) = 0$ *if and only if* g *is flat.*

Proof. For a critical point $g \in \mathring{\mathscr{M}}^p_{s,\delta}$ of m, we require

$$\mathrm{d}m(g)\cdot h = 0$$

for all

$$h \in T_g \mathring{\mathscr{M}}^p_{s,\delta} = \{h \in S^p_{s,\delta} | \Delta(\mathrm{tr}\, h) + \delta\,\delta h - h\cdot \mathrm{Ric}\,(g) = 0\}\,.$$

Thus since $\mathrm{d}m(g)\cdot h = \mathrm{d}\overline{m}(g)\cdot h$, we have

$$16\pi\,\mathrm{d}m(g)\cdot h = \int_{\mathring{R}^3} (\delta\,\delta h + \Delta\,\mathrm{tr}\, h)\mu(g) = \int_{\mathring{R}^3} h\cdot \mathrm{Ric}\,(g)\mu(g) = 0$$

for all $h \in T_g \mathring{\mathscr{M}}^p_{s,\delta}$.

For $h \in S^p_{s,\delta}$ we do not have a L_2 orthogonal splitting. However, as in sect. 4, we do have a topological splitting

$$S^p_{s,\delta} = \hat{S}^p_{s,\delta} \oplus \{fg | f \in M^p_{s,\delta}(\mathbf{R}^3)\}$$

and $\hat{S}^p_{s,\delta} = \ker \mathrm{D}R(g)$.

Thus we can split $h \in S^p_{s,\delta}$ as

$$h = \hat{h} + fg\,,$$

where $\hat{h} \in \ker \mathrm{D}R(g)$ and f is the unique solution of

$$\mathrm{D}R(g)\cdot h = \mathrm{D}R(g)\cdot fg = \Delta(\mathrm{tr}\, fg) + \delta\,\delta(fg) - (fg)\cdot \mathrm{Ric}\,(g) = 2\,\Delta f$$

(since $R(g) = 0$).

Thus, if

$$\mathrm{d}m(g)\cdot \hat{h} = \int_{\mathring{R}^3} \hat{h}\cdot \mathrm{Ric}\,(g)\mu(g) = 0$$

for all $\hat{h} \in T_g \mathring{\mathscr{M}}^s_{p,\delta} = \ker \mathrm{D}R(g)$, then for any $h \in S^p_{s,\delta}$ with a f satisfying the above, $\hat{h} = h - fg \in \ker \mathrm{D}R(g)$, and so, since $R(g) = 0$,

$$\mathrm{d}m(g)\cdot \hat{h} = \int_{\mathring{R}^3} (h - fg)\cdot \mathrm{Ric}\,(g)\mu(g) = \int_{\mathring{R}^3} h\cdot \mathrm{Ric}\,(g)\mu(g) = 0\,,$$

for all $h \in S^p_{s,\delta}$. Thus $\mathrm{Ric}\,(g) = 0$. □

By arguing as in [70], sect. 4, we see that the flat metric g_F in proposition 7.2 is isometric to γ by a diffeomorphism $f: \mathbf{R}^3 \to \mathbf{R}^3$ asymptotic to the identity, *i.e.* g_F is in the orbit of γ by $\mathscr{D}_f(\mathbf{R}^3)$.

We want to prove that there is a neighbourhood $\mathscr{U}$ of γ in $\mathscr{M}^p_{s,\delta}$ such that, for all $g \in \mathscr{U}$, $m(g) \geq 0$, and $m(g) = 0$ if and only if $g = f^*\gamma$ for $f \in \mathscr{D}_f(\mathbf{R}^3)$. For this we shall need the second derivative of m.

7.3. *Proposition.* – *The second derivative of* $\overline{m}: \mathscr{M}^p_{s,\delta} \to \mathbf{R}$, $h_1, h_2 \in S^p_{s,\delta}$, *is given by*

$$16\pi\, \mathrm{d}^2\overline{m}(g)(h_1, h_2) =$$
$$= \frac{1}{2}\int_{\mathbf{R}^3} (\nabla h_1 \cdot \nabla h_2 - 2\, \delta_g h_1 \cdot \delta_g h_2 - \mathrm{d}\,\mathrm{tr}\, h_1 \cdot \mathrm{d}\,\mathrm{tr}\, h_2 - \delta h_1 \cdot \mathrm{d}\,\mathrm{tr}\, h_2 - \mathrm{d}\,\mathrm{tr}\, h_1 \cdot \delta h_2)\mu(g) +$$
$$+ \frac{1}{2}\int_{\mathbf{R}^3} \mathrm{Ric}\,(g) \cdot (h_1 \times h_2 + h_2 \times h_1)\mu(g) - \frac{1}{2}\int_{\mathbf{R}^3} \mathrm{Ric}\,(g) \cdot (h_1\, \mathrm{tr}\, h_2 + h_2\, \mathrm{tr}\, h_1)\mu(g) +$$
$$+ \frac{1}{4}\int_{\mathbf{R}^3} R(g)(\mathrm{tr}\, h_1)(\mathrm{tr}\, h_2)\mu(g)\,.$$

Remark. Another useful expression for the second derivative is

$$16\pi\, \mathrm{d}^2\overline{m}(g) \cdot (h_1, h_2) = \frac{1}{2}\int_{\mathbf{R}^3} (h_1 \Delta \bar{h}_2 - h_1 \cdot \alpha_g \delta_g \bar{h}_2 - \mathrm{tr}\, h_1 \cdot \delta_g \delta_g h_2)\mu(g) +$$
$$+ \frac{1}{2}\int_{\mathbf{R}^3} \mathrm{Ric}\,(g) \cdot (h_1 \times h_2 + h_2 \times h_1)\mu(g) - \frac{1}{2}\int_{\mathbf{R}^3} \mathrm{Ric}\,(g) \cdot (h_1\, \mathrm{tr}\, h_2 + h_2\, \mathrm{tr}\, h_1)\mu(g) +$$
$$+ \frac{1}{4}\int_{\mathbf{R}^3} R(g)(\mathrm{tr}\, h_1)(\mathrm{tr}\, h_2)\mu(g)\,,$$

where $\bar{h} = h - \frac{1}{2}(\mathrm{tr}\, h)g$.

Proof. From proposition 7.1,

$$16\pi\, \mathrm{d}\overline{m}(g) \cdot h_1 = \int \langle \mathrm{Ein}\,(g), h_1 \rangle \mu(g)\,.$$

Hence

$$16\pi\, \mathrm{d}^2\overline{m}(g)(h_1, h_2) = \int \langle \mathrm{D}\,\mathrm{Ein}\,(g) \cdot h_2, h_1 \rangle \mu(g) -$$
$$- \int \langle \mathrm{Ein}\,(g), h_1 \times h_2 + h_2 \times h_1 \rangle \mu(g) + \frac{1}{2}\int \langle \mathrm{Ein}\,(g), h_1 \rangle\, \mathrm{tr}\, h_2\, \mu(g)$$

where the middle integral comes from variations of the g's which contract $\mathrm{Ein}\,(g)$ and h

$$\langle \mathrm{Ein}\,(g), h \rangle = g^{ik} g^{jl} (\mathrm{Ein}\,(g))_{ij} h_{kl}$$

and

$$(h_1 \times h_2)_{ij} = (h_1)_{ik} (h_2)^k{}_j$$

is the product of symmetric tensors.

From proposition 4.1 of [1]

$$\mathrm{D}\,\mathrm{Ein}\,(g)\cdot h_2 = \tfrac{1}{2}\big(\Delta_{\mathrm{L}} h_2 - \alpha_g\,\delta_g h_2 - \mathrm{Hess}\,(\mathrm{tr}\,h_2) - \\ - \big(\Delta(\mathrm{tr}\,h_2) + \delta\,\delta h_2\big)g + \tfrac{1}{2}\Big(\big(h_2\cdot\mathrm{Ric}\,(g)\big)g - R(g)\,h_2\Big)\,,$$

where

$$\Delta_{\mathrm{L}} h = \Delta h + R_i{}^k h_{kj} + R_j{}^k h_{ki} - 2R^a{}_{ibj} h_a{}^b$$

is the Lichnerowicz Laplacian and where $\Delta h = -\,g^{ab} h_{ij;a|b}$.

For a 3-dimensional manifold,

$$R^a{}_{ibj} = \delta^a{}_b R_{ij} - \delta^a{}_j R_{ib} + g_{ij} R^a{}_b - g_{ib} R^a{}_j + \tfrac{1}{2} R(g)(\delta^a{}_j g_{ib} - \delta^a{}_b g_{ij})\,.$$

Thus

$$\tfrac{1}{2}\langle h_1, \Delta_{\mathrm{L}} h_2\rangle = \tfrac{1}{2}\langle h_1, \Delta h_2\rangle + \tfrac{1}{2}\,\mathrm{Ric}\,(g)\cdot(h_1\times h_2 + h_2\times h_1) - R^a{}_{ibj}(h_2)_a{}^b(h_1)^{ij} = \\ = \tfrac{1}{2}\langle h_1, \Delta h_2\rangle + \tfrac{1}{2}\,\mathrm{Ric}\,(g)\cdot(h_1\times h_2 + h_2\times h_1) - \\ - \Big((\mathrm{tr}\,h_2)\big(h_1\cdot\mathrm{Ric}\,(g)\big) + (\mathrm{tr}\,h_1)\big(h_2\cdot\mathrm{Ric}\,(g)\big) - \mathrm{Ric}\,(g)\cdot(h_1\times h_2 + h_2\times h_1) + \\ + \tfrac{1}{2} R(g)\big(h_1\cdot h_2 - (\mathrm{tr}\,h_1)\,(\mathrm{tr}\,h_2)\big)\Big) = \tfrac{1}{2}\langle h_1, \Delta h_2\rangle + \tfrac{3}{2}\,\mathrm{Ric}\,(g)\cdot(h_1\times h_2 + h_2\times h_1) - \\ - \mathrm{Ric}\,(g)\cdot\big(h_1(\mathrm{tr}\,h_2) + h_2(\mathrm{tr}\,h_1)\big) - \tfrac{1}{2} R(g)\big(h_1\cdot h_2 - (\mathrm{tr}\,h_1)(\mathrm{tr}\,h_2)\big)\,.$$

Thus

$$\langle h_1, \mathrm{D}\,\mathrm{Ein}\,(g)\cdot h_2\rangle = \\ = \tfrac{1}{2}\big(h_1\cdot\Delta h_2 + h_1\cdot\alpha_g\,\delta_g h_2 - h_1\cdot\mathrm{Hess}\,(\mathrm{tr}\,h_2) - \mathrm{tr}\,h_1(\Delta\,\mathrm{tr}\,h_2 + \delta\,\delta h_2)\big) + \\ + \tfrac{1}{2}\Big(\big(h_2\cdot\mathrm{Ric}\,(g)\big)\,\mathrm{tr}\,h_1 - R(g)\,h_1\cdot h_2\Big) + \tfrac{3}{2}\,\mathrm{Ric}\,(g)\cdot(h_1\times h_2 + h_2\times h_1) - \\ - \mathrm{Ric}\,(g)\cdot\big(h_1(\mathrm{tr}\,h_2) + h_2(\mathrm{tr}\,h_1)\big) - \tfrac{1}{2} R(g)\big(h_1\cdot h_2 - (\mathrm{tr}\,h_1)\,(\mathrm{tr}\,h_2)\big)\,.$$

Also

$$-\,\mathrm{Ein}\,(g)\cdot(h_1\times h_2 + h_2\times h_1) = -\,\mathrm{Ric}\,(g)(h_1\times h_2 + h_2\times h_1) + R(g)\,h_1\cdot h_2$$

and

$$\tfrac{1}{2}\langle\mathrm{Ein}\,(g), h_1\rangle\,\mathrm{tr}\,h_2 = \tfrac{1}{2}\big((\mathrm{tr}\,h_2)\,\mathrm{Ric}\,(g)\cdot h_1 - \tfrac{1}{2} R(g)(\mathrm{tr}\,h_1)(\mathrm{tr}\,h_2)\big)\,.$$

Putting these three terms together gives the pointwise expression

$$\mathrm{D}\big(\langle\mathrm{Ein}\,(g), h_1\rangle\mu(g)\big)\cdot h_2 = \tfrac{1}{2}\,[h_1\cdot\Delta h_2 + h_1\cdot\alpha_g\,\delta_g h_2 - \\ - h_1\cdot\mathrm{Hess}\,(\mathrm{tr}\,h_2) - \mathrm{tr}\,h_1\,\Delta(\mathrm{tr}\,h_2) - (\mathrm{tr}\,h_1)\,\delta\,\delta h_2] + \\ + \tfrac{1}{2}\,\mathrm{Ric}\,(g)\cdot(h_1\times h_2 + h_2\times h_1) - \tfrac{1}{2}\,\mathrm{Ric}\,(g)\cdot\big(h_1(\mathrm{tr}\,h_2) + h_2(\mathrm{tr}\,h_1)\big) + \tfrac{1}{4} R(g)(\mathrm{tr}\,h_1)(\mathrm{tr}\,h_2)\,.$$

Finally, integrating by parts over $\mathbf{R}^3$ gives the proposition. Note that for these second-order terms we may freely integrate by parts because the divergence term can be converted to a surface integral which vanishes at infinity. For example

$$\int h_1 \Delta h_2 \mu(g) = \int_{R^3} \nabla h_1 \cdot \nabla h_2 \mu(g) + \int_{R^3} \delta(h_1 \cdot \nabla h_2) \mu(g)$$

and

$$\int_{R^3} \delta(h_1 \cdot \nabla h_2) \mu(g) = -\oint_{\infty} (h_1)^{ij} (h_2)_{ij|k} \mathrm{d}S^k = 0$$

because $h_1 \cdot \nabla h_2 = O(1/r^3)$.

The form of the second derivative given in the remark follows from

$$\mathrm{D}\,\mathrm{Ein}\,(g) \cdot h = \tfrac{1}{2}\left(\Delta_L \bar{h} - \alpha_\rho \delta_\sigma \bar{h} - (\delta_\rho \delta_\sigma \bar{h}) g\right) + \tfrac{1}{2}\left((\bar{h} \cdot \mathrm{Ric}\,(g)) g - R(g) \bar{h}\right)$$

and nothing that

$$\left((\bar{h} \cdot \mathrm{Ric}\,(g)) g - R(g) \bar{h}\right) = \left((h \cdot \mathrm{Ric}\,(g)) g - R(g) h\right)$$

and

$$\Delta_L \bar{h} = \Delta_L h - \tfrac{1}{2} g \Delta(\mathrm{tr}\, h) = \Delta \bar{h} + R_i{}^k h_{kj} + R_j{}^k h_{ki} - 2R^a{}_{ibj} h_a{}^b ,$$

so that the algebraic terms remain the same. □

The formula simplifies greatly in certain cases of concern. If $\mathrm{Ein}\,(g) = 0$ and $\delta_\rho \bar{h}_2 = 0$, then

$$16\pi\, \mathrm{d}^2\overline{m}(g) \cdot (h_1, h_2) = \tfrac{1}{2} \int_{R^3} h_1 \cdot \Delta \bar{h}_2 \mu(g) = \tfrac{1}{2} \int_{R^3} h_1 \cdot \Delta h_2 \mu(g) - \tfrac{1}{4} \int_{R^3} \mathrm{tr}\, h_1 \Delta(\mathrm{tr}\, h_2) \mu(g) =$$
$$= \tfrac{1}{2} \int_{R^3} \nabla h_1 \cdot \nabla h_2 \mu(g) - \tfrac{1}{4} \int_{R^3} (\mathrm{d}\,\mathrm{tr}\, h_1)(\mathrm{d}\,\mathrm{tr}\, h_2) \mu(g) .$$

If $\mathrm{Ein}\,(g) = 0$, $h_2 = 0$ and $\mathrm{tr}\, h_2 = 0$, then

$$16\pi\, \mathrm{d}^2\overline{m}(g)(h_1, h_2) = \tfrac{1}{2} \int_{R^3} \nabla h_1 \cdot \nabla h_2 \mu(g) = \tfrac{1}{2} \langle h_1, h_2 \rangle_E ,$$

which is just the « energy » inner product. The estimates in sect. 1 show that this is well defined.

Now we wish to factor out the group of co-ordinate transformations $\mathscr{D}^p_{s+1,\delta-1}$ by constructing a slice to this action. We do so by using harmonic co-ordinates.

Let $S = \{g \in \mathscr{M}^p_{s,\delta} | \Gamma^k = g^{ij} \Gamma^k_{ij} = 0\}$, *i.e.* those metrics on $\mathbf{R}^3$ for which the Euclidean co-ordinates are harmonic. We let $\Gamma = \Gamma^k = g^{ij} \Gamma^k_{ij}$, and do not decorate the S with ${}^p_{s,\delta}$.

The next step is to show that S and $S \cap \mathring{\mathscr{M}}^p_{s,\delta}$ are submanifolds and that their tangent spaces are summands in an $\langle , \rangle_E$, orthogonal decomposition of tensors. The argument is similar to that for $\mathscr{C}^p_{s,\delta}$ and $\mathring{\mathscr{M}}^p_{s,\delta}$ given in sect. 3, only considerably more technical.

7.4. *Lemma. – In a neighbourhood of the Euclidean metric γ on $\mathbf{R}^3$, $S \subset \mathscr{M}^p_{s,\delta}$ is a C^∞ submanifold of $\mathscr{M}^p_{s,\delta}$ with tangent space at γ given by the Banach space*

$$T_\gamma S = E_1 = \{h \in S^p_{s,\delta} | \delta_\gamma(h - \tfrac{1}{2}(\mathrm{tr}_\gamma h)\gamma) = 0\} .$$

Proof. We first show that $S^p_{s,\delta}$ splits as a $\langle , \rangle_\gamma$ orthogonal topological sum

$$S^p_{s,\delta} = E_1 \oplus E_2 ,$$

where

$$E_2 = \{(h_2)_{ij} = + \tfrac{1}{2}\Delta_\gamma^{-1}(\partial_i u_j - \partial_j u_i) - \tfrac{1}{2}(\Delta_\gamma^{-1}\partial_l u_l)\delta_{ij} | u_j \in M^p_{s-1,\delta+1}\} \subset S^p_{s,\delta} .$$

Indeed, let $h \in S^p_{s,\delta}$. Let

$$a_j = \partial_i h_{ij} - \tfrac{1}{2}\partial_j h_{ll} \in M^p_{s-1,\delta+1}$$

and let

$$u_j = -2a_j - \tfrac{2}{3}\partial_j \Delta_\gamma^{-1}\partial_i a_i .$$

Note that $\partial_i u_i = -\frac{4}{3}\partial_j a_j$ since $\Delta_\gamma = -\partial_i\partial_i$ and ∂_i and Δ_γ^{-1} commute since we are at flat space. (∇_i and Δ_g^{-1} do not, in general, commute for $g \neq \gamma$.) Define h_2 as above and let $h_1 = h - h_2$. Then $h_1 \in E_1$ since

$$\begin{aligned}
\delta_\gamma h_1 - \tfrac{1}{2}\mathrm{d}\,\mathrm{tr}_\gamma h_1 &= \partial_i h_{ij} - \tfrac{1}{2}\partial_j h_{ii} - \partial_i(h_2)_{ij} + \tfrac{1}{2}\partial_j(h_2)_{ii} = \\
&= \partial_i h_{ij} - \tfrac{1}{2}\partial_j h_{ii} - \tfrac{1}{2}\partial_i\Delta_\gamma^{-1}(\partial_i u_j + \partial_j u_i) + \\
&+ \tfrac{1}{2}\partial_i\Delta_\gamma^{-1}(\partial_l u_l)\delta_{ij} + \tfrac{1}{4}\partial_j\Delta_\gamma^{-1}(\partial_i u_i + \partial_i u_i) - \tfrac{1}{4}\partial_j(\Delta_\gamma^{-1}\partial_l u_l)\delta_{ii} = \\
&= a_j + \tfrac{1}{2}u_j + \tfrac{2}{3}\partial_j\Delta_\gamma^{-1}\partial_i a_i - \tfrac{2}{3}\partial_j\Delta_\gamma^{-1}\partial_i a_i - \\
&- \tfrac{2}{3}\partial_j\Delta_\gamma^{-1}\partial_i a_i + \tfrac{3}{4}\cdot\tfrac{4}{3}\partial_j\Delta_\gamma^{-1}(\partial_i a_i) = a_j + \tfrac{1}{2}u_j + \tfrac{1}{3}\partial_j\Delta_\gamma^{-1}\partial_i a_i = 0
\end{aligned}$$

by definition of u_j.

If $h_1 \in E_1$ and $h_2 \in E_2$,

$$\begin{aligned}
\langle h_1, h_2\rangle_E &= \int \nabla h_1 \cdot \nabla h_2 \,\mathrm{d}^3x = \int h_1 \cdot \Delta_\gamma h_2 \,\mathrm{d}^3x = \\
&= -\int\{-\tfrac{1}{2}h_1 \cdot L_u\gamma + \tfrac{1}{2}\mathrm{tr}_\gamma h_1\,\delta u\}\,\mathrm{d}^3x = -\int(\delta_\gamma h_1 - \tfrac{1}{2}\mathrm{d}\,\mathrm{tr}_\gamma h_1)\cdot u\,\mathrm{d}^3x = 0 .
\end{aligned}$$

Thus E_1 and E_2 are $\langle,\rangle_\gamma$ orthogonal. Thus the splitting is algebraic. Since E_2 is the image of a continuous linear map and E_1 is closed, E_2 is closed and the sum is topological (see, *e.g.*, [76], lemma 4.5).

In a neighborhood of γ any g can thus be uniquely written $g = \gamma + h_1 + h_2$, where $h_i \in E_i$. Consider the mapping ψ of a neighborhood of zero in $E_1 \oplus E_2$ to $\mathscr{X}^p_{s-1,\delta+1}$ (vector fields of type $M^p_{s-1,\delta+1}$) by

$$(h_1, h_2) \mapsto (\gamma + h_1 + h_2)\cdot\Gamma(\gamma + h_1 + h_2)\,.$$

By the multiplication and composition properties of weighted Sobolev spaces, ψ is clearly C^∞. The partial derivative with respect to h_2 at $(0, 0)$ is the map $E_2 \to \mathscr{X}^p_{s-1,\delta-1}$ given by $h \mapsto \delta_\gamma(h - \frac{1}{2}(\operatorname{tr} h)\gamma)$. This map has kernel zero since E_1 and E_2 are $\langle,\rangle_E$ orthogonal and the map is onto. Indeed, given $a_j \in \mathscr{X}^p_{s-1,\delta+1}$, define $u_j = -2a_j - \frac{2}{3}\partial_j\Delta^{-1}\partial_i a_i$ as above and let h_2 be as in the definition of E_2, then $\delta_\gamma(h - \frac{1}{2}(\operatorname{tr} h)\gamma) = a$ as in the above calculation.

Thus, since ψ is a submersion by the implicit function theorem, S is a manifold near γ. $\square$

The same type of reasoning will prove that, near γ, $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$ is a submanifold of $\mathscr{M}^p_{s,\delta}$. However, since we shall need the $\langle,\rangle_g$ orthogonal decomposition for the subspace $T_g(\mathring{\mathscr{M}}^p_{s,\delta} \cap S)$ explicitly, we shall give some additional details.

7.5. *Lemma.* – *There is a neighbourhood V of γ in $\mathscr{M}^p_{s,\delta}$ such that, for $g \in V$, $S^p_{s,\delta}$ admits the $\langle,\rangle_g$ orthogonal decomposition as a topological sum*

$$S^p_{s,\delta} = B_1(g) \oplus B_2(g)\,,$$

where

$$B_1(g) = \{h \in S^p_{s,\delta} | A_g(h) \equiv \Delta_g(\operatorname{tr}_g h) + \\ + \delta_g\delta_g h - \operatorname{Ric}(g)\cdot h = 0 \text{ and } a_g(h) \equiv \delta_g h - \tfrac{1}{2}\,\mathrm{d}\operatorname{tr}_g h - h\cdot\Gamma(g) = 0\}$$

and $B_2(g) = \{h \in S^p_{s,\delta}|$, *there is a* $U \in M^p_{s,\delta}$ *and* $u \in \mathscr{X}^p_{s-1,\delta+1}$ *such that*

$$h = h(U, u) \equiv -gU - \Delta_g^{-1}(\operatorname{Hess}_g U - \operatorname{Ric}(g)U) + \\ + \tfrac{1}{2}\Delta_g^{-1}(L_u g - g\,\delta_g u) + \Delta_g^{-1}(\Gamma(g)\cdot u)\}\,.$$

In fact, $\mathring{\mathscr{M}}^p_{s,\delta} \cap S \cap V$ is a submanifold of $\mathscr{M}^p_{s,\delta}$ with tangent space given by

$$T_g(\mathring{\mathscr{M}}^p_{s,\delta} \cap S) = B_1(g)\,,$$

i.e. $h \in T_g(\mathring{\mathscr{M}}^p_{s,\delta} \cap S)$ *if and only if*

$$\Delta_g(\operatorname{tr}_g h) + \delta\,\delta h - \operatorname{Ric}(g)\cdot h = 0$$

and

$$\delta_g h - \tfrac{1}{2}\,\mathrm{d}(\mathrm{tr}_g\, h) - h\cdot\Gamma(g) = 0\,.$$

Remark. We are working near γ since that is where we shall use the result. However, the proof works on the orbit of γ just as well. At γ itself, $T_\gamma(\mathring{\mathscr{M}}^p_{s,\delta}\cap S)$ consists just of the transverse traceless h's, as is easy to see. Thus, on this space, $32\pi\,\mathrm{d}^2 m(\gamma)(h,h) = \|h\|^2_E = \|\nabla h\|^2_{L^2}$.

Proof. First of all we note that $B_1(g)$ and $B_2(g)$ are $\langle,\rangle_g$ orthogonal. Let $h_1\in B_1(g)$ and $h_2\in B_2(g)$. Then

$$\langle h_1, h_2\rangle_g = \int_{R^3}\nabla_g h_1\cdot\nabla_g h_2\,\mu(g) = \int_{R^3} h_1\cdot\Delta_g\, h_2\mu(g) =$$

$$= -\int_{R^3}\{+\,\mathrm{tr}\, h_1\,\Delta_g U + h_1\cdot\mathrm{Hess}_g\, U - h_1\cdot\mathrm{Ric}\,(g)\,U -$$

$$-\tfrac{1}{2}(h_1\cdot L_u g - \mathrm{tr}_g\, h_1\cdot\delta_g u) - h_1\cdot(\Gamma(g)\cdot u)\}\,\mu(g) = -\int_{R^3}\{A_g(h_1)\,U + a_g(h_1)\cdot u\}\,\mu(g) = 0\,.$$

Now define for $g\in\mathscr{M}^p_{s,\delta}$ the linear mapping

$$A_g : M^p_{s,\delta}\times\mathscr{X}^p_{s-1,\delta+1}\to M^p_{s-2,\delta+2}\times\mathscr{X}^p_{s-1,\delta+1}\,,$$

$$A_g(U,u) = (L_g(U,u), l_g(U,u))\,,$$

where

$$L_g(U,u) = -\Delta_g U - \nabla^j\nabla^i\Delta_g^{-1}(\nabla_i\partial_j U) + \nabla^j\nabla^i(\Delta_g^{-1}R_{ij}\,U) +$$
$$+\,\nabla^j\nabla^i\Delta_g^{-1}(\tfrac{1}{2}\nabla_i u_j + \tfrac{1}{2}\nabla_j u_i - \Gamma^k_{ij}u_k)$$

and

$$(l_g(U,u))_j = -\nabla^i\Delta_g^{-1}(\nabla_i\partial_j U - R_{ij}\,U - \tfrac{1}{2}\nabla_i u_j - \tfrac{1}{2}\nabla_j u_i) +$$
$$+\,\nabla^i\Delta_g^{-1}(\Gamma^k_{ij}u_k) - \tfrac{1}{4}\partial_j\Delta_g^{-1}\nabla_l u^l - (\Gamma\cdot h_2(U,u))_j\,.$$

As in [70], sect. 3, one sees that A_g is a C^∞ function of g with values in the Banach space

$$\mathscr{L}(M^p_{s,\delta}\times\mathscr{X}^p_{s-1,\delta+1},\ M^p_{s-2,\delta+2}\times\mathscr{X}^p_{s-1,\delta+1})$$

of continuous linear maps.

The maps L, l are defined so that

$$A_g(h_2) = L_g(U,u)$$

and

$$a_g(h_2) = l_g(U, u) ,$$

where $h_2 = h_2(U, u)$ is given in the definition of $B_2(g)$. This can be verified by a straightforward calculation.

For $g = \gamma$ we have

$$L_\gamma(U, u) = -2\Delta_\gamma U - \partial_i u_i$$

and

$$(l_\gamma(U, u))_j = \partial_j U - \tfrac{1}{2} u_j + \tfrac{1}{4}\partial_j \Delta_\gamma^{-1} \partial_i u_i .$$

Since $L_\gamma(U, u) = F$ and $l_\gamma(U, u) = f$ have a unique solution, namely

$$U = -\Delta_\gamma^{-1}(\tfrac{3}{2}F - 2\,\delta f)$$

and

$$u_j = 2\{-\partial_j \Delta_\gamma^{-1} F + \partial_j \Delta_\gamma^{-1}\,\delta f - f_j\} ,$$

we see that A_γ is an isomorphism.

Therefore, since the isomorphisms are open, there is a neighbourhood V of γ such that, for $g \in V$, A_g is an isomorphism.

Given $h \in S^p_{s,\delta}$, $g \in V$, let $(U, u) = A_g^{-1}(A_g(h), a_g(h))$ and $h_2 = h_2(U, u) \in B_2(g)$. Then $h_1 = h - h_2$ clearly belongs to $B_1(g)$, since $A_g(h_1) = A_g(h) - A_g(h_2) = L_g(U, u) - A_g(h_2) = 0$ and, similarly, $a_g(h_1) = 0$.

Thus we have an algebraic splitting and, as in lemma 7.4, the splitting is then topological as well.

Finally, we consider the map

$$\Sigma : \mathscr{M}^p_{s,\delta} \to M^p_{s-2,\delta+2} \times \mathscr{R}^p_{s-1,\delta+1} ,$$

$$\Sigma(g) = (R(g), g\cdot\Gamma) .$$

As above, $\Sigma(g)$ is C^∞ and has derivative $D\Sigma(g)$ given by

$$D\Sigma(g) = 0 \qquad \text{on } B_1(g) ,$$

$$D\Sigma(g) = (A_g, a_g) \qquad \text{on } B_2(g) .$$

Therefore, by the above, $D\Sigma(g)$ is an isomorphism on $B_2(g)$ for $g \in V$. The lemma therefore follows. □

This argument also contains the proof that $\mathscr{M}^p_{s,\delta}$ is a manifold near γ (sect. 3). In the sequel $\mathscr{M}^p_{s,\delta} \cap S$ will stand for $\mathscr{M}^p_{s,\delta} \cap S \cap V$.

As in [58], $\mathscr{M}^p_{s,\delta}$ has a smooth connection ∇ induced by $\langle\,,\rangle_g$. From lemma 7.5 we have a smooth orthogonal projection; for $g \in \mathring{\mathscr{M}}^p_{s,\delta} \cap S$

$$P_g : T_g \mathscr{M}^p_{s,\delta} \to T_g(\mathring{\mathscr{M}}^p_{s,\delta} \cap S)\,.$$

Therefore, $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$ has a smooth connection, namely $P \circ \nabla$. Thus,

7.6. *Lemma.* – *The weak Riemannian structure* $\langle\,,\rangle_g$ *on* $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$ *has a smooth connection.*

The next step is to prove that $m \geqslant 0$ locally in $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$. This can be done in one of two ways:

Method 1. The connection guaranteed by lemma 7.6 can be used to join g near γ to γ by a geodesic $g(\tau)$.

By the mean-value theorem,

$$m(g) = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} m(g(\tau))$$

at some intermediate τ. One now has to write out this second derivative explicitly and estimate terms.

Note that $(\mathrm{d}^2/\mathrm{d}\tau^2) m(g(\tau)) = \mathrm{d}^2\overline{m}(g(\tau))\cdot(h, h) + \mathrm{d}\overline{m}(g(\tau))\cdot k$, where $h = g'(\tau)$, $k = g''(\tau)$. This expression can be written out explicitly from our earlier expressions for $\mathrm{d}\overline{m}$ and $\mathrm{d}^2\overline{m}$. The leading term is $\frac{1}{2}\|h\|^2_E = \frac{1}{2}\|\nabla h\|^2_{L^2}$.

A typical remaining term is estimated in this way:

$$\begin{aligned}
\left|\int_{R^3} f^2 R_{ij}\,\mathrm{d}^3x\right| &\leqslant \|f\|^2_{L_6}\|\mathrm{Ric}\,(g)\|_{L^3_2} \leqslant && \text{(Hölder)}\\
&\leqslant C\|f\|^2_E\|\mathrm{Ric}\,(g)\|_{L^3_2} \leqslant && \text{(see sect. 1)}\\
&\leqslant C\|f\|^2_E\|g-\gamma\|_{\mathscr{M}^p_{s,\delta}}\,.
\end{aligned}$$

The last estimate comes by writing out Ric (g) explicitly. Since $g(\tau)$ is a geodesic, $k = g''(\tau)$ can be expressed in terms of h. Putting all estimates of this sort together, one gets

$$|m''(g(\tau)) - \tfrac{1}{2}\|h\|^2_E| \leqslant C\|g-\gamma\|_{\mathscr{M}^p_{s,\delta}}\|h\|^2_E\,,$$

from which we get, if $\|g-\gamma\|_{\mathscr{M}^p_{s,\delta}}$ is small,

$$m(g) \geqslant \bar{C}\|h\|^2_E \geqslant \tilde{\bar{C}}\|g-\gamma\|^2_E\,,$$

so

$$m(g) > 0 \qquad \text{if } g \in \mathring{\mathscr{M}}^p_{s,\delta} \cap S\,,\ g \neq \gamma\,.$$

Method 2. The second method uses the Morse lemma on Banach manifolds.

Notice that $-m$ has a smooth $\langle , \rangle_g$ gradient on $\mathscr{M}^p_{s,\delta}$, namely $\Delta^{-1}\operatorname{Ric}(g)$. By projecting this to $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$ by the projection constructed above, we establish that $-m$ has a smooth $\langle , \rangle_g$ gradient on $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$. (This requires the proofs above, since in general a C^∞ function need not have a smooth gradient relative to a weak metric.) This remark together with lemma 4 enables the following to be applied:

7.7. *Lemma.* – *Let M be a Banach manifold and $\langle , \rangle$ a weak Riemannian structure on M which has a smooth connection. Let $f: M \to \mathbf{R}$ be C^2 and let Y be the $\langle , \rangle$ gradient of f; we assume Y exist and is C^∞ vector field on M. Assume that x_0 is a critical point of f and that $\mathrm{D}Y(x_0): T_{x_0}M \to T_{x_0}M$ is an isomorphism. Then there exists a co-ordinate chart about x_0 in which*

$$f(x) = f(x_0) + \tfrac{1}{2}\,\mathrm{d}^2 f(x_0)\cdot(x - x_0, x - x_0)\,.$$

This result is due to TROMBA [25]. We present a self-contained simple proof following the method of Moser-Weinstein (see [77]).

Proof. We work in an exponential chart (normal co-ordinates) and assume $x_0 = 0$ and $f(x_0) = 0$. Now join the one-forms $\mathrm{d}f = \omega_1$ and ω_2, defined by $\omega_2(x)\cdot h = \langle x, \mathrm{D}Y(0)\cdot h\rangle_0 = \langle \mathrm{D}Y(0)\cdot x, h\rangle_0$, by a straight line: $\omega_t = t\omega_1 + (1-t)\omega_2$. Note $\omega_2 = \mathrm{d}\varphi$, $\varphi(x) = \frac{1}{2}\langle x, \mathrm{D}Y(0)\cdot x\rangle_0$ and find a vector field Z_t (with $Z_t(0) = 0$) such that

$$i_{Z_t}\omega_t + (f - \varphi) = 0\,.$$

That there is such a Z_t can be seen by writing

$$\langle Y(x), Z_t(x)\rangle_x = \int_0^t \langle \tau\nabla_x \gamma(sx), \tau Z_t(x)\rangle_0\,\mathrm{d}s = \int_0^1 \langle x, \tau\nabla_{Z_t(x)} Y(sx)\rangle_0\,\mathrm{d}s\,,$$

where τ denotes parallel translation to zero, and by using invertibility of $\mathrm{D}Y$ near 0 and smoothness of the connection. If F_t is the flow of Z_t, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(F_t^*\omega_t) = F_t^*\{\mathrm{d}i_{Z_t}\omega_t + i_{Z_t}\mathrm{d}\omega_t + \omega_1 - \omega_2\} = F_t^*\,\mathrm{d}\{i_{Z_t}\omega_t + (f - \varphi)\} = 0\,,$$

so $F_1^*\omega_1 = \omega_2$. Then F_1 gives, near 0, the required co-ordinate change. □

In particular, if $\mathrm{d}^2 f(x_0) = \langle , \rangle_{x_0}$ (*i.e.* $\mathrm{D}Y(x_0) = \text{identity}$), then $f(x_0)$ is a strict local minimum of f (on this neighbourhood). However, we cannot conclude that this neighbourhood contains a ball in the $\langle , \rangle$ norm. Thus we cannot conclude that there is an $\varepsilon > 0$ such that f increases to value ε as x moves away from x_0. To do so would require the hypothesis that the Lipschitz constant

of Y varies continuously as x varies in the $\langle , \rangle$ topology. This point is one of the crucial difficulties in the global theorem, as we shall see in sect. **8**.

So far we have shown that m is positive in $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$. The final step, to complete the proof, consists of showing that S really is a slice. That is, any $g \in \mathring{\mathscr{M}}^p_{s,\delta}$ can be brought into S by a co-ordinate transformation φ. This transformation leaves $\mathring{\mathscr{M}}^p_{s,\delta}$ invariant and m unchanged, so would prove $m > 0$ from positivity on $\mathring{\mathscr{M}}^p_{s,\delta} \cap S$.

Let us summarize:

7.8. *Lemma.* – *There exists an $\varepsilon > 0$ such that, for $g \in \mathring{\mathscr{M}}^p_{s,\delta} \cap S$, $\|g - \gamma\|_{M^p_{s,\delta}} < \varepsilon$ and $g \neq \gamma$, $m(g) > 0$. If fact, $m(g) \geqslant C\|g - \gamma\|^2_E$ for a constant $C > 0$.*

7.9. *Lemma.* – If ε is sufficiently small and $\|g - \gamma\|_{M^p_{s,\delta}} < \varepsilon$, then $\varphi^* g \in S$ for some $\varphi \in \overline{\mathscr{D}}^p_{s+1,\delta-1}$.

Proof. Let, in Euclidean co-ordinates, φ have components $x^i + f^i(x^j)$. Then $\varphi^* g \in S$ if $\Delta_g \varphi^i = 0$, *i.e.* $\Delta_g f^i = \Gamma^i_{jk} g^{jk}$. Now $\Gamma^i_{jk} g^{jk}$ is not in $M^p_{s-2,\delta+2}$ necessarily, so we cannot yet apply Δ_g^{-1}. However, we can differentiate, letting F^i be the differential of f^i, so F^i is a one-form, and denoting by $\tilde{\Delta}_g$ the Laplace-DeRham operator on forms, to obtain

$$\tilde{\Delta}_g F^i = \mathrm{d}H^i ,$$

where $H^i = \Gamma^i_{jk} g^{jk}$.

Remark. φ is a harmonic map from $(\mathbf{R}^3, g)$ to $(\mathbf{R}^3, \gamma)$ so general invariant formulae are available [78]. However, it is just as easy to proceed directly in this case. Since $\tilde{\Delta}_g$ is an isomorphism of $M^p_{s,\delta}$ to $M^p_{s-2,\delta+2}$, there is a unique solution $F^i \in M^p_{s,\delta}$. However, since $s > 3$ we can assert

$$\tilde{\Delta}_g \mathrm{d}F^i = \mathrm{d}\,\tilde{\Delta}_g F^i = \mathrm{d}(\mathrm{d}H^i) = 0 ,$$

so $\mathrm{d}F^i = 0$. Thus $F^i = \mathrm{d}f^i$ for some f^i. (Explicitly, we can choose

$$f^i(x) = \int_0^1 F^i(tx) \cdot x \, \mathrm{d}t$$

from the proof of the Poincaré lemma.) Since $F^i \in M^p_{s,\delta}$, we see that $\mathrm{d}f^i \in M^p_{s,\delta}$, so, for $g^{jk}\Gamma^i_{jk}$ small in $M^p_{s-1,\delta+1}$, $\mathrm{d}f^i$ will be small in $M^p_{s,\delta}$ so φ will be a C^1 diffeomorphism. Thus $\varphi \in \overline{\mathscr{D}}^p_{s+1,\delta-1}$. □

Putting lemmas 7.8 and 7.9 together gives the main local positivity-of-mass theorem.

7.10. *Theorem.* – *Let γ be the Euclidean metric on $\mathbf{R}^3$. Then there exists a $M^p_{s,\delta}$ neighbourhood $\mathscr{U} \subset \mathring{\mathscr{M}}^p_{s,\delta}$ of γ such that, for $g \in \mathscr{U}$, $m(g) \geqslant 0$. If $m(g) = 0$, then g is flat and in fact $g = f^* \gamma$, where $f \in \overline{\mathscr{D}}^p_{s+1,\delta-1}(\mathbf{R}^3)$.*

Remark. The fact that $m(g) = 0$ implies g is flat is analogous to an isolation theorem in geometry for the compact case, except the mass case is more delicate. The geometry theorem referred to states that, if g_F is a flat metric on a compact manifold and if g is a metric near g_F with $R(g) \geq 0$, then g itself is flat. In fact, the proof we have given here uses a second variation argument and the construction of a slice in a manner similar to that given in [70].

Now we come to a space-time version of 7.10.

7.11. *Theorem.* – *There exists a* $M^p_{s,\delta}$ *neighbourhood* $\mathscr{U}$ *of the Minkowski metric* η *such that if* ${}^{(4)}g \in \mathscr{U}$ *satisfies the weak and strong energy conditions,* $\mathrm{Ric}\,({}^{(4)}g) \geq 0$ *and* $\mathrm{Ein}\,({}^{(4)}g) \geq 0$, *then on any asymptotically flat spacelike hypersurface* Σ

$$m(g_\Sigma) \geq 0 .$$

If $m(g_\Sigma) = 0$, *then* ${}^{(4)}g$ *is isometric to flat space* η.

Remark. We are assuming implicitly that $G_{\mu\nu} = 8\pi T_{\mu\nu}$, so the mass is hypersurface independent (*), and that ${}^{(4)}g$ is uniquely determined by suitable Cauchy data in a maximal development (the Cauchy data would include non-gravitational fields as well as g_Σ, k_Σ).

Proof. From sect. 2, there is a maximal hypersurface near $\{0\}$. Since m is hypersurface independent, we can therefore assume that Σ is maximal and so a whole neighbourhood of Σ is filled with a slicing by such hypersurfaces. Write $g = g_\Sigma$.

By theorem 6.3, there is a vacuum initial data set $\bar{g} = \varphi^4 g$, $\bar{k} = 0$ with $m(\bar{g}) < m(g)$. Now $(\bar{g}, 0)$ is in a neighbourhood of $(\gamma, 0)$, so by theorem 7.10 $m(\bar{g}) \geq 0$. Thus $m(g) \geq 0$. Next assume $m(g) = 0$, so $m(g) = m(\bar{g}) = 0$. Thus, from theorem 7.10 again, $\bar{g} = f^*\gamma$. Thus g is conformally flat.

From the equation

$$m(\bar{g}) = m(g) + \frac{1}{2\pi}\int_{R^3} \Delta_g \varphi\, \mu(g) = m(g) - \frac{1}{2\pi}\int_{R^3} \varphi\, R(g)\mu(g)$$

and $R(g) = k\cdot k + 2\,{}^{(4)}G_{\perp\perp} \geq 0$, $\varphi > 0$, we find $R(g) = 0$, $k = 0$, ${}^{(4)}G_{\perp\perp} = 0$. Hence the equation for φ, namely $\Delta_g \varphi = -R(g)\varphi$ gives $\varphi = 1$. Thus g is flat and $k = 0$. The result therefore follows. □

Remarks.

1) One can alternatively work, as in [24], in the space of all g's and π's and avoid the use of theorem 6.3. However, the present method seems a little simpler.

(*) The hypersurfaces here must be asymptotic to a standard $t =$ const hypersurface.

2) As a special case of the above arguments, we note the following: if Σ is a maximal asymptotically flat spacelike hypersurface in Minkowski space, then g_Σ is flat and $k_\Sigma = 0$.

3) $\mathring{\mathscr{M}}^p_{s,\delta}$ is globally a manifold. Since $\mathscr{M}$ has a critical point only at flat space, it is reasonable to conjecture that $\mathring{\mathscr{M}}^p_{s,\delta}$ is topologically trivial and has a global slice.

8. – Discussion of the global problem.

Now we make a few remarks concerning the important question: is $m \geqslant 0$ on all of $\mathring{\mathscr{M}}^p_{s,\delta}$, or its component containing γ? The local positivity and the fact that any critical point of M is flat is very suggestive that $m \geqslant 0$. As we pointed out in the introduction, this is not a proof, however.

There are some simple sufficient conditions for global positivity of a function on a manifold. Although they do not seem to directly apply to m, they may shed some light on the difficulties and emphasise the depth and importance of the work of SCHOEN and YAU [52].

Let us grant that we have a global slice, or that $\mathring{\mathscr{M}}/\mathscr{D}$ is a manifold, so we have a well-defined space to work on. In this context, the following elementary theorem seems to be suggestive:

8.1. *Theorem.* – *Let M be a connected Hilbert manifold and $m: M \to \mathbf{R}$ a C^∞ function with a single critical point at x_0, where $\mathrm{d}^2 m(x_0)$ is positive definite. Let Y be the gradient of $-m$ and assume has a complete flow and there is, for any neighbourhood U of x_0, an $\varepsilon > 0$ such that $|Y(x)| > \varepsilon$ outside U. Then x_0 is a global minimum of m.*

Proof. Let $A = \{x \in M \mid \text{the } Y\text{-trajectory starting at } x, \text{ say } x(t) \text{ tends to } x_0 \text{ as } t \to +\infty\}$. Along these trajectories m is decreasing since

$$\frac{\mathrm{d}}{\mathrm{d}t} m(x(t)) = \langle \mathrm{d}m(x(t)), x'(t) \rangle = -\langle Y(x(t)), Y(x(t)) \rangle \leqslant 0 .$$

Thus $m(x(t)) \geqslant m(x_0)$. Therefore, it suffices to show $A = M$.

We show that A is both open and closed. That A is open follows from two simple facts:

i) any trajectory of Y which enters a small neighbourhood of x_0 converges to x_0 as $t \to +\infty$ since $\mathrm{d}^2 m(x_0) > 0$,

ii) the solution curves are continuous functions of the initial data (for fixed t).

To show that A is closed, let $x_n \in A$, $x_n \to x$. Let $x_n(t)$, $x(t)$ be the Y-trajectories starting at x_n and x. Let U be a neighbourhood of x_0 inside which trajectories $\to x_0$. Outside U,

$$\frac{d}{dt} m(x_n(t)) < -\varepsilon ,$$

so m decreases at a fixed rate. Thus, after a fixed time T, $x_n(t)$ must enter U, namely after $T = m/\varepsilon$, where $m = \sup_n m(x_n)$. By continuous dependence on initial data $x(t)$ then enters U after time T, so $x \in A$. □

In the mass problem one can find a Y such that $-\,dm \cdot Y > 0$ away from flat space, such as

$$-\,Y(g)_{ij} = \Delta^{-1}R_{ij} - \tfrac{1}{2} g_{ij}\Delta^{-1}(R_{hk}\Delta^{-1}R^{hk} - \nabla_h\nabla_k\Delta^{-1}R^{hk}) .$$

Here, $-\,dm \cdot Y = \langle \Delta^{-1}\operatorname{Ric}(g), \Delta^{-1}\operatorname{Ric}(g)\rangle_g > 0$ if $\operatorname{Ric}(g) \neq 0$.

This Y is a smooth vector field in the $M^p_{s,\delta}$ topology on $\mathscr{M}^p_{s,\delta}$. There is a neighbourhood U of the orbit of flat space $\mathcal{C}_\gamma$ in which integral curves of Y converge to $\mathcal{C}_\gamma$. The problem, then, from the above theorem is:

First sufficient condition. – If $-\,dm \cdot Y \geqslant \varepsilon > 0$ uniformly outside U, then m is globally positive.

One can contemplate more sophisticated methods, such as a «minimax principle» [79].

The real difficulty seems to be the following, stated somewhat loosely.

Second sufficient condition. – If there is a $\langle , \rangle_E$ neighbourhood U of $\mathcal{C}_\gamma$ inside which trajectories of Y converge to $\mathcal{C}_\gamma$ as $t \to \infty$ and on the boundary of which $m \geqslant \varepsilon > 0$, then m is globally positive.

Actually, WINICOUR [80] has pointed out a similar fact as a crucial issue. However, the proof of local positivity does not establish this sufficient condition because of the difference between the $M^p_{s,\delta}$ topology and the $\langle , \rangle_E$ topology. It is conceivable, however, that the argument giving local positivity can be strengthened.

To back up the second sufficient condition, we present, for completeness, the minimax principle we have in mind. We thank A. TROMBA for discussions on this result.

8.2. *Theorem. – Let M be a connected Banach manifold and $\langle , \rangle$ a weak Riemannian structure with a smooth connection. Let $\|\cdot\|$ and $d(\cdot,\cdot)$ denote the norm and distance in this structure. Let $f: M \to \mathbf{R}$ be C^2 bounded below and satisfy*

i) *f has exactly one critical point at x_0; $f(x_0) = 0$, $df(x_0) = 0$;*

ii) *there is a C^∞ $\langle , \rangle$ gradient Y for f;*

iii) *there is an* $\varepsilon > 0$ *such that, if* $d(x, x_0) < \varepsilon$,

$$cd(x, x_0)^2 \leq f(x) \leq Cd(x, x_0)^2 , \qquad c, C > 0;$$

iv) *if* $\|Y(x_n)\| \to 0$, *then* $x_n \to x_0$ *in* $d(\cdot, \cdot)$ (« *condition* C »);

v) *if* $x(t)$ *is an integral curve of* $-Y$ *which is d-Cauchy as* $t \to \beta$, *then* $x(t)$ *is t-extendible beyond* β;

vi) *there is an* $\varepsilon > 0$ *such that* $d(x, x_0) < \varepsilon$ *implies that the integral curve of* $-Y$ *starting at* x *converges* (M) *to* x_0 *as* $t \to +\infty$.

Then $f(x) > 0$ *for all* $x \in M$, $x \neq x_0$.

Remarks. In the example discussed earlier (p. 399), it is condition v) which fails.

For relativity, one is to imagine $M = \mathring{\mathcal{M}}^p_{s,\delta}/\mathcal{D}$ and $x_0 = \mathcal{C}_\gamma$, $f = m$, $\langle , \rangle$ the energy inner product and Y as given above. Condition iii) appears hard, while iv) and v) seem possible to verify.

Proof of 8.2. Assume there is a $y_0 \in M$ with $f(y_0) < 0$, $y_0 \neq x_0$. Let, according to the minimax method,

$$c = \inf \{ \sup_{t \in [0,1]} f(\sigma(t)) | \sigma \text{ is a } C^0 \text{ curve joining } y_0 \text{ to } x_0 \} .$$

By iii), $c > 0$. Choose $0 < \delta < c$. Let

$$N_1 = f^{-1}(-\infty, c - \delta) \quad \text{and} \quad N_2 = f^{-1}(-\infty, c + \delta)$$

and

$$N = N_2 \setminus N_1 = f^{-1}(c - \delta, c + \delta) .$$

Lemma. – N_1 *is a deformation retract of* N_2, *i.e. there is a* C^0 *map* $H : [0, 1] \times \times N_2 \to N_2$ *such that* $H(0, x) = x$, $H(1, x) \in N_1$ *and* $H(s, x) = x$ *if* $x \in N_1$.

Proof. Let $H(s, x) = x$ if $x \in N_1$. If $x \in N_1$, let $x(t)$ be the Y-trajectory starting at x, with domain $[0, \beta[$, $\beta > 0$ chosen maximally.

Suppose $x(t)$ never meets N_1, and $\beta < \infty$. Then from

$$\frac{d}{dt} f(x(t)) = - \| Y(t) \|^2$$

we get, by integration and the Schwarz inequality,

$$f(x(t)) - f(x) \leq - \int_0^t \| x'(\tau) \| \, d\tau_\beta / \sqrt{\beta} ,$$

so

$$\int_0^t \|x'(\tau)\| \, d\tau \leq \sqrt{\beta}(2\delta)$$

and

$$d(x(t), x(s)) \leq \sqrt{|s-t|} \cdot 2\delta .$$

By v) we can continue beyond β.

Thus, if $x(t)$ does not meet N_1, $\beta = +\infty$. However, $\varphi(t) = f(x(t))$ is decreasing, bounded below, so $\varphi'(t) \to 0$, *i.e.* $\|Y(x(t))\| \to 0$. Thus $x(t) \to x_0$ in d. Using vi) we get a contradiction.

Thus $x(t)$ meets N_1. Let $\Psi(x) < \infty$ be the first time it does so and set $H(s, x) = x(s\Psi(x))$. This gives the required H. □

To complete the proof let $\sigma(t)$ be a curve joining y_0 to x_0 in N_2 and let $\varrho(t) = H(1, \sigma(t))$. Since $y_0, x_0 \in N_1$, this is a curve joining y_0, x_0 in N_1 which contradicts the definition of c. □

A simple scaling argument shows that if m is anywhere negative, then it is unbounded below. For this reason, theorem 8.2 is not useful as it stands.

9. – The mass function as a Liapunov function.

The fact that m is positive and conserved leads one to try using m as a Liapunov function for the Einstein evolution equations, *i.e.* to use it to obtain *a priori* bounds on the solution in a convenient norm. One can discuss dynamical stability of solutions to the nonlinear evolution equations.

First of all, we recall the classical role of Liapunov functions in the following

9.1. *Theorem.* – *Let* E *be a Banach space and* F_t *a local flow on* E *with* 0 *a fixed point. Suppose that for any bounded set* $B \subset \mathsf{E}$ *there is an* $\varepsilon > 0$ *such that integral curves beginning in* B *exist for a time interval* $\geq \varepsilon$.

Let $H : \mathsf{E} \to R$ *be a smooth function invariant under the flow.*

a) If $H(u) \geq C\|u\|^2$, *for some* $C > 0$, *then the flow is complete.*

b) If $H(0) = 0$, $DH(0) = 0$ *and* $D^2H(0)$ *is positive or negative definite, then there is a neighbourhood* U *of* 0 *such that any integral curve starting in* U *is defined for all* t; *moreover,* 0 *is dynamically stable.*

Proof.

a) Let $u \in \mathsf{E}$. Since H is conserved we have the *a priori* estimate $\|u\|^2 \leq$ constant, so u remains in a bounded set B. But, because of the assumption on the flow, the integral curve beginning at u can be indefinitely extended.

b) From the assumptions, there are positive constants α, β such that

$$\alpha \|u\|^2 \leqslant |D^2H(0)\cdot(u, u)| \leqslant \beta \|u\|^2 .$$

Hence, by Taylor's theorem, in a small neighbourhood U_0 of 0, we have

$$\gamma \|u\|^2 \leqslant |H(u)| \leqslant \delta \|u\|^2 .$$

Because H is conserved, this shows that there are neighbourhoods U and V of 0 such that, if $u \in U$, it remains in V as long as it is defined. Hence we have completeness as in *a*). Since V can be arbitrarily small, we also have stability. □

This theorem may be used to prove the existence for all time of solutions to certain semilinear nonlinear wave equations (see [81]).

In relativity, the difficulties center around the «strong» nonlinearities in the Einstein equations (they are quasi-linear) and the fact that the second derivative of m is only weakly positive definite; the same difficulty we encountered while proving $m \geqslant 0$. Similar difficulties occur in elasticity [82].

However, under very limited circumstances, one can pass from the linearized stability to the full nonlinear stability [83]. This general idea was first suggested by DESER in an attempt to find out if there are any vacuum solutions of Einstein's equations which are nontrivial and noncollapsing.

We state the criterion somewhat loosely. At this stage we are only concerned with the idea. The details can be nailed down when more is known about the possible examples.

9.2. *Nonsingularity criterion.* – *Suppose there is a curve* $(g(\varrho), \pi(\varrho))$ *of solutions of the (vacuum) constraint equations on* $\mathbf{R}^3$ *with* $g(0) = \gamma$, $\pi(0) = 0$, *and with tangent pointing in a nontrivial direction, i.e.* $h^{rr} \neq 0$ *where* $h = dg/d\varrho$ *at* $\varrho = 0$. *Assume that each* $g(\varrho)$, $\pi(\varrho)$ *has a global co-ordinate system which is harmonic and in which* $g(\varrho)$ *and* $\pi(\varrho)$ *depend on only one of the co-ordinates.* (*In particular, then,* $g(\varrho)$, $\pi(\varrho)$ *have two Killing fields.*) *Then for* ϱ *sufficiently small, the Cauchy development of* $g(\varrho)$, $\pi(\varrho)$ *is a geodesically complete space-time which is* C^0 *close to Minkowski space.*

Initial data sets like those desired here may or may not be obtainable. For example, if (g, π) is chosen spherically symmetric (so the one variable is r, the radius), Birkhoff's theorem will force the space-time to be Schwarzschild. However, from work of Marder [84] it appears that asymptotically flat initial data with toroidal symmetry in $\mathbf{R}^3$ may be possible (it is not clear that his solutions are everywhere regular; one could use his solutions for $t > 0$ as an initial data set and apply the above criterion to generate a singularity-free space-time) (*).

(*) Marder's $t > 0$ solution corresponds to the entire implosion, explosion of the Weber-Wheeler cylindrically symmetric wave bent into a toroidal shape.

An idea for the proof of the criterion 9.2 is as follows: by the proof of theorem 6.3, by using harmonic co-ordinates for the evolution equations, there is a neighbourhood of $(\gamma, 0)$ in which m gives a bound in the $H^1 \times L_2$ norm of (g, π). Choose ϱ small enough so $g(\varrho), \pi(\varrho)$ lies in this neighbourhood. For quasi-linear second-order hyperbolic evolution equations (such as Einstein's equations in harmonic co-ordinates) which depend on one space variable and for which the coefficients of the second-order terms do not involve derivatives of the unknown, an *a priori* bound on the $H^1 \times L_2$ norm guarantees solutions global in time (see [85]) (*). Since $C^0 \supset H^1$ in one dimension, the solution is C^0 close to Minkowski space, since it starts off that way. The higher H^s norms, however, may be unbounded (but they cannot blow up in a finite time).

Geodesic completeness follows from the uniform closeness to the Minkowski metric and the preservation of 4-lengths and angles by geodesics.

10. – The mass function as the generator of time translation.

In [1] we described the space of gravitational degrees of freedom. This space regards as equivalent all (g, π)'s which occur on slices of the same space-time. In the case of compact hypersurfaces, this process divides out all the dynamics. However, for the asymptotically flat case the structure is much richer.

We will describe this structure very briefly in this section. As we pointed out in the introduction, the process described here was suggested by WALKER with the ultimate goal of linking the ADM mass to the past limit of the Bondi mass.

Consider the space of all possible g's and π's with $1/r$ and $1/r^2$ fall-off, respectively. (We dispense with the formal $M^p_{s,\delta}$ spaces here for the sake of exposition.)

We recall from sect. 1 that there is a difficulty if we try to take the orbit of (g, π) under the group of co-ordinate transformations differing from the identity by terms of $O(1)$ at ∞; this orbit is not in general a manifold and the decomposition theorems fail.

What we have to do to fix this problem is to demand tighter asymptotic behaviour at infinity: namely, we must restrict to co-ordinate transformations F which differ from the identity by terms $O(1/r)$ at infinity. The orbit $\hat{\mathcal{O}}_{(g,\pi)}$ under these co-ordinate transformations will be a manifold and will consist of $(\bar{g}, \bar{\pi})$'s which differ from (g, π) by terms of order $1/r^2$, $1/r^3$ at ∞.

More formally, we write $(\bar{g}, \bar{\pi}) \sim (g, \pi)$ if $(\bar{g}, \bar{\pi})$ and (g, π) are solutions of the constraint equations on $\mathbf{R}^3$ with $(1/r, 1/r^2)$ fall-off at infinity and if there is an empty-space asymptotically flat space-time $(V_4, {}^{(4)}g)$ and two asympto-

(*) The evolution in harmonic co-ordinates does not preserve $\operatorname{tr}\pi = 0$, so one would have to interplay between harmonic co-ordinates and $\operatorname{tr}\pi = 0$ co-ordinates.

tically flat spacelike embeddings

$$i : \boldsymbol{R}^3 \to V_4 , \qquad \bar{i} : \boldsymbol{R}^3 \to V_4$$

which induce, respectively, the data (g, π) and $(\bar{g}, \bar{\pi})$ and which are asymptotically identical, *i.e.*

$$i(x) - \bar{i}(x) = O(1/r) \qquad \text{as } r \to \infty$$

(see fig. 1).

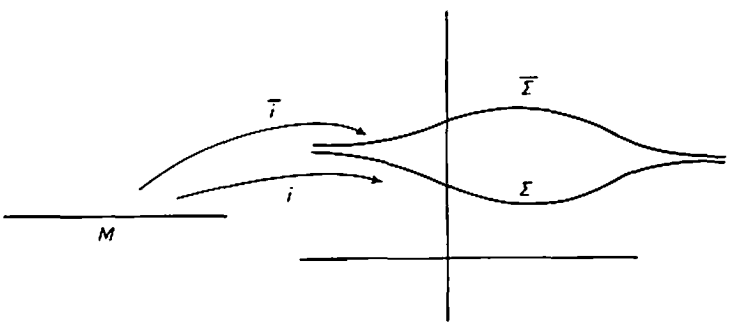

Fig. 1. – Two asymptotically flat spacelike embeddings which are asymptotically identical.

Using this notion of equivalence, we form the quotient, the space of gravitational degrees of freedom,

$$\mathscr{G}_{\text{dyn}} = \mathscr{C}_{\mathscr{H}} \cap \mathscr{C}_{\delta} / \sim$$

(as in sect. 7 of [1]), and show that it is a smooth symplectic manifold. Note that the extra fall-off at infinity is crucial in order for this space to even have a formal tangent space, since this depends on the splitting theorems.

Notice that now *not* all (g, π)'s from the same space-time are identified. In fact (cf. [11]), there is still dynamics on $\mathscr{G}_{\text{dyn}}$. Indeed, as REGGE and TEITELBOIM [12] have pointed out, diffeomorphisms which are asymptotically a Poincaré transformation, denoted $\mathscr{D}_{\text{P}}(\boldsymbol{R}^4)$, are allowed as permissible deformations of the hypersurface (*), and hence act nontrivially on $\mathscr{G}_{\text{dyn}}$. Note that an asymptotic Poincaré transformation composed with a transformation that is asymptotically the identity plus terms $O(1/r)$ has the same effect on $\mathscr{G}_{\text{dyn}}$ as the asymptotic Poincaré transformation itself. Thus, in reality, it is Poin-

(*) There is a problem with « boosts » because the maximal development of (g, π) conceivably might not be large enough for it to be defined. We ignore this here.

caré transformations at infinity, or the Poincaré group

$$P = \mathscr{D}_{\mathrm{P}}(\mathbf{R}^4)/\mathscr{D}_{\mathrm{I}}(\mathbf{R}^4)$$

that acts on $\mathscr{G}_{\mathrm{dyn}}$, where $\mathscr{D}_{\mathrm{I}}(\mathbf{R}^4)$ represents diffeomorphisms that are asymptotically the identity.

If we use the group $\mathscr{D}_{\mathrm{P}}(\mathbf{R}^4)$, the lapse and shift functions $N^\mu = (N, X)$ have the asymptotic behaviour

$$N^\mu \sim \alpha^\mu + \beta^\mu{}_a x^a + O\left(\frac{1}{r}\right),$$

where $(\alpha^\mu, \beta^\mu{}_\nu)$ represent a Poincaré transformation of $\mathbf{R}^4$. As REGGE and TEITELBOIM [12] show, supertranslations can be omitted. Also, and which is clear from the Dirac theory (see [86]), the generator of the dynamical equations is

$$G_{\mathrm{RT}} = G_{\mathrm{ADM}} - \alpha^\mu P_\mu + \tfrac{1}{2}\beta^{\mu\nu} J_{\mu\nu},$$

where P_μ and $J_{\mu\nu}$ are defined in sect. **5**. Thus a transformation $\psi \in \mathscr{D}_{\mathrm{P}}(\mathbf{R}^4)$ induces a canonical transformation on the (g, π)'s. Thus, from sect. **6** of [1], ψ also induces a canonical transformation on $\mathscr{G}_{\mathrm{dyn}}$. Consequently, the Poincaré group itself $P = \mathscr{D}_{\mathrm{P}}(\mathbf{R}^4)/\mathscr{D}_{\mathrm{I}}(\mathbf{R}^4)$ acts on $\mathscr{G}_{\mathrm{dyn}}$ by symplectic transformations. It would be of interest to link these ideas more firmly with recent work on spatial infinity, especially that of HANSEN and ASHTEKAR [87].

We can summarize the situation as follows:

10.1. *Theorem.* – *The generators of the Poincaré group P of symplectic transformations on $\mathscr{G}_{\mathrm{dyn}}$ are the momenta $P^\mu : \mathscr{G}_{\mathrm{dyn}} \to \mathbf{R}^4$ and the angular momenta $J^{\mu\nu} : \mathscr{G}_{\mathrm{dyn}} \to \Lambda^2(\mathbf{R}^4)$ (the antisymmetric 2-tensors on $\mathbf{R}^4$), where P^μ and $J^{\mu\nu}$ are defined in sect.* **5**.

In symplectic language, P^μ and $J^{\mu\nu}$ are the conserved moments of the action of the Poincaré group P and the equivariance of its moment expresses the fact that P^μ and $J^{\mu\nu}$ transform as tensors under asymptotic Lorentz transformations and are invariant under infinitesimal co-ordinate transformations asymptotic to the identity (gauge transformations).

As a special case of this result, the mass $m = P^0$ is the generator of time translations in $\mathscr{G}_{\mathrm{dyn}}$. The proof of this special case is implicit in proposition 7.1, where following REGGE and TEITELBOIM we have shown that H_{ADM} has to be supplemented by $16\pi m(g)$ to generate the Einstein equations.

Put another way, theorem 10.1 says G_{RT} generates the Einstein evolution equations, and when restricted to $\mathscr{G}_{\mathrm{dyn}}$, only the Poincaré group is left as generators.

Although it seems reasonable to suppose that the past limit of the Bondi-Sachs mass is the ADM mass, which is positive, this still does not exclude the possibility that the Bondi-Sachs mass may decrease to a negative value. In fact, such examples seem to have been constructed by MILLER [88] and STEINMÜLLER, KING and LASOTA [89].

REFERENCES

[1] A. E. FISCHER and J. E. MARSDEN: this volume, p. 322.
[2] A. EINSTEIN: *Preuss. Akad. Wiss. Sitz.*, 448 (1918).
[3] F. KLEIN: *Gott. Nachr.*, 394 (1918).
[4] P. VON FREUD: *Ann. of Math.*, **40**, 417 (1939).
[5] A. PAPAPETROU: *Proc. Roy. Irish Acad.*, **52**, 11 (1948).
[6] S. N. GUPTA: *Phys. Rev.*, **96**, 1683 (1954).
[7] J. N. GOLDBERG: *Phys. Rev.*, **111**, 315 (1958).
[8] R. ARNOWITT, S. DESER and C. W. MISNER: *The dynamics of general relativity*, in *Gravitation, an Introduction to Current Research*, edited by L. WITTEN (New York, N. Y., 1962), p. 227.
[9] A. TRAUTMAN: *Conservation laws in general relativity*, in *Gravitation, an Introduction to Current Research*, edited by L. WITTEN (New York, N. Y., 1962).
[10] P. A. M. DIRAC: *Phys. Rev.*, **114**, 924 (1959).
[11] B. DEWITT: *Phys. Rev.*, **160**, 1113 (1967).
[12] T. REGGE and C. TEITELBOIM: *Ann. of Phys.*, **88**, 286 (1974).
[13] M. S. WALKER: this volume, p. 61; *Gen. Rel. Grav.*, **6**, 489 (1975).
[14] H. BONDI: *Some special solutions of the Einstein equations*, in *Brandeis Lectures on General Relativity* (Englewood Cliffs, N. J., 1964).
[15] A. KOMAR: Ph. D. Thesis, Princeton University (1956); *Phys. Rev.*, **129**, 1873 (1963).
[16] A. TAUB: *Ann. of Math.*, **53**, 472 (1951).
[17] H. ARAKI: *Ann. of Phys.*, **7**, 456 (1959).
[18] D. BRILL: *Ann. of Phys.*, **7**, 466 (1959).
[19] R. ARNOWITT, S. DESER and C. W. MISNER: *Ann. of Phys.*, **11**, 116 (1960).
[20] P. G. BERGMAN: *Phys. Rev.*, **112**, 287 (1958).
[21] C. MOLLER: *Ann. of Phys.*, **4**, 347 (1958).
[22] C. MOLLER: *Ann. of Phys.*, **12**, 118 (1961).
[23] C. W. MISNER: *Phys. Rev.*, **130**, 1590 (1963).
[24] D. BRILL and S. DESER: *Ann. of Phys.*, **50**, 548 (1968); *Phys. Rev. Lett.*, **20**, 75 (1968); *Phys. Lett.*, **26** A, 538 (1968).
[25] A. TROMBA: *Can. Journ. Math.*, **28**, 640 (1976).
[26] A. FISCHER: *The theory of superspace*, in *Relativity*, edited by M. CARMELI, S. FICKLER and L. WITTEN (New York, N. Y., 1970).
[27] A. FISCHER and J. MARSDEN: *Bull. Amer. Math. Soc.*, **79**, 995 (1973).
[28] R. GEROCH: *General relativity*, in *Proceedings of the Symposium in Pure Mathematics*, Vol. **27**, part 2 (1975), p. 401 (see also *Journ. Math. Phys.*, **13**, 956 (1972)).
[29] N. O'MURCHADHA and J. W. YORK: *Phys. Rev. D*, **10**, 2345 (1974).
[30] D. ROBINSON and J. WINICOUR: *Journ. Math. Phys.*, **13**, 1435 (1972) (see also *Phys. Rev.*, **3**, 840 (1971)).

[31] Y. CHOQUET-BRUHAT and J. MARSDEN: *Compt. Rend.*, **282**, 609 (1976); *Comm. Math. Phys.*, **51**, 283 (1976).
[32] A. LICHNEROWICZ: *Journ. Math. Pures Appl.*, **23**, 37 (1944).
[33] Y. CHOQUET-BRUHAT: *Symposia Math.*, **12**, 317 (1973).
[34] N. O'MURCHADHA and J. W. YORK: *Phys. Rev. D*, **10**, 428 (1974); *Phys. Rev. Lett.*, **28**, 1082 (1972).
[35] A. FISCHER and J. MARSDEN: *Linearization stability of nonlinear partial differential equations*, in *Proceedings of the Symposium in Pure Mathematics*, Vol. **27**, part 2 (1975), p. 219.
[36] L. SMARR, A. CADEZ, B. DEWITT and K. EPPLEY: *Phys. Rev. D*, **14**, 2443 (1976); L. SMARR: *Spacetimes generated by computers: black holes with gravitational radiation* (preprint).
[37] Y. CHOQUET-BRUHAT: *Compt. Rend.*, **280**, 169 (1975).
[38] Y. CHOQUET-BRUHAT: *Compt. Rend.*, **281**, 577 (1975).
[39] Y. CHOQUET-BRUHAT: *Ann. Scuola Norm. Sup. Pisa*, **3**, 361 (1976).
[40] M. CANTOR, A. FISCHER, J. MARSDEN, N. O'MURCHADHA and J. YORK: *Comm. Math. Phys.*, **49**, 187 (1976).
[41] D. BRILL and F. FLAHERTY: *Comm. Math. Phys.*, **50**, 157 (1976).
[42] A. J. GODDARD: *Gen. Rel. Grav.*, **8**, 525 (1977); *Comm. Math. Phys.*, **54**, 279 (1977).
[43] F. TIPLER and J. MARSDEN: *Maximal hypersurfaces in spacetime* (to appear).
[44] A. AVEZ: *Ann. Inst. Fourier (Grenoble)*, **13**, 105 (1963).
[45] R. P. GEROCH: *Phys. Rev. Lett.*, **17**, 445 (1966); thesis (Princeton) (1967).
[46] S. DESER: *Phys. Rev. D*, **12**, 943 (1975); *Gen. Rel. Grav.*, **8**, 573 (1977).
[47] P. S. JANG: *Journ. Math. Phys.*, **17**, 141 (1976).
[48] N. O'MURCHADHA: *Positive energy and stationary solutions in general relativity* (preprint).
[49] S. DESER and C. TEITELBOIM: *Phys. Rev. Lett.*, **39**, 249 (1977).
[50] M. T. GRISARU: *Positivity of the energy in Einstein theory*, to appear in *Phys. Lett.*
[51] K. BLEULER: *Geometric methods in mathematical physics, Springer Lecture Notes in Mathematics*, No. 585 (1977).
[52] R. SCHOEN and S. T. YAU: *Positivity of the gravitational mass*, to appear in *Comp. Math.*
[53] Y. CHOQUET-BRUHAT and S. DESER: *Ann. of Phys.*, **81**, 165 (1973) (also *Compt. Rend.*, **275**, 1019 (1972)).
[54] L. NIRENBERG and H. WALKER: *Journ. Math. Anal. Appl.*, **47**, 271 (1973).
[55] M. CANTOR: *Ind. Univ. Math. Journ.*, **24**, 897 (1975).
[56] M. CANTOR: *Some problems of global analysis on asymptotically simple manifolds*, to appear in *Comp. Math.*
[57] L. NIRENBERG: *Ann. Scuola Norm. Sup. Pisa*, **13**, 115 (1959).
[58] D. EBIN: *The manifold of Riemannian matrices*, in *Proceedings of the Symposium in Pure Mathematics*, Vol. **15** (1970), p. 11.
[59] Y. CHOQUET-BRUHAT: *Compactification de variétés asymptotiquement euclidiennes*, to appear in *Compt. Rend.*
[60] Y. CHOQUET-BRUHAT: *Compt. Rend.*, **226**, 1071 (1948); *Journ. Rat. Mech. Anal.*, **5**, 951 (1956); *Cauchy problem*, in *Gravitation, an Introduction to Current Research*, edited by L. WITTEN (New York, N. Y., 1962).
[61] J. W. YORK: *Phys. Rev. Lett.*, **28**, 1082 (1972).
[62] N. O'MURCHADHA and J. N. YORK: *Journ. Math. Phys.*, **141**, 1551 (1973).
[63] D. BANCEL and J. LACAZE: *Compt. Rend.*, **283**, 405 (1976).
[64] S. HAWKING and J. ELLIS: *The Large Scale Structure of Spacetime* (Cambridge, 1973).
[65] A. LICHNEROWICZ: *Problèmes globaux en mécanique relativiste* (Paris, 1939).

[66] A. RAYCHAUDHURI: *Phys. Rev.*, **86**, 90 (1952).
[67] K. KUCHAŘ: *Journ. Math. Phys.*, **17**, 777 (1976).
[68] J. SCHWARTZ: *Nonlinear Functional Analysis* (New York, N. Y., 1969).
[69] S. Y. CHENG and S. T. YAU: *Ann. of Math.*, **104**, 407 (1976).
[70] A. FISCHER and J. MARSDEN: *Duke Math. Journ.*, **42**, 519 (1975).
[71] Y. CHOQUET-BRUHAT, A. FISCHER and J. MARSDEN: *Compt. Rend.*, **284**, 975 (1977).
[72] P. D. D'EATH: *Ann. of Phys.*, **98**, 237 (1976).
[73] M. CANTOR: *Comm. Math. Phys.*, **57**, 83 (1977).
[74] C. W. MISNER, K. S. THORNE and J. A. WHEELER: *Gravitation* (San Francisco, Cal., 1973).
[75] S. WEINBERG: *Gravitation and Cosmology* (New York, N. Y., 1972).
[76] M. BERGER and D. EBIN: *Journ. Diff. Geom.*, **3**, 379 (1969).
[77] R. PALAIS: *Bull. Amer. Math. Soc.*, **75**, 968 (1969).
[78] J. EELLS and J. SAMPSON: *Amer. Journ. Math.*, **86**, 109 (1964).
[79] R. PALAIS: *Critical point theory and the minimax principle*, in *Proceedings of the Symposium in Pure Mathematics*, Vol. **15** (1970), p. 185.
[80] J. WINICOUR: *Springer Lecture Notes in Physics*, **14**, 145 (1970).
[81] J. MARSDEN: *Applications of Global Analysis in Mathematical Physics* (Boston, Mass., 1974); *Comm. Math. Phys.*, **30**, 79 (1973).
[82] J. MARSDEN and T. J. R. HUGHES: *Topics in the mathematical foundations of elasticity*, in *Nonlinear Analysis and Mechanics*, Vol. **2**, edited by R. J. KNOPS (New York, N. Y., 1978).
[83] J. MARSDEN and A. FISCHER: *On the existence of nontrivial, complete, asymptotically flat, spacetimes*, in *Publ. Univ. Lyon*, Vol. **4** (fasc. 2) (1972), p. 182.
[84] L. MARDER: *Proc. Roy. Soc.*, **327**, 123 (1972).
[85] T. HUGHES, T. KATO and J. MARSDEN: *Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity*, to appear in *Arch. Rat. Mech. Anal.* (1976).
[86] R. ABRAHAM and J. MARSDEN: *Foundations of Mechanics*, second edition (Reading, Mass., 1978).
[87] R. HANSEN and A. ASHTEKAR: *The structure of spatial infinity* (preprint).
[88] B. D. MILLER: *Astrophys. Journ.*, **208**, 275 (1976).
[89] B. STEINMÜLLER, A. R. KING and J. P. LASOTA: *Phys. Lett.*, **51** A, 191 (1975).

Reprinted From
Isolated Gravitating Systems in General Relativity