A Phase Space for Gravitational Radiation

Adam D. Helfer

Department of Mathematics, University of Missouri, Columbia, MO 65211, U.S.A.

Abstract: We give a new definition, based on considerations of well-posedness for a certain asymptotic initial value problem, of the phase space for the radiative degrees of freedom of the gravitational field in exact General Relativity. This space fibres over the space of final states, with the fibres being the purely radiative degrees of freedom. The symplectic form is rigorously identified.

The infrared sectors are shown to be the level surfaces of a moment map of an action of the quotient group Supertranslations/Translations. A similar result holds for Electromagnetism in Minkowski space.

1. Introduction

The theory of gravitational radiation interweaves physical and mathematical progress to an unusual degree: some physical intuition is necessary to begin to pose questions which are significant; and the construction of a mathematically satisfactory theory often serves as a test of that intuition. If we define radiation mathematically as disturbances which escape to (or come in from) infinity, then gravitational radiation, which travels at the speed of light, will escape to null infinity, and ought to be analyzed there. Since what "null" is determined by the dynamical field, from the analytic point of view one is studying the asymptotics along characteristics of a system of quasilinear equations. There are several, inequivalent, mathematical meanings this might have, and the nature of the physical questions under consideration must guide the analysis.

It was Bondi and coworkers (Bondi 1960, Bondi, van der Burg and Metzner 1962, Sachs 1962a; see also Newman and Penrose 1962) who introduced the idea of analyzing the field at large null separations, and Penrose (1963) who recast their asymptotic conditions as the existence of a null hypersurface $\mathscr{I}$ at null infinity. After this work, it was clear that at a formal level the radiative degrees of freedom were represented by a function, now called the Bondi shear, on $\mathscr{I}$.

Ashtekar and Streubel (1981) seem to have been the first to realize that it would be desirable to develop, not just a space-time by space-time analysis, but a phase space of radiative degrees of freedom in terms of the asymptotics of the fields.

They gave such a construction of a phase space entirely in terms of (functions equivalent to) the Bondi shear. This resulted in a new derivation of the Bondi energy-loss formula, and a new expression for the total angular momentum emitted in gravitational waves (which has since been confirmed by independent means; see Helfer 1990). Also Ashtekar (1987) used this phase space to construct a kinematical framework which might be hoped to underly an S-matrix theory of gravitons.

It is not straightforward to decide whether such a phase space is correct. The strongest argument would be to show that it was an appropriate limit of phase spaces constructed from data on Cauchy surfaces. A program for such a proof was outlined by Ashtekar and Magnon-Ashtekar (1982); their idea was to deform a Cauchy surface to $\mathscr{I}$. They were able to show, at a formal level, that the symplectic form of Ashtekar and Streubel was correct. However, they also pointed out that there were some analytic limitations which might preclude the conclusions of the formal argument from being valid generally. It seemed likely that in many cases – perhaps generically – the symplectic form would have to be supplemented by "leakage" terms at timelike infinity. Also, if one considers a space-time which radiates in such a way as to become vacuum to the future of some asymptotically null hypersurface, then the sorts of perturbations allowed by the Ashtekar-Streubel phase space will include space-times with negative Bondi-Sachs energies, which one would like to rule out as unphysical in light of the positivity theorems (Ludvigsen and Vickers 1981, Horowitz and Tod 1982, Schoen and Yau 1982). So the successes of the Ashtekar-Streubel construction suggested that it ought to be (at least largely) correct, but it seemed difficult to justify rigorously.

I shall construct a phase space of the radiative degrees of freedom of the gravitational field in exact General Relativity. The isolation of these degrees of freedom is determined by considerations of well-posedness for a certain asymptotic initial value problem (the "M-shaped" hypersurface problem, which is well-posed by recent theorems of Friedrich 1983, 1992 and Rendall 1990, 1992.) We shall be able to make precise the sense in which the Ashtekar-Streubel symplectic form is valid, and understand why there are no leakage terms.

The most important consequence is a shift in the idea of what a radiation problem is. In order to isolate "all" of the outgoing radiative degrees of freedom, one must take two limits: one must move out to $\mathscr{I}$; and one must also consider arbitrarily late retarded times. Although for any sufficiently well-behaved space-time one can consider both limits, one cannot reverse this and solely from the radiative data reconstruct the space-time. One must give the internal state at some late retarded time, together with the radiation emitted prior.[1] From this point of view, the Ashtekar-Streubel phase space is an approximation in which radiation has wholly decoupled from the internal degrees of freedom. The coupling which does exist is rather mild (since the radiative data must match up to the internal data only at one cut of $\mathscr{I}$), so the approximation is valid for many purposes.

From this point of view, the total phase space has the structure of a fibre bundle, with the radiative degrees of freedom (prior to the late retarded time) fibering over the space of final states (at the late retarded time). This is because the radiative data must agree with the internal data at the cut of $\mathscr{I}$ corresponding to the retarded time. To a large extent, the fibres have simple structures, and most of the complications are in the base space. (This is related to the fact that $\mathscr{I}$ is a null hypersurface,

---

[1] It may at first be more intuitive to consider the time-reversed case, where one specifies a space-time by giving the internal state at some advanced time, together with the radiation incoming thereafter.

and so data on it are unconstrained.) We are concerned here only with a fibrewise analysis. An understanding of the total phase space would be a natural next step; one would expect a rigorous analysis of the linearization stability of the "M-shaped" problem to be possible (cf. Moncrief 1975, 1976).

We consider a somewhat different class of radiative data from those of previous constructions. One difference is quite minor (we work with a generalization of test functions, rather than a weighted Fréchet space) and not expected to be physically significant, but another is important: we allow "infrared" fields. These are fields which, although pure gauge at early and late retarded times, nevertheless have a gauge mismatch between these two regimes. We shall find that the residual gauge freedom at $\mathscr{I}$ gives rise to a foliation on the phase space which turns out to be a classical analog of the emergence of infrared sectors in Quantum Electrodynamics. The discovery of these was a surprise in quantum theory, but we can now see that it is presaged (logically, if not historically) by the foliation of the phase space.

The present construction seems likely to help with a similar infrared problem in the asymptotic quantization program for General Relativity. In Quantum Electrodynamics, the infrared sector for the outgoing electromagnetic field is determined by the asymptotic state of the outgoing charged particles, and this identification is important in constructing a separable Hilbert space. It had seemed that there would be no way of deciding on the infrared sector for the outgoing gravitational waves (Ashtekar and Narain 1981, Ashtekar 1987). However, it now seems that the internal state of the field at a late retarded time could determine the infrared sector. This will be considered elsewhere.

In a separate paper (Helfer 1994), I show how this phase space can be used to derive the asymptotic motions of space-time to which the angular momenta at $\mathscr{I}$, as defined by Penrose (1982), are conjugate. In this analysis, not only is knowledge of some of the internal data necessary, but also the foliation of the phase space by the infrared sectors enters crucially.

The plan of the paper is this. The next section is devoted to a review of the definitions which will be needed: spin- and boost-weighted functions; Bondi coordinates; the Bondi-Metzner-Sachs group; and the effect of an abelian gauge freedom on a fibred phase space. In Sect. 3, we sketch the application of our ideas in the case of electromagnetism, and show how the infrared sectors emerge. Section 4 reviews the results of Friedrich and Rendall, and establishes the existence of maximal globally hyperbolic solutions of the initial-value problem (Theorem 4.1). In Sect. 5, the manifold underlying the phase space is constructed as a certain function space. In Sect. 6, the main technical result, the correctness of the Ashtekar-Streubel symplectic form is established (Theorem 6.1). It is perhaps surprising that such a rigorous result is presently possible, since a direct approach to this problem would require better asymptotic estimates for Einstein's equations than are known. The last section discusses the infrared sectors for gravity. These are related to motions of the phase space in a somewhat different fashion from those for electromagnetism.

We treat only the case of pure gravity, but it should be evident that coupling to matter can be accomodated within this framework.

Conventions. Our conventions for space-time quantities are those of Penrose and Rindler (1984–6). We will be concerned with calculus on the phase space and on the space-time. It will be convenient to use the standard spin-coefficient formalism on the space-time, and the modern coordinate-free notation for the phase space. We use $\langle s, \lambda \rangle$ to stand for the duality pairing between a vector s and a covector λ in

this case. Thus the symbol d represents a gradient in the phase space except when it occurs in a line, area or volume element, du, $d\mathscr{S}$ or $d\mathscr{S}\,du$, with respect to the Bondi coordinate system on $\mathscr{I}$.

Choice of Function Spaces. In most circumstances, the particular choices of function space used in General Relativity are not of physical significance, and are made for mathematical convenience. The real question is whether the choice at hand admits a sufficiently broad class of fields to be of interest. For the analysis of gravitational radiation, although the formalism of null infinity was introduced in the early sixties, it is only recently that existence theorems for solutions to the Einstein equations with the right sort of control of the asymptotics have begun to be proved.

As of this writing, it is not clear whether there is a physically preferred degree of smoothness for the asymptotic regimes. We have chosen to work in the C^{∞} category, since the theorems of Friedrich and Rendall make this possible and in the absence of any other choice it seems the most natural. The function spaces we use are analogs of the familiar spaces of test functions, and so are unphysical approximations in that they become exactly zero (or pure gauge) in various regimes where a realistic field would only decay. However, it will be apparent that our constructions could accomodate such behavior, and fields of low differentiability, without difficulty.

2. Preliminaries

We collect here some terminology which will be assumed in what follows.

2.1. Spin- and Boost-Weighted Functions

The radiative degrees of freedom of the gravitational field will be encoded in a function called the Bondi shear. This function takes values in a certain complex line bundle over null infinity which is the pull-back of a bundle on the space of generators; in classical language, the function has spin and boost weight. We review here the essential points of this formalism; for details, see Penrose and Rindler (1984).

Spin space is a two-complex dimensional vector space. Its associated projective space may be identified with a Riemann sphere with a conformal structure and an orientation. The bundle of functions of spin-weight $(p-q)/2$ and boost-weight $(p+q)/2$ will be denoted $\{p,q\}$. Its space of smooth sections will be denoted $C^{\infty}(S^2,\{p,q\})$.

There are two important differential operators (essentially, Dirac operators) on these bundles,

$$\eth : C^{\infty}(S^2,\{p,q\}) \to C^{\infty}(S^2,\{p+1,q-1\}) \tag{1}$$

and its conjugate $\eth' : C^{\infty}(S^2,\{p,q\}) \to C^{\infty}(S^2,\{p-1,q+1\})$.

The conjugate of $\{p,q\}$ is $\{q,p\}$. In particular, elements of $\{p,p\}$ have unique SL(2, $\mathbb{C}$)-invariant decompositions into real and imaginary parts. There is also a hidden SL(2, $\mathbb{C}$)-invariant Hermitian structure for certain other spaces of sections of the line bundles. It turns out that the operator $\eth^{q+1} : C^{\infty}(S^2,\{p,q\}) \to C^{\infty}(S^2,\{p+q+1,-1\})$ is SL(2, $\mathbb{C}$)-invariant. Let us consider the case $p=q$. Then this operator is surjective and its kernel is the complexification of a real

SL(2, $\mathbb{C}$)-invariant subspace of $C^\infty(S^2, \{p, q\})$. This allows us to write each element in the image as a sum of the image of a real section and the image of an imaginary section. These are the *electric* and *magnetic* parts of $C^\infty(S^2, \{2q + 1, -1\})$ (Newman and Penrose 1966). The important case of this decomposition for us will be for the Bondi shear, $\sigma \in C^\infty(S^2, \{3, -1\})$. We can always write $\sigma = \eth^2 \lambda$ for some $\lambda \in C^\infty(S^2, \{1, 1\})$. The electric part of σ is $\eth^2 \Re \lambda$ and the magnetic part of σ is $i\eth^2 \Im \lambda$.

2.2. Bondi Coordinates

Bondi coordinates are appropriate for radiation problems. They are defined as follows. (See Penrose and Rindler 1984–6 for a fuller discussion.)

Let $(\widehat{M}, \widehat{g}_{ab})$ be an oriented, time-oriented space-time. Suppose that $\widehat{M}$ embeds in a manifold with boundary M, and there exists a smooth function Ω on M, positive on $\widehat{M}$ and zero on $\mathscr{I}^+ = \partial M$, with nowhere-vanishing gradient on $\mathscr{I}^+$. Suppose also $g_{ab} = \Omega^2 \widehat{g}_{ab}$ extends smoothly to a non-degenerate metric on M, and that all points on $\mathscr{I}^+$ are future end-points of null geodesics. Then we call $\mathscr{I}^+$ *future null infinity*.[2]

Suppose further that $\mathscr{I}^+$ is a g_{ab}-null hypersurface, diffeomorphic to $\mathbb{R} \times S^2$, with the "$\mathbb{R}$" factors being its null generators, and that the field $n_a = -\nabla_a \Omega$ is shear-free on $\mathscr{I}^+$. Then we can choose the conformal factor so that cross-sections of the fibration $\mathscr{I}^+ \to S^2$ are unit spheres, and we do so. We can also choose n^a to be g_{ab}-constant up the generators of $\mathscr{I}^+$, and we do this. Then a function u satisfying $n^a \nabla_a u = 1$ is a *Bondi retarded time coordinate*. Fixing such a function, we let (θ, φ) be standard polar coordinates on the cross-sections $u = $ constant, with $n^a \nabla_a \theta = 0 = n^a \nabla_a \varphi$. Then (u, θ, φ) form a *Bondi system* on $\mathscr{I}^+$, and $(u, \Omega, \theta, \varphi)$ form a *Bondi system* on a neighborhood of $\mathscr{I}^+$.[3] The abbreviation

$$d\mathscr{S} = \sin\theta \, d\theta \, d\varphi \tag{2}$$

will be used.

For example, for Minkowski space, we let $(u = t - r, r, \theta, \varphi)$ be retarded polar coordinates. We obtain (M, g_{ab}) by gluing to this $\{(u, \rho, \theta, \varphi) \mid u \in \mathbb{R}, \ \rho \geq 0, (\theta, \varphi) \in S^2\}$ via the relation $\rho = 1/r$, and we take $\Omega = \rho$ near $\rho = 0$.

It is often convenient to introduce a complex g_{ab}-null tetrad $(l^a, m^a, \overline{m}^a, n^a)$ at $\mathscr{I}^+$ associated with the Bondi system. Here n^a is as before, $m^a = 2^{-1/2}(\partial_\theta - i \csc\theta \partial_\varphi)$, and l^a is the unique future-pointing vector satisfying $l^a n_a = 1$, $l^a m_a = 0$.

The *Bondi shear* at $\mathscr{I}^+$ is the (appropriately conformally rescaled) shear of the $u = $ constant cuts. It takes values in $\{3, -1\}$. Notice that by its definition, the Bondi shear depends on the choice of Bondi system.

2.3. The Bondi-Metzner-Sachs Group

A Bondi system (u, θ, φ) will not be unique. There is a group, the *Bondi-Metzner-Sachs (BMS)* group, acting on the set of Bondi systems. If the null generators

---

[2] This differs from some definitions in that we do not require *every* $(\widehat{M}, \widehat{g}_{ab})$-future endless null geodesic to have an end-point on $\mathscr{I}^+$.

[3] We require $(u, -\Omega, \theta, \varphi)$ to be compatible with the orientation on $(\widehat{M}, \widehat{g}_{ab})$. The minus sign is because Ω decreases as one moves towards $\mathscr{I}^+$ from the interior.

of $\mathscr{I}^+$ are complete, that is, if a Bondi parameter u takes all real values for each $(\theta, \varphi) \in S^2$, then there is an associated "active" action by diffeomorphisms. These active motions are in a certain sense the asymptotic symmetries of the gravitational field, and we shall give an account of them. Although the following discussion is phrased in terms of the group action, it applies at the Lie algebra level even if the generators of $\mathscr{I}^+$ are not complete.

The BMS group is generated by two sorts of motions: the proper, orthochronous motions Lorentz $= O(1,3)_{\uparrow+}$, which act much the same way as they do on half of a light-cone; and the *supertranslations*, which have the form

$$u \mapsto u + \alpha(\theta, \varphi), \; \theta \mapsto \theta, \; \varphi \mapsto \varphi , \qquad \qquad . \qquad (3)$$

where α is an arbitrary smooth function. The BMS group is a semidirect product of Lorentz and Supertranslations:

$$0 \rightarrow \text{Supertranslations} \rightarrow \text{BMS} \rightarrow \text{Lorentz} \rightarrow 0 . \qquad (4)$$

The relation of the BMS group to the connected component of the isometry group of Minkowksi space, the Poincaré group, is important. The Poincaré group is also a semidirect product,

$$0 \rightarrow \text{Translations} \rightarrow \text{Poincaré} \rightarrow \text{Lorentz} \rightarrow 0 , \qquad (5)$$

and an analogy is apparent. The connection is still closer, because there is a unique four-dimensional normal abelian subgroup of BMS which may be identified with the translations:

$$\text{Translations} \cong \{\alpha \in \text{Supertranslations} \mid \eth^2 \alpha = 0\} . \qquad (6)$$

However, there is no invariant sense to a "translation-free supertranslation." Therefore there is no canonical subgroup of BMS to identify with Poincaré.

We shall be concerned with the action of the BMS group on the Bondi shear. We noted above that this shear is the shear of the $u = $ constant cuts, and so depends on the choice of Bondi system. In order to have a well-defined action, then, we must specify whether the active motion is accompanied by a passive change in Bondi system, and whether the shear is regarded as a function on the abstract manifold $\mathscr{I}^+$ or a function of the coordinates (u, θ, φ). We shall keep the Bondi system the same, and regard the shear as a function on $\mathscr{I}^+$. Then if $\phi : \mathscr{I}^+ \rightarrow \mathscr{I}^+$ is the diffeomorphism generated by the supertranslation α, the action is

$$\sigma \mapsto \sigma \circ \phi^{-1} + \eth^2 \alpha . \qquad (7)$$

Notice that it transforms like an ordinary function precisely under the translations. This may be contrasted with the result of a passive supertranslation by α, which action is

$$\sigma \mapsto \sigma + \eth^2 \alpha . \qquad (8)$$

2.4. Gauge Motions on a Fibred Phase Space

We wish to call to mind some aspects of the reduction of a phase space with fibre structure, and to fix some terminology. Accordingly, in this sketch, we ignore all technical difficulties (subtleties in infinite dimensions, whether quotients are manifolds, etc.).

Let us recall the Bergmann-Dirac theory of first-class constraints (Bergmann and Goldberg 1955, Dirac 1964). Let (Γ, ω) be a phase space, and suppose $\{C_\alpha \mid \alpha \in A\}$ is a family of constraints closed under Poisson brackets. The symplectic form on the constraint submanifold $\Gamma_{\text{constraint}} = \{p \in \Gamma \mid C_\alpha(p) = 0\}$ is degenerate, and this degeneracy is compensated by introducing an equivalence relation. Let V_α be the Hamiltonian vector field conjugate to C_α:

$$dC_\alpha = \omega(\,\cdot\,, V_\alpha)\,; \tag{9}$$

and let $\mathscr{V}$ be the Lie algebra generated by these vector fields. Then two points in $\Gamma_{\text{constraint}}$ are identified if there is a flow of a vector field in $\mathscr{V}$ taking the first to the second. The reduction of the phase space thus has two parts: restriction to $\Gamma_{\text{constraint}}$, and passage to a quotient.

We may run this backwards.[4] Suppose a connected abelian (for simplicity) gauge group acts on (Γ, ω). (That is, the action is a representation of the gauge group by symplectomorphisms.) Let A be the Lie algebra of the group, and let V_α be the Hamiltonian vector field generated by $\alpha \in A$. Furthermore, let us suppose we can find a moment map $C : \Gamma \to A^*$, so that $dC_\alpha = \omega(\,\cdot\,, V_\alpha)$. Then the level surfaces $\Gamma_c = \{p \in \Gamma \mid C_\alpha(p) = c_\alpha\}$ for $c \in A^*$ play the part of the constraint manifold in the discussion above. From this point of view, however, there is no reason for preferring one level surface to another. Therefore each such level surface, modulo the gauge action, forms a reduced phase space. We may think of the collection of these quotients, as the level surface varies, as a foliation of a space whose leaves are the reduced phase spaces.

We shall use the terms leaf *and* foliation *only to refer to the possibility of choosing different level surfaces*, not to an orbit or the space of orbits of the gauge action. We may speak of a foliation in this sense whether or not we have passed to the quotient by the gauge action.

It is the presence of such foliations that is most important in certain radiation problems. Passing to the quotient by the gauge action may or may not be convenient (and in the gravitational case is definitely undesirable). However, the foliation enters the theory by restricting the admissible vectors to those tangent to the leaves. This restriction turns out to be crucial in developing a satisfactory theory of the angular momentum of radiation.

We shall show below that the radiative modes are actually fibres in certain phase spaces, so let us see what the effects of such structure are. Let $(\Gamma_{\text{total}}, \omega)$ be a phase space; suppose there is a fibration

$$
\begin{array}{c}
\Gamma \;\; \to \;\; \Gamma_{\text{total}} \\
\downarrow \\
\Gamma_{\text{base}}
\end{array}
\tag{10}
$$

so that ω restricts to a non-degenerate form on each fibre. Suppose further that the gauge action respects the fibration, that is, there is a pair of actions $A \times \Gamma_{\text{total}} \to \Gamma_{\text{total}}$, $A \times \Gamma_{\text{base}} \to \Gamma_{\text{base}}$ such that the diagram

$$
\begin{array}{ccc}
A \times \Gamma_{\text{total}} & \to & \Gamma_{\text{total}} \\
\downarrow & & \downarrow \\
A \times \Gamma_{\text{base}} & \to & \Gamma_{\text{base}}
\end{array}
\tag{11}
$$

---

[4] Jan Segert points out this is the reduction of Marsden and Weinstein (1974).

commutes. Let $K : \Gamma_{\text{total}} \to A^*$ be the moment map. Although the gauge motions do not leave the fibres invariant, some of the information in the moment map and its level surfaces may be recovered from data on the fibres.

We can regard each fibre as a phase space, so suppose we have found a moment map $C : \Gamma \to A^*$ (depending smoothly on the base point). By this we mean

$$\langle W, dC_\alpha \rangle = \omega(W, V_\alpha) = \langle W, dK_\alpha \rangle \tag{12}$$

for W tangent to Γ. Therefore we have

$$C = K + D \tag{13}$$

for some D depending only on the projection to Γ_{base}. Thus any level surface of the total moment map will intersect each fibre in a level surface of the fibrewise moment map, and conversely. It will be enough, then, to construct the fibrewise moment maps.

3. Electromagnetism

In this section, we give a treatment of electromagnetism on Minkowski space, analogous to that of General Relativity below. The case of electromagnetism is somewhat simpler than gravity, because the gauge freedom acts "vertically," in a bundle over space-time, and because the linearity of the theory allows a cleaner isolation of the radiation field. Still, the results here will be useful in interpreting those for gravity. Our emphasis here will be conceptual, and we shall give no proofs. (These can easily be adapted from the gravitational case, below.) These results build on earlier work of Ashtekar and Narain (1981) and Sparling, Newman and coworkers (e.g. Chakravarty et al. 1986).

The electromagnetic potential will be denoted $\widehat{\Phi}_a$. (Quantities with hats refer to the physical space-time; conformally rescaled quantities are bareheaded.) The field strength is $\widehat{F}_{ab} = 2\widehat{\nabla}_{[a}\widehat{\Phi}_{b]}$. We shall suppose there are no sources, and that the field extends smoothly to $\mathscr{I}^+$ (and so has the usual peeling behavior). We may take $\Phi_a = \widehat{\Phi}_a$ to be smooth at $\mathscr{I}^+$. It is convenient to choose the gauge so that $n^a\Phi_a = 0$ at $\mathscr{I}^+$, where n^a is a null vector tangent to the generators of $\mathscr{I}^+$. Then the most important quantities are the rescaled components

$$\Phi = \lim_{r \to \infty} r\widehat{m}^a\Phi_a \quad \text{and} \quad \phi_2 = \lim_{r \to \infty} r\widehat{F}_{ab}\widehat{\overline{m}}^a\widehat{n}^b , \tag{14}$$

which are smooth on $\mathscr{I}^+$. These are related by $\phi_2 = -\partial_u\overline{\Phi}$. Note that the gauge freedom remaining at $\mathscr{I}^+$ is $\Phi \mapsto \Phi + \eth\gamma$, where γ is a smooth real-valued function of angle only.

The function ϕ_2 can be thought of as the profile of the emitted radiation. It very nearly determines the electromagnetic field throughout space-time, assuming no sources of present.[5] Indeed, if we consider electromagnetic fields which vanish in some neighborhood of future timelike infinity, then the field is determined

---

[5] If the field is assumed to have a certain regularity throughout Minkowski space, or at future timelike infinity, the determination is complete. However, these regularity assumptions are not natural in radiation problems, and analogous assumptions in the gravitational case would be ill-defined or too strong.

throughout space-time by ϕ_2. More generally, if Cauchy data for the field are given on a spacelike hyperboloidal initial surface Σ meeting $\mathscr{I}^+$ at $u = u_0$, and if ϕ_2 is specified on $\mathscr{I}^+$ to the past of u_0 (and if $\phi_2|_{u=u_0}$ agrees with the radiation field at u_0 which would be deduced from the Cauchy data), then the Maxwell field is determined throughout space-time. Since Φ determines ϕ_2, knowledge of Φ for $u \leq u_0$ can replace knowledge of ϕ_2 in this argument.

The phase space may thus be constructed from two sorts of data: the usual Cauchy data (with suitable asymptotics) on Σ; and the field Φ on $\mathscr{N} = \{p \in \mathscr{I}^+ | \ u < u_0\}$. We shall call the Cauchy data on Σ the *final states*, and the field Φ on $\mathscr{N}$ the *radiative modes*. Since Φ must agree at $Z = \{p \in \mathscr{I}^+ \mid u = u_0\}$ with what can be deduced from the final state, the phase space has a bundle structure: the radiative modes fibre over the final states. We shall give a fibrewise analysis.

A formal construction of a fibre of the phase space is as follows. Fix a gauge on $\mathscr{I}^+$ as above, and fix a final state. From this, we may work out the field at $u = u_0$, and all its u-derivatives there. For simplicity, assume the field is gauge to all orders at $u = u_0$. Let Γ be the space of fields Φ of type $\{1, -1\}$ on $\mathscr{N}$ which are smooth, agree at u_0, together with all their derivatives, with the values inferred from the data on Σ (or are gauge-equivalent to such fields), and are pure gauge sufficiently far in the past. There is a natural topology and smooth structure on Γ. The symplectic form is

$$\omega(f_1, f_2) = (4\pi)^{-1} \int\limits_{u \leq u_0} \left[(\partial_u f_1)\overline{f}_2 - f_1 \partial_u \overline{f}_2\right] du \, d\mathscr{S} + \text{conjugate} \qquad (15)$$

(Ashtekar and Streubel 1981). It is closed and weakly non-degenerate.

Now let us consider the effect of a gauge change. The admissible changes are of the form $\Phi \mapsto \Phi + \eth\gamma$ on $\mathscr{I}^+$, where γ is a smooth function on the sphere.[6] We let V_γ be the vector field on Γ generating this motion. Then one can check that

$$\omega(\ \cdot\ , V_\gamma) = dC_\gamma , \qquad (16)$$

where

$$C_\gamma = (4\pi)^{-1} \oint [\Phi]\eth'\gamma \, d\mathscr{S} \ \Big|_{u=-\infty}^{u_0} + \text{conjugate} . \qquad (17)$$

As γ varies, this detects precisely the electric part of the gauge-invariant jump

$$[\Phi] = \Phi\Big|_{u=u_0} - \Phi\Big|_{u=-\infty} . \qquad (18)$$

Therefore the leaves of the foliation of the phase space are labeled by the electric parts of these jumps.[7]

Now let us consider fields for which Φ is pure gauge (i.e., purely electric) at u_0. Those with $[\Phi] = 0$ are said to lie in the *classical* sector; the others are in *infrared* sectors, with the sector labeled by the jump. Such sectors play an important role in

---

[6] Strictly speaking, there is not a well-defined action by such functions on the space of data on Σ, since there are many gauge changes on the interior of Σ with the same asymptotic behavior. Thus a correct treatment has the space of γ's as a quotient of a larger gauge freedom. The result is the same.

[7] It should be emphasized that the term "electric" is used in the sense of spin- and boost-weighted functions, and is not obviously related to the electric part of the electromagnetic field. However, it can be shown that $C_\gamma = (2\pi)^{-1} \oint [\mathfrak{R}_{\phi 1}]\gamma \, d\mathscr{S}$ where $\mathfrak{R}_{\phi 1}$ is a certain component of the electric field in the Bondi system.

Quantum Electrodynamics, in the construction of the Hilbert space of out-states.[8] Essentially, the sector is determined by the out-state of the charged particles, and a Hilbert space of acceptable states of the electromagnetic field, given these data, is constructed from fields in the sector. It was rather a surprise to find that the electromagnetic field was so restricted (Ashtekar and Narain 1981). However, we can now see that this is foreshadowed by the foliation of the classical phase space:

Theorem 3.1. *Fix a final state for which Φ is pure gauge at Z. Let C_γ be the moment map for the gauge freedom on $\mathscr{I}^+$. Then the classical sector is the zero-set of this moment map, and the infrared sectors are the level sets for non-zero values.*

4. The Space-Times

The phase space for the outgoing radiative modes in General Relativity will be constructed as a certain function space, whose elements represent data for an initial-value problem for Einstein's equations. In this Section, we discuss the space-times which are the solutions of these problems. We first review how the data determine the solutions locally, and then give a rapid treatment of the elementary aspects of the global theory.

4.1 Local Results

It will be clearest first to describe the sort of solutions sought, and then to characterize the data.

We shall want a vacuum space-time $(\widehat{M}, \widehat{g}_{ab})$, oriented and time-oriented, with $\widehat{M}$ embedding as the interior of a manifold with boundary M, and the following properties:

(a) There exists a smooth function Ω on M, positive on $\widehat{M}$ and zero on $\mathscr{I}^+ = \partial M$, with $\nabla_a \Omega$ nowhere zero on $\mathscr{I}^+$;

(b) The metric $g_{ab} = \Omega^2 \widehat{g}_{ab}$ extends smoothly to a non-degenerate metric on M;

(c) $\mathscr{I}^+$ is a null hypersurface diffeomorphic to $\mathbb{R} \times S^2$, with the "$\mathbb{R}$" factors being the null generators;

(d) The points on $\mathscr{I}^+$ are the future end-points of null geodesics in $(\widehat{M}, \widehat{g}_{ab})$;

(e) There exists a partial Cauchy surface $\widehat{\Sigma}$ in $(\widehat{M}, \widehat{g}_{ab})$ whose closure Σ in M is compact and meets $\mathscr{I}^+$ transversely in a cut Z (i.e., a section of the fibration $\mathscr{I}^+ \to S^2$);

(f) A Bondi retarded time coordinate attains arbitrarily negative values on each generator of $\mathscr{I}^+$.

Requirements (a)–(d) are standard for a future null infinity in radiation problems. (Property (d) simply distinguishes $\mathscr{I}^+$ as future null infinity.) Property (e) will turn out to ensure that the asymptotic initial-value problem is well-posed: data will be given on Σ and the portion $\mathscr{N}$ of $\mathscr{I}^+$ to the past of Z. Property (f) is also a standard requirement; it could be weakened, however, without much change.

---

[8] In our notation, the usual construction of infrared sectors corresponds to the case $u_0 = +\infty$. It seems likely, in view of the considerations of well-posedness discussed above, that some of the difficulties in the usual construction may be removed by delaying the limit $u_0 \to +\infty$ to a later stage in the analysis. Compare Jauch and Rohrlich (1976).

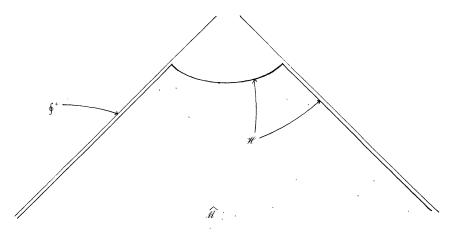**Fig. 1.** A cross-section of the initial-data surface $\mathscr{H} = \Sigma \cup \mathscr{N}$ for the space-time. Here Σ is the spacelike part of $\mathscr{H}$; this part meets $\mathscr{I}^+$ in the cut $Z = \partial\Sigma$. The set $\mathscr{N}$ is the portion of $\mathscr{H}$ on $\mathscr{I}^+$ to the past of Z

Now let us explain how such a space-time is determined by data on $\mathscr{H} = \Sigma \cup \mathscr{N}$.

Let Σ be an oriented three-manifold whose boundary Z is diffeomorphic to S^2; let $\widehat{\Sigma}$ be the interior of Σ. Let regular hyperboloidal initial data for Einstein's equations, in the sense of Friedrich (1983), be given on Σ. (Actually, we consider the time-reverse of Friedrich's case. Also, Friedrich took Σ to be the ball, but this is unnecessary here.) These data determine a space-time for which a portion of $\mathscr{I}^+$ to the future of Z exists. The past Cauchy horizon N of Σ in this space-time is (near Z) a smooth null hypersurface and the metric and its derivatives on this hypersurface attain limits which determine smooth asymptotic characteristic data. Properties (a)–(e) hold.

Fix a Bondi system at $\mathscr{I}^+$ in which Z corresponds to $u = 0$. We may then work out, from the space-time above, the Bondi shear and all its u-derivatives at Z. Let $\mathscr{N} = (-\infty, 0] \times S^2$, and let σ be any smooth extension of the Bondi shear to $\mathscr{N}$. Then a theorem of Rendall (1990) shows that the doubly-characteristic initial-value problem on $\mathscr{N} \cup N$ is well-posed locally relative to Z, and so there is a locally unique space-time inducing these data in a neighborhood of Z as well.

The space-times under consideration are then determined (at least in a neighborhood of $\mathscr{H}$) by the regular hyperboloidal initial data in Σ and the shear on $\mathscr{N}$. Therefore such data ought to form the manifold of the phase space. Since the admissible shears depend on the hyperboloidal data, we see that the phase space has a bundle structure, with the spaces of allowed shears fibering over the hyperboloidal data. We shall refer to the data on Σ as the *final state*, and the shears compatible with a final state as the *radiation data*. We are here concerned exclusively with a fibrewise analysis: we fix the final state, and consider the possible radiation data compatible with it.

At present, it is not known under what conditions a final state will admit a radiative datum for which a space-time exists in a neighborhood of $\mathscr{H}$, that is, for which condition (f) holds. However, for local analysis on the phase space, it is not necessary to know exactly which data are admissible. We simply need to know that sufficiently small perturbations of a space-time satisfying (a)–(f) will

also satisfy (a)–(f) , and that the map from data to space-times is smooth. These are guaranteed by stability results of Rendall (1992).

4.2 Developments

We now show that the data under consideration have well-defined maximal globally hyperbolic extensions. The argument is quite similar to that for the standard Cauchy problem, and so a rapid treatment will be given. However, the force of the present results is somewhat stronger from that for the standard Cauchy problem. In the standard case, a development of the data by definition contains the entire initial surface as a Cauchy surface. Once a local existence and uniqueness theorem is established, this is no loss of generality, since the finite speed of propagation guarantees the existence of such developments. In the characteristic case, on the other hand, the local existence and uniqueness results do *not* guarantee the existence of developments including the entire initial surface. Therefore one must consider developments with different domains on the initial surface, and the maximality one seeks is in part a maximality with respect to such domains.

This result is somewhat stronger than what will be needed below, since in the sequel we shall simply posit that the solution extends to the whole of the initial-data surface. However, we recall that our aim is to give a construction that is robust, and for this reason the result is of interest here.

In what follows, we assume a familiarity with the differential topology of space-times, as described in Penrose (1972). A few comments are necessary, since the standard expositions of the theory apply to finite space-time, and we shall have to consider the boundary as well.

The notions of chronological and causal precedence are conformally invariant and so have an interpretation in (M, g_{ab}) independent of the factor Ω used in defining the conformal completion. Thus the chronological and causal futures and pasts I^+, J^+, I^-, J^- have evident meanings. The concept of an *achronal set* is also well-defined. Then the domains of dependence of an achronal set are also defined. We shall soon show that $\mathscr{H}$ is an achronal set in the solutions of interest; then we shall be interested in the domains

$$D^{\pm}(\mathscr{H}) = \{x \in \widehat{M} | \text{every timelike curve in } M \text{ through } x$$
$$\text{which is past-(future)-endless in } \widehat{M} \text{ meets } \mathscr{H}\}$$

and $D(\mathscr{H}) = D^+(\mathscr{H}) \cup D^-(\mathscr{H})$.

In what follows, when condition (f) is not assumed, the sets $\mathscr{H}$ and $\mathscr{N}$ are understood to include only the portion of $\mathscr{I}^+$ which exists.

Lemma. *Let $(\widehat{M}, \widehat{g}_{ab})$ be a time-oriented space-time satisfying (a)–(e), and suppose $\mathscr{H}$ is achronal. Then $\mathscr{M} = Int\, D(\mathscr{H})$ is globally hyperbolic. Also $\widehat{\Sigma} \subset \mathscr{M}$, and every point on $\mathscr{N}$ is the endpoint of a null geodesic on $\mathscr{M}$.*

Proof. That $\mathscr{M}$ is strongly causal follows from Lemma 4.16 in Penrose (1972). That $J^+(p) \cap J^-(q)$ is compact for $p, q \in \mathscr{M}$ follows from Proposition 5.20 and Theorem 6.5 there. The remaining assertions are elementary. □

Definition. *Let a final state on Σ and a radiation datum compatible with it on $\mathscr{N}$ be given. A* development *of these data is an oriented, time-oriented space-time*

$(\widehat{M}, \widehat{g}_{ab})$ *for which properties (a)–(e) hold (and inducing the correct data on Σ and that portion of $\mathscr{I}^+$ which exists), such that $\mathscr{H}$ is g_{ab}-achronal and $\widehat{M} = D(\mathscr{H})$.*

Lemma. *For any final state and radiation datum compatible with it, a development exists.*

Proof. The results of Friedrich together with standard results on the Cauchy problem show that there is a maximal globally hyperbolic development of the data on $\widehat{\Sigma}$ for which a portion of $\mathscr{I}^+$ at and to the future of Z exists, and for which the past Cauchy horizon is (near Z) a smooth null hypersurface meeting $\mathscr{I}^+$ transversely in Z. Moreover the field attains smooth limits on this hypersurface, which are the correct characteristic initial data. Let us denote the manifold of this space-time by M_F.

Rendall's theorem shows that the initial-value problem with the data on the past horizon and those on $\mathscr{N}$ is well-posed (locally at Z). Let M_R be a space-time produced by this theorem, with S the strip of the past Cauchy horizon of $\widehat{\Sigma}$ serving as one of the characteristic data surfaces. (We do not include Z in S.)

Now $M_F \cup S \cup M_R$ is an oriented, time-oriented space-time satisfying (a)–(e). Also no point in M_R is in the future of any point in M_F. It follows that $\mathscr{H}$ is achronal in this space-time. The remaining claims will follow from the previous lemma. □

Theorem 4.1. *There exists a unique development which is maximal with respect to inclusion.*

(Here "unique" and "inclusion" are to be understood in the sense of the natural isomorphism classes of developments.)

Proof. Fix a Bondi coordinate system on $\mathscr{I}^+$ in which the cut Z is given by $u = 0$. We shall consider developments meeting $\mathscr{I}^+$ in different sets, and we shall want to compare their sets $\mathscr{N}$ of points on $\mathscr{I}^+$ to the past of Z. We begin by examining the forms of these sets.

Condition (c), that $\mathscr{N}$ be diffeomorphic to $\mathbb{R} \times S^2$ with null generators being the "$\mathbb{R}$" factors, may be expressed more formally by saying that there is a diffeomorphism $\mathbb{R} \times S^2 \to \mathscr{N}$ of the form.

$$\big(x, (\theta, \varphi)\big) \mapsto \big(u(x, \theta, \varphi), (\theta, \varphi)\big).$$

The function $\alpha : S^2 \to [-\infty, 0)$ given by $\alpha(\theta, \varphi) = \inf_x u(x, \theta, \varphi)$ is, as an infimum of continuous functions, upper semicontinuous. Conversely, it is not hard to show that for any upper semicontinuous function $\alpha : S^2 \to [-\infty, 0)$, the set

$$\mathscr{N}(\alpha) = \{(u, \theta, \varphi) \in \mathscr{I}^+ \mid 0 > u > \alpha(\theta, \varphi)\}$$

satisfies condition (c). (We do not claim that there is a development for every such set, however.)

If a development exists corresponding to a function α, then there is a unique maximal globally hyperbolic development corresponding to this function. This argument follows that for the standard Cauchy problem, cf. Hawking and Ellis (1973). (One must be a little careful because of the freedom in choosing the conformal factor. However, since there are geometric ways of fixing this in a neighborhood of $\mathscr{I}^+$, there is no real difficulty.)

Now let A be the set of α's achievable from the given data. If $\alpha_1, \alpha_2 \in A$ with $\alpha_1 \leqq \alpha_2$ (for all θ, φ), then the maximal globally hyperbolic development $(\widehat{M}, \widehat{g}_{ab})_1$ corresponding to α_1 must include the one $(\widehat{M}, \widehat{g}_{ab})_2$ corresponding to α_2. To see this, note that we can glue to $(\widehat{M}, \widehat{g}_{ab})_2$ a portion of $(\widehat{M}, \widehat{g}_{ab})_1$ on the region extending inwards from $\mathscr{N}(\alpha_1) - \mathscr{N}(\alpha_2)$ to get a development, which must then be included in $(\widehat{M}, \widehat{g}_{ab})_1$.

Let $\underline{\alpha}(\theta, \varphi) = \inf_{\alpha \in A} \alpha(\theta, \varphi)$. Since $\mathscr{N}(\underline{\alpha}) = \bigcup_{\alpha \in A} \mathscr{N}(\alpha)$, we can construct a solution to the initial-value problem in a neighborhood of $\Sigma \cup \mathscr{N}(\underline{\alpha})$ by patching together solutions in neighborhoods of the sets $\Sigma \cup \mathscr{N}(\alpha)$. From this we can construct a development corresponding to $\underline{\alpha}$, and so there is a unique maximal globally hyperbolic development $(\widehat{M}, \widehat{g}_{ab})$ corresponding to this function. By the argument of the previous paragraph, it is maximal among the developments for arbitrary $\alpha \in A$. $\square$

5. The Manifold Underlying the Phase Space

In this section, we define the function space representing the radiative modes of the gravitational field. As discussed above, this ought to be a fibre of the total phase space, with the base point representing a given final state. We must expect that there will be some restrictions necessary to have a rigorous isolation – in the sense of some sort of splitting of the symplectic form compatible with the fibre structure – of the radiative modes from the internal ones, since there *is* a coupling between the two: the data must match at the cut Z. In the realistic ("generic") physical system, one would not except an exact isolation to be possible. However, to develop a conceptual framework, the first step is to give a clear mathematical description of the case where the isolation is possible; after that one can treat what will amount to boundary (at $Z = \partial \mathscr{N} = \partial \Sigma$) effects. In this paper, we shall be explicitly concerned with the first step only. On the other hand, the boundary effects ought to be deducible from our approach: see the comments following the proof of Theorem 6.1.

In order to provide this clean isolation of the radiative modes, we shall have in mind a system which is, as far as the radiative data can reflect, stationary except for a finite range of values of the retarded time, u, which correspond to cuts prior to Z. Thus we shall assume to shear is purely electric and locally independent of u except for a compact region on $\mathscr{N}$. This restriction is really made for the sake of definiteness and could be relaxed considerably, in the following senses:
(a) In this section, we do not need the vanishing of any u-derivatives of σ at Z. For Theorem 6.1, below, we shall need only the vanishing of the first u-derivative there. (This is only used in the first paragraph of the proof.)
(b) Finite rates of decay of $\partial \sigma / \partial u$ as $u \to -\infty$ could easily be incorporated here and in all that follows.
(c) The assumption that the shear be electric near Z and in the far past could be dropped. The modification necessary to treat this are all straightforward; the only thing to bear in mind is that the "gauge freedom" is still purely electric. The significance of dropping this assumption is discussed briefly in Helfer (1994); see also Newman and Penrose (1966).

Finally, the assumption of smoothness of class C^∞ in what follows could be weakened considerably.

The allowed space Γ of σ's will therefore be the set of those agreeing at Z with that inferred from the hyperboloidal data on Σ, and which are pure gauge near Z and sufficiently far in the past. (To be pure gauge means $\sigma = \eth^2\alpha$ for some real function α independent of u; the functions α may be different near Z and in the far past.) We remark that there are many data on Σ admitting shear data which are pure gauge near Z (take the data on Σ to be those for a stationary space-time, for example). The existence theory discussed in the previous section guarantees that from any given final state and any shear data of the type under consideration, a space-time will exist inducing those shear data on a neighborhood of Z. At present, it is not known under what conditions a space-time will exist for which the generators of $\mathscr{I}^+$ are complete to the past and the shear data on all of $\mathscr{N}$ is induced. However, as noted in Sect. 4.1, the stability results of Rendall (1992) show that for any space-time for which this *does* hold, sufficiently small perturbations of the shear data also give rise to such space-times.[9] Thus the physical phase space will really be an open set in Γ. This will be adequate for the analysis in Sects. 6 and 7, which is local on the physical phase space.

In order to have a uniform treatment of the gauge freedom, it will be convenient not to fix the supertranslational freedom at first. Since we are considering shears which are pure gauge at Z, this will have the effect of allowing arbitrary such shears.

Fix a diffeomorphism $\mathscr{N} \to (-\infty, 0] \times S^2$. Let Γ be the set of all smooth functions σ of type $\{3, -1\}$ on $\mathscr{N}$ whose u-derivatives vanish outside a compact set, which are electric at $u = 0$ and for u sufficiently negative. The most convenient way to topologize Γ is as follows. Choose any function f (of type $\{0, 0\}$) on $\mathscr{N}$ which is smooth, identically -1 for sufficiently negative u and identically $+1$ close to $u = 0$. The map

$$C^\infty(S^2, \{3, -1\})_{\text{electric}} \oplus C^\infty(S^2, \{3, -1\})_{\text{electric}} \oplus C_0^\infty(\mathscr{N}, \{3, -1\}) \to \Gamma \qquad (19)$$

given by

$$(\alpha, \beta, \gamma) \mapsto \alpha f + \beta + \gamma \qquad (20)$$

is an isomorphism of (algebraic) vector spaces. (Here the electric elements of $C^\infty(S^2, \{3, -1\})$, as the kernel of a certain differential operator, form a Fréchet space. The subscript 0 denotes compact supports.) Topologize the right-hand side so that this is a homeomorphism.

Lemma. *The topology so defined is independent of the choice of f.*

Proof. Let f' be a second such function. Then we suppose that we have

$$\alpha f + \beta + \gamma = \sigma = \alpha' f' + \beta' + \gamma' . \qquad (21)$$

Considering sufficiently small u, we see that we must have $-\alpha + \beta = -\alpha' + \beta'$; similarly, for sufficiently large u, we must have $\alpha + \beta = \alpha' + \beta'$. Thus $\alpha = \alpha'$ and $\beta = \beta'$, and so $\gamma' = \gamma + \alpha(f - f')$. It follows that the transition function

$$C^\infty(S^2, \{3, -1\})_{\text{electric}} \oplus C^\infty(S^2, \{3, -1\})_{\text{electric}} \oplus C_0^\infty(\mathscr{N}, \{3, -1\})$$
$$\to C^\infty(S^2, \{3, -1\})_{\text{electric}} \oplus C^\infty(S^2, \{3, -1\})_{\text{electric}} \oplus C_0^\infty(\mathscr{N}, \{3, -1\})$$

---

[9] Here "small" is in the sense of the topology used here; in fact, we use a topology somewhat finer than necessary.

given by

$$(\alpha, \beta, \gamma) \mapsto (\alpha', \beta', \gamma') = (\alpha, \beta, \gamma + \alpha(f - f'))$$

is a homeomorphism. $\square$

Thus the inverses of the maps (19), as f varies, form an atlas for Γ. The transition function identified in the proof above is smooth, so we have given Γ the structure of a smooth manifold.

The following are immediate: (a) the manifold is modeled on a sequentially complete locally convex topological vector space; (b) this vector space is reflexive and its dual may be identified with a suitable space of distributions; (c) the sequence

$$0 \rightarrow \text{Gauge Freedom} \rightarrow \text{Unreduced Data} \rightarrow \text{Bondi News Functions} \rightarrow 0 \quad (22)$$

given by

$$0 \rightarrow C^\infty(S^2, \{3, -1\})_{\text{electric}} \rightarrow \quad \Gamma \quad \rightarrow C_0^\infty(\mathscr{N}, \{2, 0\}) \rightarrow 0$$
$$\beta \qquad\qquad \mapsto (0, \beta, 0) \qquad\qquad\qquad\qquad\qquad (23)$$
$$(\alpha, \beta, \gamma) \quad \mapsto \quad \alpha \partial_u f + \partial_u \gamma$$

is an exact sequence of topological vector spaces; (d) the projections from Γ to $C^\infty(S^2, \{3, -1\})_{\text{electric}}$ giving the future and past limits of the shear are smooth.

Property (c) implies that the smooth structure of Γ is invariant under passive supertranslations. It is not hard to show too that the smooth structure is invariant under passive Lorentz motions, so that it is invariant under all passive BMS motions. For active motions, we must consider the family of Γ's as the cut Z varies; the naturality of our construction guarantees the appropriate covariance.

6. The Symplectic Form

Ideally, one would like to identify a symplectic form on the phase space we have constructed by starting with the usual symplectic form got by integrating linearized solutions over Cauchy surfaces (Arnowitt, Deser and Misner 1962, Regge and Teitelboim 1974, Chernoff and Marsden 1974), and working out what this gives in terms of our data. Formally, the way to proceed is clear. Since the three-form which is integrated is closed, one should deform the Cauchy surface to the M-shaped surface $\mathscr{H}$. This approach was outlined in Ashtekar and Magnon-Ashtekar (1982). In our notation, the symplectic form they identified is given by

$$\omega(s_1, s_2) = (8\pi G)^{-1} \int_{\mathscr{N}} [\dot{s}_1 \bar{s}_2 - \dot{\bar{s}}_2 s_1] \, du \, d\mathscr{S} + \text{conjugate} . \quad (24)$$

Since the surfaces involved are not compact, some justification of the invariance of the integral under deformations is necessary. The difficulty occurs at $u = -\infty$. The difference between integrating over the M-shaped surface and the Cauchy surface amounts to what may be called a leakage integral (at $u = -\infty$, which presumably corresponds to spacelike infinity). A comprehensive treatment of such leakages would require stronger asymptotic estimates than are currently established.[10] Therefore we shall give a proof of the correctness of the symplectic form by less direct means. This proof has the advantage of avoiding all gauge problems.

We first verify that ω is a symplectic form.

---

[10] Roughly speaking, one would want to show that the asymptotic expansion of the metric held uniformly as $u \rightarrow -\infty$ to $O(1/r)$. The relation between null and spacelike infinity is a major outstanding problem in the existence theory for Einstein's equations.
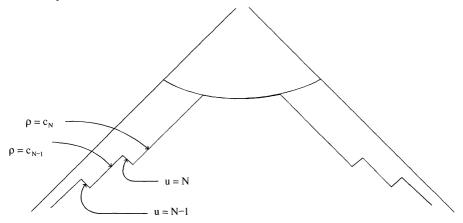
Fig. 2. Construction of the Cauchy surface in Theorem 6.1. A cross-section is shown

Proposition. *The form ω is smooth, closed and weakly non-degenerate.*

Proof. It is smooth and closed because it is constant in the chart. Integrating by parts, we have

$$\omega(s_1, s_2) = (4\pi G)^{-1} \int_{\mathscr{N}} \dot{s}_1 \overline{s_2}\, du\, d\mathscr{S} - (8\pi G)^{-1} \oint \overline{s_2} s_1\, d\mathscr{S} + \text{conjugate} ,$$

where $\oint$ denotes the difference between the boundary integrals at $u = 0$ and $u = -\infty$. Suppose, for some fixed s_1, this vanishes for all s_2. By taking s_2 to be a bump function in u times an arbitrary angular function (with values in $\{3, -1\}$), we have $\dot{s}_1 = 0$. However, then we have

$$\omega(s_1, s_2) = -(8\pi G)^{-1} \oint [\overline{s_2}] s_1\, d\mathscr{S} + \text{conjugate} ,$$

where $\oint$ now is the usual angular integral. This implies that the electric part of s_1 is zero. However, since s_1 is constant, it must be purely electric. $\square$

Theorem 6.1. *Let $(\widehat{M}, \widehat{g}_{ab})$ be the maximal globally hyperbolic development of data on $\mathscr{H}$, and let s_1, s_2 be tangent vectors at the corresponding point in Γ. Let h_1, h_2 be linearized solutions to Einstein's equations inducing the data s_1, s_2 on $\mathscr{N}$ (and pure gauge on Σ). Then there is a Cauchy surface S in $(\widehat{M}, \widehat{g}_{ab})$ such that $\omega(s_1, s_2) = \omega_S(h_1, h_2)$, where $\omega_S(h_1, h_2)$ is the usual $3 + 1$ symplectic form determined by integrating over S.*

Proof. We first show that $\omega_{\widehat{\Sigma}}(h_1, h_2) = 0$. Let $h_{1ab} = \widehat{\nabla}\xi_{1b} + \widehat{\nabla}_b \xi_{1a}$, and similarly for h_2 and ξ_2. Let α_1, α_2 be the supertranslations induced by ξ_1, ξ_2 at Z. Then $\omega_{\widehat{\Sigma}}(h_1, h_2) = \langle h_1, dP_2 \rangle$, where P_2 is the supermomentum at Z determined by α_2 (and $\langle \cdot, \cdot \rangle$ is pairing in the $3 + 1$ phase space). Then $\langle h_1, dP_2 \rangle$ is the change in this supermomentum under a supertranslation by α_1. However, the vanishing of the derivatives of σ at Z implies this is zero.

Assume the Bondi system has been chosen so that Z is given by $u = 0$. Fix the coordinate $\rho = \Omega$ as in Penrose and Rindler (1986), Sect. 9.8. (The relevant effect of this is to ensure that, for any fixed u, the surfaces of constant ρ are space-like for small enough positive ρ.) For negative integers n, consider the null hypersurface extending orthogonally inwards from $u = n$. We choose a smooth portion C_n of this

hypersurface given by $\rho \leqq c_n$ for some positive c_n. (We shall place restrictions on c_n in what follows.)

Now assume n is so negative, say $n \leqq N$, that s_1, s_2 are pure gauge for $u \leqq n$. Then $\omega(s_1, s_2) = \omega_{K_n}(h_1, h_2)$, where

$$K_n = \{p \in \widehat{\Sigma} \mid \rho(p) \geqq c_n\} \cup \{p \in \widehat{M}, \mid \rho(p) = c_n \text{ and } n \leqq u(p) \leqq 0\} \cup C_n .$$

For each $n \leqq N$, consider the contribution to the integral $\omega_{K_n}(h_1, h_2)$ from C_n. By choosing c_n small enough, we may make this contribution arbitrarily small; indeed, we may make the integral of the absolute value of the integrand arbitrarily small. Choose c_n small enough so that this latter is $< 1/|n|$.

Now construct S by starting with

$$\{p \in \widehat{\Sigma} \mid \rho(p) \geqq c_N\}$$

and for $n \leqq N$ alternately joining the surfaces

$$\{p \in \widehat{M} \mid \rho(p) = c_n, \; n - 1 \leqq u \leqq n\}$$

with

$$\{p \in \widehat{M} \mid p \in C_{n-1} \text{ and } c_{n-1} \leqq \rho(p) \leqq c_n\} .$$

This surface is by construction everywhere spacelike or null, and $\omega(s_1, s_2) = \omega_S(h_1, h_2)$. We must show it is a Cauchy surface.

A timelike curve starting from a point on $S - \Sigma$ will have ρ strictly decreasing and u strictly increasing. Since ρ is a monotonically increasing function of u on $S - \Sigma$, this curve cannot meet $S - \Sigma$ at a second point. If it meets Σ, it must do so on the set of points with $\rho < c_N$, that is, on $\Sigma - S$. So there is no timelike curve from $S - \Sigma$ to S. Since we already know Σ is achronal, the set S must be achronal.

Finally, let us show that we can ensure that every null geodesic meets S. We know every null geodesic meets $\mathscr{H}$. Let Y be the space of null geodesics, topologized as bundle over $\mathscr{H}$. (So the fibre is S^2 over a point in $\widehat{\Sigma}$ and $\mathbb{R}^2$ over a point in $\mathscr{N}$.) Write $Y = \bigcup_{n=1}^{\infty} Y_n$, where each Y_n is compact and has base contained in

$$\Sigma \cup \{p \in \mathscr{N} \mid u > -n\} .$$

Thinking of the elements of Y_n as past-directed null geodesics from $\mathscr{H}$, they must meet K_n in a compact set for which ρ is bounded away from zero. By choosing $c_n \to 0$ quickly enough as $n \to -\infty$, then, we can ensure that every null geodesic meets S. $\square$

We remark that we did not really need the vanishing of $\partial\sigma/\partial u$ in a neighborhood of Z; we used only the fact that it vanished at Z, and that fact only in the first paragraph of the proof. If this derivative did not vanish, then one would have an additional contribution $\langle h_1, dP_2 \rangle$ to the symplectic form, representing a mixing of the internal and radiative modes, which can be viewed as a boundary effect due to leakage of radiation at Z.

It should be clear that the proof could be modified to accomodate finite rates of decay of $\partial\sigma/\partial u$ as $u \to -\infty$.

Note that the Cauchy surface may depend on s_1 and s_2; on the other hand, no gauge restrictions were placed on h_1 and h_2 except that they should induce the correct data on $\mathscr{H}$. These issues are related: only if we make some choice of gauge in the interior of space-time can we expect to control the integrals uniformly.

However, gauge-fixing is not the only obstacle to getting some sort of uniform control; the relation between null infinity and spatial infinity would have to be understood.

7. Gauge Freedom

The sectors for the gravitational phase space are defined by considering the jumps $[\sigma]$, which are purely electric functions on the sphere.[11] The *classical sector* is $[\sigma] = 0$; the others are *infrared sectors* (Ashtekar 1987). They play an important role in the theory of angular momentum.

The interpretation of the sectors for gravity turns out to be subtly different from that for electromagnetism.

Theorem 7.1. *The sectors are the level surfaces of the function* $C : \Gamma \rightarrow Super$-*translations* given by*

$$C_\alpha = (8\pi G)^{-1} \oint [\sigma] \eth'^2 \alpha \, d\mathscr{S} + \text{conjugate} \ .$$

This function descends to a moment map for the group Supertranslations/Trans-lations *acting by* $\sigma \mapsto \sigma + \eth^2 \alpha$.

Proof. The first assertion follows from the fact that $C^\infty(S^2, \{3, -1\})$ and $C^\infty(S^2, \{-3, 1\})$ are dual, together with the surjectivity of the map $\eth'^2 : C^\infty(S^2, \{1, 1\}) \rightarrow C^\infty(S^2, \{-1, 3\})$.

Certainly the action $\sigma \mapsto \sigma + \eth^2 \alpha$ descends to the quotient group. Let V_α be the vector field generating this motion. Then we have

$$\omega(s, V_\alpha) = (8\pi G) \int_{\mathscr{N}} \dot{s} \eth'^2 \alpha \, du \, d\mathscr{S} + \text{conjugate}$$

$$= (8\pi G)^{-1} \oint [s] \eth'^2 \alpha \, d\mathscr{S} + \text{conjugate}$$

$$= \langle s, dC_\alpha \rangle$$

□

Notice that the action here is *not* the standard active BMS motion, Eq. (7), unless $\dot{\sigma} = 0$. Although formally the same as the result of a passive BMS motion, Eq. (8), we do *not* accompany this motion by a change in the Bondi parameter, so this is *not* a passive BMS motion, either. We shall see in a subsequent paper that this foliation, and the difference between the action here and the standard BMS actions, goes some way to resolving an old mystery: why approaches to defining angular momentum at cuts of $\mathscr{I}^+$ by BMS motions give unphysical answers.

---

[11] If shears which have magnetic components at Z and as $u \rightarrow -\infty$ are allowed, then the sectors are labeled by the electric parts of the jumps.

References

Arnowitt, R.L., Deser, S., Misner, C.W.: Canonical Analysis of General Relativity. In: Recent Developments in Relativity. New York: Pergamon (1962)

Ashtekar, A.: Asymptotic quantization. Naples: Bibliopolis (1987)

Ashtekar, A., Magnon-Ashtekar, A.: On the Symplectic Structure of General Relativity. Commun. Math. Phys. **86**, 55–68 (1982)

Ashtekar, A., Narain, K.S.: Infrared Problems and Pensrose's Null Infinity. International Conference on Mathematical Physics, Berlin (1981)

Ashtekar, A., Streubel, M.: Symplectic Geometry of Radiative Modes and Conserved Quantities at Null Infinity. Proc. R. Soc. Lond. **A376**, 585–607 (1981)

Bergmann, P.G., Goldberg, I.: Dirac Bracket Transformations in Phase Space. Phys. Rev. **98**, 531–538 (1955)

Bondi, H.: Gravitational Waves in General Relativity. Nature **186**, 535 (1960)

Bondi, H., van der Burg, M.G.J., Metzner, A.W.K.: Gravitational Waves in General Relativity. VII. Waves from Axi-symmetric Isolated Systems. Proc. R. Soc. Lond. **A269**, 21–52 (1962)

Chakravarty, S., Ivancovich, J., Newman, E.T.: "Infrared" Maxwell fields. Gen. Rel. Grav. **18**, 633–40 (1986)

Dirac, P.A.M.: Lectures on quantum mechanics. New York: Belfer Graduate School of Science, Yeshiva University (1964)

Friedrich, H.: Cauchy problems for the conformal vacuum field equations in general relativity. Commun. Math. Phys. **91**, 444–472 (1983)

Friedrich, H.: Asymptotic structure of space-time. In: A.I. Janis, J.R. Porter (eds.): Recent advances in general relativity. Boston-Basel-Berlin: Birkhäuser (1992)

Hawking, S.W., Ellis, G.F.R.: The large scale structure of space-time. Cambridge: University Press (1973)

Helfer, A.D.: The angular momentum of gravitational radiation. Phys. Lett. **A150**, 342–44 (1990)

Helfer, A.D.: Twistors and the BMS group. Class. Quantum Grav. **11**, 733–55 (1994)

Horowitz, G.T., Tod, K.P.: A relation between local and total energy in general relativity. Commun. Math. Phys. **85**, 429–447 (1982)

Jauch, J.M., Rohrlich, F.: The theory of photons and electrons. New York: Springer-Verlag (1976)

Ludvigsen, M., Vickers, J.A.G.: The positivity of the Bondi mass. J. Phys. **A14**, L389–91 (1981)

Marsden, J., Weinstein, A.: Reduction of a symplectic manifolds with symmetry. Rep. Math. Phys. **5**, 121–130 (1974)

Moncrief, V.: Spacetime symmetries and linearization stability of the Einstein equations. I. J. Math. Phys. **16**, 493–498 (1975)

Moncrief, V.: Spacetime symmetries and linearization stability of the Einstein equations. II. J. Math. Phys. **17**, 1893–1902 (1976)

Newman, E.T., Penrose, R.: An approach to gravitational radiation by a method spin coefficients. J. Math. Phys. **3**, 896–902 (Errata **4** (1963) 998) (1962)

Newman, E.T., Penrose, R.: Note on the Bondi-Merzner-Sachs group. J. Math. Phys. **7**, 863–70 (1966)

Penrose, R.: Asymptotic properties of fields and space-times. Phys. Rev. Lett. **10**, 66–8 (1963)

Penrose, R.: Techniques of Differential Topology in Relativity. Philadelphia: SIAM (1972)

Penrose, R.: Quasi-local mass and angular momentum in general relativity. Proc. R. Soc. Lond. **A381**, 53–63 (1982)

Penrose, R., Rindler, W.: Spinors and space-time. Cambridge: University Press (1984–6)

Regge, T., Teitelboim, C.: Role of surface integrals in the Hamiltonian formulation of general relativity. Ann. Phys. **88**, 286–318 (1974)

Rendall, A.D.: Reduction of the characteristic initial value problem to the Cauchy problem and its applications to the Einstein equations. Proc. R. Soc. Lond. **A427**, 221–239 (1990)

Rendall, A.D.: Stability in the characteristic initial value problem. In: Z. Perjés (ed.): Relativity today. Commack, New York: Nova Science (1992)

Sachs, R.K.: Gravitational waves in general relativity. VIII. Waves in asymptotically flat space-time. Proc. R. Soc. Lond. **A270**, 103–26 (1962a)

Sachs, R.K.: Asymptotic symmetries in gravitational theory. Phys. Rev. **128**, 2851–64 (1962b)

Schoen, R., Yau, S.-T.: Proof that the Bondi mass is positive. Phys. Rev. Lett. **48**, 369–71 (1982)