Propagation of gravitational waves through pressureless matter

J Ehlers†, A R Prasanna‡ and R A Breuer§

† Max-Planck-Institut für Physik und Astrophysik, Institut für Astrophysik, D-8046 Garching, West Germany
‡ Physical Research Laboratory, Ahmedabad, India
§ GEO Magazine, D-2000 Hamburg 36, West Germany

Abstract. We investigate the propagation of small-amplitude gravitational waves through pressureless matter ('dust'). After establishing the local linearisation stability of Einstein's equation for dust about any of its solutions we use the WKB method to study the locally plane, linearised perturbations of an arbitrary background dust spacetime asymptotically for small wavelengths. The dispersion relation exhibits two modes. One is simply degenerate and represents gravitational waves, whereas the other is doubly degenerate and describes density and vorticity perturbations. The waves are shown to propagate along the null geodesics of the background; in leading WKB order their amplitudes behave as *in vacuo*. The rays associated with the matter mode are the worldlines of the background dust. In leading order the perturbations of density and 4-velocity vanish for both modes.

1. Introduction

The propagation of gravitational waves through matter has been treated in several papers under various assumptions with different methods of approximation. Sometimes contradictory conclusions have been reached, particularly with respect to the dispersion relation and the related question of Landau damping (see, e.g. [1] and the references therein). Therefore we have begun to study these questions again systematically by means of generally applicable methods rather than with *ad hoc* approximations valid in special cases only

In this paper which is based on [2] and [3] we use the simplest matter model, dust, and investigate generally locally plane, short wave, linearised perturbations of an arbitrary background dust spacetime by means of the WKB (or geometrical optics) approach. In order to justify the use of linearised perturbations we prove in § 2 that the Einstein gravitational field equation for dust is locally linearisation stable at any of its solutions. This fact is established by means of Choquet-Bruhat's treatment of the initial value problem which is also used to exhibit which unconstrained data can serve to parametrise dust solutions.

In § 3 we first set up the equations for linearised perturbations without imposing a gauge condition, and express the matter perturbations in terms of the metric perturbation. We then show that one may, without loss of generality, impose the Landau (or radiation) gauge on the metric perturbation, and continue to work in that gauge.

In § 4 we specialise the linearised metric perturbations to rapidly oscillating, locally plane ones, and apply the WKB method to determine the dispersion relation, the phase hypersurfaces, rays and polarisation states of the two possible modes. It turns out that

the leading gravitational wave amplitudes have the same transverse-traceless structure as *in vacuo*, and that they propagate normally along the null geodesics of the background spacetime. In contrast, the spacetime dependence of the amplitudes of the matter perturbations is not determined by first-order transport equations; its determination requires the calculation to be extended beyond the leading order.

In this paper we do not take into account the back reaction of the perturbations on the background spacetime but consider the latter as given, i.e. the perturbations are treated as test fields. Let a background field be denoted as (ρ, u^a, g_{ab}), and a linearised perturbation as $(\hat{\rho}, \hat{u}^a, \hat{g}_{ab})$, and let ε denote the small ratio between the scale of variation of the perturbed variables to that of the background. According to linearisation stability, $(\rho + \delta\hat{\rho}, u^a + \delta\hat{u}^a, g_{ab} + \delta\hat{g}_{ab})$ will approximate a solution of the field equation, at least in a compact part of spacetime, provided the constant numerical factor δ is sufficiently small. Unfortunately no rigorous effective criterion appears to be known to tell us when δ is 'sufficiently' small. However, by estimating the orders of magnitude of various terms in the manner of Isaacson [4] one can guess the following. If one works in lowest WKB order, i.e. if one takes into account the leading-order amplitudes of the perturbed quantities only, δ has to be taken to be of higher order than ε^2, say $\delta = \varepsilon^3$, since for larger amplitudes $\delta \geq \varepsilon^2$ the error due to the restriction to linearised perturbations is expected to be larger than the leading WKB amplitudes. It appears to be sensible to take into account higher-order (in the WKB sense) amplitudes of linearised perturbations only if one imposes correspondingly stronger restrictions on the overall size δ of these perturbations, or if one also adds non-linear perturbations. The influence of matter on gravitational wave amplitude transport will show up at the next WKB order which we intend to work out on the basis developed below.

1.1. Notation and conventions

We set the speed of light equal to unity, $c = 1$, and write $\kappa := 8\pi G$, G denoting Newton's constant of gravity. The signature of the metric is chosen as $+++-$, and the curvature tensors, etc, are as in [5]. As usual, $A_{(bc)} := \frac{1}{2}(A_{bc} + A_{cb})$. Single and repeated covariant derivatives are indicated by ∇_a, ∇_{ba}, etc. The sign $\sim$ stands for asymptotic equality for $\varepsilon \to 0$. Latin indices range and sum over $0, 1, 2, 3$, greek ones over $1, 2, 3$.

2. Dynamics of dust

Consider a spacetime (M, g_{ab}) filled with dust, obeying the field equation

$$R^{ab} = \kappa\rho(u^a u^b + \tfrac{1}{2}g^{ab}) \tag{1}$$

with

$$\rho > 0 \tag{2}$$

and

$$u_a u^a = -1. \tag{3}$$

The velocity and density fields satisfy

$$u^b \nabla_b u^a = 0 \tag{4}$$

and

$$\nabla_a(\rho u^a) = 0. \tag{5}$$

In order to gain insight into the set of solutions of the system (1)–(3) one studies the Cauchy initial value problem. This requires the introduction of coordinate conditions. Locally it is always possible to introduce harmonic coordinates defined by

$$\mathfrak{g}^{ab}{}_{,b} = 0. \tag{6}$$

With respect to them, (1) takes the 'reduced' form

$$-\tfrac{1}{2}\mathfrak{g}^{cd}\mathfrak{g}^{ab}{}_{,cd} + H^{ab} = \kappa g\rho(u^a u^b + \tfrac{1}{2}g^{ab}) \tag{7}$$

where $g := -\det(g_{ab})$, $\mathfrak{g}^{ab} := \sqrt{g}g^{ab}$, and the expressions H^{ab} depend on $\mathfrak{g}^{cd}$, $\mathfrak{g}^{cd}{}_{,e}$ only. The non-linear system (4), (5), (7) of partial differential equations is not strictly hyperbolic and therefore not suitable for studying the initial value problem. However, Choquet-Bruhat [6] recognised long ago that this system is equivalent to a strictly hyperbolic, diagonal Leray system. (For a concise formulation of Leray's theorem see, e.g., [7].) The latter is obtained by keeping (4) and (5), but substituting for (7) the equation obtained by differentiating Einstein's equation (1) in the direction of the dust worldlines, i.e. by applying the operator $u^a \nabla_a$ to both sides of (1), and simplifying the result by means of (4)–(6). The unknowns of the new system of 15 evolution equations are the 15 functions $\mathfrak{g}^{ab}$ (or, equivalently, g_{ab}), u^a and ρ. An initial data set for this system associated with a spacelike initial hypersurface $\Sigma := x^0 = 0$ consists of

$$\mathfrak{g}^{ab}, \mathfrak{g}^{ab}{}_{,c}, \mathfrak{g}^{ab}{}_{,cd}, u^a, u^a{}_{,b}, \rho. \tag{8}$$

The solution of the evolution equations determined by the data (8) obeys the original system (1)–(3) if and only if the data satisfy the constraint equations (3), (4), (6), (7) and the standard constraints

$$G^0_a = R^0_a - \tfrac{1}{2}\delta^0_a R = \kappa\rho u^0 u_a \tag{9}$$

as well as the conditions that the density ρ is positive, $\mathfrak{g}^{ab}$ is Lorentzian, and the initial hypersurface Σ is spacelike with respect to the metric g_{ab} (determined by $\mathfrak{g}^{ab}$). (For a proof, see [6] and [8].) Such a data set will be called admissible.

Every admissible data set can be obtained as follows [8]. Take 16 functions $\mathfrak{g}^{ab}$, $\dot{\mathfrak{g}}^{\alpha\beta}$ of three spatial coordinates x^α, with $\mathfrak{g}^{ab}$ Lorentzian, the corresponding $g_{\alpha\beta}$ Riemannian, and

$$\mathfrak{g}^{ab} G^0_a G^0_b < 0. \tag{10}$$

In (10) the G^0_a, which depend on $\mathfrak{g}^{cd}$, $\mathfrak{g}^{cd}{}_{,e}$ and $\mathfrak{g}^{cd}{}_{,\alpha\beta}$, are to be evaluated (on Σ, of course) by identifying $\mathfrak{g}^{\alpha\beta}{}_{,0}$ and $\mathfrak{g}^{a0}{}_{,0}$ with the data $\dot{\mathfrak{g}}^{\alpha\beta}$ and the derivative $-\mathfrak{g}^{\alpha\beta}{}_{,\beta}$ of data, respectively, and by taking the needed spatial derivatives of data. Next, solve the system (2), (3), (9) for ρ, u^a such that $u^0 > 0$; that is uniquely possible because of (10). Then, find $\mathfrak{g}^{ab}{}_{,00}$ (and therefore all remaining $\mathfrak{g}^{ab}{}_{,cd}$ by spatial differentiation) from (7), and finally use (4) to get $u^a{}_{,b}$.

Since an admissible data set determines a solution of (1)–(3) uniquely, one can 'parametrise' dust spacetimes in terms of the unconstrained data $\mathfrak{g}^{ab}$, $\dot{\mathfrak{g}}^{\alpha\beta}$ which are subject only to inequalities that are preserved under small changes of these data. The unconstrained data consist of 16 functions of three variables. A change from one harmonic gauge to another such gauge is locally uniquely determined by initial data for four solutions of a wave equation, i.e. by eight functions of three variables. A dust solution is thus intrinsically characterised locally by eight functions of three variables, corresponding to four degrees of freedom—two for the gravitational field, and two for pressureless matter—as expected.

We have reviewed the analysis due to Choquet-Bruhat ([6], see also [8]) because it enables us to establish that *Einstein's equation for dust is locally linearisation stable at any of its solutions.* To prove this we first use the fact that every solution of (1) can locally be expressed in harmonic coordinates, and can thus be reconstructed from unconstrained initial data via a Leray system by the method just described. Every solution $\hat{g}^{ab}$, $\hat{u}^a$, $\hat{\rho}$ of the linearised system can of course be expressed in terms of the chosen (background) harmonic coordinate system, too. Moreover, by a straightforward generalisation of a well known argument from linearised GR (given, e.g., on p 75 of [5]), one may assume without loss of generality that the perturbation also obeys the harmonicity condition,

$$\hat{g}^{ab}{}_{,b} = 0. \tag{11}$$

Suppose now that $\hat{g}^{ab}$, $\hat{u}^a$, $\hat{\rho}$ is a solution of the linearised equations corresponding to (1) and (3) at a background solution g^{ab}, u^a, ρ, and assume the coordinate conditions (6) and (11) to hold. Choose a spacelike hypersurface Σ: $x^0 = 0$ and restrict the functions $g^{ab} + \varepsilon \hat{g}^{ab}$ and $(g^{\alpha\beta} + \varepsilon \hat{g}^{\alpha\beta})_{,0}$ to Σ. The resulting functions $g^{ab}(x^\lambda, \varepsilon)$, $\dot{g}^{\alpha\beta}(x^\lambda, \varepsilon)$ can be taken as unconstrained initial data. For sufficiently small (finite) values of ε and in a sufficiently small part of Σ they satisfy the signature conditions and inequalities stated above. Therefore, these data determine uniquely a one-parameter family of solutions of (1)-(3). The solution corresponding to $\varepsilon = 0$ is the original background solution, by construction. The 'tangent' to this family at $\varepsilon = 0$, restricted to Σ, corresponds to the data $(\hat{g}^{ab}, \hat{g}^{\alpha\beta}{}_{,0})_\Sigma$ of the given perturbation, also by construction. Since these initial data determine a solution of the system of linearised equations uniquely—for that system is equivalent to the linear Leray system obtained by linearising the full quasilinear Leray system—that solution is identical with the one which was assumed given at the beginning of this argument. This finishes the proof of linearisation stability of the system (1)-(3).

We remark in passing that the vacuum field equation, $R_{ab} = 0$, is also locally linearisation stable at any of its solutions [9].

We finally note that the solutions of the exact equations for dust as well as those of the corresponding linearised equations respect relativistic causality: in both cases the outer characteristics are the null hypersurfaces of the unperturbed metric.

3. Linear perturbations of dust

The equations for linear perturbations of a 'background' solution (g_{ab}, u^a, ρ) of (1) and (3) are obtained by considering a one-parameter family of such solutions passing, for the parameter value $\varepsilon = 0$, through the given solution, and taking derivatives with respect to ε at $\varepsilon = 0$. We denote these perturbations by $\hat{g}_{ab}$, $\hat{u}_a$, etc. Thus one has, e.g.,

$$\hat{g}^{ab} = -g^{ac}g^{bd}\hat{g}_{cd} \tag{12}$$

$$\hat{u}^a = g^{ab}\hat{u}_b - g^{ab}u^c\hat{g}_{bc} \tag{13}$$

and

$$(\nabla_b u^a)^{\hat{}} = \nabla_b \hat{u}^a + \hat{\Gamma}^a_{bc}u^c \tag{14}$$

where the perturbation of the connection Γ^a_{bc} is a tensor $\hat{\Gamma}^a_{bc}$ given by

$$\hat{\Gamma}^a_{bc} = \tfrac{1}{2}g^{ad}(\nabla_c \hat{g}_{bd} + \nabla_b \hat{g}_{cd} - \nabla_d \hat{g}_{bc}). \tag{15}$$

Therefore, we have

$$\hat{R}_{ab} = \nabla_c \hat{\Gamma}^c_{ab} - \nabla_b \hat{\Gamma}^c_{ac} = \nabla^c_{(a} \hat{g}_{b)c} - \tfrac{1}{2} (\nabla^c_{\ c} \hat{g}_{ab} + \nabla_{ab} \hat{g}_{cd} g^{cd}). \tag{16}$$

The linearised versions of (1) and (3) are

$$\hat{R}_{ab} = \kappa \hat{\rho} (u_a u_b + \tfrac{1}{2} g_{ab}) + \kappa \rho (2 \hat{u}_{(a} u_{b)} + \tfrac{1}{2} \hat{g}_{ab}) \tag{17}$$

and

$$2 u^a \hat{u}_a = \hat{g}_{ab} u^a u^b \tag{18}$$

respectively. These two equations imply

$$\hat{\rho} = (2\kappa^{-1} \hat{R}_{ab} + \rho \hat{g}_{ab}) u^a u^b \tag{19}$$

$$\rho \hat{u}_a = (-\kappa^{-1} h^b_a \hat{R}_{bc} + \tfrac{1}{2} \rho g_{ac}) u^c \tag{20}$$

and, with use of (15) and (16),

$$(h^c_a h^d_b - h_{ab} u^c u^d)(2 \delta^e_{(c} \nabla^f_{.d)} - \delta^e_c \delta^f_d \nabla^g_{.g} - g^{ef} \nabla_{cd}) \hat{g}_{ef} = \kappa \rho (h^c_a h^d_b + h_{ab} u^c u^d) \hat{g}_{cd}. \tag{21}$$

In these equations we have used the tensor

$$h^a_b := \delta^a_b + u_b u^a \tag{22}$$

which projects onto the subspaces of the tangent spaces that are orthogonal to u^a.

Equations (19)–(21) conversely imply (17) and (18). We may therefore consider $\hat{g}_{ab}$ as the basic perturbed field from which, for a given background, $\hat{\rho}$ and $\hat{u}_a$ can be computed via (19) and (20). The equations obtained by perturbing (4) and (5) are implied by (19)–(21). (This follows without calculation from the linearisation stability established in § 1.) Correspondingly we take (21) as the basic equation governing linearised perturbations.

In deriving (21) we have not imposed any gauge conditions on either the background solutions or their perturbations. For the further study of solutions of (21) we restrict the perturbations by requiring

$$\hat{g}_{ab} u^b = 0. \tag{23}$$

This condition slightly simplifies (18)–(21). In particular, (21) reduces to

$$(h^c_a h^d_b - h_{ab} u^c u^d)(2 \delta^e_{(c} \nabla^f_{.d)} - \delta^e_c \delta^f_d \nabla^g_{.g} - g^{ef} \nabla_{cd}) \hat{g}_{ef} = \kappa \rho \hat{g}_{ab}. \tag{24}$$

The operator on the left-hand side of this equation maps the space of symmetric 'spatial' tensor fields into itself.

It is always possible to satisfy the gauge condition (23). To prove this we recall that under the ('infinitesimal') gauge transformation generated by the vector field X^a, $\hat{g}_{ab}$ changes according to [10]

$$\hat{g}_{ab} \to \hat{g}_{ab} - (\mathscr{L}_X g)_{ab} = \hat{g}_{ab} - 2\nabla_{(a} X_{b)}. \tag{25}$$

In order to satisfy (23) one therefore needs a vector field X^a such that

$$u^b \nabla_b X_a + u^b \nabla_a X_b = \hat{g}_{ab} u^b. \tag{26}$$

This linear system of partial differential equations for X_a is not strictly hyperbolic. It can, however, be replaced by

$$u^b \nabla_b X_a - X_b \nabla_a u^b = \hat{g}_{ab} u^b - \nabla_a (u^b X_b) \tag{27}$$

which implies, because of the geodetic nature (4) of the background flow, the relation

$$(u^a X_a)^{\cdot} = \tfrac{1}{2} \hat{g}_{ab} u^a u^b. \tag{28}$$

(Here we have abbreviated the covariant differentiation along the unperturbed flow, $u^a\nabla_a$, by a dot, a notation we shall continue to use whenever it is convenient.) One can now solve (27) as follows. Prescribe arbitrary initial values for X_a on a hypersurface intersecting the unperturbed flow lines. Then find a function ϕ by integrating $\dot{\phi} = \frac{1}{2}\hat{g}_{ab}u^a u^b$ along the flow lines, using the initial values $\phi = u^a X_a$. Afterwards, compute X_a by integrating the system of ordinary differential equations $\dot{X}_a - X_b\nabla_a u^b = \hat{g}_{ab}u^b - \nabla_a\phi$ with the given initial values. Its solution obeys $u^a\dot{X}_a = (u^a X_a)^\cdot = \hat{g}_{ab}u^a u^b - \dot{\phi} = \frac{1}{2}\hat{g}_{ab}u^a u^b$ and consequently $u^a X_a = \phi$ holds everywhere. The solution X_a thus satisfies (27), which finishes the proof.

Condition (23) does not fix the gauge completely; the remaining gauge transformations are generated by vector fields satisfying the homogeneous equation

$$u^b\nabla_{(a}X_{b)} = 0 \tag{29}$$

as follows from (26). Since this equation can be solved just as (26), the restricted gauge freedom is determined by initial values for X_a, i.e. by four functions of three variables. (It is worth noting that the arguments used to establish that the gauge condition (23) can always be satisfied, and that the remaining gauge freedom is given by (29), do not depend on field equations, but only on the geodetic property of u^a. The 'radiation' gauge can therefore also be chosen for vacuum perturbations.)

The basic perturbation equation (24) restricted by (23) is an unconstrained system of six coupled equations of second order for six unknowns. A solution is thus specified by twelve functions of three variables. Since the restricted gauge freedom consists of four functions, the intrinsic freedom of the perturbations amounts to eight functions corresponding to four degrees of freedom.

The characteristics of (24) are found by replacing the operator ∇_a by a covector l_a, the gradient of prospective characteristic hypersurfaces. Using (23) one finds as the characteristic equation

$$(u^a l_a)^8 \cdot (g^{bc}l_b l_c)^2 = 0. \tag{30}$$

(To obtain this result it is convenient to choose coordinates such that, at an event, $g_{ab} = \eta_{ab}$, $u^a = \delta_0^a$, $l_2 = l_3 = 0$ so that $\hat{g}_{0a} = 0$.) The reality of the roots of (30) confirms that the system (24) (subject to (23)) is hyperbolic, its outer characteristics being the null cones; the presence of multiple roots shows that the system is not strictly hyperbolic, in accordance with the results reviewed in § 1.

4. High-frequency perturbations, gravitational waves

We now specialise the metric perturbations $\hat{g}_{ab}$ to locally plane, monochromatic, high-frequency fields, i.e. we make the WKB ansatz

$$\hat{g}_{ab}(x, \varepsilon) = \mathscr{R}\gamma_{ab}(x, \varepsilon)$$

$$\gamma_{ab}(x, \varepsilon) = e^{i/\varepsilon S(x)}f_{ab}(x, \varepsilon) \tag{31}$$

$$f_{ab}(x, \varepsilon) \underset{\varepsilon \to 0}{\sim} \sum_{n=0}^{\infty} (\varepsilon/i)^n \underset{n}{f}_{ab}(x)$$

where $\sim$ indicates that the series is to be considered as an asymptotic expansion. As usual, the 'physical' perturbation is taken to be the real part of the complex γ_{ab} introduced for computational convenience only. The parameter ε serves the usual

purpose to express that, if ε is small, the phase $\varepsilon^{-1}S$ varies rapidly in comparison with the amplitude f_{ab}: the wave covector is $\varepsilon^{-1}l_a$ where

$$l_a := \nabla_a S \tag{32}$$

and the frequency relative to the unperturbed dust flow is $-\varepsilon^{-1}u^a l_a$. (At the end of the calculation one may put $\varepsilon = 1$, absorbing scale factors into S and the f_{ab}.) In accordance with the gauge condition (23) we require

$$f_{ab}u^b = 0. \tag{33}$$

Inserting (31) into (24), requiring the coefficient of the highest-order term (ε^{-2}) of the resulting formal series to vanish, taking into account (33) and setting

$$\omega := -u^a l_a \qquad k_a := h_a^b l_b \tag{34}$$

one gets the equation

$$-2k_{(a}f_{b)c}k^c + l^2 f_{ab} + (k_a k_b + \omega^2 h_{ab})f_{cd}h^{cd} = 0. \tag{35}$$

The condition for this equation to admit a non-trivial solution f_{ab} obeying (33), the

dispersion relation, is of course identical with the characteristic equation (30). There are, consequently, *two modes*, characterised by

$$l^2 := g^{ab}l_a l_b = 0 \qquad \text{(mode I)} \tag{36}$$

and

$$-\omega = u^a l_a = 0 \qquad \text{(mode II)} \tag{37}$$

respectively. Insertion of (32) gives the *eikonal equations* whose solutions S are the phase functions of the modes. Restricting attention to real solutions we shall, without loss of generality, always take $\omega \geq 0$. Then S and the f_{ab} in (31) are uniquely determined by the one-parameter family of real fields $\hat{g}_{ab}(x, \varepsilon)$ which is asymptotically represented by the series in (31).

The *phase hypersurfaces* of the first mode are null hypersurfaces belonging to high-frequency waves propagating without dispersion or damping with the speed of light. Those of the second, zero-frequency mode are arbitrary timelike hypersurfaces generated by flow lines of the background dust; they belong to perturbations 'carried along' by the dust.

The *rays* (bicharacteristics of (24)) along which the amplitudes are transported are the spacetime projections of the solutions of Hamilton's equations

$$\frac{\partial x^a}{\partial v} = \frac{\partial H}{\partial l_a} \qquad \frac{\partial l_a}{\partial v} = -\frac{\partial H}{\partial x^a}. \tag{38}$$

Appropriate Hamiltonians are the left-hand sides of (36) and (37), respectively [11]. The rays are null geodesics for *mode I* and unperturbed dust flow lines for *mode II*.

We proceed to determine the lowest-order *amplitudes* (polarisation states). Inserting $l^2 = 0$ into (35) and choosing an orthonormal frame field (e_a^1, e_a^2) in the plane orthogonal to u^a and k^a, we get for *mode I* the general solution

$$f_{ab} = A_+ e_{ab}^+ + A_\times e_{ab}^\times$$

wherein

$$e^+_{ab} = (e^1_a e^1_b - e^2_a e^2_b) \qquad e^\times_{ab} = (e^1_a e^2_b + e^2_a e^1_b) \tag{39}$$

the standard transverse-traceless expression for a gravitational wave amplitude. For *mode II, $\omega = 0$*, one obtains from (35) the general solution

$$\text{II:} \quad \underset{0}{f}_{ab} = B_{(a}k_{b)} \qquad u^a B_a = 0. \tag{40}$$

This mode corresponds, for $B_a = Bk_a$, to density perturbations and, for B_a orthogonal to both u^a and k^a, to vorticity perturbations. (It is instructive to take Fourier components of the perturbations of an Einstein–de Sitter dust model [12] and recognise their leading high-frequency parts as special cases of the preceding general equations.)

To complete the lowest-order WKB approximation we have to determine how the amplitudes (39) and (40) are transported along the respective rays. For this purpose we substitute in (24) for $\hat{g}_{ab}$ the expression $\gamma_{ab} = e^{i/\varepsilon S}f_{ab}$, where S is some solution of one of the eikonal equations (36) and (37). This gives a partial differential equation for f_{ab} of the form

$$\left(\underset{0}{L} + \frac{\varepsilon}{i}\underset{1}{L} + \left(\frac{\varepsilon}{i}\right)^2 \underset{2}{L} \right)f = 0 \tag{41}$$

where $\underset{n}{L}$ is a linear differential operator of order n which maps spatial symmetric tensor field f_{ab} into fields of the same kind. The $\underset{n}{L}$ are given explicitly by

$$\underset{0}{L}_{ab}{}^{cd} := -[2h^{(c}_{(a}l^{d)}k_{b)} - h^{(c}_{(a}h^{d)}_{b)}l^2 + h^{cd}(\omega^2 h_{ab} - k_a k_b)]$$

$$\underset{1}{L}_{ab}{}^{cd} := i[2h^i_{(a}h^{(c}_{b)}(l^{d)}\nabla_i + \nabla^{d)}l_i) + h^{(c}_{(a}k_{b)}\nabla^{d)} - 2(h^{(c}_{(a}h^{d)}_{b)}l_i\nabla^i) - h^{(c}_{(a}h^{d)}_{b)}(\nabla_i l^i)$$

$$\qquad - g^{cd}\{h^i_{(a}h^j_{b)}\nabla_j l_i + 2h^i_{(a}k_{b)}\nabla_i + 2\omega h_{ab}u^i\nabla_i\} \tag{42}$$

$$\qquad + h_{ab}(g^{cd}u^iu^j\nabla_j l_i - 2\omega\nabla^d u^c)]$$

$$\underset{2}{L}_{ab}{}^{cd} := [h^i_{(a}h^{(c}_{b)}\nabla^{d)}_i - h^{(c}_{(a}h^{d)}_{b)}\nabla^e_e - h^{cd}h^i_{(a}h^j_{b)}\nabla_{ji}$$

$$\qquad - h_{ab}(2u^iu^{(c}\nabla^{d)}_i - u^cu^d\nabla^e_e - h^{cd}u^iu^j\nabla_{ji}) - \rho h^{(c}_{(a}h^{d)}_{b)}].$$

On introducing the series $f \sim \Sigma(\varepsilon/i)^n \underset{n}{f}$ into (41) and requiring the terms of each power of ε to vanish separately one obtains a sequence of equations. The first of these is (35), specialised to the mode to which the chosen phase function S belongs; its solutions are (39) and (40), respectively. The second one is of the form

$$\underset{0}{L}_{ab}{}^{cd}\underset{1}{f}_{cd} + \underset{1}{L}_{ab}{}^{cd}\underset{0}{f}_{cd} = 0. \tag{43}$$

This equation admits a solution $\underset{1}{f}_{cd}$ if and only if $\underset{1}{L}_{ab}{}^{cd}\underset{0}{f}_{cd}$ is annihilated by all left null 'vectors' of $\underset{0}{L}$, i.e. by all spatial symmetric tensors p^{ab} such that $p^{ab}\underset{0}{L}_{ab}{}^{cd} = 0$. The

number of such linearly independent tensors p^{ab} equals the dimension of the corresponding space of polarisation states f_{ab}, given by (39) and (40), respectively. The solvability condition for (43) thus imposes the following condition on the lowest-order amplitudes:

$$p_r^{ab} \underset{1}{L}_{ab}{}^{cd} e_{cd}^s A_s = 0. \tag{44}$$

Here r, s label a basis of left null 'vectors' p_r^{ab} and right null 'vectors' e_{ab}^s, and A_s denotes the sth component of $\underset{0}{f}_{ab}$ in the basis e_{ab}^s. $r, s = 1, \ldots, d$ if d is the dimension of the space of polarisation states; $d = 2$ for mode I, $d = 3$ for mode II.

One of the remarkable facts of the WKB method is that the principal part of the (apparently partial) differential operator acting on the 'vector' A_s in (44) always has the form

$$M_r^s \frac{\partial H}{\partial l_a} \partial_a \tag{45}$$

so that (44) reduces, in fact, to an *ordinary* differential equation for A_s along the rays given by (38). (For a proof of this assertion see the appendix which extends the general formalism developed, e.g., in [11] and [2]. For the connection of this formalism with the modern theory of partial differential equations see, e.g., [13].) If the matrix M is invertible, (44) is a transport equation for $\underset{0}{f}_{ab}$ (respectively A_s); if not, (44) in general imposes further algebraic conditions on the lowest-order amplitudes and gives transport equations for these restricted amplitudes; in exceptional cases (44) may be identically satisfied or admit no solution.

Using the expressions for $\underset{0}{L}$ and $\underset{1}{L}$ from (42) and (43) yields for both modes:

$$\{-2\delta^c{}_{(a}k_{b)}k^d + l^2 \delta_a^c \delta_b^d + (k_a k_b - \omega^2 k_{ab})h^{cd}\}\underset{1}{f}_{cd}$$
$$+ \{2h^c{}_{(a}h_{b)}{}^e(l^d\nabla_e + \nabla^d l_e) + 2h^c{}_{(a}k_{b)}\nabla^d - h^c{}_{(a}h_{b)}{}^d(\nabla_l + \tfrac{1}{2}\theta)$$
$$- h^{cd}(\delta^e{}_{(a}\delta_{b)}{}^f\nabla_e l_f + 2h^e_{(a}k_{b)}\nabla_e + h_{ab}[2\omega\nabla_u - \nabla_u\omega] - 2\omega h_{ab}\nabla^c u^d\}\underset{0}{f}_{cd} = 0 \tag{46}$$

where we have used the abbreviations

$$\nabla_l := l^a \nabla_a \qquad \nabla_u := u^a \nabla_a \qquad \theta := \nabla_a l^a. \tag{47}$$

This equation simplifies very much if specialised to one of the modes, and if suitable tetrads are used.

Let us first consider *mode I*. Then the matrix $\underset{0}{L}_{ab}{}^{cd}$ which multiplies $\underset{1}{f}_{cd}$ in (43) is annihilated if transvected with either one of the polarisation tensors occurring in (39). Transvecting (46) with e_+^{ab}, $e_\times^{ab}$ and inserting (39) leads, after some manipulation, to

$$(\nabla_l + \tfrac{1}{2}\theta)A_+ + \tfrac{1}{2}e_+^{ab}(A_+\nabla_l e_{ab}^+ + A_\times\nabla_l e_{ab}^\times) = 0$$
$$(\nabla_l + \tfrac{1}{2}\theta)A_\times + \tfrac{1}{2}e_\times^{ab}(A_+\nabla_l e_{ab}^+ + A_\times\nabla_l e_{ab}^\times) = 0. \tag{48}$$

This pair of equations shows that the pair $(A_+, A_\times)$ is indeed transported along the rays, the matrix M of (46) being the unit matrix for this mode. By specialising (e_a^1, e_a^2) we can even decouple the equations (45), and simplify each of them, as follows. Along each null geodesic ray $x^a(v)$ with tangent $\dot{x}^a = l^a$, the vectors l^a, u^a span a timelike 2-plane, and e_1^a, e_2^a span its spacelike orthogonal complement. Any vector contained in the (e_1^a, e_2^a) plane at one point of a ray can be transported uniquely along the ray in such a way that it remains in the (e_1^a, e_2^a) planes and its covariant rate of change has no component in those planes. This 'quasiparallel transport' (which generalises Fermi transport) is analytically expressed by

$$\nabla_l X^a = -X^b \nabla_l q^a{}_b \tag{49}$$

if $q^a{}_b$ denotes the tensor which projects orthogonally onto the (l^a, u^a) plane; it preserves inner products and linear relations between vectors [2]. We now *choose* e_a^1, e_a^2 such that they are *propagated quasiparallel* along the rays. Then (48) simplifies further; we obtain

$$(\nabla_l + \tfrac{1}{2}\theta) \begin{bmatrix} A_+ \\ A_\times \end{bmatrix} = 0. \tag{50}$$

This equation means that the change of the complex vector $(A_+, A_\times)$ along a ray consists solely of a rescaling by a positive factor proportional to the square root of the cross sectional area of a small bundle of rays, just as in the case of gravitational or electromagnetic waves *in vacuo*. As in those cases, the transport preserves linear, circular, elliptic polarisation, helicity and ellipticity. (50) also implies the conservation law

$$\nabla_a \hat{T}^{ab} = 0 \tag{51}$$

for the effective energy-momentum tensor

$$\hat{T}^{ab} = \frac{1}{4\kappa} (|A_+|^2 + |A_\times|^2) l^a l^b \tag{52}$$

of the wave, defined as *in vacuo* (Isaacson tensor [4]). (51) and (52) and the geodesic law for l^a (or (50) directly) also give the conservation law

$$\nabla_a N^a = 0 \tag{53}$$

for the effective 'graviton number' current

$$N^a = \frac{1}{4\kappa\hbar} (|A_+|^2 + |A_\times|^2) l^a. \tag{54}$$

Of course, the bending of gravitational rays by the background field is taken into account by the geodesic law.

We now turn to *mode II*. Then $\omega = 0$, $k_a = l_a$, and $\underset{0}{f_{ab}}$ is given by (40). It is easy to verify that the left and right null 'vectors' for the matrix $\underset{0}{L}$ in this case are given by

$$\begin{aligned} &p_r^{ab}: e_1^{(a} e_3^{b)}, \; e_2^{(a} e_3^{b)}, \; h^{ab} - 2e_3^a e_3^b, \\ &e_{ab}^s: e_{(a}^1 e_{b)}^{3)}, \; e_{(a}^2 e_{b)}^{3)}, \; e_a^3 e_b^3. \end{aligned} \tag{55}$$

Using (40) in (46) and then transvecting it with p_r^{ab} we get the three equations ($e_3^a = k^a/k$, $B_a = B_1 e_a^1 + B_2 e_a^2 + B_3 e_a^3$):

$$e_1^c(B_3 e_3^d - 2B^d)(\nabla_c k_d - \nabla_d k_c) = 0$$

$$e_2^c(B_3 e_3^d - 2B^d)(\nabla_c k_d - \nabla_d k_c) = 0 \qquad (56)$$

$$ku^c \nabla_u B_c + B^c e_3^d(\nabla_c k_d - \nabla_d k_c) = 0.$$

However, as $\omega = 0$ implies $k_c = l_c = \nabla_c S$, these three equations reduce to the single equation

$$u^c \nabla_u B_c = 0. \qquad (57)$$

Recalling that $B_c u^c = 0$ and that u^c generates geodesics, equation (57) is trivially satisfied. This result means that for mode II, the matrix M in (45) vanishes and (44) is identically satisfied. Thus in this case there is no first-order transport equation for the leading amplitude, (43) can be solved for f_{ab} irrespective of the way f_{ab} is

transported, and the transport of f_{ab} is governed by higher-order equations. Finally

the matter perturbations $\hat{\rho}$, $\hat{u}_a$ which at this stage can only be determined in leading ε^{-2} order, may easily be verified by means of (16), (19), (20) and (39) to be zero for both modes.

Acknowledgments

Appendix. The transport equation for degenerate modes

The purpose of this appendix is to show in general how transport equations arise from the solvability condition (44) for the next-to-lowest-order WKB equation (43). The argument generalises the 'lemma on bicharacteristic directions' ([11] ch 6, § 3, no 11) to systems of arbitrary order with degenerate modes; see also [2].

We use matrix notation as in (41). Suppose $H(x, l)$ is a factor of the left-hand side of the dispersion relation $\det L = 0$ which characterises a mode via $H = 0$, and assume

that on this dispersion branch $\partial H/\partial l \neq 0$ so that H is a Hamiltonian for the corresponding rays (cf (38)). Let m denote the number of components of the vector f in (41) and let r ($< m$) denote the rank of L for the mode considered. (For the case considered

in this paper, $m = 6$, and $r = 4$ for *mode I*, $r = 3$ for *mode II*.) Then, on the subset $H = 0$ of phase space, L admits $m - r =: p$ linearly independent left null vectors N_j and

equally many such right null vectors R_j. The mode is simple if $p = 1$, degenerate

otherwise. (In the example treated in this paper, both modes are degenerate.) Thus one has, on the branch $H = 0$,

$$N_j \cdot \underset{0}{L} = 0 \qquad \underset{0}{L} \cdot R_k = 0. \tag{A1}$$

Varying the second set of equations (A1) at a fixed point x with respect to the wave covector l gives

$$\underset{0}{L} \cdot \delta R_k + (\partial \underset{0}{L}/\partial l) \cdot R_k \delta l = 0 \tag{A2}$$

and therefore, because of the first set of equations (A1),

$$N_j \cdot \frac{\partial \underset{0}{L}}{\partial l} \cdot R_k \delta l = 0. \tag{A3}$$

Since this holds for all variations δl characterised by $(\partial H/\partial l) \cdot \delta l = 0$ it follows that there exist functions M_{jk} such that

$$N_j \cdot \frac{\partial}{\partial l} \underset{0}{L} \cdot R_k = M_{jk} \frac{\partial H}{\partial l}. \tag{A4}$$

The rest of the argument concerns the relation between $\underset{0}{L}$ and $\underset{1}{L}$. From the way in which (41) is constructed from the original perturbation equation (24) it follows that the principal (first-order) part $\underset{1}{L'}$ of $\underset{1}{L}$ is given by

$$\underset{1}{L'} = \left(\frac{\partial}{\partial l_a} \underset{0}{L} \right) \partial_a. \tag{A5}$$

This relation holds generally, not only for the example considered in this paper [2, 11]. If (A5) is inserted into (43) of the main text and (A4) is used, the expression (44) for the transport operator arises, except for an obvious change of notation.

References

[1] Gayer S and Kennel C F 1970 *Phys. Rev.* D **19** 1979
[2] Breuer R and Ehlers J 1980 *Proc. R. Soc.* A **370** 589; A **374** 65
[3] Breuer R 1982 *Gen. Rel. Grav.* **14** 757
[4] Isaacson R A 1968 *Phys. Rev.* **166** 1263
[5] Wald R M 1984 *General Relativity* (Chicago: Chicago University Press)
[6] Choquet-Bruhat Y 1958 *Bull. Soc. Math. France* **86** 155–75
[7] Choquet-Bruhat Y, DeWitt-Morette C and Dillard-Bleich M 1977 *Analysis, Manifolds and Physics* (Amsterdam: North-Holland)
[8] Lichnerowicz A 1967 *Relativistic Magnetohydrodynamics* (New York: Benjamin)
[9] Brill D R, Reula O and Schmidt B G 1986 MPA 245
[10] Sachs R K 1964 *Relativity, Groups and Topology* ed B De Witt and C De Witt p 523
[11] Courant R and Hilbert D 1962 *Methods of Mathematical Physics* vol 2 (New York: Interscience)
[12] Sachs R K and Wolfe A M 1967 *Astrophys. J.* **47** 74
[13] Taylor M R 1981 *Pseudo differential operators* (Princeton, NJ: Princeton University Press)