Gravitational Radiation in the Limit of High Frequency. I. The Linear Approximation and Geometrical Optics*

RICHARD A. ISAACSON†
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 14 July 1967)

A formalism is developed for obtaining approximate gravitational wave solutions to the vacuum Einstein equations of general relativity in situations where the gravitational fields of interest are quite strong. To accomplish this we assume the wave to be of high frequency and expand the vacuum field equations in powers of the correspondingly small wavelength, getting an approximation scheme valid for all orders of $1/r$, for arbitrary velocities up to that of light, and for all intensities of the gravitational field. To lowest order in the wavelength, we obtain a gauge-invariant linearized equation for gravitational waves which is just a covariant generalization of that for massless spin-2 fields in a flat background space. This wave equation is solved in the WKB approximation to show that gravitational waves travel on null geodesics of the curved background geometry with their amplitude, frequency, and polarization modified by the curvature of space-time in exact analogy to light waves.

1. INTRODUCTION

OBJECTS such as neutron stars, collapsing supernovae, and quasars may endow regions of space with gravatitional fields which seem enormous by terrestrial standards and may provide us with natural sources of intense gravitational radiation. In order to describe mathematically the waves from such objects, we must use the full apparatus of general relativity. This, however, runs us up against the notorious complexity of the nonlinear Einstein field equations. If we do not wish to be limited to unphysically overspecialized models with high symmetry, we must abandon any practical hope of getting exact solutions to these equations, and must content ourselves with just finding good approximations to the true radiative solutions. Here other problems arise, for the conventional technique of obtaining approximate wave solutions via linearization of the field equations[1] has its own drawbacks. First of all, linearization is accomplished by assuming that field strengths are weak and that space-time is essentially flat to lowest order. From the very start, then, this procedure is manifestly inapplicable to the really interesting strong-field problems, where gravitational waves can be expected to impart huge curvature to the fabric of space-time. Secondly, linearization of the field equations inherently destroys the possibility of describing the interaction of the gravitational field with itself. Therefore, within any linear approximation, we cannot explore the resulting secular changes in geometry as energy is transported away from regions in space containing purely gravitational fields. Finally, it is not at all clear that the weak-field approximation procedure can be extended beyond the linear first step to include the higher-order nonlinear terms as corrections.[2–4] We seem to be challenged to find a new method of obtaining approximate solutions representing strong gravitational radiation propagating through a highly curved space, and capable of including at least some nonlinear features of the theory as well.

Multipole-approximation procedures at null infinity overcoming some of these limitations have already been developed.[5–7] These depend on using light cones as coordinates and expanding the metric in powers of the luminosity distance from an isolated source which is embedded in an asymptotically flat space and emitting outgoing radiation. Even with these hypotheses, much algebra must be done before basic results are uncovered.

An entirely different method of approximation is appropriate for gravitational fields of high frequency, for in this interesting limit it is possible to see the local distribution of gravitational energy. This is accomplished by rewriting the equations of general relativity in such a way as to exhibit an effective contribution to the total stress energy which comes from gravity itself. From a mathematical viewpoint, this amounts to the trivial shifting of terms from the left to the right side of the field equations, but from a physical viewpoint, it significantly contributes to our insight into certain classes of geometries. These are geometries which consist of a smoothly changing "background" metric which has been altered by "perturbations" of small amplitude but of high frequency. Wheeler[8] has used this outlook in estimating the possible energy density present in gravitational waves (perturbations) moving through the large scale structure of the universe (background). Energy was regarded as localized in the

* Work supported in part by NASA Grant No. NsG-436; based on the author's doctoral dissertation at the University of Maryland.
† National Academy of Sciences Postdoctoral Resident Research Associate. Present address: Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio 45433.
[1] A. Einstein, Sb. Preuss. Akad. Wiss. **1916**, 688 (1916).
[2] A. Trautman, Lectures on General Relativity, King's College, London, 1958 (unpublished).
[3] A. Trautman, in Preceedings of the International Conference on Relativistic Theories of Gravitation, London, 1965 (unpublished).
[4] V. Fock, Rev. Mod. Phys. **29**, 325 (1957).
[5] H. Bondi, Nature **186**, 535 (1960).
[6] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
[7] R. K. Sachs, Proc. Roy. Soc. (London) **270**, 103 (1962).
[8] J. A. Wheeler, *Geometrodynamics* (Academic Press Inc., New York, 1962), p. 117.

high-frequency waves and, when averaged over many wavelengths, served as a source for the curvature of the cosmos. Brill and Hartle (BH)[9,10] published a study of spherical gravitational geons in which they presented the details of an exciting new approximation method for treating high-frequency gravitational waves in a strongly curved space and emphasized the reality of the effective stress energy carried by such fields. Still another example of the importance of high-frequency fields was given by Misner,[11] who pointed out that the Arnowitt-Deser-Misner[12] canonical decomposition of the metric involves only purely local operations for the case of high-frequency radiation.

In this and an accompanying paper,[13] we will study the powerful BH approximation method for gravitational waves in a self-consistent background field, but with a more detailed discussion of the basis they gave, paying attention to gauge (coordinate) invariance and developing some new applications. The essential assumption we will make is that the gravitational waves are to have a high frequency, and our plan will be to expand the Einstein equations in powers of their wavelength, the small parameter this assumption supplies. We find an approximation scheme correct to all orders of $1/r$, all magnitudes of the field strength, and valid for arbitrary velocities up to that of light. To lowest order we will have a linear wave equation for the high-frequency gravitational field which will tell how the curvature of space interacts with and modifies the wave. Later, we will incorporate higher-order terms in the expansion to see how the wave reacts back on the geometry of space in a nonlinear feedback process. The higher-order nonlinearities will be left as the subjects of a second paper,[13] while the present one will treat the basic expansion procedure, its gauge invariance, and the linear approximation in great detail. We will analyze the linear equations in such a way that the strong analogy between gravitation and electromagnetism in the geometrical-optics limit clearly emerges.

What exactly do we mean when we say the wave is of "high frequency," and under what circumstances can we expect such fields to be important? We will call the frequency "high" whenever the wavelength of the gravitational field is small compared to the radius of curvature of the background geometry. This assumption is already seen to hold in the conventional weak-field linearization, where the background is flat and thus has infinite radius of curvature, and so all weak-field results will just be a special case of our general theory. Besides the binary star systems, fissioning stars, oscillating and rotating spheroids, and other conventional weak-field sources,[14–16] there exist sources of gravitational radiation at optical frequencies which may be among the more important sources which persist for long periods of time. For example, the predominant source of gravitational waves from the sun is in the thermal motion of matter causing gravitational bremsstrahlung.[17–19] Also, gravitational waves of optical frequency should arise as photons are converted into gravitons of the same frequency in the presence of a constant electromagnetic field.[20] Moreover, all gravitational radiation from isolated systems is of high frequency when it gets far enough away from its source, for, assuming that the wavelength λ remains approximately constant, as we increase the distance r from the source of mass m, the ratio of wavelength to radius of curvature of space is of order $(\lambda^2 m/r^3)^{1/2}$, which becomes negligible for large r. In fact, even wavelengths on the scale of galactic diameters or intergalactic distances are seen to be short when compared to the average background cosmological curvature of the universe. These examples should give some idea of the scope to which the high-frequency approximation can be applied.

While the hypothesis of high frequency will hold throughout this paper, we will sometimes combine it with an additional hypothesis. This will be to assume the existence of a simple asymptotic expansion of the exact solution, valid when the frequency becomes very large and wavefront curvatures are negligible, in which the leading term is locally a single plane wave. Asymptotic methods are often applied to physical problems containing a parameter in order to give results where exact solutions are difficult to obtain. Moreover, even if such exact solutions are available, it is invariably simpler to obtain the asymptotic expansion directly than to first find the exact solution and then to extract its asymptotic behavior. We shall call the existence of such an expansion the "WKB assumption," after its best-known application.

2. EXPANSION OF THE RICCI TENSOR

Let us agree at the start to use the name "gravitational wave" to describe both radiation and induction fields, while the term "gravitational radiation" will be reserved for those fields which escape to null infinity. We picture a gravitational wave as a small ripple in the geometry of space-time running through a highly curved, slowly changing background. The frequency of the ripple is high but its amplitude quite small, since

[9] D. Brill and J. B. Hartle, Phys. Rev. **135**, B271 (1964).

[10] D. R. Brill, Nuovo Cimento Suppl. **II**, No. 1 (1964).

[11] C. W. Misner, in *Proceedings on Theory of Gravitation* (Gauthier-Villars, Paris, 1964).

[12] R. Arnowitt, S. Deser, and C. W. Misner, Phys. Rev. **121**, 1556 (1961).

[13] R. A. Isaacson, following paper, Phys. Rev. **166**, 1272 (1968).

[14] C. V. Vishveshwara, University of Maryland Technical Report on Grant NSG-436, 1964 (unpublished).

[15] C. W. Chin, Phys. Rev. **139**, B761 (1965).

[16] W. Y. Chan, Astrophys. J. **147**, 664 (1967).

[17] V. N. Mironovskii, Zh. Eksperim. i Teor. Fiz. **48**, 358 (1965) [English transl.: Soviet Phys.—JETP **21**, 236 (1965)].

[18] S. Weinberg, Phys. Rev. **140**, B516 (1965).

[19] M. Carmeli, Phys. Rev. **158**, 1243 (1967).

[20] M. E. Gertsenshtein, Zh. Eksperim. i Teor. Fiz. **41**, 113 (1961) [English transl.: Soviet Phys.—JETP **14**, 84 (1962)].

we do not want pathological situations where physicists struck by waves change size even faster than Alice in Wonderland. However, this by no means implies that the energy content of the wave is small. Quite the contrary, the energy carried along by the gravitational wave is pictured as a major (if not the only) cause for the background geometry to curve up. We assume that the total metric $g_{\mu\nu}$ takes the form postulated by BH,[9]

$$g_{\mu\nu}=\gamma_{\mu\nu}+\epsilon h_{\mu\nu},$$

where $\gamma_{\mu\nu}$ represents the background metric which is a slowly varying function of space-time, $h_{\mu\nu}$ is the high-frequency ripple which changes singificantly over a much smaller distance, and ϵ is a smallness parameter which insures that laboratory geometry has only microscopic fluctuations.

We introduce estimates of how fast the metric components vary by saying that typically their derivatives are of order

$$\partial\gamma\sim\gamma/L,\quad \partial h\sim h/\lambda,$$

where L and λ are characteristic lengths over which the background and ripple change significantly, and where L is assumed very much larger than λ. The effective energy density contained in the wave is of order $(c^4/G)(\epsilon/\lambda)^2$, while the curvature of the background is of order L^{-2}. The Einstein equations then tell us that the background curvature is equal to G/c^4 times the total energy density curving the background, or $L^{-2}\geq(G/c^4)(c^4/G)(\epsilon/\lambda)^2$, i.e., $\epsilon\leq\lambda/L$. The most interesting case occurs when no other sources of energy besides gravitational waves are present, and the two dimensionless numbers are equal, i.e., $\epsilon=\lambda/L\ll 1$. We will make this assumption in order to simplify matters, and will only need to concern ourselves with the one small parameter ϵ. Once this is done, the total metric remains slowly changing on a macroscopic scale, and the total curvature will be entirely due to the microscopic wave. We are now in a position to formalize our order-of-magnitude arguments and give an axiomatic characterization to the types of metrics of interest. All we need do is regard L as a constant (say, of order unity) and λ as a parameter which is to be replaced by $\epsilon=\lambda/L$, since $O(\lambda)=O(\epsilon)$.[21] We then see that we are studying the one-parameter class of geometries differing infinitesimally by a high-frequency field which serves as a source for the background metric common to all. We will say that a metric contains a high-frequency wave if and only if there exist a family of coordinate systems (called steady coordinates), related by infinitesimal coordinate transformations, in which the total metric takes the form

$$g_{\mu\nu}(x)=\gamma_{\mu\nu}(x)+\epsilon h_{\mu\nu}(x,\epsilon), \tag{2.1}$$

$$\epsilon\ll 1,\quad \gamma_{\mu\nu}=O(1),\quad h_{\mu\nu}=O(1), \tag{2.2}$$

[21] By definition, $f(x)=O(\epsilon^n)$ means that there exists a constant M such that $f(x)<M\epsilon^n$ as ϵ approaches zero.

TABLE I. Magnitude of terms in expansion of Ricci tensor.

Term	Symbolic form	Order of magnitude
$R_{\alpha\beta}^{(0)}$	$\gamma^{-1}\partial^2\gamma$	$L^{-2}=\epsilon^2\lambda^{-2}=O(1)$
$\epsilon R_{\alpha\beta}^{(1)}$	$\gamma^{-1}\partial^2(\epsilon h)$	$\epsilon\lambda^{-2}=O(\epsilon^{-1})$
$\epsilon^2 R_{\alpha\beta}^{(2)}$	$\epsilon h\gamma^{-2}\partial^2(\epsilon h)$	$\epsilon^2\lambda^{-2}=O(1)$
$\epsilon^3 R_{\alpha\beta}^{(3+)}$	$\epsilon^2 h^2\gamma^{-3}\partial^2(\epsilon h)$	$\epsilon^3\lambda^{-2}=O(\epsilon)$

$$\gamma_{\mu\nu,\alpha}=O(1),\quad h_{\mu\nu,\alpha}=O(\epsilon^{-1}), \tag{2.3}$$

$$\gamma_{\mu\nu,\alpha\beta}=O(1),\quad h_{\mu\nu,\alpha\beta}=O(\epsilon^{-2}). \tag{2.4}$$

It should be noted that (2.1)–(2.4) imply a highly curved space, since in steady coordinates the Riemann tensor is $R_{\alpha\beta\gamma\delta}=O(\epsilon^{-1})$. Let $R_{\alpha\beta}(m_{\mu\nu})$ denote the Ricci tensor formed out of the metric $m_{\alpha\beta}$ (for sign conventions, see Appendix A). We may expand the Ricci tensor for the total metric in powers of ϵ to obtain[9]

$$R_{\alpha\beta}(\gamma_{\mu\nu}+\epsilon h_{\mu\nu}) \equiv R_{\alpha\beta}^{(0)}+\epsilon R_{\alpha\beta}^{(1)}+\epsilon^2 R_{\alpha\beta}^{(2)}+\epsilon^3 R_{\alpha\beta}^{(3+)}, \tag{2.5}$$

where

$$R_{\alpha\beta}^{(0)}=R_{\alpha\beta}(\gamma_{\mu\nu}), \tag{2.6}$$

$$R_{\alpha\beta}^{(1)}=\tfrac{1}{2}\gamma^{\rho\tau}(h_{\rho\tau;\alpha\beta}+h_{\alpha\beta;\rho\tau}-h_{\tau\alpha;\beta\rho}-h_{\tau\beta;\alpha\rho}), \tag{2.7}$$

$$\begin{aligned}R_{\alpha\beta}^{(2)}=-\tfrac{1}{2}\big[&\tfrac{1}{2}h^{\rho\tau}{}_{;\beta}h_{\rho\tau;\alpha}\\ &+h^{\rho\tau}(h_{\tau\rho;\alpha\beta}+h_{\alpha\beta;\tau\rho}-h_{\tau\alpha;\beta\rho}-h_{\tau\beta;\alpha\rho})\\ &+h_\beta{}^{\tau;\rho}(h_{\tau\alpha;\rho}-h_{\rho\alpha;\tau})\\ &-(h^{\rho\tau}{}_{;\rho}-\tfrac{1}{2}h^{;\tau})(h_{\tau\alpha;\beta}+h_{\tau\beta;\alpha}-h_{\alpha\beta;\tau})\big].\end{aligned} \tag{2.8}$$

Here the semicolons denote covariant differentiation with respect to the background metric, which is also used to raise or lower all indices. The remainder term $R_{\alpha\beta}^{(3+)}$ is now fully defined by (2.5)–(2.8). Even if we allow for the manifest powers of ϵ in (2.5), the quantities defined by (2.6)–(2.8) are not intrinsically of the same magnitude. Symbolically, the Ricci tensor is

$$R(g)=g^{-1}\partial^2 g,$$

and the metric and its inverse are

$$g=\gamma+\epsilon h,$$
$$g^{-1}=\gamma^{-1}+\epsilon h\gamma^{-2}+\epsilon^2h^2\gamma^{-3}+\cdots.$$

Although γ and h are the same order, by (2.3) and (2.4) their derivatives are very different. Therefore, when second derivatives are applied to γ, the result is very much smaller than when they are applied to h. The results are summarized in Table I. We see that the dominant term is $\epsilon R_{\alpha\beta}^{(1)}=O(\epsilon^{-1})$. Smaller than this by a factor ϵ are both $R_{\alpha\beta}^{(0)}$ and $\epsilon^2R_{\alpha\beta}^{(2)}=O(1)$. Smallest of all is the remainder term $\epsilon^3R_{ab}^{(3+)}=O(\epsilon)$, down by ϵ^2 from the dominant term.

3. EXPANSION OF THE RIEMANN TENSOR

Just as we did for the Ricci tensor, we may expand the Riemann tensor $R_{\alpha\beta\gamma\delta}(g_{\mu\nu})$ in powers of ϵ to obtain

$$R_{\alpha\beta\gamma\delta}(\gamma_{\mu\nu}+\epsilon h_{\mu\nu}) =R_{\alpha\beta\gamma\delta}^{(0)}+\epsilon R_{\alpha\beta\gamma\delta}^{(1)}+\epsilon^2R_{\alpha\beta\gamma\delta}^{(2)}+\epsilon^3R_{\alpha\beta\gamma\delta}^{(3+)}, \tag{3.1}$$

where

$$R_{\alpha\beta\gamma\delta}^{(0)}=R_{\alpha\beta\gamma\delta}(\gamma_{\mu\nu}), \tag{3.2}$$

$$R_{\alpha\beta\gamma\delta}^{(1)}=\tfrac{1}{2}(h_{\alpha\gamma;\beta\delta}+h_{\beta\delta;\alpha\gamma}-h_{\beta\gamma;\alpha\delta}-h_{\alpha\delta;\beta\gamma}+R_{\alpha\sigma\gamma\delta}^{(0)}h^{\sigma}{}_{\beta}-R_{\beta\sigma\gamma\delta}^{(0)}h^{\sigma}{}_{\alpha}). \tag{3.3}$$

As before, $\epsilon R_{\alpha\beta\gamma\delta}^{(1)}=O(\epsilon^{-1})$ is dominant in magnitude, $R_{\alpha\beta\gamma\delta}^{(0)}$ and $\epsilon^2R_{\alpha\beta\gamma\delta}^{(2)}=O(1)$ are smaller by a factor of ϵ, and $\epsilon^3R_{\alpha\beta\gamma\delta}^{(3+)}=O(\epsilon)$ is smaller by ϵ^2.

The expressions defined in (2.7) and (3.3) will be shown to be gauge invariant in the high-frequency limit. Because of this, $R_{\alpha\beta\gamma\delta}^{(1)}$ will play a central role in distinguishing the presence of coordinate waves from true gravitational effects. In the absence of waves (i.e., $h_{\mu\nu}=0$), the Riemann tensor reduces in value to $R_{\alpha\beta\gamma\delta}^{(0)}$. However, if gravitational waves are present, the total curvature of space-time grows enormously in magnitude to the dominant $\epsilon R_{\alpha\beta\gamma\delta}^{(1)}$ term. If space is empty of waves, but a gauge transformation has mixed in coordinate effects, the total curvature is still only of order $R_{\alpha\beta\gamma\delta}^{(0)}$ and hence easy to distinguish.

$R_{\alpha\beta\gamma\delta}^{(1)}$ satisfies the same symmetries as the total Riemann tensor $R_{\alpha\beta\gamma\delta}(g_{\mu\nu})$, but the "Bianchi identities" hold only in the limit of zero wavelength. That is, we find

$$R_{\alpha\beta\gamma\delta}^{(1)}=-R_{\alpha\beta\delta\gamma}^{(1)}, \quad R_{\alpha\beta\gamma\delta}^{(1)}=R_{\gamma\delta\alpha\beta}^{(1)},$$
$$R_{\alpha\beta\gamma\delta}^{(1)}+R_{\alpha\gamma\delta\beta}^{(1)}+R_{\alpha\delta\beta\gamma}^{(1)}=0,$$
$$R_{\mu\nu\alpha\beta;\gamma}^{(1)}+R_{\mu\nu\gamma\alpha;\beta}^{(1)}+R_{\mu\nu\beta\gamma;\alpha}^{(1)}\doteq\epsilon^2.$$

We have introduced a new symbol $\doteq$, which can be read "is down by a factor ϵ^2 (from *a priori* expectations)," whose use and definition can be seen by the following. The terms on the left side of the last equation each involve three derivatives of h. Thus, if we multiply this side by successive powers of ϵ, we would expect that a factor of ϵ^3 is necessary to yield a finite limit for the left side as $\epsilon\rightarrow 0$. In reality, we find that the last expression is of the form $R_{\mu\nu\alpha\beta;\gamma}^{(1)}+\cdots=\sum R^{(0)}\partial h$. The right side is seen to remain finite if we multiply through by a factor of ϵ, two orders less than we would expect. This is the meaning of $\doteq\epsilon^2$. In a general case, $f\doteq\epsilon^p$ means that while $f=O(\epsilon^n)$ is expected by counting the number of derivatives of h in f, actually, because of some internal cancellation, $f=O(\epsilon^{n+p})$. This new notation saves us from having to write confusing equations like $\epsilon^{-n}f=O(\epsilon^p)$.

4. GAUGE TRANSFORMATIONS AND INVARIANCE

We now wish to study infinitesimal coordinate transformations and the gauge transformations they induce, in order to establish the gauge invariance of individual terms in the expansions of the total Riemann and Ricci tensors.

Consider an infinitesimal coordinate transformation from one steady coordinate system to another,

$$x^{\alpha}\rightarrow\bar{x}^{\alpha}=x^{\alpha}+\epsilon\xi^{\alpha}. \tag{4.1}$$

In the new coordinate system, we find,[22] neglecting terms of order ϵ^2,

$$g_{\alpha\beta}=\gamma_{\alpha\beta}+\epsilon(h_{\alpha\beta}-\xi_{\alpha;\beta}-\xi_{\beta;\alpha}). \tag{4.2}$$

Since $\epsilon h_{\alpha\beta}$ is defined as the difference between the background and total perturbed metric, we have

$$\bar{h}_{\alpha\beta}=h_{\alpha\beta}-\xi_{\alpha;\beta}-\xi_{\beta;\alpha} \tag{4.3}$$

for the perturbation in the new coordinate system.

We interpret this to mean that the infinitesimal coordinate change induces a "gauge transformation" similar to those found in flat-space spin-1 and spin-2 fields. Under this change in gauge, quantities dependent upon $h_{\mu\nu}$ also undergo transformations. Either by direct calculation or from the definition of the Lie derivative, we find that under a gauge transformation the dominant parts of the Ricci and Riemann tensors become

$$R_{\alpha\beta}^{(1)}\rightarrow\bar{R}_{\alpha\beta}^{(1)}=R_{\alpha\beta}^{(1)}-\pounds_{\xi}R_{\alpha\beta}^{(0)},$$
$$R_{\alpha\beta\gamma\delta}^{(1)}\rightarrow\bar{R}_{\alpha\beta\gamma\delta}^{(1)}=R_{\alpha\beta\gamma\delta}^{(1)}-\pounds_{\xi}R_{\alpha\beta\gamma\delta}^{(0)}, \tag{4.4}$$

where the Lie derivatives are explicitly given by

$$\pounds_{\xi}R_{\alpha\beta}^{(0)}=R_{\alpha\beta;\sigma}^{(0)}\xi^{\sigma}+R_{\sigma\beta}^{(0)}\xi^{\sigma}{}_{;\alpha}+R_{\alpha\sigma}^{(0)}\xi^{\sigma}{}_{;\beta},$$
$$\pounds_{\xi}R_{\alpha\beta\gamma\delta}^{(0)}=R_{\alpha\beta\gamma\delta;\sigma}^{(0)}\xi^{\sigma}+R_{\sigma\beta\gamma\delta}^{(0)}\xi^{\sigma}{}_{;\alpha}+R_{\alpha\sigma\gamma\delta}^{(0)}\xi^{\sigma}{}_{;\beta}+R_{\alpha\beta\sigma\delta}^{(0)}\xi^{\sigma}{}_{;\gamma}+R_{\alpha\beta\gamma\sigma}^{(0)}\xi^{\sigma}{}_{;\delta}. \tag{4.5}$$

In (4.2), the additional terms $\epsilon\xi_{(\alpha;\beta)}$ can only be regarded as resulting from an infinitesimal coordinate transformation if they still allow us to call $\gamma_{\alpha\beta}$ the background metric unambiguously. This implies that $\epsilon\xi_{(\alpha;\beta)}$ is truly a small quantity compared to $\gamma_{\alpha\beta}$. If ξ^{α} is to be the generator of an infinitesimal coordinate transformation, it may be assumed to satisfy

$$\xi_{\mu}=O(1), \quad \xi_{\mu;\nu}=O(1), \tag{4.6}$$

a form sufficiently general to admit both high- and low-frequency coordinate waves.

If we insert (4.6) and (2.1)–(2.4) into (4.4) and (4.5), we then see that

$$R_{\alpha\beta}^{(1)}-\bar{R}_{\alpha\beta}^{(1)}\doteq\epsilon^2,$$
$$R_{\alpha\beta\gamma\delta}^{(1)}-\bar{R}_{\alpha\beta\gamma\delta}^{(1)}\doteq\epsilon^2. \tag{4.7}$$

In the high-frequency limit $\epsilon\rightarrow 0$, we see that the "perturbations" of the Riemann and Ricci tensors are gauge invariant to an extremely good approximation and therefore meaningful entities, capable of physical measurement. The basic reason behind this is that on a scale of distance of order λ, space appears locally flat, and curvature is locally gauge invariant as in weak-field linear theory. As long as $\lambda\ll L$, perturbations do not have any long-wavelength components, and this local behavior carries over to curved backgrounds to give a global gauge invariance. The fact that our expansion is gauge invariant is extremely important, since any physically observable effects cannot be co-

[22] L. Landau and E. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley Publishing Co., Inc., Reading, Mass., 1962), Sec. 94.

ordinate-dependent. Finally, it should be emphasized that our gauge invariance has resulted only from the assumption of high frequency.

5. LINEAR APPROXIMATION

We have just found a remarkable degree of freedom in the choice of a gauge, and so it is natural to exploit this in order to either simplify the labor involved in future calculation or exhibit interesting results most effectively. To decide on a convenient gauge, let us briefly review the theory of massless spin-2 fields in flat space. It is well known[23] that such a field may be described by a real symmetric tensor field $\psi^{\mu\nu}$ satisfying

$$\psi^{\mu\nu,\beta}{}_{,\beta}=0\,, \tag{5.1a}$$

$$\psi^{\mu\nu}{}_{,\nu}=0\,, \tag{5.1b}$$

$$\psi\equiv\eta_{\mu\nu}\psi^{\mu\nu}=0\,, \tag{5.1c}$$

where $\eta_{\mu\nu}\equiv\mathrm{diag}(1,-1,-1,-1)$. The supplementary conditions (5.1b) and (5.1c) insure that the field energy is positive definite and the field pure spin-2 without spin-0 or spin-1 components. The number of degrees of freedom of $\psi^{\mu\nu}$ is reduced from the five implied by (5.1) to just two because (5.1) is left unaltered by the gauge transformation

$$\psi^{\mu\nu}\rightarrow\bar{\psi}^{\mu\nu}=\psi^{\mu\nu}-\xi^{\mu,\nu}-\xi^{\nu,\mu}\,, \tag{5.2}$$

where ξ^μ is a vector field satisfying

$$\xi^{\mu,\alpha}{}_{,\alpha}=0\,, \tag{5.3a}$$

$$\xi^{\mu}{}_{,\mu}=0\,. \tag{5.3b}$$

The standard weak-field linearization of gravity can be put into the form described in (5.1) by means of infinitesimal coordinate transformations, and it retains this form under the class of coordinate transformations $x^\mu\rightarrow\bar{x}^\mu=x^\mu+\xi^\mu$ satisfying (5.3).[23,24]

We are now in a position to proceed to examine the strong-field case, using the "correspondence principle" that the theory we derive must reduce back to Eqs. (5.1) when field strengths become negligible. Applying the decomposition of the Ricci tensor to the Einstein equations *in vacuo*, we find to lowest order

$$R_{\alpha\beta}{}^{(1)}=0\,, \tag{5.4}$$

and to the next order

$$R_{\alpha\beta}{}^{(0)}=-\epsilon^2R_{\alpha\beta}{}^{(2)}\,. \tag{5.5}$$

The remainder of this paper will be devoted to the analysis of the linear equation (5.4). This was derived by Lanczos[25] and used by Regge and Wheeler[26] to study the stability of Schwarzschild solution. The nonlinear equation (5.5) (which we will treat in the accompanying paper) was first written by BH[9] to show that gravitational waves can be thought to have an effective stress energy which can produce the background curvature.

Let us define

$$\psi_{\mu\nu}\equiv h_{\mu\nu}-\tfrac{1}{2}\gamma_{\mu\nu}h\,,\quad \psi\equiv\gamma^{\alpha\beta}\psi_{\alpha\beta}\,, \tag{5.6}$$

where $h\equiv\gamma^{\alpha\beta}h_{\alpha\beta}$. We regard the $\psi_{\mu\nu}$ as our basic field quantities and rewrite (5.4) as

$$\psi_{\mu\nu}{}^{;\beta}{}_{;\beta}-\tfrac{1}{2}\gamma_{\mu\nu}\psi^{;\beta}{}_{;\beta}-\psi_{\mu\beta}{}^{;\beta}{}_{;\nu}-\psi_{\nu\beta}{}^{;\beta}{}_{;\mu}+2R_{\sigma\nu\mu\beta}{}^{(0)}\psi^{\beta\sigma}+R_{\mu\sigma}{}^{(0)}\psi^{\sigma}{}_{\nu}+R_{\nu\sigma}{}^{(0)}\psi^{\sigma}{}_{\mu}=0\,. \tag{5.7}$$

When we compare this to the flat-space equations (5.1), we have the strong temptation to impose as our choice of gauge that

$$\psi^{\mu\nu}{}_{;\nu}=0\,, \tag{5.8}$$

$$\psi=0\,. \tag{5.9}$$

If we succumb to this temptation, we must, as usual, pay a price. For the case of massless spin-1 or spin-2 fields in a flat-background geometry, the dynamical equations are rigorously gauge invariant and so allow a convenient change of gauge to simplify computations. This, most emphatically, is not the case for the approximate dynamical equations (5.7) which concern us. From Eq. (4.4), we see that our wave equation changes form under an arbitrary gauge transformation (except in the important special case where the Ricci tensor for the background vanishes). This violation of strict gauge invariance should not be a cause for despair, since the magnitude of the terms which destroy rigorous invariance is extremely small. The linearized wave equation (5.7) has terms present of order ϵ^{-2}, ϵ^{-1}, and 1. Consequently, it would be fortuitous if gauge invariance could be extended below the dominant $O(\epsilon^{-2})$ terms. It is therefore quite surprising that we can show [using Eq. (4.4)] that both the $O(\epsilon^{-2})$ and $O(\epsilon^{-1})$ terms in the wave equation are left unaltered under a gauge change, and so we should expect them to be of fundamental physical importance. The remaining $O(1)$ terms are inextricably mixed in with the (coordinate-dependent) Lie derivative terms which create a fog obscuring any intrinsic significance which might be contained in these lowest-order components of the curved-space wave equation.

We now investigate the possibility of choosing a gauge in which the field satisfies the conditions given by Eqs. (5.8) and (5.9). We will always drop the negligible Lie derivative terms which arise from gauge changes, since they only generate higher-order corrections to (5.7). We will, however, retain all terms in this wave equation which have the form $R_{\cdots}{}^{(0)}\psi^{\cdots}$, which are also quite small. This will allow our discussion to apply rigorously to gauge transformations in background geometries which have vanishing Ricci tensor.

Under an infinitesimal coordinate transformation

[23] G. Wentzel, *Quantum Theory of Fields* (Interscience Publishers, Inc., New York, 1949), Sec. 22.

[24] W. Pauli, *Theory of Relativity* (Pergamon Press, Inc., London, 1958), p. 173.

[25] C. Lanczos, Z. Physik **31**, 112 (1925).

[26] J. Regge and J. A. Wheeler, Phys. Rev. **108**, 1063 (1957).

$x^\alpha \to \bar{x}^\alpha = x^\alpha + \epsilon\xi^\alpha$, we find that $\psi^{\mu\nu} \to \bar{\psi}^{\mu\nu}$ and

$$\psi_{\mu\alpha}{}^{;\alpha} \to \bar{\psi}_{\mu\alpha}{}^{;\alpha} = \psi_{\mu\alpha}{}^{;\alpha} - \xi_\mu{}^{;\alpha}{}_{;\alpha} + R_{\mu\alpha}{}^{(0)}\xi^\alpha, \quad (5.10)$$

$$\psi \to \bar{\psi} = \psi + 2\xi^\alpha{}_{;\alpha}. \quad (5.11)$$

If we choose ξ^μ to be the solution to the inhomogeneous system

$$\xi_\mu{}^{;\alpha}{}_{;\alpha} - R_{\mu\alpha}{}^{(0)}\xi^\alpha = \psi_{\mu\alpha}{}^{;\alpha}, \quad \xi^\alpha{}_{;\alpha} = -\tfrac{1}{2}\psi,$$

we find that (5.8) and (5.9) hold in the new coordinate system. (For simplicity, we now drop the bars over all symbols in the new gauge.) We ignore the Lie derivatives introduced by this coordinate change and find that (5.7) becomes

$$-\Delta\psi_{\mu\nu} \equiv \psi_{\mu\nu}{}^{;\beta}{}_{;\beta} + 2R_{\sigma\mu\nu\beta}{}^{(0)}\psi^{\sigma\beta} + R_{\mu\sigma}{}^{(0)}\psi^\sigma{}_\nu + R_{\nu\sigma}{}^{(0)}\psi^\sigma{}_\mu = 0. \quad (5.12)$$

The operator Δ is precisely the generalization of the flat-space d'Alembertian which Lichnerowicz[27] introduced for symmetric tensors following deRham's[28] definition of a similar operator for symmetric ones. We may contract (5.12) to see that $\psi^{;\alpha}{}_{;\alpha} = 0$, so that (5.9) is consistent with (5.12). However, if we differentiate $\Delta\psi_{\mu\nu}$ and use (5.8), we find

$$(-\Delta\psi_{\mu\alpha})^{;\alpha} = \psi_{\mu\alpha}{}^{;\alpha\beta}{}_{;\beta} + R_{\mu\alpha}{}^{(0)}\psi^{\alpha\beta}{}_{;\beta} + \psi^{\alpha\beta}(2R_{\mu\alpha}{}^{(0)}{}_{;\beta} - R_{\alpha\beta;}{}^{(0)}{}_\mu) \doteq \epsilon^3, \quad (5.13)$$

which contradicts (5.12). Thus there is a (small) inconsistency between the wave equation (5.12) and the gauge condition (5.8). Once again, we disregard this as being the inconsequential result of the higher-order corrections to $\psi_{\mu\nu}$ which are of no physical interest. Moreover, there is no inconsistency at all for background metrics of constant curvature such as Schwarzschild's. This is an important advantage derived from using the well-behaved wave equation $\Delta\psi_{\mu\nu} = 0$ instead of some other expression (such as $\psi_{\mu\nu}{}^{;\alpha}{}_{;\alpha} = 0$), which differs from it by terms down by ϵ^2.

In summary, then, high-frequency gravitational waves are approximated to lowest order by the linearized curved-space wave equation (5.12) subject to the gauge conditions (5.8) and (5.9). These are left unchanged by further gauge transformations generated by ξ^μ, satisfying

$$\xi_\mu{}^{;\alpha}{}_{;\alpha} - R_{\mu\alpha}{}^{(0)}\xi^\alpha = 0, \quad \xi^\mu{}_{;\mu} = 0.$$

This final gauge freedom can be pinned down, if desired, by requiring additional conditions (for example, $\psi_{\mu 0} = 0$).

The wave equation (5.7) without gauge specialization is derivable from a variational principle with Lagrangian density;

$$\mathfrak{L} = -(c^4/32\pi G)(\tfrac{1}{2}h^{\alpha\beta;\rho}h_{\alpha\beta;\rho} - \tfrac{1}{2}h^{;\rho}h_{;\rho} + h_{;\alpha}h^{\alpha\beta}{}_{;\beta} - h_{\rho\beta;\alpha}h^{\alpha\beta;\rho})(-\gamma)^{1/2}. \quad (5.14)$$

However, $\mathfrak{L}$ is not gauge invariant and contains explicit x^μ dependence. Therefore the canonical energy-momentum tensor

$$t_\alpha{}^\beta = \mathfrak{L}\delta_\alpha{}^\beta - (\partial\mathfrak{L}/\partial h_{\gamma\delta,\beta})h_{\gamma\delta,\alpha}$$

is neither symmetric, conserved, nor gauge invariant, and $t_0{}^0$ is not necessarily positive definite. A much better tensor than $t_\alpha{}^\beta$ will be shown to describe the energy content of our gravitational wave when we consider the higher-order nonlinearities in the approximation.[13]

6. WKB ANALYSIS OF THE WAVE EQUATION

In the limit of flat space, the wave equation reduces to (5.1a), and it is sufficient to consider only one Fourier component of the solution, which may be written as

$$\psi_{\mu\nu} = A_{\mu\nu}e^{ik_\alpha x^\alpha}, \quad (6.1)$$

where $A_{\mu\nu}$ and k_μ are constants, and the exponential is a rapidly fluctuating function of position. When we deal with a space containing gravitational fields, we know that geometry may be considered locally flat over distances of order L (remember $R_{\alpha\beta\gamma\delta}{}^{(0)} \sim L^{-2}$), as can be seen by introducing normal Riemannian coordinates. On this scale of distance, (6.1) should remain an approximate solution to the wave equation; however, because of the slowly varying geometry, the previously constant $A_{\mu\nu}$ and k_μ can be expected to slowly change in value over a characteristic distance of order L. Thus we may expect to try for a solution to (5.12) of the form

$$\psi_{\mu\nu} = A_{\mu\nu}e^{i\phi} \quad (6.2)$$

[actually only the real part of (6.2) is to be used], where $A_{\mu\nu}$ is a slowly changing real function of position, and ϕ is a real function with a large first derivative but no larger derivatives beyond this to correspond to a slowly changing k_α. Solutions of this form are frequently assumed in mathematical physics for both theoretical insight and computational ease. Perhaps the most familiar example for modern physicists is the WKB approximation for one-dimensional problems in quantum mechanics where a solution in the form (6.2) is sought for an ordinary differential equation. For this reason we will call (6.2) the WKB approximation, although its application to partial differential equations predates quantum mechanics.[29,30] This method was successfully used by Sommerfeld and Runge[31] to establish the transition from Maxwell's equations to classical geometrical optics as the wavelength of light approaches zero. Since that time, the WKB approximation has been used in many fields besides electromagnetism and quantum mechanics, such as acoustics, plasma

[27] A. Lichnerowicz, in *Relativity, Groups and Topology*, edited by C. DeWitt and B. DeWitt (Gordon and Breach Science Publishers, Inc., New York, 1964), p. 827.

[28] G. de Rham, *Variétés Différentiables* (Hermann et Cie., Paris, 1955), p. 131.

[29] J. Liouville, J. Math. (Paris) **2**, 418 (1837).

[30] Lord Rayleigh, Proc. Roy. Soc. (London) **A86**, 207 (1912).

[31] A. Sommerfeld and I. Runge, Ann. Physik **35**, 277 (1911).

physics, elasticity, and hydrodynamics.[32,33] In the varied literature, this approximation is sometimes also known as the eikonal approximation, or the method of stationary phase; however, it is really just the first term in an asymptotic expansion of the exact solution in the limit of vanishing wavelength[34,35]

Returning to (6.2), let us now introduce ray vectors (normals to surfaces of constant phase ϕ) by

$$k_\alpha \equiv \phi_{,\alpha}. \tag{6.3}$$

We may estimate the order of various derivatives needed in calculations as

$$\begin{aligned} R_{\alpha\beta\gamma\delta}^{(0)} &= O(1), \quad A^{\mu\nu}{}_{;\tau} = O(1), \\ k_\mu &= O(\epsilon^{-1}), \quad k_{\mu,\nu} = O(\epsilon^{-1}). \end{aligned} \tag{6.4}$$

Note that (6.2)–(6.4) are compatible with (2.1)–(2.4). Now substitute (6.2) into (5.8)–(5.9) to get

$$\gamma^{\mu\nu} A_{\mu\nu} = 0, \tag{6.5}$$

$$ik_\beta A^{\alpha\beta} + A^{\alpha\beta}{}_{;\beta} = 0. \tag{6.6}$$

In (6.6), the first term is of order ϵ^{-1}, while the second is of order unity and must be neglected for a consistent approximation. This then gives

$$k_\beta A^{\alpha\beta} = 0. \tag{6.7}$$

Similarly substituting (6.2) into (5.12) and grouping terms of the same size, we find

$$\begin{aligned} &[-k_\beta k^\beta A_{\mu\nu}] + i[2k_\beta A_{\mu\nu}{}^{;\beta} + k^\beta{}_{;\beta} A_{\mu\nu}] \\ &\quad + [A_{\mu\nu}{}^{;\beta}{}_{;\beta} + 2R_{\sigma\nu\mu\beta}^{(0)} A^{\beta\sigma} + R_{\mu\sigma}^{(0)} A^\sigma{}_\nu + R_{\nu\sigma}^{(0)} A^\sigma{}_\mu] = 0. \end{aligned} \tag{6.8}$$

The terms in the first, second, and third brackets are of order ϵ^{-2}, ϵ^{-1}, and 1, respectively. To lowest order, (6.8) is

$$k_\beta k^\beta = 0. \tag{6.9}$$

This gives us the important result that gravitational-wave rays are null vectors. Alternatively, we may write (6.9) as an eikonal equation

$$\gamma^{\mu\nu} \phi_{,\mu} \phi_{,\nu} = 0. \tag{6.10}$$

We may introduce a congruence of curves with rays as tangents by

$$dx^\mu/dl = k^\mu. \tag{6.11}$$

The solution curves $x^\mu(l)$ are null geodesics, and l is a preferred affine parameter, since by differentiating (6.9) and remembering that k_α is a gradient, we get

$$0 = k_{\alpha;\beta} k^\alpha = k_{\beta;\alpha} k^\alpha. \tag{6.12}$$

Thus the rays of the gravitational field are parallel-propagated tangentially along null geodesics, just as are the rays for electromagnetic waves.[33,36–38]

Consequently, high-frequency gravitational waves in the WKB approximation are red-shifted and deflected in direction exactly the same way as light when, for example, passing through localized strong gravitational fields or when traveling across the universe.

Now we may proceed to the second-order terms in (6.8) which give us

$$A_{\mu\nu}{}^{;\beta} k_\beta + \tfrac{1}{2} A_{\mu\nu} k^\beta{}_{;\beta} = 0. \tag{6.13}$$

It is convenient to separate the behavior of the amplitude from that of the polarization of the gravitational field. At each point in space where there are waves, we define a polarization tensor field $e_{\mu\nu}$ proportional to $A_{\mu\nu}$, and whose arbitrary magnitude is fixed by normalizing, so that $e_{\mu\nu}e^{\mu\nu}=1$. This is given by $e_{\mu\nu} = (A^{\mu\nu}A_{\mu\nu})^{-1/2} A_{\mu\nu}$. As a shorthand, define the amplitude $\mathcal{A}$ of the wave by $\mathcal{A} = (A_{\mu\nu}A^{\mu\nu})^{1/2}$. This is a real, positive scalar measure of the intensity of the field, and vanishes only when no waves are present. Now substitute $A_{\mu\nu} = \mathcal{A} e_{\mu\nu}$ into (6.13) to obtain

$$(\mathcal{A}_{,\beta} k^\beta + \tfrac{1}{2}\mathcal{A} k^\beta{}_{;\beta}) e_{\mu\nu} + \mathcal{A} e_{\mu\nu;\beta} k^\beta = 0. \tag{6.14}$$

Multiplying by $e_{\mu\nu}$, we have

$$(\ln \mathcal{A})_{,\beta} k^\beta + \tfrac{1}{2} k^\beta{}_{;\beta} = 0. \tag{6.15a}$$

This is just an ordinary differential equation along the null ray, and gives the amplitude once the geometry of the ray congruence is known. This is easily seen if we rewrite (6.15a) using (6.11) to get

$$(d/dl) \ln \mathcal{A} = -\tfrac{1}{2} k^\beta{}_{;\beta}, \tag{6.15b}$$

showing how the field decreases as the rays diverge.

Now, from (6.14) and (6.15a), we see that

$$e_{\mu\nu;\beta} k^\beta = \delta e_{\mu\nu}/\delta l = 0. \tag{6.16}$$

This gives the important physical result that the polarization tensor is parallel-transported along the null geodesic $x^\mu(l)$. At a fixed point along $x^\mu(l)$, we may impose the initial conditions

$$e_{\mu\nu} k^\nu = 0, \tag{6.17a}$$

$$e_{\mu\nu} \gamma^{\mu\nu} = 0. \tag{6.17b}$$

Since both $e_{\mu\nu}$ and k_μ are parallel-transported, these conditions consequently hold everywhere along $x^\mu(l)$, guaranteeing consistency with (6.5) and (6.7).

We see the remarkable similarity between light and gravitation, since the geometrical optics of light tells us that its amplitude satisfies exactly the same transport equation (6.15a), and its polarization is also

[32] See the bibliography in J. B. Keller, R. M. Lewis, and B. D. Seckler, Commun. Pure Appl. Math. 9, 207 (1956).
[33] See also Appendix B.
[34] M. Kline, J. Rat. Mech. Anal. 3, 315 (1954).
[35] R. M. Lewis, J. Math. Mech. 7, 593 (1958).
[36] J. Kristian and R. K. Sachs, Astrophys. J. 143, 379 (1966).
[37] D. Zipoy, Phys. Rev. 142, 825 (1966).
[38] A. Trautman, in *Lectures in General Relativity, Brandeis Summer Institute in Theoretical Physics, 1964*, edited by S. Deser and K. W. Ford (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1965), Vol. I.

parallel-propagated along the null geodesics with ray vectors k_μ as tangents.[33,36,37]

We may rewrite (6.15a) in still another form,

$$(\mathcal{A}^2 k^\beta)_{;\beta}=0\,, \qquad (6.15c)$$

which may be interpreted as a conservation law for the total number of gravitons present in our field, since, if we let

$$N=\int_S k^0 \mathcal{A}^2(-\gamma)^{1/2} d^3x\,,$$

where S is a spacelike hypersurface, x^0 is a timelike coordinate, and (x^1, x^2, x^3) are three spacelike coordinates, then $N_{,0}=0$ for localized wave pulses. Alternately, (6.15c) shows that $\mathcal{A}^{-1}$ is a luminosity distance.

Additional insight into (6.15c) follows from work by Kristian and Sachs[36] intended for light, but which holds equally well for gravitational waves. They consider two observers located at different points along a given ray and moving in such a fashion that the frequency ω which they measure is the same. Each observer measures intensity $I=\mathcal{A}^2\omega^2$ and cross-sectional area dA for the same bundle of rays. Then (6.15c) implies that the measured energy flux through dA is the same for either observer,

$$I_1 dA_1 = I_2 dA_2\,.$$

This shows that energy is transported with the rays serving as guidelines.

We may also find the dominant part of the total Riemann tensor using the WKB approximation. Inserting (6.2) into (3.3), we find

$$R_{\alpha\beta\gamma\delta}{}^{(1)}-2k_{[\alpha}h_{\beta][\gamma}k_{\delta]}\doteq\epsilon\,,$$

where square brackets denote the antisymmetric part. Then $k^\delta R_{\alpha\beta\gamma\delta}(g_{\mu\nu})\doteq\epsilon$. This tells us that the ray vectors serve as principal null vectors of multiplicity four for the Riemann tensor of the total metric (at least to order ϵ), and in the high-frequency limit the total metric is Petrov-type N to lowest order in ϵ. The invariant classification of our metric agrees with the results of Sachs[7] for radiation at large distances from bounded sources, and so this serves as additional confirmation that our decomposition of the metric into a background plus small high-frequency WKB ripple does indeed correspond to the presence of gravitational radiation for spaces which asymptotically become flat. It should be noted that the high-frequency WKB assumption can also be applied to spaces which are not asymptotically flat, and so suggests a limit in which it is possible to extend the notion of radiation to more general space-times.

7. LIMITS OF VALIDITY

While we have seen some of the power of the WKB and high-frequency approaches, they do have their limitations. These are essentially the same limits one finds in the geometrical-optics approach to light.[39,40] Thus we may show that if a beam of gravitational waves is initially convergent, it will collapse, causing the energy density to become infinite at a finite parameter distance along the ray. Before this singularity arises, our WKB approximation must break down, since the radius of curvature of wavefronts no longer is large, and $A_{\mu\nu}$ becomes a rapidly varying function of position. In general, whenever the ray congruence has a caustic (i.e., a manifold of dimension equal to or less than three, on which any neighborhood contains a point with more than one ray of the congruence passing through it), we expect this to happen.

ACKNOWLEDGMENT

It gives me great pleasure to express my deep gratitude to Professor C. W. Misner for suggesting this topic and for his continued patient guidance and helpful criticism.

APPENDIX A: NOTATION AND CONVENTIONS

We assume space-time to be described by a four-dimensional normal-hyperbolic Riemannian manifold with first fundamental form

$$dS^2=g_{\mu\nu}dx^\mu\, dx^\nu \qquad (A1)$$

and signature $(+---)$. We denote the determinant of $g_{\mu\nu}$ by g. Greek indices take on the values $\mu, \nu, \cdots =0, 1, 2, 3$, Latin indices take the values $i, j, \cdots =1, 2, 3$, and summation over repeated indices is implied.

Partial derivatives are indicated by a comma, e.g.,

$$f_{,\nu}\equiv\partial f/\partial x^\nu\,.$$

Covariant derivatives with respect to the total metric $g_{\mu\nu}$ are *never* used; however, covariant derivatives with respect to the background metric $\gamma_{\mu\nu}$ will be indicated by a semicolon (as in $T_{\mu\nu;\alpha}$).

Christoffel symbols for a metric $m_{\mu\nu}$ are defined by

$$\begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}\equiv\tfrac{1}{2}m^{\alpha\tau}(m_{\tau\beta,\gamma}+m_{\tau\gamma,\beta}-m_{\beta\gamma,\tau})\,, \qquad (A2)$$

and the Riemann tensor for this metric is given by

$$R^\alpha{}_{\beta\gamma\delta}(m_{\mu\nu})\equiv\begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}_{,\delta}-\begin{Bmatrix}\alpha\\ \beta\ \delta\end{Bmatrix}_{,\gamma}+\begin{Bmatrix}\alpha\\ \tau\ \delta\end{Bmatrix}\begin{Bmatrix}\tau\\ \gamma\ \beta\end{Bmatrix}-\begin{Bmatrix}\alpha\\ \tau\ \gamma\end{Bmatrix}\begin{Bmatrix}\tau\\ \delta\ \beta\end{Bmatrix}. \qquad (A3)$$

The Ricci tensor is $R_{\mu\nu}(m_{\alpha\beta})=m^{\tau\sigma}R_{\tau\mu\sigma\nu}(m_{\alpha\beta})$.

[39] M. Kline and I. W. Kay, *Electromagnetic Theory and Geometrical Optics* (Interscience Publishers, Inc., New York, 1965), p. 326.

[40] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., New York, 1959), Sec. 3.2.3.

Pairs of indices, respectively, symmetrized or antisymmetrized are denoted as

$$T_{(\mu\nu)}=\tfrac{1}{2}(T_{\mu\nu}+T_{\nu\mu}), \quad T_{[\mu\nu]}=\tfrac{1}{2}(T_{\mu\nu}-T_{\nu\mu}).$$

The alternating symbol is defined by

$$\epsilon_{\mu\nu\sigma\tau}=\epsilon_{[\mu\nu\sigma\tau]}, \quad \epsilon_{0123}=1.$$

Units are chosen so that the speed of light c and the gravitational constant G both equal unity, and the flat-space metric assumes the Lorentzian form

$$\eta_{\mu\nu}=\begin{bmatrix}1 & & & 0\\ & -1 & & \\ & & -1 & \\ 0 & & & -1\end{bmatrix}.$$

APPENDIX B: GEOMETRICAL-OPTICS LIMIT FOR LIGHT

In this Appendix, we present a brief treatment of the geometrical optics of light moving through a curved-background geometry. The analysis is based upon class lectures given by Professor C. W. Misner at the University of Maryland in 1965. It should be read in conjunction with Sec. 6, and so a similar notation has been adapted to emphasize the close relation between light and gravity.

In vacuum, the Maxwell equations have the form

$$F_{\mu\nu}=\psi_{\nu;\mu}-\psi_{\mu;\nu}, \tag{B1a}$$

$$F_{\mu\nu}{}^{;\nu}=0. \tag{B1b}$$

We may combine these and specialize to the Lorentz gauge, where

$$\psi_{\mu}{}^{;\alpha}{}_{;\alpha}+R_{\mu\sigma}{}^{(0)}\psi^{\sigma}=0, \tag{B2a}$$

$$\psi^{\mu}{}_{;\mu}=0 \tag{B2b}$$

[compare (5.8), (5.9), and (5.12)].

In flat space, (B2a) has solution

$$\psi_{\mu}=A_{\mu}e^{ik_{\alpha}x^{\alpha}}, \tag{B3}$$

where A_μ, k_μ are constants; however, in a curved space this is no longer true. If we consider only high-frequency light waves with wavelength λ, and the geometry varies over a characteristic distance $L\gg\lambda$, then locally the wave finds itself moving in an approximately flat domain. The flat-space solution (B3) should then be good if A_μ and k_α are assumed to vary slowly over a distance L. We may therefore assume a trial solution of the WKB form (see Sec. 6)

$$\psi_{\mu}=A_{\mu}e^{i\phi}. \tag{B4}$$

We define

$$\epsilon\equiv\lambda/L, \quad k_{\mu}\equiv\phi_{,\mu},$$

and assume

$$R^{(0)}{}_{\alpha\beta}=O(1), \quad A^{\mu}=O(1), \quad A^{\mu}{}_{;\tau}=O(1), \quad k_{\mu}=O(\epsilon^{-1}), \quad k_{\mu;\tau}=O(\epsilon^{-1}). \tag{B5}$$

When these are substituted into (B2), we find

$$[-k_{\alpha}k^{\alpha}A^{\mu}]+i[2k_{\alpha}A^{\mu;\alpha}+k_{\alpha}{}^{;\alpha}A^{\mu}] +[A^{\mu}{}_{;\alpha}{}^{;\alpha}+R^{\mu}{}_{\alpha}A^{\alpha}]=0, \tag{B6a}$$

$$ik_{\mu}A^{\mu}+A^{\mu}{}_{;\mu}=0. \tag{B6b}$$

In (B6a) the terms in the various brackets are of order ϵ^{-2}, ϵ^{-1}, and 1, respectively. When we set terms of the same size equal to zero, to lowest order we have

$$k_{\beta}k^{\beta}=0, \tag{B7}$$

which implies the eikonal equation

$$\gamma^{\alpha\beta}\phi_{,\alpha}\phi_{,\beta}=0,$$

and that k^α is tangent to null geodesics since $k_{\alpha;\beta}k^{\alpha}=k_{\beta;\alpha}k^{\alpha}=0$.

To lowest order, (B6b) yields

$$k_{\mu}A^{\mu}=0, \tag{B8}$$

or that the vector potential ψ_μ is orthogonal to the direction of propagation of the wave.

The second-order terms in (B6a) imply

$$A_{\mu;\beta}k^{\beta}+\tfrac{1}{2}A_{\mu}k_{\beta}{}^{;\beta}=0. \tag{B9}$$

We may introduce an amplitude $\mathcal{A}$ and polarization vector e_μ by the decomposition

$$A_{\mu}=\mathcal{A}e_{\mu},$$

where $e_{\mu}e^{\mu}=-1$, $\mathcal{A}=(A_{\mu}A^{\mu})^{1/2}$. Equation (B9) may be solved to give

$$e^{\mu}{}_{;\beta}k^{\beta}=0, \tag{B10a}$$

$$(\mathcal{A}^{2}k^{\beta})_{;\beta}=0. \tag{B10b}$$

Since the polarization is parallel-propagated along null geodesics with k^α as tangent, we may impose $k_{\alpha}e^{\alpha}=0$ as an initial condition at one point, and it will be true along the entire curve, guaranteeing consistency with (B8).

Electromagnetic radiation in the geometrical-optics limit travels along null geodesics, with its polarization parallel-propagated and its amplitude satisfying (B10b), just as does gravitational radiation [see (6.9), (6.16), and (6.15c)].