Energy and directional signatures for plane quantized gravity waves

Donald E. Neville*
Department of Physics, Temple University, Philadelphia, Pennsylvania 19122
(Received 2 April 1997; published 19 December 1997)

Solutions are constructed to the quantum constraints for planar gravity (fields dependent on z and t only) in the Ashtekar complex connection formalism. A number of operators are constructed and applied to the solutions. These include the familiar ADM energy and area operators, as well as new operators sensitive to intrinsic spin and directionality ($z+ct$ vs $z-ct$ dependence). The directionality operators are quantum analogs of the classical constraints proposed for unidirectional plane waves by Bondi, Pirani, and Robinson (BPR). It is argued that the quantum BPR constraints will predict unidirectionality reliably only for solutions which are semiclassical in a certain sense. Schwinger has proved that a unidirectional plane electromagnetic wave is stable, even in the presence of the quantum zero point fluctuations of the vacuum. A preliminary calculation (preliminary, because not regulated) indicates that the corresponding gravitational wave may be destabilized by zero point fluctuations. The ADM energy and area operators are likely to have imaginary eigenvalues, unless one either shifts to a real connection, or allows the connection to occur other than in a holonomy. In classical theory, the area can evolve to zero. A quantum mechanical mechanism is proposed which would prevent this collapse. [S0556-2821(98)04102-2]

PACS number(s): 04.60.Ds, 04.30.Nk

I. INTRODUCTION: CLASSICAL RADIATION CRITERIA

The connection-triad variables introduced by Ashtekar [1] have simplified the constraint equations of quantum gravity; further, these variables suggest that in the future we may be able to reformulate gravity in terms of non-local holonomies rather than local field operators [2,3,4]. However, the new variables are unfamiliar, and it is not always clear what they mean physically and geometrically. In particular, it is not clear what operators or structures correspond to gravity waves. Although the quantum constraint equations are much simpler in the new variables, and solutions to these equations have been found [2,5,6], it is not clear whether any of these solutions contain gravitational radiation.

This is the fifth of a series of papers which search for operator signatures for gravitational radiation by applying the Ashtekar formalism to the problem of plane gravitational waves. Paper I in the series [7] constructed classical constants of the motion for the plane wave case, using the more familiar geometrodynamics rather than Ashtekar connection dynamics. Papers II and III switched to connection dynamics and proposed solutions to the quantum constraints [8,9]. The constraints annihilate the solutions of paper II except at boundary points, and annihilate the solutions of paper III everywhere. Paper IV constructs an operator L_Z which measures total intrinsic spin around the z axis [10]. The present paper proposes operator signatures which are sensitive to the directionality of gravitational radiation ($z-ct$ vs $z+ct$ dependence) and applies those operators (as well as the spin, energy, and area operators) to the solutions constructed in papers II and III.

It is not easy to detect the presence of radiation, even when the problem is formulated classically, using the more familiar metric variables. One would like to define gravitational radiation using an energy criterion (such as radiation is a means of transporting energy through empty space, etc.). Gravitational energy is notoriously difficult to define, however, since there is no first-order-in-derivatives-of-the-metric quantity which is a tensor. Accordingly, in the period 1960–1970 several authors developed an algebraic criterion involving transverse components of a second-order quantity, the Weyl tensor [11,12,13,14]. To use the criterion, one needs to know which directions are ‘‘transverse’’; hence the criterion is most useful when the direction of propagation is clear from the symmetry: e.g. radial propagation (for spherical symmetry) or z-axis propagation (for planar symmetry, the case studied in the present paper). The planar metrics considered here [15,16] admit two null vectors k and l which have the right hypersurface orthogonality properties to be the propagation vectors for right-moving (k) and left-moving (l) gravitational waves along the z axis, so that the propagation direction is especially easy to identify. The Weyl criterion is derived and discussed in Appendix D.

A second, more group-theoretical criterion was developed by Bondi, Pirani, and Robinson (BPR) [17]. It is applicable when the plane wave is unidirectional, that is, when the wave is either right moving (depending only on $z-ct$) or left moving (depending only on $z+ct$). The unidirectional case is especially intruguing. It is relatively simple, since no scattering occurs [18,19]. Nevertheless the full complexity of gravity is already present; the unidirectional case is not simply waves propagating in an inert background. In particular, no one has been able to cast the Hamiltonian into a free-field form in terms of variables (π_i, q^j) which commute in the canonical $[\pi_i, q^j] = -i\hbar\delta_i^j$ manner characteristic of non-interacting, non-gravitational theories. Also, the BPR criterion for unidirectional radiation requires *three* amplitudes to vanish, rather than the two one would naively expect from counting the two polarizations associated with unidirectional radiation. I shall argue that the remaining vanishing ampli-

*Electronic address: nev@vm.temple.edu

tude represents a constraint on the background geometry, a constraint which must be satisfied in order for the waves to propagate without backscattering off the background. The BPR group is derived and discussed in Sec. II.

Note that one criterion, that based on the Weyl tensor, is relegated to Appendix D, while the BPR criterion is discussed in the body of the paper. I have done this first because the BPR amplitudes are much simpler than the Weyl amplitudes. Second, the Poisson brackets between the classical Weyl amplitudes and the constraints do not vanish. Since the constraints generate coordinate transformations, this suggests that the amount of Weyl amplitude is coordinate dependent. This plane wave metric does admit coordinates that could be considered as preferred or natural (coordinates which are constant on the hypersurfaces picked out by the null vectors k and l). But even in such a coordinate system, the scalar constraint occurs as part of the Hamiltonian. (This system has a Hamiltonian because the space is non-compact.) The non-zero Poisson bracket would mean that the amount of Weyl amplitude changes with time.

It is not possible to ignore the Weyl amplitudes completely, however; they are central to the literature of the 1960s. Further, expressions which appear complex at one time may appear simple at a later time. At one time the traditional scalar constraint was thought to be too complex because it contained a factor of $1/\sqrt{g}$. Then Thiemann proposed a regularization of this constraint which actually requires that factor [20]. The present ''complexity'' of the Weyl amplitudes may disappear once a quantum regularization is constructed.

Also, the time dependence of these amplitudes may not be a serious drawback. The amount of Weyl amplitude can change with time, presumably because the transverse components of the Weyl tensor can self-interact and evolve into non-transverse components; nevertheless it may be of interest to know that some transverse amplitude was present initially. Therefore I have devoted Appendix D to the Weyl amplitudes.

Now return to the BPR criteria. It is of some interest to reexpresss the classical BPR criteria in the Ashtekar language, even if one does not go on to consider the quantum case. However, one would really like to construct from each classical criterion a corresponding quantum operator. I construct such operators in Sec. III.

The Poisson brackets between the classical BPR operators and the constraints have the form

$$\{\text{classical BPR operator,constraint}\} = \Sigma(\cdots)(\text{classical BPR operators}). \tag{1}$$

Following a terminology used widely in the literature, I will refer to a quantity as physical if its Poisson brackets with the constraints vanish (or equal linear combinations of the constraints). Since the constraints generate transformations of the arbitrary coordinate labels, a physical quantity is coordinate independent. From Eq. (1), the classical BPR amplitudes are not physical. However, they are consistent, meaning that it is consistent with dynamics to demand that BPR amplitudes vanish for all time (to demand that the system is unidirectional for all time): From Eq. (1), if a BPR amplitude is set to zero initially, then its Poisson brackets with the Hamiltonian also vanish, and the amplitude remains zero at later times.

On physical grounds, it is not obvious that the quantum BPR operators should be consistent. The BPR operators enforce unidirectionality, and the quantized system is never purely unidirectional: There are always zero point fluctuations traveling counter to the wave. If the nonlinearities of the system cause the initially unidirectional wave to scatter off these fluctuations, then the BPR amplitudes will evolve in time.

The analogous effect in quantum electrodynamics was investigated by Schwinger [21]. A classical, unidirectional electromagnetic wave does not scatter. In the quantum theory, conceivably the wave might break up into a number of softer photons, because the wave can interact with zero-point photons via exchange of virtual electron-positron pairs. However, Schwinger proved that a unidirectional wave is stable even in the quantum theory.

I have been unable to prove that the corresponding gravitational wave is also stable. If the classical Poisson brackets on the left-hand side of Eq. (1) are replaced by a quantum commutator, then for consistency the commutator should have the form

$$[\text{quantum BPR operator,constraint}] = \Sigma(\cdots)(\text{quantum BPR operators}), \tag{2}$$

where the (quantum BPR operators) must stand to the right, on the right-hand side of Eq. (2). I am assuming that the wave functional stands to the right of the operators, so that the quantum BPR operators must be commuted to the far right, in order to annihilate the wave functional. I have computed the commutators of Eq. (2) in Appendix E, but it is not possible to bring all the quantum BPR operators to the far right, due to factor ordering problems.

Since the commutator calculation in Appendix E involves four degrees of freedom, it is possible on first reading the appendix to miss key features because of the details. Accordingly, in this Introduction I describe an imaginary commutator calculation which involves only two degrees of freedom, yet has all the essential complications of the calculation in Appendix E. Suppose the classical criterion for no left-moving wave has the form $\pi+q,_zQ=0$. Here π and q are a canonical pair ($\pi=-i\hbar\,\delta/\delta q$ after quantization), and Q is the second degree of freedom. All fields depend on two variables (z,t). In a linear theory such as free-field QED, Q is unity, since the classical condition for no left-moving wave is $\pi+q,_z=(\partial_t+\partial_z)q=0$. Now check consistency by commuting this expression with the Hamiltonian (equivalently, with the constraints). Suppose the result is

$$[\pi+q,_zQ,H]=\cdots+(\pi)^2-(q,_zQ)^2, \tag{3}$$

where the ellipsis denotes terms with $\pi+q,_zQ$ to the right. The last term on the right may be rewritten using the identity

$$(\pi)^2-(q,_zQ)^2=(\pi+q,_zQ)(\pi-q,_zQ)/2 + (\pi-q,_zQ)(\pi+q,_zQ)/2. \tag{4}$$

The last term on the right has the correct form. The first term on the right cannot be brought to correct form, because the two factors do not commute:

$$[\pi(z)+q,_z Q(z),\pi(z')-q,_{z'}Q(z')]=-2i\hbar\,\delta(z-z')Q,_z. \tag{5}$$

This is essentially what happens in the actual calculation, Appendix E.

Note that the commutator, Eq. (3), involves products of two operators evaluated at the same point. It is therefore badly defined, and the calculation needs to be redone after a regularization. Discussion of regularization is beyond the scope of this paper, and the conclusion drawn here is provisional: The stability of unidirectional plane gravitational waves has not been established. A historical note on the importance of regularization to calculations of this type: Schwinger's paper continues to be quotable today, in part for its result on plane waves, but primarily because it introduces the elegant method of proper-time renormalization.

The proof of consistency requires a computation of commutators between BPR operators and constraints. Originally I considered doing the calculation in the Ashtekar formalism. However, the Ashtekar calculation is unconvincing, I believe, for both quantitative and qualitative reasons. Quantitatively, the algebra is difficult to follow or reproduce, because the algebra is lengthy and complex. There is also the qualitative objection that (as pointed out above) factor ordering issues are important. Quantum commutator and classical Poisson brackets will differ only if factor ordering issues are signifigant. But it is not clear how to factor order the Ashtekar scalar constraint. There is the usual ambiguity about functional derivatives: Do they go to the right or to the left? In addition, it is not even clear that a single factor ordering suffices: If one insists that the scalar constraint should be self-adjoint, then it should be a sum of two terms having two different factor orderings. Therefore in Appendix E, I compute the commutators using a geometrodynamic formalism. In this formalism, the algebra is signifigantly easier to follow. In addition, there is no ambiguity about how to factor order the scalar constraint, because, in the plane wave case, every term in the scalar constraint is a product of commuting factors.

So far I have assumed that a classical constraint of the form (classical BPR amplitude)$=0$ translates into a quantum constraint (quantized BPR amplitude)$\psi=0$. In Appendix E I require, not only that the commutator [BPR amplitude,constraint] be proportional to a BPR amplitude, but also that the BPR amplitude can be commuted to the far right, where it annihilates the wave functional, assumed to be standing to the right. Actually, the classical theory is recovered if the quantum theory obeys the weaker condition $\langle$quantized BPR amplitude$\rangle=0$, where the $\langle\ \rangle$ denotes an average over an appropriate semiclassical state. In Sec. III, I argue that the weaker condition is the correct one to use with the BPR amplitudes.

Suppose for the moment that this is true; does switching to the weaker condition affect the argument that unidirectional plane waves are unstable? The calculation of Appendix E remains relevant, even if one uses the weaker condition. Just as for the stronger condition, one must continue to check that the weaker condition is preserved in time, by computing commutators with the constraints. However, if the weaker condition is correct, then the commutator, Eq. (3), should be interpreted in a different manner. There is no need to commute the BPR amplitudes to the right. Instead, one should average both sides of Eq. (3) over an appropriate semiclassical state (sandwich both sides between the brackets $\langle\ \rangle$) and look for a zero on the right. Since the $(\pi)^2$ and $(q,_z Q)^2$ terms on the right in Eq. (3) resemble kinetic and potential energies of an oscillator, it would not be surprising if the two energies averaged to the same value. The average of the right-hand side would then vanish, implying consistency on average, $\langle\partial_t(\pi+q,_z Q)\rangle=0$.

But this does not settle the question whether left-moving waves are being produced from the vacuum. Suppose the quantity $\pi+q,_z Q$ is initially zero. Even if this amplitude becomes non-zero because of scattering off the vacuum, the *average* of $\pi+q,_z Q$ probably continues to be zero, because of the randomness inherent in the zero-point fluctuations; similarly for the time derivative of the amplitude. If the weaker condition is the correct one to use, then the only safe way to see growth in left-moving waves is to look at an amplitude which cannot fluctuate in sign, for example $(\pi+q,_z Q)^2$. In free-field QED, where $Q=1$, this is the (divergent) energy of the left-moving zero-point oscillations. Forget the divergence for the moment and proceed: In QED the time derivative, $[H,(\pi+q,_z)^2]$ is proportional to $\partial_z(\pi+q,_z)^2$. One expects the zero-point fluctuations of the vacuum to be z independent by translation invariance, so that the QED commutator vanishes. I have computed the corresponding commutator for the gravitational case in Appendix E. In addition to z-derivative terms, analogous to those encountered in QED, there are derivative-free terms proportional to $(\pi+q,_z Q)^2$ itself; these terms should not vanish when both sides are averaged over a semiclassical state. Again, the calculation is heuristic; regularization is needed.

Even though the unidirectional amplitudes may evolve away from zero, they remain of interest. It is of interest to know that a wave is unidirectional, at least initially, even if it later departs from that state. However, if one must use the weaker condition (amplitudes shoud be averaged over a semiclassical state), rather than the strong condition (amplitudes should annihilate the state), the weaker condition is more difficult to use. How does one recognize when a state is ''semiclassical?'' Also, the weaker criterion involves an average and hence seems to require knowledge of the measure in Hilbert space. The situation is difficult, but may not be impossible; these issues are discussed in Sec. III. Section III in fact discusses several issues which arise on replacing a classical criterion with a quantum one, but the issues having to do with the weaker criterion are the most important. Section III also reports progress made toward the goal of constructing a measure.

In Sec. IV, I construct additional solutions to the quantum constraints. In Sec. V, I apply the Arnowitt-Deser-Misner (ADM) energy operator, the area operator, and the operator L_Z for total intrinsic spin to the wave functional solutions obtained in papers II–III, as well as the new solutions constructed in Sec. IV. Also in Sec. V, I apply the BPR quantum operators to these solutions; the results are informative even when the wave functionals do not correspond to states which are semiclassical.

There are five appendixes. Two of the appendixes (A and D) cover calculational details and the details of the Weyl criterion. Appendix B elucidates the L_Z intrinsic spin content of the BPR amplitudes. Appendix C considers the ADM energy. There is a modest surprise here: Normally the ADM energy is considered to be given by the surface term in the Hamiltonian, but in the quantum case it is possible for the volume term to contribute also. Appendix E studies the consistency of the BPR constraints.

My notation is typical of papers based upon the Hamiltonian approach with concomitant $3+1$ splitup. Uppercase indices $A,B,\ldots,I,J,K,\ldots$ denote local Lorentz indices [''internal'' SU(2) indices] ranging over X,Y,Z only. Lowercase indices $a,b,\ldots,i,j,\ldots$ are also three dimensional and denote global coordinates on the three-manifold. Occasionally the formula will contain a field with a superscript (4), in which case the local Lorentz indices range over X,Y,Z,T and the global indices are similarly four dimensional, or a superscript (2), in which case the local indices range over X,Y (and global indices over x,y) only. The (2) and (4) are also used in conjunction with determinants; e.g., g is the usual 3×3 spatial determinant, while ${}^{(2)}e$ denotes the determinant of the 2×2 X,Y subblock of the triad matrix e_a^A. I use Levi-Cività symbols of various dimensions: $\epsilon_{TXYZ}=\epsilon_{XYZ}=\epsilon_{XY}=+1$. The basic variables of the Ashtekar approach are an inverse densitized triad $\tilde{\mathrm{E}}_A^a$ and a complex SU(2) connection A_a^A.

$$\tilde{\mathrm{E}}_A^a = e e_A^a, \tag{6}$$

$$[\tilde{\mathrm{E}}_A^a, \mathrm{A}_b^B] = \hbar\, \delta(x-x')\, \delta_A^B \delta_b^a. \tag{7}$$

The planar symmetry (two spacelike commuting Killing vectors ∂_x and ∂_y in appropriate coordinates) allows Husain and Smolin [22] to solve and eliminate four constraints (the x and y vector constraint and the X and Y Gauss constraint) and correspondingly eliminate four pairs of $(\tilde{\mathrm{E}}_A^a, \mathrm{A}_a^A)$ components. The 3×3 $\tilde{\mathrm{E}}_A^a$ matrix then assumes a block diagonal form, with one 1×1 subblock occupied by $\tilde{\mathrm{E}}_Z^z$ plus one 2×2 subblock which contains all the ''transverse'' $\tilde{\mathrm{E}}_A^a$, that is, those with $a=x,y$ and $A=X,Y$. The 3×3 matrix of connections A_a^A assumes a similar block diagonal form. None of the surviving fields depends on x or y.

The local Lorentz indices are vector rather than spinor; strictly speaking the internal symmetry is O(3) rather than SU(2), gauge fixed to O(2) rather than U(1). Often it is convenient to shift to transverse fields which are eigenstates of the surviving gauge invariance O(2):

$$\tilde{\mathrm{E}}_\pm^a = [\tilde{\mathrm{E}}_X^a \pm i \tilde{\mathrm{E}}_Y^a]/\sqrt{2}, \tag{8}$$

where $a=x,y$, and similarly for $\mathrm{A}_a^\pm$.

In papers I–III, I use the letter H to denote a constraint (scalar, vector, or Gauss). In the present paper I adopt what is becoming a more common convention in the literature and use the letter C to denote a constraint, while reserving the letter H for the Hamiltonian. The quantity denoted C_S in the present paper is identical to the constraint denoted H_S in papers II and III. This convention underscores the fact that every gravitational theory has constraints, but not every gravitational theory has a Hamiltonian.

In three spatial dimensions it is usual to place the boundary surface at spatial infinity. Bringing the surface at infinity in to finite points is a major change, because at infinity the metric goes over to flat space, and flat space is a considerable simplification. In the present case (effectively one dimensional because of the planar symmetry) the space does *not* become flat as z goes to infinity, and nothing is lost by considering an arbitrary location for the boundary surface. The ''surface'' in one dimension is of course just two points (the two end points of a segment of the z axis). The notation z_b denotes either the left or right boundary point z_l or z_r, $z_l \leqslant z \leqslant z_r$. The result that the space does not become flat as z goes to infinity was established in paper II. Note that this result agrees with one's intuition from Newtonian gravity, where the potential in one spatial dimension due to a bounded source does not fall off, but grows as z at large z.

If a certain solution does not satisfy the Gauss constraint (or other constraint) at the boundary, this does not mean that necessarily there is something wrong with the solution. In classical theory the solutions satisfy the constraints everywhere. In quantum theory, however, when the constraints are imposed after quantization, in the Dirac manner, it is only necessary that the *smeared* constraint annihilate the solution. The Hamiltonian contains a sum of constraints of the form $\int N(z)C(z)dz$, where N is a smearing function or Lagrange multiplier. Since N has no physical significance, one must require that arbitrary small changes in N, δN, annihilate the wave functional:

$$\int dz\, \delta N(z) C(z) \psi = 0. \tag{9}$$

N and δN are not totally arbitrary, however. If N obeys a boundary condition of the form $N \to$constant at boundaries z_b, then Eq. (9) must respect this boundary condition, which means

$$\delta N(z_b) = 0. \tag{10}$$

Equation (10) implies that $C(z_b)\psi$ does not have to vanish. A statement that ''this solution does not obey the constraint at the boundaries'' does not mean necessarily that the solution is flawed.

II. BONDI-PIRANI-ROBINSON SYMMETRY

Bondi, Pirani, and Robinson argue that the metric of a unidirectional plane gravitational wave should be invariant under a five-parameter group of symmetries. Their argument proceeds essentially as follows. First they point out that a plane *electromagnetic* wave moving in the $+z$ direction is invariant under a five parameter group. [Besides the obvious ∂_x, ∂_y, and ∂_v symmetries, there are two ''null rotations'' which rotate the $v=(t+z)/\sqrt{2}$ direction into the x or y direction.] Then for gravitational plane waves they construct five Killing vectors which have the same Lie algebra as the corresponding Killing vectors for the electromagnetic case.

(More precisely they construct ten Killing vectors, one set of five for $ct+z$ waves and a similar set of five for $ct-z$ waves.)

This section constructs the five $ct-z$ vectors and then imposes the usual symmetry requirement that the Lie derivative of the basic Ashtekar fields must vanish in the direction of the Killing vectors. In this way one finds that the fields must obey certain constraints; the section closes with a discussion of the physical meaning of these constraints in the classical theory.

It is convenient to do the proofs in a gauge which has been simplified as much as possible using the ∂_v symmetry and then afterwards transform the results to a general gauge. The plane wave metrics we consider here possess two hypersurface orthogonal null vectors; if the two hypersurfaces are labeled $u=$const and $v=$const [u and $v=(ct\mp z)/\sqrt{2}$], then one can always transform the metric to a conformally flat form in the (z,t) sector by using u and v as coordinates [16]:

$$ds^2=-2dudvf(u,v)+\sum^{(2)} g_{ab}dx^a dx^b. \tag{11}$$

The sums over $a,b,c,\ldots$ extend over x,y only. If one now invokes the symmetry under ∂_v, then $f(u,v)$ depends on u only, and one can remove the function f by transforming to a new u coordinate. In this gauge, the Rosen gauge [23], the metric is (not just conformally flat, but) flat in the (z,t) sector and non-trivial only in the (x,y) sector:

$$ds^2=-2dudv+\sum^{(2)} g_{ab}dx^a dx^b. \tag{12}$$

In the Rosen gauge, the five BPR Killing vectors are $\partial_x,\partial_y,\partial_v$, and

$$\xi^{(c)\lambda}=x^c\delta^\lambda_v+\int^u g^{cd}(u')du'\,\delta^\lambda_d, \tag{13}$$

where $c=x$ or y. These two Killing vectors are the gravitational analogues of the electromagnetic ''null rotations'' which mix the x or y direction with the v direction.

The constraints imposed by the first three ∂_x, ∂_y, and ∂_v Killing vectors are satisfied already, because of the choice of gauge. I now work out the constraints which the last two Killing vectors, Eq. (13), impose on the Ashtekar variables (in the Rosen gauge first, than in a general gauge). I summarize the highlights of the calculation in this section, and move the algebraic details to Appendix A.

It is necessary to calculate the symmetry constraints on the four-dimensional tetrads and Ashtekar connections first, since the Killing vectors are intrinsically four dimensional, and then carry out a $3+1$ decomposition to obtain the constraints on the usual three dimensional densitized triad and connection. At the four-dimensional level, the three local Lorentz boosts have been gauge fixed by demanding that three of the tetrads vanish:

$$e^t_M=0, M=\text{space}. \tag{14}$$

The gauge condition of Eq. (14) is the standard choice, used with all metrics [24]. In addition, for the special case of the plane wave metric, the gauge fixing of the XY Gauss constraint and xy spatial diffeomorphism constraints imply that four more tetrads vanish [22]:

$$e^z_X=e^z_Y=e^x_Z=e^y_Z=0. \tag{15}$$

At the four-dimensional level, the requirement of a vanishing Lie derivative in the direction of the Killing vector gives

$$0=\xi^\lambda\partial_\lambda e^\alpha_I-\partial_\beta\xi^\alpha e^\beta_I-L^{I'}_{.I}e^\alpha_{I'}, \tag{16}$$

$$0=\xi^\lambda\partial^{(4)}_\lambda A^{IJ}_\alpha+\partial_\alpha\xi^{\lambda(4)}A^{IJ}_\lambda+\mathcal{L}^I_{.I'}{}^{(4)}A^{I'J}_\alpha + \mathcal{L}^J_{.J'}{}^{(4)}A^{IJ'}_\alpha-\partial_\alpha\mathcal{L}^{IJ}. \tag{17}$$

These equations are not quite the usual Lie derivatives because of the L and $\mathcal{L}$ terms. L and $\mathcal{L}$ are local Lorentz transformations. If $\xi=\partial_x$, ∂_y, or ∂_v, no L or $\mathcal{L}$ is required. If $\xi=\xi^{(c)}$, one of the two Killing vectors defined at Eq. (13), then a Lorentz transformation L is required in Eq. (16) for the tetrads; otherwise the symmetry destroys the gauge conditions, Eqs. (14) and (15). If the tetrads are Lorentz transformed, then for consistency, ${}^{(4)}A$ in Eq. (17) must undergo the same Lorentz transformation. Since ${}^{(4)}A$ is self-dual, the Lorentz transformation $\mathcal{L}$ in Eq. (17) must be the self-dual version of the Lorentz transformation L:

$$2\mathcal{L}^{IJ}=L^{IJ}+i\,\delta(\epsilon^{IJ}_{..MN}/2\epsilon_{TXYZ})L^{MN}. \tag{18}$$

The phase $\delta/\epsilon_{TXYZ}=\pm1$ is the duality eigenvalue which determines whether the theory is self-dual or anti-self-dual. Because I include the extra factor of ϵ_{TXYZ}, Eq. (18) contains two factors of ϵ and so is independent of one's choice of phase for this quantity. After the four-dimensional theory is rewritten in $3+1$ form, all results will depend only on δ. In the body of the paper I choose $\delta=+1$, but Appendix A indicates what happens for the opposite choice $\delta=-1$.

It is a straightforward matter to determine the Lorentz transformation L which will preserve the gauge conditions of Eqs. (14) and (15), then solve Eqs. (16) and (17). This is done in Appendix A. From Eq. (A23) in that appendix, Eqs. (16) and (17) imply the following constraints on the connection A:

$$0=A^-_a,$$

$$0=-A^+_a+2\text{``Re''}A^+_a \quad \text{(right-moving)}, \tag{19}$$

where $a=x,y$ only. The connection A is now the usual $3+1$ connection, not the four-dimensional connection ${}^{(4)}A$. Re A^X_a without the quotes is the usual real part, containing no factors of i, while

$$\text{``Re''}A^+_a\equiv(\text{Re }A^X_a+i\ \text{Re }A^Y_a)/\sqrt{2}. \tag{20}$$

''Re''A contains a factor of i, because of the i in the definition of the $(X\pm iY)$ O(2) eigenstates, and is no longer real. If one writes out the "Re" and "Im" parts of $A^\pm_a$, it is easy to see that the two constraints of Eq. (19) are just the complex conjugates of each other. To obtain the constraints for left-moving waves, interchange + and − in Eq. (19).

Equation (19) can be interpreted physically by using the classical equations of motion to prove theorems about the spin behavior of the BPR fields. Again, the required calculations are done in an appendix (Appendix B), and this section summarizes the main conclusions.

To interpret the spin content of the four amplitudes which vanish, one should express the total spin angular momentum L_Z of the gravitational wave in terms of the BPR amplitudes [Eq. (19) for right-moving waves, plus two more amplitudes with $+\leftrightarrow-$ for left-moving waves]. Note L_Z cannot be simply an integral over these amplitudes: They are not weight one, and therefore integrating them over dz will not produce a diffeomorphism scalar. In fact the integrand of L_Z is a weighted average over the four weight one combinations $\tilde{\mathrm{E}}^a_A\mathrm{A}^{\mp}_a$ and $\tilde{\mathrm{E}}^a_B[-\mathrm{A}^{\pm}_a+2\text{``Re''}\mathrm{A}^{\pm}_a]$. One can always recover the original four BPR amplitudes from these four, because it is always possible to invert the 2×2 matrix formed from the transverse components of the densitized triad, $\tilde{\mathrm{E}}^B_b$ with $B=X,Y$ and $b=x,y$. Equation (B4) of Appendix B expresses L_Z in terms of these weight one combinations:

$$L_Z=i\int dz\{e^y_+e_{+x}\tilde{\mathrm{E}}^a_-[\mathrm{A}^-_a+(\mathrm{A}^-_a-2\text{``Re''}\mathrm{A}^-_a)]$$
$$+e^y_-e_{-x}\tilde{\mathrm{E}}^a_+[\mathrm{A}^+_a+(\mathrm{A}^+_a-2\text{``Re''}\mathrm{A}^+_a)]\}-(x\leftrightarrow y), \tag{21}$$

where e_{Aa} and e^a_A are triad and inverse triad fields, respectively. Out of the four possible combinations $\tilde{\mathrm{E}}^a_A\mathrm{A}^-_a$ and $\tilde{\mathrm{E}}^a_B[-\mathrm{A}^+_a+2\text{``Re''}\mathrm{A}^+_a]$, only two combinations $\tilde{\mathrm{E}}^a_-\mathrm{A}^-_a$ and $\tilde{\mathrm{E}}^a_+[-\mathrm{A}^+_a+2\text{``Re''}\mathrm{A}^+_a]$ contribute to the spin angular momentum. Since these are the two amplitudes with O(2) helicity ±2 in the local Lorentz frame, it is natural to interpret $\tilde{\mathrm{E}}^a_\pm\mathrm{A}^\pm_a$ as an amplitude for a wave having helicity ±2. Both helicity ±2 combinations must vanish in order to eliminate the two polarizations moving in the $ct+z$ direction.

Two more combinations, $\tilde{\mathrm{E}}^a_+\mathrm{A}^-_a$ and $\tilde{\mathrm{E}}^a_-[-\mathrm{A}^+_a+2\text{``Re''}\mathrm{A}^+_a]$, are helicity zero and also contain the fields of Eq. (19). How does one interpret these helicity zero amplitudes? Using the Gauss constraint plus the classical equations of motion in a conformally flat gauge, one can prove that these two constraints collapse to a single constraint, Eq. (B10) of Appendix B:

$$0=(\partial_t+\partial_z)\tilde{\mathrm{E}}^z_Z. \tag{22}$$

$\tilde{\mathrm{E}}^z_Z=e^z_Ze={}^{(2)}e$, where ${}^{(2)}e$ is the determinant of the 2×2 transverse sector of the triad matrix, a scalar function of (z,t). This function characterizes the background geometry, rather than the wave. In general relativity, however, ''background'' and wave are inseparable, in the sense that the ''background'' is not inert. The wave will scatter off the background, in general, unless it obeys the constraint given by Eq. (22).

This completes the survey of the constraints predicted by BPR symmetry and their physical interpretation in the classical context. In the next section these classical expressions are promoted to quantum operators.

III. TRANSITION FROM THE CLASSICAL TO QUANTUM CRITERION

This section lists three issues which arise when converting a classical expression into a quantum criterion. I summarize and discuss each issue and then show the application to the BPR criteria.

A. Factor ordering

I choose a factor ordering which is natural and simple within the complex connection formalism. The ordering is ''functional derivatives to the right'' [25]. That is, I quantize in a standard manner, by replacing one-half of the fields by functional derivatives,

$$\tilde{\mathrm{E}}^z_Z\rightarrow-\hbar\,\delta/\delta\mathrm{A}^Z_z,$$
$$\mathrm{A}^A_a\rightarrow+\hbar\,\delta/\delta\tilde{\mathrm{E}}^a_A\quad(\text{for }a=x,y\text{ and }A=X,Y), \tag{23}$$

and then order the functional derivatives to the right in every operator or constraint. This approach has the virtue of consistency, since I have used it in two previous papers on quantization of plane waves.

B. Semiclassicality

This section discusses the point raised in the Introduction: The classical constraint (classical BPR amplitude) $=0$ probably does not imply the strong constraint (quantized BPR amplitude)$\psi=0$, but rather the weaker constraint $\langle$quantized BPR amplitude$\rangle=0$, where the average is taken over a semiclassical state.

An example from QED will be helpful in explaining why the constraint must be weaker. The BPR criteria are essentially field strengths for waves moving in a given direction with a given polarization. An analogous quantity from flat space QED is

$$\mathcal{F}\equiv\mathcal{F}^{\mu\nu}m_\mu k_\nu. \tag{24}$$

Here $\mathcal{F}^{\mu\nu}$ is the self-dual QED field strength and $(k,l,m,\overline{m})$ is the usual flat space null tetrad: k and l are null vectors with space components along $\pm z$; m and $\overline{m}$ are transverse polarization vectors along $(\hat{x}\pm i\hat{y})/\sqrt{2}$. Classically, the criterion for absence of radiation along k with polarization $\overline{m}$ is $\mathcal{F}=0$. The corresponding quantum criterion for absence of radiation, obtained after replacing classical A fields by quantum operators $\hat{A}$, is *not*

$$\hat{\mathcal{F}}\psi=0,\ \ (\text{wrong}), \tag{25}$$

but rather

$$\langle\text{semicl}|\hat{\mathcal{F}}|\text{semicl}\rangle=0. \tag{26}$$

Since the QED field strength $\hat{\mathcal{F}}$ contains creation as well as annihilation operators, it cannot annihilate any state, and Eq. (25) is too strong. The classical statement $\mathcal{F}=0$ merely implies the existence of a corresponding semiclassical state such that Eq. (26) holds. I have deliberately used the term ''semiclassical'' rather than ''coherent'' to describe the state in Eq. (26), because the latter term conventionally denotes a state which is an eigenfunction of the annihilation operator, and annihilation operators usually are not available in quantum gravity. While annihilation operators may not exist, certainly semiclassical states will, because of the correspondence principle.

Clearly the criterion, Eq. (26), is more difficult to apply than Eq. (25), but let us survey the damage; the situation may not be hopeless. To define ''semiclassical'' without invoking coherent states or annihilation operators, one can study

$$\hat{F}|\text{semicl}\rangle = f(z)|\text{semicl}\rangle + |\text{remainder}\rangle, \tag{27}$$

where $\hat{F}$ stands for a typical BPR field. $|\text{semicl}\rangle$ is normed to unity, although $|\text{remainder}\rangle$ need not be. I define semiclassical, not by requiring $f(z)$ to be large, but rather by requiring the $|\text{remainder}\rangle$ to be small. If I require $f(z)$ to be large (perhaps reasoning that ''classical'' means large quantum numbers), then I exclude the vanishing amplitude case, $f(z)=0$, where $|\text{semicl}\rangle$ is the vacuum with respect to a given radiation mode. The vacuum is a well-defined state, classically, and one expects it to have a quantum analogue. For $|\text{semicl}\rangle$ to approximate a classical state, therefore, it is not necessary that $f(z)$ be large, only that the fluctuations away from this state be small. In a theory with three spatial dimensions, these fluctuations are measured by

$$\langle \hat{F}\dagger \hat{F}\rangle - \langle \hat{F}\dagger\rangle\langle \hat{F}\rangle = \langle \text{remainder}|\text{remainder}\rangle$$
$$\leqslant (l_p/l)^2/l^2 \quad (3\text{D}). \tag{28}$$

Here the fluctuations are assumed to be small compared to the size of typical matrix elements one gets when $|i\rangle$ and $\langle f|$ are few-graviton states and $\hat{F}$ is a canonical degree of freedom in the linearized theory:

$$\langle f|\hat{F}|i\rangle \approx l_p/l^2 \quad (3\text{D}). \tag{29}$$

l_p is the Planck length, and l is a typical length or wavelength. The right hand side of Eq. (28) is the square of the right hand side of Eq. (29). In the planar case (one rather than three spatial dimensions) remove one power of length l from the denominator of Eq. (29), and two powers of l from the denominator of Eq. (28).

I interrupt the flow of the discussion to present the dimensional analysis needed to establish Eq. (29). [This paragraph and the next could perhaps be skipped on a first reading.] Start by estimating the order of magnitude of the matrix element in Eq. (29) when $\hat{F}$ is replaced by a massless scalar field ϕ, and the dynamics is small perturbation around free field theory. The usual dimensional analysis applied to the quadratic terms in the Lagrangian gives

$$\phi \sim 1/l \quad (3\text{D}) \tag{30}$$

Since ϕ is a massless field, it contains no built-in length scale; l will come from the length scales associated with the initial and final states. If the initial state is the vacuum and the final state is a one-particle wavepacket,

$$\langle f|\phi|i\rangle = \int d^3k f(k)\langle k|\phi|0\rangle$$
$$\sim 1/l \quad (3\text{D}). \tag{31}$$

There is no change from Eq. (30) because the packet is normed to unity; therefore $\int d^3k f(k)\langle k|$ is dimensionless. Now switch from ϕ to the gravitational field h, h a small fluctuation of the tetrad away from background.

$$e_i^I \sim \delta_i^I + l_p h_i^I. \tag{32}$$

The explicit factor of Planck length ensures that the overall factor of $1/G \sim 1/l_p^2$ in the Lagrangian cancels out of the quadratic terms. Therefore the dimensional analysis for h is identical to that for ϕ: $h \sim 1/l$.

Now make the transition from h to the BPR amplitudes $\hat{F}$. Those amplitudes are linear in GA, A the Ashtekar connection and G the Newtonian constant (usually set to unity in this paper). GA is linear in the Lorentz connection ω_a^{IJ}, so that

$$\hat{F} \sim GA \sim \omega \sim \partial e$$
$$\sim l_p/l^2 \quad (3\text{D}), \tag{33}$$

which is Eq. (29). (Again, taking matrix elements does not change dimension.) The discussion for the one dimensional planar case is identical to the discussion just given for the three dimensional case, except for the very first step: ϕ has dimension 1 rather than $1/l$. Therefore the final answer, Eq. (29) or Eq. (33), should be l_p/l rather than l_p/l^2.

In a classical theory, or in a theory quantized on a flat background, l is a typical length associated with the initial and final wavepackets. One might ask what is meant by a typical length l, in a quantized and diffeomorphism invariant theory where no background metric is available. In such a theory, even in the absence of a background metric, length, area, and volume operators can be defined [26,27,20], and the eigenvalues of these geometric operators are dimensionless functions of spins j_i, times factors of l_p to give the correct dimension. The spins label the irreducible representations of SU(2) associated with each holonomy (if the wave functional ψ is in connection representation, so that ψ is a product of holonomies) or associated with each edge of a spin network (if the wavefunctional is a spin network state). Thus one expects $l = l_0(j_i)l_p$, l_0 dimensionless and $\gg 1$. [In the planar case, SU(2) is gauge fixed to O(2) and presumably the SU(2) eigenvalues j will be replaced by the O(2) eigenvalues m = spin angular momentum component along z.] Evidently, then, to check that Eq. (28) is satisfied, one should apply the geometric operators to the state first, in order to estimate l. One also needs the measure in Hilbert space, but in favorable situations one might be able to tell that $|\text{remainder}\rangle$ is small simply by inspection of Eq. (27).

Since the criterion of Eq. (28) is an inequality, it cannot be used to draw sharp distinctions between states. For example, in the linearized limit, if $|N\rangle$ denotes an eigenstate of the number operator having N quanta of a given polarization and direction, then $f(z)$ will be zero, while the norm equation (28) will be order $N(l_p/l)^2/l^2$. There will be uncertainties in estimating l, so that the criterion cannot distinguish sharply between the vacuum state and a number eigenstate having small occupation number N.

The semiclassical criterion works best if one has a measure. Here I report some initial steps taken toward construction of a measure. The discussion will be in the nature of a progress report (results only, no proofs), because more needs to be done. However, it is surprisingly easy to construct a measure which preserves the reality conditions which must be obeyed by the Ashtekar connection.

The wavefunctionals constructed in papers II and III and Sec. IV depend on a complete set of commuting observables consisting of the four $\widetilde{\mathrm{E}}$ in the 2×2 X,Y sector, plus the complex connection A_z^Z. This suggests that one should take the dot product to have the form

$$\langle \phi|\psi\rangle = \int \phi^* \psi \mu d^4\widetilde{\mathrm{E}} d^2A, \tag{34}$$

where $d^2A \equiv d\ \mathrm{ReA}_z^Z d\ \mathrm{ImA}_z^Z$. The measure μ must satisfy several requirements. (i) It must guarantee the quantum form of the reality constraints on the connection:

$$\langle \phi|A\psi\rangle + \langle A\phi|\psi\rangle = 2\langle \phi|\mathrm{ReA}\psi\rangle. \tag{35}$$

(ii) It must guarantee the invariance of $\mu d^4\widetilde{\mathrm{E}} d^2A$ under transformations generated by the scalar, vector, and Gauss constraints. (iii) It must contain enough gauge-fixing delta functions to remove the usual unbounded integrations over infinite numbers of gauge copies. Note that (ii) requires only invariance under the constraints, not invariance under four-dimensional diffeomorphisms. In a $3+1$ formalism, one does not have the proper set of fields to implement the latter invariance, essentially because all fields are evaluated on a constant time hyperslice, whereas four-dimensional diffeomorphisms move fields off the hyperslice [28].

It is possible to construct a μ which guarantees the reality constraints [requirement (i)]. Set

$$\mu = \delta[\mathrm{A}_z^Z + \mathrm{A}_z^{Z}{}^* + 2\omega_z^{XY}]. \tag{36}$$

ω_z^{XY} is the Lorentz connection, (-1) times the real part of A_z^Z, so that this delta function enforces the A_z^Z reality constraint. The surprising fact is that it also enforces the reality constraints on the remaining, transverse A_a^A as well, those with $A=X,Y$ and $a=x,y$. I sketch the proof. From the quantization rule, Eq. (23), the first term in Eq. (35) contains a functional derivative $\delta/\delta\widetilde{\mathrm{E}}_A^a$ acting on the ket wavefunctional. When this is functionally integrated by parts, one gets $-\delta/\delta\widetilde{\mathrm{E}}_A^a$ acting on the bra wave functional [second term in Eq. (35)], plus a term which can be rewritten as

$$[-\delta\mu/\delta\mathrm{A}_z^Z]\delta[2\omega_z^{XY}]/\delta\widetilde{\mathrm{E}}_A^a. \tag{37}$$

The $\delta/\delta\mathrm{A}_z^Z$ can be integrated by parts onto the ket. (The bra depends only on A_z^{Z*}.) It is then (lengthy but) straightforward to show that

$$[\delta[2\omega_z^{XY}]/\delta\widetilde{\mathrm{E}}_A^a]\delta/\delta\mathrm{A}_z^Z = 2\ \mathrm{ReA}_a^A. \tag{38}$$

The $\delta/\delta\mathrm{A}_z^Z$ in the quantum expression corresponds to a factor of $\widetilde{\mathrm{E}}_Z^z$ in the classical expression for ReA_z^Z. Further progress will involve choosing a specific gauge and verifying that requirements (ii) and (iii) above have been met.

The following three paragraphs contain material which at first glance may seem to be of only historical interest, but will be needed later in Sec. V. I interpret the BPR criteria in a semiclassical sense, as $\langle \hat{F}\rangle = 0$ rather than $\hat{F}\psi = 0$. (From now on $\langle\ \rangle$ is understood to indicate an average over a semiclassical state, unless explicitly indicated otherwise.) Yet the Hamiltonian constraints are always imposed strongly, as $\hat{C}_i\psi = 0$, even though (in the linearized limit, at least) the $\hat{C}_i$ are sums of creation and destruction operators, like the BPR field strengths. Why this difference in treatment? This same question was posed and answered in a different context, Lorentz gauge QED, many years ago, and it is worthwhile to take a moment here to review that discussion [29]. In Lorentz gauge QED, the analogue of the C_i is (the usual Gauss constraint, plus) the four-divergence $\partial A = 0$. The analogue of the strong requirement $\hat{C}_i\psi = 0$ would be $\partial\hat{A}\psi = 0$, and the analogue of the semiclassical requirement $\langle \hat{C}_i\rangle = 0$ would be $\partial\hat{A}^+\psi = 0$, where the superscript $+$ denotes positive frequency components. (Since a splitup into positive and negative frequencies is available in QED, there is no need to introduce a semiclassical average.)

Both constraints, the strong and the positive frequency–semiclassical, are used in the Lorentz gauge literature. Authors who employ the positive frequency constraint tend to treat the ‘‘unphysical’’ part of the Hilbert space with more respect. (Remember that the Lorentz gauge condition is designed to eliminate the effects of the unphysical, longitudinal and timelike components.) Heitler [30] is a typical proponent of this approach: He gives a very careful treatment of the unphysical sector, including a full discussion of the Gupta-Bleuler formalism. The payoff is that dot products over the full Hilbert space are well defined, including dot products of longitudinal and timelike photons. Authors who employ the stronger constraint [31] pay a price: It is possible to find states which are annihilated by both the annihilation *and* the creation parts of $\partial\hat{A}$, but these states are not normalizable in the unphysical sector [29]. This result is not particularly surprising: Since the creation operators in $\partial\hat{A}$ create a state with one more timelike or longitudinal photon, ψ must be a sum over an unbounded, infinite number of longitudinal and timelike occupation numbers. This infinte sum leads to the divergence in the norm. The authors who use the strong constraint are well aware of this difficulty, and they circumvent it by requiring that the dot product in Hilbert space be taken over physical excitations only.

Returning to the gravitational case, one can now see why the strong criterion will work for the C_i, but not for the BPR operators. For the moment, imagine the gravitational theory to be linearized, so that the analogy to QED is strongest. The creation operators for the BPR operators create physical quanta, not unphysical. If I impose the strong criterion, I get states which have an infinite norm in the physical sector. There is no way of avoiding this by restricting the measure at a later step. If I now pass from the linearized to the full theory, there is no reason why the strong criterion should suddenly become applicable. I must use the semiclassical criterion, which is justified using the correspondence principle.

C. Non-polynomiality

The BPR operators, Eq. (19), occur in complex conjugate pairs, and one member of the BPR pair involves ReA_a^A, which is a known, but non-polynomial function of $\tilde{\mathrm{E}}_A^a$. In particular, ReA_a^A contains factors of $1/{}^{(2)}\tilde{\mathrm{E}}$, where ${}^{(2)}\tilde{\mathrm{E}}$ is the 2×2 determinant formed from the $\tilde{\mathrm{E}}_A^a$ with internal indices $A=X,Y$ and global indices $a=x,y$. I have dealt with a similar operator, $1/\tilde{\mathrm{E}}_Z^z$, in a previous paper [8], but would just as soon not do so here.

One can use the fact that the BPR constraints come in complex conjugate pairs, plus semiclassicality, to prove the following theorem (and then one uses the theorem to avoid dealing with the non-polynomiality). Theorem:

$$\langle -\mathrm{A}_a^+ + 2\,\mathrm{ReA}_a^+\rangle = \langle \mathrm{A}_a^-\rangle *, \tag{39}$$

and similarly for the other BPR pair. This result is just what one would expect from the corresponding result for expectation values of complex operators in ordinary quantum mechanics (for instance $\langle p+iq\rangle * = \langle p-iq\rangle$) except that here the basic operators are not Hermitean, so that the proof is slightly longer. Proof: Expand out the $\mathrm{A}_a^\pm$ operators using

$$\begin{aligned}\mathrm{A}_a^\pm &= (\mathrm{A}_a^X \pm i\mathrm{A}_a^Y)/\sqrt{2}\\ &= \hbar(\delta/\delta\tilde{\mathrm{E}}_X^a \pm i\,\delta/\delta\tilde{\mathrm{E}}_Y^a)/\sqrt{2}.\end{aligned} \tag{40}$$

Integrate by parts the functional derivative on the left side of Eq. (39), using

$$\begin{aligned}\int \mu\psi*\hbar\,\delta\psi/\delta\tilde{\mathrm{E}}_A^a &= \int \mu[(-\hbar)\,\delta/\delta\tilde{\mathrm{E}}_A^a\psi*]\psi\\ &\quad + \int \mu\psi*2\,\mathrm{ReA}_a^A\psi.\end{aligned} \tag{41}$$

Here ψ is the semiclassical state and $\int\mu$ is the measure, a path integral over the fields in ψ. μ need not be known in detail, except that it enforces the reality condition in Eq. (41) [via $(-\hbar)\,\delta\mu/\delta\tilde{\mathrm{E}}_A^a = 2\,\mathrm{ReA}_a^A\mu$]. Also, μ must be real ($\mu* = \mu$) in order for norms to be real. If one inserts Eq. (41) into the left-hand side of Eq. (39) and carries out the complex conjugation (using $\mu*=\mu$), the result is the right-hand side of Eq. (39).□

IV. ADDITIONAL SOLUTIONS

In the previous section I derived quantum BPR operators. In the present section I construct new solutions to the constraints. In the next section I apply the BPR, ADM energy, and L_Z operators to the solutions constructed in papers II and III and this section.

I start from the solutions considered in paper III. These are strings of transverse $\tilde{\mathrm{E}}_A^a$ operators, ordered along the z axis, and separated by holonomies:

$$\begin{aligned}\psi = \Bigg[\prod_{i=1}^{n}\int_{z0}^{z_{n+1}} dz_i\,\Theta(z_{i+1}-z_i)M(z_{i+1},z_i)\\ \times\tilde{\mathrm{E}}_{A_i}^{a_i}(z_i)S_{Ai}\Theta(z_1-z_0)\Bigg]M(z_0,z_{n+1}).\end{aligned} \tag{42}$$

The M are holonomies along z,

$$M(z_{i+1},z_i) = \exp\left[i\int_{z_i}^{z_{n+1}} \mathrm{A}_z^Z(z')S_Z dz'\right], \tag{43}$$

and the S_M are the usual Hermitian SU(2) generators. These can be $2j+1$ dimensional; they need not be Pauli matrices. The Θ functions in Eq. (42) are Heaviside step functions which path-order the integrations, $z_0 \leq z_1 \cdots \leq z_{n+1}$. For this section only, the boundary points z_l and z_r are relabeled z_0 and z_{n+1}. Although the metric is not flat at the boundaries, it can be taken as conformally flat at the boundaries, with any radiation present confined to a wavepacket near the origin [8].

Since the full SU(2) invariance has been gauge fixed to O(2), it is convenient to use basis fields introduced in Eq. (8), fields which are irreducible representations of O(2). These are one dimensional, labeled by the eigenvalue of S_Z, e.g.,

$$\tilde{\mathrm{E}}_\pm^a = (\tilde{\mathrm{E}}_X^a \pm i\tilde{\mathrm{E}}_Y^a)/\sqrt{2},$$

$$\tilde{\mathrm{E}}_A^a S_A = \tilde{\mathrm{E}}_+^a S_- \text{ or } \tilde{\mathrm{E}}_-^a S_+. \tag{44}$$

Because the irreducible representations are one-dimensional, there is no need to sum over both values of $A_i = \pm$ in Eq. (42), in order to obtain a Gauss-invariant expression; nor is it necessary to take the trace in that equation. However, one must be sure to have an equal number of S_+ and S_- matrices in the chain, in order to form a closed loop of flux with no open ends violating Gauss invariance. That is, if one visualizes each holonomy $M(z_{i+1},z_i)$ as a flux line along z from z_i to z_{i+1}, then the factor in the square brackets, Eq. (42), may be visualized as a flux line from z_0 to z_{n+1}. The line varies in thickness (varies in S_Z eigenvalue) because of the $S_\pm$ operators encountered along the way, but the final S_Z value at z_{n+1} must equal the initial S_Z value at z_0 (there must be an equal number of S_+ and S_- matrices in the chain). Then the final holonomy in Eq. (42), $M(z_0,z_{n+1})$, can join the two ends at z_0 and z_{n+1} and turn the open flux line into a closed flux loop. As shown in paper III, the wave functional of Eq. (42) can be made to satisfy all the constraints, by suitable choice of the a_i and A_i.

A set of wave functionals is said to form part of a kinematical basis if the wavefunctionals are annihilated by the Gauss and spatial diffeomorphism constraints; the wave functionals are physical if they are (kinematical and) annihilated by the scalar constraint as well. In order to obtain a larger, kinematical space, as well as more physical solutions, one can simplify Eq. (42) by dropping all Θ functions. This removes the path ordering, or (in visual terms) this allows flux lines to double back on themselves.

$$\psi_{\rm kin}=\prod_{i=1}^{n}\int_{z0}^{z_{n+1}}dz_i M(z_{i+1},z_i)\tilde{\rm E}^{a_i}_{A_i}(z_i)S_{Ai}M(z_0,z_{n+1}). \tag{45}$$

Again, the expression is Gauss invariant even if $A_i=\pm$ is not summed over, and no trace is needed.

To check the spatial diffeomorphism and scalar constraints, one must first obtain these constraints from the Hamiltonian, written out in an O(2) eigenbasis:

$$H_T={\rm N}'\left[i^{(2)}\tilde{\rm E}(\tilde{\rm E}^z_Z)^{-1}\epsilon_{ab}{\rm A}^+_a{\rm A}^-_b+\sum_\pm(\pm i)\tilde{\rm E}^b_\pm{\rm F}^\mp_{zb}\right]+iN^z\sum_\pm \tilde{\rm E}^b_\pm{\rm F}^\mp_{zb}-i{\rm N_G}\left[\partial_z\tilde{\rm E}^z_Z-\sum_\pm(\pm i)\tilde{\rm E}^a_\pm{\rm A}^\mp_a\right]+ST$$
$$\equiv{\rm N}'C_S+N^zC_z+{\rm N_G}C_G+ST, \tag{46}$$

where

$${\rm F}^\mp_{zb}=[\partial_z\mp i{\rm A}^Z_z]{\rm A}^\mp_b. \tag{47}$$

${}^{(2)}\tilde{\rm E}$ is the determinant of the 2×2 transverse subblock of the matrix $\tilde{\rm E}^a_A$. ST denotes surface terms (terms evaluated at the two end points on the z axis, z_0 and z_{n+1}). The detailed form of these terms is worked out in paper II but will not be needed here. The primed lapse N' equals the usual lapse N multiplied by a factor of $\tilde{\rm E}^z_Z$, and correspondingly the scalar constraint C_S is the usual constraint divided by $\tilde{\rm E}^z_Z$. As shown in paper II, this renormalization leads to a much simpler constraint algebra, but again the details of this will not be relevant here. The system is quantized by replacing transverse ${\rm A}^A_a$ and $\tilde{\rm E}^z_Z$ by functional derivatives, these being the fields conjugate to the fields in ψ:

$${\rm A}^\pm_a\rightarrow\hbar\,\delta/\delta\tilde{\rm E}^a_\mp,$$
$$\tilde{\rm E}^z_Z\rightarrow-\hbar\,\delta/\delta{\rm A}^Z_z. \tag{48}$$

The operator ordering [already adopted in Eq. (42)] is functional derivatives to the right. The first term in Eq. (46) contains an inverse operator $(\tilde{\rm E}^z_Z)^{-1}$; this is well defined provided $\tilde{\rm E}^z_Z M$ never vanishes, that is, provided the S_Z in Eq. (43) never has eigenvalue zero.

As discussed at Eq. (9), the physics must be invariant under small *changes* δN in the lapse and shift, so that the constraint should be written as

$$0=\int dz[\delta{\rm N}'C_S+\delta N^zC_z]\psi_{\rm kin}, \tag{49}$$

$$0=\delta{\rm N}'(z_b)=\delta N^z(z_b), \tag{50}$$

where z_b is either boundary point, z_0 or z_{n+1}. Equation (50) guarantees that the boundary conditions at z_b are left unchanged by the transformation of Eq. (49).

Both the C_S and C_z constraints in Eq. (49) contain terms proportional to ${\rm F}^\mp_{zb}$, the field strength defined in Eq. (47). When a typical term of this type acts on ψ, the result is (up to constants)

$$\begin{aligned}\int dz\,\delta N(z)\tilde{\rm E}^a_+{\rm F}^-_{zb}[\psi_{\rm kin}]&=\int dz\,\delta N(z)\tilde{\rm E}^a_+{\rm F}^-_{zb}\left[\cdots\int dz_iM\tilde{\rm E}^{a_i}_+S_-M\cdots\right]\\&=\int dz\,\delta N\sum_i\tilde{\rm E}^{a_i}_+\left[\cdots\int dz_iM(\partial_z-i{\rm A}^Z_z)\delta(z-z_i)S_-M\cdots\right]\\&=\cdots\left[\cdots\int dz_i\delta(z-z_i)(\partial_{zi}-i{\rm A}^Z_z)MS_-M\cdots\right]\\&=\cdots\left[\cdots\int dz_i\delta(z-z_i)(-iM[S_Z,S_-]M{\rm A}^Z_z-i{\rm A}^Z_zMS_-M)\cdots\right]\\&=0.\end{aligned} \tag{51}$$

On the third line I have changed the ∂_z to ∂_{zi} and integrated by parts with respect to z_i. The surface terms at $z_i=z_b$ vanish because the $\delta N(z)\delta(z-z_i)$ yields a factor of $\delta N(z_b)$, which vanishes at boundaries. Thus $\psi_{\rm kin}$ is annihilated by all constraint terms containing field strengths ${\rm F}^\mp_{zb}$. This is already enough to prove that the spatial diffeomorphism constraint C_z annihilates $\psi_{\rm kin}$, and hence $\psi_{\rm kin}$ is at least part of a kinematical basis, if not a physical basis.

The state would be physical if it were also annihilated by the first term in C_S, which I call C_E because of the ${}^{(2)}\tilde{\rm E}$ factor which it contains. The following is a sufficient condition for $\psi_{\rm kin}$ to be physical: C_E annihilates the state if it contains only two out of the four transverse fields:

$$\text{(either)}\quad\tilde{\rm E}^x_+\ \text{and}\ \tilde{\rm E}^x_-\ \text{only},$$
$$\text{(or)}\quad\tilde{\rm E}^y_+\ \text{and}\ \tilde{\rm E}^y_-\ \text{only}. \tag{52}$$

If ψ contains $\tilde{\rm E}^x_+$ and $\tilde{\rm E}^x_-$ only, for instance, the connection determinant $\epsilon_{ab}{\rm A}^+_a{\rm A}^-_b$ in C_E will necessarily annihilate the wavefunctional, since the indices a and b cannot both equal x. Note that the wavefunctional must have an *equal* number of $\tilde{E}^x_+$ and $\tilde{E}^x_-$ fields, because Gauss invariance requires an equal number of S_- and S_+ operators in the chain.

The result given in Eq. (52) can be generalized: $\psi_{\rm kin}$ is physical if every $\tilde{E}^{a_i}_{A_i}$ is the same linear combination

$$\tilde{E}^{a_i}_{A_i} = \alpha_x \tilde{E}^x_{A_i} + \alpha_y \tilde{E}^y_{A_i}, \tag{53}$$

where α_x and α_y are constants independent of i. In particular, every $\tilde{E}^{a_i}_{A_i}$ can be either an $\tilde{E}^+$ or an $\tilde{E}^-$, where

$$\tilde{E}^{\pm}_{A_i} = [\tilde{E}^x_{A_i} \pm i \tilde{E}^y_{A_i}]/\sqrt{2}. \tag{54}$$

These are eigenstates of global O(2) rotations mixing x and y. [All *local* transformations mixing x and y have been gauge fixed, but the Hamiltonian continues to be invariant under global O(2) rotations.] These linear combinations will be used in the next section to construct the eigenstates of the L_Z operator.

One can generate additional physical states, starting from those described by Eqs. (45) and (52), by applying operators G^x_y or G^y_x constructed by Husain and Smolin [22]. The G^a_b $(a,b=x,y)$ are integrals over z of weight one objects:

$$G^a_b = \int_{z0}^{z_{n+1}} dz \tilde{E}^a_A A^A_b. \tag{55}$$

Husain and Smolin have shown that the operators G^a_b commute with $\mathbf{H}$; they are physical. Hence application of m factors of G^y_x to a functional $\psi\,[\tilde{E}^x_+, \tilde{E}^x_-]$ replaces mx superscripts in the chain by y superscripts, but leaves ψ physical. The operators G^a_b are essentially raising and lowering operators for total intrinsic spin [10].

I have labeled the generators S_{Ai} in Eq. (42) using two quantum numbers j and m. (If the generator S_A is one which changes m, the m label can be the initial m value, say.) Once the SU(2) symmetry is broken to O(2), however, the j quantum number loses significance. I could replace the S_{Ai} by any other matrix with the same m (and Δm) but different j, and ψ would change only by a constant factor.

Even though the j has no physical significance, it is mathematically convenient to use S_A having definite j. One can then employ the familiar commutation relations of the S_A in calculations. Also, the planar state ψ is presumably a limit of some three-dimensional state for which the label j has meaning. The fact that states of different j are equivalent in the planar limit presumably means that the correspondence between three-dimensional and planar states is many to one.

V. APPLICATION TO SOLUTIONS

In this section I study the solutions constructed so far by applying several operators to them: the BPR operators (from Sec. III), the ADM energy operator (from paper III and Appendix C), the area operator $\tilde{E}^z_Z$ for areas in the xy plane (from paper III), and the operator L_Z giving the total spin angular momentum around the z axis (from paper IV and Appendix B). All the solutions (those constructed in papers II and III as well as the new solutions constructed in Sec. IV) have the form of strings of transverse $\tilde{E}^a_{\pm}(z_i)S_{\mp}$ operators separated by holonomies $M(z_{i+1},z_i)$. The solutions in papers II and III have additional step functions $\theta(z_{i+1}-z_i)$ which path order the integrations over the z_i, but the θ factors will play a minor role in the considerations of the present section.

Consider first the L_Z operator. From paper IV, this may be expanded as

$$\begin{aligned} L_Z &= 2\int dz[\tilde{E}^+_+ A^-_- - \tilde{E}^-_- A^+_+] \\ &= 2\hbar \int dz[\tilde{E}^+_+ \delta/\delta\tilde{E}^+_+ - \tilde{E}^-_- \delta/\delta\tilde{E}^-_-], \end{aligned} \tag{56}$$

where the fields $\tilde{E}^{\pm}_A$ are the global O(2) eigenstates introduced in Eq. (54). From Eq. (56), L_Z is determined by counting the number of $\tilde{E}^{\pm}_{\pm}$ fields in the wave functional ψ. Each such field contributes an amount $\pm 2\hbar$ to L_Z, while fields $\tilde{E}^{\pm}_{\mp}$ contribute nothing. Similarly when a connection representation is used, $\psi=\psi[A]$, each $A^{\pm}_{\pm}$ field contributes $\pm 2\hbar$.

Next consider the ADM energy operator. Often this is identified with the surface term in the Hamiltonian, but, as discussed in Appendix C, the volume term can also contribute. In the present case the volume term typically does contribute, but its only effect is to double the size of the surface term, and I will ignore volume contributions. The surface term is, from Eq. (C4),

$$\begin{aligned} H_{st} &= -\epsilon_{MN}\tilde{E}^b_M A^N_b|^{z_r}_{z_l} \\ &= i\hbar[\tilde{E}^b_- \delta/\delta\tilde{E}^b_- - \tilde{E}^b_+ \delta/\delta\tilde{E}^b_+]^{z_r}_{z_l}. \end{aligned} \tag{57}$$

When this operator acts upon a factor of $\tilde{E}^a_{\pm}(z_i)dz_i$ in the wave functional, it gives $\mp i\hbar \tilde{E}^a_{\pm} dz_i$ times a factor of $\delta(z_i - z_r)$ or $\delta(z_i - z_l)$. Obviously none of the solutions is an eigenfunction of the ADM energy, since the δ function deletes one integration dz_i. One could perhaps construct an eigenfunction by summing over an infinite number of solutions, each containing one more dz_i integration. Each additional integration should be multiplied by an additional factor of i, to cancel the i in Eq. (57) and make the eigenvalue real. Investigation of such sums is beyond the scope of the present paper. Without a measure one does not know whether such a sum converges to a normalizable result.

If the Gauss constraint

$$C_G = -i[\partial_z \tilde{E}^z_Z - \epsilon_{MN}\tilde{E}^a_M A^N_a] \tag{58}$$

vanishes at the boundaries, it can be used to simplify the ADM surface term to

$$\begin{aligned} H_{st} &= -\tilde{E}^z_{Z,z}|^{z_r}_{z_l} \\ &= \hbar(\delta/\delta A^Z_z)_{,z}. \end{aligned} \tag{59}$$

Unwanted factors of i are now more of a problem. A^Z_z occurs only in holonomies, where it is always multiplied by (a real matrix S_Z times) a factor of i, and the $\tilde{E}^z_Z$ will bring down this factor of i. Except for the solutions constructed in paper II, there is always at least one boundary where the Gauss constraint holds, so that factors of i will be a generic problem. Of course one could eliminate the problem by discarding the holonomy structure, but this is a solution almost as unattractive as the problem.

The operator $\tilde{\mathrm{E}}^z_Z{}_{,z}$ occurs in the ADM energy, while $\tilde{\mathrm{E}}^z_Z$ itself is the area operator for areas in the xy plane. ($\tilde{\mathrm{E}}^z_Z = e e^z_Z = {}^{(2)}e$.) Therefore the area operator also has pure imaginary eigenvalues, a situation already noted in Appendix D of paper II; see also DePietri and Rovelli [32].

Even if one for the moment ignores the factors of i, there is another problem with the area operator: At any boundary where the Gauss constraint is satisfied, the area operator will always give zero. If the Gauss constraint is satisfied, say, at the left boundary z_l, then there is no net flux exiting at z_l. One can regroup the holonomies until there are no $M(z_i,z_l)$, only $M(z_r,z_i)$; or until every $M(z_i,z_l)$ is paired with an $M(z_l,z_j)$ to give $M(z_i,z_l)\,M(z_l,z_j) = M(z_i,z_j)$. Either way, there is no holonomy depending on $\mathrm{A}^Z_z(z_l)$, and the area operator $\tilde{\mathrm{E}}^z_Z(z_l)$ gives zero.

Next consider the action of the BPR operators. The solutions given in paper II (and some of those considered in Sec. IV) contain either $\tilde{\mathrm{E}}^a_+$ operators or $\tilde{\mathrm{E}}^a_-$ operators, but not both. A wave functional which contains only $\tilde{\mathrm{E}}^a_-$ operators (for example) will be annihilated by $\mathrm{A}^-_a = \hbar\,\delta/\delta\tilde{\mathrm{E}}^a_+$, even before any semiclassical average is taken:

$$\mathrm{A}^-_a\,\psi[\tilde{\mathrm{E}}^a_-] = 0. \tag{60}$$

In classical theory, $\mathrm{A}^-_a = 0$ is a signal that the solution is purely left moving, and from this one might expect that $\psi[\tilde{\mathrm{E}}^a_-]$ is unidirectional. Condition (60) is too strong, however. As discussed in Eq. (26) of Sec. III, one expects at most a vanishing semiclassical average $\langle \mathrm{A}^-_a \rangle = 0$. In fact from the remarks on Lorentz gauge QED in the concluding paragraphs of Sec. III, Eq. (60) implies that ψ is probably not a normalizable state.

The solutions of paper III and most of those from Sec. IV contain both $\tilde{\mathrm{E}}^a_+$ and $\tilde{\mathrm{E}}^a_-$, and hence are not annihilated by any BPR operator. One cannot conclude that these solutions are infinite norm, therefore. However, they do suffer from the problems described previously in this section, those associated with the ADM energy and area operators.

VI. CONCLUSIONS

Papers II and III proposed new solutions to the constraints, and Sec. IV of the present paper proposes still more solutions. However, the investigations of Sec. V have demonstrated that these solutions are less than satisfying in several respects.

This outcome is perhaps not surprising. In earlier work, simplicity was the primary criterion for choosing the method of quantization (polarization and factor ordering), as well as the primary criterion for constructing solutions. Simplicity is the criterion one uses when information is scarce, however. In the present paper several operators of physical significance were available, and the theory and its solutions can be held to a standard more demanding than simplicity, the standard of a reasonable physical interpretation. As a result, solutions are called into doubt, but this happens for a reason which is fundamentally positive: More is known about how to interpret and understand the solutions.

The difficulties with imaginary eigenvalues encountered in Sec. V seem to require fundamental revisions in the theory, since the difficulties are closely linked to the complex nature of the connection. As discussed in Sec. V, the operator $\tilde{\mathrm{E}}^z_Z$, present in both the area operator and the ADM energy operator, has complex eigenvalues because of the i in the holonomies, $\exp[i\mathrm{A}^Z_z\cdots]$. Getting rid of the i entails dropping or modifying the holonomic structure, not a pleasant prospect. Recently Thiemann [20] has proposed an alternative formalism based upon a real connection. Thiemann's alternative is motivated primarily by issues of regularization, but has the desirable side effect of producing real eigenvalues for the area operator.

It is a little harder to see how switching to a real connection will cure the problem of zero eigenvalues of the area operator at boundaries, also uncovered in Sec. V. The zero eigenvalues occur only at boundaries where Gauss invariance is satisfied. Since $\tilde{\mathrm{E}}^z_Z$ is a Gauss invariant, at first sight any connection between area and Gauss invariance seems strange. The connection is indirect, via the structure of the complex connection $\mathrm{A}^Z_z = i\,\mathrm{ImA}^Z_z + \mathrm{ReA}^Z_z$. ImA^Z_z is the part which does not commute with $\tilde{\mathrm{E}}^z_Z$; therefore its presence in the wave functional gives rise to nonzero area. ReA^Z_z is the part which transforms like a connection under Gauss rotations; therefore it is needed in the wavefunctional for gauge invariance. The notions of area and gauge invariance are linked only because ImA and ReA are linked together to form a single (complex) connection. At a boundary where Gauss invariance is satisfied, if there is no net flux exiting, then there is no dependence on the connection and hence no area. In the Thiemann scheme one still joins ImA and ReA together to form a single (real) connection; in fact the Thiemann connection is just the Ashtekar connection without the factor of i. The real and imaginary parts of the connection are separated when constructing the Thiemann constraints. Must they be separated when constructing the wave functional as well? If the answer is yes, the wave functional would not be purely a product of holonomies.

Before doing anything as drastic as dropping the holonomic structure, it is a good idea to investigate what happens to the zero area argument when it is extrapolated from the planar case to the full, three-space-dimensional case. The argument that Gauss invariance leads to zero area depends on properties of the wavefunctional at boundaries, and the behavior at boundaries changes markedly with spatial dimension.

In the full three-dimensional case, the smearing function for the Gauss constraint must vanish at spatial infinity, so that there is no need for the Gauss constraint to annihilate the wave functional there. As a result, net flux may pass through the boundary at infinity, and there is no difficulty obtaining finite area at the boundary, even when the wave functional is purely a product of holonomies. This suggests that the zero area may be a problem which occurs only in the planar limit.

In fact the area problem may not exist in the planar limit, if one takes this limit correctly. Imagine the generic three dimensional flux configuration which is well approximated by planar symmetry: Near the origin, the flux lines corresponding to holonomies containing $A^Z_a dx^a$ are finite in cross section and well collimated along the z axis. The planar wave functionals constructed in papers II and III contain factors of $S_\pm$, presumably relics of Clebsch-Gordan coefficients coupling $A^Z_a dx^a$ holonomies to holonomies containing $A^X_a dx^a$ and $A^Y_a dx^a$. The latter are represented by flux lines lying in planes z=const. Although $A^Z_a dx^a$ flux lines are well collimated near the origin, they must diverge far out along z into the past or future. A sketch of the $A^Z_a dx^a$ flux lines resembles a drawing depicting radial geodesics near a wormhole: The flux lines come in from radial infinity, pass through a narrow ''throat'' oriented along z, and then diverge once more to radial infinity. (The wavefronts perpendicular to these rays are constructed from $A^X_a dx^a$ and $A^Y_a dx^a$ holonomies.) Alternatively, the $A^Z_a dx^a$ flux lines at infinity may not exit through the surface at infinity, but may loop back and close on themselves, resembling the flux lines of a solenoid in magnetostatics. Either outcome is allowed by the boundary conditions on the Gauss smearing function at infinity, and for either behavior at infinity, the behavior at the throat is the same. If one takes a cross section through two points z_l and $z_r > z_l$ at the throat, one finds net $A^Z_a dx^a$ flux through both boundary points z_l and z_r. If this picture of the three dimensional flux is correct, then in the planar limit one should *not* impose Gauss invariance at the boundaries. Planar solutions would resemble the ''open flux'' solutions studied in paper II. If there is $A^Z_a dx^a$ flux through the boundaries, the zero area problem at boundaries disappears.

Even though the $A^Z_a dx^a$ flux lines now extend throughout the entire range $z_l \leqslant z \leqslant z_r$, one can still construct localized wave packets. In paper II, I reviewed the geometrodynamical treatment of the planar problem, and introduced the Szekeres scalar fields B, W, and A. (It is a little easier to work out the boundary conditions for B, W, and A, rather than work directly with the Ashtekar fields; as shown in paper II, the boundary conditions on B, W, and A then imply corresponding boundary conditions on the Ashtekar variables.) When only the field A is present, the forces on a cloud of test particles are isotropic in the transverse (xy) direction; the elliptical distortions characteristic of gravitational waves appear only when the B and W fields are non-zero. (In a more covariant language, the components of the Weyl tensor which give rise to transverse deviations of geodesics are present only when B and W are non-zero.) One can impose wavepacket boundary conditions on B and W (equivalently, on transverse components of the Weyl tensor), requiring these quantities to vanish at the boundaries z_b, but this requirement tells us nothing about the behavior of A at the boundaries. The variable A determines geometrical quantities such as areas: The Ashtekar area operator $\tilde{E}^z_Z$ is just exp(A). Perhaps one has a ''wavepacket,'' but not a ''geometry packet.'' One may require localized, wavepacket behavior for B and W, but not for the more geometrical quantity A.

To summarize, there are two possible solutions to the zero area problem. The first splits the connection, abandoning the strict holonomic form for the wave functional. The second allows the Gauss constraint to be non-zero at boundaries and in effect assumes that the ''open flux'' boundary conditions used in paper II are generic. Further information and thought is needed before one can decide between these two alternatives.

Apart from the zero area difficulty, there is another reason why the transition from three to one space dimension needs more attention. Ultimately one would like to use the planar case as a guide to the behavior of radiation in the full, three dimensional case. In the full case, one expects the connections to occur in holonomies, so as to preserve gauge invariance. In the planar case, the gauge fixing allows the transverse connections to occur outside of holonomies. Two key radiative properties of the planar solutions, their directionality and spin, are associated with the transverse sector, which least resembles the three-dimensional case. It will probably be necessary to recast the transverse sector in a more holonomic language, in order to understand more clearly what happens on passing to the full theory.

For the moment let us overlook any possible difficulties with zero area and suppose that one shifts to a real connection, in order to eliminate the problem with imaginary eigenvalues. One can ask whether the solutions constructed in papers II and III and Sec. IV are likely to survive the shift to a real connection formalism. The present solutions may not survive, if the factor ordering is changed, and it is easy to imagine a reason why one might want to change the factor ordering. The scalar constraint usually must be taken to be non-Hermitian, in a complex connection formalism, whereas with a real connection one may wish to factor order so as to make the constraint Hermitian.

It may be helpful to comment briefly on why the scalar constraint is difficult to make Hermitian in a complex connection formalism. The usual recipe for constructing a self-adjoint operator is to factor order it, and then, if the operator is not self-adjoint, form the average $(C+C\dagger)/2$. $C\dagger$ is C, with the order of all operators reversed and the connections A replaced by $A\dagger$; in turn the $A\dagger$ are replaced by $-A + 2\,\mathrm{Re}A$. This last step introduces the unwanted non-polynomial expressions ReA into the $C\dagger$ term. In the case of the Gauss and spatial diffeomorphism constraints, identities may be used to eliminate the ReA contributions, and $C\dagger$ is well behaved, in fact identical to C. In the case of the scalar constraint, the unwanted ReA terms do not go away. Within a real connection framework, the $A\dagger$ is just A, and the traditional $(C+C\dagger)/2$ recipe is easier to implement.

Although the present solutions may not survive as exact solutions, they may constitute approximate solutions to the new, modified constraints, solutions valid in the limit $\hbar \to 0$. This would happen because the new and old scalar constraints presumably will differ only by a reordering of factors and hence will differ by terms of order $\hbar$. Also, the G^a_b operators, defined in Eq. (55) and shown to be constants of the motion by Husain and Smolin [22], are likely to remain constants of the motion, in any transition to a new operator ordering, because of the close connection between the G^a_b and total spin [10].

Even though the complex connection formalism may not be appropriate for the dynamics, this formalism is the natural one to use when constructing a criterion for the presence of radiation. Note that the BPR operators were derived in Sec. II using only symmetry considerations; no assumptions were made about the factor ordering or dynamics. Even if one dropped the Ashtekar connection and used the Thiemann connection, one would have to reintroduce the Ashtekar connection in order to express the results of Sec. II succinctly. Anyone familiar with the classical results on radiative criteria will not be surprised at this: Much of that work is most conveniently expressed using the language of complex connections. (See for example the work on the Weyl tensor quoted in Sec. I and Appendix D.)

Classically, exact plane wave solutions are known in which the area operator $\tilde{\mathrm{E}}^z_Z = {}^{(2)}e$ evolves to zero [16]. In fact this collapse behavior appears to be generic; solutions which do not collapse are rare and are unstable under small perturbations [33]. The zero area cannot be removed by a change of coordinates, since typically there is an accompanying singularity in a scalar polynomial quadratic in components of the Weyl curvature tensor. One can ask whether a quantum-mechanical effect might prevent this collapse. It is not possible to answer this question definitively within the present context, because the area operator has imaginary eigenvalues. Nevertheless, one can see the outlines of a possible quantum solution which would avoid a collapse. The quantum area operator $\tilde{\mathrm{E}}^z_Z$ acts on holonomies $\exp(i\int \mathrm{A}^Z_z S_Z dz)$; as long as S_Z is not allowed to assume the value zero, $\tilde{\mathrm{E}}^z_Z$ cannot have the eigenvalue zero. In the solutions constructed in papers II and III and Sec. IV, the S_Z value in each holonomy does not evolve dynamically and therefore remains non-zero if chosen to be non-zero initially. It remains to be seen whether this happy state of affairs will persist to a new formalism with real eigenvalues for the area operator.

APPENDIX A: DETAILS OF THE BPR CALCULATION

This appendix solves Eqs. (16) and (17) for the connection and tetrads obeying BPR symmetry, the invariance group for unidirectional plane gravitational waves. When setting up a complex connection formalism, it is necessary to choose three phases: When defining the Lagrangian at the four-dimensional level, one must choose the duality phase δ and the phase of ϵ_{TXYZ} [see for example Eq. (18)], and an additional phase comes in when rewriting the four-dimensional formalism in $3+1$ canonical form. These phases are explained in Appendix A of paper II, and I use the same phase choices here as in that paper. I begin at the four-dimensional level by solving Eq. (16) for the Lorentz transformation L;

$$L_{I'I} = -\partial_\beta \xi^\alpha e_{\alpha I'} e^\beta_I . \tag{A1}$$

I have dropped a $\xi^\lambda \partial_\lambda$ term which is zero because $\xi^{(c)}$, Eq. (13), is a linear combination of ∂_x, ∂_y, and ∂_v, all of which annihilate e^α_I. In the Rosen gauge, from Eq. (13),

$$\partial_\beta \xi^\alpha = \delta^u_\beta g^{c\alpha} - \delta^c_\beta g^{u\alpha}. \tag{A2}$$

Therefore

$$L_{I'I} = -e^c_{I'} e^u_I + e^u_{I'} e^c_I . \tag{A3}$$

L is antisymmetric, as it should be. Equation (16) determines L, Eq. (A3), but otherwise imposes no new constraints on the tetrads beyond those already imposed by ∂_x, ∂_y, and $\partial_v = 0$.

Next consider Eq. (17). In principle, one should be able to determine all the constraints on the A^A_a by solving these equations directly, but they are awkward, and it is easier to adopt an indirect approach. Given the tetrads, compute the Lorentz connection ω^{IJ}_α; then compute ${}^{(4)}A$, which is just the self-dual version of ω. In this way one finds that many components of ${}^{(4)}A$ are identically zero. When this information is inserted into Eq. (17), that equation reduces to the trivial statement $0=0$, for most values of the indices; for a small number of index values the equation is non-trivial and can be solved with moderate effort.

The equation relating ω to the tetrads is

$$\omega_{ija} \equiv e_{iI} e_{jJ} \omega^{IJ}_a$$
$$= [-g_{aj,i} + g_{ai,j} + e_{jK} \overleftrightarrow{\partial}_a e^K_i]/2. \tag{A4}$$

From this equation, at least one of i, j, or a must be u, since derivatives with respect to x,y,v are zero. Further, the tetrad matrix in the Rosen gauge is 2×2 block diagonal, with the zt to ZT block containing constants only. This implies that *at most* one of i, j, or a must be u. After stripping off the (block diagonal) tetrads, one finds that the only non-zero ω are

$$\omega^{XY}_u , \omega^{VA}_a , \tag{A5}$$

where $a=x,y$ only and $A=X,Y$ only. Now compute ${}^{(4)}A$ from ω, using

$$2\,{}^{(4)}\mathrm{A}^{IJ}_a = \omega^{IJ}_a + i\,\delta(\epsilon^{IJ}_{..MN}/2\epsilon_{TXYZ})\,\omega^{MN}_a , \tag{A6}$$

where the duality eigenvalue $\delta\epsilon_{TXYZ}$ equals $+1$, given my conventions $\delta = \epsilon_{TXYZ} = 1$. The only non-zero ${}^{(4)}A$ components are

$${}^{(4)}\mathrm{A}^{XY}_u , {}^{(4)}\mathrm{A}^{TZ}_u , {}^{(4)}\mathrm{A}^{V+}_a . \tag{A7}$$

At first glance one might think this list is too short; there should be non-zero ${}^{(4)}\mathrm{A}^{U\pm}_a$ as well, because the $\epsilon^{IJ}_{..MN}$ term in Eq. (A6) will map the VA indices on ω^{VA}_a into UB ($B = X,Y$ or $\pm$). Duality maps VX into VY, *not* UY, however, because of the identities

$$\epsilon_{TXYZ} = \epsilon_{UVXY} = \epsilon^{VX}_{..VY} , \tag{A8}$$

etc., where the last, mixed index tensor with two V's is the one which occurs in duality relations. This explains why there are no ${}^{(4)}\mathrm{A}^{U\pm}_a$; to see what happened to ${}^{(4)}\mathrm{A}^{V-}_a$, note

$$2\,{}^{(4)}\mathrm{A}^{V\pm}_a = \omega^{V\pm}_a + i\,\delta(\epsilon^{V\pm}_{..MN}/2\epsilon_{TXYZ})\,\omega^{MN}_a$$
$$= \omega^{V\pm}_a + i\,\delta(\epsilon^{V\pm}_{..V\mp}/\epsilon_{TXYZ})\,\omega^{V\pm}_a$$
$$= \omega^{V\pm}_a (1 \pm \delta). \tag{A9}$$

For my phase choice $\delta = +1$, ${}^{(4)}\mathrm{A}^{V-}_a$ vanishes.

From Eq. (A7), there is no need to consider $\alpha = v$ in Eq. (17). I consider first $\alpha = a = x,y$. The first term in Eq. (17) is a linear combination of ∂_x, ∂_y, and ∂_v, which vanishes for any α. The second term vanishes from Eq. (A2) and the absence of any $\lambda=v$ component of ${}^{(4)}A$. Then Eq. (17) collapses to

$$0=0+0+\mathcal{L}^{I}_{.I'}{}^{(4)}\mathrm{A}_a^{I'J}+\mathcal{L}^{J}_{.J'}{}^{(4)}\mathrm{A}_a^{IJ'}-0. \tag{A10}$$

From Eqs. (A3) and (18), the only non-zero elements of L, are

$$\mathcal{L}_{UX}=i\mathcal{L}_{UY}=(e^c_X+ie^c_Y)/2=e^c_+/\sqrt{2} \tag{A11}$$

or

$$\mathcal{L}_{U+}=e^c_+,\quad \mathcal{L}_{U-}=0. \tag{A12}$$

Inserting this and Eq. (A7) into Eq. (A10), one finds that all the $\alpha=a=x,y$ equations are trivially $0=0$.

Finally, consider $\alpha=u$. The first term in Eq. (15) vanishes as before. The only IJ index pair which does not give $0=0$ is $IJ=VA$, $A=X,Y$ only, which gives

$$0=0+\partial_u\xi^{\lambda}\,{}^{(4)}\mathrm{A}_\lambda^{VA}+\mathcal{L}^{V}_{.I'}{}^{(4)}\mathrm{A}_u^{I'A}+\mathcal{L}^{A}_{.J'}{}^{(4)}\mathrm{A}_u^{VJ'}-\partial_u\mathcal{L}^{VA}. \tag{A13}$$

Use Eq. (A2) to simplify the first term; use Eq. (A12) to simplify the remaining terms and to show that the $A=-$ equation is trivial. Then

$$\begin{aligned}0&=0+g^{ca}\mathrm{A}_a^{V+}-2e^c_+\mathrm{A}_u^{-+}+\partial_u e^c_+\\&=g^{ca}\mathrm{A}_a^{V+}-2e^c_B\mathrm{A}_u^{B+}+\partial_u e^c_+.\end{aligned} \tag{A14}$$

I have used the duality relation $A^{VU}=-iA^{XY}$ and $A^{-+}=iA^{XY}$. On the second line, recall that a ''plus'' index always pairs with a ''minus'' index to form the two-dimensional dot product: $e^c_BA^{B+}=e^c_+A^{-+}+0$. The A_u field in Eq. (A14) is a linear combination of A_z and A_t fields, and the A_t fields are non-dynamical Lagrange multipliers for the Gauss constraints. I therefore eliminate the A_u field in order to obtain a constraint on the dynamical field A_a^{V+}. From duality and Eqs. (A4) and (A5) for ω,

$$\begin{aligned}2e_{aB}^{(4)}\mathrm{A}_u^{B+}&=e_{aB}[\omega_u^{B+}+i(\delta/2\epsilon_{TXYZ})\epsilon^{B+}_{MN}\omega_u^{MN}]\\&=[e^{j+}\omega_{aju}+0]\\&=e^{j+}[e_{jK}\overleftrightarrow{\partial}_u e^K_a]/2.\end{aligned} \tag{A15}$$

I solve Eq. (A14) for A^{VA} and insert Eq. (A15):

$$\begin{aligned}{}^{(4)}\mathrm{A}_a^{V+}&=e^{j+}[e_{jK}\overleftrightarrow{\partial}_u e^K_a]/2-g_{ca}\partial_u e^c_+\\&=e^{j+}[2e_{jK}\partial_u e^K_a-\partial_u g_{ja}]/2-\partial_u e_{a+}+\partial_u g_{ca}e^c_+\\&=\partial_u g_{ca}e^c_+/2.\end{aligned} \tag{A16}$$

The right-hand side of Eq. (A16) is proportional to the part of A_a^{V+} which contains no $i\delta$ factor:

$$\begin{aligned}2\,\text{``Re''}\,{}^{(4)}\mathrm{A}_a^{V+}&\equiv(\omega_a^{VX}+i\omega_a^{VY})/\sqrt{2}\\&=e^{iV}(e^{Xj}+ie^{Yj})\omega_{ija}/\sqrt{2}\\&=e^{+j}\partial_u g_{aj}/2.\end{aligned} \tag{A17}$$

Therefore

$$\begin{aligned}0&=-{}^{(4)}\mathrm{A}_a^{V+}+2\,\text{``Re''}\,{}^{(4)}\mathrm{A}_a^{V+}\\&={}^{(4)}\mathrm{A}_a^{V+}\quad\text{for }\ \delta=-1.\end{aligned} \tag{A18}$$

The second line means that the first line is the four dimensional connection computed with the opposite choice for the duality eigenvalue, $\delta=-1$ rather than $\delta=+1$.

This result is very easy to transform from the Rosen to a general gauge in the z,t sector, since the ''minus'' and ''a'' indices in the x,y sector remain invariant under such a transformation. In order to maintain the gauge conditions (14) and (15) on the tetrads, it is necessary to combine any four-dimensional diffeomorphism $(z,t)\rightarrow(z't')$ with a Lorentz transformation, as in Eq. (16). A short calculation shows that a very simple Lorentz transformation $L^T_{.Z}=\partial t'/\partial z$ will maintain all the gauge conditions of Eqs. (14) and (15). For transforming ${}^{(4)}A$ one needs the corresponding $\mathcal{L}$; the only non-zero matrix elements will be $\mathcal{L}^T_{.Z}$ and $\mathcal{L}^X_{.Y}$ or, equivalently, $\mathcal{L}^U_{.U}$, $\mathcal{L}V^V_{.V}$, and $\mathcal{L}^{\pm}_{.\mp}$. Thus the coordinate transformation to the general gauge amounts to a Lorentz transformation which multiplies Eq. (A18) by an overall factor:

$${}^{(4)}\mathrm{A}_a^{V+\prime}=\mathcal{L}^V_{.V}\mathcal{L}^+_{.-}\,{}^{(4)}\mathrm{A}_a^{V+}\quad\text{for }\ \delta=-1, \tag{A19}$$

where every $\mathcal{L}$ and every ${}^{(4)}A$ is to be calculated using the $\delta=-1$ convention. From Eq. (A19), the quantity in Eq. (A18) vanishes in every gauge. Similarly, from Eq. (A7), the following quantities vanish in every gauge:

$$\begin{aligned}0&=-{}^{(4)}\mathrm{A}_a^{U+}+2\,\text{``Re''}\,{}^{(4)}\mathrm{A}_a^{U+}\\&={}^{(4)}\mathrm{A}_a^{V-}\\&={}^{(4)}\mathrm{A}_a^{U-}.\end{aligned} \tag{A20}$$

Equations (A20) and (A18) may be rewritten as

$$\begin{aligned}0&={}^{(4)}\mathrm{A}_a^{T+}-2\,\text{``Re''}\,{}^{(4)}\mathrm{A}_a^{T+}\\&={}^{(4)}\mathrm{A}_a^{Z+}-2\,\text{``Re''}\,{}^{(4)}\mathrm{A}_a^{Z+}\\&={}^{(4)}\mathrm{A}_a^{T-}\\&={}^{(4)}\mathrm{A}_a^{Z-},\end{aligned} \tag{A21}$$

where every connection in Eq. (A21) is evaluated using the $\delta=+1$ convention. For the opposite duality convention, $\delta=-1$, exchange $(+\leftrightarrow-)$ everywhere in Eq. (A21). For $\delta=+1$ but left-moving rather than right moving waves, again exchange $(+\leftrightarrow-)$ in Eq. (A21).

So far the calculation has been carried out entirely at the four-dimensional level. The four-dimensional connection ${}^{(4)}A$ is related to the usual $3+1$ connection A by the following equation from Appendix A of paper II:

$$\mathrm{A}_a^S=-\epsilon^{(4)}_{MNS}\mathrm{A}_a^{MN}. \tag{A22}$$

Then the four-dimensional equation (A21) implies the $3+1$ dimensional equations

$$\begin{aligned}0&=\mathrm{A}_a^-,\\0&=-\mathrm{A}_a^++2\,\text{``Re''}\,\mathrm{A}_a^+.\end{aligned} \tag{A23}$$

Again, exchange $(+\leftrightarrow-)$ for the opposite duality convention or left-moving waves.

APPENDIX B: KINEMATICS OF THE $A_A^{\pm}$ FIELDS: SPIN

From paper IV, the integral

$$L_Z=i\int dz[\tilde{E}_I^y(A_x^I-\mathrm{Re}A_x^I)-(x\leftrightarrow y)] \tag{B1}$$

gives the total spin angular momentum of the wave, and is a constant of the motion [10]. The integral is over the entire wavepacket, that is, from z_l to z_r. As in paper II, the fields and Weyl tensor components which produce transverse displacements of test particles are assumed to vanish at the boundaries, with support only in the region $z_l<z<z_r$. It is at first sight surprising that any conserved quantity associated with the Lorentz group should be given by a volume integral (integral over z) rather than by a surface term (term evaluated at the end points z_l and z_r). However, in the one-dimensional planar case, the extensive gauge fixing in the x,y plane removes all gauge freedom, except for rigid rotations around z, and the x,y sector of the theory resembles special relativity rather than general relativity.

This appendix rewrites the integrand of L_Z in terms of the unidirectional fields [Eq. (19) for right-moving waves, and two more amplitudes with $+\leftrightarrow-$ for left-moving waves], in order to understand the spin content of these fields. Introduce triads e_a^A and inverse triads e_A^a, and write the integrand of L_Z as

$$\begin{aligned}\tilde{E}_I^y\ \mathrm{Im}\ A_x^J-(x\leftrightarrow y)&=[(e_J^y e_{Kx}-(x\leftrightarrow y)]\tilde{E}_K^a\ \mathrm{Im}A_a^J\\ &=[(e_{[J}^y e_{K]x}+e_{(J}^y e_{K)x}-(x\leftrightarrow y)]\tilde{E}_K^a\ \mathrm{Im}A_a^J.\end{aligned} \tag{B2}$$

In the last line the term antisymmetric in J,K is proportional to $\tilde{E}_K^a\ \mathrm{Im}A_a^J\epsilon_{JK}$. This expression is part of the Gauss constraint $\partial_z\tilde{E}_Z^z+\tilde{E}_K^a A_a^J\epsilon_{JK}=0$, which implies $\tilde{E}_K^a\ \mathrm{Im}A_a^J\epsilon_{JK}=0$. Hence the term antisymmetric in J,K can be dropped. The term symmetric in J,K can be expanded in O(2) eigenstates, keeping in mind that every $+$ index must be contracted with a $-$ index. The $J\neq K$ terms are proportional to

$$\begin{aligned}e_+^y e_{-x}+e_-^y e_{+x}&=e_B^y e_{Bx}\\ &=\delta_x^y\\ &=0.\end{aligned} \tag{B3}$$

Hence these terms can be dropped also. The surviving terms are products of tensors with J=K and therefore helicity ±2 in the local Lorentz frame, a reassuring result:

$$\begin{aligned}L_Z&=-\int dz[e_+^y e_{+x}\tilde{E}_-^a\,``\mathrm{Im}"\,A_a^-\\ &\quad+e_-^y e_{-x}\tilde{E}_+^a\,``\mathrm{Im}"\,A_a^+-(x\leftrightarrow y)]\\ &=i\int dz\{e_+^y e_{+x}\tilde{E}_-^a[A_a^-+(A_a^--2``\mathrm{Re}"\,A_a^-)\\ &\quad+e_-^y e_{-x}\tilde{E}_+^a[A_a^++(A_a^+-2``\mathrm{Re}"\,A_a^+)\}-(x\leftrightarrow y).\end{aligned} \tag{B4}$$

In the last line I have written L_Z in terms of the weight one combinations of unidirectional BPR fields introduced in Eq. (21). This expression for gravitational spin angular momentum possesses the same coordinate times momentum structure as the corresponding expression for electromagnetic spin angular momentum:

$$\begin{aligned}\vec{L}_{em}&=(1/4\pi)\int d^3x[\vec{E}\times\vec{A}]\\ &=-\int d^3x[\vec{\Pi}\times\vec{A}].\end{aligned} \tag{B5}$$

That is, one can interpret the unidirectional quantities $\tilde{E}$ A and $\tilde{E}$ (A-2"Re"A) in Eq. (B4) as momenta associated with waves of definite helicity. This parallel with QED does not extend too far: These "momenta" have nothing like free-field commutation relations with each other or with the triad "coordinates."

It is now clear why one wants the two combinations $\tilde{E}_+^a A_a^+$ and $(-A_a^-+2``\mathrm{Re}"A_a^-)\tilde{E}_-^a$ to vanish: These two constraints remove left-moving helicity $\pm2$ contributions from $L_Z$. Why must the remaining helicity zero combinations vanish? The helicity zero combinations are $\tilde{E}_-^a A_a^+$ and $(-A_a^-+2``\mathrm{Re}"A_a^-)\tilde{E}_+^a$. They are complex conjugates of each other, so that by adding and subtracting them from each other one gets pure imaginary and pure real constraints

$$0=\tilde{E}_B^a(A_a^B-\mathrm{Re}A_a^B)-i\epsilon_{AB}\ \mathrm{Re}A_a^A\tilde{E}_B^a, \tag{B6}$$

$$0=-i\epsilon_{AB}(A_a^A-\mathrm{Re}A_a^A)\tilde{E}_B^a+\tilde{E}_B^a\ \mathrm{Re}A_a^B. \tag{B7}$$

Now consider the classical equation of motion

$$\begin{aligned}0&=-i\tilde{E}_{Z_t}^z-\delta H/\delta A_z^Z\\ &=-i\tilde{E}_{Z_t}^z-\tilde{E}_B^a A_a^B.\end{aligned} \tag{B8}$$

On the second line I use the Hamiltonian of Eq. (46). I also use the unidirectionality assumption (for the first time in this section; L_Z is the spin operator also for the scattering case) and evaluate the Hamiltonian in Rosen (or at least conformally flat) gauge. The metric in a general gauge has the form

$$\begin{aligned}ds^2=&\{[-(N')^2+(N^z)^2]dt^2+2N^z dz dt+dz^2\}g_{zz}\\ &+x,y\ \ \text{sector},\end{aligned} \tag{B9}$$

so that to obtain conformal gauge, one must take $N^z=0$ and $N'=1$ in the Hamiltonian, where N^z is the shift and N' is the renormalized lapse defined following Eq. (46). From the real part of Eq. (B8), the $\widetilde{\mathrm{E}}$ ReA term in Eq. (B7) vanishes. The rest of this equation is just the imaginary part of the Gauss constraint, $\partial_z \widetilde{\mathrm{E}}^z_Z + \epsilon_{AB} \mathrm{A}^A_a \widetilde{\mathrm{E}}^a_B = 0$, and vanishes also. This leaves Eq. (B6). The ϵ Re$\mathrm{A}\widetilde{\mathrm{E}}$ term may be simplified using the real part of the Gauss constraint; the $\widetilde{\mathrm{E}}(\mathrm{A} - \mathrm{Re}\, A)$ term may be simplified using the imaginary part of the equation of motion, Eq. (B8). The result is simply

$$0 = -i(\partial_t + \partial_z) \widetilde{\mathrm{E}}^z_Z, \tag{B10}$$

in any conformally flat gauge. This equation is discussed further in Eq. (22) of Sec. II.

APPENDIX C: THE ADM ENERGY

It is a worthwhile exercise to express the ADM energy in terms of BPR operators. In the usual three-space dimensional case, the Hamiltonian expressed in terms of the original ADM variables as the sum of a volume integral plus a surface term H_{st} [34,35],

$$H = \int d^3x [N C_{sc} + N^i C_i] + H_{st}. \tag{C1}$$

In the classical theory, the ADM energy is just the surface term, since the constraints in the volume term must vanish everywhere when the solution obeys the classical equations of motion. Often one says that in the quantum case the ADM energy is just the surface term also, but this is not quite right, as we shall see in a minute.

In the planar, one-space dimensional case, the expression for the Hamiltonian in terms of Ashtekar variables looks superficially much the same as Eq. (C1) [8],

$$\int dz [N' C_S + N^z C_z + N_G C_G] + H_{st}, \tag{C2}$$

except for the additional Gauss constraint and the prime on N'. (The prime means I have renormalized the usual Ashtekar lapse by absorbing a factor of $\widetilde{\mathrm{E}}^z_Z$ into the lapse, as explained in paper II.) In both one and three spatial dimensions, one might be tempted to drop the volume terms, in the quantum mechanical case, because the constraints C_i are required to annihilate the wavefunctional. However, the statement that the scalar constraint (say) annihilates the wavefunctional means, not $C_S \psi = 0$, or even $\int dz C_S \psi = 0$, but rather

$$\int dz\, \delta N' C_S \psi = 0, \tag{C3}$$

where $\delta N'$ is a small change in the lapse. The arbitrary change $\delta N'$ must preserve the boundary condition at spatial infinity, $N' \to 1$. Hence $\delta N' \to 0$ there. On the other hand, when C_S occurs in the Hamiltonian of Eq. (C2), it is multiplied by N' rather than $\delta N'$. There is no need for N' to vanish at the boundaries (in fact it becomes unity there). Now suppose the evaluation of the action of C_S on ψ requires an integration by parts with respect to z. In Eq. (C3), when the constraint acts upon ψ, surface terms at $z = z_b$ will vanish because of the boundary condition on $\delta N'$. When H acts upon ψ, however, the C_S in H is smeared by N' rather than $\delta N'$; the former does not vanish at boundaries. Consequently the volume term can contribute to the ADM energy in the quantum case. In the planar case both the Gauss and scalar constraints in the volume term can contribute to the ADM energy; neither N' nor N_G is required to vanish at the boundaries. In the usual $3+1$ dimensional case with flat space boundary conditions at infinity, only the scalar constraint in the volume term can contribute; the remaining constraints are smeared by N_i which are required to vanish at spatial infinity.

For the plane wave case, the surface terms in the Hamiltonian were computed in Sec. 4 of paper II:

$$\begin{aligned} H_{st} &= -\epsilon_{MN} \widetilde{\mathrm{E}}^b_M \mathrm{A}^N_b \big|^{z_r}_{z_l} \\ &= I[\widetilde{\mathrm{E}}^b_- \mathrm{A}^+_b - \widetilde{\mathrm{E}}^b_+ \mathrm{A}-_b] \\ &= i\hbar [\widetilde{\mathrm{E}}^b_- \delta/\delta \widetilde{\mathrm{E}}^b_- - \widetilde{\mathrm{E}}^b_+ \delta/\delta \widetilde{\mathrm{E}}^b_+]^{z_r}_{z_l} \end{aligned} \tag{C4}$$

To simplify the boundary term quoted in paper II, I have invoked the boundary conditions $N^z \to 0$, $N' \to 1$ on the shift and renormalized lapse. Evidently the ADM energy contains the BPR operators which are sensitive to the long-range scalar potential, which suggests that these operators may play a role even in the presence of waves which are not unidirectional.

The operator of Eq. (C1) gives a finite result when applied to the solutions of papers II and III; there is no need to renormalize. However, the solutions are not eigenfunctions of this operator. A typical solution involves n integrations dz_i over the locations of the $n\widetilde{\mathrm{E}}^a_A(z_i)$ operators contained in the wavefunctional, and the ADM operator acts as a "lowering operator," removing one integration. Hence an eigenfunction would have to be an infinite sum over wavefunctionals of all possible values of n. It is beyond the scope of this paper to investigate the finiteness of the norm of such a sum.

APPENDIX D: THE TRANSVERSE WEYL CRITERION

1. Classification of Weyl tensors

The Weyl tensor is the part of the Reimann tensor which can be non-zero even in empty space, and certain of its components induce transverse vibrations when inserted into the equation of geodesic deviation [36]. It is therefore a natural object to work with when constructing a criterion for the presence of radiation [37]. The construction proceeds in two steps. The first step is a straightforward mathematical problem: Classify Weyl tensors using their algebraic properties. In the second step, one uses physical arguments to determine the Weyl class(es) most closely associated with radiation.

To begin with the mathematical problem, there are 10 independent real components of the Weyl tensor, and from these one can construct 5 independent complex components which have simple duality properties.

$$\mathcal{C}_{abcd} = [C_{abcd} + i(\delta/2\epsilon_{TXYZ}) \epsilon_{abmn} C^{mn}_{cd}]/2, \tag{D1}$$

$$\mathcal{C}_{abcd} = i(\delta/2\epsilon_{TXYZ}) \epsilon_{abmn} \mathcal{C}^{mn}_{cd}. \tag{D2}$$

Lowercase Roman indices $a,b,c,\ldots$ are global; uppercase Roman indices $A,B,C,\ldots$ are local Lorentz. ϵ_{abmn} is the totally antisymmetric global tensor, while ϵ_{TXYZ} is the corresponding local Lorentz quantity, the Levi-Civitá constant tensor. The duality eigenvalue is $\delta/\epsilon_{TXYZ}=\pm 1$. There is another Levi-Civitá tensor hidden in the ϵ_{abmn},

$$\epsilon_{abmn}=e_a^A e_b^B e_c^C e_d^D \epsilon_{ABCD};$$

therefore ϵ_{TXYZ} and its associated sign convention drop out after the conversion to Ashtekar variables and the $3+1$ splitup. The final $3+1$ Hamiltonian contains only the phase δ. My convention is $\delta=+1$, but in this paper all results are stated in a manner which facilitates a conversion to the opposite convention. Of course the combinations with simple duality properties also have simple transformation properties in the local Lorentz frame, which is why one chooses to work with $\mathcal{C}$ rather than C, when attempting a classification.

Petrov was the first to classify Weyl tensors by their algebraic properties [11], but for present purposes the equivalent classification scheme due to Debever [12,13] is more convenient. A null vector k is said to be a principal null vector (Debever vector) of $\mathcal{C}$ if

$$k_{[a}\mathcal{C}_{b]mn[c}k_{d]}k^m k^n=0; k_a k^a=0. \tag{D3}$$

Debever proved that a Weyl tensor can have up to four distinct principal null vectors, and he classified Weyl tensors by the number of degeneracies among these vectors. If [1111] denotes the Weyl tensors which have 4 distinct Debever vectors, [211] the Weyl tensors with two vectors degenerate and the rest distinct, etc., then the five classes are [1111], [211], [22], [31], and [4]. (The corresponding five Petrov classes are I, II, D, III, and N, respectively.)

Now suppose k is a Debever vector obtained by solving Eq. (D3). Make it one leg of a null tetrad k, l, m, $\overline{m}$. Choose the Z axis of a local free fall frame so that k and l have spatial components along $\pm Z$, while m and $\overline{m}$ are transverse:

$$-k_a l^a = m_a \overline{m^a}=1,$$

$$k_a k^a = l_a l^a = m_a m^a$$

$$= k_a m^a = l_a m^a = 0. \tag{D4}$$

C may be expanded in this basis, and (not surprisingly) one gets five possible terms:

$$\mathcal{C}_{abcd}=C_1 V_{ab}V_{cd}+C_2(V_{ab}M_{cd}+M_{ab}V_{cd})+C_3(M_{ab}M_{cd} - U_{ab}V_{cd}-V_{ab}U_{cd})+C_4(U_{ab}M_{cd}+M_{ab}U_{cd}) + C_5 U_{ab}U_{cd}, \tag{D5}$$

where

$$V_{ab}=2k_{[a}m_{b]},$$

$$U_{ab}=2l_{[a}\overline{m_{b]}},$$

$$M_{ab}=2k_{[a}l_{b]}+2m_{[a}\overline{m_{b]}}. \tag{D6}$$

The five combinations in Eq. (D5) are the only ones allowed by the duality convention $\delta=+1$. The expansion for the opposite duality convention may be obtained from Eqs. (D5) and (D6) by interchanging m and $\overline{m}$ in Eq. (D6). Equation (D5) is essentially the expansion given by Szekeres [36], after a relabeling of the basis vectors $(k,l,m,\overline{m}) \rightarrow (k,-m,t,\overline{t})$. Since the expansion treats k and l quite symmetrically, it is valid also for the case that l, rather than k is the principal null vector.

At this point one turns from the mathematical to the physical: Which Petrov/Debever class(es) or which term(s) in Eq. (D5) are most closely associated with radiation? Consider first which of the five tensors in Eq. (D5) distorts a cloud of test particles in the manner expected for gravitational radiation. Szekeres finds that only the C_1 and C_5 terms produce the transverse displacements in the XY plane characteristic of gravitational radiation in the linearized theory. C_2 and C_4 produce longitudinal displacements in the XZ or YZ planes. C_3 produces a Coulomb (or tidal force) displacement: C_3 distorts a sphere of particles into an ellipsoid of revolution with axis along Z. These facts suggest that the C_1 and C_5 terms signal the presence of radiation.

There is another set of arguments which suggest that the C_1 and C_5 terms are closely associated with the presence of radiation. If the Weyl tensor contains *only* a C_1 or C_5 term, the tensor is type N. (A type N tensor with k as principal null vector contains only a C_1 term; a type N tensor with l as principal null vector contains only a C_5 term.) Type N is closely associated with radiation. Along characteristic curves, when the metric is discontinuous, the discontinuity in the Weyl tensor is type N [38]. In the linearized theory, the tensor associated with unidirectional gravitational radiation is type N. At large distances from bounded sources, the surviving components of the Weyl tensor are type N ("peeling theorem" [13]).

Although the C_1 and C_5 terms are closely associated with type N, it would be better to call these terms transverse Weyl components, rather than type N components, since a tensor which is not type N can nevertheless contain C_1 or C_5 terms. A [22] (type II) field contains C_1 plus some admixture of longitudinal component, while [1111] (type I) contains C_1, C_5, and C_3 terms. In a theory as non-linear as general relativity, one can expect that a collision between C_1 and C_5 transverse waves will produce some C_3 (Coulomb) component, and the tensor will be type I rather than type N. In asymptotic regions, after the transverse wave has "outrun" its Coulomb companion, presumably the tensor will revert to type N, but in general one should be looking for C_1 and C_5, rather than type N or any other specific Petrov class. One should describe this radiation criterion as the transverse Weyl criterion, rather than the type N criterion.

I now construct operators which project out the C_1 and C_5 terms:

$$C_1=\mathcal{C}_{abcd}l^a\overline{m^b}l^c\overline{m^d},$$

$$C_5=\mathcal{C}_{abcd}k^a m^b k^c m^d, \tag{D7}$$

$$C_3=-\mathcal{C}_{abcd}l^a\overline{m^b}k^c m^d.$$

For completeness I have included also the expression for the pure Coulomb component C_3. The plane wave case has no longitudinal components; C_2 and C_4 vanish identically.

2. Transition to Ashtekar variables

At this point I specialize to the case of plane waves along the Z axis. By definition, the plane wave metric has two hypersurface orthogonal null vectors which may serve as normals to right- and left-moving wavefronts $U=(cT-Z)/\sqrt{2}=$const and $V=(cT+Z)/\sqrt{2}=$const. I identify these normals with k (right-moving) and l (left-moving), so that a small change in the wave phase will look like

$$k_a dx^a=(-dT+dZ)/\sqrt{2}=-dU,$$

$$l_a dx^a=(-dT-dZ)/\sqrt{2}=-dV. \tag{D8}$$

[Hypersurface orthogonality demands $k_a dx^a \propto dU$, etc.; the normalization conditions force the constants of proportionality to be as shown in Eq. (D8), and the overall phase of k_a and l_a is fixed by the requirement that k^0 and l^0 be positive, i.e. future pointing.] Lowercase x denotes a global coordinate; uppercase (T,Z,U,V) denotes a coordinate in a local Lorentz frame. From the expression for the (inverse) tetrads, $e_b^A dx^b=dX^A$, k and l may be identified with the tetrads

$$k^a=-e^{aU}=+e_V^a,$$

$$l^a=e_U^a. \tag{D9}$$

Similarly,

$$m^a=e_+^a,$$

$$\overline{m^a}=e_-^a. \tag{D10}$$

Evidently the quantities C_i are (global scalars and) tensors in a Local Lorentz frame.

Equation (D8) places quite a strong restriction on the null basis, going beyond what is required to maintain the normalization equation (D4). The choice equation (D9) certainly is not unique. For example, k_a remains null if it is rescaled by an arbitrary function. (Simultaneously l_a must be rescaled by the inverse function, in order to maintain the normalization condition $-k_a l^a=1$.) Similarly m and $\overline{m}$ may be rescaled. The choice equations (D9) and (D10) facilitate calculations and lead to highly symmetric formulas for C_1 and C_5. [In this basis, C_1 is just C_5 with some plus and minus indices interchanged; see Eq. (D20) below.] For further discussion of the effect of choice of basis, see the remarks following Eq. (D20).

Conversion to the Ashtekar language is straightforward. $\mathcal{C}$ is essentially the (four dimensional) Ashtekar field strength, since the $\mathcal{C}$ tensor is self-dual, and in empty space the Weyl tensor is the full Riemann tensor:

$${}^{(4)}F_{cd}^{AB}=e^{Aa}e^{Bb}\mathcal{C}_{abcd}. \tag{D11}$$

The four-dimensional field strengths ${}^{(4)}F$ may be replaced by $3+1$ quantities F by using standard formulas:

$${}^{(4)}F_{cd}^{TM}=-i\sigma\delta F_{cd}^{M}/2,$$

$${}^{(4)}F_{cd}^{MN}=\sigma\epsilon_{MNS}F_{cd}^{S}/2. \tag{D12}$$

$M,N,S=$space only. σ is a new phase which appears at the $3+1$ reduction step. I choose $\sigma=-1$, for reasons explained in Appendix A of paper II. This phase (unlike δ) merely changes the overall sign of the C_i, and I shall not keep track of the σ dependence in the future. Useful corollaries of Eq. (D12) are

$${}^{(4)}F_{cd}^{V\pm}=(i/\sqrt{2})F_{cd}^{\pm}[(\delta\pm1)]/2,$$

$${}^{(4)}F_{cd}^{U\pm}=(i/\sqrt{2})F_{cd}^{\pm}[(\delta\mp1)]/2. \tag{D13}$$

If tetrad equations (D9) and (D10) and field strength equation (D13) are inserted into Eq. (D7) for C_1, the result for $\delta=+1$ is

$$C_1=i[-F_{cd}^{+}e_T^c+F_{cd}^{+}e_Z^c]e_+^d/2. \tag{D14}$$

(For $\delta=-1$ replace $+$ by $-$.) For any metric, typically the Lorentz boosts are gauge fixed by demanding that three of the tetrads vanish:

$$e_M^t=0,\quad M=\text{space}. \tag{D15}$$

For the special case of the plane wave metric, the gauge fixing of the XY Gauss constraint and xy spatial diffeomorphism constraints imply that four more tetrads vanish:

$$e_X^z=e_Y^z=e_Z^x=e_Z^y=0. \tag{D16}$$

The tetrad matrix reduces to two 2×2 subblocks which link x,y to X,Y (or $\pm$) and z,t to Z,T. Therefore the first term in Eq. (D14) (and only the first term) contains an F_{td}^{+} term, $d=x$ or y, with unacceptable time derivatives of the "coordinate" A_d^+. I eliminate this term using the classical equations of motion, which are

$${}^{(4)}F_{cd}^{AB}e_A^c=0, \tag{D17}$$

or after $3+1$ splitup, and setting $B=+$,

$$\begin{aligned}0&={}^{(4)}F_{cd}^{A+}e_A^c\\&={}^{(4)}F_{cd}^{-+}e_+^c+{}^{(4)}F_{cd}^{U+}e_U^c+{}^{(4)}F_{cd}^{V+}e_V^c\\&=F_{cd}^{+}e_V^c=F_{cd}^{+}e_T^c+F_{cd}^{+}e_Z^c.\end{aligned} \tag{D18}$$

In the second line the ${}^{(4)}F^{U+}$ term vanishes because of Eq. (D13), and the ${}^{(4)}F_{cd}^{-+}e_+^c$ may be dropped because at the next step the entire term will be contracted with e_+^d. When Eq. (D18) is inserted into Eq. (D14), the result is

$$C_1=iF_{cd}^{+}e_Z^c e_+^d. \tag{D19}$$

This is not quite Ashtekar form, because the triads must be densitized. Also, the c index can equal z only, because of the gauge conditions (D15) and (D16). The final result is (for C_5 and C_3 also, since they are calculated similarly)

$$C_1 = iF^{+}_{zd}\tilde{\mathrm{E}}^{d}_{+}/{}^{(2)}\tilde{\mathrm{E}},$$

$$C_5 = iF^{-}_{zd}\tilde{\mathrm{E}}^{d}_{-}/{}^{(2)}\tilde{\mathrm{E}},$$

$$C_3 = F^{Z}_{xy}/2\tilde{\mathrm{E}}^{z}_{Z}. \tag{D20}$$

${}^{(2)}\tilde{\mathrm{E}}$ is the determinant of the 2×2 XY subblock of the tetrad matrix. The results for $\delta=-1$ are the same, except for overall phases, and interchange of $+$ and $-$ everywhere.

In the case that the wave is unidirectional, the results (D20) are consistent with the BPR constraints. For example, if the wave is right moving, then the principal vector is k, associated with the tensor C_1. From Eq. (19), A^{-}_{a} vanishes, implying that (C_1 is finite, while) C_5 vanishes.

The C_i of Eq. (D20) were calculated in a specific basis. In particular the factors of ${}^{(2)}\tilde{\mathrm{E}}$ in Eq. (D20) are basis dependent, and will change if the basis is changed. For example, suppose one shifts from the tetrad basis, Eq. (D9), to a basis in which k^a is affinely parametrized ($k^b k_{a;b} = \lambda k_a$, $\lambda \neq 0$). Then the ${}^{(2)}\tilde{\mathrm{E}}$ factor in C_1 disappears, replaced by a factor of $\tilde{\mathrm{E}}^{z}_{Z}$. Could such a change of basis make C_1 kinematical, physical, or consistent? [Kinematical, C_1 commutes with the Gauss and spatial diffeomorphism constraints; physical, C_1 commutes with all the constraints; consistent, the commutator (C_1, constraint) is proportional to C_1. Consistency is relevant only if C_1 is to be set equal to zero.]

It is unlikely that C_1 or any transverse operator could be made consistent or physical. Gravitational radiation is closely identified with transversality only in the linearized theory. In the full classical theory, scattering of two transverse waves produces a C_3 Coulomb component [16]. Presumably, then, the commutator or Poisson brackets of any purely transverse operator with the Hamiltonian will not be especially simple, even in the classical theory. (This is one reason why the main body of the paper concentrates on the BPR operators, rather than the Weyl tensor.)

Although it is unlikely that any transverse criterion could be made physical or consistent, a kinematical criterion for transversality should be feasible in some cases and useful, for example, to monitor the amount of C_1 amplitude present initially, for either case $C_1=0$ or $C_1\neq0$. One could rescale C_1 by change of basis, until it became density weight unity and then integrate it over z; the resulting expression would commute with the diffeomorphism constraint. It is less clear how to make an expression that is Gauss invariant. If the wave were unidirectional, one could multiply $C_1(z)$ by a holonomy stretching from z to the right-hand boundary (flux tube open to the right). Since the wave is right moving, no signal has reached the right-hand boundary as yet, and the boundary condition on the Gauss smearing function is flat space, $\delta N_G=0$. This means the Gauss constraint does not have to commute with C_1 at the boundary, and the flux exiting through the right boundary causes no problems. I do not know how to handle the scattering case, in which waves run in both directions.

APPENDIX E: CONSISTENCY OF THE BPR CONSTRAINTS

This appendix calculates the commutators between the BPR operators and the scalar constraint. The formalism used is geometrodynamical [7]. One starts from the three-metric g_{ij} and canonical momentum π^{ij}, gauge-fixes two canonical pairs (g_{xz}, π^{xz}) and (g_{yz}, π^{yz}) to zero, and then carries out a canonical transformation to a set of four variables A, D, B, W with canonical momenta π_A, π_D, $\pi_B \cosh W$, and π_W. Note the unexpected canonical pair ($\pi_B \cosh W$,B). Also, since the constraints contain e^A more often than A, it is convenient to use ($\pi_A e^{-A}$,e^A) rather than (π_A,A) as the canonical pair. It may be helpful to note that in the linearized limit the variables A, D disappear, while the variables B, W represent the two polarizations of gravitational radiation. In terms of these variables the metric is

$$\begin{aligned} ds^2 = {} & \{[-(N')^2+(N^z)^2]dt^2 + 2N^z dz dt + dz^2\}\exp(D-A/2) \\ & + e^A[e^B \cosh W dx^2 - 2 \sinh W dx dy \\ & + e^{-B}\cosh W dy^2]. \end{aligned} \tag{E1}$$

Two constraints survive the gauge fixing: the scalar constraint C_S and the constraint C_z generating spatial diffeomorphisms along z:

$$\begin{aligned} C_S = {} & 2(e^A)_{,zz} - (e^A)_{,z}D_{,z} - \pi_A(e^{-A})_{,z}\pi_D \\ & + e^A[(B_{,z}\cosh W)^2 + (W_{,z})^2]/2 \\ & + e^{-A}[(\pi_B)^2 + (\pi_W)^2]/2, \end{aligned} \tag{E2}$$

$$\begin{aligned} C_z = {} & -2(\pi_D)_{,z} + (e^A)_{,z}\pi_A(e^{-A})_{,z} + D_{,z}\pi_D \\ & + B_{,z}\cosh W \pi_B + W_{,z}\pi_W. \end{aligned} \tag{E3}$$

I will not check explicitly the commutator with the C_z constraint, since the commutator is trivially consistent: [BPR amplitude, C_z] $\sim \partial_z$ (BPR amplitude), just what one would expect, since C_z generates z diffeomorphisms. In fact, since the BPR amplitudes are scalar densities, their integral over z will commute with C_z.

At a later point in the discussion I will need to eliminate time derivatives of $A,D,\ldots$ in favor of momenta $\pi_A,\pi_D,\ldots$ by using the classical equations of motion for $A,D,\ldots$. These follow from the Hamiltonian, which is a sum of the above constraints.

$$H = \int dz[N'C_S + N^z C_z] + \text{S.T.} \tag{E4}$$

S.T. denotes the surface term, which was worked out in [7] but will not be needed explicitly here, since it does not contribute to the classical equations of motion. N' is not quite the usual lapse N (and C_S is not quite the usual scalar constraint):

$$N' = N/\sqrt{g_{zz}}. \tag{E5}$$

In the planar case this renormalization of the scalar constraint allows the constraints themselves to be consistent; that is, a commutator of two constraints is always a sum of constraints [8].

I now have the constraints; the next step is to calculate the BPR unidirectionality constraints in terms of the A, D, B, W variables and their conjugate momenta. One must compute the Lie derivatives of the four-dimensional quantities $g_{\mu\nu}$ along the direction of the BPR Killing vector equation (13) for right-moving waves, and set these derivatives equal to zero. In the Rosen gauge equation (12) these equations are easy to solve.

The only non-trivial components are the transverse space-space components, which obey

$$\partial_v g_{ab}=0, a,b=x,y. \tag{E6}$$

The two Killing vectors of equation (13) therefore give no new information in the Rosen gauge (and therefore in any gauge), since we know already that one of the five BPR Killing vectors is ∂_v. Something similar happens in the Ashtekar formalism, where Eq. (16) for the tetrads gives no new information about the tetrads, but merely determines the Lorentz transformation L. Note that we would be almost done at this point, if we were working in a covariant, second order formalism ($g_{\mu\nu}$ and $\partial_t g_{\mu\nu}$ as variables, rather than g_{ij} and π_{ij}, $i,j=$space). It would only be necessary to transform Eq. (E6) from Rosen coordinates (u,v) to a general gauge (z',t'), and this is easy. The transverse space-space components g_{ab} are scalars under the transformation; it is only necessary to transform ∂_v.

Since it is so easy to do the calculation in a covariant second order formalism, I will continue with this formalism a bit longer. Shift from the Rosen $g_{\mu\nu}$ to A, B, W, and D as variables; Eq. (E6) becomes

$$\partial_v A=\partial_v B=\partial_v W=0. \tag{E7}$$

The functions g_{ab} contain only A, B, and W in the Rosen gauge; hence there is no constraint $\partial_v D=0$. In fact D is identically zero in the Rosen gauge. Since D is totally an artifact of the gauge transformation from Rosen to general coordinates, it can depend on both u and v.

Next, transform ∂_v to general coordinates. Note that

$$\begin{aligned}0&=g_{vv}=\partial_v x'^{\mu}\partial_v x'^{\nu} g'_{\mu\nu}\\ &=(\partial_v t')^2[-N'^2+(N^z)^2]+2\partial_v t'\partial_v z' N^z+(\partial_v z')^2,\end{aligned} \tag{E8}$$

where the second line follows from Eq. (E1) (after priming the z,t variables in that equation). The second line is a product of two factors, so that

$$N'\partial_v t'=N^z\partial_v t'+\partial_v z', \tag{E9}$$

where I have chosen the root that reduces correctly in the Rosen limit $N'\to 1$, $N^z\to 0$. Now note that

$$\begin{aligned}\partial_v&=\partial_v t'\partial_{t'}+\partial_v z'\partial_{z'}\\ &\propto\partial_{t'}+[\partial_v z'/\partial_v t]'\partial_{z'}\\ &=\partial_{t'}+(N'-N^z)\partial_{z'},\end{aligned} \tag{E10}$$

where the last line follows from Eq. (E9). From this result, Eq. (E7) becomes, in the general gauge,

$$\begin{aligned}0&=[\partial_t+(N'-N^z)\partial_z]A\\ &=[\partial_t+(N'-N^z)\partial_z]B\\ &=[\partial_t+(N'-N^z)\partial_z]W.\end{aligned} \tag{E11}$$

I have dropped primes on z and t.

The constraints of Eq. (E11) are still second order, but it is straightforward to eliminate the time derivatives in favor of the momenta, by using the classical equations of motion $\dot{q}=\{q,H\}$ to derive

$$\begin{aligned}0&=(e^A),_t-N^z(e^A),_z+N'\pi_D,\\ 0&=B,_t-N^zB,_z-N'\pi_B e^{-A}/\cosh W,\\ 0&=W,_t-N^zW,_z-N'\pi_W e^{-A}.\end{aligned} \tag{E12}$$

(Remember that $\pi_B\cosh W$ is the momentum conjugate to B.) Eliminating the time derivatives between Eqs. (E11) and (E12) gives the BPR constraints for right-moving waves, in first order form:

$$\begin{aligned}0&=\pi_B+B,_z\cosh We^A,\\ 0&=\pi_W+W,_z e^A,\\ 0&=-\pi_D+(e^A),_z.\end{aligned} \tag{E13}$$

For left-moving waves use g_{uu} in place of g_{vv} in Eq. (E8), replace N' by $-N'$ in Eq. (E11), and change the sign of all z derivatives in Eq. (E13). One reassuring feature of Eq. (E13) is their gauge independence. All factors of N' and N^z have dropped out.

Now both the scalar constraint and the BPR amplitudes have been expressed in geometrodynamical variables, and all that remains is to commute the constraint equation (E2) with the amplitude equation (E13). The first two amplitudes of Eq. (E13) are the ones which give difficulties. For example,

$$\begin{aligned}&[\pi_W+W,_z e^A, C_S(z')]\\ &=[-i\hbar\,\delta(z-z'),_{z'}+i\hbar\,\delta(z-z')e^{-A}\pi_D]\\ &\quad-i\hbar\,\delta(z-z')e^{-A}\pi_W[(e^A),_z-\pi_D]\\ &\quad+i\hbar\,\delta(z-z')\tanh We^{-A}[(\pi_B)^2-(e^AB,_z\cosh W)^2].\end{aligned} \tag{E14}$$

The right-hand side of this commutator is of the correct form (BPR amplitudes to the right) except for the square brackets in the last line. This has the $(\pi)^2-(q,_zQ)^2$ form discussed following Eq. (2). As shown there, this term cannot be rewritten with BPR amplitudes to the right, so that the strong form of consistency (BPR amplitudes annihilate the wavefunctional) does not hold.

Before going on to consider the weaker form of consistency, note that this troublesome term is multiplied by a factor of tanh W, so that if the wave is linearly polarized ($W=\pi_W=0$), one might suppose that the BPR constraints are consistent. However, the next commutator to be studied will suggest that a unidirectional wave is unstable even if initially linearly polarized, since a non-zero W can be created out of the vacuum.

As discussed in the Introduction, in order to check for the consistency of the weaker condition, one should examine an amplitude which cannot fluctuate in sign. The amplitude $(\pi_W(z)+W,_z e^A)^2$ is the simplest choice, because its commutator can be computed in terms of the known commutator, Eq. (E14):

$$\begin{aligned}[(\pi_W(z)+W,_z e^A)^2, C_S(z')] &= (\pi_W+W,_z e^A)[\pi_W+W,_z e^A, C(z')]+[\pi_W+W,_z e^A, C(z')](\pi_W+W,_z e^A)\\ &= -2i\hbar\,\delta(z-z')_{z'}(\pi_W(z)+W,_z e^A)^2+i\hbar\,\delta(z-z')\partial_z(\pi_W(z)+W,_z e^A)^2\\ &\quad -2i\hbar\,\delta(z-z')(\pi_W(z)+W,_z e^A)^2\pi_D - i\hbar\,\delta(z-z')[\pi_W e^{-A}(\pi_W+W,_z e^A)\\ &\quad +(\pi_W(z)+W,_z e^A)\pi_W e^{-A}][(e^A),_z-\pi_D]+\cdots\,. \end{aligned} \tag{E15}$$

The terms indicated by the ellipsis contain the $(\pi_B)^2-(e^A B,_z \cosh W)^2$ term in Eq. (E14) and will presumably average to zero. The first two terms of the commutator resemble the corresponding free-field QED commutator, which has no production of left-moving waves from the vacuum. Several terms are cross terms, products of two BPR amplitudes, rather than perfect squares. Again, these lines are likely to average to zero. This leaves a term on the second line from the end, which is a perfect square times π_D. Unless the latter function fluctuates, this term is unlikely to average to zero.

At least in classical theory, π_D is unlikely to fluctuate in sign, in the manner that left-moving π_B, π_W, B, and W fields fluctuate, because π_D is quadratic in those fields. For orientation, consider calculating π_D in classical theory, in a conformally flat gauge $N^z=0$, $N'=1$. Since two gauges have been fixed, it is necessary to eliminate two coordinates and their canonical momenta. It is natural to eliminate A and D and keep the coordinates B and W, since these represent the two polarizations in the linear limit. If the wave is unidirectional, it is consistent with the equations of motion in this gauge to choose $D=\pi_A=0$. Then e^A and π_D may be eliminated by solving the constraints (E3). The constraint which determines π_D is $C_z=0$. This constraint relates $(\pi_D),_z$ to a difference of *squares* of BPR amplitudes for left- and right-moving waves:

$$\begin{aligned}C_z = &-2(\pi_D),_z+e^{-A}(B,_z \cosh W e^A+\pi_B)^2/4\\ &-e^{-A}(B,_z \cosh W e^A-\pi_B)^2/4+e^{-A}(W,_z e^A+\pi_W)^2/4\\ &-e^{-A}(W,_z e^A-\pi_W)^2/4.\end{aligned} \tag{E16}$$

Hence π_D is unlikely to fluctuate in sign even if the left-moving BPR amplitudes fluctuate. Therefore the $(\pi_W+W,_z e^A)^2\pi_D$ term in Eq. (E15) is unlikely to average to zero, and the weaker condition also suggests that the amount of left-moving BPR amplitude will change with time. Note that this term creates non-zero W amplitude even when the wave is initially linearly polarized ($W=0$ initially).

It is instructive to repeat these heuristic arguments for the interacting QED case, where Schwinger has worked out the correct answer. The condition for no left-moving x-polarized wave is $\pi_x+A^x,_z=0$. The QED Hamiltonian in radiation gauge is [39]

$$H=\int dz[(\pi_x)^2+(A^x,_z)^2-2A^xJ^x+\cdots]/2, \tag{E17}$$

where the ellipsis indicates terms which commute with the BPR amplitude, and $J^x=e\overline{\psi}\gamma^x\psi$ is the usual current. Then

$$[\pi_x+A^x,_z, H]=i\hbar[\partial_z(\pi_x+A^x,_z)+J^x], \tag{E18}$$

$$[(\pi_x+A^x,_z)^2, H]=i\hbar[\partial_z(\pi_x+A^x,_z)^2+2J^x(\pi_x+A^x,_z)]. \tag{E19}$$

The interaction adds additional J^x terms which are *not* perfect squares. One expects

$$\langle J^x\rangle \sim \langle \chi\dagger\sigma^x\eta\rangle, \tag{E20}$$

where χ and η are positron and electron two-spinors, and the right-hand side has the correct parity because the particle and antiparticle have opposite intrinsic parity. The spin of the vacuum should average to zero.

[1] A. Ashtekar, Phys. Rev. D **36**, 1587 (1987).

[2] C. Rovelli and L. Smolin, Nucl. Phys. **B331**, 80 (1990).

[3] For applications of loops to gauge theories in general, see R. Gambini and A. Trias, Phys. Rev. D **22**, 1380 (1980); Nucl. Phys. **B278**, 436 (1986).

[4] C. Rovelli and L. Smolin, Phys. Rev. D **52**, 5743 (1995).

[5] B. Brügmann, R. Gambini, and J. Pullin, Phys. Rev. Lett. **68**, 431 (1992).

[6] R. Gambini, J. Pullin, Phys. Rev. D **54**, R5935 (1996).

[7] D. E. Neville, Class. Quantum Grav. **10**, 2223 (1993). This paper is referred to as paper I in the text.

[8] D. E. Neville, Phys. Rev. D **55**, 766 (1997). This paper is referred to as paper II in the text.

[9] D. E. Neville, Phys. Rev. D **55**, 2069 (1997). This paper is referred to as paper III in the text.

[10] D. E. Neville, Phys. Rev. D **56**, 3485 (1997). This paper is

referred to as paper IV in the text.
[11] A. Z. Petrov, *Einstein Spaces* (Fizmatgiz, Moscow, 1961).
[12] R. Debever, C. R. Acad. Sci. (Paris) **249**, 1324 (1959).
[13] R. Sachs, Proc. R. Soc. London **A205**, 463 (1962).
[14] R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960).
[15] J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).
[16] P. Szekeres, J. Math. Phys. **13**, 286 (1972).
[17] H. Bondi, F. Pirani, and I. Robinson, Proc. R. Soc. London **A251**, 519 (1959).
[18] W. B. Bonnor, Commun. Math. Phys. **13**, 163 (1969).
[19] P. C. Aichelburg, Acta Phys. Austriaca **34**, 279 (1971).
[20] T. Thiemann, Phys. Lett. B **388**, 257 (1996); ‘‘Quantum spin dynamics,’’ I and II, gr-qc/9606089-90; ‘‘Closed formula for the matrix elements of the volume operator,’’ gr-qc/9606091; ‘‘A length operator for canonical quantum gravity,’’ gr-qc/9606092.
[21] J. Schwinger, Phys. Rev. **82**, 664 (1951); reprinted in *Quantum Electrodynamics*, edited by J. Schwinger (Dover, New York, 1958).
[22] V. Husain and L. Smolin, Nucl. Phys. **B327**, 205 (1989).
[23] N. Rosen, Phys. Z. Sowjetunion **12**, 366 (1937).
[24] P. Peldan, Class. Quantum Grav. **11**, 1087 (1994).
[25] B. Brügmann, R. Gambini, and J. Pullen, Nucl. Phys. **B385**, 587 (1992); B. Brügmann and J. Pullen, *ibid.* **B390**, 399 (1993).
[26] C. Rovelli and L. Smolin, Nucl. Phys. **B442**, 593 (1995); A. Ashtekar, C. Rovelli, and L. Smolin, Phys. Rev. Lett. **69**, 237 (1992).
[27] R. Loll, Phys. Rev. Lett. **75**, 3048 (1995).
[28] C. J. Isham and K. Kuchař, Ann. Phys. (N.Y.) **164**, 288 (1985); J. J. Halliwell and J. B. Hartle, Phys. Rev. D **43**, 1170 (1991).
[29] J. Belinfante, Phys. Rev. **76**, 226 (1949); S. T. Ma, *ibid.* **75**, 535 (1949).
[30] W. Heitler, *The Quantum Theory of Radiation*, 3rd ed. (Clarendon Press, Oxford, 1954).
[31] P. A. M. Dirac, *Lectures on Quantum Field Theory* (Belfer Graduate School of Science, Yeshiva University, New York, 1966).
[32] R. De Pietri and C. Rovelli, Phys. Rev. D **54**, 2664 (1996).
[33] U. Yurtsever, Phys. Rev. D **36**, 1662 (1987).
[34] B. DeWitt, Phys. Rev. **160**, 1113 (1967).
[35] T. Regge and C. Teitelboim, Ann. Phys. (N.Y.) **88**, 286 (1974).
[36] P. Szekeres, J. Math. Phys. **7**, 751 (1966).
[37] For a review of classical criteria for the presence of radiation, see the monograph by V. D. Zakharov, *Gravitational Waves in Einstein’s Theory*, Israel program for scientific translation (Halstead Press and Wiley, New York, 1973).
[38] F. Pirani, in *Gravitation: An Introduction to Current Research*, edited by Louis Witten (Wiley, New York, 1962).
[39] J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, MA, 1967).