PHYSICAL REVIEW D **73,** 124004 (2006)

Volume operator for spin networks with planar or cylindrical symmetry

Donald E. Neville*

Department of Physics, Temple University, Philadelphia 19122, Pennsylvania, USA

(Received 7 December 2005; published 5 June 2006)

This paper constructs a kinematic basis for spin networks with planar or cylindrical symmetry, by exploiting the fact that the basis elements are representations of an O(3) subgroup of O(4). The action of the volume operator on this basis gives a difference equation for the eigenvalues and eigenvectors of the volume operator. For basis elements of low spin, the difference equation can be solved readily on a computer, yielding the eigenvalues and eigenvectors. For higher spins, I solve for the eigenvalues using a WKBJ method. This paper considers only the case where the gravitational wave can have both polarizations. The single polarization case is considered in a separate paper.

DOI: 10.1103/PhysRevD.73.124004 PACS numbers: 04.60.Pp, 04.30.−w

I. INTRODUCTION

This paper calculates the eigenvalues of the volume operator for spaces with planar or cylindrical symmetry. The volume operator plays a key role in Thiemann's construction of a finite Hamiltonian for spin network canonical quantum gravity [1]

The quantization of systems with planar or cylindrical symmetry has been studied using both local field theory [2–5] and spin networks [6,7]. Before one can construct a spin network Hamiltonian for these systems, one must construct a volume operator. Bojowald has given a discussion of some of the difficulties involved in constructing this operator [8].

Systems with planar or cylindrical symmetry are the simplest where gravitational wave propagation can occur. Simpler systems, such as homogeneous cosmologies and black holes with spherical symmetry, have higher symmetry and fewer degrees of freedom, but do not allow gravitational waves. In a homogeneous cosmology [9], for example, every point is equivalent to every other point. One may choose any one point, or vertex, of the spin network as representative. The Hamiltonian changes intertwiners and holonomies at that one vertex only. Since the action of the Hamiltonian is purely local, there can be no propagation of gravitational waves from point to point. For black holes with spherical symmetry, with or without matter [8,10–14], spherical symmetry rules out gravity waves.

The construction of a loop formalism typically proceeds in two steps. The states of the Hilbert space must satisfy seven constraints. In the first step one constructs a basis for the Hilbert space which satisfies the six constraints which are easiest to treat (Gauss and spatial diffeomorphism constraints). The hardest constraint, the scalar or Hamiltonian constraint, is left for the second step. The basis constructed in the first step is the kinematical basis, and the dot product constructed in this step is the kinematical dot product. This structure is enough to determine the eigenvalues of the volume operator.

This paper completes the first step. It constructs the kinematical basis and dot product, and writes down equations which allow the eigenvalues and eigenvectors of the volume operator to be constructed numerically. The paper does not go on to the second step and construct a Hamiltonian.

The rest of this paper proceeds as follows. Section II describes the topology of the spin network. Section III defines the volume operator and sets up the kinematical basis. Section IV writes out the eigenvalue equation for the volume operator in this basis and derives an analytical solution for eigenvectors with zero eigenvalue. For nonzero eigenvalues, and for basis functions of low spin, the equation can be programmed readily and solved on a computer. Section V derives a WKBJ method to find approximate eigenvalues of the volume operator. This approach was intended for high spins, but WKBJ surprises (as it often does). It yields accurate eigenvalues even for basis functions of relatively low spin.

This paper does not impose the constraints which limit the gravitational waves to a single polarization. In classical gravitational theory, and in quantum field theories of gravitation, the single polarization case is easier. However, in the spin network case I found the single polarization case to be harder. I will leave a discussion of that case for a separate paper.

II. TOPOLOGY OF THE SPIN NETWORK

A system with planar or cylindrical symmetry has two commuting, spacelike Killing vectors. In the planar case, if one shifts to the coordinates x, y suggested by the Killing vectors, the seven constraints simplify considerably, because derivatives with respect to x and y may be dropped. (For the cylindrical case, replace the planar coordinates x, y, z by coordinates ϕ, z, r. For simplicity in what follows, I shall discuss primarily planar symmetry, and devote only an occasional remark to the cylindrical case). In particular, the x, y diffeomorphism constraints, and the X, Y Gauss

*Electronic address: dneville@temple.edu

constraints simplify enough that they can be solved and eliminated from the theory. Following Husain and Smolin, I fix these four constraints by imposing the four gauge fixing conditions [7]

$$\tilde{E}^x_Z = \tilde{E}^y_Z = 0; \qquad \tilde{E}^z_X = \tilde{E}^z_Y = 0. \tag{1}$$

Lower case letters $x, y, z, \ldots$ denote global coordinates; upper case $X, Y, Z, \ldots$ denote local coordinates rotated by the Gauss constraints. Setting the four constraints equal to zero and solving yields four more equations,

$$A^Z_x = A^Z_y = 0; \qquad A^X_z = A^Y_z = 0. \tag{2}$$

Both the triad and connection arrays are now block diagonal. A 2×2 block contains fields with indices x, y and X, Y; a 1×1 block contains the field with indices z, Z. (For cylindrical symmetry substitute ϕ, z, r and Φ, Z, R for x, y, z and X, Y, Z). The SU(2) local gauge symmetry has been reduced to local U(1), although the theory still contains all three generators S_i of su(2). In order to express the consequences of the U(1) symmetry as clearly as possible, I choose the usual matrix representation of su(2) where S_z is diagonal. I suppose that the reduction in components, Eqs. (1) and (2) has been carried out classically, and I now set up a spin network formalism to quantize the reduced theory.

In classical general relativity, the theory allows one to choose virtually any coordinate system; but it is not always clear which coordinates lead to the simplest equations. The choice of coordinates is not obvious unless the system has a lot of symmetry.

Similarly in the spin network approach, the choice for the topology for the network is not obvious unless the system has a lot of symmetry. I will arrive at the topology for the planar case via the group theoretical approach favored by practitioners in the field of quantum cosmology. This approach has been discussed extensively in the literature [15,16], and for the most part my discussion will be a summary of results.

As a first step in constructing the symmetry reduced theory, one uses group theory to construct a connection which embodies the symmetry. Modern treatments manage to avoid anything so *déclassé* as solving the Killing equations, but in the end results are the same. The group theory determines the number of independent components and their form. It also determines the support of the connection fields. (In effect this step has been carried out already at Eqs. (1) and (2), where we determined the number of independent components, and determined their support to be the z axis).

However, group theory by itself does not supply the topology of the spin network. For example, in a homogeneous cosmology, every point of the space is equivalent to every other point. Not surprisingly then, the group theory predicts support for the homogeneous connection is limited to a single point. We know that, in the full theory, each vertex is connected to some number of edges. In effect, the group theory supplies a vertex (the point) but no edges. Motivated by the full theory, one introduces edges, and promotes the connections to holonomies integrated over these edges.

In the cosmological case, one can justify the introduction of edges *á posteriori* by constructing a Bohr compactified Hilbert space which has some very unusual and useful properties [17]. But it is unlikely anyone would have thought of doing this, had they not been motivated by the existence of edges and holonomies in the full, nonsymmetric, theory.

Similarly, in the planar case, the group theory predicts that support for the connections is the z axis, but produces no edges. Motivated by the full theory, one introduces vertices along this axis, as well as two edges radiating from each vertex in x and y directions. The connections are then promoted to holonomies integrated over the edges.

Holonomies on the z axis look like holonomies in the full theory. Each holonomy is integrated from one vertex to the next. Holonomies on the x and y edges are treated differently. These edges are given the topology of a circle; the two ends of each edge are identified.

This identification is as first sight somewhat puzzling. (Most of the rest of this section records my initial doubts, and eventual acceptance of circular topology for the xy edges. Readers who are comfortable with this choice of topology may wish to skip to the last three paragraphs of this section). One would like to think of the symmetric theory as a reduction of a full theory. That is, whenever one has a symmetry, one expects to start from the action of the full Hamiltonian on the full space, and end with the action of a simpler Hamiltonian on a smaller space. The ultimate example of such a reduction is a homogeneous cosmology, where one starts from a Hamiltonian acting on every point, and ends with a simpler Hamiltonian acting on a space consisting of a single, representative point.

It is therefore natural to visualize the full spin network as constructed from a basic unit, or module, which is repeated over and over to create the full space. For the homogeneous cosmology the basic module would be a unit cell containing only a single vertex, plus three edges. For the edges one could choose (say) three edges extending in positive coordinate directions, with the midpoint of each edge at the vertex. The basic cell would then be six half-edges radiating from a central point. Repeating this cell over and over generates the full space. (If the system possesses additional symmetries such as isometry, one may need fewer than three edges. A change in the number of edges does not affect the present argument, and I will ignore this possibility).

To return to my point: the modular picture does not immediately suggest an S_1 topology for each of these edges. Each basic cell contains an intertwiner at the vertex, plus six holonomies associated with the half-edges. When

the full space is obtained by multiplying these basic units together, one index on each holonomy is already contracted with an index on an intertwiner; as for the remaining index, one might expect to contract it with the corresponding index on a neighboring cell. It is not obvious one should contract this index with another index in the same cell, which is what happens when one identifies ends and imposes an S_1 topology on each edge.

If one does not identify ends and contract indices, each cell would have six "dangling" SU(2) indices, indices not contracted with any other index. Those indices can be ignored when Gauss-rotating the cell, however. The uncontracted indices may be viewed as merely an artifact of splitting the full space up into identical cells.

However, the "dangling index" picture does not hold up well when we consider reduction from the full to the symmetry reduced theory; circular topology seems essential. In more detail (and continuing with the homogeneous example for simplicity), label each cell by its vertex v and label each half-edge e by an index j_e, which denotes the total spin of the holonomy on that edge (if one is using $2j_e + 1$ dimensional irreducible representations of SU(2) on each edge); or j_e denotes the number of spin $1/2$ holonomies (if one is using a product of spin $1/2$ matrices on each edge). By homogeneity, the two halves of a given edge must have the same j_e. Then the spin network wavefunctional for the full space is

$$\psi = \Sigma_{je} \Pi_v \psi_v(j_e)_m. \tag{3}$$

ψ_v denotes the wavefunctional of a single cell. The full space is a product of such cells (the Π); and one must sum over possible assignments of edge spins (the Σ). The subscript m is shorthand for the six SU(2) indices on ψ_v. These indices are contracted with corresponding indices on neighboring cells.

Since every point is equivalent to every other, the modules must identical, initially, and the dynamics must keep them identical. Dynamics consists of identical transformations applied to every cell. In this spirit, I write the Hamiltonian constraint as a sum of constraints, one for each unit cell.

$$H = \Sigma_v H_v.$$

The constraint $0 = H\psi$ implies

$$0 = \Sigma_{je} H_v \psi_v(j_e)_m [\Pi_{v' \neq v} \psi_{v'}(j_e)]_m \tag{4}$$

I would like this to reduce to

$$0 = \Sigma_{je} H_v \psi_v(j_e)_m,$$

i.e. to a single term in H acting on a single cell of ψ. However, this will not happen, in general, because the m indices on the square brackets in Eq. (4) have ranges which depend on the j_e. The square brackets therefore do not factor out of the sum over j_e.

I can, of course, eliminate the sum over j_e by confining myself to a single set of j_e; but this is too restrictive. I am left with the standard choice in the literature: identify opposite ends of each edge; trace over indices at opposite ends. Each cell then has no SU(2) indices to contract with neighboring cells. The square brackets in Eq. (4) are now identical, and they factor out of the sum. Equivalently, the traces change ψ, Eq. (3) to

$$\psi \to \Pi_v \Sigma_{je} \psi_v(j_e), \tag{5}$$

where now there are no SU(2) indices on ψ_v. Each cell $\psi_v(j_e)$ is an SU(2) scalar.

I treat the planar case in the same way as the homogeneous case just discussed. The basic module is now the z line, plus two x and y edges at each vertex, with midpoints located at the vertex. The transverse edges are given the S_1 topology. For example, if I write out only the x edges and x indices at a given vertex,

$$\psi_v = \cdots h[A_x, e_{xi}/2]_{m',mi} I_{mi,mf} h[A_x, e_{xf}/2]_{mf,m'} \cdots$$

Here $e_{xi}/2$, for example, is the half-edge entering the vertex with intertwiner I. This structure may be rewritten in a manner which hides the traced indices.

$$\psi_v = \cdots I_{mi,mf} h[A_x, e_{xf}]_{mf,mi} \cdots.$$

Those indices usually play no dynamical role anyway, since the volume operator acts at vertices, i.e. its action affects only the m_i, m_f indices.

Loosely, one can describe the rewritten holonomy as a "loop" holonomy which leaves from and returns to the same vertex. I will sometimes use this way of speaking, but note the holonomy is integrated over a line, not over a loop in the xy plane. The holonomy depends only on the connection A_x.

The final modular structure allows communication in the longitudinal direction, but no communication in transverse directions. Gravitational waves can propagate only along z.

III. SPIN NETWORK BASIS STATES

In the previous section I proposed a spin network topology consisting of an infinite line dotted with vertices; at each vertex two edges extend in the positive x and y directions with ends identified. I now associate a holonomy with each segment of the line and each transverse edge. For example, for the transverse x edge,

$$h[A_x, e_{xf}] = \exp \int i A_x^B S_B dx. \tag{6}$$

The index B ranges over X and Y only. The expression on the right is not a trace. The $h[A_x, e_{xf}]$ matrix has two SU(2) indices. Both are contracted with corresponding indices on the intertwiner at the vertex; I am using the loop viewpoint described in the last three paragraphs of the previous section. The holonomies along z edges are similar to those

in Eq. (6), except that the su(2) valued connection is $A_z^Z S_Z$ and the integration is from one vertex to the next.

The next step is to determine the selection rule which the holonomies must obey, because they preserve the residual gauge symmetry U(1) at each vertex. U(1) transformations are generated by the surviving Gauss constraint,

$$G[\Lambda] = \int \Lambda(\partial_z \tilde{E}_Z^z + \epsilon_{ZAB} A_a^A \tilde{E}_B^a). \tag{7}$$

A typical vertex, located at coordinate z, will have two z-edge holonomies, one beginning and one ending at z. The constraint Eq. (7) generates the infinitesimal transformation $A_z^Z \to A_z^Z - \partial_z \Lambda$ which leads to the following finite transformation of the z holonomies.

$$\begin{aligned} H[A_z] &:= h[A_z, e_{zi}] h[A_z, e_{zf}] \\ &= \exp\left(i \int^z m_{zi} A_z^Z dz\right) \exp\left(i \int_z m_{zf} A_z^Z dz\right) \\ &\to \exp[-i\Lambda(z)(m_{zf} - m_{zi})] H[A_z]. \end{aligned} \tag{8}$$

Here I have replaced $S_Z \to m_{zf}$ or m_{zi}; there is no point to retaining the S_Z, since U(1) does not mix different eigenvalues of S_Z. m_{zf} and m_{zi} are integers or half-integers. f and i label the outgoing and incoming z edges, respectively, beginning and ending at the vertex.

The same vertex will have two transverse edge holonomies beginning at the vertex. For the transverse connection A_x, the infinitesimal transformation is

$$A_x^C S_C \to A_x^C S_C + [A_x^C S_C, i\Lambda S_Z],$$

which implies the following finite transformation, when the connection is promoted to a holonomy

$$\begin{aligned} h[A_x, e_{xf}]_{mf'mf} &\to (\exp[-i\Lambda S_Z] h[A_x, e_{xf}] \\ &\quad \times \exp[+i\Lambda S_Z])_{mf'mf} \\ &= \exp[-i\Lambda(m_{f'} - m_f)] h[A_x, e_{xf}]_{mf'mf}, \end{aligned} \tag{9}$$

and a similar equation for $h[A_y, e_{yf}]_{nf'nf}$.

Collecting together the x, y, and z contributions, I find that a U(1) transformation multiplies a vertex by the overall phase

$$\exp[-i(m_{zf} - m_{zi} + m_{f'} - m_f + n_{f'} - n_f)\Lambda(z)]. \tag{10}$$

U(1) invariance requires the relation

$$2F := m_f - m_{f'} + n_f - n_{f'} = m_{zf} - m_{zi} \tag{11}$$

Note that the intertwiner at each vertex no longer has to be a product of SU(2) $3J$ symbols. The intertwiner can be any product of Kronecker deltas assigning specific values to the m's and n's, provided the selection rule Eq. (11) is obeyed at the vertex.

It is straightforward to concoct a kinematic Hilbert space for the z holonomies, Eq. (8). The functions $\exp[im\theta]$ constitute a complete set of functions periodic on $[0, 4\pi]$, for m integer or half-integer.

However, I cannot use the SU(2) Haar measure to formulate a dot product for the transverse edge holonomies. To see this, I construct the simplest edge holonomy, that given by the spin $1/2$ representation of h. Denote the axis of rotation by $\hat{n}$. Since $\hat{n}$ must lie in the XY plane, it has the form

$$\hat{n} = (\cos\beta, \sin\beta, 0). \tag{12}$$

for some angle β. Then the spin $1/2$ holonomy becomes

$$\begin{aligned} h^{(1/2)} &= \exp[i\hat{n} \cdot \sigma\theta/2] \\ &= \cos(\theta/2) + i\hat{n} \cdot \sigma \sin(\theta/2) \\ &= \exp[-i\sigma_z(\beta - \pi/2)/2] \exp[i\sigma_y \theta/2] \\ &\quad \times \exp[+i\sigma_z(\beta - \pi/2)/2]. \end{aligned} \tag{13}$$

On the last two lines I have written the usual Euler angle decomposition for this rotation. It is clear that the rotation with axis confined to the XY plane depends on only two Euler angles, rather than the generic three. Let h^j denote the $2j + 1$ dimensional matrix representation of SU(2). Then integration over only two angles

$$\int (h^{j'*})_{a'b'} (h^j)_{ab} \sin\theta d\theta d\beta$$

guarantees only $a - b = a' - b'$ but not the stronger constraints $a' = a$, $b' = b$; the latter constraints must be satisfied in order for the θ integration to yield $j' = j$.

So far, I have been trying to follow a logical order: first construct an orthonormal basis; later, investigate the action of the volume operator on this basis. Logical order has not revealed a suitable dot product for the loop holonomies.

It turns out it is better to reverse logical order. Start by investigating the action of the volume operator (on the spin $1/2$ holonomy constructed above). This action suggests a basic structure which can be used to build higher order polynomials in the matrix elements of the spin $1/2$ holonomy, while maintaining simple behavior under the action of the volume operator. Finally, construct the kinematic dot product suggested by the structure of these polynomials.

I will need to define the action of the volume operator (more precisely the square of the volume operator). This operator is the product

$$(V_3)^2 = \epsilon_{ZBC} \tilde{E}_Z^z \tilde{E}_B^x \tilde{E}_C^y. \tag{14}$$

Each $\tilde{E}$ operator must be integrated over an area, in order to make the volume invariant under spatial diffeomorphisms. In particular $\tilde{E}_B^x$ must be integrated over an area in the yz plane. For clarity I have suppressed these areas in Eq. (14). However, I cannot ignore them completely. Their precise extent is needed in order to determine the location on the yz area where $\tilde{E}_B^x$ can find an A_x^B to grasp.

That location is easy to find in the context of the full theory. The A_x^B field has its support on an x edge, whereas

the $\tilde{E}^x_B$ field is confined to the yz plane. The yz plane intersects the x edge only at the vertex. Therefore $\tilde{E}^x_B$ acts where the x edge holonomy meets a vertex. In the reduced theory, this means $\tilde{E}^x_B$ acts at the two ends of the holonomy, Eq. (6), where the holonomy intersects the vertex.

I choose the yz area narrow enough in the y direction that the area includes only one vertex. When the full theory reduces to the symmetric one, neighbors in the y direction disappear anyway.

In the z direction, I could choose an area which overlaps two or more vertices. Then $\tilde{E}^x_B$ could act on one vertex at one z, while the other two triads in the volume operator act on a different vertex at a different z. I assume the yz area may extend in the z direction halfway to the next vertices, but not all the way, so that the volume operator can grasp lines exiting from only a single vertex. Equivalently, I assume all three triads grasp edges exiting from the same vertex. This assumption has the advantage of simplicity. It is also reasonable, since propagation (which demands operators that change more than one vertex) is associated with components of the Riemann tensor, not with the volume.

The $\tilde{E}$ operators in a canonical quantization act like functional derivatives with respect to the corresponding A operators. When acting on an edge holonomy, the $\tilde{E}$ bring down an su(2) generator at each end of the holonomy (never in the middle, because the volume operator acts only where the holonomy meets a vertex). For example the x holonomy is replaced by the following anticommutator.

$$\begin{aligned}\tilde{E}^x_A h[A_x, e_{xf}] &= \tilde{E}^x_A \exp[i \int A^B_x S_B dx] \\ &= (1/2)[S_A h[A_x, e_{xf}] + h[A_x, e_{xf}]S_A)]\gamma\kappa.\end{aligned} \tag{15}$$

Subscripts $A, B = X, Y$ only. I have omitted the triple delta functions, because they are always cancelled by the area and line integrals associated with $\tilde{E}^x_A$ and A^B_x. The overall factor $1/2$ is a relic of the integrals over the delta's, $1/2$ because the deltas occur at the endpoints of the edge integration. $\gamma\kappa$ is the product of Immirzi parameter times $8\pi G$; $\hbar = c = 1$.

The anticommutator means that the volume operator generates an infinitesimal O(4) transformation. To see this, specialize Eq. (15) to the spin $1/2$ case: $h[A_x, e_{xf}] \to h[A_x, e_{xf}]^{(1/2)}$ (see Eq. (13)); and $S_A \to \sigma_A/2$. Now consider the following two sets of SU(2) transformations,

$$\begin{aligned}R &:= \{U^{-1} h[A_x, e_{xf}]^{(1/2)} U\}; \\ B &:= \{U h[A_x, e_{xf}]^{(1/2)} U\}.\end{aligned} \tag{16}$$

The R's are just ordinary rotations; their infinitesimal form is a commutator. The infinitesimal form of the B's, however, is an anticommutator. The sets R and B are special cases of the set of transformations $\{U' h[A_x, e_{xf}]^{(1/2)} U, U' \neq U\}$. The set is $\mathrm{SU}(2) \bigotimes \mathrm{SU}(2)$.

Modulo fine points about covering groups, $\mathrm{SU}(2) \times \mathrm{SU}(2)$ is $\mathrm{O}(3) \bigotimes \mathrm{O}(3) = \mathrm{O}(4)$. To exhibit the O(4) structure, introduce the suggestive notation

$$ih[A_x, e_{xf}]^{(1/2)} := \begin{bmatrix} x_3 + ix_4 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 + ix_4 \end{bmatrix} \tag{17}$$

Since the matrices U' and U are unimodular, the transformation $U' h[A_x, e_{xf}]^{(1/2)} U$ preserves the determinant, which is

$$-(x_1^2 + x_2^2 + x_3^2 + x_4^2).$$

The transformations R turn out to be ordinary rotations; they leave x_4 invariant. A transformation B which has axis along direction i rotates x_i into x_4 $(1 \le i \le 3)$. (The B stands for boost. The B's are of course rotations, not Lorentz boosts, but they become Lorentz boosts when their angular parameter is continued to a pure imaginary value).

If possible, I would like to avoid using O(4) spherical harmonics as a kinematic basis. O(4) is double trouble: the harmonics are products of two rotation matrices, and the Clebsch-Gordan coefficients are products of two SU(2) Clebsch-Gordan coefficients. Fortunately, it is possible to use harmonics of an O(3) subgroup of O(4).

To identify the O(3) subgroup, I calculate the components x_i of $h[A_x, e_{xf}]^{(1/2)}$, Eq. (17). The following traces give components x_1 through x_4.

$$\begin{aligned}x_1 + ix_2 &= \mathrm{Tr}[ih[A_x, e_{xf}]^{(1/2)} \sigma_+]/\sqrt{2} \\ &= -\sin(\theta/2)\exp[i\beta];\end{aligned} \tag{18}$$

$$ix_4 = \mathrm{Tr}[ih[A_x, e_{xf}]^{(1/2)} 1]/2 = i\cos(\theta/2); \tag{19}$$

$$\begin{aligned}x_1 - ix_2 &= \mathrm{Tr}[h[A_x, e_{xf}]^{(1/2)} \sigma_-]/\sqrt{2} \\ &= -\sin(\theta/2)\exp[-i\beta];\end{aligned} \tag{20}$$

$$x_3 = \mathrm{Tr}[ih[A_x, e_{xf}]^{(1/2)} \sigma_z]/2 = 0. \tag{21}$$

I have introduced the usual raising and lowering operators

$$\sigma_\pm = (\sigma_x \pm i\sigma_y)/\sqrt{2}. \tag{22}$$

In computing the above traces, it is convenient to write Eq. (17) as $\vec{x} \cdot \vec{\sigma} + ix_4$.

From Eq. (21), the component x_3 vanishes. This is not an accident. There is no σ_3 in the expansion of Eq. (13) because the rotation is confined to the XY plane. Further, A_z is the only S_z valued connection left in the theory, after the gauge fixing, Eq. (2); and neither the Gauss constraint nor the volume operator have the power to change A_z into an $A_{x,y}$. Hence the relevant operators maintain $x_3 = 0$. This suggests we do not need the full O(4), but only the little

group which leaves invariant the vector

$$(x_1, x_2, x_3, x_4) = (0, 0, 1, 0). \tag{23}$$

The following theorems will make this idea more precise. First I recall some standard O(4) theory in order to identify the generators of boosts and rotations. For the SU(2) $\bigotimes$ SU(2) transformation $U'h[A_x e_{xf}]U$, let s' and s denote the generators of U' and U respectively, so that

$$[s'_i, s'_j] = i\epsilon_{ijk}s'_k; \qquad [s_i, s_j] = i\epsilon_{ijk}s_k; \qquad [s'_i, s_j] = 0. \tag{24}$$

Then the generators of boosts and rotations are given by

$$b_i = (s_i + s'_i)/2; \qquad r_i = (s_i - s'_i)/2. \tag{25}$$

Proof: write s and s' as

$$\begin{aligned} s &= (s + s')/2 + (s - s')/2; \\ s' &= (s + s')/2 - (s - s')/2. \end{aligned} \tag{26}$$

Demand that the general transformation $U'h[A_x, e_{xf}]U$ reduce to the special transformations defined in Eq. (16). For the boosts, the matrices U' and U must have the same generator, which means only $(s + s')/2$ can contribute; and similarly for the rotations, only $(s - s')/2$ can contribute.

Next I need to identify the little group. Let (r_1, r_2, r_3) be generators of su(2) and let $\vec{V}$ transform like a vector under the r_i, that is

$$[r_i, r_j] = i\epsilon_{ijk}r_k; \qquad [r_i, V_j] = i\epsilon_{ijk}V_k.$$

Then

$$\begin{aligned} (r_1, r_2, r_3) &\cong (b_1, b_2, r_3); \\ (V_1, V_2, V_3) &\cong (-x_2, +x_1, x_4). \end{aligned} \tag{27}$$

That is, the generators (b_1, b_2, r_3) of o(4) have the same Lie algebra as o(3); and the quantities $(-x_2, +x_1, x_4)$ rotate like a vector under the action of these generators. Note that x_3 does not rotate; (b_1, b_2, r_3) generate the little group.

On the second line of Eq. (27), the subscripts 1 and 2 have been interchanged, equivalent to a 90 ° rotation of the V's. The rotation is necessary because b_1 (for example) is not quite r_1. b_1 rotates x_1 into x_4, whereas r_1 rotates V_2 into V_3. A rotation is needed to exchange V_2 and V_1 before the isomorphism will work. For later convenience I rewrite this isomorphism in terms of eigenstates of the 90 ° rotation:

$$(V_\pm, V_3) \cong (\pm ix_\pm, x_4). \tag{28}$$

Proof of Eq. (27): it is straightforward to verify that the (b_1, b_2, r_3) obey the Lie algebra of o(3), by using the definitions Eq. (25) of the b's and r's, together with the commutation relations Eq. (24) for the s and s'. To verify that the x's rotate like a vector, note that the U in $U'h[A_x, e_{xf}]U$ is $\exp[(i(b_1, b_2, r_3) \cdot \vec{\alpha}]$, while the U' is the same, except for $\alpha_3 \rightarrow -\alpha_3$. Therefore for the b_i, the infinitesimal transformation is an anticommutator (with $\sigma_i/2$, since $h[A_x, e_{xf}]$ is spin 1/2); for r_3 the infinitesimal transformation is a commutator. b_1 (for example) generates the following infinitesimal transformation $\delta[ih[A_x, e_{xf}]^{(1/2)}$:

$$[ih[A_x, e_{xf}]^{(1/2)}, i\sigma_1/2]_+ = \begin{bmatrix} ix_1 & -x_4 \\ -x_4 & ix_1 \end{bmatrix}, \tag{29}$$

or

$$(\delta x_1, \delta x_2, \delta x_3, \delta x_4) = (-x_4, 0, 0, +x_1). \tag{30}$$

Compare this to $[ir_i, V_j]_- = -\epsilon_{ijk}V_k$, or

$$\delta V_2 = -V_3; \qquad \delta V_3 = +V_2.$$

After relabeling these equations as required by Eq. (27), they become Eq. (30). One proceeds in similar manner to prove the rest of the isomorphism.

I will refer to the vector on the right in Eq. (28) or Eq. (27) as the basic vector. By inserting Eq. (18) through Eq. (20) into Eq. (28), one can express the basic vector in terms of the angles β and θ.

$$\begin{aligned} (V_\pm, V_3) &\cong (\pm ix_\pm, x_4) \\ &= (\sin(\theta/2)\exp[\pm i(\beta - \pi/2)], \cos(\theta/2)) \end{aligned} \tag{31}$$

Evidently the basic vector is a unit vector. Also, we can construct Condon-Shortley spherical harmonics Y_{1m} from the components of the basic vector [18], but these spherical harmonics will have nonstandard arguments: $Y_{1m}(\theta/2, \beta - \pi/2)$ rather than the usual $Y_{1m}(\theta, \beta)$.

The next step is to build up more complex holonomies at the vertex: multiply together L matrix elements of $h[A_x, e_{xf}]^{1/2}$ to form all possible homogeneous polynomials of order L in the matrix elements. These polynomials form a rank L reducible representation of the little group O(3). Since we are multiplying together L identical vectors, we must break into irreducible representations by symmetrizing and taking traces; antisymmetrization gives zero. The irreducible representations are then just the $Y_{L'm}$ with $L' = L, L - 2, L - 4, \ldots$.

Therefore the set $\{Y_{Lm}(\theta/2, \beta - \pi/2)\}$ is complete, and I adopt it as a basis. Because these Y_{Lm} have argument $\theta/2$ rather than θ, the usual dot product for the spherical harmonics

$$\int_0^{2\pi} d\beta \int_0^{\pi} \sin\theta d\theta$$

will have to be modified slightly.

$$\langle L'm' \mid Lm\rangle = \int_0^{2\pi} d\beta \int_0^{2\pi} \sin(\theta/2)d(\theta/2)Y^*_{L'm'}Y_{Lm}. \tag{32}$$

I should perhaps emphasize that I have actually constructed two bases and two dot products, one for x holonomies with angles θ_x, β_x, and another for y holonomies with angles θ_y, β_y. These angles are unrelated except in the

one-polarization case, where classically, the two axes of rotation are at right angles, $\beta_x = \beta_y \pm \pi/2$. In that case, the kinematic Hilbert space and dot product must be reconsidered from scratch.

Note that the Y_{Lm} basis is far easier to use than (say) the set $\{\mathcal{D}^j(-\beta + \pi/2, \theta, \beta - \pi/2)_{0m}\}$, which would be a straightforward generalization of Eq. (13) to a rotation matrix of higher spin. The $\mathcal{D}^j_{0m}$ are orthonormal, but they are not representations of the O(3) little group, and the set $\{\mathcal{D}^j_{0m}\}$ is not closed under a grasp by the volume operator.

As for the Y_{Lm}, a grasp by $\tilde{E}^x_+$, for example, multiplies the basic vector by the spin 1 representation of S_+; this basic map induces a map of the higher harmonics into themselves, a map given by the spin L representations of S_+. Symbolically,

$$\tilde{E}^x_+ Y_{Lm} = \Sigma_{m'} Y_{Lm'} \langle L, m' | S_+ | L, m \rangle, \tag{33}$$

with [18]

$$\langle L, m \pm 1 | S_A = S_\pm | L, m \rangle = \sqrt{(L \mp m)(L \pm m + 1)/2}. \tag{34}$$

I have now completed the construction of the kinematic basis and dot product; and described the action of the volume operator triads on this basis.

IV. EIGENVALUE EQUATION FOR THE VOLUME OPERATOR

To summarize the results of the previous section: I now have a basis set of holonomies; at each vertex there are two z-edge holonomies

$$H[A^Z_z] = \exp\left[i \int^z m_{zi} A^Z_z dz \right] \exp\left[i \int_z m_{zf} A^Z_z dz \right] \tag{35}$$

plus two loop holonomies

$$Y_{L_x m_x}(\theta_x/2, \beta_x - \pi/2) Y_{L_y m_y}(\theta_y/2, \beta_y - \pi/2). \tag{36}$$

At Eq. (11) I worked out the conseqences of the surviving U(1) gauge invariance, but for standard holonomies $h_{m'm}$ rather than the new Y_{Lm} basis. It is easy to see from Eq. (12) that a gauge rotation changes only β while leaving θ alone; therefore to discover the gauge behavior of the Y's we need to study their β dependence. The holonomies $h_{m'm}$ have β dependence $\exp[i(m - m')\beta]$ (see, for example, Eq. (13)). Except for a normalization, $Y_{L_x m_x}$ is the rotation matrix $\mathcal{D}^{Lx}_{0mx}$, which has β dependence $\exp[i m_x \beta]$ Therefore at Eq. (11) the differences $m - m'$ should be replaced by m_x; and there is a similar replacement for the y indices. The U(1) selection rule, Eq. (11) simplifies to

$$2F = m_x + m_y = m_{zf} - m_{zi}. \tag{37}$$

The z holonomies of Eq. (8) are eigenfunctions of the $\tilde{E}^z_Z$ factor in the volume operator; and the remaining $[\tilde{E}]$ operators in the volume map the loop holonomies Eq. (36) into themselves as at Eq. (34) Therefore the volume operator will not change m, m_{zf}, L_x, or L_y. It will not change $F = (m_x + m_y)/2$, F for "fixed", because of the selection rule Eq. (37). The volume operator can change the quantity

$$D := (m_x - m_y)/2,$$

D for "difference". Therefore an eigenfunction of the volume operator will be the product of the two z-edge holonomies Eq. (8) times a sum

$$|\lambda; L_x L_y F\rangle = \Sigma_D Y_{L_x m_x} Y_{L_y m_y} c(m_x, m_y) \tag{38}$$

For simplicity I have suppressed the L and F dependence of the c's. The z dependence of the volume operator Eq. (14) acts on each of these basis elements as follows.

$$(V_3)^2 H[A_z] |\lambda; L_x L_y F\rangle = (\gamma\kappa/2)(m_{zi} - m_{zf}) H[A_z] \times (V_2) |\lambda; L_x L_y F\rangle. \tag{39}$$

$H[A_z]$ is defined at Eq. (8); the constants $\gamma\kappa/2$ are as at Eq. (15). $(V_2)^2$ is the determinant of the 2×2 subblock.

To complete the action of V_3, Eq. (39), I must determine how V_2 acts on the expansion Eq. (38). I shift from X, Y to the combinations $X \pm iY$ in V_2, to simplify later matrix elements.

$$\begin{aligned} \tilde{E}^a_\pm &:= (\tilde{E}^a_X \pm i \tilde{E}^a_Y)/\sqrt{2}; \\ (V_2)^2 &= \epsilon_{ZAB} \tilde{E}^x_A \tilde{E}^y_B = \epsilon_{Z-+} \tilde{E}^x_+ \tilde{E}^y_- + \epsilon_{Z+-} \tilde{E}^x_- \tilde{E}^y_+ \\ &= i(\tilde{E}^x_+ \tilde{E}^y_- - \tilde{E}^x_- \tilde{E}^y_+). \end{aligned} \tag{40}$$

The i and the minus sign come from $\epsilon_{Z\mp\pm} = \pm i$. Note how plus indices are always contracted with minus indices. For example,

$$\tilde{E}^x_X \sigma_X + \tilde{E}^x_Y \sigma_Y = \tilde{E}^x_+ \sigma_- + \tilde{E}^x_- \sigma_+ \tag{41}$$

$\tilde{E}^x_+$ is essentially the functional derivative with respect to A^-_x, therefore when it acts on the basic holonomies (Eq. (13) and (18) through Eq. (20)), the A^-_x functional derivative replaces each holonomy by its anticommutator with σ_+. (Again, plus indices always pair with minus). The basic holonomies transform as a spin 1 representation under this anticommutation. When acted on by $\tilde{E}^x_\pm$, therefore, the Y's transform as the $2L + 1$ dimensional representation of $S_\pm$.

$$(V_2)^2|\lambda; L_x L_y F\rangle = i(\gamma\kappa/2)^2 \Sigma c(m_x, m_y)[Y_{L_x m_x+1}\langle L_x m_x + 1|S_+|L_x m_x\rangle Y_{L_y m_y-1}\langle L_y m_y - 1|S_-|L_y m_y\rangle - (x \leftrightarrow y)]$$
$$= (\gamma\kappa/2)^2 \lambda|\lambda; L_x L_y F\rangle \tag{42}$$

On the last line I have assumed the state is an eigenstate of $(V_2)^2$. Multiplying both sides by $Y_{L_x m_x} Y_{L_y m_y}$, using the kinematic dot product introduced at Eq. (32), and the matrix elements given at Eq. (34), I find the following eigenvalue equation for λ.

$$2\lambda c(m_x, m_y) = ig(L_x, m_x, L_y, m_y)c(m_x - 1, m_y + 1) - ig(L_y, m_y, L_x, m_x)c(m_x + 1, m_y - 1);$$
$$g(L_x, m_x, L_y, m_y) = \sqrt{(L_x - m_x + 1)(L_x + m_x)} \times \sqrt{(L_y + m_y + 1)(L_y - m_y)}. \tag{43}$$

In terms of λ the eigenvalues λ_3 of the original operator $(V_3)^2$ are, from Eqs. (39) and (42),

$$\lambda_3 = (\gamma\kappa/2)^3(m_{zi} - m_{zf})\lambda \tag{44}$$

Because $F = (m_x + m_y)/2$ is held fixed, only the quantity

$$D := (m_x - m_y)/2 \tag{45}$$

is incremented in Eq. (43). The equation is an ordinary difference equation masquerading as a partial difference equation. To exhibit the ordinary difference character, I make the following replacements

$$m_x = F + D; \qquad m_y = F - D; \qquad c(m_x \pm 1, m_y \mp 1) = c(D \pm 1). \tag{46}$$

With these replacements, Eq. (43) becomes

$$2\lambda c(D) = i\sqrt{(L_x - F - D + 1)(L_x + F + D)} \times \sqrt{(L_y + F - D + 1)(L_y - F + D)}c(D - 1) - i\sqrt{(L_x + F + D + 1)(L_x - F - D)} \times \sqrt{(L_y + F - D)(L_y - F + D + 1)}c(D + 1) \tag{47}$$

I have been unable to find a compact, analytic solution to the above equation. The next three paragraphs, which describe my efforts to find such a solution, are perhaps of interest only to readers who are already familiar with the standard literature on ordinary difference equations [19]. Some readers may wish to skip these paragraphs on first reading.

To make contact with the standard literature, which treats primarily equations with rational coefficients, I need to get rid of the square roots. This is easily done by a change of dependent variable.

$$c(D) = \sqrt{(L_x + m_x)!/(L_x - m_x)!} \times \sqrt{(L_y + m_y)!/(L_y - m_y)!}d(D) = \sqrt{(L_x + F + D)!/(L_x - F - D)!} \times \sqrt{(L_y + F - D)!/(L_y - F + D)!}d(D) \tag{48}$$

If this is inserted into Eq. (43) and square roots cancelled, the resulting equation for d is

$$2\lambda d(D) = i(L_y + F - D + 1)(L_y - F + D)d(D - 1) - i(L_x + F + D + 1)(L_x - F - D)d(D + 1). \tag{49}$$

The coefficients are now rational, but they are quadratic functions of the independent variable D. Solutions when the coefficients are linear are already available in the literature, but for quadratic coefficients one must construct a solution.

One approach is to assume the solution is a series of factorials $d(D) = \Sigma_n a_n/(D - n)!$. (Factorials play the same role in the theory of difference equations that powers do in the theory of differential equations). The problem of determining the d's is turned into the problem of determining the coefficients a_n; for an equation with quadratic coefficients the new problem generally is as hard, or harder than the original problem.

One may also try Laplace's method: write $d(D)$ as the integral transform of a kernel, then show that solving the original difference equation is equivalent to solving a differential equation obeyed by the kernel. The differential equation corresponding to Eq. (49) is of a type unknown to me. It has four regular singular points. If one attempts to solve the differential equation with a series solution, the recurrence relation for the coefficients in the series is as hard to solve as the original difference equation, Eq. (49).

Although analytic solutions are hard to find, numerical solutions are easy to implement. For low values of the L's, one may write Eq. (47) as a matrix equation, $M \cdot \vec{c} = \lambda\vec{c}$, and solve for the eigenvalues of M. For larger values of the L's, I develop a WKBJ technique in the next section. For the rest of this section, I will discuss exact symmetries of Eq. (47).

Equation (47) possesses the following symmetry, which is easy to prove. Write Eq. (47) in a matrix notation, $M \cdot \vec{c} = \lambda\vec{c}$, and let $\{c(D; \lambda)\}$ be the components of a vector $\vec{c}$ which satisfies Eq. (47) with eigenvalue λ. Then the vector with components $\{(-1)^D c(D; \lambda)\}$ also satisfies the equation, with eigenvalue $-\lambda$. It follows that the eigenvalues

occur in pairs $(\lambda, -\lambda)$, except for possibly the zero eigenvalues.

This is a good place to mention that the operator needed by Thiemann [1] for his construction of the spin network formalism is actually the absolute value, $|(V_3)^2|$, rather than $(V_3)^2$ itself; hence the relevant eigenvalues are $|\lambda_3|$, and the theorem just proven states that eigenvalues of $|(V_3)^2|$ are at least doubly degenerate (except possibly the zero eigenvalues).

It is also possible to prove that there is at most one zero eigenvalue, for given values of L_x, L_y and F, and to construct the zero eigenfunction explicitly. When $\lambda = 0$, Eq. (47) collapses from a second order recurrence relation to a first order relation. (The relation connects every other value of D, relating $c(D+1)$ to $c(D-1)$; it is first order, with increment 2 rather than 1). Square $c(D+1)/c(D-1)$ to get rid of the square roots; one then has a first order equation for the squares of the c's, with rational coefficients. This is a standard form, with solution known up to a normalization constant N [19].

$$c(D) = N\sqrt{f(D-1)/f(D)}; \qquad f(D) = \left(\frac{L_y - F + D}{2}\right)!\left(\frac{L_y + F - D}{2}\right)!\left(\frac{L_x - F - D}{2}\right)!\left(\frac{L_x + F + D}{2}\right)!, \tag{50}$$

for $D = \max D, \max D - 2, \max D - 4, \ldots, \min D$; and $c(D) = 0$ otherwise.

Equation (50) must satisfy the boundary conditions that $c(D)$ vanishes outside the limits $\max D \geq D \geq \min D$. To find the limits on D, I note that (although the symmetry is U(1)) I am using a basis of SU(2) spherical harmonics. Therefore D is constrained by the SU(2) limits $-L_i \leq m_i \leq +L_i$. These limits may be turned into limits on $F \pm D$ by using the definitions $F = (m_x + m_y)/2$, $D = (m_x - m_y)/2$. It is then straightforward to derive the following limits on D.

$$\max(-L_x - F, -L_y + F) \leq D \leq \min(+L_x - F, +L_y + F) \tag{51}$$

From these limits, the denominator $f(D)$ in Eq. (50) becomes infinite for $D = \max D + 2, \max D + 4, \ldots$ and for $D = \min D - 2, \min D - 4, \ldots$, which enforces the boundary condition. Note the numerator is finite at those points.

Since the recurrence relation connects only every other value of D, one might suppose there is another zero eigenvalue, with $c(D)$ given by Eq. (50) for $D = \max D - 1, \max D - 3, \ldots, \min D + 1$; and $D = 0$ otherwise. However, this solution does not obey the boundary conditions; e.g. $c(\max D + 1)$ is nonzero.

Note also that the boundary condition requires the series $D = \max D, \max D - 2, \ldots$ to terminate at $\min D$, rather than $\min D + 1$. If the series terminates at $\min D + 1$, then from Eq. (50) $c(\min D - 1)$ will be nonzero, which violates the boundary condition. It follows that $\min D$ and $\max D$ must differ by an even integer, and the total number of allowed values of D, $\max D - \min D + 1$, must be an odd integer. This result is consistent with our earlier result that nonzero eigenvalues always occur in pairs $(\lambda, -\lambda)$. For given values of (L_x, L_y, F) and therefore given values of $(\max D, \min D)$, there will be one zero eigenvalue if the number of allowed values of D is odd; otherwise there will be no zero eigenvalues.

The Eq. (47) links c's all having the same value of the parameter F. Put another way, there is one set of equations Eq. (47) for each value of F. A symmetry relates the equations for F to the equations for $-F$, however, so that there is no need to solve both sets of equations. Let $\{c(D; F)\}$ be a solution to Eq. (47) with eigenvalue λ; then $\{c(-D; F)\}$ is a solution to the equations Eq. (47) with $F \rightarrow -F$ and eigenvalue $-\lambda$. (Temporarily I have restored the suppressed F dependence of the c's, for clarity). The proof is straightforward, because changing (D, F) to $(-D, -F)$ in the coefficients of Eq. (47) interchanges the two terms on the left, therefore changes the sign of the left hand side. (The c's on the left must be relabeled correctly;, for example, $c(D+1)$ becomes $c(-D-1)$, not $c(-D+1)$).

V. WKBJ

In this section I use results from a previous paper on WKBJ solutions to recurrence equations [20], and I obtain a WKBJ solution for the recurrence relation Eq. (47). Since derivations were given in the earlier paper, for the most part I shall avoid derivations and motivate results using physical arguments. However, the earlier paper applied the theory to the $6J$ symbols, where there is no need to quantize an eigenvalue, hence no need to derive quantization conditions, formulas analogous to $\int p dx = (n + 1/2)h$ in standard quantum mechanics. Later in this section I include some detail from the previous paper, enough to extend the theory slightly and derive a quantization condition.

It is perhaps not surprising that the recurrence relation Eq. (47) has a solution of WKBJ type,

$$c(D) = A \exp(iS). \tag{52}$$

A recurrence relation may be turned into a second order difference equation, since the c's may be replaced by central differences,

$$\begin{aligned}
&\delta^1 c := c(D+1) - c(D);\\
&\delta^2 c := c(D+1) - 2c(D) + c(D-1);\\
&c(D+1) = c(D) + \delta^1 c;\\
&c(D-1) = -c(D) - \delta^1 c + \delta^2 c.
\end{aligned} \tag{53}$$

Differences are very close to derivatives, therefore recurrence relations are very close to differential equations.

For a WKBJ solution to be possible, the recurrence relation does not have to contain the small parameter $\hbar$. For example, WKBJ has been applied to the classical equations describing waves moving through an inhomogeneous medium,

$$d^2\psi/dx^2 + k(x)^2\psi = 0. \tag{54}$$

A necessary condition for validity of WKBJ in this classical context is small derivatives, $d^n k/dx^n =$ order kL^{-n}, where L is a (large) length characterizing the rate of variation of the dielectric constant. In the present application, derivatives are replaced by differences; the quantities analogous to k are the square root functions in the recurrence relation Eq. (47); and the large parameter(s) L are the quantities $L_x \pm F$, $L_y \pm F$. There is no $\hbar$, but we are definitely in the limit of large quantum numbers.

I now state the requirements for a second order recurrence relation

$$\begin{aligned}
g_-(D,L)e(D-1) + g_0(D,L)e(D)\\
+ g_+(D,L)e(D+1) = 0
\end{aligned} \tag{55}$$

to have a WKBJ solution [20]. The parameters L must be large and the coefficients g must satisfy the following conditions.

$$(g_+ - g_-)/g_- \leq \text{order } 1/L; \tag{56}$$

$$g_0/g_\pm \leq \text{order unity}; \tag{57}$$

$$\delta^n g/g \leq \text{order } L^{-n}. \tag{58}$$

Consider first Eqs. (56) and (57). If these conditions are not satisfied initially, often a change of independent variable will lead to g's which satisfy these conditions. The coefficients in Eq. (47) do not satisfy Eq. (56), but I can remedy this by changing the dependent variable,

$$c(D) = (i)^D e(D). \tag{59}$$

This brings Eq. (47) to the standard form Eq. (55), with

$$\begin{aligned}
g_-(D) &= \sqrt{(L_x - F - D + 1)(L_x + F + D)}\\
&\quad\times\sqrt{(L_y + F - D + 1)(L_y - F + D)};\\
g_+(D) &= g_-(D+1);\\
g_0(D) &= -2\lambda
\end{aligned} \tag{60}$$

Next consider the third condition, Eq. (58). This is the analog of the "small derivatives of k" requirement and is the most important condition. It implies that the coefficients behave like polynomials under differencing, rather than like sinusoids, say.

To test whether Eq. (58) is satisfied, it is convenient to approximate differences by derivatives. If derivatives are falling off as $1/L$, then so are the differences.

$$\delta^1 g \approx (dg/dD)(\Delta D = 1) + \text{order}\,(d^2 g/dD^2). \tag{61}$$

When a square root in Eq. (60) is differentiated with respect to D, in effect the square root is divided by factors of order $(L \pm F) \pm D$. Therefore the difference is down by a factor of order $1/(L \pm F)$ (not order $1/D$; D does not have to be large).

The $g_\pm$ will not obey Eq. (58) (will not behave like polynomials under differencing) whenever D is near zeros of the square roots, where the $g_\pm$ are non analytic and differences $\simeq$ derivatives can be badly behaved.

Where are these zeros, and what is their physical interpretation? At Eq. (50) I derived limits on D using the SU(2) constraints $-L_i \leq m_i \leq +L_i$. I recall these limits here:

$$\begin{aligned}
\max(-L_x - F, -L_y + F) \leq D\\
\leq \min(+L_x - F, +L_y + F)
\end{aligned} \tag{62}$$

From a comparison of Eqs. (60) and (62) the zeros of the g's occur at points where D is approaching the limits imposed by SU(2). I will refer to the limits on D given in Eq. (62) as the SU(2) limits on D, and the corresponding zeros of the g's as SU(2) zeros. I cannot expect a WKBJ solution to work near the SU(2) limits, and I will have to check that these values of D occur inside classically forbidden regions, where the solution is negligible anyway. I therefore need to locate the turning points, values of D where the solution shifts from classically allowed (sinusoidal) to classically forbidden (exponential), in order to check that the SU(2) limits are outside the turning points.

To find the turning points, I apply formulas derived in [20]. In classically allowed regions, where the solution is sinusoidal, the amplitude and phase in Eq. (52) are given by

$$A^2 = \text{const.}[(g_+ + g_- - 2\lambda)(-g_+ + 3g_- + 2\lambda)]^{-1/2}; \tag{63}$$

$$\delta^1 S = \arccos[(g_- - g_+ + 2\lambda)/2g_-] \tag{64}$$

$$= 2\arcsin\sqrt{[(g_- + g_+ - 2\lambda)/(4g_-)]}. \tag{65}$$

When using these formulas for initial orientation, it is permissible to replace $g_+ + g_- = 2g_-$, because from Eq. (60) $g_+ - g_- = \delta^1 g_-$, which is assumed $\ll g_\pm$. With this replacement, the above formulas simplify to

$$A^2 \cong \text{const.}[(g_-^2 - \lambda^2)]^{-1/2}; \tag{66}$$

$$\delta^1 S \cong \arccos[\lambda/g_-] \tag{67}$$

$$\cong 2\arcsin\sqrt{((g_- - \lambda)/(2g_-)}. \tag{68}$$

For the qualitative discussions in this paper I shall use these approximate formulas. However, for numerical work one should use the more exact formulas. Neglecting the difference between g_- and g_+ is equivalent to setting $L(L+1) \cong L^2$, which introduces significant errors even when the L's are as large as 10.

In order to visualize the classically allowed region and the SU(2) zeros, it is helpful to make a rough plot of the function g_-^2 versus D. From Eq. (60), g_-^2 is a quartic, therefore the sketch looks like the usual Mexican hat potential. g_-^2 has four zeros at the four SU(2) limits, Eq. (62). Because of the min and max in Eq. (62), the SU(2) allowed region for D is the segment on the D axis lying between the two zeros closest to the center of the hat. Now draw a horizontal line at a height λ^2 above the D axis. From Eq. (66), the turning points are the two values of D (closest to the center of the hat) where this line cuts the hat. As $\lambda \to 0$, these turning points approach the SU(2) limits and the WKBJ approximation breaks down.

There is also a problem as the λ^2 line approaches the top of the hat. Recall the physical interpretation of the quantum number n in the usual quantum mechanical WKBJ formula $\int p dx = (n + 1/2)h$. n counts the number of half wavelengths which fit between the two turning points. In our case, as λ^2 grows large, the two turning points coalesce, and n becomes small; consequently WKBJ will be inaccurate. The "large" quantum number is not λ^2, but λ^2—order (L^4), since g_-^2 is order (L^4) near the top of the hat. It is perhaps best to think of the Mexican hat as an upside-down potential. The "well" of the potential is at the top of the hat.

From Eq. (68), I must take $\lambda \geq 0$, in order for the sine to have a zero at the same time as the cosine Eq. (67) becomes ± 1. I seem to have lost the negative λ eigenvalues. I can recover them if I use $(-i)^D$ instead of $(i)^D$ in Eq. (59). This yields Eq. (55), except $g_0 \to -g_0$. From Eq. (59) this replaces λ by $-\lambda$ everywhere. I have rederived a theorem from the previous section: the amplitudes for λ and $-\lambda$ differ by $(-1)^D$.

I must now derive a solution near turning points, since the eigenvalues λ are quantized by the usual requirement that the WKBJ solution connects to exponentially decaying solutions at left and right turning points. Reference [20] derives the necessary connection formulas. I change independent variable,

$$e(D) = Z(D)/\sqrt{g_-} \tag{69}$$

and insert this form into Eqs. (55) and (59). I find that Z obeys

$$\delta^2 Z + 2(g_- - \lambda)Z/g_- \cong 0. \tag{70}$$

In deriving Eq. (70) I have assumed that I am near $g_-^2 - \lambda^2 = 0$, but far from the SU(2) zeros of g_-, so that I can neglect higher differences of the $\sqrt{g_-}$ in Eq. (69).

From the Mexican hat plot, there are two turning points $D_<$ and $D_>$, to the right and left of the center of the hat. Consider first the smaller turning point $D = D_<$. I expand

$$\begin{aligned} 0 &\cong \delta^2 Z + (D - D_<)Z/k_<; \\ 1/k_< &= [2d(g_-)/dD](D_<)/\lambda. \end{aligned} \tag{71}$$

I have assumed that the zero of the second term in Eq. (70) is linear rather than quadratic. This means $1/k_<$ cannot vanish. $1/k_<$ is essentially the slope of the Mexican hat potential at the smaller turning point, so is indeed nonzero, and also positive. At the smaller turning point, and in the classically allowed region,

$$z_< := D - D_< \geq 0; \qquad 1/k_< \geq 0. \tag{72}$$

$k_<$ is order L, assuming that the λ and g_- in the definition of $k_<$, Eq. (71), are coefficients in the same recurrence relation, therefore are of the same order in L.

Equation (71) is a recurrence relation for the Bessel and Neumann functions. The Neumann solution can be discarded, because it has exponentially diverging behavior in the unphysical region, and I get

$$\begin{aligned} e(D) &= |c_<| J_{-z_< + 2k_<}(2k_<)/\sqrt{g_-} \\ &\to |c_<| \cos[-2\sqrt{z_<^3/k_<}/3 + \pi/4] \\ &\quad \times 1/\sqrt{[k_< z_< - (z_</2)^2]\pi g_-}. \end{aligned} \tag{73}$$

The $\sqrt{g_-}$ comes from Eq. (69). On the second line I have used the Debye asymptotic limit, rather than the usual Hankel limit, because both the index and the argument of the Bessel function are large [21]. The form given is valid for $2k_< \gg z_< > 3\sqrt[3]{2k_<}$, z large but not too large. I am free to choose the overall phase of the wavefunction, and I have chosen the constant $c_<$ to be positive.

Now consider the larger turning point, $D = D_>$, and again expand

$$\begin{aligned} 0 &\cong \delta^2 Z + (D - D_>)Z/k_>; \\ 1/k_> &= [2d(g_-)/dD](D_>)/\lambda. \end{aligned} \tag{74}$$

The slope on the right side of the Mexican hat, $1/k_>$, is now negative, and so is $z_> = D - D_>$ in the classically allowed region. Again, the solution is a Bessel function

$$\begin{aligned} e(D) &= c_> J_{+z_> - 2k_>}(-2k_>)/\sqrt{g_-} \\ &\to c_> \cos[-2\sqrt{(-z_>)^3/(-k_>)}/3 + \pi/4] \\ &\quad \times 1/\sqrt{[k_> z_> - (z_>/2)^2]\pi g_-} \end{aligned} \tag{75}$$

The constant $c_>$ may have either sign, since I have already

used up all the phase arbitrariness in the solution when I chose $c_<$ to be positive.

Equations (73) and (75) must be matched up with the WKBJ solution Eqs. (63) and (65). The difference equation Eq. (65) has the following exact solution for S.

$$S(D) = S(D_<) + \sum_{x=D_<}^{D} \delta^1 S(x)$$
$$= S(D_<) + \sum 2\arcsin[\sqrt{(g_- - \lambda)/2g_-}] \quad (76)$$

$$\cong S(D_<) + \sum \sqrt{[2(g_- - \lambda)/g_-]}$$
$$\cong S(D_<) + \sum_{z=0}^{z} \sqrt{z_</k_<}. \quad (77)$$

The third line is $\arcsin x \cong x$, valid near turning points. The last line uses the expansion of Eq. (71).

I would like to replace the sum in Eq. (77) by an integral, then integrate to obtain S (equivalently, replace

$$\delta^1 S \to dS/dD = dS/dz,$$

then integrate). It is not immediately clear I may do this, because the square root in Eq. (77) is not polynomial-like. However, consider the ratio

$$r(z) = \Sigma_{m=0}^{z}\sqrt{m}/(2\sqrt{z^3}/3), \quad (78)$$

The numerator is the sum on the last line of Eq. (77); the denominator is the approximation to the sum obtained by replacing the sum by an integral. For $D - D_< = z = 3$, $r = 1.2$; for $z = 10$, $r = 1.07$. In words: as I move away from the turning point, the integral is dominated by regions where the function is polynomial-like, and the integral becomes a better approximation to the sum. Therefore replacing sum by integral in Eq. (77) is valid near $z = D - D_<$ (but not too near). I integrate Eq. (77) to get S and then substitute into Eq. (52):

$$\mathrm{Re}\, A\exp(iS) = A\cos\left[\int_{D_<}^{D} \sqrt{z_</k_<} + S(D_<)\right]$$
$$= A\cos[2\sqrt{z_<^3/k_<}/3 + S(D_<)]. \quad (79)$$

Comparing this to Eq. (73) I get

$$S(D_<) = -\pi/4. \quad (80)$$

Now consider the WKBJ solution near the larger turning point $D = D_>$.

$$\mathrm{Re}A\exp(iS) = A\cos\left[\sum_{D_<}^{D}(2\arcsin) - \pi/4\right]$$
$$= A\cos\left[\sum_{D_<}^{D_>}(2\arcsin) + \sum_{D_>}^{D}(2\arcsin) - \pi/4\right]$$
$$= A\cos\left[\sum_{D_<}^{D_>}(2\arcsin) + \int_0^{z_>}\sqrt{(-z_>)/(-k_>)} - \pi/4\right]$$
$$= A\cos\left[\sum_{D_<}^{D_>}(2\arcsin) - 2\sqrt{(-z_>)^3/(-k_>)}/3 - \pi/4\right]. \quad (81)$$

"arcsin" denotes the arcsin function from Eq. (65). Comparing Eqs. (75) and (81), I find

$$\int_{D_<}^{D_>} 2\arcsin\sqrt{(g_- + g_+ - 2\lambda)/(4g_-)}dD = (n + 1/2)\pi. \quad (82)$$

I have used the more accurate expression, Eq. (65), for the arcsin. The formula contains $n\pi$, rather than $2n\pi$, because $c_>$ can have either sign. Except for the sum and the unfamiliar arcsin, the connection formula has the usual form. Again, the turning points are the two roots of $g_-^2 - \lambda^2 = 0$ which are closest to the center of the Mexican hat. (For more accuracy, use Eq. (63) and find the two roots of $g_- + g_+ - 2\lambda = 0$).

This is a good point to revisit the case $\lambda^2 \to 0$ briefly. If one thinks of the Mexican hat as an upside-down potential, then, because n is large at $\lambda^2 \to 0$, one should expect the form Eq. (52) to work well, not badly. However,in the WKBJ approach, one always solves the equations several times: once away from turning points and once at each turning point. From Eq. (71), $1/k$ blows up at a turning point which is also an SU(2) zero, since g_- has a square root zero there. The problem as $\lambda^2 \to 0$ is therefore at the turning points. They need a more careful discussion which I do not give here.

If convenient for numerical purposes, I can replace the sum by an integral, as I did when computing S near turning points. The arcsin is monotonic and positive, so that (by the same arguments as those used to bound series by integrals in the Cauchy integral test for convergence) one can bound the sum above and below by two integrals with lower limits differing by one unit. These bounds should be fairly tight, if the integral extends over the entire classically allowed region of order L. (I needed to do numerical work at Eq. (78), only because the integral was confined to a limited region and near turning points, where presumably the integral bounds are least restrictive).

I attempted to obtain an analytic form for λ, first replacing sum by integral in Eq. (82), then carrying our the integral. When an integrand contains an arcsin of complicated argument, usually it is best to integrate by parts, which replaces the arcsin by a square root (also complicated, but not as bad as the arcsin).

$$\int_{D_<}^{D_>} 2\arcsin\sqrt{(g_- - \lambda)/2g_-}\,dD = -\lambda \int_{D_<}^{D_>} dDDd(g_-)^2/dD \div [2(g_-)^2\sqrt{(g_-)^2 - (\lambda)^2}] \tag{83}$$

(The integrate by parts surface term vanishes). Equation (83) is reducible to elliptic integrals, but I believe this fact is of little use for extracting an analytic form for λ. The arguments of the elliptic integrals will be functions of the zeros of $g_-^2 - \lambda^2$, so that the λ dependence will be implicit.

As a numerical check on calculations of this section, I solved the exact difference equation on a computer for the case $F = 0$ and $L_x = L_y = L = 5$. I obtained the eigenvalues

$$\lambda = 26.6, 19.6, 13.3, 7.99, 3.59, \tag{84}$$

plus five more which were the negatives of these eigenvalues, plus one zero eigenvalue. I then used the FindRoot command on Mathematica to solve the WKBJ quantization condition Eq. (82). FindRoot must be given n, plus two initial guesses as to the value of λ; FindRoot then iterates, Newtonian style, until the program converges to a solution. I obtained

$$\lambda = 26.2, 19.3, 13.1, 7.86. \tag{85}$$

The four values correspond to $n = 0$ (largest) through $n = 3$ (smallest). I was unable to compute the $n = 4$ eigenvalue (the $\lambda^2 \to 0$ eigenvalue). The FindRoot program would not converge. The agreement between Eqs. (84) and (85) is surprisingly good for this small value of L.

[1] T. Thiemann, Class. Quant. Grav. **15**, 839 (1998).
[2] Karel Kuchar, Phys. Rev. D **4**, 955 (1971).
[3] A. Ashtekar and M. Pierri, J. Math. Phys. (N.Y.) **37**, 6250 (1996).
[4] Guillermo A. Mena Marugan and Manuel Montejo, Phys. Rev. D **58**, 104017 (1998).
[5] D. Korotkin and H. Samtleben, Phys. Rev. Lett. **80**, 14 (1998).
[6] Donald E. Neville, Phys. Rev. D **57**, 986 (1998), and references therein.
[7] V. Husain and L. Smolin, Nucl. Phys. **B327**, 205 (1989).
[8] M. Bojowald, Class. Quant. Grav. **21**, 4495 (2004).
[9] M. Bojowald, Class. Quant. Grav. **19**, 2717 (2002); **20**, 2595 (2003).
[10] T. Thiemann and H. A. Kastrup, Nucl. Phys. **B399**, 211 (1993).
[11] A. Ashtekar and M. Bojowald, Class. Quant. Grav. **23**, 391 (2006).
[12] L. Modesto, gr-qc/0411032; gr-qc/0509078.
[13] M. Bojowald, Phys. Rev. Lett. **95**, 061301 (2005); in *Trends in Quantum Gravity Research*, edited by David C. Moore (Nova Science, Hauppaughe, New York, 2005); www.Novapublishers.com.
[14] M. Bojowald and R. Swiderski, Class. Quant. Grav. **23**, 2129 (2006).
[15] M. Bojowald and H. A. Kastrup, Class. Quant. Grav. **17**, 3009 (2000).
[16] Martin Bojowald, Ph.D. thesis, Shaker Verlag, Aachen, 2000.
[17] A. Ashtekar, M. Bojowald, and J. Lewandowski, Adv. Theor. Math. Phys. **7**, 233 (2003).
[18] E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge University Press, New York, 1959).My definition of $S_{\pm}$ differs from Condon and Shortley's by a factor of $\sqrt{2}$.
[19] L. M. Milne-Thompson, *Calculus of Finite Differences* (MacMillan, London, 1933).
[20] Donald E. Neville, J. Math. Phys. (N.Y.) **12**, 2438 (1971).
[21] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea, New York, 1954), Debye's series are on p. 23.