Statistics of the excitations of the resonating-valence-bond state

N. Read and B. Chakraborty
Section of Applied Physics, Becton Center, Yale University, P.O. Box 2157, Yale Station, New Haven, Connecticut 06520
(Received 18 July 1988; revised manuscript received 15 February 1989)

By extending recent proposals for choosing the phases in resonating-valence-bond ground states to the case of excited states, we show for these wave functions that the hole excitations are charged, spinless fermions and the spin excitations are neutral, spin-$\frac{1}{2}$ bosons. We also show that for a system with periodic boundary conditions, all states are at least fourfold degenerate.

I. INTRODUCTION

There has recently been some controversy concerning the statistics of the excitations of the resonating-valence-bond (RVB) state of the two-dimensional square-lattice spin-$\frac{1}{2}$ Heisenberg-Hubbard model originally proposed by Anderson.[1] Kivelson *et al.*[2] argued in the case of the original "*s*-wave" RVB state that there are neutral spin-$\frac{1}{2}$ excitations ("spinons") which are fermions (as earlier suggested by Anderson) and spinless charged excitations ("holons") which are bosons. More recently, Laughlin[3] has suggested, motivated by wave functions originally proposed for the *triangular* lattice case, that the two types of excitation obey neither Fermi nor Bose statistics but obey $\frac{1}{2}$ statistics, i.e., exchange of identical particles produces a factor $\pm i$. (The general possibility of fractional statistics in two dimensions is discussed in Sec. III.) This conclusion is apparently supported by Marston[4] in quite a different approach. The purpose of this paper is to reexamine this question by a Berry phase calculation using the best RVB wave functions presently available. We find, in contrast to earlier proposals, that the spinons are bosons and the holons are fermions. We also give a new bosonic representation of the RVB states, which may be useful in generalizations to higher spin, and show that when the system has periodic boundary conditions, there is a four-fold degeneracy of the RVB ground and low-lying excited states.

II. TRIAL WAVE FUNCTIONS

A. Ground state

We begin by discussing some representations which have been used in the literature for states with short (i.e., nearest-neighbor) valence bonds. Sutherland[5] took up Anderson's suggestion of using, as a ground state for the square-lattice spin-$\frac{1}{2}$ Heisenberg model, a linear combination of all valence-bond configurations with equal weight, but with a specific choice for the relative *phases* of the configurations. More precisely, for a pair of spins at lattice sites i,j, a *valence bond* is a singlet state

$$[i,j]=(\uparrow_i\downarrow_j-\downarrow_i\uparrow_j)/\sqrt{2} \qquad (1)$$

in an obvious notation. A nearest-neighbor valence-bond configuration is then a state in which each site is paired, as in (1), with a neighboring site, thus forming a pattern of bonds which do not touch (Fig. 1). Since (1) is antisymmetric on interchanging i and j, a direction must be specified for each bond; we do this, slightly modifying Sutherland's prescription, by taking each bond to be directed from left to right, or down to up, where the arrow for (1) is directed from j to i. Sutherland's wave function is then a superposition of these bond states with equal weight. Since valence-bond states are generally not orthogonal, calculation of the norm or spin correlations of this wave function is nontrivial. The calculation of the overlap of two given valence-bond configurations $|a\rangle$, $|b\rangle$ is, however, quite straightforward and gives[5]

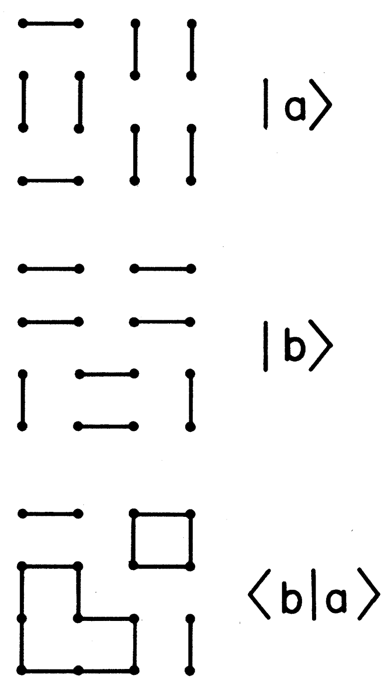

FIG. 1. Two valence-bond configurations $|a\rangle,|b\rangle$ and the "transition graph" for their overlap $\langle b|a\rangle$.

$$\langle b|a\rangle = \prod_u (2)^{1-L(u)} , \quad (2)$$

where u labels the (nonintersecting) loops in the "transition graph" (Fig. 1) formed by drawing all the bonds in both $|a\rangle$ and $|b\rangle$, and $2L(u)$ is the length of loop u (if a bond appears in the same position in both $|a\rangle$ and $|b\rangle$, it forms a dimer in the transition graph, i.e., a loop with zero area and $L=1$). Spin correlations can be calculated similarly. The result for the energy[5] is that it takes the form of a sum of negative terms; this was the motivation for the sign choice.

Liang *et al.*[6] used a slightly different way of directing the arrows on the bonds, always taking them to point from (say) a site on the B sublattice to one on the A sublattice. This gives the same result for overlaps (2) as Sutherland's and can be seen to give the same RVB state by comparing, in each approach, overlaps of each bond configuration with a reference state[7] whose sign is seen to be the same in each method. This method can be easily generalized[6] to include bonds of arbitrary length that connect the two sublattices. The motivation[6] for this choice was that each bond configuration individually has the Marshall sign; it is known[8] that the whole wave function must have this sign if it is to be the ground state of the nearest-neighbor Heisenberg model. The states with longer bonds can give very low energies,[6] even when there is no long-range Néel order, although the lowest energy (and the true ground state) of the nearest-neighbor Hamiltonian is believed to be obtained for states having Néel order.

The nearest-neighbor bond state can also be written using a fermionic representation[9] for the spins. A bosonic representation is also useful and will be presented later in this paper.

A spin of $\frac{1}{2}$ at each site can be built up by acting on an empty lattice with fermion creation operators $c_{i\sigma}^\dagger$ carrying spin $\sigma=\pm\frac{1}{2}$, with only one operator at each site. A singlet pair is then

$$(ij)=(c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger - c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger)/\sqrt{2} , \quad (3)$$

and is now symmetric under $i \leftrightarrow j$, so the bonds no longer carry arrows. Bond configurations now have overlaps[9]

$$\langle b|a\rangle = \prod_u (-2)^{1-L(u)} . \quad (4)$$

The sign may be restored if each bond configuration is multiplied by a configuration-dependent phase factor. From the mean-field solutions[10] of $s+id$ or "flux" symmetry, we guess that a possible choice is to build up the phase by multiplying $|a\rangle$ by a factor of g equal to 1 for each "horizontal" (i.e., parallel to the x axis) bond, and i for each "vertical" (y axis) bond.[11] The Sutherland wave function then resembles a Gutzwiller projected Bardeen-Cooper-Schrieffer (BCS) type pair wave function of $s+id$ symmetry

$$|\psi\rangle = \sum_{\text{configurations}} g_{i_1j_1}(i_1j_1)g_{i_2j_2}(i_2,j_2)\cdots|0\rangle , \quad (5)$$

where $g_{ij}=0$ if i,j are not nearest neighbors.

That the inclusion of the phases g_{ij}, as in (5), cancels the phases (4) from the overlaps of valence-bond configurations can be established as follows. For a dimer of length $2L=2$, the phase in (4) is positive and the factors $|g_{ij}|^2=1$ are also positive. The next simplest case is a square, $L=2$, where the product of g_{ij}'s and g_{ij}^*'s around the loop must provide a minus sign. It is easy to see that this is the case. For longer loops, one works by induction, showing that increasing the area (number of square plaquettes) bounded by the loop by 1 changes the sign by -1. Thus, for a loop u, the product of g's and g^*'s is

$$(-1)^{A(u)} ,$$

where $A(u)$ is the area enclosed by u. In general, $A(u)$ is not necessarily equal to $1-L(u) (\text{mod } 2)$ but one can show for non-self-intersecting loops traced on the square lattice that

$$A(u)=S(u)+L(u)-1 , \quad (6)$$

where $S(u)$ is the number of sites lying strictly inside u. For the ground state (5), a loop can only enclose other loops, and so $S(u)$ is even, which completes the proof that (5) has the correct phases. Note that (5) is real even though individual g's may be complex.

B. Excited states: Holons and vortices

We now wish to extend this construction to simple excited states. The basic excitations of our nearest-neighbor RVB state are "holons" in which a spin is missing from one site and all other spins form nearest-neighbor bonds as before and "spinons" in which one spin is left unpaired, all other spins forming bonds. Our basic assumption is that a good low-energy trial wave function for states with a single spinon or holon can be found by again applying Sutherland's prescription for choosing the phases.[12] This seems particularly reasonable for the holon, since the Marshall sign argument applies as well to a Heisenberg model with vacancies as to the original (nearest-neighbor) model. (We neglect for the time being the Hubbard hopping term for the hole.)

Consider a static holon first. In Sutherland's language (1), the result (2) still applies and we have our holon wave function. In the fermionic picture, however, we must reexamine, in the presence of the hole, the argument that (5) gives the correct phases with $g_{ij}=1$, i for horizontal, vertical bonds. We see that a loop enclosing the hole encloses an odd number of sites, and so the phases in (5) are not generally correct as they stand; the overlaps are wrong by a factor of -1 for each loop surrounding the hole. This problem may be corrected by a nontranslationally invariant choice of the g_{ij}'s which amounts to a "vortex" in the phases, centered on a plaquette next to the hole. A simple choice is to take such a plaquette and make a "cut" along links of the dual lattice from that plaquette to infinity (dashed line in Fig. 2). On each link of the original lattice that crosses the cut, g_{ij} for that position is multiplied by -1. Then, in the self-overlap of the hole-valence bond state, a loop surrounding the hole must cross the cut an odd number of times and so there is

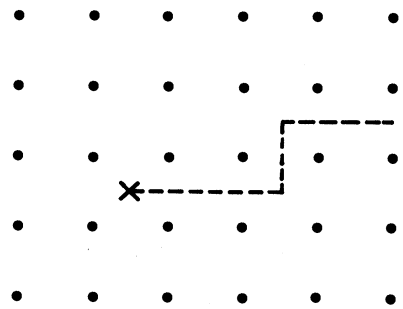

FIG. 2. X is the center of a vortex, in a plaquette of the lattice; the dashed line is the "cut": the g_{ij} factor for a link crossing the cut is multiplied by -1.

an extra factor of -1 for each loop, while loops not surrounding the hole cross the cut an even number of times and are multiplied only by $+1$. This then corrects the phases, making all terms in the selfoverlap positive, as in Sutherland's prescription for the ground state.

The reason for the term "vortex" becomes clearer if we make a gauge transformation. Since all sites except for the hole site are occupied by a fermion, a position-dependent gauge transformation of the fermion operators just multiplies the state vector by an overall phase; it does not change the relative phases. The phase change of the fermions can be absorbed into the g_{ij}'s giving

$$g_{ij} \to g_{ij} e^{i\theta_i + i\theta_j} , \quad (7)$$

where θ_i, θ_j are the gauge parameters for the sites i,j. In this way one can obtain different choices of g_{ij}'s giving the same relative phases and hence the same wave function up to an overall phase. For example, one can move the cut of a vortex to a different location, but not eliminate it. The position of the cut has no physical significance, and there is no tension or other energy associated with it.

A particular choice of gauge transformation can spread the factor of -1 for a loop surrounding the hole smoothly around the loop. Take $\theta_i = \theta(\mathbf{R}_i)$, where

$$\theta(\mathbf{R}) = \tfrac{1}{2} \tan^{-1} \left[\frac{R_y - R_y^{(h)}}{R_x - R_x^{(h)}} \right] , \quad (8)$$

where $\mathbf{R}_i$ is the position of site i, $\mathbf{R}^{(h)}$ is the position of the hole, and the discontinuity of θ (of magnitude π) is taken to lie on the cut. By (7), the phase of, e.g., a horizontal bond now varies "smoothly" (in small steps) as the position of the bond is moved around the hole, and changes by 2π on completing a circuit. This shows that we indeed have a vortex, with an effective flux of $\frac{1}{2}$ quantum since a bond consists of two charges.[13]

C. Spinons and vortices: Orthogonality questions

The above prescription obviously works equally well for a "spinon" in which the hole is replaced by an unpaired spin $\frac{1}{2}$, ensuring that the self-overlap of the sum of valence-bond configurations including the vortex is a sum of positive terms. However, in the spinon case, we must confront a nonorthogonality problem. We wish to have a set of orthogonal trial spinon states, one associated with each site in the lattice. Unfortunately, a valence-bond configuration with an unpaired spin at one position is not generally orthogonal to another configuration with the spin at another position; the overlap analogous to (4) now contains an open line segment of length $2L$ terminating at the free-spin positions, contributing a factor

$$(-2)^{-L} \quad (9)$$

to (4) (note that the free spins must be on the same sublattice, giving a line of length $2L$=even, to give a nonzero overlap). On summing over bond configurations, with or without the vortex in the phases, it appears that the overlap cannot fall less rapidly than exponentially with the separation of the unpaired spins.

From these nonorthogonal spinon states, we wish to construct orthogonal, localized spinon states (Wannier states). This may be done using standard methods, based on the Fourier transform of the overlap as a function of separation. Given that the overlap falls exponentially with separation of the unpaired spins, the Wannier states have an unpaired spin localized near the center of the Wannier wave packet. This localization is exponential if the Fourier transform of the overlap function has no zeros (as may be the case) or it is power law otherwise. Localization in the sense of square integrability of the unpaired spin (and accompanying vortex) near the nominal center of the Wannier wave packet is essential for the following calculation of statistics, which is otherwise not defined (see Haldane and Wu[14] for an example of this effect). We will assume this localization holds, in which case we need not work explicitly with the Wannier states.

Similarly, we may ask whether the holon state including the vortex is orthogonal to the holon state with the vortex omitted. If it is not, the Hamiltonian will connect these two states, and since, as we will see, the vortex contributes to the statistics, a mixture of the two states would not have well-defined statistics. Also, even in the absence of holes, we may consider vortices as excitations of the ground state, and ask whether vortices at different points are orthogonal, analogously to the spinon states.

To answer these questions, consider a state with a pair of vortices (and no holes), the vortex locations (on the dual lattice) being connected by a "cut" on which bonds have reversed sign. In the overlap with the Sutherland ground state, we see that any loop configuration containing a loop surrounding one vortex core gives zero contribution, while many configurations of loops crossing the cut give negative or zero terms. We see that the cut does not drop out, unlike the case of self-overlaps and the overlap is gauge dependent, though it only varies by an overall phase under a gauge transformation of the two-vortex state. Then choosing the cut to minimize the number of links it crosses, we compare the overlap with the "loop-gas" partition function[5,6] representing the self-overlap of either the ground state or the two-vortex state. Since some terms on the cut have been replaced by a negative number or zero, we expect that the overlap is reduced in magnitude, and the reduction factor relative to the self-overlap of either state falls exponentially with the length of cut. This occurs because the loop gas[5,6] has a finite correlation length, and its partition function (the self-overlap of the ground state) can be written roughly as a product of factors for each correlation area; each such area lying on the cut is reduced in magnitude by the zero terms and cancellations of positive and negative terms, leading to the exponentially decaying result. Thus, an isolated vortex in an infinitely large system is orthogonal to the ground state, and similarly when a hole is present. A similar calculation for the overlap of a state with a vortex at one point, with a state with a vortex at another point, shows that they are not generally orthogonal, but the overlap falls exponentially with separation; one may then construct Wannier wave-packet vortex states, as for the spinon. This suggests that the energy of the trial holon and spinon states could be lowered by not placing the vortex immediately next to the hole or unpaired spin; we will not consider this further since it leads to negative terms in the self-overlaps, contrary to the spirit of Sutherland's wave function.

III. STATISTICS

A. General discussion

We are now ready to calculate the statistics of the holon and spinon states constructed above. We regard statistics of indistinguishable particles as a group-theoretic property of the system, the symmetry operation being an interchange of particles *along a specified path;* for the spatial dimension $d>2$, all paths give the same result, and the group reduces to the permutation group, while for $d=2$, this does not occur, and we have the braid group.[15] The wave functions of the system will then transform under an interchange operation as some unitary irreducible representation of the group; these representations are usually one dimensional, i.e., an interchange can, at most, produce a phase factor (see Ref. 16 for a more general discussion). For the permutation group there are only two one-dimensional representations, namely the trivial representation, where an interchange produces a phase of $+1$ (Bose statistics) and the alternating representation, where a single interchange gives a phase of -1 (Fermi statistics). In contrast, for the braid group, i.e., in $d=2$ there is an irreducible representation for each value of the phase factor for the elementary interchange, giving "fractional statistics." [It is assumed that the particles are interchanged with all separations remaining large, and that the representation of the interchange group is well defined in (converges rapidly towards) this limit; this corresponds to the square integrable property in $d=2$, mentioned above. In $d=1$, of course, this assumption cannot hold, and our notion of statistics is never defined. Note that for the braid group the phase, in general, depends on the choice of path; we take an elementary interchange to be one for which the path does not enclose any other particles.]

Since any two particles are indistinguishable, the dynamics will generally be symmetric under interchanges, so the state of the system will be representations of the interchange group; however, it is possible that an effective Hamiltonian, say, for the dynamics, may not be explicitly invariant under interchanges of particle coordinates, just as a result of a choice of basis or gauge for the N-particle states as functions of the N-particle coordinates. In other words, multiplication of (say) the two-particle wave functions (a position-space basis state) by a phase that depends on the relative orientation of the two particles leads to a corresponding change in the Hamiltonian in this basis, which may render it no longer invariant under interchanges. An analogous situation for translational invariance would be to take a single particle in free continuous space and make an arbitrary gauge transformation of the wave function. The Hamiltonian than contains a vector potential (with zero curl) which breaks translational invariance. The symmetry of the underlying problem is recovered in this gauge, however, by noticing that translation of the particle coordinate, followed by a gauge transformation, returns the Hamiltonian to its original form. The eigenstates of this Hamiltonian will be momentum eigenstates of this modified translation operator, recovering the structure which can, of course, be obtained more straightforwardly by working in the gauge where the vector potential is zero.

A similar argument applies to the particle interchange problem, where Wilczek and others[15,17] have discussed the effect of particular singular gauge transformations on the realization of the interchange symmetry. In this case, the important part of the vector potential is that which depends on the relative orientation of the two particles, which can be picked out by examining the Hamiltonian when the particles are well separated. Again, the simplest picture is obtained by first gauging away this vector potential, so that the Hamiltonian is explicitly invariant under interchanges. In this case, the statistical phase can be read off from the change in the wave function under the interchange of particles along an anticlockwise path in $d=2$ not enclosing any other particles; in the Bose or Fermi case, the result is path independent. Even if one uses a different gauge, one still recovers the same result by ensuring that the transformation leaves the Hamiltonian invariant. Particle statistics in $d\geq 2$ are therefore a unique and well-defined property of the system (given the caveat mentioned earlier). Note that the only use of the adiabatic argument (a'la Berry) in the present discussion will be to ensure that an N-body effective Hamiltonian *exists* for our many-body "particle" excitations, working in the *low-energy* subspace of the trial states constructed earlier (thus, for example, the kinetic energy of the holon or spinon should not destroy the character of the static wave functions constructed earlier).

We note that the above arguments work equally well on the lattice as in free space, and that the "dangerous" parts of the effective Hamiltonian are those which move the particles around, such as the kinetic energy, since these will contain any vector potentials. We therefore take a definite basis of two-holon states, each basis state being specified by the position of the two holons, and examine the hopping matrix elements of the Hamiltonian between these basis states, looking for any orientation-dependent phase factors. If the long-distance part of these can be gauged out of the Hamiltonian, the statistics can be read off from the wave functions alone. Since motion of a particle around another is equivalent to two interchanges, we can calculate the phase factor for the former and take the square root; since we follow the change in phase under small moves, there is no ambiguity in this. A single particle moving in a loop may pick up a phase due to an effective magnetic field penetrating the loop; we are therefore interested only in the *change* in phase for such a path due to inclusion of another particle inside the path.

B. The calculation

As our "particles" we first take the holon states described earlier. We will now show how a certain modification of these states amounts to the desired singular gauge transformation that renders the effective Hamiltonian invariant under interchanges.

In the construction of the holon states, the hole was a site left unoccupied by a Fermi operator, and the vortex was created by the operation of introducing the line of minus signs for bonds lying on the cut (or any gauge-equivalent prescription). The latter operation may be regarded as a singular (multivalued) gauge transformation on the Fermi operators, which is double valued since the flux in the vortex is $\frac{1}{2}$ a flux quantum as pointed out earlier. Now let us introduce as a bookkeeping device a spinless creation operator $b_i^\dagger$ at the hole site, so that at every site a creation operator acts. Since destroying a spin is described by a Fermi (electron) operator, it is simplest to take $b_i^\dagger$ to obey Bose commutation relations, so that the destruction operator for physical electrons is $b_i^\dagger c_{i\sigma}$; $b_i^\dagger$ is a slave boson creation operator.[18] So far we have done nothing. Now we will make a further singular gauge transformation, this time acting on the locations of the holes only, by letting the previously described vortex-creation operation act on the $b_i^\dagger$,'s as well as the $c_i^\dagger$'s. Then, from the double valuedness of this transformation *for unpaired sites*, it follows that when two or more holes and their vortices are present, motion of one hole (with vortex) around another produces a phase change of 2π. This value is found since each "charge" moving around the other vortex gives a factor of -1, and there are two of these factors (when moving holes and vortices, we keep the relative location of a hole and the attached vortex center fixed). This corresponds to a phase change of π, i.e., a factor of -1, for interchanging the holons.

To complete the calculations for the holon, we must check that these changes in the holon location coordinates leave the effective Hamiltonian (in this gauge) invariant. The relevant part of the effective Hamiltonian is the Hubbard hopping term with amplitude t, which moves a hole from one site to a neighbor. We will assume that t is small compared with the Heisenberg coupling J, so that the earlier energetic arguments are likely to be valid. [In a realistic Hamiltonian, $t \gtrsim J$. This would substantially complicate the holon wave function; for the sake of a clear calculation, we would prefer to avoid this, postponing it for further study. (See also Sec. VI.)] Since it is small, it is sufficient to examine matrix elements of the Hamiltonian between holon states of different position. It is easy to see that a hop by one lattice spacing generates a state with one bond connecting next nearest neighbors, which is orthogonal to our basis states. Thus, we must go to second order to obtain a hopping amplitude of order t^2/J for hopping of a holon remaining on the same sublattice. This double hop of a holon produces either zero or two next-nearest neighbors, which is not orthogonal to our basis set. Motion of the vortex is accomplished simply by adjusting the phases of the bond states so that the vortex is in the correct position, any excess phase contributing to the effective hopping element itself. The key part of the calculation for two holons is the phase change due to moving a hole in the presence of a vortex connected to the other hole which does not move, or vice versa. Since the vortex is now introduced by a singular gauge transformation acting on hole and spin alike, the hopping simply interchanges the hole with a spin, and so the extra phase due to the distant vortex cancels; the phases of the effective hopping elements are completely insensitive to the presence of other distant holons. Thus the Hamiltonian is invariant under interchanges, and we conclude that the holons as constructed here are *spinless fermions*.

The calculation for the spinons is now very similar. The vortex-creation operation is redefined to act on the unpaired spin as well as the singlet pairs. The "hopping" now arises from the exchange part of the Hamiltonian, and the spinon remains on the same sublattice when it hops. The effective Hamiltonian is again invariant under interchanges. The phases from interchanges performed on the wave function are as for the holon, but in this case the interchange also involves changing the order of the unpaired Fermi operators, giving an extra minus. Thus, our spinons are *neutral spin-$\frac{1}{2}$ bosons*.

To summarize, we have argued that a vortex of $\frac{1}{2}$ flux quantum must be included in the fermionic representation of the spins to give a low energy for the holon and spinon. This flux couples to the unpaired spin and hole, and reverses the statistics of the holon and spinon relative to the result obtained[2] if all vortices are omitted. (It is well known[17] that a $\frac{1}{2}$-flux quantum bound to a charged particle turns bosons into fermions and vice versa.)

IV. BOSONIC REPRESENTATION

We now discuss the relation of our results with a simple bosonic representation of the spins. Here we interchange the statistics of the spin-$\frac{1}{2}$ and spinless- (hole) creation operators, so that a spin is represented by the presence of a Bose operator $a_{i\sigma}^\dagger$ acting on the empty lattice, while a "hole" is representing by the spinless Fermi operator $f_i^\dagger$. Again, the physical electron-destruction operator is $f_i^\dagger a_{i\sigma}$, which obeys the correct anticommutation relations. A singlet pair is now

$$[ij]=(a_{i\uparrow}^\dagger a_{j\downarrow}^\dagger - a_{i\downarrow}^\dagger a_{j\uparrow}^\dagger)/\sqrt{2} , \tag{10}$$

which is given a direction as for (1). In this representation, the Sutherland wave function for nearest-neighbor bonds can be obtained by directing all arrows from the B sublattice to the A sublattice, and this also works for longer bonds, as in Liang *et al.*; no factors of $1,i$ are needed. In this case, the self-overlaps of states with holes or unpaired spins included are all positive as in (2) *without* any extra vortices in the phases. It follows immediately that the holons are fermions and the spinons are bosons. If we do attach a vortex to each particle, even though this probably raises the energy, the statistics are again reversed. The important point is that the states including a vortex in the fermion representation correspond to states without a vortex in the boson representation, and vice versa; there is complete consistency between the pictures. As a bonus, the bosonic picture allows us to say that the results are unchanged by including longer bonds, though we cannot tell what happens if we go through the phase transition to the Néel state. Also, since longer bonds can be regarded as singlet pairs of spinons, the vortices should be useful in building longer bonds into the ground state in the fermionic picture, which had previously been thought to be difficult.

Before concluding this section, we note that the holon hopping amplitude can be calculated more easily in the bosonic picture since we do not have to contend with all the phase factors—all bonds have positive weight. One then finds that all contributions to the second-order effective hopping element have an extra minus sign, so that the minimum energy cannot occur at holon wave vector (0,0) or (π,π). This is in agreement with the results of several authors that the minimum occurs at $(\pm\pi/2,\pm\pi/2)$ for a hole moving in either the long-range Néel ordered state[19] or the mean-field ($s+id$ or flux) RVB state.[10]

V. DEGENERACY OF STATES FOR PERIODIC BOUNDARY CONDITIONS

So far in this paper, there has been no need to specify boundary conditions. We now wish to examine some global properties of the RVB states when periodic boundary conditions are present. For simplicity, we take the system to be a rectangle of sides L_x,L_y with periodic boundary conditions in each direction, and L_x,L_y even. The total number of holons and spinons together must then be even, and because a cut can end only at a vortex, so must the number of vortices. The argument in this section applies in either the fermionic or bosonic representation.

The degeneracy of the ground and excited states in the nearest-neighbor RVB language follows by first noting that we could insert into the ground state a cut running in, say, the y direction, which wraps once around the system, instead of ending at a vortex. Since cuts may be moved around by (nonsingular, topologically trivial) gauge transformations, there is essentially only one way of doing this, and since pairs of such cuts will cancel themselves when they run along the same path, the only gauge-inequivalent possibilities are either zero or one cut running in the y direction, and zero or one running in the x direction, i.e., four states in all. As seen earlier, adding a cut has no effect on the expectation of the energy locally, and here can only produce a minus sign for a loop in a transition graph that runs around the system in the direction orthogonal to the cut. Such long loops are very improbable in a large system, and so we expect the energies of our four ground states to be equal when $L_x,L_y\rightarrow\infty$, with exponentially small finite-size corrections.

It remains to be shown that these four states are orthogonal. Taking the overlap of two such (normalized) states, the effect of a cut which appears in one state but not the other is similar to its effect in the discussion of the vortex state (Sec. II C), i.e., the overlap decays exponentially with the length of the cut, here L_x or L_y. These arguments apply without modification to excited states with, say, a finite number of holon, spinon, and vortex excitations, so that these states are also four-fold degenerate.

This result is in agreement with a recent prediction of Haldane[20] that a disordered state of a half-integer spin system will have four-fold degeneracy in its ground state. Earlier, Affleck[21] had shown that any ground state of a half-integer spin system with periodic boundary conditions and an odd number of rows must have either gapless excitations or at least a two-fold degeneracy. Our result, though restricted to special states of a spin-$\frac{1}{2}$ system (we do anticipate generalizations to a higher half-integer spin and to longer bonds), does not require that the number of rows (or columns) be odd, and gives a definite result for the degeneracy (none of our excitations is expected to be gapless.[2]).

We note that the degeneracy we have found is a global, topological property of the system, which does not appear to reflect any symmetry breaking. There seems to be no local operator whose expectation value would specify which of the four states we are in, and so there can be no "domain-wall" excitations at which the state changes to another of the set of four. Therefore, we do not expect important physical consequences of this degeneracy, which is a consequence of the topology of the boundary conditions, rather analogous to the degeneracy of the states in the fractional quantum Hall effect with periodic boundary conditions.[22]

VI. FURTHER DISCUSSION: $d>2$ AND UNQUANTIZED VORTICES

Since submitting the original version of this paper, some further facts have come to light, which will be discussed here. Following our work, Kivelson[23] has reexamined the hard-core dimer model of Ref. 7 and applied our results in that model. He argues that when the hole hopping amplitude t is large enough, a vortex unbinds from (or binds to[13]) the hole, which again becomes a boson. This is clearly in agreement with the general principles expressed above. He also shows that the existence of $\frac{1}{2}$-flux vortices of finite energy leads to (physical) flux quantization in units of $hc/2e$. In addition, he argues that in three dimensions or more, the holon is always a boson, since there is no simple way to change statistics as there is in two dimensions (by binding a vortex or flux tube). While we agree with the latter point, we disagree with the former, since the methods of this paper can be used to show that holons are fermions and spinons are bosons in all dimensions ≥ 3. First we note that, in three dimensions or more, there seems to be no simple, local way to introduce phases into the fermionic picture of the valence bonds to recover Sutherland's sign prescription, in contrast to our discussion of two dimensions in Sec. II.[24] Therefore, we will adopt the bosonic picture described in Sec. IV for all dimensions ≥ 3. Since in this picture the holon involves adding a fermion, and no changes in the phase factors are needed to maintain the sign prescription, the holon is a fermion and similarly the spinon is a boson. The spin-statistics theorem[23] need not apply, since "spin" here is an internal quantum number similar to isospin, and not necessarily related to spatial rotations (this is also emphasized by Haldane and Levine,[25] who have arrived at the same conclusion as this present paper, by a different method).

Returning to two dimensions, we have also discovered some further results on ground-state degeneracy, vortices, and statistics. These stem from the observation by Rokhsar and Kivelson[7] that nearest-neighbor valence-bond configuration with periodic boundary conditions possess a pair of integer-valued winding numbers, one for each of the x and y directions. These may be defined by attaching an arrow to each valence bond, directed from the B to the A sublattice. Then, for a given bond configuration on a rectangular system, we count the number of arrows entering, and subtract the number of arrows leaving, the system at the left-hand boundary, and call this Ω_x. We do the same at the lower boundary and call it Ω_y. If we calculate the overlap of (or Hamiltonian matrix elements between) two states of different winding numbers (Ω_x, Ω_y) we find that the transition graph necessarily contains topologically nontrivial loops that run around the system, the winding numbers being equal to the differences in the Ω_x's and the Ω_y's between the two states. Then in a Sutherland-type superposition of configuration, such loops are very improbable in a large, two-dimensional system, and so the different winding number sectors are not mixed by the Hamiltonian; if nonorthogonality of valence bonds is omitted,[7] the conservation of winding numbers is exact for all sizes of system.

It now follows, by an argument similar to that in Sec. V, that there is an effectively *continuous* degeneracy of RVB ground states. To obtain this, we have only to weight the configuration with a pair of phases that are conjugate to the winding numbers, since the conservation of winding numbers and the mixing of winding numbers in the ground state should mean that different phases give orthogonal states and that the energy does not depend on the phases. An explicit construction again places a cut at the left-hand and lower boundaries of the system, and multiplies the g_{ij} for a bond on the cut by $e^{i\alpha_x}$ ($e^{i\alpha_y}$) for a bond "entering" the system, and the conjugate for a bond "leaving" the system at the left (lower) boundary. The phases α_x, α_y are arbitrary, and we can see that these factors are indeed conjugate to the winding numbers. Following the argument of Sec. V, one can show that in the self-overlap (or energy) of a state, the phases cancel for loops in the transition graph which are topologically trivial, because of the bipartite nature of the square lattice, and the fact that the bonds always connect opposite sublattices. Overlaps of states with different (α_x, α_y) give factors like $\exp[i(\alpha_z - \alpha_z')]$ $(z = x,y)$ and so by the arguments of Sec. II C these states are orthogonal in the limit of an infinite system. The special cases $\alpha = 0, \pi$ give the result of Sec. V.

A vortex can be made by having a cut similar to that in the ground-state degeneracy argument, but ending in a "core" at a plaquette center (Fig. 2). In this case, bonds crossing the cut are given a factor $e^{\pm i\alpha}$ depending on the sense in which their arrow traverses the cut (which is a well-defined notion irrespective of the path of the cut). Again, we can show that this new type of vortex state has only a finite energy, due to the core, and no energy per unit length of the cut, and that it is orthogonal to the ground state. If we attach the vortex to a holon or a spinon, then by repeating the argument of Sec. III, we now find *arbitrary* (fractional) statistics, given by the phase α. (Of course, $\alpha = \pi$ recovers the previous result.) Note, however, that these states give weight $\cos\alpha$ to loops surrounding the vortex core, and that Sutherland's prescription still leads to unique states as shown earlier.

These results are rather unpleasant, since a degeneracy of order $L_x L_y$, and the possible existence of bound states of arbitrary statistics seem unphysical. We believe that one possible answer is as follows: The nearest-neighbor bond states (and other states[6] with short-range bonds always connecting opposite sublattices) have a *larger symmetry,* namely their transition graphs connect sites which alternate between A and B sublattices, allowing the introduction of the cuts with factors $e^{\pm i\alpha}$ depending on the direction of a bond, as in this section. A more realistic state would presumably have some admixture of bonds connecting sites on the same sublattice (a small admixture need not affect the overall Marshall sign of the state). In this case, winding numbers are only conserved mod 2, and phases $e^{\pm i\alpha}$ on a cut do not generally cancel, so that the small admixture will eventually lead to a (small) energy per unit length of the cut, unless $\alpha = 0$ or π. The phases ± 1 on the cut, as used previously, thus correspond to the mod 2 winding numbers.[26] In this way we arrive back at the quantization of vortices and only fourfold ground-state degeneracy. Another possibility, which is realized for some nearest-neighbor Heisenberg Hamiltonians,[27] is that the large symmetry is unbroken, but the ground state has spin-Peierls order, again giving fourfold degeneracy, this time due to a broken lattice rotational symmetry.

VII. CONCLUSION

We have found that Sutherland's prescription for the phases of the valence-bond configurations in the short-range RVB state leads to holons which are fermions and spinons which are bosons. We have emphasized the importance of first establishing good trial wave functions, before statistics of excitations can be determined. Thus, we cannot rule out the possibility of the statistics being reversed again if improving the ground-state and excited-state wave functions leads to some essential change, such as the unbinding of vortices from the particles.[23] Nonetheless, we believe that our wave functions, with the emergence of vortices as excitations, set the scene for further investigations.

ACKNOWLEDGMENTS

We are grateful to P. A. Lee, R. Shankar, and S. Sachdev for useful discussions. The work of B. C. has been funded in part by the National Science Foundation Grant No. DMR 8658135.

---

[1]P. W. Anderson, Science **235**, 1196 (1987); G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).

[2]S. A. Kivelson, D. S. Rokhsar, and J. P. Sethna, Phys. Rev. B **35**, 8865 (1987).

[3]R. B. Laughlin, Phys. Rev. Lett. **60**, 2677 (1988); V. Kalmeyer and R. B. Laughlin, *ibid.* **59**, 2095 (1987).

[4]J. B. Marston, Phys. Rev. Lett. **61**, 1914 (1988).

[5]B. Sutherland, Phys. Rev. B **37**, 3786 (1988).

[6]S. Liang, B. Doucot, and P. W. Anderson, Phys. Rev. Lett. **61**, 365 (1988).

[7]D. S. Rokhsar and S. A. Kivelson, Phys. Rev. Lett. **61**, 2376 (1988).

[8]W. Marshall, Proc. R. Soc. London, Sect. A **232**, 48 (1955).

[9]P. Lederer and Y. Takahashi, Z. Phys. B **71**, 311 (1988).

[10]G. Kotliar, Phys. Rev. B **37**, 3664 (1988); I. Affleck and J. B. Marston, *ibid.* **37**, 3774 (1988).

[11]This has been noticed for the ground state independently by P. Lederer and Y. Takahashi, Z. Phys. B **71**, 415 (1988); in Ref. 6; and by ourselves.

[12]This is confirmed numerically for a hole in a 5×5 system by Lederer and Takahashi in Ref. 11; the resonance energy for a state equivalent to our hole with vortex state was found to be lower than for the state without the vortex.

[13]One may attempt to make a vortex of arbitrary flux by replacing the -1's on the cut by $e^{i\alpha}$, α arbitrary, but it seems that the state has a "defect energy" per unit length of the cut unless $\alpha=0$ or π, so that only quantized vortices can appear singly. Note that two vortices can annihilate.

[14]F. D. M. Haldane and Y. S. Wu, Phys. Rev. Lett. **55**, 2887 (1985).

[15]Y. S. Wu, Phys. Rev. Lett. **52**, 2013 (1984).

[16]G. A. Goldin, R. Menikoff, and D. H. Sharp, Phys. Rev. Lett. **54**, 603 (1985).

[17]F. Wilczek, Phys. Rev. Lett. **48**, 1144; **49**, 957 (1982).

[18]P. Coleman, Phys. Rev. B **29**, 3055 (1984).

[19]S. Trugman, Phys. Rev. B **37**, 1597 (1988); S. Sachdev, Phys. Rev. B **39**, 12 232 (1989); S. Schmitt-Rink, C. Varma, and A. Ruckenstein, Phys. Rev. Lett. **60**, 2793 (1988); C. L. Kane, P. A. Lee, and N. Read, Phys. Rev. B **39**, 6880 (1989).

[20]F. D. M. Haldane, Phys. Rev. Lett. **61**, 1029 (1988).

[21]I. Affleck, Phys. Rev. B **37**, 5186 (1988).

[22]F. D. M. Haldane, Phys. Rev. Lett. **55**, 2095 (1985).

[23]S. A. Kivelson, Phys. Rev. B **39**, 259 (1989).

[24]M. Ma [Phys. Rev. B **39**, 2952 (1989)] shows that in dimension $d>4$ it is not even possible to choose g_{ij}'s in the fermionic picture to make every elementary square transition graph on a (hyper-)cubic lattice give the correct sign. For longer loops, we find no solution already in $d=3$, because $S(u)$ in Eq. (6) can be odd.

[25]F. D. M. Haldane and H. Levine (unpublished).

[26]N. Bonesteel (private communication).

[27]N. Read and S. Sachdev, Nucl. Phys. **B316**, 609 (1989); Phys. Rev. Lett. **62**, 1694 (1989).