Large-n limit of the Hubbard-Heisenberg model

J. Brad Marston
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544

Ian Affleck
Canadian Institute for Advanced Research and Physics Department, University of British Columbia, Vancouver, British Columbia, Canada V6T 2A6

(Received 23 December 1988)

To gain insight into the behavior of the Hubbard model, we define a SU(n) invariant generalization of the Hubbard-Heisenberg model and, in the large-n limit, solve it in one dimension and in two dimensions on a square lattice. In one dimension the ground state is completely dimerized near half filling. We show that this behavior agrees with a renormalization-group solution of the one-dimensional SU(n) Hubbard model. In two spatial dimensions we find several different ground states depending on the size of the hopping term t, the doping δ, and the biquadratic spin interaction $\tilde{J}$. In particular, the undimerized "flux" or "$s+id$" phase is the ground state at half filling for sufficiently large t or $\tilde{J}$. We study the electronic and spin excitations of the various phases and comment on the relevance of the large-n problem to the high-T_c superconductors.

I. INTRODUCTION

The Hubbard model has been studied for over 25 years and a variety of approximate and exact techniques have been developed to understand its behavior. However, only the one-dimensional version is well understood. For that special case, an exact Bethe ansatz solution exists and powerful field-theory techniques such as bosonization and the renormalization group may be used to study it.

The behavior of the two-dimensional Hubbard model became a pressing issue after Anderson used it to describe the square copper oxide planes in the high-T_c superconductors.[1] Of particular importance are the following questions: What are the phases of the Hubbard model? Does the ground state exhibit long-range spin order? What are the low-energy excitations? Can superconductivity occur?

We attempt here to address some of these problems with a mean-field theory for the Hubbard model. (A brief outline of our results was given in Ref. 2. Some recent related work can be found in Refs. 3–5.) In spirit it is similar to the Green's function or Hartree-Fock approximations that have been applied before but differs in several respects. First, one of the two mean fields that we introduce is proportional to the amplitude of a valence bond lying on a given lattice link; the other mean field is simply the on-site electron number. Second, the approximation that we make to solve the model is more systematic than previous ones. Namely, we formally allow the number of "flavors" of electrons to be an arbitrary even integer n (the realistic model corresponds to $n=2$, the two spin directions of the electron). Higher values of n do not correspond to higher spin but rather to a higher symmetry group SU(n). [Recall that higher spins correspond to larger representations of the same SU(2) group.] Thus, our work generalizes certain SU(n) antiferromagnets introduced earlier.[6] We solve the model in the large-n limit. It is then possible to calculate systematic corrections in powers of $1/n$. Of course in the real world the expansion parameter is of order one, which is not very small. However, the Hubbard model has no other expansion parameters. (Weak-coupling perturbation theory is infrared divergent and a strong-coupling expansion just leads to another difficult problem—the Heisenberg model.) Perhaps a coupling constant of order 1 is better than none at all since at least we can investigate whether a qualitative change in the behavior of the model occurs as n decreases from large values down to two.

In the limit of infinite on-site repulsion, the Hubbard model reduces to the nearest-neighbor Heisenberg antiferromagnet. An important clue about the behavior of half-integer-spin antiferromagnets comes from a theorem proved by Lieb, Schultz, and Mattis (LSM) and generalized in Refs. 7 and 8. Apparently this theorem implies that there must either be broken translational symmetry or gapless excitations. The LSM theorem also applies to the SU(n) Heisenberg models that we discuss here. We are reassured to find that the theorem is obeyed by our large-n solutions.

We studied the two-dimensional Hubbard-Heisenberg model in the large-n approximation for different values of the hopping parameter, antiferromagnetic exchange constants, doping, and temperature. A variety of phases exist in this parameter space. In the Heisenberg limit (no electron hopping and a half-filled band) we find two zero-temperature phases. For small biquadratic coupling the ground state is highly degenerate and breaks translational symmetry: It is a bond-centered charge-density wave in which each site forms a dimer with one of its nearest neighbors. We call this phase the "Peierls" phase because the electrons are localized on individual dimers and the electronic spectrum is completely gapped. (Another localized state reported in Ref. 5 has the same ground-state energy as the Peierls phase for the case of zero biquadratic exchange. It will play an important role in our stability analysis.) At larger values of the biquadratic exchange, a "flux" phase is the ground state. It has full translational symmetry and the gap vanishes at discrete points in

momentum space. It is particularly interesting because the low-energy sector is described by a 2+1 dimensional relativistic field theory.

Passing to the more general Hubbard-Heisenberg model (but with the biquadratic spin interaction turned off), we find that the Peierls and flux phases exist at small doping depending on the size of the hopping parameter. For low values of the hopping term the Peierls state remains the ground state but there is a transition to the flux phase at sufficiently large hopping. The bond-centered charge-density wave vanishes at this transition; however, in both phases diamagnetic currents flow around the unit cells in alternating directions—"plaquette-centered antiferromagnetism." These currents break the symmetry of translation by one site down to translations across the diagonals of the unit cells. [The currents do not flow in the *pure* Heisenberg model. We will show that their disappearance is connected with the appearance of a local U(1) gauge symmetry in the Heisenberg limit.]

For larger doping, two new phases emerge. At sufficiently large doping, the ground state has unbroken symmetry and a gapless Fermi surface. We call this phase the "uniform" phase. We also find a peculiar phase at intermediate doping. Again there is a bond-centered charge-density wave, but unlike the Peierls phase, these waves are continuous ridges that extend through the entire system. All of these states conduct since no charge gaps occur at nonzero doping.

In Sec. II we define the models and give some qualitative discussion of their behavior. In Sec. III we solve the one-dimensional $n\rightarrow\infty$ Hubbard-Heisenberg model and compare the solution with exact results for the $n=2$ Hubbard model and with renormalization-group results for general SU(n). In Sec. IV we present the details of the two-dimensional large-n solution. We discuss the phase diagram and characterize each phase by its symmetries and electronic behavior. We show that these phases are stable by considering the effect of small, but spatially arbitrary, fluctuations of the order parameters. (The details of this calculation are presented in the Appendix.) Section V contains a calculation of the spin-spin correlation function for the uniform and flux phases. In Sec. VI we discuss the extrapolation from large n down to $n=2$ for the case of the pure Heisenberg model. Finally, we conclude by discussing the relevance of our results to the new high-T_c superconductors.

II. SU(n) HUBBARD, HEISENBERG, AND HUBBARD-HEISENBERG MODELS

The SU(2) Hubbard model Hamiltonian can be written as

$$H=t\sum_{\langle \mathbf{x},\mathbf{y}\rangle}(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{y}\sigma}+\text{H.c.})+(U/2)\sum_{\mathbf{x}}(n_{\mathbf{x}}-1)^2 .$$

Here $c_{\mathbf{x}\sigma}$ is the electron destruction operator and H.c. denotes Hermitian conjugate; the first sum is over nearest neighbors (on a square lattice, unless otherwise noted) and the repeated spin index σ is summed over the two spin states. Also, $n_{\mathbf{x}}=c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}$, the number of electrons on site $\mathbf{x}$. The usual Hubbard model interaction, $n_{\mathbf{x}\downarrow}n_{\mathbf{x}\uparrow}$, can be written in this manifestly SU(2) invariant way due to the Fermi statistics identity $(n_{\mathbf{x}\sigma})^2=n_{\mathbf{x}\sigma}$. It is well known that for $U/|t|\rightarrow\infty$, the Hubbard model at half filling ($\langle n_{\mathbf{x}}\rangle=1$) becomes a nearest-neighbor Heisenberg antiferromagnet:

$$H=J\sum_{\langle \mathbf{x},\mathbf{y}\rangle}\mathbf{S}_{\mathbf{x}}\cdot\mathbf{S}_{\mathbf{y}} .$$

Here the exchange coupling $J=4t^2/U$ and the electron-spin operators are

$$\mathbf{S}_{\mathbf{x}}=\tfrac{1}{2}c_{\mathbf{x}}^{\dagger\alpha}\boldsymbol{\sigma}_{\alpha}^{\beta}c_{\mathbf{x}\beta} , \tag{2.1}$$

where the $\boldsymbol{\sigma}$'s are the Pauli matrices. This exchange occurs because in the large-U limit, the low-energy states have one electron on each site but second-order perturbation theory in t produces a spin-spin interaction. Higher-order terms are negligible if $U\rightarrow\infty$ with t^2/U held fixed.

Alternatively, we may obtain the Heisenberg model as the limiting case of an itinerant electron model. We first set $t=0$ and write an explicit electron spin-spin interaction that is independent of U:

$$H=(J/4)\sum_{\langle \mathbf{x},\mathbf{y}\rangle}(c_{\mathbf{x}}^{\dagger\sigma}\boldsymbol{\sigma}_{\alpha}^{\beta}c_{\mathbf{x}\beta})\cdot(c_{\mathbf{y}}^{\dagger\gamma}\boldsymbol{\sigma}_{\gamma}^{\delta}c_{\mathbf{y}\delta})+(U/2)\sum_{\mathbf{x}}(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-1)^2 . \tag{2.2}$$

Using a Pauli-matrix identity, the Heisenberg interaction can be written

$$(c_{\mathbf{x}}^{\dagger\alpha}\boldsymbol{\sigma}_{\alpha}^{\beta}c_{\mathbf{x}\beta})\cdot(c_{\mathbf{y}}^{\dagger\gamma}\boldsymbol{\sigma}_{\gamma}^{\delta}c_{\mathbf{y}\delta})=-2(c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha})(c_{\mathbf{y}}^{\dagger\beta}c_{\mathbf{x}\beta})+2c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{x}\sigma}-(c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{x}\alpha})(c_{\mathbf{y}}^{\dagger\beta}c_{\mathbf{y}\beta}) .$$

We have grouped together SU(2) invariant bilinears. The second term on the right-hand side simply shifts the chemical potential. The third term is a nearest-neighbor Coulomb interaction that occurs in the extended Hubbard model. It reduces to a constant ($=-1$) for $U\rightarrow\infty$ so in this limit we consider only the first term. The Hamiltonian now reads

$$H=-(J/2)\sum_{\langle \mathbf{x},\mathbf{y}\rangle}(c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha})(c_{\mathbf{y}}^{\dagger\beta}c_{\mathbf{x}\beta})+(U/2)\sum_{\mathbf{x}}(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-1)^2 .$$

Including the hopping term, we have a hybrid Hubbard-Heisenberg Hamiltonian which reduces to the Hubbard model when $J=0$ and the Heisenberg model when $U\rightarrow\infty$ and $t\rightarrow 0$:

$$H=\sum_{\langle \mathbf{x},\mathbf{y}\rangle}[t(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{y}\sigma}+\text{H.c.})-(J/2)(c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha})(c_{\mathbf{y}}^{\dagger\beta}c_{\mathbf{x}\beta})]+(U/2)\sum_{\mathbf{x}}(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-1)^2 . \tag{2.3}$$

We wish to develop a general method for solving the Hubbard-Heisenberg model at arbitrary values of t, J, and U. Our approach is based on the functional integral method and the Hubbard-Stratonovich transformation. The partition function for this model can be written as a Feynman path integral

$$Z=\int[dc][dc^{\dagger}]\exp\left(-\int_0^{\beta}d\tau L\right) . \tag{2.4}$$

Here β is the inverse temperature and the (imaginary

time) Lagrangian is

$$L=\sum_{\mathbf{x}}[c_{\mathbf{x}}^{\dagger\sigma}(d/d\tau)c_{\mathbf{x}\sigma}]+H\,.$$

Antiperiodic boundary conditions at times $\tau=0$ and $\tau=\beta$ are imposed on the Grassmanian integration variables. This functional integral is, of course, nontrivial due to the quartic interaction terms in H. However, the Hubbard-Stratonovich transformation allows us to rewrite the Lagrangian quadratically in the fermions, at the expense of introducing two new bosonic fields, $\phi_{\mathbf{x}}$ (real) and $\chi_{\mathbf{xy}}$ (complex). We effect this transformation by adding two terms to the Lagrangian that do not change the equations of motion for the electrons:

$$L\rightarrow L+(2/U)\sum_{\mathbf{x}}[\tfrac{1}{2}\phi_{\mathbf{x}}+i(U/2)(c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{x}}-1)]^2 +(2/J)\sum_{\langle\mathbf{x},\mathbf{y}\rangle}|\chi_{\mathbf{xy}}+(J/2)c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha}|^2\,. \tag{2.5}$$

That these two terms leave the dynamics unchanged may be seen in two different ways. First, the equations of motion for the auxiliary fields ϕ and χ are trivial:

$$i\phi_{\mathbf{x}}=U(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-1)\,, \tag{2.6a}$$

$$\chi_{\mathbf{xy}}=-(J/2)c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{y}\sigma}\,. \tag{2.6b}$$

By using these equations to eliminate the auxiliary fields in the Lagrangian (2.5), we return to the original form expressed solely in terms of the electron fields. Alternatively, the functional integration over the ϕ and χ fields can be performed exactly. (The factor of i in the first Hubbard-Stratonovich term is necessary to make the functional integral over ϕ convergent.) These integrals simply convert the new terms to irrelevant constants.

However, upon adding the new pieces to the original Lagrangian we eliminate the quartic Fermi interactions. We now have a Lagrangian which is only quadratic in the electron operators, though new cubic interactions emerge:

$$L(c^{\dagger},c,\chi,\phi)=\sum_{\langle\mathbf{x},\mathbf{y}\rangle}\{(2/J)|\chi_{\mathbf{xy}}|^2+[(t+\chi_{\mathbf{yx}})c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{y}\sigma}+\text{H.c.}]\} +\sum_{\mathbf{x}}[(1/2U)\phi_{\mathbf{x}}^2-i\phi_{\mathbf{x}}+c_{\mathbf{x}}^{\dagger\sigma}(d/d\tau+i\phi_{\mathbf{x}})c_{\mathbf{x}\sigma}]\,.$$

The partition function for this new Lagrangian now equals the old partition function (2.4) up to a constant numerical factor:

$$Z=(\text{const})\int[dc][dc^{\dagger}][d\chi][d\phi] \times\exp\left(-\int_0^{\beta}L(c^{\dagger},c,\chi,\phi)d\tau\right)\,. \tag{2.7}$$

The advantage of this representation is that we may (at least formally) integrate out the electron fields and obtain an effective action written entirely in terms of the bosonic ϕ and χ variables:

$$\exp[-S_{\text{eff}}(\chi,\phi)]=\int[dc][dc^{\dagger}]\exp\left(-\int_0^{\beta}L(c^{\dagger},c,\chi,\phi)d\tau\right)\,.$$

Thus,

$$S_{\text{eff}}(\chi,\phi)=\int d\tau\left(\sum_{\langle\mathbf{x},\mathbf{y}\rangle}(2/J)|\chi_{\mathbf{xy}}|^2+\sum_{\mathbf{x}}[(1/2U)\phi_{\mathbf{x}}^2-i\phi_{\mathbf{x}}]\right) -2\,\text{Tr}\ln(d/d\tau+i\phi_{\mathbf{x}}+\{[t+\chi_{\mathbf{e}}(\mathbf{x})]\Delta_{\mathbf{e}}+\text{H.c.}\})\,.$$

Here $\Delta_{\mathbf{e}}$ is a lattice derivative operator and $\chi_{\mathbf{e}}$ represents the χ variables on the links running in the $\mathbf{e}$ direction. The repeated $\mathbf{e}$ index implies a sum over the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ directions. The factor of 2 in front of the Tr ln term results from the identical contributions which occur for each spin of the electron fields. Arbitrary electron Green's functions can now be expressed in terms of the bosonic auxiliary variables; for example, the time-ordered electron two-point Green's function is given by

$$\langle\{d/d\tau+i\phi_{\mathbf{x}}+[t+\chi_{\mathbf{e}}(\mathbf{x})]\Delta_{\mathbf{e}}\}^{-1}\rangle\,,$$

where $\langle\,\rangle$ denotes a Boltzmann average using the effective action. Note that the equations of motion tell us that $i\langle\phi_{\mathbf{x}}\rangle$ is proportional to the charge density at site $\mathbf{x}$ (relative to one electron per site) and $\langle\chi_{\mathbf{xy}}\rangle$ corresponds to the hopping amplitude for electrons between the sites $\mathbf{x}$ and $\mathbf{y}$. We have therefore rewritten the Hubbard-Heisenberg model using the number density and hopping amplitude as order parameters.

So far our transformations have been exact. To continue, we now must make some approximation. We will perform the integration over ϕ and χ in a saddle-point approximation which we may regard as an *approximate* technique for solving the *exact* model. Alternatively, we can see it as the exact solution of an approximate model. Let us define a generalized SU(n) invariant Hubbard-Heisenberg model by making the spin index σ in Eq. (2.3) run from 1 to n (where n is even). Note that this generalization is *not* the same as letting the electrons have a higher spin. The coupling between the various spin components would not, in that case, have the simple form of Eq. (2.3). We should instead think of the existence of n different "flavors" of electrons, all of which are equivalent. Note that the Hamiltonian conserves the total number of electrons of each flavor. It is convenient to adjust the chemical potential to be zero for a half-filled band by rewriting the Hubbard interaction as $(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-n/2)^2$. We also rescale both U and J by a factor of $2/n$ so that the large-n limit is smooth—the spacing of energy levels should be $O(1)$ as $n\rightarrow\infty$. The ground-state energy and free energy are, however, $O(n)$. Thus H becomes

$$H=\sum_{\langle\mathbf{x},\mathbf{y}\rangle}[t(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{y}\sigma}+\text{H.c.})-(J/n)(c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha})(c_{\mathbf{y}}^{\dagger\beta}c_{\mathbf{x}\beta})] +(U/n)\sum_{\mathbf{x}}(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-n/2)^2\,. \tag{2.8}$$

Apart from the desire to study the more general hybrid model, we have a technical reason for keeping the explicit Heisenberg interaction in our Hamiltonian even though that term would be generated by the exchange of electrons. We will be studying the limit $n\rightarrow\infty$ with U held fixed. We may then take U/t large if we wish. However, these limits do not commute. An expansion in powers of t/U contains higher and higher powers of n, so we cannot justify truncating the series of $O(t^2/U)$. Conversely, in the large-n limit, a large U eliminates fractional fluctua-

tions in the number of electrons on a site, but the cost in energy to add or remove a single electron is only $O(U/n)$. Taking $n\to\infty$ first and then taking $U/t\to\infty$ therefore does not give the large-n limit of the Heisenberg model. Thus, in order to have a model which interpolates smoothly between the Hubbard and Heisenberg models at $n\to\infty$, we are forced to consider the hybrid Hubbard-Heisenberg model. For n finite and even and a half-filled band (now meaning an average of $n/2$ electrons per site), the $U\to\infty$ limit forces exactly $n/2$ different flavor electrons per site.

Upon making the Hubbard-Stratonovich transformation, the SU(n) (imaginary time) Lagrangian associated with the Hamiltonian (2.8) becomes

$$L(c^\dagger,c,\phi,\chi)=\sum_{\langle x,y\rangle}\{(n/J)\,|\chi_{xy}|^2+[(t+\chi_{yx})c_x^{\dagger\sigma}c_{y\sigma}+\text{H.c.}]\}+\sum_x[(n/4U)\phi_x^2-i(n/2)\phi_x+c_x^{\dagger\sigma}(d/d\tau+i\phi_x)c_{x\sigma}]\,. \tag{2.9}$$

The equations of motion for the auxiliary fields now read

$$i\phi_x=(2U/n)(c_x^{\dagger\sigma}c_{x\sigma}-n/2)\,,$$

$$\chi_{xy}=-(J/n)c_x^{\dagger\sigma}c_{y\sigma}\,,$$

and the effective action is

$$S_{\text{eff}}(\chi,\phi)=n\int d\tau\left[\sum_{\langle x,y\rangle}(1/J)\,|\chi_{xy}|^2+\sum_x[(1/4U)\phi_x^2-(i/2)\phi_x]\right]-n\,\text{Tr}\ln(d/d\tau+i\phi_x+\{[t+\chi_e(x)]\Delta_e+\text{H.c.}\})\,. \tag{2.10}$$

The factor of n outside the Tr ln term now reflects the integration over the n flavors of electrons. Thus, the entire effective action is proportional to n and taking the $n\to\infty$ limit corresponds to the classical $\hbar\to 0$ limit.

We will also consider a further generalization of the pure Heisenberg problem by including the biquadratic spin-spin interaction $(\tilde{J}/4)(\mathbf{S}_x\cdot\mathbf{S}_y)^2$. In fact, we will show that it is necessary to include such a term to stabilize the flux state in the Heisenberg limit $t\to 0$. For the SU(2) case, this term simply renormalizes the bilinear interaction (J) since $(\mathbf{S}_x\cdot\mathbf{S}_y)^2=-\frac{1}{2}\mathbf{S}_x\cdot\mathbf{S}_y+\text{const}$ but for $n>2$ it corresponds to new processes that interchange four fermions simultaneously. Biquadratic interactions are generated (along with next-nearest-neighbor and plaquette terms) at order t^4/U^3 in the large-U expansion of the SU(n) Hubbard model. However, we will treat $\tilde{J}$ as a free parameter since these multiple-exchange processes are quite natural from the large-n viewpoint. We may then study the usual SU(2) Heisenberg problem using the $1/n$ expansion as long as we restrict our choice of $\tilde{J}$ such that the renormalized value of J remains positive (antiferromagnetic). Thus, we have a one-parameter family of SU(n) spin models (characterized by different values of $\tilde{J}/J$) that all have the same SU(2) limit.

We start by writing down the Hamiltonian with a ferromagnetic biquadratic coupling that is properly scaled so that the $n\to\infty$ limit remains well defined:

$$H=\sum_{\langle x,y\rangle}[-(J/n)\,|c_x^{\dagger\alpha}c_{y\alpha}|^2+(\tilde{J}/n^3)\,|c_x^{\dagger\alpha}c_{y\alpha}|^4]+(U/n)\sum_x(c_x^{\dagger\sigma}c_{x\sigma}-n/2)^2\,. \tag{2.11}$$

In fact, this Hamiltonian is antiferromagnetic for arbitrary n as long as $\tilde{J}<2J$ in the sense that the ground state of the two site problem is a SU(n) singlet for $\tilde{J}<2J$. To prove this result, note that the different SU(n) representations that occur in the two-site problem can be labeled by Young's tableau with one column of $(n-p)$ boxes and another with $0\le p\le n/2$ boxes. (The singlet state corresponds to $p=0$.) Using the SU(n) quadratic Casimir,[9] we find

$$|c_1^{\dagger\alpha}c_{2\alpha}|^2=n^2/4+n/2-p(n+1-p)\equiv n^2x_p\,.$$

Thus, the energy of the two-site problem (with $n_x=n/2$) becomes

$$E_p=n(\tilde{J}x_p^2-Jx_p)\,.$$

Clearly, the minimum of the energy occurs at the value of x_p as close as possible to $J/2\tilde{J}$. Since x_p is a decreasing function of p, the ground state has $p=0$ for sufficiently small $\tilde{J}$. Upon increasing $\tilde{J}$, the value of p for the ground state increase by unit increments. The $p=0$ and $p=1$ states are degenerate at the boundary of the antiferromagnetic region. Using $x_0=\frac{1}{4}+1/2n$ and $x_1=\frac{1}{4}-1/2n$ we see that this boundary occurs at $\tilde{J}=2J$ regardless of n.

The Hubbard-Stratonovich factorization of the spin-spin interaction terms must now be done in two stages. First we introduce a new real field Φ_{xy} to break up the biquadratic term:

$$H\to H+n\tilde{J}\sum_{\langle x,y\rangle}[(1/\tilde{J})\Phi_{xy}+(i/n^2)\,|c_x^{\dagger\alpha}c_{y_\alpha}|^2]^2\,.$$

Note the factor of i in this expression. Again, it is needed to make the functional integral over Φ converge. We shall see, however, that Φ acquires a pure imaginary value at the saddle point and therefore the saddle-point free-energy remains real.

After adding the first Hubbard-Stratonovich term, the Hamiltonian still contains a four-Fermi interaction. This term can now be factorized by introducing the complex χ fields:

$$H\to H+nJ\sum_{\langle x,y\rangle}(\{(1/J)\chi_{xy}+(1/n)[1-(2i/J)\Phi_{xy}]^{1/2}c_x^{\dagger\alpha}c_{y\alpha}\}\{(1/J)\chi_{xy}^\dagger+(1/n)[1-(2i/J)\Phi_{xy}]^{1/2}c_y^{\dagger\beta}c_{x\beta}\})\,.$$

It is important to recognize that the added term is not Hermitian. In fact, the factors of i will complicate the discussion of the stability of the saddle points. (We will return to this point in Sec. IV and the Appendix.) The new Hamiltonian is now quadratic in the fermions and can be studied by our saddle-point method. Upon including the ϕ fields which act as Lagrange multipliers to constraint the system to $n/2$ fermions per site in the $U\rightarrow\infty$ limit, the Lagrangian reads

$$L=\sum_{\mathbf{x}}[c_{\mathbf{x}}^{\dagger\sigma}(d/d\tau+i\phi_{\mathbf{x}})c_{\mathbf{x}\sigma}-i(n/2)\phi_{\mathbf{x}}]+\sum_{\langle\mathbf{x},\mathbf{y}\rangle}\{(n/\tilde{J})\Phi_{\mathbf{xy}}^2+(n/J)|\chi_{\mathbf{xy}}|^2+[1-(2i/J)\Phi_{xy}]^{1/2}(\chi_{\mathbf{xy}}c_{\mathbf{y}}^{\dagger\alpha}c_{\mathbf{x}\alpha}+\text{H.c.})\}\,. \quad (2.12)$$

Before proceeding further, we note that there exists another way of enlarging the SU(2) symmetry to the larger SU(n) group. Instead of demanding that each site have $n/2$ electrons, we could instead require one electron per site on the even sublattice and $(n-1)$ per site on the odd sublattice. This condition can be enforced by the $U\rightarrow\infty$ limit of a different Hubbard interaction that reads

$$(U/n)\sum_{\mathbf{x}\text{ even}}(c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-1)^2+(U/n)\sum_{\mathbf{x}\text{ odd}}[c_{\mathbf{x}}^{\dagger\sigma}c_{\mathbf{x}\sigma}-(n-1)]^2\,.$$

(Note that this term explicitly breaks the symmetry of translation by one site.) In this limit, we obtain the SU(n) generalization of an antiferromagnet discussed previously in the literature.[6]

To emphasize the differences between the two large-n limits, consider the SU(n) analog of the usual SU(2) operator $\mathbf{S}^2=\frac{3}{4}$. For the case of m fermions on a given site ($m\leq n$) the Casimir takes the value $(n+1)\times(n-m)m/2n$. Thus, $\mathbf{S}^2=n(n+1)/8$ when $m=n/2$ but $\mathbf{S}^2=(n^2-1)/2n$ for either $m=1$ or $n-1$. In the former case the "spin" is of order n, while in the later case it is of order $n^{1/2}$.

The ground state in the large-n limit of the $m=1$, $n-1$ model can be constructed out of nearest-neighbor valence bonds. A valence bond for general n is the state

$$|\mathbf{x}_\mathbf{y}\rangle\equiv c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha}|0\rangle\,,$$

where $|0\rangle$ has no electrons on site $\mathbf{x}$ and all n on site $\mathbf{y}$. In fact, the most general way of making SU(n) singlets out of the particle and hole indices—fundamental and antifundamental representations of SU(n)—is to contract arbitrary pairs of particle and hole indices. In general, the valence bonds can have arbitrary length; however, they must run between the even and odd sublattices. In the following discussion, we assume that the ground state is an SU(n) singlet. (For $n=2$, the ground state of the finite system has been proven rigorously to be a singlet.[10]) Long-range order occurs when the amplitudes for arbitrarily long valence bonds drop off as some inverse power of their length.[11]

To study these valence bonds, it is convenient to make a particle-hole transformation on the odd sublattice:

$$c_{\mathbf{y}}^{\alpha}\equiv c_{\mathbf{y}}^{\dagger\alpha}\quad\text{for }\mathbf{y}\text{ on the odd sublattice}\,.$$

The bond operator χ that we introduced earlier can then be regarded a valence-bond creation or annihilation operator:

$$\chi_{\mathbf{xy}}=(J/n)c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha}^{\dagger}\quad\text{for }\mathbf{x}\text{ even and }\mathbf{y}\text{ odd}\,.$$

In this basis, the Heisenberg Hamiltonian takes the following form:

$$H=-(J/n)\sum_{\langle\mathbf{x},\mathbf{y}\rangle}(c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{y}\alpha}^{\dagger})(c_{\mathbf{x}\beta}c_{\mathbf{y}}^{\beta})\quad\text{for }\mathbf{x}\text{ even and }\mathbf{y}\text{ odd}\,.$$

Consider the action of H on a valence bond:

$$H|\mathbf{x}_\mathbf{y}\rangle=-J|\mathbf{x}_\mathbf{y}\rangle\,.$$

Obviously, the bond has energy $-J$. How does H act on a lattice link not containing a valence bond? The sites at the two ends of the link must form bonds with two other sites somewhere in the lattice [see Fig. 1(a)]. We can write this state, mentioning only the four sites, as $|\mathbf{x}_\mathbf{z}\rangle|\mathbf{w}_\mathbf{y}\rangle$. Acting with H on the $\mathbf{x}_\mathbf{y}$ link yields an $O(1/n)$ amplitude for a different configuration of valence bonds: $-(J/n)|\mathbf{x}_\mathbf{y}\rangle|\mathbf{w}_\mathbf{z}\rangle$. In other words, H creates a valence bond on the link on which it acts, breaks the other two valence bonds which had terminated there, and creates a new non-nearest-neighbor bond between the two left over sites.

We can now see the simplification that ensues when $n\rightarrow\infty$. Those processes that create non-nearest-neighbor bonds have amplitudes which are of order $1/n$. The ground state in the $n\rightarrow\infty$ limit can therefore be constructed just from nearest-neighbor valence bonds. In the one-dimensional case there are only two such states (see Fig. 2). In two (or higher) dimensions, there are an

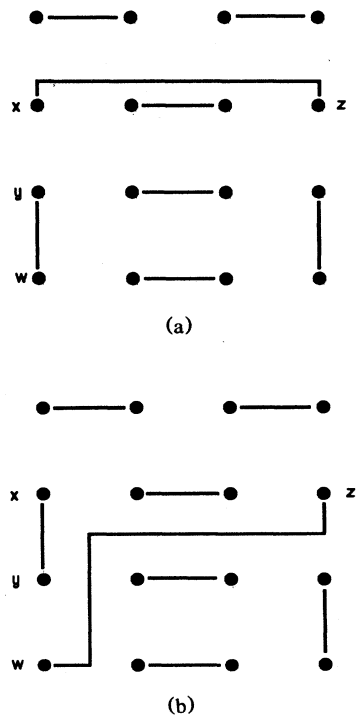


FIG. 1. The effect of H on a valence-bond configuration (a) before H acts on the link $\langle\mathbf{x},\mathbf{y}\rangle$, and (b) after.

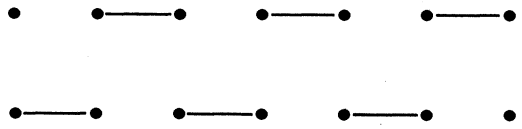


FIG. 2. The two dimerized ground states of the antiferromagnetic chain.

infinite number of such states. $O(1/n)$ perturbations mix these states and pick out a unique ground state (or a set of ground states if some symmetry is broken).

In $O(1/n)$ degenerate perturbation theory we need only consider matrix elements of H between nearest-neighbor valence-bond states. The only nonzero matrix elements are between two states that differ by the rotation of a pair of parallel valence bonds through an angle of 90° around a plaquette. (See Fig. 3.) This matrix element has the value $-J/n$. It is easy to check that the ground state is not simply an equal-weight sum of all nearest-neighbor valence-bond states with some simple assignment of phases. The weighting factor for a given diagram must depend on the number of plaquettes containing parallel valence bonds (as well as on how that number changes upon making the 90° rotation, etc.). Thus the ground state has quite a bit of information built into it about how the bonds "resonate." Recently, Read and Sachdev examined the $1/n$ corrections to this state using this Hamiltonian approach.[4] One could try to solve this problem by the path-integral methods developed here for the problem of $n/2$ electrons per each site. However, we have encountered some problems with this approach and will not discuss it further.

The two SU(n) models have an important difference: the first generalization (with $n/2$ electrons per site) preserves lattice-translation symmetry, whereas the second (with one electron on each even site) is obviously not invariant under translation by one site. Since the Lieb-Schultz-Mattis theorem only applies to spin Hamiltonians which preserve translation symmetry, it is useless in the second case. The theorem implies that either the spectrum contains gapless excitations or the ground state exhibits broken symmetry of translation by one site. In fact, the second generalization could exhibit a ground state with no broken symmetry *and* a gap, but this phenomena should not happen in the first model.

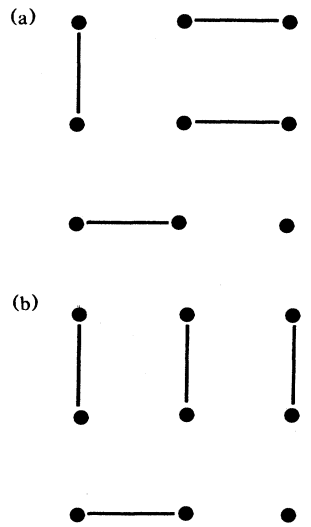


FIG. 3. The two valence-bond configurations that mix in degenerate perturbation theory.

Returning to the model with $n/2$ electrons per site, it is clear that singlet states have $n/2$ valence bonds emanating from each point. (Again we define a valence bond as a contraction between a particle index on an even site and a hole index on an odd site. Odd sites contain $n/2$ holes in this case.) In general, a large number of valence bonds, of $O(n)$, exists on each link. The probability for H to break a valence bond is now $O(1)$ because there are $O(n)$ different valence bonds to choose from. However, this model is also solvable in the large-n limit because the fractional fluctuation in the number of valence bonds on each link is suppressed at large n. Indeed, the number of valence bonds on any link becomes a classical quantity as $n \to \infty$.

More formally, the overall factor of n in the effective action (2.10) means that the saddle-point approximation becomes exact as $n \to \infty$. The Feynman diagram loop expansion is thus equivalent to a power series in $1/n$. Higher-order terms in the $1/n$ expansion can be calculated; we will only consider the lowest order term here (see, however, the Appendix). This approach has several advantages over other mean-field theories of the Hubbard model:

(1) It is possible to systematically calculate corrections to our approximation.

(2) The approximation is justified insofar as we may regard $1/n$ as a small parameter. How small $1/n$ must be to make the approximation good is of course not obvious.

(3) By thinking about the differences between large-n and $n=2$ we may be able to understand the strengths and limitations of our approximation.

With this justification, we now proceed to discuss the details of solving the model in the large-n limit by making the saddle-point approximation. We must find the appropriate minimum action saddle point of the effective action of Eq. (2.10). This problem is still not completely trivial. However, we expect that the minimizing configuration will have a high degree of symmetry reflecting the unbroken symmetries of the equilibrium state. In particular, the minimizing configuration should be time independent since we wish it to describe an equilibrium state. If we assume ϕ and χ to be time independent, then we can write S_{eff} in a more familiar way:

$$S_{\text{eff}}(\phi,\chi) = \beta F(\phi,\chi;\mu)$$

where $F(\phi,\chi;\mu)$ is the free energy of the electrons in the static classical fields ϕ and χ and chemical potential μ:

$$F(\phi,\chi;\mu) \equiv n\left[\sum_{\langle x,y\rangle}(1/J)\,|\chi_{xy}|^2+\sum_x[(1/4U)\phi_x^2-(i/2)\phi_x]\right.$$
$$\left.-(1/\beta)\sum_m \ln\{1+\exp[-\beta(E_m-\mu)]\}\right]. \tag{2.13}$$

Here E_m are the energies of a single electron interacting with the static, classical fields ϕ and χ. In other words, E_m are the eigenvalues of the one-particle Hamiltonian

$$H_1 = \sum_{\langle \mathbf{x},\mathbf{y}\rangle} [(t+\chi_{\mathbf{yx}})c_{\mathbf{x}}^{\dagger}c_{\mathbf{y}} + \text{H.c.}] + i\sum_{\mathbf{x}} \phi_{\mathbf{x}} c_{\mathbf{x}}^{\dagger} c_{\mathbf{x}} .$$

Thus the large-n limit amounts to a type of Hartree-Fock approximation in which the quartic interaction terms are factorized. We now seek the lowest saddle point of $F[\phi,\chi]$ where the ϕ and χ fields in principle can have arbitrary spatial dependence. Anticipating a high degree of symmetry in the equilibrium state, we will however limit our search to a subspace of ϕ's and χ's with a certain specified symmetry.

Turning our attention to the bosonic auxiliary fields, we see that nonuniform $|\chi_{\mathbf{xy}}|$ corresponds to a bond-centered charge-density wave since $\chi_{\mathbf{xy}}$ is related to the electron-hopping amplitude. In the Heisenberg limit, a site-centered charge-density wave is impossible; however, a bond-centered charge-density wave can still occur. It also corresponds to a modulation of the strength of the nearest-neighbor spin-spin correlation $\langle \mathbf{S}_{\mathbf{x}} \cdot \mathbf{S}_{\mathbf{y}} \rangle$ (a spin-Peierls phase). This relationship between the correlator and the χ fields can be seen by evaluating $\langle |\chi_{\mathbf{xy}}|^2 \rangle$ in the functional approach

$$\langle |\chi_{\mathbf{xy}}|^2 \rangle = (1/Z)\int [dc^{\dagger}][dc][d\phi][d\chi]\,|\chi_{\mathbf{xy}}|^2 \times \exp - \int d\tau L(c^{\dagger},c,\phi,\chi) .$$

Here the Lagrangian L is given by Eq. (2.9). The Gaussian integration over the χ fields can be performed by shifting the integration variable to $\chi'_{\mathbf{xy}} = \chi_{\mathbf{xy}} + (J/n)c_{\mathbf{x}}^{\dagger a} c_{\mathbf{y}a}$. With this shift, it follows that

$$\langle |c_{\mathbf{x}}^{\dagger a} c_{\mathbf{y}a}|^2 \rangle = (n/J)^2 (\langle |\chi_{\mathbf{xy}}|^2 \rangle - \langle |\chi'_{\mathbf{xy}}|^2 \rangle) ,$$

since $\langle \chi'_{\mathbf{xy}} \rangle = 0$. Now the second term on the right-hand side is simply a constant, so

$$(n/J)^2 \langle |\chi_{\mathbf{xy}}|^2 \rangle = \langle |c_{\mathbf{x}}^{\dagger a} c_{\mathbf{y}a}|^2 \rangle + \text{const} \propto \langle \mathbf{S}_{\mathbf{x}} \cdot \mathbf{S}_{\mathbf{y}} \rangle + \text{const} .$$

Thus, modulation in the magnitude of the χ fields does indeed correspond to a bond-centered spin-density wave.

This bond-centered wave can also be discussed in the language of valence bonds. A uniform value of $|\chi_{\mathbf{xy}}|^2$ corresponds to equal numbers $(n/4)$ of valence bonds on each lattice link. In two spatial dimensions, such a uniform state is not possible for $n=2$, but only for n equal to a multiple of 8. At large n, this constraint effectively disappears since the number of bonds can be equal up to relative corrections to $O(1/n)$. Thus, though a uniform phase can occur at large n, the valence bonds must "resonate" at $n=2$. Apparently, the large-n limit represents a sort of static limit of the resonating valence bonds. On the other hand, a simple spin-Peierls configuration with $|\chi_{\mathbf{xy}}|$ nonzero on only one link emanating from each point *can* be realized for $n=2$.

The phase of χ is not physical in the Heisenberg limit. In particular, if we take $t \to 0$ and $U \to \infty$ (with J held constant) then the global charge symmetry $c_{\mathbf{x}\sigma} \to e^{i\Theta} c_{\mathbf{x}\sigma}$ of the Hamiltonian (2.3) is promoted to a local (or gauge) symmetry under which the electron phases can be rotated by a different amount at each point in space and time: $c_{\mathbf{x}\sigma}(\tau) \to e^{i\Theta(\mathbf{x},\tau)} c_{\mathbf{x}\sigma}(\tau)$. (The gauge transformations should be continuous in the time direction, however.) This local symmetry simply reflects the fact that the pure Heisenberg Hamiltonian conserves the number of electrons at each site. For the Lagrangian (2.9) to be invariant under such a gauge transformation, χ and ϕ must also transform in the following way:

$$\chi_{\mathbf{xy}}(\tau) \to \exp\{i[\Theta(\mathbf{y},\tau) - \Theta(\mathbf{x},\tau)]\}\chi_{\mathbf{xy}}(\tau) , \quad (2.14a)$$

$$\phi_{\mathbf{x}}(\tau) \to \phi_{\mathbf{x}}(\tau) - d\Theta(\mathbf{x},\tau)/d\tau . \quad (2.14b)$$

In fact, the χ fields transform in the same way as link variables of a lattice Abelian gauge theory. Thus, we may identify the phase of χ as a spatial gauge field $A_{\mathbf{e}}$:

$$\chi_{\mathbf{x},\mathbf{x}+\mathbf{e}} \equiv |\chi_{\mathbf{x},\mathbf{x}+\mathbf{e}}| \exp[iA_{\mathbf{e}}(\mathbf{x})] . \quad (2.15)$$

Again $\mathbf{e}$ is a unit vector pointing in either the $\hat{\mathbf{x}}$ or $\hat{\mathbf{y}}$ directions. Clearly, ϕ transforms as the temporal component of the gauge field. Finally, note that the $\Phi_{\mathbf{xy}}$ fields that we introduced above to describe biquadratic interactions are invariant under gauge transformations since they couple only to the charge singlet operator $\chi_{\mathbf{xy}} c_{\mathbf{y}}^{\dagger a} c_{\mathbf{x}a}$.

Compact gauge symmetries cannot break spontaneously, so χ cannot have an expectation value in the Heisenberg limit. However, the magnitude of χ *may* have an expectation value. In fact, fluctuations in the magnitude are suppressed in the large-n limit. In contrast, pure gauge fluctuations are *not* suppressed, even at $n \to \infty$, since the effective action depends on the gauge fields only through gauge-invariant combinations. Thus, while no gauge symmetry breaking occurs, the bond-centered charge-density waves associated with modulation of the *magnitude* of χ can arise. Of course, at nonzero t, the phase of χ is meaningful and does have an expectation value in the large-n limit. But, as $t \to 0$, the fluctuations in this phase become stronger and eventually restore the gauge symmetry.

The spatial plaquette operator $\Pi \equiv \chi_{12}\chi_{23}\chi_{34}\chi_{41}$ (1, 2, 3, and 4 are sites on the corners of a unit square) is another gauge-invariant object that will interest us. It obviously exists in spatial dimensions greater than one and is gauge invariant because the net phase from the gauge transformations on the four sites cancels out. From Eq. (2.15) it is clear that the phase θ of the plaquette is related to the gauge fields by $\theta = A_{12} + A_{23} + A_{34} + A_{41}$. By Stoke's theorem, θ is the flux due to a fictitious magnetic field penetrating the plaquette.

We will discuss saddle points with nonzero $\chi_{\mathbf{xy}}$. We emphasize that all gauge-equivalent saddle points have exactly the same action (or free energy). However, a consistent saddle-point evaluation of any observable must always average over all gauge-equivalent saddle points. Consequently, only gauge-invariant observables will be nonzero. We must be careful to classify saddle points in a gauge invariant way. For example, the single-particle spectrum of H_1 can be extracted once we know the saddle point. However, we should interpret this spectrum with care, since it is, in general, gauge dependent.

Finally, while lattice symmetries can break in the large-n limit, the global spin rotational [SU(n)] symmetry

must be maintained. This symmetry is protected because the available order parameters ϕ and χ are both singlets. However, Néel order apparently occurs at sufficiently small n (see Sec. VI) and breaks the SU(n) symmetry. It may be possible to see a tendency towards long-range ordering at finite n by examining higher-order corrections to the spin-spin correlation function in the $1/n$ expansion.

III. RESULTS IN ONE DIMENSION

We study the one-dimensional case (at zero temperature) first to gain some understanding of the behavior of the large-n solutions. In one dimension we can understand the qualitative behavior of the large-n saddle-point equation without resorting to the computer. We will compare the large-n results with the Bethe ansatz solution for the $n=2$ Hubbard-Heisenberg model and the renormalization group results for the general SU(n) Hubbard model.

To begin, we must first choose a reasonable ansatz for the equilibrium configurations of the χ and ϕ fields. Anticipating spontaneous dimerization of the chain, we choose a configuration that permits different values of χ on alternating links and different ϕ's on the two sublattices (see Fig. 4). We choose to consider only real values of χ. For the pure Heisenberg case we can always make the χ real by a gauge transformation since in one dimension there are no spatial plaquettes to frustrate this choice. For $t\neq 0$ we have no such gauge freedom, but impose the reality condition as part of the configuration ansatz.

With this choice for the fields, we can now evaluate the free energy. Since to leading order the χ and ϕ fields assume their classical expectation values, the electrons behave as fermions in a tight-binding model. The energy of a single electron of momentum k is given by

$$E(k)-i\phi=\pm[\,|\,2t\cos(k)+\chi_0\exp(ik)+\chi_e\exp(-ik)\,|^2-(\Delta\phi)^2]^{1/2}.$$

Here, $\Delta\phi\equiv(\phi_0-\phi_e)/2$ and $\phi\equiv(\phi_0+\phi_e)/2$ with the subscripts denoting odd and even sites. The evaluation of the total free energy [Eq. (2.13)] can now be performed by summing over the reduced Brillouin zone (now half the size of the original zone). The chemical potential μ must also be chosen by requiring the total number of electrons to equal $N(1-\delta)/2$ where δ is the doping below a half-filled band and N is the number of sites on the chain.

To find the minima of the free energy with respect to the two χ and two ϕ fields we first make the free energy completely real by continuing $\phi_{\mathbf{x}}$ into the complex plane with the redefinition $\phi_{\mathbf{x}}\rightarrow i\phi_{\mathbf{x}}$. Now we must *maximize* the free energy with respect to ϕ and $\Delta\phi$ since the factor of i changes the sign of the saddle-point curvature in the $\phi_{\mathbf{x}}$ directions. The (zero-temperature) free energy now reads

$$F(\phi,\Delta\phi,\chi_0,\chi_e;\mu)=Nn[(1/2J)(\chi_0^2+\chi_e^2)-(1/4U)(\phi^2+\Delta\phi^2)+\tfrac{1}{2}\phi]+n\sum_k[E(k)-\mu],$$

where

$$E(k)=\pm[\,|\,2t\cos(k)+\chi_0\exp(ik)+\chi_e\exp(-ik)\,|^2+(\Delta\phi)^2]^{1/2}-\phi,$$

and the sum over momenta k runs only over the filled states. Next, note that the total free energy is minimized in the χ directions by choosing the $(-)$ sign in the one-particle energy. We should now set $\Delta\phi=0$ (for $U>0$) to maximize the free energy with respect to $\Delta\phi$. Thus, there are no site-centered charge-density waves in the large-n limit. Finally, ϕ can acquire a positive expectation value at the saddle point for nonzero doping. However, its behavior is not interesting since it merely shifts the saddle-point free energy by a constant regardless of the χ fields. We will therefore ignore it in the following expressions. (For $\delta=0$ it is clear that $\phi=0$ at the saddle point since the two terms linear in the ϕ cancel out in the free energy.) Furthermore, at $t=0$ the Hamiltonian possesses the particle-hole symmetry $c_{\mathbf{x}}^{\dagger\sigma}\rightleftarrows c_{\mathbf{x}\sigma}$ that forbids nonzero expectation values for $\phi_{\mathbf{x}}$.

Defining $\Delta\chi\equiv(\chi_0-\chi_e)/2$ and $\chi=(\chi_0+\chi_e)/2$ we arrive at the single-particle energies

$$E(k;\chi,\Delta\chi)=-2[(t+\chi)^2\cos^2(k)+(\Delta\chi)^2\sin^2(k)]^{1/2}.$$

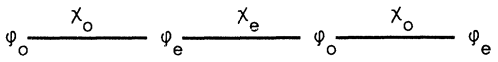


FIG. 4. Ansatz for the symmetry-breaking pattern on the chain.

Therefore, the ground-state free energy per site for an infinite chain is now

$$F(\chi,\Delta\chi)=n\left(1/(2J)[\chi^2+(\Delta\chi)^2]+\int(dk/\pi)E(k)\right).$$

The integral is over the filled region of the Brillouin zone and runs over the range $-\pi/2<k<\pi/2$ at half filling ($\delta=0$).

Focus attention on $\Delta\chi$. If we minimize $F(\chi,\Delta\chi)$ with respect to $\Delta\chi$ we can see how it behaves. The first derivative is

$$\partial F/\partial(\Delta\chi)=(1/J)\Delta\chi-\Delta\chi\int(dk/\pi)\sin^2(k)\times|\,E(k;\chi,\Delta\chi)\,|^{-1}.$$

Clearly, $\Delta\chi=0$ is a saddle point. However, it is unstable since the second derivative $\partial^2F/\partial(\Delta\chi)^2$ is negative and logarithmically singular because the single-particle energy vanishes at the two Fermi points. Nonzero $\Delta\chi$ opens a charge gap of $\Delta E=4\Delta\chi$ at the Fermi points that eliminates the singularity. Therefore, the system does have a bond-centered charge-density wave. (In fact, for $t=0$, the chain dimerizes completely.) Near half filling this dimerization persists; but if the doping is large enough a uniform

phase with $\Delta\chi=0$ will supersede the dimerized phase as the lowest energy state. In either case, there is no gap if $\delta>0$.

The behavior of the charged sector of the large-n Hubbard-Heisenberg model agrees qualitatively with the Bethe-ansatz solution for the $n=2$ Hubbard model. Lieb and Wu[12] find that at half filling, a charge gap opens for any $U>0$ and the system becomes an insulator. However, at any nonzero doping there is no charge gap and the system remains metallic. Had we considered the Hubbard model alone and not its generalization to the Hubbard-Heisenberg model, we would have found no dimerization (and no charge gap) since $\Delta\phi$ would still be zero. This behavior would be at odds with the exact solution but we see that a charge gap does develop for any J, no matter how small. Apparently, the pure large-n Hubbard model is unstable and an explicit spin-spin coupling should be included. We can view this instability as further justification for considering the hybrid Hubbard-Heisenberg model.

The Bethe-ansatz solution also shows no *spin* gap for $U>0$. However, the dimerized phase found above has a gap for any single-particle excitations, including spin excitations. To understand this conflict better, we apply renormalization-group methods to the general SU(n) Hubbard model. In particular, we will consider weak-coupling perturbation theory. (See Ref. 13 for a more detailed explanation of this method.) The SU(n) Hamiltonian in one dimension can be written as

$$H=\sum_x [t(c_x^{\dagger a}c_{x+1,a}+\text{H.c.})+U/n(c_x^{\dagger a}c_{xa}-n/2)^2] .$$

For small $U/(nt)$, we can take the continuum limit by considering only low-energy excitations near the two Fermi points. The dispersion relation of free electrons near the Fermi points is linear so the electrons there behave as chiral left- and right-moving relativistic fermions. Expressing the lattice electron operators in terms of these chiral fields by $c_{xa}=(+i)^x\psi_{La}(x)+(-i)^x\psi_{Ra}(x)$, we see that the noninteracting piece of the Hamiltonian density becomes $H_0=i(\psi_L^{\dagger a}\partial/\partial x\psi_{La}-\psi_R^{\dagger a}\partial/\partial x\psi_{Ra})$ (we have set the Fermi velocity to one). The corresponding Lagrangian density is $L_0=i(\psi_L^{\dagger a}\partial_-\psi_{La}-\psi_R^{\dagger a}\partial_+\psi_{Ra})$ where we have introduced the light-cone coordinates $x_\pm=(t\pm x)/2$. The low-momentum modes of the interacting four Fermi term may now be expressed as the sum of three terms: A U(1) charge current-current interaction, a SU(n) spin current-current interaction, and an umklapp piece. We define the currents as

$$\begin{aligned}&\text{U(1)},\quad J_{L,R}\equiv{:}\psi_{L,R}^{\dagger a}\psi_{L,Ra}{:}\,,\\&\text{SU}(n),\quad \mathbf{J}_{L,R}\equiv\psi_{L,R}^{\dagger a}\mathbf{T}_a^\beta\psi_{L,R\beta}\,,\end{aligned}$$

where the colons denote normal ordering so that $\langle J_L\rangle=\langle J_R\rangle=0$ and the T^a's ($a=1$ to n^2-1) are the generators of SU(n) that satisfy the relations

$$\begin{aligned}&\mathrm{Tr}(T^aT^b)=\delta^{ab}/2\,,\\&(\mathbf{T}_\beta^\alpha)\cdot(\mathbf{T}_\delta^\gamma)=\tfrac{1}{2}[\delta_\delta^\alpha\delta_\beta^\gamma-(1/n)\delta_\beta^\alpha\delta_\delta^\gamma]\,.\end{aligned}\tag{3.1}$$

The interaction Lagrangian density now reads

$$L_{\text{int}}=\lambda_1J_LJ_R+\lambda_2\mathbf{J}_L\cdot\mathbf{J}_R-2\lambda_3[(\psi_L^{\dagger a}\psi_{Ra})^2+\text{H.c.}]\,.$$

Left-left and right-right terms such as J_LJ_L do not appear here because they can be absorbed into a redefinition of the Fermi velocity in L_0 which we will continue to set equal to one.[13] The values of the bare coupling constants are $\lambda_1=2(1/n^2-1/n)U$, $\lambda_2=4U/n$, $\lambda_3=-U/2n$.

First-order perturbative renormalization of the coupling constants λ_m is now straightforward. Defining $s\equiv\ln(L)$ where L is the length scale at which the couplings are defined, we find

$$\begin{aligned}&d\lambda_1/ds=(32/\pi)(1/n-1)\lambda_3^2\,,\\&d\lambda_2/ds=-(n/2\pi)\lambda_2^2+(32/\pi)(2-n)\lambda_3^2\,,\\&d\lambda_3/ds=-(4/\pi)\lambda_1\lambda_3+(1/\pi)(1-n+2/n)\lambda_2\lambda_3\,.\end{aligned}$$

Note that for $n=2$ the renormalization of λ_2 decouples from λ_3 (it never couples to λ_1 because $\mathbf{J}_L\cdot\mathbf{J}_R$ is a charge singlet). This decoupling actually holds to all orders in perturbation theory and occurs only for $n=2$ since the umklapp term for that special case can be written as the spin-singlet object $[(\epsilon_{\alpha\beta}\psi_L^{\dagger\alpha}\psi_L^{\dagger\beta})(\epsilon^{\gamma\delta}\psi_{R\gamma}\psi_{R\delta})+\text{H.c.}]$. Thus, the SU(2) case behaves differently than the SU($n\neq2$) case. In particular, for $n=2$ and $U>0$, $\lambda_2\to0$ and $\lambda_1\to\infty$ as $L\to\infty$; therefore a charge gap, but no spin gap, opens up in agreement with the Lieb and Wu solution. However, for $n>2$, a simple numerical integration shows that both the charge and spin coupling constants renormalize to large values so both sectors develop a gap. Although this analysis was based on weak coupling, we expect it to remain qualitatively valid right up to $U=\infty$. For $n=2$ the gapless spin excitations describe those of the Heisenberg model. For $n>2$, there is no reason to have gapless spin excitations for any U. We *should* therefore expect the spin gap that our large-n result displays.

At nonzero doping the four-Fermi umklapp term disappears because umklapp processes no longer result in a change of 2π in the lattice momentum. Setting $\lambda_3=0$ in the renormalization-group equations we see that the other coupling constants no longer renormalize to large values so no gaps develop. Again, this conclusion is consistent with the large-n solution and the exact result for the Hubbard model.

Thus we see that the large-n limit is actually misleading in one dimension because the $n>2$ case behaves differently than the $n=2$ case. The Lieb-Schultz-Mattis theorem required either vanishing gap or broken translational symmetry (i.e., dimerization). We see that the former occurs at $n=2$ but the latter for $n>2$. Dimerization *can* occur at $n=2$ with the addition of a next-nearest-neighbor coupling:

$$H=J\sum_i\mathbf{S}_i\cdot\mathbf{S}_{i+1}+J_2\sum_i\mathbf{S}_i\cdot\mathbf{S}_{i+2}\,.$$

This system is dimerized for J_2/J greater than approximately $\frac{1}{4}$. In fact, for $J_2/J=\frac{1}{2}$, the ground state consists of disconnected dimers (like those depicted in Fig. 2).[14]

IV. RESULTS IN TWO DIMENSIONS

To study the two-dimensional case, we again make an ansatz for the minimal symmetry of saddle-point con-

figurations. The assumption of constant ϕ and χ is simplest since the translational symmetry cannot then break spontaneously. However, as in the one-dimensional case, such a configuration can be proven *not* to be the minimum of the free energy for small doping (i.e., close to half filling) because of a Fermi-surface instability connected with the perfect nesting of the Fermi surface for nearest-neighbor tight-binding models on a square lattice. It appears that the minimum amount of symmetry breaking which will allow a locally stable minimum is an increase of the unit-cell length by a factor of $\sqrt{2}$. Thus the square lattice breaks up into even and odd sublattices and the Brillouin zone is reduced to half its original size [see Fig. 5(a)]. ϕ_x may now have different values on each sublattice (ϕ_0 and ϕ_e), corresponding to a site-centered charge-density-wave state. Also, χ_{xy} can have four different values (χ_{12}, χ_{23}, χ_{34}, and χ_{41}) [Fig. 5(b)]. We will allow the χ fields to be complex, and therefore oriented. Finally, when the biquadratic interaction is present, we will permit the Φ_{xy} fields to have the same symmetry breaking pattern as the χ_{xy} fields. We will check the stability of the saddle-point configurations we find using this ansatz. Of course, energy barriers must exist between our saddle points and any other possible saddle points that break the $\sqrt{2}$ unit cell to insure stability.

For the pure Heisenberg model ($t=\delta=0$) we can always gauge transform a given saddle-point state and still have a saddle point. However, unlike the one-dimensional case, for $d=2$ we cannot in general find a gauge transformation that will make all the χ's purely real because the plaquets $\Pi\equiv\chi_{12}\chi_{23}\chi_{34}\chi_{41}$ could have a nontrivial phase. In other words, if the plaquette is not purely real, no gauge transformation can make all four χ's real. Therefore, if the flux θ does not equal 0 or π (mod2π), at least one of the four χ fields must have an imaginary component. Of course, when $t\neq 0$, the "gauge" transformations no longer leave the Hamiltonian invariant and therefore generate distinct states that are generally not saddle points.

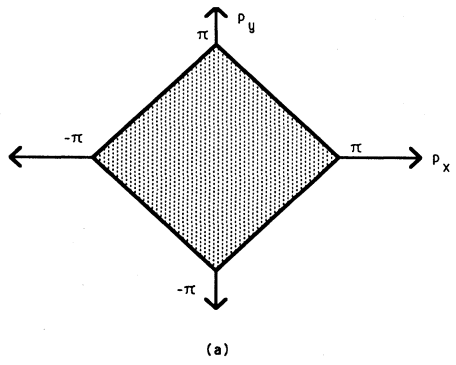
For simplicity, we first consider the problem with no biquadratic interactions ($\tilde{J}=0$). With the above ansatz for the χ and ϕ fields at the saddle point, we can now evaluate the free energy. The electrons have single-particle energies:

$$E(k_x,k_y)-i\phi=\pm\{|2t[\cos(k_x)+\cos(k_y)]+\chi_{12}\exp(ik_x)+\chi_{34}\exp(-ik_x)$$
$$+\chi_{23}^{*}\exp(ik_y)+\chi_{41}^{*}\exp(-ik_y)|^2-(\Delta\phi)^2\}^{1/2}$$

and as before the total free energy is given by Eq. (2.13).

Again we should set $\Delta\phi\equiv(\phi_0-\phi_e)/2=0$ for any filling factor because of the quadratic terms $(\Delta\phi)^2$ in the free energy. Also, the dependence on ϕ is again trivial. We can now calculate the free energy numerically and employ the Metropolis Monte Carlo algorithm to find the minima. That is, we set an artificial Monte Carlo (MC) temperature β_{MC} and evaluate the free energy for a given starting set of values for the four real and four imaginary components of the χ fields. A new point in this parameter space is then chosen by moving in a random direction away from the old point. After each move we must recalculate the chemical potential by requiring that the average electron density equal $(1-\delta)/2$ per site. (Again, the chemical potential is strictly zero for the $\delta=0$ case.) The new point is accepted with conditional probability one if the free energy decreases ($\Delta F<0$) and with probability $\exp(-\beta_{MC}\Delta F)$ if it increases. This process is repeated and as we approach the equilibrium values of the χ fields, we decrease the temperature (simulated annealing) and make smaller random jumps until the fluctuations away from the equilibrium point are sufficiently small to identify the phase.

We will classify the saddle-point solutions by their symmetries. For this purpose, it is sufficient to consider parity (P), charge conjugation (C), translation by one site (T),



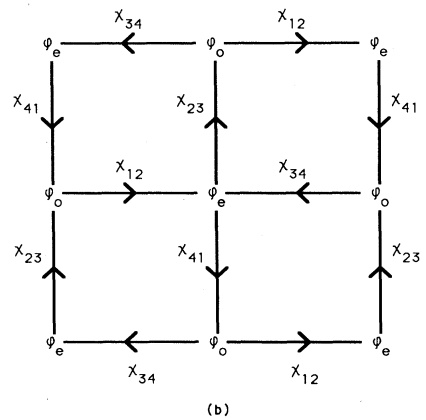


FIG. 5. (a) The reduced Brillouin zone for the $\sqrt{2}$ unit cell. (b) Ansatz for the symmetry-breaking pattern in the two-dimensional Hubbard-Heisenberg model.

and rotations through 90° (R). Since we are considering a lattice problem, we must specify these operations more precisely. By parity we will mean a reflection across a line that cuts through the midpoint of a link. Implicit in our discussion is the fact that we could reflect about either vertical or horizontal lines. We will say that parity is broken if either operation changes the saddle-point solution. Similarly, we could study translations by one site in either the vertical or horizontal directions. Again, a saddle-point solution that breaks either translational symmetry will be said to break T. R will be a rotation of 90° about the center of a plaquette. Finally, by C we mean complex conjugation of the link fields χ_{xy}. We must remember, however, that the gauge-invariance of the pure Heisenberg model allows us a great amount of freedom in our choice of the χ fields. For this special case, we instead consider the effect of complex conjugation on gauge-invariant quantities such as the plaquette Π.

We treat the Heisenberg ($t=\delta=0$) case separately here from the more general Hubbard-Heisenberg model because the gauge invariance of the former model leads to qualitatively different behavior. At zero temperature our ansatz yields two saddle points.

(1) *Peierls.* Only one of the four χ's is nonzero and it may be made real by a gauge transformation. The nonzero link has $|\chi| = J/2$ and the saddle-point free energy is $F=-nJ/8$ per site. It violates P, R, and T but preserves C.

(2) *Flux.* All of the χ's are equal in magnitude and the flux $\theta=\pi$ since Π is real and negative. There is only one gauge nonequivalent state of this kind and it has a free energy about 10% higher than the Peierls phase: $F\approx -0.115(nJ)$ per site. This state preserves P, T, R, and C.

The Peierls state corresponds to the dimerized state we found in one dimension because each site forms a dimer with one of its neighbors. The electrons are completely localized on each bond and the spectrum has a gap of $\Delta E = J$. The state obviously breaks T symmetry and the gap is consistent with the LSM theorem. In our ansatz, the dimers form staggered columns. However, since the electrons are localized, the saddle point is actually highly degenerate. Any configuration of dimers will be a large-n solution as long as every lattice site is attached to a nearest-neighbor site by a dimer. However, this degeneracy is broken by $1/n$ corrections. In fact, Read and Sachdev[4] find that the dimers arrange into regular columns at next order in the $1/n$ expansion [see Fig. 6(a)]. This column state breaks the $\sqrt{2}$ unit cell into a 2×1 cell that can be described by three different χ fields [Fig. 6(a)]. Curiously, the column state no longer breaks P.

By considering a larger 2×2 unit cell, Dombre and Kotliar demonstrated that the flux saddle point is unstable.[5] Indeed, it collapses into a new "box" phase that has the same free energy as the Peierls phase but is separated from it by energy barriers. This new phase is depicted in Fig. 7. It consists of disconnected plaquettes and the χ fields are zero on half of the available links. The flux passing through the nonzero plaquettes is π. This phase breaks T and R but preserves P and C (it is invariant under C since Π is real for every plaquette).

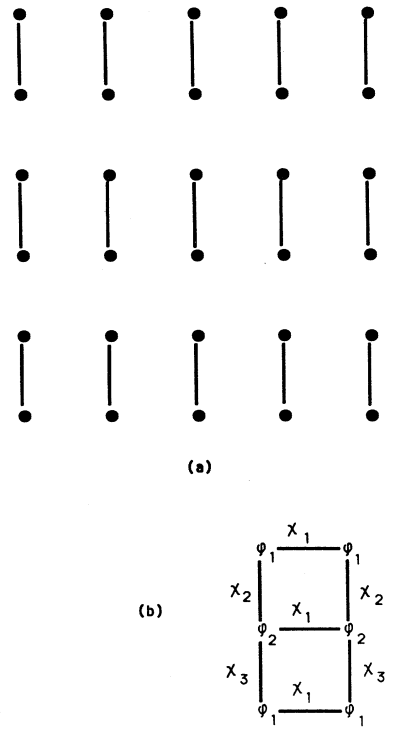


FIG. 6. (a) The ground state determined by $1/n$ corrections to the Peierls saddle point. The dimers form columns. (b) The 2×1 unit cell needed to describe the column state.

Apparently, it is important to test the local stability of the saddle points in the most general possible way. Only then can we be sure that the saddle points we find with a given ansatz are completely stable to small perturbations. (Of course, a saddle point can be locally stable and still not be the *global* minima of the free energy. We will not address this possibility directly, but rather argue that the configurations we consider are sufficiently general to accommodate the global minima.) To accomplish that end, we consider the effect of adding small but spatially arbitrary perturbations to the χ_{xy} fields and examine the change in the free energy. This problem is equivalent to finding the curvature of the saddle point in the infinite-dimensional functional space of the χ_{xy} fields. Fortunately, in momentum-space perturbations with different wave vectors decouple and the problem becomes tractable. The

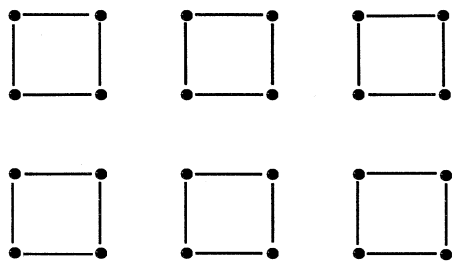


FIG. 7. The box phase that is degenerate in the $n\rightarrow\infty$ limit with the Peierls phases. χ is nonzero only on the depicted links.

details of this calculation are presented in the Appendix; here we note that the flux phase is unstable for a range of momenta near $(0,\pi)$ and $(\pi,0)$. Indeed, these modes correspond to the instability that moves the flux phase into the box phase.

Since the flux phase has great intrinsic interest, we seek to stabilize it by adding the biquadratic interaction. To explore this possibility, we calculate the saddle-point free energy of the Peierls, flux, and box phases for $\tilde{J}\neq 0$. We first must analytically continue the saddle-point value of Φ_{xy} into the complex plane with the redefinition: $\Phi_{xy} \rightarrow i\Phi_{xy}$. Now the free energy associated with the Lagrangian (2.12) is completely real. Next, we will assume that the Φ_{xy} fields are nonzero only on the links that have nonzero χ_{xy}. There is now no need to recalculate the fermionic contribution to the free energy since we already did that above. Setting $J=1$ for simplicity, the free energy is

$$F(\chi,\Phi)=n\{2a[|\chi|^2-(1/\tilde{J})\Phi^2]-(b/2)(1+2\Phi)^{1/2}|\chi|\},$$

where $|\chi|$ and Φ denote their values on the nonzero links and a equals the fractional number of such links. Specifically, $a=\frac{1}{4}$ (Peierls), $\frac{1}{2}$ (box), and 1 (flux). The fermionic contribution to the free energy is encapsulated in the number b which equals 1 (Peierls), $\sqrt{2}$ (box), and ~ 1.916 (flux). The saddle-point equations $\partial F/\partial\chi=0$ and $\partial F/\partial\Phi-0$ then give $\Phi=-\tilde{J}(b/8a)^2$ and $\chi=(b/8a)\times(1+2\Phi)^{1/2}$. A simple calculation now shows that the flux-free energy does indeed drop below both the Peierls and box free energies for $\tilde{J}\gtrsim 1.134$.

Two possible objections must be addressed before we can be sure that this flux state is sensible. First, is the critical value $\tilde{J}$ so large that the system is now a ferromagnet? The answer to this question is no, as long as $\tilde{J}<2$ since (as mentioned earlier) the two-site problem can be shown to be antiferromagnetic in that case.

The other objection concerns the stability of the new flux state. Is it truly stable to all perturbations in the auxiliary fields? This question is complicated by the fact that the saddle-point free energy should be a *maximum* with respect to the redefined Φ fields and a *minimum* with respect to the χ fields. We sort out this problem by integrating out the Φ fields (exactly in the $n\rightarrow\infty$ limit). We then calculate the saddle-point curvature of the effective action that now depends only on the χ fields and which must be positive in all directions for stability. We conclude (see Appendix) that the unstable directions disappear for $\tilde{J}\gtrsim 1.45$. (Some intermediate phase apparently interpolates between the box and flux states in the range $1.134\lesssim\tilde{J}\lesssim 1.45$.)

The flux phase appears to break translational symmetry since χ_{12} and χ_{34}, for example, are oriented in opposite directions. This orientation is a gauge artifact, however. Because the plaquette Π is real, we can always find a gauge transformation that makes all of the χ's real and therefore unoriented. By the Lieb-Schultz-Mattis theorem, we expect gapless spin excitations. Similarly, this phase also respects C and P because Π is invariant under these operations.

The electronic spectrum for the flux state is not gauge invariant. For example, gauge transformations that shift the phases of the four χ's also move the spectrum around in momentum space:

$$E(k_x,k_y)\rightarrow E(k_x+\alpha,k_y+\beta).$$

However, this gauge dependence does not lead to any inconsistencies since the physical excitations consist of particle-hole pairs confined together on the same site. For example, the operator $c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{x}\beta}$ produces a spin excitation at site $\mathbf{x}$ and is clearly gauge invariant. Thus, the dispersion relation for spin waves should not depend on the choice of gauge. Apparently, the higher-order processes (in the $1/n$ expansion) that lead to particle-hole confinement play a crucial role in maintaining gauge invariance. (We mention one possible mechanism below.)

In the gauge with $\chi_{12}=\chi_{34}=i|\chi|$ and $\chi_{23}=\chi_{41}=|\chi|$ the single-particle spectrum is

$$E(\mathbf{k})=2|\chi|[\cos^2 k_x+\cos^2 k_y]^{1/2}.$$

Note that the gap vanishes only at the four Fermi points $\mathbf{k}=(\pm\pi/2,\pm\pi/2)$. Thus, the particle-hole excitations have zero-energy modes at momenta $(0,0)$, $(0,\pi)$, $(\pi,0)$, and (π,π) in accord with the LSM theorem.

The gap vanishes linearly at the Fermi points, so the low-energy theory is a $2+1$ dimensional (massless) relativistic free-fermion quantum-field theory. Let us examine the low-energy excitations more closely. With the above gauge choice, the Hamiltonian becomes

$$H=|\chi|\sum_{\mathbf{r},\text{even}}\{c_{\mathbf{r}}^{\dagger}[i(c_{\mathbf{r}+\hat{x}}+c_{\mathbf{r}-\hat{x}})+(c_{\mathbf{r}+\hat{y}}+c_{\mathbf{r}-\hat{y}})]+\text{H.c.}\}.$$

Here the sum over $\mathbf{r}$ runs only over the even sublattice points. In momentum space this Hamiltonian reads

$$H=2|\chi|\int[d^2\mathbf{k}/(2\pi)^2]\times\{[i\cos(k_x)+\cos(k_y)]c_e^{\dagger}(\mathbf{k})c_o(\mathbf{k})+\text{H.c.}\}. \tag{4.1}$$

We take the two nonequivalent points where the gap vanishes in the reduced Brillouin zone to be $(\pi/2,\pm\pi/2)$. The low-energy Hamiltonian near these points can be written in terms of shifted momenta $k_x=\pi/2+k_x'$ and $k_y=\pm\pi/2+k_y'$ as

$$H\approx -2|\chi|\int[d^2\mathbf{k}'/(2\pi)^2]\{[(ik_x'+k_y')c_{e1}^{\dagger}(\mathbf{k}')c_{o1}(\mathbf{k}')+\text{H.c.}]+[(ik_x'-k_y')c_{e2}^{\dagger}(\mathbf{k}')c_{o2}(\mathbf{k}')+\text{H.c.}]\}.$$

We may now introduce two-component Dirac spinors, $\psi_1\equiv(c_{e1},c_{o1})$ and $\psi_2\equiv(c_{o2},c_{e2})$. In this basis, the continuum Hamiltonian acquires a particularly simple form

$$H=-2i|\chi|\int d^2\mathbf{x}\psi_a^{\dagger}(\sigma_1\partial_1+\sigma_2\partial_2)\psi_a.$$

Here we are implicitly summing over the $a=1,2$ fermions. Defining $\bar{\psi}=\psi^{\dagger}\gamma_0$ and $\gamma_\mu=(i\sigma_3,\sigma_1,\sigma_2)$, the continuum limit Lagrangian density is also simple:

$$\mathcal{L}=-i\bar{\psi}_a\gamma_0\partial_0\psi_a-\mathcal{H}=i\bar{\psi}_a\gamma^\mu\partial_\mu\psi_a\,.$$

Note that the γ matrices obey the Clifford algebra: $\{\gamma_\mu,\gamma_\lambda\}=2g_{\mu\lambda}$. We also have adopted the standard index raising convention so $\gamma^\mu=(-\gamma_0,\gamma_1,\gamma_2)$ and set the speed of "light" $c=2|\chi|$ to one. It follows that L is invariant under Lorentz transformations and describes two Dirac fermions for each spin component. Furthermore, this theory possesses a global SU($2n$) symmetry since it is invariant under transformations that mix the n flavor and the two Fermi-point indices. Lattice effects, however, break this SU($2n$) $\rightarrow$ SU(n).

For $n\rightarrow\infty$ we can ignore fluctuations of the gauge fields. The effective action is of course proportional to n so that any fluctuations away from the saddle point are suppressed by powers of $1/n$. This means that the square of the effective U(1) coupling constant g^2 is $O(1/n)$. Thus in the large-n limit we obtain a system of free fermions. However, at next order in $1/n$ we must include interactions with the ϕ, χ, and Φ fields. These interactions could qualitatively change the low-energy excitations, even at very large n. Indeed, the compact U(1) gauge symmetry cannot break so the gauge interactions must confine the fermions. Thus, the leading order spectrum consists of noninteracting particle-hole pairs but at finite n we should find that the spectrum consists of confined particle-hole pairs (mesons), and possibly other collective excitations. It is not at all obvious that the massless excitations that we found in the flux phase at momenta (0,0), $(0,\pi)$, $(\pi,0)$, and (π,π) will survive at finite n. For example, a gap of order $1/n$ might be produced. If a gap forms the Lieb-Schultz-Mattis theorem suggests that either translational symmetry has broken or else some other types of gapless spin excitations exist.

We can include the background field fluctuations away from the flux saddle point in the continuum theory. For this purpose it is convenient to introduce the link fields $\sigma_e(\mathbf{x})$ and $A_e(\mathbf{x})$:

$$\chi_{\mathbf{x},\mathbf{x}+\hat{\mathbf{x}}}\equiv Ji(-1)^{x+y}\sigma_{\hat{\mathbf{x}}}\exp(igA_{\hat{\mathbf{x}}})\,,$$

$$\chi_{\mathbf{x},\mathbf{x}+\hat{\mathbf{y}}}\equiv J\sigma_{\hat{\mathbf{y}}}\exp(igA_{\hat{\mathbf{y}}})\,.$$

Here, $\mathbf{x}\equiv(x,y)$. With the above choice of gauge, the saddle point corresponds to setting $A_e=0$ and $\sigma_e=$const. Of course, the A_e are the spatial components of the U(1) gauge field mentioned earlier and we can combine them with the time component $\phi(\mathbf{x})$ to form the three-vector field A_μ. If we ignore the σ and Φ excitations (which have a mass gap), the continuum theory assumes the usual relativistic form:[15]

$$L=i\bar{\psi}_a\gamma^\mu(\partial_\mu+igA_\mu)\psi_a\,.$$

(Note that the gauge field does not couple $\bar{\psi}_1$ to ψ_2 or $\bar{\psi}_2$ to ψ_1 in the continuum approximation.) Apparently, the low-energy excitations in the flux phase of the Heisenberg model are described by a 2+1 dimensional relativistic U(1) gauge theory of massless fermions.

Let us return to the hybrid Hubbard-Heisenberg model problem with no biquadratic interactions ($\tilde{J}=0$). We find two new phases in the parameter space of different values of the doping δ and the ratio t/J. These regions are depicted in the phase diagram of Fig. 8. The Peierls state of the Heisenberg model persists for small doping and hopping term t. However, the flux state becomes the ground state at small doping and larger values of t. We again characterize the various phases by their symmetries.

(1) *Peierls.* The χ's are complex and $\chi_{12}=\chi_{23}=\chi_{34}$ but $|\chi_{41}|>|\chi_{12}|$. (There are three other equivalent states corresponding to 90° rotations of this state.) This phase breaks P, T, R, and C. (However, C is a symmetry when $\delta=0$ because the four χ's are then real.)

(2) *Flux.* All the χ fields are equal and have imaginary components. This state breaks P, T, and even C since the orientation of the links now cannot be removed by a gauge transformation. The 90° rotations remain a symmetry, however.

(3) *Uniform.* All the χ fields are equal and real. This state is invariant under translations, reflections, rotations, and complex conjugation (i.e., it has the full tetragonal symmetry).

(4) *Kite.* The χ's are nonzero, real (invariant under C), and are equal in pairs. All three kite states break R but only two break P and T: $\chi_{12}=\chi_{23}\neq\chi_{34}=\chi_{41}$ and $\chi_{12}=\chi_{41}\neq\chi_{23}=\chi_{34}$. The third state respects these symmetries ($\chi_{12}=\chi_{34}\neq\chi_{23}=\chi_{41}$).

Near half filling ($\delta=0$) and small t/J the partially dimerized Peierls state has the lowest energy. It is just partially dimerized now because (for $t\neq0$) all the χ's are nonzero even though the number of valence bonds is larger on one of the four links. Thus, this state remains a staggered bond-centered charge-density wave. The electronic spectrum is completely gapped at $\delta=0$ so this state insulates at half filling. To test its stability, we also calculated the free energy at half filling of the "column" Peierls phase [using the unit cell depicted in Fig. 6(b) with real χ_1, χ_2, and χ_3]. Although it is degenerate (at $n\rightarrow\infty$) with our "staggered" Peierls phase for $t=0$, we find that it has a *higher* free energy for $t\neq0$. Finally, we checked that the saddle-point curvature of the staggered Peierls state is positive, at least for $\delta=0$ and $0\leq t\lesssim0.4$ (see the Appendix).

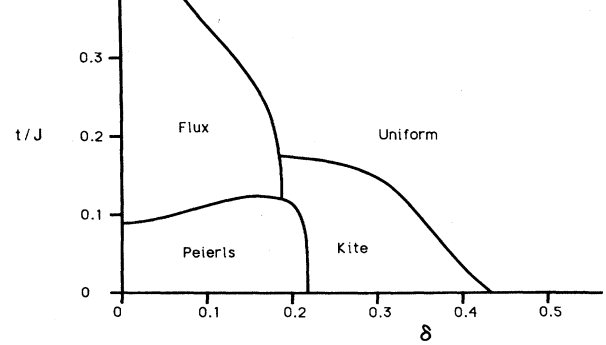


FIG. 8. Phase diagram for the saddle points of the two-dimensional Hubbard-Heisenberg model.

For larger values of t/J the flux state has the lowest energy and is the genuine large-n ground state. In fact, for the case of zero doping we have verified (again by the saddle-point curvature calculation) that it is stable to further symmetry breaking for $t/J \gtrsim 0.15$ so that it is at least a local minima. Physically, we might expect a phase of this type for sufficiently large t because, unlike the one-dimensional problem, the flux phase has no Fermi-surface instability since the density of states near the Fermi points vanishes linearly with the energy. [The electronic spectrum of the flux state still has four Fermi points at momenta $(\pm\pi/2, \pm\pi/2)$ at half filling.] Thus, the saddle-point equation can be solved in two dimensions without gapping the entire surface. Also, we know that the uniform phase must occur as $t/J \to \infty$. Thus, the flux phase is the logical candidate to interpolate between the Peierls and uniform phases at intermediate values of t.

The flux phase is still undimerized in the hybrid model since the number of valence bonds (proportional to $|\chi_{\mathbf{xy}}|^2$) is the same on all links. However, diamagnetic currents flow around the plaquettes in an alternating sense and break the tetragonal symmetry. In fact, the current on the link $\mathbf{xy}$ is

$$j_{\mathbf{xy}} = \langle (c_{\mathbf{x}}^{\dagger\alpha} c_{\mathbf{y}\alpha} - c_{\mathbf{y}}^{\dagger\alpha} c_{\mathbf{x}\alpha}) \rangle = (2n/J)\,\mathrm{Im}\{\langle \chi_{\mathbf{xy}} \rangle\} .$$

(These currents also flow in the $\delta > 0$, $t > 0$ Peierls phase since it too violates C.) Note that the current becomes an unobservable gauge artifact in the pure Heisenberg limit since all of the χ fields can be made real by a gauge transformation.

The new uniform state corresponds to the undimerized state found in one dimension for large doping. It occurs for sufficiently large doping since the Fermi-surface singularity disappears. Because the χ's are real and equal, they simply renormalize t in the single-particle spectrum. Thus there are no gaps and this state corresponds to the spin-liquid state proposed by Baskaran, Zou, and Anderson.[16]

Finally, we find "kite" states at intermediate doping that, like the Peierls phase, consist of bond-centered change-density waves. The links with a higher valence-bond density form lines that run through the lattice. For example, the state with $\chi_{12} = \chi_{23}$ makes zigzags through the lattice. These strong bonds should have an excess electron density and hence a net negative charge. This state is actually highly degenerate since any combination of straight lines and zigzags that have the same connectivity will have the same free energy. Again, we expect $1/n$ corrections to break this degeneracy and pick out the true ground state.

The phase diagram of Fig. 8 is unaffected by small nonzero temperatures. This behavior respects the Coleman-Mermin-Wagner theorem for two spatial dimensions because the different phases are characterized by the breaking of *discrete* lattice symmetries. [In contrast, *continuous* symmetries such as the SU(n) spin symmetry can break only at zero temperature in two spatial dimensions.]

For the pure Heisenberg system (with $\tilde{J} = 0$), one might expect a transition from the Peierls phase to the flux phase at a nonzero temperature because entropy favors the delocalization of the electrons. Instead, the Peierls phase changes directly to the uniform $\chi_{\mathbf{xy}} = 0$ phase at a temperature $T = J/4$. This transition can be shown to be second order by the usual mean-field methods. Clearly, a transition to a uniform phase will always occur in the hybrid model at sufficiently high temperatures because thermal fluctuations effectively remove the Fermi-surface instability.

V. SPIN-SPIN CORRELATIONS IN THE HEISENBERG MODEL

The spin-spin correlation function for the copper-oxide planes in the high-T_c superconductors has been determined experimentally by neutron scattering (see, for example, Ref. 17). To make contact between our large-n solution and experiment, in this section we study the correlator in the uniform and flux phases at zero doping. (There are no low-energy spin excitations in the Peierls phase.) Of course, the uniform saddle point is unstable at low temperatures, but we will investigate its behavior anyway. We present results for both the low-energy (quasistatic) and instantaneous correlators.

Since $|\chi_{\mathbf{xy}}|^2$ only gives the nearest-neighbor correlator, we instead work directly with the electron fields. Relations (3.1) for the SU(n) generators allow us to express the spin-spin correlator in terms of the fermions:

$$\langle \mathbf{S}_{\mathbf{x}}(t) \cdot \mathbf{S}_{\mathbf{0}}(0) \rangle = \tfrac{1}{2} \langle c_{\mathbf{x}}^{\dagger\alpha}(t) c_{\mathbf{x}\beta}(t) c_{\mathbf{0}}^{\dagger\beta}(0) c_{\mathbf{0}\alpha}(0) \rangle - (1/2n) \langle c_{\mathbf{x}}^{\dagger\alpha}(t) c_{\mathbf{x}\alpha}(t) c_{\mathbf{0}}^{\dagger\beta}(0) c_{\mathbf{0}\beta}(0) \rangle .$$

To proceed further, we transform to momentum and frequency space. For the uniform phase, the odd and even sublattices can be treated together and the correlator becomes

$$\langle \mathbf{S}_{\mathbf{Q}}(\omega) \cdot \mathbf{S}_{-\mathbf{Q}}(0) \rangle = \tfrac{1}{2} \int d^2p\, d\eta/(2\pi)^3 [\langle c_{\mathbf{Q}+\mathbf{p}}^{\dagger\alpha}(\omega+\eta) c_{\mathbf{p}\beta}(\eta) c_{\mathbf{p}}^{\dagger\beta}(\eta) c_{\mathbf{Q}+\mathbf{p}\alpha}(\omega+\eta) \rangle - (1/n) \langle c_{\mathbf{Q}+\mathbf{p}}^{\dagger\alpha}(\omega+\eta) c_{\mathbf{p}\alpha}(\eta) c_{\mathbf{p}}^{\dagger\beta}(\eta) c_{\mathbf{Q}+\mathbf{p}\beta}(\omega+\eta) \rangle] .$$

Since the electrons are free in the $n \to \infty$ limit, these four-point functions can be decomposed exactly into products of two-point functions using Wick's theorem. The two-point functions are simply

$$\langle c_{\mathbf{p}}^{\dagger\alpha}(\eta) c_{\mathbf{p}\beta}(\eta) \rangle = \delta_{\beta}^{\alpha} \delta[\eta - E(\mathbf{p})] f[E(\mathbf{p})]$$

where $f[E] = [\exp(\beta E) + 1]^{-1}$ is the Fermi function and $E(\mathbf{p}) = |\chi|[\cos(p_x) + \cos(p_y)]$ is the single-particle energy. ($|\chi| = 2J/\pi^2$ at the *unstable* uniform saddle point.) Upon contracting and performing the integral over the frequencies

η we find

$$\langle \mathbf{S}_{\mathbf{Q}}(\omega)\cdot\mathbf{S}_{-\mathbf{Q}}(0)\rangle = (n^2-1)/2\int d^2p/(2\pi)^2 f[E(\mathbf{p}+\mathbf{Q})]\{1-f[E(\mathbf{p})]\}\delta[\omega+E(\mathbf{q})-E(\mathbf{Q}+\mathbf{p})]\,.$$

The instantaneous correlator is easy to calculate:

$$\langle \mathbf{S}_{\mathbf{Q}}(t=0)\cdot\mathbf{S}_{-\mathbf{Q}}(t=0)\rangle = \int d\omega \langle \mathbf{S}_{\mathbf{Q}}(\omega)\cdot\mathbf{S}_{-\mathbf{Q}}(0)\rangle$$
$$= (n^2-1)/2\int d^2p/(2\pi)^2 f[E(\mathbf{p}+\mathbf{Q})]\{1-f[E(\mathbf{p})]\}\,. \tag{5.1}$$

Since some of the neutron-scattering experiments integrate over a range of exchange energies, it is also interesting to compute the finite-energy correlator:

$$\langle \mathbf{S}_{\mathbf{Q}}\cdot\mathbf{S}_{-\mathbf{Q}}\rangle_{\Delta} \equiv \int d\omega\,[\exp-(\omega^2/2\Delta^2)]\langle \mathbf{S}_{\mathbf{Q}}(\omega)\cdot\mathbf{S}_{-Q}(0)\rangle$$
$$= (n^2-1)/2\int d^2p/(2\pi)^2 f[E(\mathbf{p}+\mathbf{Q})](1-f[E(\mathbf{p})]\exp(-\{[E(\mathbf{q})-E(\mathbf{Q}+\mathbf{p})]^2/2\Delta^2\})\,.$$

Here Δ controls the range of integration over the neutron energies. In practice, Δ is much smaller than J so the correlation function only measures quasistatic low-energy processes near the Fermi surface.

For the purpose of evaluating the correlators, we now set $n=2$. We can check the SU(2) quantum-mechanical sum rule $\mathbf{S}_{\mathbf{x}}\cdot\mathbf{S}_{\mathbf{x}}=\frac{3}{4}$ by integrating the correlator (5.1) over all momenta $\mathbf{Q}$. However, this integral yields only $\frac{3}{8}$. Arovas and Auerbach note that the sum rule fails because the factorization of four-point functions into products of two-point functions is valid only in the $n\to\infty$ limit. Higher-order terms in the $1/n$ expansion apparently lead to the recovery of the sum rule.[3]

The correlators peak at wave vectors (0,0) and (π,π) because low-energy processes can occur at these momenta. [See Fig. 9 for an example of a process that contributes to the (π,π) correlator.] Figure 10 displays the low-temperature ($T=0.02J$), low-energy correlator evaluated at $\Delta=0.14J$. The (π,π) scattering peak is clearly evident. Note that the magnitude of the peak is quite small compared to that of the instantaneous correlator (Fig. 11). It is small because only a few scattering processes are energetically allowed. The instantaneous correlator is peaked at (π,π) though the width is half the zone size. It is quite insensitive to the details of the single-particle spectrum since all processes are energetically possible.

Turning now to the flux phase, we must distinguish between even and odd sites. The eigenstates of the flux Hamiltonian (4.1) are created by the single-particle operator $\psi^{\dagger}_{\pm}(\mathbf{k})$:

$$H\psi^{\dagger}_{\pm}(\mathbf{k})\,|0\rangle = \pm E(\mathbf{k})\psi^{\dagger}_{\pm}(\mathbf{k})\,|0\rangle$$

where now $E(\mathbf{k})=2|\chi|[\cos^2(k_x)+\cos^2(k_y)]^{1/2}$ and $\chi\approx 0.239J$. The $\psi^{\dagger}_{\pm}(\mathbf{k})$ operators may be expressed in terms of the even- and odd-site creation operators by

$$\psi^{\dagger}_{\pm}(\mathbf{k}) = 2^{-1/2}[c_e^{\dagger}(\mathbf{k}) \pm g(\mathbf{k})c_o^{\dagger}(\mathbf{k})]\,.$$

Here the phase factor

$$g(\mathbf{k}) = [\cos(k_x)-i\cos(k_y)]/[\cos^2(k_x)+\cos^2(k_y)]^{1/2}\,.$$

The various two-point functions that contribute to the correlator such as $\langle c_o^{\dagger}(\mathbf{k},\omega)c_o(\mathbf{k},\omega)\rangle$ and $\langle c_e^{\dagger}(\mathbf{k},\omega)c_o(\mathbf{k},\omega)\rangle$ can now be expressed as linear combinations of the two $\psi_{\pm}$ two-point functions:

$$\langle \psi^{\dagger}_{-}(\mathbf{k},\omega)\psi_{-}(\mathbf{k},\omega)\rangle = \delta[\omega+E(\mathbf{k})]f[-E(\mathbf{k})]\,,$$

$$\langle \psi^{\dagger}_{+}(\mathbf{k},\omega)\psi_{+}(\mathbf{k},\omega)\rangle = \delta[\omega-E(\mathbf{k})]f[E(\mathbf{k})]\,.$$

Also,

$$\langle \psi^{\dagger}_{+}(\mathbf{k},\omega)\psi_{-}(\mathbf{k},\omega)\rangle = \langle \psi^{\dagger}_{-}(\mathbf{k},\omega)\psi_{+}(\mathbf{k},\omega)\rangle = 0$$

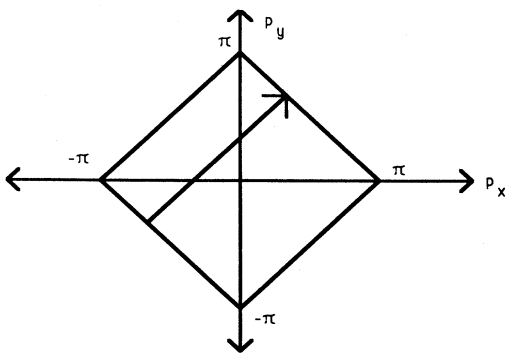


FIG. 9. Fermi surface of the uniform phase and a (π,π) excitation that contributes to the peak in the spin-spin correlator.

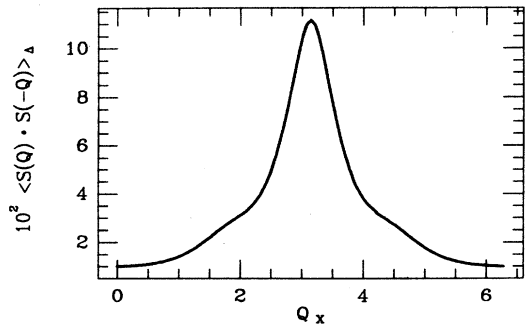


FIG. 10. The low-energy ($\Delta=0.14J$) spin-spin correlator in the uniform phase at momenta $Q_y=\pi$. The temperature is $T=0.02J$.

at zero temperature since all the positive energy states are empty. For this case, the spin-spin correlator is particularly simple:

$$\langle \mathbf{S}_{\mathbf{Q}}(\omega)\cdot\mathbf{S}_{-\mathbf{Q}}(0)\rangle = (n^2-1)/8\int d^2p/(2\pi)^2[1-g^*(\mathbf{p}+\mathbf{Q})g(\mathbf{p})]\delta[\omega+E(\mathbf{p})+E(\mathbf{p}+\mathbf{Q})]\,.$$

The low energy ($\Delta=0.14J$) and instantaneous correlators are plotted in Figs. 12 and 13. Again there is a peak in the low-energy correlator at momenta (π,π) but there are also smaller peaks at $(0,\pi)$ and $(\pi,0)$ corresponding to the other processes that link Fermi points. As expected, the instantaneous correlator differs little from that of the uniform phase.

VI. SU(2) HEISENBERG MODEL

To gain some understanding of relevance of our large-n solutions to the real SU(2) Heisenberg model, we consider several aspects of the $n\rightarrow 2$ limit. We first note that long-range Néel order probably occurs in the two-dimensional SU(2) system. We then discuss a variational approach to understanding the ground state that makes use of the large-n saddle-point wave functions. Recent numerical calculations of the energy of the Gutzwiller projected wave functions support the conjecture that the flux phase is a better saddle point than the Peierls phase for describing the SU(2) case. Finally, we mention the local SU(2) *gauge* symmetry that appears in the $n=2$ system.

Recent exact,[18] numerical,[19] variational,[11] and renormalization-group calculations suggest that Néel order occurs in the zero-temperature two-dimensional spin-$\frac{1}{2}$ Heisenberg model. In fact, it has been rigorously proven to occur for $s>\frac{1}{2}$ in two dimensions and in the spin-$\frac{1}{2}$ case with weak spin-spin coupling between the two-dimensional planes.[18] Experimenters also see long-range spin order in good two-dimensional spin-$\frac{1}{2}$ antiferromagnets such as the undoped La_2CuO_4 material.[20] Apparently, a transition at some finite value of n (possibly equal to 2) separates the large-n phases with short-range order from the Néel ordered $n=2$ case. By considering finite n corrections to our large-n solutions, it may be possible to see the onset of long-range order. In this regard, the flux phase is probably a better starting point since it has gapless excitations that could lead to long-range order at finite n. Some support for this point of view comes from the method of Gutzwiller projection. Let us briefly review this method.

In our analysis of the large-n Hubbard-Heisenberg ground states, the saddle-point solutions had a trivial dependence on the Hubbard repulsion U. Indeed, for positive U (regardless of its numerical value) we found that $\phi_{\mathbf{x}}$ was constant and consequently no site-centered charge density wave formed. It is easy to understand why the Hubbard repulsion is important only at next order in the $1/n$ expansion. The saddle-point states have an average of $n/2$ electrons per site but the fluctuation in this number is of order $n^{1/2}$ since the electrons are basically free. Thus, in the $n\rightarrow\infty$ limit, the fractional number fluctuations vanish as $n^{-1/2}$ regardless of the value of U.

For the pure SU(2) Heisenberg model, $U\rightarrow\infty$ so the number fluctuations must be completely suppressed by higher-order corrections to the saddle point. Gutzwiller projection allows us to account for some of these correlations. To implement this procedure, we examine the wave functions for the electrons moving in the background χ fields and project out the parts that correspond to doubly occupied sites. The expectation value of the Heisenberg Hamiltonian for the resulting projected state can then be evaluated. Thus, Gutzwiller projection is in some sense equivalent to integrating over the $\phi_{\mathbf{x}}$ and $\chi_{\mathbf{xy}}$ fields in the functional integral (2.7) with a fixed-fermion background wave function.

The Gutzwiller projection of the completely dimerized states may be performed exactly because the electrons are completely localized on particular bonds. Consider the one-particle Hamiltonian for the dimer between sites 1 and 2: $H_1=|\chi|(c_1^{\dagger\alpha}c_{2\alpha}+\text{H.c.})$. The lowest-energy state for two electrons on this dimer is given by

$$|12\rangle=(c_1^{\dagger\downarrow}-c_2^{\dagger\downarrow})(c_1^{\dagger\uparrow}-c_2^{\dagger\uparrow})|0\rangle\,.$$

The doubly-occupied sites can be eliminated by hand, yielding the singlet state $|\psi\rangle$:

$$|\psi\rangle\equiv 2^{-1/2}P_G|12\rangle=2^{-1/2}(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle)\,.$$

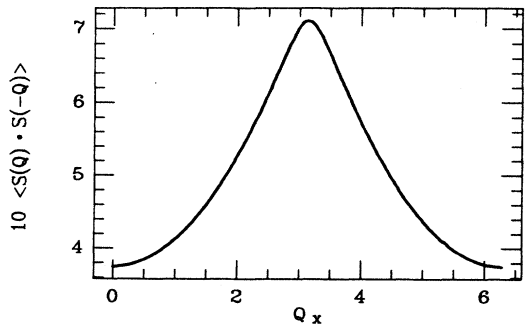


FIG. 11. The instantaneous correlator in the uniform phase for $Q_y=\pi$. The temperature is $T=0.02J$.

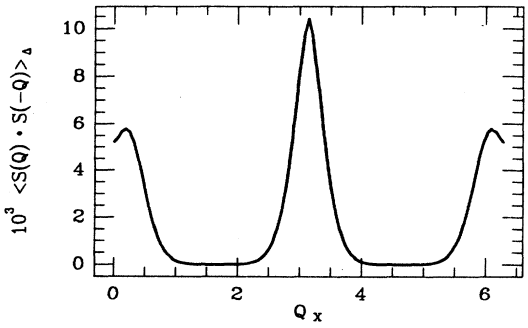


FIG. 12. Same as Fig. 10, except for the zero-temperature flux phase.

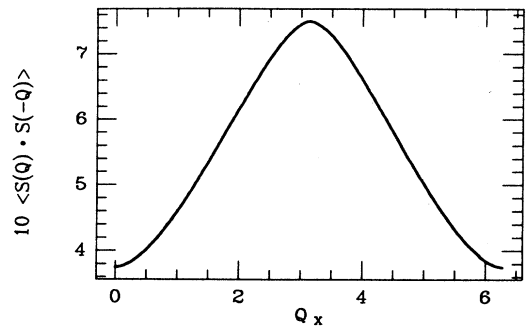


FIG. 13. Same as Fig. 11, except for the zero-temperature flux phase.

In fact, a projected dimerized state is a valence-bond solid (VBS) since the singlet bonds are frozen on the lattice. Clearly, $\langle\psi|\mathbf{S}_1\cdot\mathbf{S}_2|\psi\rangle=-\frac{3}{4}$ so each dimer contributes $-\frac{3}{4}J$ to the expectation value of the energy.

In Sec. III we found that the saddle point of the one-dimensional Heisenberg model was dimerized. Since the dimers cover only $\frac{1}{2}$ of the lattice links, the expectation value of the energy for the projected state is $E=-\frac{3}{8}J=-0.375J$. This value compares poorly with the exact Bethe-ansatz ground-state energy of $E=(\frac{1}{4}-\ln 2)J\approx -0.443\,147J$. It is interesting to also consider the uniform state (the electrons in this state are the usual tight-binding wave functions with momenta in the range $-\pi/2<k<\pi/2$). The projected uniform state must resonate, and this resonance complicates the calculation. However, resonance also lowers the energy. Gros, Joynt, and Rice[21] have calculated the energy of the projected uniform state numerically and find $E=-(0.442\,17\pm 0.000\,06)J$ which is only 0.2% higher than the exact energy. Thus, the dimerized large-n saddle point is probably a poor starting point for understanding the $n=2$ physics via a perturbative expansion in powers of $1/n$. Our renormalization-group analysis of the Heisenberg chain in Sec. III supports this idea since the $n>2$ systems had a spin gap whereas the SU(2) system remained gapless.

Turning now to the two-dimensional case, we also expect that resonance can lower the energy of the undimerized states. Consider the energies of the projected Peierls, uniform, and flux states. The Peierls dimers cover only $\frac{1}{4}$ of the lattice links, so the energy of the projected state is $E=-\frac{3}{16}J=-0.1875J$. The calculation for the undimerized states again must be done numerically. Gros and coworkers[22] have evaluated $\langle\psi P_G|\mathbf{S}_x\cdot\mathbf{S}_y|P_G\psi\rangle$ for both states and find that $E=-(0.267\pm 0.003)J$ for the uniform state and $E\approx -0.32J$ for the flux phase. Apparently, both undimerized states have lower energy than the Peierls VBS. Indeed, the energy of the projected flux phase is only about 5% higher than that of the best variational estimate (using a Néel ordered state) of the ground-state energy:[11] $E=-0.3344J$.

For the special case of $n=2$, the Heisenberg model has a larger local gauge symmetry that contains the U(1) symmetry.[23,24] This symmetry mixes particles of spin σ and holes of spin $-\sigma$. Consider the expression for the spin operator (2.1). This operator is invariant under the local transformations:

$$c_{x\uparrow}^{\dagger}\rightarrow \alpha_x c_{x\uparrow}^{\dagger}+\beta_x c_{x\downarrow},$$

$$c_{x\downarrow}\rightarrow -\beta_x^{*} c_{x\uparrow}^{\dagger}+\alpha_x^{*} c_{x\downarrow},$$

subject to the constraint $|\alpha_x|^2+|\beta_x|^2=1$. The three independent degrees of freedom of these local transformations correspond to the three generators of SU(2). [Note that the U(1) gauge symmetry appears here as the subgroup of the SU(2) symmetry defined by setting $\beta_x=0$.] A Lagrange multiplier can now be included to constrain the system to one particle per site and it transforms as the time component of a SU(2) gauge field. Finally, a Hubbard-Stratonovich transformation may be used to introduce spatial SU(2) gauge fields. This version of the Heisenberg model was used in Ref. 23 to show the equivalence of the large-n problem and the mean-field method of Baskaran, Zou, and Anderson.[16]

VII. CONCLUSION

Having solved the Hubbard-Heisenberg model in the large-n limit, we now want to comment on its relevance to the high-temperature superconductors. Only the outer d orbitals of the copper ions are considered in Anderson's simplified description of the copper-oxygen planes. The Hubbard-Hamiltonian models the hopping of electrons between different Cu atoms and the highly screened Coulomb repulsion between the electrons. The number of electrons in the CuO_2 planes of the high-T_c superconductors is controlled by the atoms above and below the planes. For example, strontium doping (x) in $La_{2-x}Sr_xCuO_4$ or oxygen addition (y) in $YBa_2Cu_3O_{6.5+y}$ lowers the number of valence electrons. Zero doping or $y=0$ corresponds to approximately one valence electron per copper atom. Curiously, a structural transition from tetragonal-to-orthorhombic phases occurs in the lanthanum-based compounds at $x\approx 0.15$ with the more symmetric phase existing at higher doping.[25] The superconducting transition temperature also hits a maximum of about 40 K at approximately the same doping. The question of whether this structural transition is related to the electronic behavior, and in particular to superconductivity, has generated a great deal of interest.

The orthorhombic-tetragonal transition could be driven by a Fermi-surface instability. Indeed, our large-n solutions do show this sort of behavior. For large doping, the uniform phase (which exhibits the full tetragonal symmetry) is stable. At small doping, two phases are candidates for the observed orthorhombic symmetry: the flux and "zigzag" kite states. However, only the flux state has the full-orthorhombic symmetry. In the "zigzag" kite state, the two sides of the unit cell with larger $|\chi|$ should have different lengths than the other two sides. However, no such charge-density waves are observed. The sides of the unit cell in the flux state *do* have the same length but it is unclear how the rhombic distortion would arise.

The spin-spin correlator provides additional contact between experiment and theory. Neutron-scattering experiments[17] in the lanthanum compounds show a narrow

$\mathbf{Q}=(\pi,\pi)$ scattering peak with a width of less than 5% of the zone size. These experiments typically involve energy exchanges of at most tens of meV. Since this energy scale is small compared to J (which is of order 1000 K or about 100 meV) we should compare these measurements to the low-energy correlators calculated in Sec. V. These correlators do have peaks at momenta (π,π) though the flux phase also has peaks at $(0,\pi)$ and $(\pi,0)$ that are not seen experimentally. However, the widths of the calculated peaks depend strongly on the cutoff Δ. Since no such sensitivity is seen in the experiments, the observed (π,π) peak probably corresponds instead to spin excitations about true long-range Néel order.[26]

Finally, the presence of superconductivity in the Hubbard-Heisenberg model cannot be addressed in the $n\rightarrow\infty$ limit. Off-diagonal long-range order (ODLRO) requires large particle-number fluctuations, but all such fluctuations are suppressed at large n. Furthermore, the only available order parameters (χ and ϕ) are electrically neutral and cannot break the U(1) gauge invariance. For finite n the conventional superconducting order parameter generalizes to the charged SU(n) singlet operator: $\epsilon_{\alpha\beta\ldots\omega}c_{\mathbf{x}}^{\dagger\alpha}c_{\mathbf{x}}^{\dagger\beta}\ldots c_{\mathbf{x}}^{\dagger\omega}$ and n tuples of particles might condense together. The search for ODLRO at finite n would make an interesting, but difficult, project.

ACKNOWLEDGMENTS

We would like to thank P. W. Anderson, D. Arovas, G. Baskaran, P. Coleman, C. Gros, T. Hsu, S. Kivelson, G. Kotliar, D. H. Lee, D. Rokhsar, S. Sachdev, J. Wheatley, P. Wiegmann, and Z. Zou for many helpful discussions. One of us (J.B.M.) was supported by the National Science Foundation and by a grant from International Business Machines. Another of us (I.A.) was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

APPENDIX: THE SADDLE-POINT CURVATURE

To compute the curvature of the saddle points, we consider the effect of small but spatially arbitrary perturbations of the bosonic fields $\chi_{\mathbf{xy}}$ on the free energy (we postpone the study of fluctuations in the $\Phi_{\mathbf{xy}}$ fields until the end). We will restrict ourselves to time-independent perturbations; the extension to dynamic fluctuations is straightforward. Also, for simplicity we only consider the case of a half-filled band ($\delta=0$). Finally, we will ignore the fluctuations in the $\phi_{\mathbf{x}}$ fields. For the pure Heisenberg model, $\phi_{\mathbf{x}}$ cannot have a nonzero saddle-point value by particle-hole symmetry; therefore $\phi_{\mathbf{x}}=0$ is always a stable saddle point. But, for nonzero t, we will simply assume that the $\phi_{\mathbf{x}}$ directions remain stable.

Suitably generalized, this calculation provides the foundation for calculating higher-order terms in the $1/n$ expansion. In fact, the $O(1)$ correction to the free energy can be obtained by integrating out the fluctuations around the saddle point in the Gaussian approximation. To this order, we only need to know the curvature (no higher moments).

We will calculate the free energy by integrating out the electrons in the background fields:

$$\exp(-\beta F[\chi])=\int[dc][dc^{\dagger}]\exp(-S[c^{\dagger},c,\chi]).$$

We first decompose the action into a purely bosonic piece and a term that includes the fermions: $S[c^{\dagger},c,\chi]=\beta F_B[\chi]+S_F[c^{\dagger},c,\chi]$. Here,

$$F_B[\chi]=n\sum_{\langle\mathbf{x},\mathbf{y}\rangle}|\chi_{\mathbf{xy}}|^2,$$

and

$$S_F(c^{\dagger},c,\chi)=\int_0^{\beta}d\tau\left[\sum_{\mathbf{x}}c_{\mathbf{x}}^{\dagger\alpha}(d/d\tau)c_{\mathbf{x}\alpha}+\sum_{\langle\mathbf{x},\mathbf{y}\rangle}(\chi_{\mathbf{xy}}c_{\mathbf{y}}^{\dagger\alpha}c_{\mathbf{x}\alpha}+\text{H.c.})\right].$$

(We set $J=1$ and take $\beta\rightarrow\infty$ at the end of the calculation.) Thus, $F[\chi]=F_B[\chi]+F_F[\chi]$ where $F_F[\chi]$ is defined by

$$\exp(-\beta F_F[\chi])=\int[dc][dc^{\dagger}]\exp[-S_F(c^{\dagger},c,\chi)].$$

To proceed, we now decompose the χ fields into a saddle-point value plus a perturbation $\chi_{\mathbf{xy}}=\bar{\chi}_{\mathbf{xy}}+\delta\chi_{\mathbf{xy}}$. Then

$$\exp(-\beta F[\chi])=\exp[-\beta F_B(\bar{\chi}+\delta\chi)]\int[dc][dc^{\dagger}]\exp\{-[S_F(c^{\dagger},c,\bar{\chi})+\delta S_F(c^{\dagger},c,\delta\chi)]\}.$$

Here we have introduced a term that couples the fermions to the perturbation

$$\delta S_F(c^{\dagger},c,\delta\chi)=\int_0^{\beta}d\tau\left(\sum_{\langle\mathbf{x},\mathbf{y}\rangle}(\delta\chi_{\mathbf{xy}}c_{\mathbf{y}}^{\dagger\alpha}c_{\mathbf{x}\alpha}+\text{H.c.})\right).$$

Upon expanding the exponential to second order, we have

$$\begin{aligned}\exp(-\beta F[\chi])&=\exp[-\beta(F_B[\bar{\chi}]+F_B[\delta\chi])]\int[dc][dc^{\dagger}]\exp[-S_F(c^{\dagger},c,\bar{\chi})]\{1+\tfrac{1}{2}(\delta S_F[c^{\dagger},c,\delta\chi])^2+O[(\delta\chi)^3]\}\\&=\exp[-\beta(F_B[\bar{\chi}]+F_B[\delta\chi])]\exp(-\beta F_F[\bar{\chi}])\{1+\tfrac{1}{2}n\beta W(\bar{\chi},\delta\chi)+O[(\delta\chi)^3]\}\\&=\exp[-\beta(F[\bar{\chi}]+F_B[\delta\chi])]\{1+\tfrac{1}{2}n\beta W(\bar{\chi},\delta\chi)+O[(\delta\chi)^3]\}.\end{aligned}$$

Note that the terms linear in $\delta\chi$ cancel out at the saddle point. By taking the logarithm of both sides of this equation we can now extract the total second-order correction

to the free energy:

$$F[\chi] = F[\bar{\chi}] + F_2[\bar{\chi}, \delta\chi] + O[(\delta\chi)^3] .$$

$$F_2[\bar{\chi}, \delta\chi] = n\left(\sum_{\langle \mathbf{x},\mathbf{y}\rangle} |\delta\chi_{\mathbf{xy}}|^2 + \tfrac{1}{2} W(\bar{\chi}, \delta\chi)\right) .$$

To actually compute $W(\bar{\chi}, \delta\chi)$ we transform to momentum space. Let us perturb the four χ_a fields ($a = 12$, 23, 34, or 41) of the $\sqrt{2}$ unit cell with fluctuations of momenta $\pm\mathbf{p}$:

$$\chi_a(\mathbf{x}) = \bar{\chi}_a + \delta\chi_a \cos(\mathbf{p}\cdot\mathbf{x}) . \tag{A1}$$

Here, $\mathbf{x}$ is on the even sublattice and the four $\chi_a(\mathbf{x})$ are attached to it as shown in Fig. 14. Defining $\psi_a(\mathbf{k}) \equiv [c_{e\alpha}(\mathbf{k}), c_{o\alpha}(\mathbf{k})]$ we can write the unperturbed action using the 2×2 matrix inverse propagator for the fermions $G^{-1}(\omega, \mathbf{k})$:

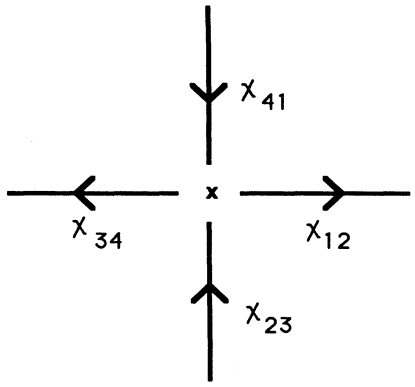


FIG. 14. The identification of the four $\chi_a(\mathbf{x})$ fields with the even site $\mathbf{x}$.

$$S_F(\psi^\dagger, \psi, \bar{\chi}) = \int [d\omega\, d^2k/(4\pi^3)]\, \psi^{\dagger\alpha}(\omega, \mathbf{k})\, G^{-1}(\omega, \mathbf{k})\psi_\alpha(\omega, \mathbf{k}) ,$$

where

$$G^{-1}(\omega, \mathbf{k}) = i\omega \mathbf{1} + \bar{\chi}(\mathbf{k})\sigma^+ + \bar{\chi}^*(\mathbf{k})\sigma^- .$$

Here

$$\bar{\chi}(\mathbf{k}) \equiv 2t[\cos(k_x) + \cos(k_y)] + \bar{\chi}_{12}\exp(ik_x) + \bar{\chi}_{34}\exp(-ik_x) + \bar{\chi}^*_{23}\exp(ik_y) + \bar{\chi}^*_{41}\exp(-ik_y) ,$$

and $\sigma^{+(-)}$ is the 2×2 matrix with only one nonzero entry ($=1$) in the upper right (lower left) corner. Likewise, δS_F can be expressed as

$$\delta S_F(\psi^\dagger, \psi, \delta\chi) = \tfrac{1}{2}\int [d\omega\, d^2k/(4\pi^3)][\psi^{\dagger\alpha}(\omega, \mathbf{k}+\mathbf{p})\delta L(\omega, \mathbf{k}; \mathbf{p})\psi_\alpha(\omega, \mathbf{k}) + \psi^{\dagger\alpha}(\omega, \mathbf{k}-\mathbf{p})\delta L(\omega, \mathbf{k}; -\mathbf{p})\psi_\alpha(\omega, \mathbf{k})]$$

where

$$\delta L(\omega, \mathbf{k}; \mathbf{p}) \equiv \delta\chi(\mathbf{k})\sigma^+ + \delta\chi^*(\mathbf{k}+\mathbf{p})\sigma^-$$

and

$$\delta\chi(\mathbf{k}) \equiv \delta\chi_{12}\exp(ik_x) + \delta\chi_{34}\exp(-ik_x) + \delta\chi^*_{23}\exp(ik_y) + \delta\chi^*_{41}\exp(-ik_y) .$$

The one-loop Feynman diagram that corresponds to the fermionic contribution to the second-order term is displayed in Fig. 15. It is given by

$$W(\bar{\chi}, \delta\chi; \mathbf{p}) = \tfrac{1}{2}\int [d\omega\, d^2k/(4\pi^3)]\, \mathrm{Tr}[G(\omega, \mathbf{k})\delta L(\omega, \mathbf{k}; -\mathbf{p})G(\omega, \mathbf{k}+\mathbf{p})\delta L(\omega, \mathbf{k}; \mathbf{p})] .$$

Fortunately, the integral over ω can be performed analytically and we have

$$W(\bar{\chi}, \delta\chi, \mathbf{p}) = \tfrac{1}{2}\int [d^2k/(2\pi^2)][\mathrm{Re}[\bar{\chi}^*(\mathbf{k})\delta\chi(\mathbf{k}-\mathbf{p})\bar{\chi}^*(\mathbf{k}+\mathbf{p})\delta\chi(\mathbf{k}+\mathbf{p})]\{|\chi(\mathbf{k})\chi(\mathbf{k}+\mathbf{p})|[|\chi(\mathbf{k})| + |\chi(\mathbf{k}+\mathbf{p})|]\}^{-1}$$
$$- |\delta\chi(\mathbf{k})|^2[|\chi(\mathbf{k})| + |\chi(\mathbf{k}+\mathbf{p})|]^{-1}) .$$

Further evaluation of $W(\bar{\chi}, \delta\chi; \mathbf{p})$ must generally be done numerically. However, for the $t = 0$ Peierls case we can do it exactly since the dispersion relation for the disconnected dimers is so simple. It is instructive to check our calculation in this way and we proceed by setting $\bar{\chi}_{12} = \bar{\chi}_{23} = \bar{\chi}_{34} = 0$ and $\bar{\chi}_{41} = \frac{1}{2}$. Then it is easy to see that

$$F_2(\bar{\chi}, \delta\chi; \mathbf{p}) = (n/2)\{|\delta\chi_{12}|^2 + |\delta\chi_{23}|^2 + |\delta\chi_{34}|^2 + 2[\mathrm{Re}(\delta\chi_{41})]^2\} .$$

This result agrees with a direct calculation of the free energy away from the saddle point. Also, note that F_2 does not depend on $\mathrm{Im}(\delta\chi_{41})$. The curvature is flat in this direction because this mode corresponds to a gauge direction that simply changes the phase of $\bar{\chi}_{41}$. Apparently, $F_2 > 0$ for all the physical modes, so the Peierls phase is locally stable. For $t > 0$, we have evaluated F_2 numerically and find that the Peierls phase remains locally stable for $t \lesssim 0.4$.

As another test, we look for flat gauge directions in the

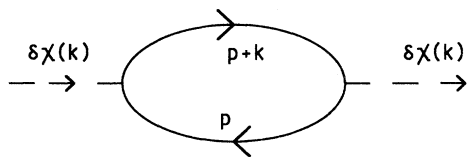


FIG. 15. Feynman diagram that contributes to the saddle-point curvature. The solid lines represent fermion propagators.

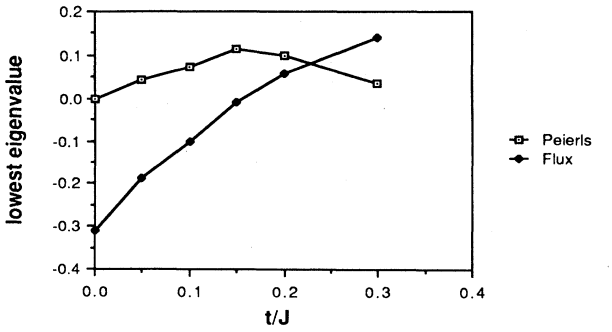


FIG. 16. The lowest eigenvalue of the saddle-point curvature vs t/J. At $t=0$, note the zero gauge mode of the Peierls state and the unstable mode of the flux state.

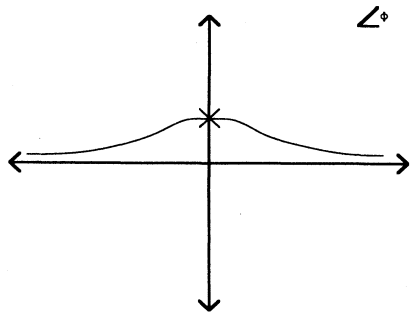


FIG. 17. The steepest-descent trajectory in the complex plane passing through the saddle point.

flux phase of the Heisenberg model. For $\mathbf{p}=\mathbf{0}$ there should be three such zero modes since there are a total of eight degrees of freedom (the real and imaginary components of the χ_a) but only five of these are physical (the magnitudes $|\chi_a|$ and the plaquette angle θ). Upon diagonalizing the 8×8 matrix corresponding to $F_2(\bar{\chi},\delta\chi;\mathbf{0})$ we do in fact find three zero modes. At nonzero $\mathbf{p}$ the situation is rather more complicated. In general, we expect only *one* gauge mode. Two modes disappear because our choice (A1) for the perturbations causes θ to fluctuate from plaquette to plaquette for these modes when $\mathbf{p}\neq\mathbf{0}$. In particular, consider the gauge in which the saddle point is $\bar{\chi}_{12}=\bar{\chi}_{23}=\bar{\chi}_{34}=|\chi|$ and $\bar{\chi}_{41}=-|\chi|$. Only the mode $\delta\chi_{12}=\delta\chi_{34}=-\delta\chi_{23}=-\delta\chi_{41}=i\epsilon$ (for ϵ infinitesimal and real) maintains $\theta=\pi$ in each plaquette. However, a special case does occur at momenta $(0,\pi)$ or $(\pi,0)$: two zero modes exist there. We have checked that the eigenvalues of F_2 exhibit these features.

The instability in the flux phase shows up at momenta $\mathbf{p}=(0,\pi)$ or $(\pi,0)$. The two negative modes are (with the above gauge choice): $\delta\chi_{12}=-\delta\chi_{34}=\epsilon$ and $\delta\chi_{23}=\delta\chi_{41}=\epsilon$. Indeed, by examining the various links it is easy to see that these modes are responsible for the decay of the flux state into the box state. The unstable modes have a nonzero width in momentum space; for example, $\mathbf{p}=(0,2.2)$ is a marginally unstable direction.

This instability can be cured in two different ways. First, it disappears for $t\gtrsim 0.15$ in the hybrid Hubbard-Heisenberg model. This transition can be seen in Fig. 16 which displays the lowest eigenvalue of F_2 as a function of t/J for both the flux and Peierls phases. (A boxlike phase may be the correct saddle point for the narrow range $0.08\lesssim t\lesssim 0.15$ since the flux state has a lower free energy than the Peierls state for $t\gtrsim 0.08$ yet is still unstable towards the box phase for $t\lesssim 0.15$.)

Alternatively, we can include the biquadratic interaction to raise the free energy of the box phase above that of the flux phase. However, the introduction of the Φ fields complicates the study of the saddle-point curvature so we proceed by integrating them out. The effective action then depends only on the χ fields and the meaning of the curvature remains unambiguous.

We recall the Lagrangian (2.12) for the biquadratic problem (we continue to set $J=1$ and $\phi_{\mathbf{x}}=0$):

$$L=\sum_{\mathbf{x}} c_{\mathbf{x}}^{\dagger\sigma}(d/d\tau)c_{\mathbf{x}\sigma} + \sum_{\langle \mathbf{x},\mathbf{y}\rangle}[(n/\tilde{J})\Phi_{\mathbf{xy}}^2+n|\chi_{\mathbf{xy}}|^2 + (1-2i\Phi_{\mathbf{xy}})^{1/2}(\chi_{\mathbf{xy}}c_{\mathbf{y}}^{\dagger\alpha}c_{\mathbf{x}\alpha}+\text{H.c.})]\,.$$

It is convenient to use the variables $\Omega_{\mathbf{xy}}$ and $\tilde{\Omega}_{\mathbf{xy}}$:

$$\Omega_{\mathbf{xy}}\equiv(1-2i\Phi_{\mathbf{xy}})^{1/2}\chi_{\mathbf{xy}}\,,$$
$$\tilde{\Omega}_{\mathbf{xy}}\equiv(1-2i\Phi_{\mathbf{xy}})^{1/2}\chi_{\mathbf{xy}}^{*}\,.$$

Note that $\Omega_{\mathbf{xy}}^{*}\neq\tilde{\Omega}_{\mathbf{xy}}$ except when $\Phi_{\mathbf{xy}}$ is purely imaginary (such as at the saddle point). We can now integrate out the fermions and find the free energy.

$$F(\chi,\Phi)=n\sum_{\langle \mathbf{x},\mathbf{y}\rangle}[(1/\tilde{J})\Phi_{\mathbf{xy}}^2+|\chi_{\mathbf{xy}}|^2]+nF_F(\Omega,\tilde{\Omega})\,.$$

Here, $F_F(\Omega,\tilde{\Omega})$ is the fermionic contribution defined above.

To proceed, we expand $F(\Omega,\tilde{\Omega})$ around the saddle point to second order in the perturbations $\delta\chi_{\mathbf{xy}}$ and $\delta\Phi_{\mathbf{xy}}$ using the definitions for the Ω fields and the function $W(\bar{\Omega},\delta\Omega;\mathbf{p})$. The Gaussian integral over the $\delta\Phi$ fluctuations can now be performed and the path follows the steepest descent trajectory through the imaginary saddle point depicted in Fig. 17. (We can continue the integration contour off the real axis because the integrand contains no poles. The boundary conditions $\Phi\rightarrow\pm\infty$ along the real axis make the functional integral convergent.) Note that the Gaussian approximation becomes exact at $n\rightarrow\infty$. The crucial point is that this integral renormalizes the curvature in the χ directions. After a rather tedious but straightforward calculation, we find that the χ modes are stabilized when $\tilde{J}\gtrsim 1.45$—i.e., $\tilde{J}$ is sufficiently large that the flux phase has lower free energy than the box phase. Indeed, all the χ directions now have positive curvature, so we conclude that the flux phase is at least locally stable.

[1]P. W. Anderson, Science **235**, 1196 (1987).
[2]I. Affleck and J. B. Marston, Phys. Rev. B **37**, 3774 (1988).
[3]D. P. Arovas and A. Auerbach, Phys. Rev. B **38**, 316 (1988); Phys. Rev. Lett. **61**, 617 (1988).
[4]N. Read and S. Sachdev (unpublished).
[5]T. Dombre and G. Kotliar, Phys. Rev. B **38**, 855 (1989).
[6]I. Affleck, Phys. Rev. Lett. **54**, 966 (1985).
[7]I. Affleck and E. Lieb, Lett. Math. Phys. **12**, 57 (1986).
[8]I. Affleck, Phys. Rev. B **37**, 5186 (1988).
[9]A. M. Perelomov and V. M. Popov, Pis'ma Zh. Eksp. Teor. Fiz. **1**, 15 (1965) [JETP Lett. **1**, 160 (1965)].
[10]E. Lieb and D. Mattis, J. Math. Phys. **3**, 749 (1962).
[11]S. Liang, B. Doucot, and P. W. Anderson, Phys. Rev. Lett. **61**, 365 (1988).
[12]E. Lieb and F. Wu, Phys. Rev. Lett. **20**, 1445 (1968).
[13]I. Affleck and J. Brad Marston, J. Phys. C **21**, 2511 (1988).
[14]R. Jullien and F. D. M. Haldane, Bull. Am. Phys. Soc. **28**, 34 (1983); I. Affleck, D. Gepner, H. Schulz, and T. Ziman, J. Phys. A **22**, 511 (1989).
[15]J. B. Marston, Phys. Rev. Lett. **61**, 1914 (1988).
[16]G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).
[17]G. Shirane *et al.*, Phys. Rev. Lett. **59**, 1613 (1987); Y. Endoh *et al.*, Phys. Rev. B **37**, 7443 (1988).
[18]T. Kennedy, E. H. Lieb, and B. S. Shastry, J. Stat. Phys. **53**, 1019 (1988); E. Jordao Neves and J. Fernando Peres, Phys. Lett. **114A**, 331 (1986).
[19]J. Oitmaa and D. D. Betts, Can. J. Phys. **56**, 897 (1978); E. Manousakis and R. Salvador, Phys. Rev. Lett. **60**, 840 (1988).
[20]D. Vaknin *et al.*, Phys. Rev. Lett. **58**, 2802 (1987).
[21]C. Gros, R. Joynt, and T. M. Rice, Phys. Rev. B **36**, 381 (1987).
[22]C. Gros, Phys. Rev. B **38**, 931 (1988); F. C. Zhang *et al.*, Supercond. Sci. Technol. **1**, 36 (1988).
[23]I. Affleck, Z. Zou, T. Hsu, and P. W. Anderson, Phys. Rev. B **38**, 745 (1988).
[24]E. Fradkin and A. Moreo, Phys. Rev. B **38**, 2926 (1988).
[25]J. D. Jorgensen *et al.*, Phys. Rev. Lett. **58**, 1024 (1987).
[26]S. Chakravarty, B. I. Halperin, and D. R. Nelson, Phys. Rev. Lett. **60**, 1057 (1988).