Renormalized perturbation theory of magnetic instabilities in the two-dimensional Hubbard model at small doping

Andrey V. Chubukov
P. L. Kapitza Institute for Physical Problems, Academy of Sciences, 117334, ul. Kosygina, 2, Moscow, Russia

David M. Frenkel
Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801
(Received 5 March 1992)

We consider magnetic instabilities in the two-dimensional Hubbard model at small doping. We find that the renormalization of the effective interaction prevents an immediate instability of a commensurate antiferromagnetic state upon doping. At increased doping levels, our calculations indicate the occurrence of an instability in the channel of transverse spin fluctuations. This instability is known to lead to a spiral magnetic phase. We also consider a dynamical spin susceptibility near the instability and find that conventional spin waves play no role in the transition. Instead, the incommensurate instability is governed by collective fermionic excitations coupled to the spin background.

I. INTRODUCTION

The discovery of high-T_c superconductivity initiated a search for a theoretical model that would be simple yet hold a promise to describe the crucial features of the new superconductors. Many scientists working in this field believe that a simple Hubbard model already contains at least some of the important physics governing the properties of high-T_c materials.[1] This model appears to be particularly relevant to the description of the interesting magnetic properties of cuprate oxides.

The Hubbard model is described by the Hamiltonian

$$H=-t\sum_{\substack{\langle ij\rangle\\ \alpha}}(c_{i\alpha}^{\dagger}c_{j\alpha}+c_{j\alpha}^{\dagger}c_{i\alpha})+U\sum_{i}n_{i\uparrow}n_{i\downarrow}\ , \tag{1}$$

where t is the nearest-neighbor hopping and U mimics the effects of the Coulomb repulsion. The zero-temperature properties of this model are characterized by a single dimensionless ratio $8t/U$, which compares the interaction strength U with the bandwidth. From the experimental point of view, the most relevant values of $8t/U$ are of order unity. Theoretically, it is possible to perform expansions in the limiting cases of $t \ll U$ and $t \gg U$.

Traditionally, the large- and small-U limits were studied by completely different methods. For small U, the customary tool is the random phase approximation[2] (RPA), while in the strong-coupling limit, one would perform a canonical transformation reducing the Hubbard model to the so-called t-J model.[3–6] This model includes the Heisenberg exchange interaction as well as the nearest-neighbor hopping. A strong nonlinear feature of the t-J model is the constraint which forbids a double occupancy of a single site.

In spite of such a different methodology, it was shown by several authors that for many physical properties there is a smooth crossover between the two limits.[2,7] In view of this one may expect to describe the experimental situation by starting either with the RPA or the t-J model approach.

In this paper, we choose the RPA description, renormalized—when necessary—to account for strong-coupling effects. This will allow us to compare the situation at small and large U. We note that the comparison would not be possible if we started with the t-J model, which in principle cannot apply to the weak-coupling case because of the no double occupancy constraint.

At half filling, the Hubbard model, given by Eq. (1), has an antiferromagnetically (AFM) ordered ground state.[8,9] The magnetic ordering leads to a gap in the electronic spectrum and thus at half filling the system is an insulator. The goal of this paper is to study what happens with the AFM ordering away from half filling. This question was studied by many authors, mainly in the mean-field approximation.[10–20] It has been argued that in two dimensions, immediately away from half filling the commensurate AFM ordering becomes unstable.

For small U, it was suggested by Schulz[13] that a small density of holes doped into the AFM state discommensurates it by forming domain walls. When the doping increases, the magnetic structure gradually transforms into a linearly polarized incommensurate spin-density wave (SDW). In a mean-field approach, the transition to this type of incommensurate state is driven by an instability in the longitudinal spin susceptibility. For large U, the analysis performed by Shraiman and Siggia pointed to an immediate instability upon doping in the transverse spin susceptibility.[10] This leads to an incommensurate spin structure, which is usually referred to as a spiral phase. An alternative suggestion for the possible low-doping phase was put forward by Schrieffer, Wen, and Zhang (SWZ).[2] Based on the idea of "spin bags," they argued that the commensurate antiferromagnetism survives at sufficiently small density of holes.[2]

In this paper we will study the question of the instabili-

ties outside the limits of the mean-field description. We will show that the vacuum renormalization of the effective interaction suppresses the instability of the commensurate AFM state in two dimensions for sufficiently small deviations from half filling. For large doping concentrations the incommensurate instability may occur; however, its possibility in both the large- and small-U limits is not governed by a large parameter, but rather depends on numerical factors.

The vacuum renormalization comes from a summation of a ladder sequence of diagrams for the interaction between two particles. It converts the interaction potential between the holes into a scattering amplitude, which is the proper quantity to describe the effects of interactions in a dilute Fermi gas.[21] In two dimensions, the scattering amplitude is known to vanish logarithmically in the low-energy limit. Consequently, any effect of a small density of holes will be weakened by this reduction of the effective coupling.

We will also study the dynamics of the instability. Specifically, we will show that the spin-wave excitations do not take part in the transition, in contrast to what one would normally expect based on the knowledge of the incommensurate transitions in purely magnetic systems.[22–25] The transition in the doped Hubbard model is instead driven by a collective fermionic mode coupled to the spin background.[10]

The paper is organized as follows. In Sec. II, we review a standard mean-field approach to the two-dimensional (2D) Hubbard model at half filling. We also show how one can obtain analytically the first quantum corrections to the sublattice magnetization and some other magnetic properties in the Heisenberg antiferromagnet starting from the large-U limit of the Hubbard model. In Sec. III, we will perform the RPA study of the transverse and longitudinal magnetic susceptibilities away from half filling. Section IV is devoted to the analysis of how the mean-field theory should be modified when the vertex and self-energy corrections are taken into account. Here we also consider the effects of vacuum renormalization in 2D and the situation at finite doping. The dynamics of the transition is studied in Sec. V. Finally, Sec. VI is devoted to the conclusions.

II. HUBBARD MODEL AT HALF FILLING

We start by reviewing a standard mean-field approach to the two-dimensional Hubbard model at half filling.[2,8,9] This section also fixes our notation, which we use throughout this paper. We start with the Hubbard model Hamiltonian with nearest-neighbor hopping. In momentum space, the Hamiltonian of Eq. (1) is

$$H=\sum_{k\sigma}\epsilon_k a^\dagger_{k\sigma}a_{k\sigma} + \frac{U}{2}\sum_{k_i\sigma_i} a^\dagger_{k_1\sigma_1}a^\dagger_{k_2\sigma_2}a_{k_3\sigma_3}a_{k_4\sigma_4}\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}\,, \tag{2}$$

where $\epsilon_k=-2t(\cos k_x+\cos k_y)$. It is well known that at half filling, the 2D Hubbard model has a commensurate antiferromagnetic ground state. A way to see this is to introduce spin-density operators,

$$\vec{S}(q)=\frac{1}{2}\sum_{k\alpha\beta} a^\dagger_{k+q,\alpha}\vec{\sigma}_{\alpha\beta}a_{k\beta}\,, \tag{3}$$

where $\vec{\sigma}$ are the usual Pauli matrices. Antiferromagnetism at $q=(\pi,\pi)$ can be described by a finite expectation value of

$$S_z(q=Q)=\frac{1}{2}\sum_k (a^\dagger_{k+Q,\uparrow}a_{k\uparrow}-a^\dagger_{k+Q,\downarrow}a_{k\downarrow})\,, \tag{4}$$

where we defined $Q=(\pi,\pi)$. If, as we suppose, only a spin-density wave is present in the system, then

$$\left\langle\sum_k a^\dagger_{k+Q,\uparrow}a_{k\uparrow}\right\rangle=-\left\langle\sum_k a^\dagger_{k+Q,\downarrow}a_{k\downarrow}\right\rangle=\alpha\,. \tag{5}$$

Hence, $\langle S_z\rangle=\alpha$.

In a mean-field (MF) approach one uses Eq. (5) to decouple the quartic term. After decoupling, the quadratic Hamiltonian takes a form

$$H_{\rm MF}=\sum_{k\sigma}{}' \epsilon_k(a^\dagger_{k\sigma}a_{k\sigma}-a^\dagger_{k+Q\sigma}a_{k+Q\sigma}) - U\alpha\sum_{k\sigma}{}' \mathrm{sgn}(\sigma)(a^\dagger_{k\sigma}a_{k+Q,\sigma}+a^\dagger_{k+Q,\sigma}a_{k\sigma})\,. \tag{6}$$

Primes to the summation signs indicate that the sums are over the reduced Brillouin zone. This effective quadratic Hamiltonian can be diagonalized by the Bogolyubov transformation

$$\begin{aligned} a_{k\sigma}&=u_k c_{k\sigma}+v_k d_{k\sigma}\,,\\ a_{k+Q,\sigma}&=\mathrm{sgn}(\sigma)(-v_k c_{k\sigma}+u_k d_{k\sigma})\,, \end{aligned} \tag{7}$$

where

$$\begin{aligned} u_k&=\left[\frac{1}{2}\left[1+\frac{\epsilon_k}{E_k}\right]\right]^{1/2},\\ v_k&=\left[\frac{1}{2}\left[1-\frac{\epsilon_k}{E_k}\right]\right]^{1/2}, \end{aligned} \tag{8}$$

and $E_k=(\epsilon_k^2+\Delta^2)^{1/2}$ with $\Delta=U\alpha$ is the energy of the quasiparticles. The diagonalization reduces Eq. (7) to

$$H=\sum_{k\sigma}{}' E_k(c^\dagger_{k\sigma}c_{k\sigma}-d^\dagger_{k\sigma}d_{k\sigma})\,. \tag{9}$$

Throughout the paper we will refer to $c_{k\sigma}$ and $d_{k\sigma}$ as the conduction and valence band quasiparticles, respectively.

The self-consistency condition of the diagonalization procedure requires

$$\langle S_z(Q)\rangle=\sum_k{}' u_k v_k\,, \tag{10}$$

which leads to the "gap equation"

$$\frac{1}{U}=\sum_k{}'\frac{1}{E_k}\,. \tag{11}$$

At large U we have a conventional Mott-Hubbard insulator and $\Delta\simeq U/2$. Correspondingly $\langle S_z\rangle\simeq\frac{1}{2}$. This agrees with a known fact that the Hubbard model at half

filling reduces to the nearest-neighbor Heisenberg model with $J=4t^2/U$ in the large-U limit. At small U, a nonzero solution of the gap equation is possible due to Van Hove singularities as well as the nesting properties of the Fermi surface for noninteracting particles at half filling. The combination of the two effects leads to a doubly logarithmic singularity[26] in the right-hand side of Eq. (11) at small Δ. In explicit form

$$\Delta \simeq t e^{-2\pi\sqrt{t/U}} . \tag{12}$$

The antiferromagnetic ordering also implies the existence of the low-energy bosonic excitations, namely, the transverse spin waves, as well as the longitudinal spin fluctuations. The spin waves are gapless by the Goldstone theorem since they reflect a spontaneously broken spin rotation invariance. These bosonic degrees of freedom can be analyzed by studying the dynamical spin susceptibility,

$$\bar{\chi}^{ij}(q,q';\omega)=\int dt \left[\frac{i}{2N} \langle TS_q^i(t)S^j_{-q'}(0)\rangle \right] e^{i\omega t} . \tag{13}$$

In the RPA, the total susceptibility $\bar{\chi}^{ij}(q,q';\omega)$ is given by a sum of "bubble" diagrams (Fig. 1). Each bubble stands for a "noninteracting" susceptibility evaluated over the mean-field ground state.

The antiferromagnetic ordering at $q=(\pi,\pi)$ doubles the unit cell and makes the RPA calculations somewhat peculiar because $\bar{\chi}^{ij}(q,q';\omega)$ is generally nonzero either when $q=q'$ or when $q=q'+Q$. Thus the RPA summation is a 2×2 matrix problem in this case. It was solved in Ref. 2. For the transverse susceptibility, the result is

$$\bar{\chi}^{+-}(q,q;\omega)=\frac{\chi_0^{+-}(q,\omega)[1-U\chi_0^{+-}(q+Q,\omega)]+U[\chi_Q^{+-}(q,\omega)]^2}{[1-U\chi_0^{+-}(q,\omega)][1-U\chi_0^{+-}(q+Q,\omega)]-U^2[\chi_Q^{+-}(q,\omega)]^2} , \tag{14}$$

$$\bar{\chi}^{+-}(q,q+Q;\omega)=\frac{\chi_Q^{+-}(q,\omega)}{[1-U\chi_0^{+-}(q,\omega)][1-U\chi_0^{+-}(q+Q,\omega)]-U^2[\chi_Q^{+-}(q,\omega)]^2} . \tag{15}$$

In these equations, $\chi_0^{+-}(q,\omega)$ and $\chi_Q^{+-}(q,\omega)$ are the mean-field susceptibilities (single bubbles) with the momentum transfer $q-q'=0$ and Q, respectively. Explicitly,

$$\chi_0^{+-}(q,\omega)=\frac{1}{2N}\sum_k{}' \left[1-\frac{\epsilon_k\epsilon_{k+q}-\Delta^2}{E_kE_{k+q}}\right]\left[\frac{1}{E_k+E_{k+q}-\omega}+\frac{1}{E_k+E_{k+q}+\omega}\right] , \tag{16}$$

$$\chi_Q^{+-}(q,\omega)=\frac{1}{2N}\sum_k{}' \frac{\Delta(E_k+E_{k+q})}{E_kE_{k+q}}\left[\frac{1}{E_k+E_{k+q}-\omega}-\frac{1}{E_k+E_{k+q}+\omega}\right] . \tag{17}$$

Note that $\chi_Q(q,\omega)\sim\omega$ and thus contributes only to the dynamical part of susceptibility.

In agreement with the general requirements, the poles of $\bar{\chi}^{+-}(q,q',\omega)$ define gapless bosonic excitations. The vanishing of the spin-wave gap at $q=0$ or Q follows from the condition

$$1-U\chi_0^{+-}(q=Q,\omega=0)=0 , \tag{18}$$

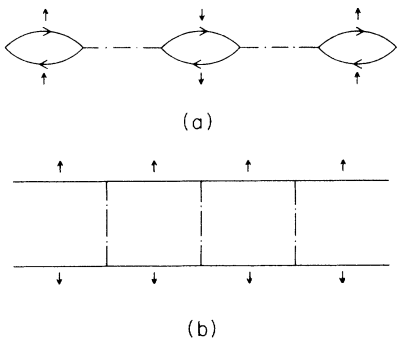

FIG. 1. Representative members of the RPA series for (a) the longitudinal channel susceptibilities $\bar{\chi}^{zz}$ and $\bar{\chi}^{\rho\rho}$ and (b) the transverse susceptibility $\bar{\chi}^{+-}$. The dash-dotted lines represent the Hubbard U interaction. The electronic Green's functions are evaluated in the spin-density-wave mean-field ground state.

which, as observed in Ref. 2, coincides with the gap equation [cf. Eq. (11)]. This provides a check on the validity of the RPA approach.

To find the spin-wave velocity c, one expands the denominator of Eqs. (14) and (15) around $q,\omega=0$ up to quadratic order. This procedure leads to

$$c^2=\frac{4t^2\left[\sum_k{}' \frac{\epsilon_k^2}{E_k^3}\right]\left[\sum_k{}' \frac{\sin^2 k_x}{E_k^3}\right]}{\left[\sum_k{}' \frac{\epsilon_k^2}{E_k^3}\right]\left[\sum_k{}' \frac{1}{E_k^3}\right]+\Delta^2\left[\sum_k{}' \frac{1}{E_k^3}\right]^2} . \tag{19}$$

For large U this expression reduces to $c^2\simeq 2J^2$, where $J=4t^2/U$. This agrees with the spin-wave solution of the corresponding Heisenberg model. Moreover, in this limit one can calculate the spectrum for all q and also find an agreement with the spin-wave calculations. Indeed, the solution for the spectrum is

$$\omega_q=2J\sqrt{1-\gamma_q^2} , \tag{20}$$

where

$$\gamma_q=\tfrac{1}{2}(\cos k_x+\cos k_y) . \tag{21}$$

In the small-U limit the second term in the denominator

of Eq. (19) is small by a factor of $1/\ln(t/\Delta)\sim(U/t)^{1/2}$ compared to the first one, and the spin-wave velocity is

$$c^2\simeq\left[\frac{4}{\pi}\right]t^2\left[\frac{U}{t}\right]^{1/2}. \tag{22}$$

This expression differs from $c^2\sim t^2$ obtained in Ref. 2. Such an answer would follow from Eq. (19) if one retained the second term in the denominator instead of the first one.

The fact that one can neglect the second term in the denominator of Eq. (19) in any limit seems somewhat peculiar. This term comes from $\chi_Q^{+-}(q,\omega)$, which links pairs of canonically conjugate variables, say $S_x(q)$ and $S_y(q+Q)$. Coupled vibrations of these two variables are known to produce spin waves in Heisenberg antiferromagnets.[27,28] However, we have checked that Eq. (19) exactly follows from the hydrodynamic theory of spin waves.[27,28] According to hydrodynamics, the spin-wave velocity is given by

$$c^2=\rho_s/\chi_\perp, \tag{23}$$

where $\chi_\perp$ is the transverse susceptibility

$$\chi_\perp=\chi^{xx}(q=0,\omega=0) \tag{24}$$

and ρ_s is the spin stiffness defined through

$$\chi^{xx}(q\simeq Q,\omega=0)=\frac{(N_0)^2}{\rho_s(q-Q)^2}. \tag{25}$$

Here, N_0 is the antiferromagnetic order parameter

$$N_0=\langle S_z(q=Q)\rangle=\Delta/U. \tag{26}$$

The evaluation of the transverse susceptibility near $q=0$ and $q=Q$ leads to exactly the same expression for c^2 as the one found from dynamic susceptibility [cf. Eq. (19)] for all values of U. We also note that in the large-U limit the static transverse susceptibility $\chi_\perp\simeq 1/8J$, which coincides with the mean-field expression for the Heisenberg model.

Actually, the RPA approach to the Hubbard model in the large-U limit allows one not only to find an agreement with the zeroth-order spin-wave theory for the Heisenberg model, but also correctly reproduces the first quantum corrections. This was demonstrated numerically by SWZ for the order-parameter renormalization. One can show this analytically as well. Indeed, in the large-U limit $\langle S_z\rangle$ is directly related to the density of valence band quasiparticles

$$\langle S_z\rangle\simeq\frac{2}{N}\sum_k{}'\langle d_{k\uparrow}^\dagger d_{k\uparrow}\rangle-\frac{1}{2}. \tag{27}$$

The deviation of $\langle S_z\rangle$ from the mean-field value of $\frac{1}{2}$ has to do with the fact that the quasiparticles interact with each other through collective modes. The lowest-order diagram that contributes to the self-energy of valence quasiparticles is shown in Fig. 2. The calculation of this diagram requires a knowledge of the coupling between the fermions and the spin waves, as well as an expression for the susceptibility for all frequencies and momenta. In the large-U limit, $\bar\chi^{+-}(q,\omega)$ becomes [cf. Eqs. (14)–(17)],

$$\bar\chi^{+-}(q,q;\omega)=\frac{1}{U}\frac{1}{2\omega_q}[\omega^2+8t^2(1-\gamma_q)]\times\left[\frac{1}{\omega+\omega_q-i\delta}-\frac{1}{\omega-\omega_q+i\delta}\right], \tag{28}$$

$$\bar\chi^{+-}(q,q+Q;\omega)=\frac{\omega}{2\omega_q}\left[\frac{1}{\omega+\omega_q-i\delta}-\frac{1}{\omega-\omega_q+i\delta}\right], \tag{29}$$

where

$$\omega_q=2J\sqrt{1-\gamma_q^2}. \tag{30}$$

To obtain the interaction vertex, one should reexpress the Hubbard interaction in terms of the valence and conduction band operators. After this procedure, the constant U gets multiplied by the coherence factors (see Ref. 2 and Sec. IV below). However, for the interaction vertex relevant to Fig. 1, this combination becomes unity in the large-U limit. As a result, the expression for the self-energy takes a simple form,

$$\Sigma(q,\Omega)=-i\frac{U^2}{N}\sum_k{}'\int\frac{d\omega}{2\pi}\frac{1}{\omega+\Omega-E_{k+q}+i\delta}[\bar\chi^{+-}(k,k;\omega)+\bar\chi^{+-}(k+Q,k+Q;\omega)+2\bar\chi^{+-}(k,k+Q;\omega)]. \tag{31}$$

The reason for having more than one term in the brackets is that the integration over internal momenta should be performed over the reduced Brillouin zone.

In the large-U limit, $E_{k+q}\simeq E_k\simeq\Delta\gg\omega_k$. With this condition, the integration is elementary and the result is

$$\Sigma(q,\Omega)=\frac{U^2}{\Omega-\Delta}\left[\frac{1}{N}\sum_k{}'Q_k\right], \tag{32}$$

where

$$Q_k=\frac{1-\sqrt{1-\gamma_k^2}}{\sqrt{1-\gamma_k^2}}. \tag{33}$$

This self-energy produces a correction to the quasiparticle spectrum, but also results in a wave-function renormalization factor Z being different from unity:

$$Z=1-\sum_p{}'Q_p. \tag{34}$$

This Z factor is directly related to the expression for the quasiparticle density and hence to $\langle S_z\rangle$,

$$\langle S_z\rangle=Z-\frac{1}{2}=\frac{1}{2}\left[1-\frac{2}{N}\sum_k{}'\frac{1-\sqrt{1-\gamma_k^2}}{\sqrt{1-\gamma_k^2}}\right], \tag{35}$$

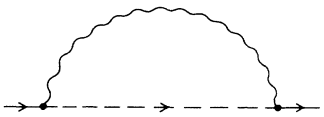

FIG. 2. The self-energy correction to the valence electrons of the spin-density-wave (SDW) state. The solid lines represent the valence electrons, the dashed line corresponds to conduction electrons, and the wavy line is the spin-wave excitation of the SDW state.

which is the usual answer known from the spin-wave theory.[29] A similar expression for $\langle S_z \rangle$ was obtained in Ref. 16. However, the authors of Ref. 16 related the correction to $\langle S_z \rangle$ to the renormalization of the gap in the quasiparticle spectrum. We do not believe this is a proper way of calculating $\langle S_z \rangle$.

That Eq. (35) and the spin-wave result should agree is not obvious. Actually, in the conventional spin-wave theory one performs an expansion in the inverse powers of the spin $(1/S)$. This expansion is inappropriate for the Hubbard model which describes electrons with $S=\frac{1}{2}$. However, there exists another way to construct a perturbation theory for magnetic systems. One can suppose that the underlying lattice has a large coordination number and perform a perturbation expansion in the inverse powers of z, where z is a number of nearest neighbors. The zeroth order in this expansion corresponds to a mean-field solution of the problem.

The expansion in powers of $1/z$ can be applied to the Hubbard model as well. Indeed, the RPA solution of Eq. (2) is essentially the mean-field result, while the expression for the self-energy [Eq. (32)] contains a factor $1/z$ due to $\gamma_k \sim 1/z$ and thus represents the first correction in the expansion in the inverse coordination number.

Strictly speaking, when performing a $1/z$ expansion, one should expand Eq. (33) in powers of γ_k^2, and retain the leading-order term. However, it is known[30] that in Heisenberg systems the formal expression for the one-loop correction to the mean-field result [i.e., Eq. (33)] exactly coincides with the leading term in the spin-wave expansion taken for $S=\frac{1}{2}$. In view of this, it is not surprising that the correction to the RPA theory for the Hubbard model coincides with the spin-wave calculations for the Heisenberg model.

In fact, with a little more effort one can also reproduce the spin-wave results for the static susceptibility and spin-wave velocity. To do this, one should retain terms of order J in the self-energy. This in turn requires the coherence factors (which we had previously set equal to unity) to be incorporated into Eq. (31).

To leading order in $1/z$, the single-particle spectrum preserves its form, but Δ and t change to $\bar{\Delta}$ and $\bar{t}$, where

$$\bar{\Delta}=\Delta/Z^2, \quad \bar{t}=t\left[1+\frac{1}{N}\sum_k{}' \frac{\gamma_k^2}{\sqrt{1-\gamma_k^2}}\right], \tag{36}$$

where Z is given by Eq. (34). After this, a simple inspection of Eq. (14) with $\chi_0^{\pm}$ calculated with the renormalized Green's functions shows that corrections to the static susceptibility come solely from the renormalization of t:

$$\chi_\perp=\frac{1}{8J}\left[\frac{t}{\bar{t}}\right]^2 \simeq \frac{1}{8J}\left[1-\frac{2}{N}\sum_k{}' \frac{\gamma_k^2}{\sqrt{1-\gamma_k^2}}\right], \tag{37}$$

which coincides with the spin-wave result.

The susceptibility near $q=Q$ is renormalized in the same way and with the use of Eqs. (23)–(26) one obtains for the spin-wave velocity

$$c=\sqrt{2}J\left[1+\frac{2}{N}\sum_k{}'(1-\sqrt{1-\gamma_k^2})\right], \tag{38}$$

which also coincides with the known expression from the spin-wave theory.

Finally, we consider charge and longitudinal spin-density fluctuations. In the RPA theory, the corresponding susceptibilities are represented by sequences of particle-hole bubbles with parallel spins. At half filling, both the $\bar{\chi}^{zz}$ and $\bar{\chi}^{\rho\rho}$ are given by simple expressions, familiar from the paramagnon theory

$$\bar{\chi}^{zz}(q,q,\omega)=\frac{\chi_l(q,\omega)}{1-U\chi_l(q,\omega)} \tag{39}$$

and

$$\bar{\chi}^{\rho\rho}(q,q,\omega)=\frac{\chi_l(q,\omega)}{1+U\chi_l(q,\omega)} . \tag{40}$$

Here $\chi_l(q,\omega)$ is a particle-hole bubble which for parallel spins is given by

$$\chi_l(q,\omega)=\frac{1}{2N}\sum_k{}'\left[1-\frac{\epsilon_k\epsilon_{k+q}+\Delta^2}{E_kE_{k+q}}\right]\times\left[\frac{1}{E_k+E_{k+q}-\omega}+\frac{1}{E_k+E_{k+q}+\omega}\right]. \tag{41}$$

Similarly to the transverse case, the longitudinal spin susceptibility $\bar{\chi}^{zz}(q,q,\omega)$ is peaked at $q=Q$. However, unlike $\bar{\chi}^{+-}(q,q,\omega)$, it does not contain any gapless bosonic modes.

Note that Eqs. (39) and (40) contain neither a mixing between charge and spin fluctuations nor a contribution from bubbles with the momentum transfer $q'-q=Q$. It is generally true for symmetry reasons that $\bar{\chi}^{zz}(q,q+Q,\omega)$ and $\bar{\chi}^{\rho\rho}(q,q+Q,\omega)$ are identically zero. However, the absence of mixing between the spin and charge channels is not a general property and is true only at half filling. Indeed, the symmetry of the antiferromagnetic ordering implies that the ground state is invariant under the combination of translation by one lattice period and rotation by π with respect to, say, the $\hat{x}$ axis. Under this operation

$$S_z(q)\rightarrow -e^{iqa}S_z(q) , \tag{42}$$

while

$$\rho(q') \rightarrow + e^{iq'a}\rho(q') \ . \tag{43}$$

These transformation properties imply that $\bar{\chi}^{zz}(q,q',\omega)$ and $\bar{\chi}^{\rho\rho}(q,q',\omega)$ are nonzero only for a momentum transfer $q-q'=0$, but they do allow a nonzero value for $\bar{\chi}^{\rho z}(q,q+Q,\omega)$. At half filling, $\bar{\chi}^{\rho z}$ turns out to be zero, which allows us to decouple the charge and spin channels in this case. However, away from half filling $\bar{\chi}^{\rho z}$ acquires a finite value and plays an important role in the study of the incommensurate spin-density-wave instabilities in the doped case.

We are now ready to study the system at small but finite concentration of holes. We first focus on the mean-field solution and later discuss how it should be modified in the case of a strong on-site repulsion, when the effective interaction between the holes is expected to undergo a significant renormalization.

III. MEAN-FIELD THEORY AT FINITE DOPING

Just as in the case of half filling, we first consider the transverse susceptibility and then the longitudinal one. Obviously, an arbitrarily small density of holes forces the chemical potential to lie within the valence band. Away from half filling, the RPA expression which links the total susceptibility $\bar{\chi}^{+-}(q,q',\omega)$ with the bare ones [see Eqs. (14) and (15)] remains the same. However, there are now two types of bubbles which define the bare susceptibilities $\chi_0^{+-}(q,\omega)$ and $\chi_Q^{+-}(q,\omega)$. The bubbles of the first type have one fermion in the valence and one in the conduction band, just as at half filling. The new bubbles have both fermions in the valence band. Accordingly, each of the expressions for $\chi_0^{+-}(q,\omega)$ and $\chi_Q^{+-}(q,\omega)$ now have two contributions,

$$\chi_0^{+-}(q,\omega)=\frac{1}{2N}\sum_{E_{k+q}>|\mu|}{}' \left[1-\frac{\epsilon_k\epsilon_{k+q}-\Delta^2}{E_kE_{k+q}}\right]\left[\frac{1}{E_k+E_{k+q}-\omega}+\frac{1}{E_k+E_{k+q}+\omega}\right]$$
$$+\frac{1}{2N}\sum_{\substack{E_{k+q}>|\mu| \\ E_k<|\mu|}}{}' \left[1+\frac{\epsilon_k\epsilon_{k+q}-\Delta^2}{E_kE_{k+q}}\right]\left[\frac{1}{E_{k+q}-E_k-\omega}-\frac{1}{E_k-E_{k+q}-\omega}\right] \tag{44}$$

and

$$\chi_Q^{+-}(q,\omega)=\frac{1}{2N}\sum_{E_{k+q}>|\mu|}{}' \frac{\Delta}{E_kE_{k+q}}\left[\frac{1}{E_{k+q}+E_k-\omega}-\frac{1}{E_{k+q}+E_k+\omega}\right](E_k+E_{k+q})$$
$$-\frac{1}{2N}\sum_{\substack{E_{k+q}>|\mu| \\ E_k<|\mu|}}{}' \frac{\Delta}{E_kE_{k+q}}\left[\frac{1}{E_{k+q}-E_k-\omega}-\frac{1}{E_{k+q}-E_k+\omega}\right](E_{k+q}-E_k) \ , \tag{45}$$

where μ (which is negative) is the chemical potential. The finite density of holes also modifies the self-consistency condition of the mean-field solution, which should now involve only the filled electronic levels in the valence band,

$$\sum_{E_k>|\mu|}{}' \frac{1}{E_k}=\frac{1}{U} \ . \tag{46}$$

For small U, this reduces the gap to $\bar{\Delta}=\Delta_0(1-2\epsilon_F/\Delta_0)^{1/2}$, where $\epsilon_F^2=\mu^2-\bar{\Delta}^2$ and Δ_0 is the gap value at half filling. It is easy to check that the self-consistency condition still ensures the divergence of the static susceptibility at $q\rightarrow Q$. However, the nonzero doping significantly modifies the spin stiffness. To study this effect, we first focus on the expansion of the static part of the total susceptibility near $q=Q$. In this limit, χ_Q^{+-} disappears, and the denominator of $\bar{\chi}^{+-}(q,q',\omega=0)$ is given by

$$1-U\chi_0^{+-}(q,\omega=0)$$
$$=\frac{U}{4N}\sum_{E_k>|\mu|}{}' \frac{(\epsilon_k+\epsilon_{k+q})^2}{E_k^3}$$
$$-\frac{U}{2N}\sum_{\substack{E_{k+q}>|\mu| \\ E_k<|\mu|}}{}' \frac{(\epsilon_k+\epsilon_{k+q})^2}{E_kE_{k+q}}\frac{1}{E_{k+q}-E_k} \ . \tag{47}$$

At low doping, the first term remains practically the same as at half filling, and for $q\rightarrow q+Q$ reduces to αq^2, where

$$\alpha=\frac{Ut^2}{N}\sum_{E_k>|\mu|}{}' \frac{\sin^2 k_x}{E_k^3} \ . \tag{48}$$

In the two limiting cases, $\alpha\sim Ut/\pi^2\Delta^2$ for $U\ll t$ and $\alpha=2t^2/U^2$ for $U\gg t$. The second contribution basically has the structure of Pauli susceptibility for a Fermi gas. It is nonzero only away from half filling, when the chemical potential moves into the valence band and produces a finite density of states at the Fermi level. To make this analogy more explicit, it is convenient to rewrite the second term in a form $q^2(Ut^2/\Delta^2)\tilde{\chi}^{+-}$, where

$$\tilde{\chi}^{+-}=\frac{2}{N}\sum_{\substack{E_{k+q/2}>|\mu| \\ E_{k-q/2}<|\mu|}}{}' \frac{1}{E_{k+q/2}-E_{k-q/2}}$$
$$\times\frac{(q_x\sin k_x+q_y\sin k_y)^2}{q^2} \ . \tag{49}$$

Thus $\tilde{\chi}^{+-}$ is actually Pauli susceptibility modified by a momentum-dependent structure factor.

Since both contributions to Eq. (47) are proportional to

q^2, they combine to produce a renormalized spin stiffness

$$\rho_s=\rho_s^0(1-\Gamma^{+-}\tilde{\chi}^{+-}) , \tag{50}$$

where Γ^{+-} is given by $Ut^2/\Delta^2\alpha$ and in the two limiting cases reduces to $\Gamma^{+-}\sim t$ when $U\ll t$, and $\Gamma^{+-}\sim U$ when $U\gg t$.

The calculation of $\tilde{\chi}^{+-}$ strongly depends on the specific structure of the hole spectrum. In the mean-field approximation, the Fermi surface stretches along the whole reduced Brillouin zone boundary. At the same time the Fermi velocity is zero right on the zone boundary. This results in a very large ($\sim 1/\epsilon_F$) density of states, which in turn leads to Pauli susceptibility which diverges as one approaches the limit of zero doping,[31]

$$\tilde{\chi}^{+-}\sim\frac{\Delta}{t}\frac{1}{\epsilon_F} . \tag{51}$$

The divergence in $\tilde{\chi}^{+-}$ results in a negative spin stiffness for arbitrarily small doping levels. This result was obtained in Ref. 31. We will discuss the validity of the divergence in $\tilde{\chi}^{+-}$ in the next section. At the moment we wish to complete the mean-field description and so we turn to the treatment of the longitudinal fluctuations.

As we discussed above, away from half filling there is no decoupling between the charge and longitudinal spin fluctuations. Thus the calculation of the total susceptibility requires a solution of a 2×2 matrix equation. The final expression is similar to the one previously given for $\bar{\chi}^{+-}$,

$$\bar{\chi}^{zz}(q,q,\omega)=\frac{\chi_0^{zz}(q,\omega)[1+U\chi_0^{\rho\rho}(q+Q,\omega)]+U[\chi_Q^{z\rho}(q,\omega)]^2}{[1-U\chi_0^{zz}(q,\omega)][1+U\chi_0^{\rho\rho}(q+Q,\omega)]+U^2[\chi_Q^{z\rho}(q,\omega)]^2} . \tag{52}$$

Here, $\chi_0^{zz}(q,\omega)$ and $\chi_0^{\rho\rho}(z,\omega)$ are given by

$$\chi_0^{zz}=\chi_0^{\rho\rho}=\frac{1}{2N}\sum_{E_{k+q}>|\mu|}{}' \left[1-\frac{\epsilon_k\epsilon_{k+q}+\Delta^2}{E_kE_{k+q}}\right]\left[\frac{1}{E_k+E_{k+q}-\omega}+\frac{1}{E_k+E_{k+q}+\omega}\right]$$
$$+\frac{1}{2N}\sum_{\substack{E_{k+q}>|\mu| \\ E_k<|\mu|}}{}' \left[1+\frac{\epsilon_k\epsilon_{k+q}+\Delta^2}{E_kE_{k+q}}\right]\left[\frac{1}{E_{k+q}-E_k-\omega}+\frac{1}{E_{k+q}-E_k+\omega}\right] , \tag{53}$$

while $\chi_Q^{\rho z}(q,\omega)$ only has a contribution from the valence band fermions

$$\chi_Q^{\rho z}(q,\omega)=\frac{\Delta}{2N}\sum_{\substack{E_{k+q}>|\mu| \\ E_k<|\mu|}}{}' \frac{E_k+E_{k+q}}{E_kE_{k+q}}\left[\frac{1}{E_{k+q}-E_k-\omega}+\frac{1}{E_{k+q}-E_k+\omega}\right] . \tag{54}$$

The absence of the Goldstone poles in these channels allows us to limit ourselves to considering only the static susceptibility. Near $q=Q$, the denominator in Eq. (52) is modified from its value at half filling by a factor which involves Pauli-like susceptibility, much as we found for $\bar{\chi}^{+-}$. This factor is given by $1-\Gamma^{zz}\tilde{\chi}^{zz}(q)$ where $\tilde{\chi}^{zz}(q)$ differs from the corresponding expression in the transverse spin channel only by the absence of the momentum-dependent structure factor, and Γ^{zz} is related to the bare susceptibility at half filling,

$$\Gamma^{zz}=\frac{U^2\chi_0^{zz}(Q)}{1-U\chi_0^{zz}(Q)} , \tag{55}$$

where $\chi_0^{zz}(Q)\equiv\chi_0^{zz}(Q,\omega=0)$.

In the limiting cases, $\Gamma^{zz}\sim t(U/t)^{1/2}$ for $t\ll U$, and $\Gamma^{zz}\sim J=4t^2/U$ for $t\gg U$.

Just as in the transverse channel, Pauli-like susceptibility $\tilde{\chi}^{zz}$ is divergent at small dopings because the density of states scales as $1/\epsilon_F$ anywhere along the continuous Fermi surface. However, there is also an additional logarithmic divergence of the density of states from the van Hove singularities near the corners of the Brillouin zone. These singularities were suppressed in the $\bar{\chi}^{+-}$ channel by the structure factor. In explicit form, the expression for $\tilde{\chi}^{zz}$ is

$$\tilde{\chi}^{zz}(q\to Q)=\frac{1}{\pi^2}\frac{\Delta}{t}\frac{1}{\epsilon_F}\ln\left[\frac{4t}{\epsilon_F}\right] . \tag{56}$$

It thus follows that in mean-field approximation there are immediate instabilities away from half filling in both the transverse and the longitudinal channels.

IV. RENORMALIZATION EFFECTS

A. Effective Hamiltonian

Now we turn to a discussion of the validity of the results obtained in the mean-field approach. The first point to be addressed is the divergence of the Pauli-like susceptibility.

We have seen that the power-law singularities in $\tilde{\chi}^{+-}$ and $\tilde{\chi}^{zz}$ are related to the fact that the mean-field Fermi surface stretches all around the Brillouin zone even while the doping goes to zero. This effectively makes the problem one-dimensional. However, power-law divergencies in one dimension are often artificial and disappear upon vertex renormalization. We will show in Sec. IV C that this is the case for the present problem as well, i.e., for an "open" Fermi surface the vertex corrections cancel the power-law divergence both in the large- and small-U limits. Meanwhile, there are other reasons that make the

question of the power-law divergencies somewhat unphysical. It was pointed out by many authors that the open Fermi surface for arbitrarily small doping concentrations is an artifact of the mean-field approximation. From the symmetry point of view, there are no reasons for the hole energy to be exactly constant along the Brillouin zone boundary. Numerical,[32] variational,[33–35] as well as some perturbative[36] calculations suggest that the actual band for a single hole has minima at $q=(\pm\pi/2,\pm\pi/2)$, at least for moderately large values of U/t. Thus, at low dopings the problem is actually two-dimensional since then the Fermi surface forms separate small pockets near the band minima. In this situation, the susceptibility is given by a two-dimensional expression which does not depend on the Fermi energy. Indeed, near its minima the band structure is well approximated by a quadratic form[10]

$$E_k \simeq E_{\min}+\frac{k_\perp^2}{2m_\perp}+\frac{k_\parallel^2}{2m_\parallel}\ , \tag{57}$$

where k is measured with respect to the corresponding minimum, while $m_\perp$ and $m_\parallel$ are the effective masses normal and tangent to the Brillouin zone boundary.

In this case, the density of states at the Fermi surface is independent of the density of holes, and the susceptibility is

$$\tilde{\chi}=\frac{\sqrt{m_\parallel m_\perp}}{2\pi} \tag{58}$$

for both the transverse and longitudinal spin channels, since the structure factor in Eq. (49) can be set to unity at $k=(\pm\pi/2,\pm\pi/2)$.

The finiteness of the two-dimensional Pauli susceptibility reduces the question of immediate instability to the investigation of the scales of the effective masses and the effective interaction between the holes. The basic scale of the effective masses is determined by the bandwidth. For small U, the bandwidth is of order t, since the hopping term in the Hubbard model is dominant.[37] For large U, the gap $\Delta\simeq U/2\gg t$, and the mean-field expression for the energy of the valence band fermions can be expanded for any k to be

$$E_k-\Delta\simeq\frac{\epsilon_k^2}{2\Delta}=\frac{2t^2(\cos k_x+\cos k_y)^2}{\Delta}\ . \tag{59}$$

From this formula the bandwidth of excitations is of the order of $J=4t^2/U$.

Certainly, the mean-field expression for E_k should be taken with caution at large U since the self-energy corrections are by no means small. However, we will see that while the mean-field description of the Hubbard model in the large-U limit should be significantly modified, the bandwidth of the fermionic spectra remains of the order of J in the more sophisticated theories as well. This scale for the bandwidth has been confirmed by the variational[33–35] and numerical[38,32] studies of the t-J model, with $m_\parallel$ and $m_\perp$ being of order $1/t$ or $1/J$ in the small- and large-U limits, respectively.

There is now a clear difference between the longitudinal and the transverse channels. The effective coupling Γ^{zz} which multiplies Pauli-like susceptibility in the longitudinal channel is less than the bandwidth in the small-U limit and is of the same order of magnitude as the bandwidth at large U. Accordingly there is no longer a large parameter which would force the instability in $\bar{\chi}^{zz}$ at low doping.

On the contrary, the effective coupling Γ^{+-} in the transverse spin channel is of the order of the bandwidth for small couplings, but for strong coupling is large by a factor of $(U/t)^2$ compared to the bandwidth. Thus the improvements on the mean-field theory we have incorporated so far still suggest an immediate instability in $\bar{\chi}^{+-}$ governed by a large parameter. However, this result also looks suspicious. Indeed, we have a situation where the coupling is much larger than the bandwidth. Such a theory is expected to have a strong renormalization of both the self-energy and the coupling constant. To clarify this point we introduce a diagrammatic description of the problem.

In the diagrammatic language, the RPA expressions for the total susceptibilities result from summations of infinite sequences of the ladder (for $\bar{\chi}^{+-}$) and bubble (for $\bar{\chi}^{zz}$) diagrams (see Fig. 1). We will focus on the $\bar{\chi}^{+-}$ channel, but will keep using the term "bubble" for the ladder sequence constituents as well. Away from half filling, each bubble has either one fermion from the conduction and one from the valence band or both fermions from the valence band [see Fig. 3(a)]. There are terms in the expansion which have adjacent bubbles with only valence fermions in both of them. However, we will show below that whenever two such bubbles are next to

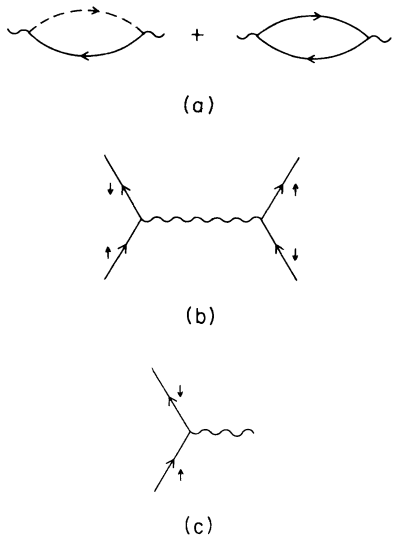

FIG. 3. Reduction of the electronic problem to the system described as holes interacting by the exchange of collective excitations. (a) Each "bubble" in the RPA sequence for susceptibilities can be broken up into a diagram with both fermions in the valence band (solid lines) as well as a diagram with one fermion in the valence and one in the conduction (dashed line) band. (b) The valence-conduction bubbles of (a) are summed up to represent spin waves whose exchange mediates hole-hole interactions, as shown here. (c) A triple vertex for a single hole interacting with a spin wave.

each other, the diagram acquires an extra factor of q^2, where q is measured with respect to $Q=(\pi,\pi)$. Thus in the low-q limit that we are interested in, one should retain only those diagrams in which the "valence-valence" bubbles are separated by subsequences of the "valence-conduction" ones. Each such subsequence can be summed up to give what is essentially a susceptibility at half filling. The poles of this susceptibility correspond to spin-wave excitations and we thus come to a picture where the dominant interaction between the valence fermions is an exchange of spin waves. In diagrammatic language, this means that in perturbation theory we deal with the interaction vertex of the type shown in Fig. 3(b), where the solid lines represent the Green's functions of the holes and the wavy line is the total static transverse susceptibility near the antiferromagnetic wave vector. Incidentally, this formulation of the problem closely corresponds to the approach based on the t-J model[10,11,39] and is convenient for comparisons with it.[40]

We now demonstrate our assertion that a q^2 factor arises whenever two "valence-valence" bubbles are adjacent in the PRA ladder. To see this, we calculate the direct interaction between valence electrons. It is given by a Hubbard U term multiplied by the coherence factors which arise from projecting the interaction onto the valence band. The calculation is straightforward and results in a four-fermion interaction given by

$$\Lambda(k_1,k_2;k_3,k_4) = U\{(u_{k_1}u_{k_2}-v_{k_1}v_{k_2})(u_{k_3}u_{k_4}-v_{k_3}v_{k_4})\}\sigma^{+}_{\alpha_1\alpha_2}\sigma^{-}_{\alpha_3\alpha_4} . \tag{60}$$

Here, k_1,k_3 and k_2,k_4 are the momenta of the outgoing and incoming valence fermions with opposite spins, respectively, and u_i,v_i are the coherence factors given by Eq. (7). Expanding the Bogolyubov coefficients, we obtain

$$\Lambda(k_1,k_2;k_3,k_4)\sim U(\epsilon_{k_1}+\epsilon_{k_2})(\epsilon_{k_3}+\epsilon_{k_4})\sim q^2 \tag{61}$$

if the momentum transfer $k_1-k_2=k_4-k_3$ is near $Q=(\pi,\pi)$.

We would also like to show that if we incorporate the momentum-dependent structure factor in Eq. (49) into the coefficient Γ^{+-} of Eq. (50), it will exactly coincide with the four-fermion vertex of Fig. 3(b), taken in the limit $k_1=k_4$, $k_2=k_3$. This vertex can be represented as a product of the static susceptibility $\bar{\chi}^{+-}$ and two triple vertices [Fig. 3(c)]. The interaction between the fermions and spin waves is again a Hubbard U term multiplied by the corresponding coherence factors. In the original RPA sequence this vertex comes from an interaction term between a "conduction-valence" bubble and an adjacent "valence-valence" one. The coherence factors from the "conduction-valence" side are included in the total static susceptibility. What is left contributes to the triple vertex

$$\Phi(k_1,k_2)=U(v_{k_2}v_{k_1}-u_{k_2}u_{k_1})=U\frac{\epsilon_{k_1}+\epsilon_{k_2}}{2\sqrt{E_{k_1}E_{k_2}}} . \tag{62}$$

For the momentum transfer k_2-k_1 close to Q, this expression is linear in q. However, the total fourfold vertex is finite in this limit since the q^2 term which comes from two threefold vertices is cancelled by the $1/q^2$ dependence of the static susceptibility[7]

$$\bar{\chi}^{+-}(q+Q,q+Q;\omega=0)\simeq 1/\alpha U q^2 , \tag{63}$$

where α is given by Eq. (48). The final result for the four-fermion interaction for the momentum transfer close to Q and $k_1=k_4$ is

$$\Gamma^{+-}=\Phi^2(k_1,k_2)\bar{\chi}^{+-}=\left[\frac{Ut^2}{\Delta^2\alpha}\right]\frac{(q_x\sin k_x+q_y\sin k_y)^2}{q^2} , \tag{64}$$

where $k_1=k_4=k+q/2$ and $k_2=k_3=k-q/2$. This is exactly what we obtained within the RPA approach [cf. Eqs. (49) and (51)]. Below we will keep the notations Γ^{+-} and $\bar{\chi}^{+-}$, but assume that the structure factor is incorporated in Γ^{+-}.

The diagrammatic representation of the problem is a convenient way to discuss the validity of the mean-field approximation. For small U, the effective interaction between quasiparticles is of the order of the bandwidth in the transverse channel and so both vertex and self-energy corrections are of the order of unity. In this limit one might expect that the mean-field treatment of the problem leads to at least qualitatively correct results. On the contrary, for large U the interaction is much larger than the bandwidth of both the holes and the spin waves. This is an inherently strong-coupling situation in which we are bound to have strong vertex and self-energy renormalization. For example, the first corrections to the vertex function and self-energy are given by the diagrams in Figs. 4(a) and 4(b). When calculated with bare vertices and Green's functions they produce a correction of order $(t/J)^4$ to the vertex and of order $(t/J)^2$ to the self-energy. Hence the mean-field solution cannot be directly used,

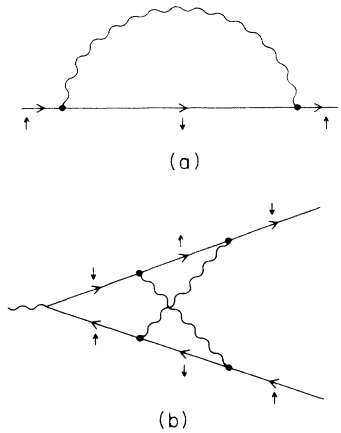

FIG. 4. (a) Self-energy of holes; (b) Leading-order vertex corrections to the hole–spin-wave interaction vertex. In (b), because of the presence of the spin indices, the leading-order graph has two, rather than one, spin-wave lines. Both diagrams here are to be understood as applying to a *single* hole introduced into an antiferromagnetic background.

and the only possible way to proceed is to look for some effective theory which will describe the system at small energy scales of order J.

Some information about the general structure of the spectral weight $A(k,\omega)$ of the exact hole Green's function can be inferred from the available numerical[38,32] and variational[33–35,41] treatments of the t-J model for moderately large values of t/J. The picture which emerges is that $A(k,\omega)$ has a single δ-function peak at the frequencies of order J and an incoherent part whose dominant weight is located at the energy scales which are larger than J by some fractional power of t/J. In other words, $G(k,\omega)$ can be approximated by

$$G(k,\omega)\simeq\frac{Z}{\omega-E_k+i\delta}+G_{\rm inc}(k,\omega)\ , \tag{65}$$

where $G_{\rm inc}(k,\omega)$ is the incoherent part. A nontrivial consequence of these studies is that the basic energy scale of the coherent hole band remains of the same order of magnitude ($E_k\sim J$) as in the mean-field treatment.

Our goal is to obtain a low-energy theory at the energy scales of the order of J. At these scales the interaction reduces to some function of momenta and, generally, the frequency. The momentum dependence of the interaction is determined by symmetry considerations, which require the interaction Φ to scale linearly with q at small $q-Q$, just as the bare vertex given in Eq. (62). For example, if both holes are located near $k=(\pm\pi/2,\pm\pi/2)$, then the momentum dependence of the vertex is $\Phi\sim(q_x\pm q_y)$. The frequency dependence of the vertex is unlikely to be singular on the scale of J and so the only important parameter in the problem is the overall scale of the interaction at these frequencies.

Actually, we do not need to know the separate values of the vertex $\Phi(k_1,k_2,\omega)$ and the Z factor in the Green's function. The physically relevant coupling at low energies is determined by a product $\Phi^R(k_1,k_2,\omega)=Z\Phi(k_1,k_2,\omega)$. Strictly speaking, any precise statement about the magnitude of this quantity requires solving a strong-coupling problem. However, we observe that with $\Phi^R(p_1,p_2;\omega\sim J)$ being of order J, $\Gamma^{+-}\sim J$ and the low-energy theory becomes "self-consistent" in a sense that it does not generate any new energy scales besides J. Indeed, if one calculates the same diagrams as in Fig. 4, using renormalized Green's functions and vertices, then the self-energy and vertex corrections contain factors $(\Phi^R/J)^2$ and $(\Phi^R/J)^4$, respectively, which are simply numerical factors if $\Phi^R\sim J$.

The assumption that at large U the physical interaction between the holes is likely to be of order J had been put forward by Shraiman and Siggia.[39] Kane, Lee, and Read[42] calculated the self-energy correction with nonrenormalized triple vertices $\sim t$ and obtained $Z\sim J/t$. This will also give $\Phi^R\sim\Gamma^{+-}\sim J$. We take a viewpoint that it is hard to separate self-energy renormalization from the renormalization of the vertex. At the same time, the result that $\Phi^R\sim\Gamma^{+-}\sim J$ seems very plausible, and we will adopt it in the rest of the paper.

A similar treatment can be applied to the χ^{zz} channel. The effective interaction can be represented as an exchange of longitudinal spin fluctuations. In this case the validity of such treatment is not, however, based on the smallness of q, but rather on the smallness of ϵ_F/Δ, which works at low dopings.

The renormalization of the effective interaction in the longitudinal channel is less important since the mean-field value is already of the same order of magnitude as the bandwidth. We have not been able to reach a definite conclusion as to whether the effective coupling constant in this channel is significantly renormalized down from a value of order J. In any case, we will assume that the interaction in the $\bar{\chi}^{zz}$ channel is smaller than that in the channel of transverse spin fluctuations. This is also consistent with the mean-field results in the small-U limit.

B. Vacuum renormalization

With the effective vertex Φ^R being of the order of the bandwidth, there is no longer a large parameter which would govern an instability in the $\bar{\chi}^{+-}$ channel at arbitrarily small doping levels. It was argued in Ref. 10 that in a strictly 2D situation, the instability is still likely to occur immediately away from half filling if the ratio of the two effective masses, $m_\parallel/m_\perp$ is sufficiently large. This would be the case if the hole band retained some of its mean-field features and, although it now had minima at $(\pm\pi/2,\pm\pi/2)$, remained flat along the Brillouin zone boundary. As was said above, the susceptibility of a 2D Fermi gas with a parabolic dispersion is given by

$$\tilde{\chi}=\frac{\sqrt{m_\parallel m_\perp}}{2\pi}=m_\perp\frac{\sqrt{m_\parallel/m_\perp}}{2\pi}\ . \tag{66}$$

The transverse mass $m_\perp$ was of the order of the bandwidth in the mean-field solution, and is expected to remain that way, while $m_\parallel$ was infinite in the mean-field treatment and may be considerably larger than $m_\perp$ in the actual band structure. Note that this question is important only for large U. For small U, the large ratio of $m_\parallel/m_\perp$ was already taken into account when defining the basic scale for the effective mass.[37]

The conclusion of Ref. 10 is based on the fact that the 2D susceptibility has a discontinuity at $\epsilon_F=0$, i.e., $\Gamma^{+-}\tilde{\chi}^{+-}$ is a constant even at arbitrarily small ϵ_F. Actually, this is true only within the Hartree-Fock treatment. The authors of Ref. 10 pointed to the limitation of the mean-field approach, but resorted to finite temperature and the effects due to the third dimension to delay the instability. We have found that the physics in two dimensions at $T=0$ already does not have such a discontinuity. The reason is that the vacuum renormalization which transforms the interaction potential into a scattering amplitude is singular at low energies in the 2D case.[43] This renormalization comes from a ladder sequence of diagrams in the particle-particle channel, shown in Fig. 5. For the interaction of Eq. (64), which is only weakly momentum dependent, each such diagram is logarithmically divergent in the limit of low energies. The sum of the ladder diagrams depends crucially on the sign of the effective coupling. As follows from Eq. (64), the four-fermion interaction in the $\bar{\chi}^{+-}$ channel contains a dominant short-range repulsive part (also see Ref. 7 and the Appendix). The question of the sign of this interaction

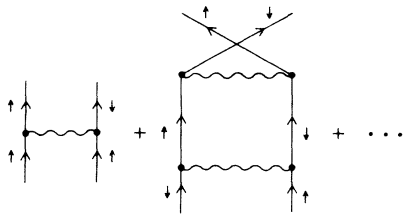

FIG. 5. The lowest-order ladder diagrams for the scattering amplitude. Summations of the whole sequence leads to an effective interaction which is much weaker than the "bare" one in the repulsive case in 2D at low doping.

has caused some confusion in the literature and is further clarified in the Appendix.

In the case of repulsion, the summation of the ladder sequence of diagrams leads to a scattering amplitude which logarithmically goes to zero with the Fermi energy. The explicit summation for the anisotropic mass case yields

$$T^{+-}(k,\omega)=\frac{\Gamma^{+-}}{1+(\Gamma^{+-}/4\pi)\sqrt{m_\perp m_\parallel}\ln(1/\omega m)} . \tag{67}$$

$T^{+-}(k,\omega)$ is a scattering amplitude and ω is a characteristic energy scale in the particle-hole channel which at low doping is of the order of ϵ_F^2/Δ. The cutoff parameter $1/m$ is determined by the energy scale up to which the approximation of the effective interaction by leading order in q is valid. For large U, it obviously coincides with the bandwidth, while for small U, spin-wave excitations are present only at low energies and with a logarithmic accuracy $1/m$ is of order Δ.

The quantity governing the instability is the product $T^{+-}\tilde{\chi}^{+-}$. This product now is

$$T^{+-}\tilde{\chi}^{+-}\simeq\frac{(\Gamma^{+-}/2\pi)\sqrt{m_\perp m_\parallel}}{1+(\Gamma^{+-}/4\pi)\sqrt{m_\perp m_\parallel}\ln(1/\omega m)} . \tag{68}$$

It thus follows that the logarithmic singularity also makes the question about the ratio of the effective masses irrelevant, since at small doping

$$T^{+-}\tilde{\chi}^{+-}\sim\frac{2}{\ln(1/\omega m)} \tag{69}$$

independently of the mass ratio.

Thus we conclude that the corrections to the spin stiffness are actually small at low doping, and the commensurate antiferromagnetic state is stable against the transverse fluctuations at a sufficiently small concentration of holes. Now we discuss what happens at finite doping.

C. Instability at finite doping

The vacuum renormalization is a powerful way to weaken the interaction between the holes when the exchange of a transverse spin wave leads to a nearly δ-function-type repulsion. This is definitely the case when the holes are located in pockets near $(\pm\pi/2,\pm\pi/2)$. However, when ϵ_F grows enough to open the Fermi surface along the whole Brillouin zone boundary, the momentum dispersion of the interaction prevents the vacuum renormalization from significantly changing the interaction potential between the holes.

We have shown in Sec. III that in a mean-field approximation, which produces an open Fermi surface, the Pauli-like susceptibility of holes diverges as $\tilde{\chi}^{+-}\sim(1/t)(\Delta/\epsilon_F)$ and $\tilde{\chi}^{zz}\sim 1/t(\Delta/\epsilon_F)\ln(t/\epsilon_F)$, when one approaches the limit of zero doping. Actually, on the basis of a previous one can expect that the Fermi surface spans the whole Brillouin zone boundary when the Fermi momenta $k_\parallel^F$ become of the order of unity. This happens when ϵ_F becomes of order Δ for $U\ll t$ and of order t for $U\gg t$. At these values of ϵ_F, $\tilde{\chi}^{+-}$ is of the order of the inverse bandwidth for both small and large U, while $\tilde{\chi}^{zz}\sim(1/t)(t/U)^{1/2}$ for small U, and $\tilde{\chi}^{zz}\sim 1/J$ for large U.

Since for $U\gg t$, $\tilde{\chi}^{+-}\sim\tilde{\chi}^{zz}\sim 1/J$, the transverse channel is more favorable for the magnetic instability in the large-U limit because of a larger interaction between the holes (see Sec. IV A). On the other hand, $\Gamma^{+-}\tilde{\chi}^{+-}\sim O(1)$ and so even for a moderate density of holes there is no large parameter which would allow one to make a definite conclusion about whether the transition to a spiral phase actually occurs.

For small U, the factor $(t/U)^{1/2}$ in $\tilde{\chi}^{zz}$ compensates the relative smallness of Γ^{zz} [see Eq. (55)] and the product $\Gamma\tilde{\chi}$, which governs the possible magnetic instability, turns out to be of the order of unity in both the transverse and longitudinal channels. This makes the situation even more indefinite.

At the same time, the photoemission studies of the high-T_c materials[44] show that the Fermi surface is continuous over the whole reduced Brillouin zone already at relatively small dopings. In this situation, it might well be that for the doping concentrations of interest the Pauli-like susceptibilities are numerically much larger than the corresponding bandwidths. On the one hand, this again implies that the renormalization effects are large. On the other hand, when $\tilde{\chi}$ is large, there is a way to distinguish between the transverse and longitudinal channels in the small-U limit. We have already mentioned that a power-law ($\sim 1/\epsilon_F$) divergence in $\tilde{\chi}$ is a direct consequence of quasi-one-dimensionality which is present in the mean-field approach to the problem. As in many other 1D problems, this divergence is likely to disappear upon the vertex renormalization induced by a finite density of holes. This renormalization comes from the diagrams where additional bubbles are inserted into a given one.

The leading renormalization of Γ^{zz} comes from the diagram of Fig. 6(a). The integration over the internal momenta on both sides of the effective interaction gives $(\tilde{\chi}^{zz})^2$ provided the interaction inside the bubble reduces to some constant at the momentum transfers equal to Q. This requirement is related to the fact that in order to obtain $\ln(1/\epsilon_F)$ in $\tilde{\chi}^{zz}$, the momenta of the fermions should be close to the corners of the Brillouin zone.

The effective interaction which is inserted into the bubble is a total interaction between the holes with parallel spins. It has contributions from $\bar{\chi}^{zz}$ and $\bar{\chi}^{\rho\rho}$ channels,

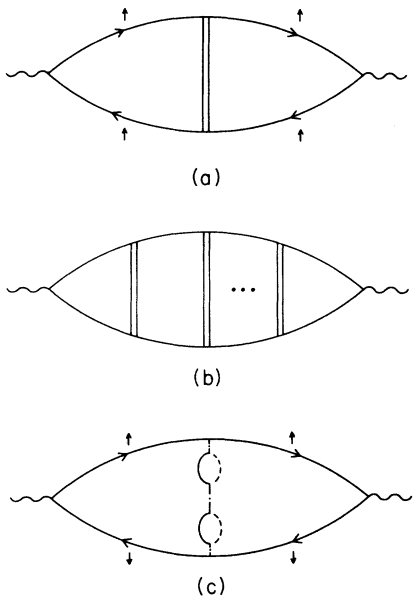

FIG. 6. Vertex corrections at small U. (a) The leading-order vertex correction in the longitudinal spin channel: the double line represents an exchange of the longitudinal fluctuations given by a sequence of Fig. 1(a). (b) A ladder obtained by iterating the correction in (a). (c) The lowest-order vertex correction in the transverse channel; solid lines are the valence band fermions, dashed lines represent the conduction band fermions, and the dash-dotted line is the Hubbard U interaction.

and in the RPA treatment has a finite value

$$\frac{1}{4}[\Gamma^{zz}(Q)+\Gamma^{\rho\rho}(Q)]=\frac{U^2\chi_0^{zz}(Q)}{2\{1-[U\chi_0^{zz}(Q)]^2\}} \tag{70}$$

when the momentum transfer is equal to Q.

For small U, $\Gamma^{zz}(Q)\gg\Gamma^{\rho\rho}(Q)$ and the renormalization multiplies $\tilde{\chi}^{zz}$ by a factor $\frac{1}{4}\Gamma^{zz}\tilde{\chi}^{zz}$, i.e., the product $\Gamma^{zz}\tilde{\chi}^{zz}$ acquires a correction of the same order as itself. One can sum up the ladder sequence of diagrams for a total vertex $\bar{\Gamma}^{zz}$ [Fig. 6(b)] and obtain that the product $\bar{\Gamma}^{zz}\tilde{\chi}^{zz}$, which governs the instability in a longitudinal spin channel, reduces to a constant even in the limit when $\Gamma^{zz}\tilde{\chi}^{zz}\gg 1$:

$$\bar{\Gamma}^{zz}\tilde{\chi}^{zz}=\frac{\Gamma^{zz}\tilde{\chi}^{zz}}{1+\frac{1}{4}\Gamma^{zz}\tilde{\chi}^{zz}} . \tag{71}$$

A similar renormalization can be performed for Γ^{+-}. The lowest-order diagram is shown in Fig. 6(c). It necessarily involves two additional particle-hole bubbles to fit the requirements on spin indices.

In contrast to the bare interaction between the holes with antiparallel spins, which scaled as $\Lambda\sim\epsilon_F^2$ for particles on the Fermi surface [see Eq. (61)], the renormalized interaction $\bar{\Lambda}$ remains finite in the limit $\epsilon_F\to 0$. Hence, the diagram of Fig. 6(c) contains $(\chi^{+-})^2$ and after summing up the ladder sequence of diagrams, one again obtains that the product $\bar{\Gamma}^{+-}\tilde{\chi}^{+-}$ remains finite even if one deals with the open Fermi surface at very small ϵ_F. However, in the transverse channel, there is no restriction on the transferred momenta and the integration over the internal momentum in each bubble contributes only a single power of $\ln\Delta\sim(t/U)^{1/2}$. As a result, $\bar{\Gamma}^{+-}\tilde{\chi}^{+-}$ is given by

$$\bar{\Gamma}^{+-}\tilde{\chi}^{+-}=\frac{\Gamma^{+-}\tilde{\chi}^{+-}}{1+\beta(U/t)^2\Gamma^{+-}\tilde{\chi}^{+-}} , \tag{72}$$

where β is a numerical constant.

It thus follows that when $\Gamma^{+-}\tilde{\chi}^{+-}\gg 1$, the effective dimensionless coupling in the transverse channel is large: $\bar{\Gamma}^{+-}\tilde{\chi}^{+-}\sim(t/U)^2\gg 1$. Hence, in the small-$U$ limit, the "strength" of the magnetic fluctuations in the transverse channel is much larger than that in the longitudinal spin fluctuation channel. This favors the occurrence of the instability in the transverse channel, just as we argued for large U.

Of course, the above arguments are not rigorous since we cannot prove that $\Gamma^{+-}\tilde{\chi}^{+-}$ is actually large in some region of the parameter space. However, the result that there is no qualitative difference between the situation at small and large U seems to be rather plausible since the low-energy physics at zero and small doping is essentially the same in the two limits.

We also note that if the spiral instability occurs at the doping levels at which the Fermi surface is open, it is no longer essential whether the band minima are at $(\pm\pi/2,\pm\pi/2)$. Thus while on the one hand the instability depends on an unknown numerical factor, on the other it appears possible under more general conditions than the original derivation may lead one to believe.

V. DYNAMICAL PROPERTIES AT FINITE DOPING

The consequences of the instability in the $\bar{\chi}^{+-}$ channel were considered in detail by Shraiman and Siggia.[10] They have shown that on a mean-field level the instability leads to an incommensurate magnetic structure known as the spiral phase. Several authors have looked into the question of stability of this phase against quantum fluctuations.[45–47] We will not discuss this point but rather focus on the dynamical properties of the transition. It turns out that the nature of the instability is completely different from the conventional picture of incommensurate transitions in purely magnetic systems.[21–24] In the latter case, the appearance of the incommensurate structure is inevitably preceded by a softening of spin-wave excitations. We will show that in the present case the transition does not affect the spin-wave excitations, which in the doped case remain at practically the same frequencies as at half filling. The instability of the commensurate antiferromagnetic state is instead driven by collective fermionic modes. Specifically, when the static susceptibility changes its sign, the dynamical susceptibility acquires a second pole at an imaginary frequency with a residue which goes to zero at the transition point.

To see how this occurs, consider an RPA expression for the dynamic susceptibility, given in Eq. (14). Expanding the denominator for small q and ω, we obtain $(\bar{\chi}^{+-})^{-1}$ is proportional to

$$\left[\frac{1}{4N}\sum_{E_k>|\mu|}'\frac{(\epsilon_{k+q}+\epsilon_k)^2}{E_k^3}-\frac{1}{2N}\sum_{\substack{E_{k+q}>|\mu|\\E_k<|\mu|}}'\frac{(\epsilon_{k+q}+\epsilon_k)^2(E_{k+q}-E_k)}{E_kE_{k+q}[(E_{k+q}-E_k)^2-\omega^2]}\right]$$
$$-\omega^2\left\{\left[\frac{1}{4N}\sum_k'\frac{1}{E_k^3}-\frac{1}{2N}\sum_{\substack{E_{k+q}>|\mu|\\E_k<|\mu|}}'\frac{(E_{k+q}-E_k)}{E_kE_{k+q}[(E_{k+q}-E_k)^2-\omega^2]}\right]\right.$$
$$\left.+\gamma\left[\frac{1}{2N}\sum_{E_k>|\mu|}'\frac{1}{E_k^3}-\frac{1}{2N}\sum_{\substack{E_{k+q}>|\mu|\\E_k<|\mu|}}'\frac{(E_{k+q}-E_k)}{E_kE_{k+q}[(E_{k+q}-E_k)^2-\omega^2]}\right]^2\right\}, \tag{73}$$

where $\gamma=U\Delta^2/(1-U\chi_0(q=\omega=0))$. At half filling, this expression vanishes at spin-wave energies $\omega_q=cq$, where c is given by Eqs. (19) and (22). Away from half filling, Eq. (73) contains contributions from the fermionic excitations whose energy scale is of order v_Fq, where v_F is a Fermi velocity which is much less than the spin-wave velocity. Hence, when calculating how the spin-wave pole shifts upon doping, one can neglect $(E_{k+q}-E_k)^2$ compared to ω^2 in the denominators occurring in Eq. (73). As a result the spin-wave pole occurs at nearly the same frequencies as at half filling.

Within the RPA approximation with renormalized interactions and for an open Fermi surface without pockets, the expressions for the renormalized spin-wave velocity are

$$\omega^2=\begin{cases}c^2q^2\{1+(U_{\rm eff}/3\pi^2J)(\epsilon_F/t)(1+\sin^4 2\phi)\}, & U\gg t\\ c^2q^2\{1+(4\pi/3)(t/U)^{1/2}(\epsilon_F/t)(1+\frac{1}{4}\sin^4 2\phi)\}, & U\ll t\end{cases} \tag{74}$$

where c is a spin-wave velocity at half filling, the angle ϕ is measured from the $\hat{x}$ axis, and $U_{\rm eff}$ is the effective interaction, which at large U is of order J, as discussed at length above. A nontrivial feature of these expressions is the appearance of an angular dependence of the spin-wave velocity. This reflects the fact that away from half filling the holes mediate an additional long-range dipolar interaction between the spins. Furthermore, we can see that the spin-wave velocity increases upon doping in both large- and small-U limits.

A drastic effect of finite doping can be seen at very small frequencies, $|\omega|\ll cq$. An explicit analysis of the mean-field expression [Eq. (73)] shows that immediately away from half filling $\bar{\chi}^{+-}$ acquires a second pole at

$$\omega^2=-\frac{16}{3\pi^2}\left[\frac{t^3}{\alpha\Delta^2}\right]\left[\frac{\epsilon_F}{\Delta}\right]q^2, \tag{75}$$

where α is defined by Eq. (48). For simplicity, the calculations were performed for $q_x=q_y$.

The appearance of a pole at $\omega^2<0$ obviously signals an instability of the commensurate state. However, we have already discussed that an immediate instability upon doping in the RPA approach is related to an unphysical one-dimensional divergence of Pauli-like susceptibility. In view of this it is preferable to describe the transition by a model where the Fermi surface consists of pockets near $(\pm\pi/2,\pm\pi/2)$ and the interaction between the fermions is described by a scattering amplitude T^{+-}. This model follows from the description in Sec. IV.

Within this model, the denominator in $\bar{\chi}^{+-}$ at $\omega\ll cq$ is given by

$$\bar{\chi}^{-1}\sim q^2\left[1-T^{+-}\frac{2}{N}\sum_{\substack{E_{k+q}>|\mu|\\E_k<|\mu|}}'\frac{E_{k+q}-E_k}{(E_{k+q}-E_k)^2-\omega^2}\right], \tag{76}$$

where

$$T^{+-}=\begin{cases}4\pi^2t, & t\ll U\\ \sim J, & t\gg U\end{cases} \tag{77}$$

and E_k is given by Eq. (57). For simplicity, we evaluate the integral in the isotropic mass limit $m_\parallel=m_\perp=m$, and obtain

$$\frac{2}{N}\sum_{\substack{E_{k+q}>|\mu|\\E_k<|\mu|}}'\frac{E_{k+q}-E_k}{(E_{k+q}-E_k)^2-\omega^2}=\frac{m}{2\pi}\left\{1-\left[\frac{\delta^2}{\delta^2-1}\right]^{1/2}\right\}, \tag{78}$$

where $\delta^2=\omega^2m^2/q^2k_F^2$. Hence

$$\bar{\chi}^{-1}\sim q^2\frac{mT^{+-}}{2\pi}\left[\frac{2\pi-mT^{+-}}{mT^{+-}}+\left[\frac{\delta^2}{\delta^2-1}\right]^{1/2}\right]. \tag{79}$$

It follows from Eq. (79) that if the scattering amplitude is small in comparison with the bandwidth (this is definitely the case for a very small doping), the transverse susceptibility has no poles besides a spin-wave one. However, when the scattering amplitude exceeds a critical

value $T_{cr}^{+-}=2\pi/m$, the static stiffness changes its sign and Eq. (79) acquires a root at $\omega^2<0$. Close to the instability, $\omega=\pm i\omega_0(q)$, where

$$\omega_0(q)=\frac{qk_F}{m}\left[\frac{T^{+-}}{T_{cr}^{+-}}-1\right]. \tag{80}$$

The finiteness of the spin-wave velocity at the transition point leads to a rather peculiar structure of the low-energy excitations in a spiral phase. From general symmetry considerations, this phase should have three gapless bosonic modes, related to a breakdown of $SO(3)$ symmetry. One of the Goldstone modes is at $k=0$ and the two others are at $\pm Q_0$ where Q_0 is a pitch of a spiral.[10] The system thus has two spin-wave velocities, c_0 and C_{Q_0}. In conventional magnets, both velocities go to zero as one approaches a transition between commensurate and incommensurate phases. In the present case, however, nothing special happens at $k=0$ and the low-energy excitations near this point scale with momenta as $\omega\sim c_0 q$ both above and below the transition. However, near $k=Q$, the low-energy dynamics below the transition is governed by a second pole in $\bar{\chi}^{+-}$, i.e., by the collective mode of the holes. It produces the Goldstone excitations near $\pm Q_0$ with $C_{Q_0}\sim(k_F/m)(T^{+-}/T_{cr}^{+-}-1)$, while the would be zero spin-wave mode at $k=Q$ acquires a finite gap of the order of ω_0. Such structure of the low-energy excitations was also found in a mean-field treatment of the spiral phase working from the t-J model.[10]

Another important issue is the stability of an ordered spiral phase against quantum fluctuations. One may think that fluctuations would definitely destroy the long-range order close to the transition since the velocity of the Goldstone bosons at $q=Q_0$ can be made arbitrarily small. However, by comparing susceptibilities at $\omega\sim|\omega_0(q)|$ with that near the spin-wave pole, we found an additional smallness of the residue in $\bar{\chi}^{+-}$ near $\omega\sim|\omega_0(q)|$ which goes to zero at $T^{+-}=T_{cr}^{+-}$, and neutralizes the $1/|C_{Q_0}|$ divergence in the corrections to the sublattice magnetization. Thus, to have a definite opinion about whether the fluctuations destroy the long-range order close to the spiral instability requires further considerations.

VI. CONCLUSIONS

In this paper we studied the stability of the commensurate AFM phase of the 2D Hubbard model at a small but finite density of holes. The main conclusions of our work are as follows.

(1) At a small density of holes, the Fermi surface is known to form small pockets which are likely to be located near $q=(\pm\pi/2,\pm\pi/2)$. In this case, the vacuum renormalization of the effective interaction between holes inevitably violates an immediate instability upon doping, and thus commensurate AFM survives until a finite concentration of holes is reached.

(2) At larger concentrations, the incommensurate instability may occur, however, its possibility is not related to the presence of any large parameter, but rather to the interplay of various numerical factors. At these concentrations, the Fermi surface may no longer consist of pockets, but we argue that this in no way precludes the instability.

(3) We compared the "relative strength" of the magnetic fluctuations in the transverse and longitudinal channels. We found that for both small and large values of the interaction, the first instability is likely to occur in the transverse channel. When treated on a mean-field level, this instability leads to a spiral magnetic phase.[10]

(4) We considered the dynamical susceptibility near the possible instability and found that spin waves play no role in the transition. In contrast, the incommensurate instability is governed by collective Fermionic excitations coupled to the spin background.

The preceding discussion was focused on the magnetic properties of the doped AFM. Another important issue that we wish to address briefly is the possibility for a pairing instability that would lead to superconductivity in the 2D Hubbard model. The authors of Ref. 2 considered the interaction between holes produced by the exchange of longitudinal spin fluctuations and found it to be attractive when two holes share the same pocket in the reduced Brillouin zone. In the case where the Fermi surface has pockets near $(\pm\pi,0)$ and $(0,\pm\pi)$, this would be a dominant interaction and superconducting instability might well compete with the magnetic one.

Throughout our article we focused on the situation where the pockets on a Fermi surface are located around $(\pm\pi/2,\pm\pi/2)$. In this case, there is a large contribution to the total interaction between the holes with opposite spins which comes from the transverse channel. This interaction is repulsive and overshadows the effect of attraction from the longitudinal channel. For small U, the repulsion due to the $\bar{\chi}^{+-}$ channel is of order t, while the attractive part, which comes from $\bar{\chi}^{zz}$, is of order $t(U/t)^{1/2}$. Note that in this limit $\bar{\chi}^{zz}(q)$ is strongly enhanced near $q\sim Q$ compared to charge fluctuation channel $\bar{\chi}^{\rho\rho}$, which can be neglected. For large U, both $\bar{\chi}^{+-}$ and $\bar{\chi}^{zz}$ produce the interaction of order J. However, in this case $\bar{\chi}^{zz}$ is no longer peaked at $q\sim Q$, and we have to consider the combined effect of $\bar{\chi}^{zz}$ and $\bar{\chi}^{\rho\rho}$. They turn out to nearly cancel each other. The combined attractive contribution from the $\bar{\chi}^{zz}$ and $\bar{\chi}^{\rho\rho}$ channels is of order $J(t/U)^2$, which is again overshadowed by $\bar{\chi}^{+-}$, which produces a repulsive interaction of order J.

The cancellation between $\bar{\chi}^{zz}$ and $\bar{\chi}^{\rho\rho}$ is a natural consequence of nearly perfect long-range order. Close to the transition to the disordered state, the longitudinal fluctuations are quite soft, and $\bar{\chi}^{zz}$ acquires a well-defined peak at (π,π). In this case, the $\bar{\chi}^{\rho\rho}$ channel no longer reduces the interaction in the longitudinal channel, thus leaving the possibility for a competition between magnetic and superconducting instabilities.

We found that a magnetic instability may occur well before the long-range order is destroyed. In principle, this leaves open the possibility of a superconducting instability at larger doping concentrations. Without further belaboring this point, we merely note that the superconducting phase, if it is caused by magnetic fluctuations, should have $d_{x^2-y^2}$ symmetry.[48,49] The gap in such a

state vanishes near $(\pm\pi/2,\pm\pi/2)$, which is a natural way to reduce the repulsion which comes from the $\bar{\chi}^{+-}$ channel and is peaked at those points in the Brillouin zone.

ACKNOWLEDGMENTS

We would like to thank Eduardo Gagliano and David Pines for useful discussions and Boris Shraiman and Eric Siggia for sending us a copy of their work prior to publication. This work was supported by Grant No. DMR 88-09854 through the Science and Technology Center for Superconductivity at the University of Illinois at Urbana-Champaign.

APPENDIX: HOLE-HOLE INTERACTION DUE TO SPIN-WAVE EXCHANGE

A key ingredient in the calculations at nonzero doping is the interaction between two holes in an AFM background due to exchange of spin waves. This interaction has a somewhat involved structure in both the momentum and spin variables. This has resulted in some confusion in the literature with regard to such seemingly innocent properties of the interaction as its overall sign and the structure of the superconducting gap equation that it might lead to. We clarify the issue of the hole-hole interaction in this appendix.

The existing derivations of the interaction have been performed both in the RPA (Ref. 2) and the t-J model (Ref. 11) formalisms. We will focus on the RPA approach of Ref. 2. Equation (3.8) of that paper is the expression for the effective hole-hole interaction due to exchange of spin fluctuations. Rewritten slightly to conform to the notation of our paper, it reads

$$H_{+-}=-\frac{1}{N}\sum_{kk'}{}'\sum_{\substack{\alpha\alpha'\\ \beta\beta'}}[V_{+-}(k-k')n^2(k,k') - V_{+-}(k-k'+Q)p^2(k,k')] \times\sigma^{+}_{\alpha'\alpha}\sigma^{-}_{\beta'\beta}d^{\dagger}_{k'\alpha'}d^{\dagger}_{-k'\beta'}d_{-k\beta}d_{k\alpha}\ , \quad \text{(A1)}$$

where the coherence factors $p(k,k')$ and $n(k,k')$ are

$$n(k,k')=u_k u_{k'}-v_k v_{k'}\ , \qquad p(k,k')=u_k v_{k'}-v_k u_{k'}\ , \quad \text{(A2)}$$

and u_k, v_k are defined in Eq. (8) of the main text. Also note that the momenta of incoming and outgoing particles are chosen to be opposite, as one does when analyzing the pairing aspects of the problem. This special choice is sufficient for analyzing the sign of the interaction.

In Eq. (A1) all the momenta lie in the reduced Brillouin zone. A cavalier "analytic continuation" of this expression to a full Brillouin zone can result in errors.

To analyze the sign of the interaction, consider the process shown in Fig. 7. It corresponds to a matrix element for a "hop" of a two-particle state, $(k\downarrow,-k\uparrow)\rightarrow(k'\uparrow,-k'\downarrow)$. Both k and k' are inside the reduced Brillouin zone. In Fig. 7(a), k and k' are both near $(\pi/2,\pi/2)$, while in Fig. 7(b), k is near $(-\pi/2,-\pi/2)$. This corresponds to the momentum

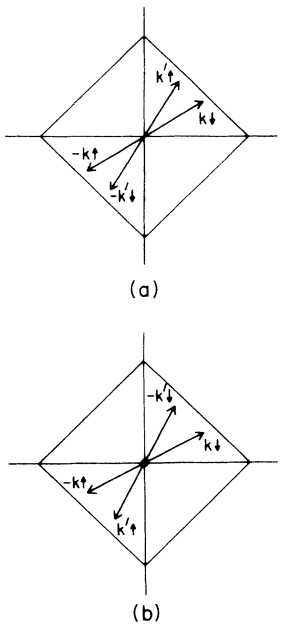

FIG. 7. The hole-hole interaction due to the spin-wave exchange in the reduced Brillouin zone representation. When the hole band minima are near $(\pm\pi/2,\pm\pi/2)$, we have to distinguish two kinds of processes. In (a) k and k' are both near $(+\pi/2,+\pi/2)$ and the process $|k\uparrow,-k\downarrow\rangle\rightarrow|k'\downarrow,-k'\uparrow\rangle$ implies the momentum transfer $q=k'-k\simeq Q$, where $Q=(\pi,\pi)$. In (b) k is near $(+\pi/2,+\pi/2)$ but k' is near $(-\pi/2,-\pi/2)$. The same process as in (a) now has $q\simeq 0$.

transfers $q=k-k'$ of nearly zero and $Q=(\pi,\pi)$, respectively. The dominant terms in H_{+-} in the two cases are $+V_{+-}(q+Q)p^2(k,k')$ and $-V_{+-}(q)n^2(k,k')$, since $V_{+-}(p)$ is peaked at $p=Q$. A rash interpretation would suggest that "the sign of the interaction" differs for the two processes. This is not a viable viewpoint, as we will see shortly.

Note that $(\pi/2,\pi/2)$ and $(-\pi/2,-\pi/2)$ are physically identical points. Thus, in the processes of Fig. 7, all the initial and final momenta stay near each other in the physically relevant sense. However, the operators corresponding to these nearby points do not evolve continuously when crossing the Brillouin zone boundary. To see this, consider the valence band operators that are obtained by inverting Eq. (7) of the main text,

$$d_{k\uparrow}=v_k a_{k\uparrow}+u_k a_{k+Q\uparrow}\ , \qquad d_{k\downarrow}=v_k a_{k\downarrow}-u_k a_{k+Q\downarrow}\ . \quad \text{(A3)}$$

Under the replacement $k\rightarrow k+Q$ in these equations, the up- and down-spin operators transform differently,

$$d_{k\uparrow}=+d_{k+Q\uparrow},\quad d_{k\downarrow}=-d_{k+Q\downarrow}\ . \quad \text{(A4)}$$

In particular, we have $d_{(\pi/2,\pi/2),\downarrow}=-d_{(-\pi/2,-\pi/2),\downarrow}$. Thus, we have opposite signs for two operators which represent exactly the same physical state. The choice of this sign depends on which side of the Brillouin zone we approach that physically unique point. We now show that such a sign change modifies the way in which one should interpret the interaction properties.

The general information about the "sign of the interaction" usually comes from considering its real space form, $V(r)$. To connect the real-space and momentum-space formulations, it is convenient to consider well-localized wave packets, made out of the valence band fermions, whose band minima are at $(\pm\pi/2,\pm\pi/2)$. The key point is that such a wave packet will not have a smooth envelope in the momentum space because of the sign discontinuity of the fermion operators.

As an illustration, consider a particle obeying a Schrödinger equation in one dimension. A typical wave packet would be a Gaussian in k space, $f(k)\sim e^{-a^2k^2}$. However, this standard result applies to the usual choice of the momentum eigenstates, $\psi_k(x)=e^{ikx}$. Let us now perform a "gauge" (in this case, sign) transformation and define, $\tilde{\psi}_k(x)=\mathrm{sgn}(k)e^{ikx}$. Then the envelope for a localized wave packet has to be $\tilde{f}(k)\sim\mathrm{sgn}(k)e^{ikx}$. Were we to use $\tilde{f}(k)\sim e^{-a^2k^2}$, we would *not* obtain a localized wave packet. This is exactly what happens in Eq. (A1). We have a discontinuous (in the sense explained above) set of basis states, and thus cannot simply take the signs in front of $V_{+-}(q)$ and $V_{+-}(q+Q)$ as indicative of the attractive or repulsive nature of the interaction. Furthermore, the reduced Brillouin zone, the Fermi surface at, say, $(\pm\pi/2,\pm\pi/2)$, which is physically a single object, is broken up into two semicircular pieces near $(\pi/2,\pi/2)$ and $(-\pi/2,-\pi/2)$. It would be best to work in the extended zone, and confine our attention to a single (i.e., circularly shaped) Fermi surface at, say, $(\pi/2,\pi/2)$.

This is accomplished by a transformation to the extended zone, which we analyze separately for the situations shown in Figs. 7(a) and 7(b). In the case of Fig. 7(a), we wish to replace the operators acting near $(-\pi/2,-\pi/2)$ by those acting near $(\pi/2,\pi/2)$, i.e., we want to substitute $d_{-k},\uparrow$ and $d^{\dagger}_{-k'},\downarrow$ for $d_{-k+Q},\uparrow$ and $d^{\dagger}_{-k'+Q},\downarrow$. Since we simultaneously shift two opposite spin operators, by Eq. (84) we pick up a minus sign, which should be absorbed into the V_{+-} term. Thus, the effective interaction now is $-V_{+-}(q+Q)p^2(k,k')$, where k,k' and $q=k'-k$ remain unchanged.

In the case of Fig. 7(b), it is $-k$ and k' momenta that are near $(-\pi/2,-\pi/2)$. Thus, in this case we have to replace $d_{-k},\uparrow$ and $d^{\dagger}_{k'},\uparrow$ by $d_{-k+Q},\uparrow$ and $d^{\dagger}_{k'+Q},\uparrow$. The difference with the previous situation is that we are now moving two *identical* spins, and thus no sign change is needed in front of the effective interaction $-V_{+-}(q)n^2(k,k')$. However, this time we moved $+k'$, and we ought to redefine this expression in terms of $k'_{\mathrm{new}}=k'+Q$ and $q_{\mathrm{new}}=k'_{\mathrm{new}}-k=q+Q$. Making the required substitutions, we obtain

$$-V_{+-}(q)n^2(k,k')=-V_{+-}(q_{\mathrm{new}}+Q)p^2(k,k'_{\mathrm{new}}) ,$$

where the equality $n(k,k')=p(k,k'_{\mathrm{new}})$ can be verified from Eq. (A2). We will omit the subscript "new" in the following.

We see that the leading terms in the interaction Hamiltonian near $(\pi/2,\pi/2)$ are now the same in both cases,

$$H_{+-}\simeq -V_{+-}(q+Q)p^2(k,k')\sigma^{+}_{\alpha'\alpha}\sigma^{-}_{\beta'\beta} \times d^{\dagger}_{k'\alpha'}d^{\dagger}_{-k'\beta}d_{-k\beta}d_{k\alpha} . \quad \text{(A5)}$$

We stress that "$-k$" is now defined as an opposite of k with respect to $(\pi/2,\pi/2)$, and as a result, we have $q\ll 1$ in Eq. (A5).

We have thus disposed of the unusual momentum dependence of the initial expressions. We still do not have an entirely conventional situation because of the peculiar spin dependence in Eq. (A5). In particular, it is not yet time to conclude that this is an attractive interaction. In the usual basis of "up" and "down" hole spins, our interaction has no *diagonal* matrix elements. Furthermore, it is not invariant with respect to spin-space rotations.

As has been shown in Ref. 7, greater clarity is achieved by going into real-space representation. In Eq. (A5) there are no discontinuities in the definition of the fermion operators, and an ordinary Fourier transform can be taken without any need for further interpretation. The final answer is

$$V_{\mathrm{eff}}=-\bar{V}(r_1-r_2)(\sigma_1^{+}\sigma_2^{-}+\sigma_1^{-}\sigma_2^{+}) , \quad \text{(A6)}$$

where $\bar{V}(r_1-r_2)$ is a Fourier transform of the interaction in Eq. (A5).

The Fourier transform is analyzed in Ref. 7, and the result is

$$V_{\mathrm{eff}}\sim -\left[\delta(r_1-r_2)+\frac{1}{2\pi}\frac{2xy}{r^4}\right](\sigma_1^{+}\sigma_2^{-}+\sigma_1^{-}\sigma_2^{+}) . \quad \text{(A7)}$$

When the interaction is written in this form, we can clearly see that it is diagonalized not in the "up-down" spin state basis, but in the basis made of the spin-symmetric and spin-antisymmetric states, which are $1/\sqrt{2}(|\uparrow\downarrow\rangle+|\downarrow\uparrow\rangle)$ and $1/\sqrt{2}(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle)$, respectively. The interaction does not couple to the $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ states at all. This is consistent with the observation we made earlier that it is not rotationally invariant in spin space.

We note that the sign of the interaction depends on the relative orientation of holes, as seen from Eq. (A7) because of a long-range dipolar piece in it. Shraiman and Siggia argued that one could simply drop this term and work with the δ-function term alone.[7,50] This is based on the expectation that when electrons form a degenerate Fermi sea, this term gets averaged out.[50] We follow Ref. 11 and drop the dipolar term. Actually, one can raise questions about the stability of the spiral phase against quantum fluctuations that might be induced by the dipolar coupling, but we do not address them in this paper.

Once we decide to keep the δ-function term only, we see that it actually does not have any effect when the two holes are in the spin-symmetric state, since due to the Pauli exclusion principle they cannot then be at one point in space at the same time.

The only effect left is a *repulsion* between two holes in the spin-antisymmetric channel, since the eigenvalue of $(\sigma_1^{+}\sigma_2^{-}+\sigma_1^{-}\sigma_2^{+})$ is then equal to -1. In an abuse of language we could say that "opposite spin" holes repel each other, which makes the situation similar to Stoner model of ferromagnetism, an analogy emphasized in Refs. 10 and 11. Thus there is no attraction due to the transverse spin fluctuations, and therefore no superconductivi-

ty from this interaction in the usual spin-antisymmetric channel.

We conclude with a comment on the results of Ref. 31, in which they obtain an effective interaction between holes with the sign opposite of that in Eq. (A7). This result was obtained by using the *negative* static susceptibility $\bar{\chi}^{+-}$ to mediate interactions. The negative susceptibility $\bar{\chi}^{+-}$ was a result of considering the system below the spiral magnetic instability, but not in its proper new ground state. We believe this procedure is not well justified. The negative susceptibility means that the magnetic fluctuations used to mediate interactions are exponentially growing in time. Instead, the static susceptibility in a stable state should remain positive, and at sufficiently low doping levels we expect the interaction between two holes in the new ground state to be nearly the same as in Eq. (A7). Moreover, the magnitude of the interaction, which is of order J, as we argued in Sec. IV, is also different from that in Ref. 31, but agrees with Ref. 10.

[1] See, e.g., P. W. Anderson and J. R. Schrieffer, Phys. Today **44** (6) 54 (1991), and references therein.

[2] J. R. Schrieffer, X. G. Wen, and S. C. Zhang, Phys. Rev. B **39**, 11 663 (1989).

[3] L. N. Bulaevskii and D. I. Khomskii, Zh. Eksp. Teor. Fiz. **52**, 1603 (1967) [Sov. Phys. JETP **25**, 1067 (1967)].

[4] W. Brinkman and T. M. Rice, Phys. Rev. B **2**, 1324 (1970).

[5] K. A. Chao, J. Spałek, and A. M. Oleś, J. Phys. C **10**, L271 (1977); Phys. Rev. B **18**, 3453 (1978).

[6] C. Gros, R. Joynt, and T. M. Rice, Phys. Rev. B **36**, 381 (1987).

[7] D. M. Frenkel and W. Hanke, Phys. Rev. B **42**, 6711 (1990).

[8] A. W. Overhauser, Phys. Rev. **128**, 1437 (1962).

[9] D. R. Penn, Phys. Rev. **142**, 350 (1966).

[10] B. I. Shraiman and E. D. Siggia, Phys. Rev. Lett. **62**, 1564 (1989).

[11] B. I. Shraiman and E. D. Siggia, Phys. Rev. B **40**, 9162 (1989).

[12] T. Dombre, J. Phys. (Paris) **51**, 847 (1990).

[13] H. J. Schulz, Phys. Rev. Lett. **64**, 1445 (1990); **65**, 2462 (1990).

[14] C. Jayaprakash, H. R. Krishnamurthy, and S. Sarker, Phys. Rev. B **40**, 2610 (1989).

[15] B. Chakraborty, N. Read, C. Kane, and P. A. Lee, Phys. Rev. B **42**, 4819 (1990).

[16] A. Singh and Z. Tešanović, Phys. Rev. B **41**, 614 (1990).

[17] C. L. Kane, P. A. Lee, T. K. Ng, B. Chakraborty, and N. Read, Phys. Rev. B **41**, 2653 (1990).

[18] J. Gan, N. Andrei, and P. Coleman, J. Phys. Condens. Matter **3**, 3537 (1991).

[19] J. Gan and F. Mila, Phys. Rev. B **44**, 12 624 (1991).

[20] Z. Y. Weng and C. S. Ting, Phys. Rev. B **42**, 803 (1990).

[21] V. M. Galitskii, Zh. Eksp. Teor. Fiz. **7**, 151 (1958) [Sov. Phys. JETP **34**, 104 (1958)].

[22] L. B. Ioffe and A. I. Larkin, Int. J. Mod. Phys. B **2**, 203 (1988).

[23] N. Read and S. Sachdev, Phys. Rev. Lett. **66**, 1773 (1991).

[24] F. Mila, D. Poilblanc, and C. Bruder, Phys. Rev. B **43**, 7891 (1991).

[25] A. Chubukov, Phys. Rev. B **44**, 392 (1991).

[26] J. E. Hirsch and D. J. Scalapino, Phys. Rev. Lett. **56**, 2732 (1986).

[27] B. I. Halperin and P. C. Hohenberg, Phys. Rev. **188**, 898 (1969).

[28] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin/Cummings, Reading, MA, 1975).

[29] P. W. Anderson, Phys. Rev. **86**, 694 (1952).

[30] V. G. Vaks, A. I. Larkin, and S. A. Pikin, Zh. Eksp. Teor. Fiz. **53**, 281 (1967) [Sov. Phys. JETP **26**, 188 (1968)].

[31] A. Singh, Z. Tešanović, and Ju. H. Kim, Phys. Rev. B **44**, 7757 (1991).

[32] V. Elser, D. Huse, B. I. Shraiman, and E. D. Siggia, Phys. Rev. B **41**, 6715 (1990).

[33] S. Trugman, Phys. Rev. B **37**, 1597 (1988).

[34] S. Trugman, Phys. Rev. B **41**, 892 (1990).

[35] S. Sachdev, Phys. Rev. B **39**, 12 232 (1989).

[36] G. Vignale and M. R. Hedayati, Phys. Rev. B **42**, 786 (1990).

[37] Actually, defining the effective masses for small U is somewhat difficult. Close to the minima, one may use the expansion of Eq. (59) and obtain $m_\perp \sim \Delta/t^2$. However, the expansion holds only up to $|\pi/2-k| \sim \Delta/t$. For larger deviations from the minima, the whole spectrum is linear in $|\pi/2-k|$, and so $m_\parallel$ by no means coincides with the inverse bandwidth. The longitudinal mass is formally infinite in Eq. (59), but the self-energy corrections make $m_\perp$ finite. These corrections are of order Δ^2/t^2 and so $m_\parallel \sim 1/\Delta$. As a result, the geometrical average $\sqrt{m_\parallel m_\perp}$ scales as the inverse bandwidth of the hole spectrum.

[38] E. Dagotto, R. Joynt, A. Moreo, S. Bacci, and E. Gagliano, Phys. Rev. B **41**, 9049 (1990).

[39] B. I. Shraiman and E. D. Siggia, Phys. Rev. Lett. **61**, 467 (1988).

[40] We also note that this procedure has a parallel with some of the formal manipulations in Landau's Fermi-liquid theory (L. D. Landau, Zh. Eksp. Teor. Fiz. **35**, 97 (1959) [Sov. Phys. JETP **8**, 70 (1959)]). In the derivation of that theory, one splits a product of two Green's functions from a single bubble into "singular" and "regular" parts, $G(p)G(p+q) = (2\pi i Z^2/v)[\mathbf{v}\cdot\mathbf{q}/(\omega-\mathbf{v}\cdot\mathbf{q})]\delta(v)\delta(|p|-p_0)+\phi_{\rm reg}(p)$. Then $\phi_{\rm reg}(p)$ is summed up, via a ladder with "irreducible vertices" serving as interactions, to result eventually in the interaction function $f(p,p')$ acting on a singular part of the Green's function product, i.e., on the quasiparticle poles of $G(p)$. A similar philosophy is used in Galitskii's work on dilute Fermi gases (see Ref. 21). In that case, the product of two Green's functions in a medium is rewritten as $GG = G^{(0)}G^{(0)} + (GG - G^{(0)}G^{(0)})$ where $G^{(0)}$ are the vacuum Green's functions. When the term $G^{(0)}G^{(0)}$ is summed up via a ladder, one obtains a vacuum T matrix instead of the bare potential acting as an interaction. Thus, the procedure we employ appears to be rather general.

[41] B. I. Shraiman and E. D. Siggia, Phys. Rev. B **42**, 2485 (1990).

[42] C. L. Kane, P. A. Lee, and N. Read, Phys. Rev. B **39**, 6880 (1989).

[43] A. M. Afanas'ev and Yu. Kagan, Zh. Eksp. Teor. Fiz. **43**, 1456 (1962) [Sov. Phys. JETP **16**, 1030 (1963)].

[44] J. C. Campuzano *et al.*, Phys. Rev. Lett. **64**, 2308 (1990).

[45] B. I. Shraiman and E. D. Siggia, Phys. Rev. B **46**, 8305 (1992).

[46] A. Auerbach and B. E. Larson, Phys. Rev. B **43**, 7800 (1991).

[47]Z. Y. Weng, Phys. Rev. Lett. **66**, 2156 (1991).
[48]J. Annett, N. Goldenfeld, and S. Renn, in *Physical Properties of High Temperature Superconductors,* edited by D. M. Ginsberg (World Scientific, Singapore, 1990).
[49]P. Montoux, A. V. Balatsky, and D. Pines, Phys. Rev. Lett. **67**, 3448 (1991).
[50]B. I. Shraiman (private communication).