Gapless fermions and gauge fields in dielectrics

L. B. Ioffe and A. I. Larkin
Landau Institute for Theoretical Physics, 117940, Kosygina 2, Moscow V-334, U.S.S.R.
(Received 20 June 1988; revised manuscript received 28 November 1988)

To study the nonmagnetic dielectric state and Mott transitions we consider an example of a two-dimensional modified Hubbard model with a large number of colors. Low-energy excitations in this phase are fermionic excitations and Bose excitations described by gauge fields of the U(1) group. The transition into the metal state has little effect on the fermionic spectrum, but it results in the local U(1) symmetry being broken and fermions becoming able to transfer charge excitations. Apart from the half-filling, scalar Bose excitations also appear. Due to the presence of additional gauge fields the physical conductivity is determined by the lowest conductivity of the Fermi or Bose subsystems.

I. INTRODUCTION

Interest in the Mott-Hubbard metal-dielectric transition has been revived recently by the discovery of high-temperature superconductivity in copper oxides.[1] Whereas the properties of metals can be described (at least phenomenologically) by the Fermi-liquid theory, the properties of the dielectric are still beyond complete understanding. Diverse dielectric phases can exist with the choice being governed by the microscopic interaction constants. In the simplest case this phase is a Néel antiferromagnet. The nature of a dielectric state without any magnetic long-range order is a far more complicated and long-standing problem. The first formulation of this problem dates back to 1941 when Pomeranchuk[2] supposed that in a $S=\frac{1}{2}$ Heisenberg antiferromagnet the long-range order is destroyed by quantum fluctuations and in the resulting spin liquid the excitations are gapless fermions. Later and independently Anderson[3] hypothesized that the appropriate reference state for both the $S=\frac{1}{2}$ Heisenberg antiferromagnet and the Mott-insulating Hubbard model is a "resonating valence bond" (RVB) state, and he gave the qualitative description of its ground-state wave function. Recently Anderson conjectured[4] the importance of the RVB state for high-temperature superconductivity and proposed[4-6] a different way for its description based on the idea that in a RVB state pairing occurs between the electrons (or an electron and a hole) on adjacent sites. We note in passing that the analogous pairing between electrons on the adjacent wires of quasi-one-dimensional superconductors was introduced and studied in.[7] In this paper we discuss the properties of the dielectric RVB state and the Mott-transition RVB metal employing the simplest model. It is generally accepted that the Mott transition is described by the Hubbard model:

$$H_H=\sum_{i,j} t_{ij}c_{i\alpha}^{\dagger}c_{j\alpha}+\sum_i (\tfrac{1}{2}Un_i^2-\mu n_i),\quad n_i=c_{i\alpha}^{\dagger}c_{i\alpha}\ , \tag{1}$$

where $\alpha=1,2$ stands for the electron spin index, μ is the chemical potential which should be chosen so that $\langle n_i\rangle=n_0$, where n_0 is the mean number of electrons per site ($n_0=1$ corresponds to a half-filled band), and t_{ij} is the tunneling amplitude between sites i,j. In the interaction-constant region $U/t\gg 1$ (dielectric state) the Hubbard model (1) is equivalent to the Heisenberg antiferromagnet with a spin $S=\frac{1}{2}$:[8]

$$H_{\mathrm{AF}}=\sum_{i,j}\mathcal{T}_{ij}S_i^{(p)}S_j^{(p)},\quad \mathcal{T}_{ij}=\frac{2t_{ij}^2}{U},\quad S_i^{(p)}=\tfrac{1}{2}c_{i\alpha}^{\dagger}\sigma P_{\alpha\beta}c_{i\beta}\ . \tag{2}$$

Quantal fluctuations in this antiferromagnet are strong enough that only numerical simulations can provide reliable investigations of its ground state. If the spin S is large ($S\gg 1$) quantal fluctuations are weak and the ground state is a Néel antiferromagnet. The antiferromagnetic interaction of the next-nearest neighbors frustrates the interaction and results in the helicoid antiferromagnet. At the boundary between the regions of Néel and helicoid antiferromagnets quantal fluctuations are strong: They destroy long-range magnetic order and form a spin-liquid state[9] in the exponentially small [$\exp(-1/S)$] region. Presumably this region becomes rather large for small S, but the absence of a small parameter prevents the analysis in this case. Therefore, in this paper we study the modification of the original Hubbard model (1), where the index α runs over N values $\alpha=1,\ldots,N$. Mainly (if not stated explicitly otherwise) we discuss the case $n_0=N/2$ (N even) corresponding to a half-filled band. Generalization for the case $n_0\sim N/2$, n_0 an integer, is straightforward and does not present any difficulties. The case of a fractal n_0 (doped material) is more subtle. We consider it in Sec. V, in the discussion of the Mott transition and in Sec. VIII, where we discuss the influence of doping on electromagnetic properties. In the study of the Mott transition it will be more convenient for us to study the generalization of the Hubbard model (1) with an additional four-fermion interaction:

$$H=H_H-\sum_{i,j} I'_{ij}(c_{i\alpha}^{\dagger}c_{j\alpha})(c_{j\beta}^{\dagger}c_{i\beta})\ . \tag{3}$$

The term of the same form (with $I_{ij}=t_{ij}^2/U$) appears in the second-order perturbation expansion over U in the

original model (1). Recently this model was studied[10,11] within the framework of the mean-field approximation (MFA).

In the MFA the order parameter $\Delta_{ij}=\langle c^{\dagger}_{i\alpha}c_{j\alpha}\rangle$ is formed at low temperatures and fermionic excitations are present in this model. Various phases and states are possible; the exact phase is determined by the interaction constants (t_{ij},I_{ij}) of the model (3). In the molecular crystal phase the order parameter differs from zero only for one bond per each site; in this phase the lattice period doubles and fermionic excitations acquire a gap. In the other uniform phase the order parameter does not depend on the bond position and fermions fill the Fermi sea to the Fermi surface. In the third phase (flux phase), which was proposed in Refs. 10 and 12, the amplitude of the order parameter also does not depend on the bond position, whereas the sum of the phases ϕ_{ij} $[\Delta_{ij}=|\Delta|\exp(i\phi_{ij})]$ over the smallest plaquette is π. In this phase fermions have a linear (relativistic) spectrum, and their Fermi surface collapses to a point.

The lowest order (over N^{-1}) and the most important corrections to the effective action obtained in the MFA are due to the quantal fluctuations of the phase ϕ_{ij} of the order parameter. These fluctuations can be described by two fields: the field ψ_i (defined at a lattice point i) and the field a_{ij} (defined at bonds ij) with the latter being an analog of the vector potential in the lattice electrodynamics. The Mott-transition metal dielectric corresponds to the disappearance of $\langle\cos\psi_i\rangle$ and restoration of the gauge U(1) symmetry: $\psi_i\rightarrow\psi_i+\varphi_i$. In the dielectric state the symmetry is restored, and the field a_{ij} becomes (at least in the long-wavelength limit) a genuine gauge field. The averaging over fluctuations of ψ_i results in the exponential decrease of the fermionic Green's functions corresponding to the charge transfer at large distances, as it should be in the dielectric state. However, the spectrum of the fermionic excitations and fermion correlation functions that correspond to the processes without the charge transfer remain qualitatively the same as they were in the corresponding phase before averaging over ψ_i. Thus, in this state the long-wave excitations are fermions and have a U(1) gauge field. In the metal state the symmetry U(1) is broken and the averaging over ψ_i does not result in such dramatic changes in the behavior: The spectrum of the fermionic excitations and their Green's functions do not change qualitatively, the gauge field acquires a mass, and its interaction with fermions is small so that the long-wavelength excitations in this state are fermions near the Fermi surface, as it should be. In spite of being small the fluctuations of the gauge field a_{ij} turn out to be very important for electromagnetic properties since in the state in which the gauge U(1) symmetry is restored they screen out the real electromagnetic field A.

The plan of the paper is as follows: in Sec. II we discuss the mean-field solution and possible phases, in Sec. III the fermionic spectrum in these phases, in Sec. IV we study the main influence of the phase fluctuations (ψ fluctuations) in the dielectric phase ($U/t\gg1$), in Sec. V we discuss how these fluctuations result in the Mott transition, in Sec. VI we return to the dielectric state and study the influence of the a-field fluctuations, in Sec. VII we discuss the role of instantons of a field, and, finally, in Sec. VIII we discuss the electromagnetic properties of the phases with a low density of holes.

II. DIELECTRIC PHASES

We start with the dielectric state at large U. In the limit $U\rightarrow\infty$ the number of electrons at each site is fixed and is equal to n_0, and interaction between them arises in the first order of the perturbation series in t/U. The resulting Hamiltonian in the subspace of states with exactly n_0 electrons per site has the form

$$\begin{aligned}H&=-\sum_{i,j}I_{ij}(c^{\dagger}_{i\alpha}c_{j\alpha})(c^{\dagger}_{j\beta}c_{i\beta})\\&\quad-\tilde{\mu}\sum_i c^{\dagger}_{i\alpha}c_{i\alpha}+i\sum_i\varphi_i(n_i-n_0)\ ,\\I_{ij}&=I'_{ij}+t^2_{ij}/U\ ,\end{aligned}\tag{4}$$

where to ensure the condition $n_i=n_0$ we introduce an auxiliary field variable φ_i. Using the representation of the density matrix $\exp(-\beta H)$ in the path-integral form we perform a Hubbard-Stratanovich transformation and introduce other auxiliary Hermitian fields Δ_{ij} which are defined on each bond ij with $t_{ij}\neq0$:

$$\begin{aligned}\rho&=\exp(-\beta H)=\int\mathcal{D}\Delta_{ij}\exp\left[-\int_0^{\beta}L\,d\tau\right]\ ,\\L\{\Delta\}&=\sum_{i,j}(\tfrac{1}{4}|\Delta_{ij}|^2I^{-1}_{ij}-\Delta^{*}_{ij}c^{*}_{i\alpha}c_{j\alpha})\\&\quad+\sum_i[(i\varphi_i-\tilde{\mu})n_i-i\varphi_in_0]\ .\end{aligned}\tag{5}$$

In the leading mean-field approximation the functional integration over Δ_{ij} is replaced by a simple substitution in the Hamiltonian (5); the value of Δ_{ij} in the saddle point

$$\Delta_{ij}=2I_{ij}\langle\bar{c}_{i\alpha}c_{j\alpha}\rangle\ ,\tag{6}$$

where $\langle\cdots\rangle$ means averaging over fermionic degrees of freedom with the density matrix

$$\rho_f=\exp\left[-\int_0^{\beta}L_F d\tau\right]\ ,\tag{7}$$

$$L_F=\sum_i\bar{c}_{i\alpha}\partial_\tau c_{i\alpha}-\sum_{i,j}\Delta^{*}_{ij}\bar{c}_{i\alpha}c_{j\alpha}+\sum_i[n_i(i\varphi_i-\tilde{\mu})-i\varphi_in_0]\ .$$

If the value of chemical potential $\tilde{\mu}$ is chosen so that the mean number of electrons per site equals n_0 even at $\varphi=0$, then, as we show below (Secs. IV and VI), the fluctuations of the field φ are small and do not change the solution of Eqs. (6) and (7). Meanwhile, we neglect these fluctuations and discuss the properties of solutions of (6) and (7).

At high temperatures the solution is unique and trivial: $\Delta_{ij}=0$. At lower temperatures other solutions appear, and the critical temperature T_c can be determined by inserting, into the right-hand side (rhs) of (6), the Green's function expression:

$$\hat{D}=2NT\sum_n[i\omega_n+\tilde{\mu}+\hat{\Delta}]^{-1}\ ,\tag{8}$$

where $\hat{\Delta}$ is the matrix Δ_{ij}, and $\hat{D}$ is the matrix $D_{ij}=\Delta_{ij}I_{ij}^{-1}$. Expanding the rhs over $\hat{\Delta}$ we get the equation for the critical temperature

$$\Delta_{ij}=[I_{ij}N/(2T_c)]\Delta_{ij} \ . \tag{9}$$

Equation (9) allows many solutions. The concrete form of the solution should be determined from the condition of the minimum of free energy. In the vicinity of the transition point the free energy can be expanded over powers of Δ and equals

$$F=\tfrac{1}{4}\sum_{i,j}\left[I_{ij}^{-1}-\frac{N}{2T}\right]|\Delta_{ij}|^2+\frac{N}{192T_c^3}\sum_{ijke}\Delta_{ij}\Delta_{jk}\Delta_{kl}\Delta_{li} \ . \tag{10}$$

In the simplest case the interaction $I_{ij}\neq 0$ only for the nearest neighbors, then the minima of the energy (10) corresponds to three solutions:[10,12] $\Delta_{ij}=\Delta_u$ for all bonds between the nearest neighbors (uniform phase), $\Delta_{ij}=\Delta_F\exp(i\phi_{ij})$ with $\phi_{ij}+\phi_{jk}+\phi_{kl}+\phi_{li}=\pi$ (flux phase), and $\Delta_{ij}=\Delta_p$ for only one bond per each site (molecular crystal phase) (Fig. 1). Inserting the forms of the solutions into the free energy (10) and minimizing over $|\Delta|$ we get the energies of these phases per site:

$$\begin{aligned} E_U&=-(1/6)I\tau^2N^2 \ , \\ E_F&=-(3/10)I\tau^2N^2 \ , \\ E_p&=-(3/8)I\tau^2N^2 \ , \end{aligned} \tag{11}$$

where $\tau=(T_c-T)/T$. The molecular phase has the lowest energy of all. In the lowest-order approximation over N^{-1} which we consider here this phase is degenerate with respect to different positions of the valence bonds. This degeneracy is lifted by the next orders in N^{-1} if the next orders in t/U are also taken into account in agreement with the result of the paper[13] in which it was shown that degeneracy of the energy of the trial wave function corresponding to the molecular phase is lifted if the next order over t/U is taken into account. In the molecular phase the period of the lattice doubles and fermionic excitations have a large gap. Other phases are far more interesting. These phases can have a lower (than the molecular phase) energy if the interaction (4) has a more complex form than the nearest-neighbors interaction that leads to (11). For instance, the next-nearest-neighbor interaction lowers the energy of the uniform phase, leaving the energies of the flux phase and molecular phase intact, whereas the interactions of more distant sites belonging to different sublattices lowers the energy of both flux and uniform phases (note that interaction between the sites of the same sublattice has no influence on the energy of the flux phase). A low ratio n_0/N (with n_0,N being integers) also favors the uniform phase which becomes more stable than the Peierls phase at $n_0/N\leq\alpha_{c1}\simeq 0.39$, and the most stable phase at $n_0/N\leq\alpha_{c2}\simeq 0.27$. In this connection we discuss the influence of the next-order over t/U correction to the Hamiltonian (4).

We consider as before the case of $t_{ij}=t\neq 0$ only for the nearest neighbors. The first correction to (4) arises in the third order of the perturbation expansion and equals

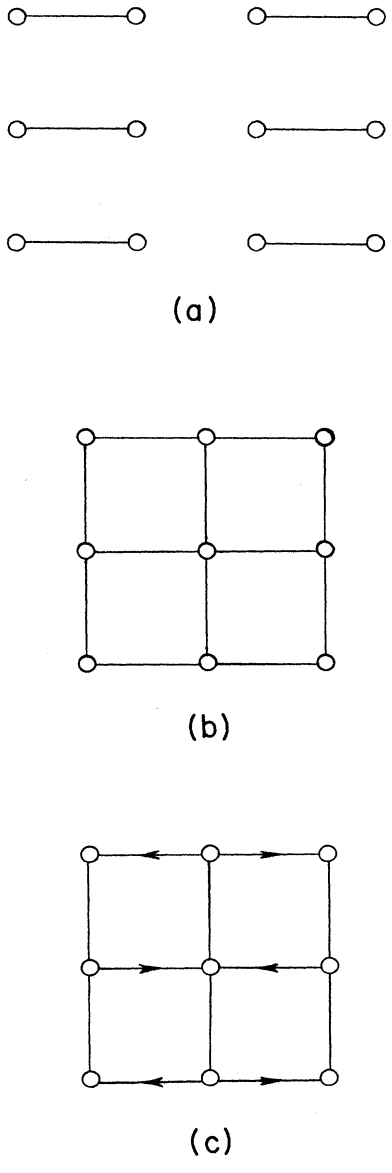

FIG. 1. Different ways of pairing: ○——○ stands for $\Delta_{ij}=\Delta_0$, ○⟶○ stands for $\Delta_{ij}=i\Delta_0$. (a) molecular crystal phase, (b) uniform spin liquid phase, (c) flux phase of spin liquid.

$$V^{(4)}=-\frac{t^4}{U^3}\left[5\sum_{[ijkl]}(c^\dagger_{i\alpha}c_{j\alpha})(c^\dagger_{j\beta}c_{k\beta})(c^\dagger_{k\gamma}c_{l\gamma})(c^\dagger_{l\delta}c_{i\delta})+\tfrac{4}{3}\sum_{i,j\neq k}(c^\dagger_{i\alpha}c_{j\alpha})(c^\dagger_{j\beta}c_{i\beta})(c^\dagger_{i\gamma}c_{k\gamma})(c^\dagger_{k\delta}c_{i\delta})-\tfrac{7}{4}\sum_{i,j}(c^\dagger_{i\alpha}c_{j\alpha})^2(c^\dagger_{j\beta}c_{i\beta})^2\right], \quad (12)$$

where the notation $[ijkl]$ means a sum over $ijkl$ which constitutes a plaquette.

The analogous correction for the original Hubbard model was obtained in Ref. 13 which differs slightly from Eq. (12) by the values of the coefficients in the second and the third terms since, in the present model, the virtual processes with intermediate states in which two electrons (or holes) occupy the same site are not prohibited, as they are in the Hubbard model with $N=2$. In the leading mean-field approximation in the vicinity of T_c the term (12) is proportional to Δ^4 and thus does not change the transition temperature T_c. To take into account its influence on the ground-state energy in the leading order over Δ_{ij} we can replace $c^\dagger_{i\alpha}c_{j\alpha}$ in it by its average $\langle c^\dagger_{i\alpha}c_{j\alpha}\rangle=\Delta_{ij}/(2I)$. We get

$$V^{(4)}=-\frac{U}{(2t)^4}\left[5\sum_{[ijkl]}\Delta_{ij}\Delta_{jk}\Delta_{kl}\Delta_{li}+\tfrac{4}{3}\sum_{i,j\neq k}|\Delta_{ij}|^2|\Delta_{kl}|^2-\tfrac{7}{4}\sum_{i,j}|\Delta_{ij}|^4\right]. \quad (13)$$

Adding this term to the free energy (10) and minimizing again over $|\Delta|$ we get new values of the ground-state energy

$$\begin{aligned}E_U&=-[6-49x]^{-1}I\tau^2/N^2\,,\\ E_F&=-3[10+93x]^{-1}I\tau^2N^2\,,\\ E_p&=-3[8+84x]^{-1}I\tau^2N^2\,,\\ x&=(tN/U)^2\,.\end{aligned} \quad (14)$$

Comparing the energies (14) we conclude that at large tN/U ($x\geq 10/231$) the uniform phase has the lowest energy of all.

Now we discuss the validity of our disregard of all higher-order terms in t/U at $tN/U\sim 1$. Using formula (4) we infer that the main contribution to the density matrix comes from the states with exactly n_0 electrons per site, which is justified if the excitation energy $U\gg T$, i.e., (using the estimate $T\approx T_c=Nt^2/2U$), at $N^{1/2}t/U\ll 1$, which is compatible with $x\approx 1$ at $N\gg 1$. Taking into account the virtual states with $n\neq n_0$ electrons per site results in the considered series over t/U.

The next order of this series, as can be easily verified, is proportional to Δ^6, so we conclude that in the vicinity of T_c the parameter of expansion is $Nt\Delta/U\approx x\sqrt{\tau}\ll 1$ which justifies our consideration of the third term and disregard of higher-order terms. We will come back to this point once again in our discussion of the Mott transition.

III. SPECTRUM OF FERMIONIC EXCITATIONS IN DIELECTRIC PHASES

Within the framework of the mean-field approximation the fermionic spectrum is governed by L_F (7) with fixed Δ_{ij}. In the molecular phase the energy of one-particle fermionic excitation does not depend on its momentum and equals $|\Delta_p|$. In fact, these excitations violate the conservation of a number of particles at each site and are absent. Eventually, the spectrum of excitations has a gap $2|\Delta_p|$ and begins with two-particle excitations. In the uniform phase Δ_{ij} can be chosen real and equal for equivalent bonds, and the excitation spectrum follows from L_F (7):

$$\epsilon(k)=-\sum_j\Delta_{ij}\cos[k(z_i-z_j)]-\tilde{\mu}\,. \quad (15)$$

In the case of nearest-neighbor interaction it is simplified to

$$\epsilon(k)=-2\Delta[\cos(k_x)+\cos(k_y)]-\tilde{\mu}\,. \quad (16)$$

This form of spectrum implies a nonzero density of states at the Fermi surface, which, in turn, means that the heat capacity is linear with T at low temperatures. The one-particle Green's function in this phase is

$$G_{\alpha\beta}(\omega,k)=-\int_0^\beta\sum_i\exp[i\omega\tau-k(z_i-z_j)]\langle Tc_{i\alpha}c^\dagger_{j\beta}\rangle=\frac{\delta_{\alpha\beta}}{i\omega-\epsilon(k)}\,. \quad (17)$$

In the flux phase it is convenient to choose phases of the order parameter Δ_{ij} so that $\phi_{ij}=0$ at all vertical bonds and $\phi_{ij}=\pm\pi/2$ at all horizontal bonds with signs $\pm$ alternating in chess order (Fig. 1). This choice of Δ_{ij} makes a distinction between different sublattices; therefore, it will be convenient to introduce an auxiliary index η which equals 1 in the first sublattice and 2 in the second one, and performs the Fourier transform of Fermi fields in the first and in the second sublattices separately:

$$c_{k\eta}=\sum_{ic\{\eta\}}c_i\exp(-ikz_i)\,.$$

In this notation the Hamiltonian for the Fermi fields acquires a simple form:

$$H_F=\int dk[-c^\dagger_{k\eta}\Delta_{\eta\nu}(k)c_{k\nu}-\tilde{\mu}c^\dagger_{k\eta}c_{k\eta}]\,, \quad (18)$$

where the integral over k is performed over a new Brillouin zone which is half of the original one. Generally the matrix $\Delta_{\eta\nu}$ has the form

$$\hat{\Delta}=\begin{bmatrix}\Delta_1(k) & \Delta_2(k)\\ \Delta_2^*(k) & \Delta_1(k)\end{bmatrix},$$

where $\Delta_1(k)$ is the Fourier transform of Δ_{ij} between the

sites of one sublattice, and $\Delta_2(k)$ is that between the sites of different sublattices.

Diagonalization of the matrix $\hat{\Delta}$ yields the spectrum of Fermi excitations:

$$\epsilon_{1,2}=-\Delta_1(k)-\tilde{\mu}\pm|\Delta_2(k)| \ , \tag{19}$$

which consists of two branches jointed only at points k^* at which $|\Delta_2(k^*)|=0$.

In the simplest case in which $\Delta_{ij}\neq 0$ only for the nearest neighbors

$$\Delta_1(k)=0, \quad \Delta_2(k)=2\Delta(\cos k_y+i\cos k_x) \ .$$

Thus $k^*_{1,2}=(\pm\pi/2,\pi/2)$. At half-filled band the chemical potential $\tilde{\mu}=0$, and the Fermi surface collapses to two points $k^*_{1,2}$.

In the vicinity of these points the spectrum (19) is linear, the heat capacity is proportional to T^2 at low temperatures, and the Hamiltonian (18) can be rewritten in explicitly relativistically invariant form. For this purpose we denote the pair of Fermi fields $(c_{\alpha k1},c_{\alpha k2})$ by $\chi_{\alpha k}$ and treat the regions of k space in the vicinity of points $k^*_{1,2}$ separately. In these regions we introduce a new momentum variable $p=k_1-k$ in the vicinity of k_1 or $p=k_2-k$ in the vicinity of k_2. Using this notation we get

$$H_F=2\Delta\int(dp)\bar{\chi}^\dagger_{\alpha a p}p_i\alpha_i\bar{\chi}_{\alpha a p} \ , \tag{20}$$

where $a=1,2$ denotes the regions of k space, α_i are relativistic matrices: $\alpha_1=-\sigma_2$, $\alpha_2=\sigma_1$ (σ_i are the usual Pauli matrices), and we redefine spinor field in the vicinity of k_2:

$$\bar{\chi}_{\alpha 1 p}=\chi_{\alpha 1 p}, \quad \bar{\chi}_{\alpha 2 p}=\sigma_1\chi_{\alpha 2 p} \ .$$

Hamiltonian (20) is evidently relativistically invariant. Moreover, it possesses a number of other symmetries: It is invariant under group U($2N$) which acts upon the pair of indices (αa), and it is invariant under parity transformations. Two types of parity transformations can be defined for a relativistic Hamiltonian (20): the "physical" parity transformation which transforms $k_x\to -k_x$, $k_y\to k_y$ (in the original notation) and which in the notation of (20) becomes $\bar{\chi}_{\alpha 1 p}\to\sigma_1\bar{\chi}_{\alpha 2 p'}$, $\bar{\chi}_{\alpha 2 p}\to\sigma_1\chi_{\alpha 1 p'}$, $p'=(-p_x,p_y)$; and the "unphysical" parity transformation $\bar{\chi}_{\alpha a p}\to\sigma_1\bar{\chi}_{\alpha a p'}$ which really changes particle and antiparticle. The origin of these parity transformations is quite different: The "physical" parity symmetry is always present in the initial Hamiltonian (1),(3); it can be broken only spontaneously by some order parameter, whereas the particle-antiparticle symmetry which is implied by the "unphysical" parity transformation is certainly absent away from half-filled band.

In Secs. VI–VIII we study the interaction of these fermions with gauge fields which appear naturally in this problem. To study this interaction the bare spectrum of fermions described by Hamiltonian (20) should be regularized by an infinitesimally small mass. Two types of mass terms are possible which regularize Hamiltonian (20) in quite different ways.[14] The first type of mass term has a form $m'(\bar{\chi}^\dagger_{\alpha 1 p}\alpha_0\bar{\chi}_{\alpha 1 p}-\bar{\chi}^\dagger_{\alpha 2 p}\alpha_0\bar{\chi}_{\alpha 2 p})$ and preserves the physical parity but breaks the "chiral" U($2N$) symmetry to its subgroup $[\mathrm{SU}(N)\times\mathrm{U}(1)]^2$. The second type of mass term has a form $m(\bar{\chi}^\dagger_{\alpha a p}\alpha_0\bar{\chi}_{\alpha a p})$, which preserves the chiral U($2N$) symmetry, but breaks the physical parity. Since the chiral transformations imply symmetry between real particles and antiparticles (bear in mind that $\bar{\chi}_{\alpha 2 p}=\sigma_1\chi_{\alpha 2 p}$), this symmetry is broken by an infinitesimally small doping which serves as a natural regulator in the present problem. In this paper we shall study only the Hamiltonian (20) regularized by the first type of mass terms:

$$H'_F=H_F+m'\int dp(\bar{\chi}^\dagger_{\alpha 1 p}\alpha_0\bar{\chi}_{\alpha 1 p}-\bar{\chi}^\dagger_{\alpha 2 p}\alpha_0\bar{\chi}_{\alpha 2 p}) \ . \tag{21}$$

IV. FLUCTUATIONS OF THE PHASE OF THE ORDER PARAMETER

Hitherto we have discussed the results of the mean-field approximation which implies the replacement of the path integral over the fields Δ_{ij},φ_i in (4) by a substitution of their values at the saddle point. Fluctuations of the amplitude of Δ are small everywhere except the narrow ($\tau^2\propto N^{-1}$) region in the vicinity of the transition point which we do not consider here.

Now we turn to the study of phase fluctuations of Δ_{ij}. Usually in a two-dimensional case thermal fluctuations destroy the mean order parameter. In our case these fluctuations are much stronger; they destroy the order parameter even at zero temperature.

It is convenient to represent the phase field ϕ_{ij} ($\Delta_{ij}=\Delta^{(0)}_{ij}\exp(i\phi_{ij})$ (where $\Delta^{(0)}_{ij}$ is the solution of the mean-field (MF) equations which we discussed in Sec. II) as a sum $\phi_{ij}=\psi_j-\psi_i+a_{ij}$, where the fields ψ_i are defined in the lattice sites, whereas ϕ_{ij} and $a_{ij}=-a_{ji}$ are defined at bonds (ij). This separation of ϕ_{ij} into ψ_i and a_{ij} is ambiguous, but it can be determined from some auxiliary condition imposed on a (choice of gauge); we discuss the most convenient form of this condition below. The integration over phase fluctuations ϕ_{ij} can be performed in two steps: at first over the field ψ_i and then over the field a_{ij} with the auxiliary condition (to prevent double counting). Now we discuss the effective action of these fields. To begin with we note that the obtained free energy $F\{\Delta\}=\ln\mathrm{Tr}\rho_f$ (10),(13) does not depend on ψ_i in the vicinity of the transition temperature T_c. To prove this fact in the general case we represent $\mathrm{Tr}\rho_f$ as a path integral over the fermion field:

$$\mathrm{Tr}\rho_f=\int\mathcal{D}\bar{c}\mathcal{D}c\exp\left[-\int_0^\beta[\bar{c}\partial_\tau c+H_F(\bar{c},c)]d\tau\right]$$

and perform in this integral the variable transformation

$$\begin{aligned} &c_i\to c_i\exp(i\psi_i) \ , \\ &\varphi_i\to\varphi_i-\dot{\psi}_i \ , \end{aligned} \tag{22}$$

which leads to the additional term in H_F

$$\delta H_F=in_0\dot{\psi}_i \ ,$$

that is, the total derivative is over time and does not change the effective action $\ln\mathrm{Tr}\rho_f$ at the integer n_0. Thus, in new variables H_F has the same form as before,

but Δ_{ij} changes to $\tilde{\Delta}_{ij}=\Delta_{ij}\exp(i\psi_i-i\psi_j)$.

Therefore, the effective action does not depend on the field ψ. In Sec. VI we discuss the effective action of the fields a_{ij},φ_i, while now we discuss the integration over the field ψ_i.

To compute any fermion correlator we should average its expression with the weight ρ_f over the fermion field and ψ field, then the variable transformation (21) in such an integral changes all variables $\bar{c},c$ in this way: $c\to c\exp(i\psi_i)$. Thus, the averaging over ψ results in the cancellation of any fermion correlator that corresponds to a charge-transfer process (e.g., one-particle Green's function). This averaging results also in a zero mean of the order parameter. The averaging over ψ does not change the fermion correlators that correspond to the processes without charge transfer, e.g., the two-particle correlator

$$\langle T(c^{\dagger}_{i\alpha}c_{i\beta})_{\tau}(c^{\dagger}_{j\gamma}c_{j\delta})_0\rangle=\left[\delta_{\alpha\delta}\delta_{\beta\gamma}-\frac{1}{N}\delta_{\alpha\beta}\delta_{\gamma\delta}\right]G^2_{ij}(\tau)\ . \quad (23)$$

The second term in this expression is a small correction to the first (over N^{-1}); it is due to fluctuations of the a_{ij} field (Sec. VI). In the case $N=2$ this correlation function means that the spin correlator is described in the RVB state by the ordinary metal Ruderman-Kittel-Kasuya-Yosida expression. The averaging over ψ does not change the partition sum, thus leaving intact the free energy so that all statements about the low-temperature behavior of specific heat (Sec. III) are still valid. A similar phenomenon that integration over bosonic field fluctuations changes the fermionic Green's function, but not the spectrum of fermionic excitations, was observed in Ref. 15.

V. METAL-DIELECTRIC TRANSITION

Above we considered the dielectric state at large U so that all states containing any other than n_0 electrons per site make a negligible contribution to the partition sum. Thus we take into account these states as virtual states only. With a decrease of the repulsion U the transition into the metal state occurs. In this transition region and in the metal state we cannot regard only the states with n_0 electrons per site. A small doping also results (even at large U/t) in the appearance of states with $n\neq n_0$ electrons per site. To describe these transitions it is far more convenient to discuss the modified Hubbard model (3) with an additional interaction $\mathcal{T}_{ij}$:

$$H=\sum_{i,j}[-\mathcal{T}_{ij}(c^{\dagger}_{i\alpha}c_{j\alpha})(c^{\dagger}_{j\beta}c_{i\beta})+t_{ij}c^{\dagger}_{i\alpha}c_{j\alpha}]$$
$$-\mu\sum_i n_i+\tfrac{1}{2}U\sum_i n_i^2\ . \quad (24)$$

Unlike the case of a large U (4), where the interaction appeared naturally in the first order of the perturbation expansion over t/U and was equal to t_{ij}^2/U, now $\mathcal{T}_{ij}$ is an independent parameter; moreover, at half-filled band we shall consider only the region $\mathcal{T}_{ij}\gg t_{ij}^2/U$, whereas at low doping and $U/t\gg 1$ we can treat also the region $\mathcal{T}\lesssim t_{ij}^2/U$. Proceeding then analogously to Sec. II we make the Hubbard-Stratanovich transformation in the path-integral representation of the density matrix and introduce the auxiliary field φ_i uncoupling the interaction Un_i^2 and the field Δ_{ij} uncoupling $\mathcal{T}_{ij}c^{\dagger}_{i\alpha}c_{j\alpha}c^{\dagger}_{j\beta}c_{i\beta}$. We get

$$L=\sum_i\bar{c}_{i\alpha}\partial_{\tau}c_{i\alpha}+\sum_{i,j}[\tfrac{1}{4}|\Delta_{ij}|^2\mathcal{T}_{ij}^{-1}+(t_{ij}-\Delta^{*}_{ij})c^{\dagger}_{i\alpha}c_{j\alpha}]$$
$$+\sum_i\left[(i\varphi_i-\mu)n_i+\frac{1}{2U}\varphi_i^2\right]\ . \quad (25)$$

New terms present in this Lagrangian [cf. (5)] affect slightly the mean-field equations (6) and (7) for Δ_{ij}, but these terms violate the symmetry of the action with respect to ψ transformations. To find the effective action of the ψ field we again make the variable transformation (21). Now this transformation changes the effective action and helps us to pick out nonvariant terms. After the variable transformation (21) the Lagrangian becomes

$$L=\sum_i\bar{c}_{i\alpha}\partial_{\tau}c_{i\alpha}+\sum_{i,j}\bar{c}_{i\alpha}c_{j\alpha}[\,t_{ij}\exp(-i\psi_i+i\psi_j)$$
$$+\exp(-ia_{ij})\Delta^{(0)}_{ij}]$$
$$+\sum_i\left[(i\varphi_i-\mu)\bar{c}_{i\alpha}c_{i\alpha}+\frac{1}{2U}(\varphi_i-\dot{\psi}_i)^2\right]\ . \quad (26)$$

It is convenient to choose the gauge transformations (9) so that the resulting φ_i does not depend on time and replaces the path integral over φ_i by a path integral over $\dot{\psi}_i(t)$ and an ordinary integral over $\phi_i=\phi_0$. The constant ϕ_0 should be determined from the condition of the free-energy minimum:

$$i\langle c^{\dagger}_{i\alpha}c_{i\alpha}\rangle=\frac{1}{U}\langle\dot{\psi}_i-\phi_0\rangle\ , \quad (27)$$

where $\langle\cdots\rangle$ means an average with weight $\exp[-\int L\,d\tau]$ with L being defined by (26). The fluctuations of the a_{ij} field have a small parameter N^{-1}. We postpone this discussion until Sec. VI. Here we only state that their effect on the effective action if the ψ field is small.

The last term in the effective action (26) describes a system of noninteracting rotators governed by the Lagrangian L_0:

$$L_0=\frac{1}{2U}\sum_i(\dot{\psi}_i-i\varphi_0)^2\ , \quad (28)$$

where we use real-time representation. The presence of the cross term $\dot{\psi}_i\varphi_0$ in the Lagrangian L_0 distinguishes these rotators from the ordinary ones. To obtain their energy spectrum we employ the Schrödinger representation. The canonical momentum M conjugate to the variable ψ_i is

$$M=\frac{1}{U}(\dot{\psi}_i-i\varphi_0)\ . \quad (29)$$

Inserting the expression (29) for a canonical momentum into Hamiltonian $H=M\dot{\psi}-L$ we get $H=\frac{1}{2}U(M+i\varphi_0/U)^2$. The wave function $Y(\psi)$ should be periodic over ψ so that eigenvalues of the operator M

are integers m and the corresponding energy levels are

$$\epsilon_m = \tfrac{1}{2}U(m + i\varphi_0/U)^2 \ . \tag{30}$$

In the rotator ground state $m = m_0$, where m_0 is the integer which is closest to $-i\varphi_0/U$. (It can be shown that φ_0 obeying the condition (27) is purely imaginary so that $i\varphi_0$ is real.)

Now the Eq. (27) acquires a simple meaning: It ensures that the mean number of electrons in the system equals the mean value of the operator $-M$.

The second term in the effective action (26) describes the interaction of rotators with each other and with fermions. Using the large-N approximation we can replace $c^\dagger_{i\alpha}c_{j\alpha}$ in it by its mean $\langle c^\dagger_{i\alpha}c_{j\alpha}\rangle = \Delta_{ij}(2\mathcal{T}_{ij})^{-1}$. Thus, in the leading order of the large-N approximation the effective action of the Bose field ψ_i becomes

$$\begin{aligned} &L\{\psi\} = L_0\{\psi\} + L_{\rm int}, L_{\rm int} = \sum_{i,j} \tilde{t}_{ij}\cos(\psi_i - \psi_j) \ , \\ &\tilde{t}_{ij} = t_{ij}\Delta_{ij}(2\mathcal{T}_{ij})^{-1} \ , \end{aligned} \tag{31}$$

where $L_{\rm int}$ describes the interaction between neighboring quantum rotators. The effective action for the system of interacting Josephson junctions also has the form of (31).[16]

In the absence of doping we should choose $i\varphi_0/U = n_0$, which ensures that each rotator spectrum is symmetric around $m_0 = -n_0$ so that the condition (27) is satisfied. In this case the energy spacing between the rotator ground state and its first excited level is $U/2$, whereas the interaction energy is of the order of tN. Thus, at small tN/U long-range order is absent in this system even at zero temperature. At larger t (such that $tN/U \approx 1$), the transition into the ordered state takes place in this system. Unfortunately, we are not aware of any reliable analytical method which allows one to study the properties of the system (31) near the transition into the ordered state (at $tN/U \approx 1$). However, the properties of the transition can be studied analytically for the systems with strong repulsion $tN/U \ll 1$ and low doping.

In this case the interaction between rotators is small. At nonzero doping the mean value of M is fractional, which implies that the rotator wave function is a superposition of wave functions with $m_1 = -n_0$ and with $m_2 = -(n_0 - 1)$. Thus, in this state we may choose $i\varphi_0/U$ to be close to $n_0 - \frac{1}{2}$ so that the level spacing between levels m_1 and m_2 is of the order of t. The level spacing between other levels remains $U \gg tN$. Therefore, in this state each rotator can be described by a two-level system. The excitations to the higher level are Bose particles, they can be described by the operators S_i which are equal to the operator $\exp(-i\psi_i)$ projected on the remaining two levels. If the density of holes is small then the number of Bose excitations is also small and it is convenient to represent the operators $S_i, S_i^\dagger$ as a series over creation-annihilation operators of the Bose field. We make use of the Holstein-Primakoff transformation and get

$$S_i^\dagger = b_i^\dagger(1 - b_i^\dagger b_i) \ , \tag{32}$$

where we retain only the leading and the next term of the expansion over the boson density. The projection of $-M$ on the remaining two levels equals $n_0 - b_i^\dagger b_i$ describing the number of bosons. Thus in the Bose representation the condition (27) acquires a simple form, $\langle c^\dagger_{i\alpha}c_{i\alpha}\rangle + \langle b_i^\dagger b_i\rangle = n_0$, which is exactly the constraint inserted by hand in the slave-boson approach.[5] It means that the number of bosons equals the number of physical holes. In the Bose representation the Hamiltonian of the rotators becomes

$$\begin{aligned} &H\{b\} = H_0\{b\} + H_{\rm int}\{b\} \ , \\ &H_0\{b\} = \sum_{i,j} \tilde{t}_{ij} b_i^\dagger b_j + \epsilon \sum_i b_i^\dagger b_i \ , \\ &H_{\rm int}\{b\} = -\sum_{i,j} \tilde{t}_{ij}(b_i^\dagger b_j^\dagger b_i b_i + b_j^\dagger b_j^\dagger b_i b_j) - \epsilon_i \, b_i^\dagger b_i^\dagger b_i b_i \ , \end{aligned} \tag{33}$$

where ϵ is the level spacing between m_1 and m_2 levels:

$$\epsilon = i\phi_0 - (n_0 - \tfrac{1}{2})U \ . \tag{34}$$

If the hole density is small then the gas approximation can be employed to study the Bose system (33). In this case only long-wavelength bosons are important; their bare spectrum follows from the Hamiltonian $H_0\{b\}$:

$$\epsilon_0(k) = \tilde{t}(k) + \epsilon = \epsilon - t_0 + t_1 k^2 \ . \tag{35}$$

We note here that the long-wave properties of the Bose spectrum $\epsilon_0(k)$ are insensitive to the fermionic state: In both the uniform state and the flux state $t_0 \sim t_1 \sim tN$ with only numerical coefficients in t_0, t_1 being different in these states.

At $\epsilon > t_0$ there are no bosons in the ground state. At $\epsilon < t_0$ the bare spectrum becomes unstable, but the Bose interaction described by $H_{\rm int}\{b\}$ corresponds to repulsion between long-wavelength bosons, therefore in this case a Bose condensate is formed, $\langle b\rangle \neq 0$. In the gas approximation the density of this condensate is governed by the scattering amplitude Γ of two bosons:

$$n_c = |\langle b\rangle|^2 = \frac{t_0 - \epsilon}{2\Gamma(\omega = 0, \ k = 0)} \ . \tag{36}$$

The scattering amplitude Γ at zero frequency and zero momentum can be obtained with logarithmic accuracy from a summation of ladder diagrams:

$$\Gamma = \frac{4\pi t_1}{\ln[t_0/(t_0 - \epsilon)]} \ . \tag{37}$$

In the leading order over $\{\ln[t_0/(t_0 - \epsilon)]\}^{-1}$ the total density of bosons n coincides with n_c. The difference between n and n_c is proportional to the next order of the small parameter $\{\ln[t_0/(t_0 - \epsilon)]\}^{-1}$:

$$\begin{aligned} n - n_c &\simeq (t_0 - \epsilon)/4\pi t_1 \\ &\simeq n_c\{\ln[t_0/(t_0 - \epsilon)]\}^{-1} \ . \end{aligned} \tag{38}$$

Thus in the present system the total number of bosons (and thus the hole density) is zero at $n_c = 0$, which means that at $n_c = 0$ the band is half filled. This is a general result for the Bose system with repulsion; attraction be-

tween bosons results in a macroscopic collapse. In Ref. 17 it was shown that for a system with one type of bosons there is no alternative to a one-particle condensate other than collapse. However, another scenario is possible if there are several types of Bose fields in the system. In this case a pair Bose condensate can be formed: $\langle b^{\alpha}b^{\beta}\rangle\neq 0$, $\langle b^{\alpha}\rangle=0$. For instance, this is the case for a system consisting of a number of planes in which a weak tunneling between planes results in an effective attraction between bosons on adjacent planes, whereas bosons on the same plane still repel one another. (This mechanism of boson attraction was proposed in Ref. 18.) The appearance of these Bose condensates results in the breakdown of the local gauge symmetry related to the fields a_{ij},φ.

The additional interaction

$$\sum_{ij}\mathcal{T}_{ij}(c^{\dagger}_{i\alpha}c_{j\alpha})(c^{\dagger}_{j\beta}c_{i\beta})$$

in the model considered in this section results in pairing of an electron and a hole at adjacent sites. An analogous interaction

$$\sum_{ij}\tilde{\mathcal{T}}_{ij}(c^{\dagger}_{i\alpha}c^{\dagger}_{j\alpha})(c_{i\beta}c_{j\beta})$$

results in electron-electron pairing at adjacent sites. The properties of the dielectric state with this interaction differs slightly from the properties of the dielectric state in the model (24), but in the metal state the formation of the order parameter $\langle\Delta_{ij}\rangle$ means superconductivity with the usual electron spectrum

$$\epsilon^2(k)=t^2(k)+\Delta^2(k)\ . \tag{39}$$

Far more interesting is the combined effect of the interaction

$$\sum_{ij}\mathcal{T}_{ij}(c^{\dagger}_{i\alpha}c_{j\alpha})(c^{\dagger}_{j\beta}c_{i\beta})$$

and pairing

$$\sum_{ij}t'_{ij}(c^{\dagger}_{i\alpha}c^{\dagger}_{i\beta})(c_{j\alpha}c_{j\beta})$$

which corresponds to a two-electron tunneling process from the i to j sites. Along with a one-electron tunneling process, this term violates local U(1) symmetry and results in the effective action of the ψ field with potential energy

$$\sum_{ij}t'_{y}\Delta^2_{ij}(2\mathcal{T}_{ij})^{-2}\cos[2(\psi_i-\psi_j)]\ .$$

The important difference from the potential energy (31) is the factor 2 inside the cosine function. If the repulsion U is small so that this term violates the symmetry and $\langle\cos(2\psi)\rangle$ is formed, one symmetry is still left: It is the transformation $c\rightarrow -c\,(\psi=\pi)$, that corresponds to the local Z_2 gauge group. In this state one-fermion charge transfer excitations still cannot exist, but two-fermion excitations can transfer charge. Thus, in this state normal current is impossible, whereas a genuine superconductive state can be formed.

VI. COLLECTIVE EXCITATIONS IN THE SPIN LIQUID

Hitherto we averaged over phase fluctuations ϕ_{ij} of the order parameter Δ_{ij} and over the field φ_i restricting these fluctuations to the sector that can be parametrized by the ψ field: $\phi_{ij}=\psi_i-\psi_j$, $\varphi_i=\dot{\psi}_i$. In this section we discuss the effect of the averaging over the remaining fluctuation sector.

First of all we choose the parametrization of these fluctuations. We note that the representation

$$\phi_{ij}=\psi_j-\psi_i+a_{ij} \tag{40}$$

corresponds to the gauge transformation of the a field. Thus the independence of the effective action $\ln \mathrm{Tr}\rho_f$ under the variable transformation (22) means that the effective action $\ln \mathrm{Tr}\rho_f\{\phi_{ij},\varphi_i\}$ does not change under the gauge transformations if the field φ_i (scalar potential) is also transformed:

$$\varphi_i=\tilde{\varphi}_i-\dot{\psi}_i\ . \tag{41}$$

The averaging over ψ fields means in these terms also averaging over the gauge transformation, and the remaining fluctuation sector corresponds to fluctuations of the fields $F_{\mu\nu}$.

Now we discuss the form of the effective action of these fields. In this section we shall consider only the dielectric state neglecting the influence of fluctuations of the ψ field on the action of the gauge fields (a,φ). The bare action (without fermions) is zero, but the integration over fermion fields results in a large action (it contains a factor N) so that the resulting fluctuations of these fields are small and we can safely retain only terms quadratic over a,φ in action. As before (Sec. VI) it will be convenient to choose the gauge imposing the condition $\varphi_i=\varphi_0$. Expanding then the action (26) over small a_{ij} and keeping only the terms of second order, we perform the averaging over fermions and get

$$\frac{1}{N}S\{a\}=-\tfrac{1}{2}\int_0^{\beta}d\tau\,d\tau'\sum_{ijkl}a_{ij}(\tau)\Delta^{(0)}_{ij}a_{kl}(\tau')\times\Delta^{(0)}_{kl}G_{il}(\tau-\tau')G_{kj}(\tau'-\tau)$$
$$-\tfrac{1}{2}\int_0^{\beta}d\tau\sum_{ij}a^2_{ij}\Delta^{(0)}_{ij}G_{ij}(0)\ . \tag{42}$$

Since the fluctuations of a_{ij} have a small parameter, only long-wave fluctuations of the a field can result in dangerous consequences. For these fluctuations we can replace the link field a_{ij} by its continuum limit: $a_{ij}=(z_i^{\alpha}-z_j^{\alpha})a_{\alpha}[(z_i+z_j)/2]$, the effective action of these fluctuations becomes

$$\frac{1}{N}S\{a\}=\frac{T}{2}\int\sum_{\omega}a_\alpha(k,\omega)a_\beta(k,\omega)\Pi_{\alpha\beta}(k,\omega)(d^2k)$$

$$\Pi_{\alpha\beta}=\int(d^2p)\left[\frac{n(p+k/2)-n(p-k/2)}{\omega-[\epsilon(p+k/2)-\epsilon(p-k/2)]}\frac{\partial\epsilon}{\partial p_\alpha}\frac{\partial\epsilon}{\partial p_\beta}+\frac{\partial^2\epsilon}{\partial p_\alpha\partial p_\beta}n(p)\right] \tag{43}$$

$$n(p)=(\exp\{[\epsilon(p)-\mu]/T\}+1)^{-1}\ .$$

In the uniform phase only the first terms in the expansion of $\Pi_{\alpha\beta}(\omega,k)$ over ω and k are required for the following. We find

$$\Pi_{\alpha\beta}=\left[\delta_{\alpha\beta}-\frac{k_\alpha k_\beta}{k^2}\right]\left[\nu_1(\hat{k})\frac{|\omega|}{|k|}+\nu_2(\hat{k})k^2\right]+\frac{k_\alpha k_\beta}{k^4}\omega^2\nu_F(\epsilon_F)\ , \tag{44}$$

where $\nu_1(\hat{k})$ and $\nu_2(\hat{k})$ are smooth functions of $\hat{k}=k/|k|$ and are due to anisotropy of the spectrum $\epsilon(k)$. Their explicit form is unimportant for the following estimates in Sec. VII, where we put

$$\nu_1(\hat{k})=1,\quad \nu_2(\hat{k})=\Delta,\quad \nu_F(\epsilon_F)=\Delta^{-1}\ .$$

In the flux phase the electron spectrum is relativistically invariant near the Fermi surface so that the effective action of the gauge field a_α should also be relativistically invariant. Generally to two types of relativistically invariant terms are possible: apart from a purely transverse term with $\Pi_{\alpha\beta}$ given by

$$\Pi_{\alpha\beta}=\frac{1}{8}\frac{\omega^2\delta_{\alpha\beta}+(Sk)^2\delta_{\alpha\beta}-S^2k_\alpha k_\beta}{[\omega^2+S^2k^2]^{1/2}},\quad S=2\Delta\ , \tag{45}$$

the other term is also possible which acquires a more simple form in the real-space representation:

$$S'=\frac{i}{16}\tilde{\gamma}N\int d\tau\,d^2z\,\epsilon_{\alpha\beta\gamma}a_\alpha f_{\beta\gamma}\ ,$$
$$f_{\beta\gamma}=\partial_\beta a_\gamma-\partial_\gamma a_\beta\ . \tag{46}$$

The coefficient $\tilde{\gamma}$ of this term depends crucially on the regularization of the Fermi spectrum in the vicinity of points k^* (Sec. III). If the regularization is provided by the parity-conserving mass term, $\gamma=0$ and this term is absent as it should be since it violates parity. On the other hand, if the regularization is provided by chirality-conserving mass, $\gamma=2$ and the parity is violated. In this paper we discuss only the former case. We only note that this term, if it exists, would have serious consequences for the statistics of low-energy excitations, and, in fact, may inhibit the Bose condensation of holons discussed in Sec. V, so that this problem deserves careful study.

The form of the effective action (44) means that collective low-energy bosonic excitations of the gauge field are purely relaxational in the uniform phase with the decay rate $\omega=i\nu_2k^3\nu_1$. In contrast, in the flux phase, low-energy excitations are purely oscillatory with the spectrum $\omega=Sk$.

The effective action (43) contains a large factor N so that the correlators of these fields are small and the corrections to the spinon correlators that originate from the bosonic exchange are also small ($\propto N^{-1}$). For instance, in the correlation function (23) the last term is due to the exchange of the a_α field. (It can be easily obtained in the gauge div$a+\partial_\tau\varphi=0$.) For fixed indices α,β this term is small over N^{-1}; however, it cancels the other term in the trace of this correlation function over pairs of indices (α,β) and (γ,δ). This cancellation means that the correlation function of the density operator n_i is exactly zero as it should be in a dielectric.

VII. THE ACTION OF INSTANTON

In the lattice (2+1)-dimensional electrodynamics the interaction of photons and instantons result in an exponentially small mass [$m\propto\exp(-1/e^2)$] for photon and quark confinement.[19] In this theory instantons form a very dilute gas. In our case the singularity of the effective action (44) and (45) at small momenta which originates from the interaction of the gauge field with fermions changes the situation dramatically and results in a zero-density instanton gas and zero photon mass.

To prove it we compute the action of one instanton. The action is a periodic function of a_{ij}. Thus it is also the periodic function of the lattice "magnetic" field

$$b_z=a_{ij}+a_{jk}+a_{kl}+a_{li}$$

(z is the center of a plaquette $[ijkl]$). the simplest way to take into account this periodicity is to introduce the auxiliary integer-valued field n_z and rewrite the action in the form

$$S=\tfrac{1}{2}\int d\tau'd\tau\left[\sum_{ijkl}\epsilon_{ijkl}(\tau'-\tau)e_{ij}(\tau)e_{kl}(\tau')+\sum_{z,z'}\mu_{zz'}(\tau-\tau')[b_z(\tau)-2\pi n_z][b_{z'}(\tau')-2\pi n_{z'}]\right]\ , \tag{47}$$

where $e_{ij}=\partial_\tau a_{ij}$ is the "electric" field: in the long-wavelength limit $e_{ij}=e_\alpha(z_i^\alpha-z_j^\alpha)$, $e_\alpha=\partial_\tau a_\alpha$, $b=\nabla\times a$, ϵ is the dielectric constant, and μ is the magnetic permeability; in the long-wavelength limit they acquire a simple form in the momentum representation:

$$\epsilon_{\alpha\beta}(\omega,k)=\left[\delta_{\alpha\beta}-\frac{k_\alpha k_\beta}{k^2}\right]\epsilon_1+\frac{k_\alpha k_\beta}{k^2}\epsilon_2\ ,$$
$$\epsilon_1=\frac{\nu_1N}{|\omega|k},\quad \epsilon_2=\frac{\nu_FN}{k^2}\quad\text{(uniform)}\ ,$$

$$\mu=\nu_2 N \ , \tag{48}$$

$$\epsilon_{\alpha\beta}=\delta_{\alpha\beta}\epsilon, \quad \mu=S^2\epsilon=\frac{NS^2}{8(\omega^2+S^2k^2)} \quad \text{(flux)} \ .$$

One instanton centered at the point 0 corresponds to the n-field configuration in which the sum of n_z over any surface enclosing [in (2+1)-dimensional space] the point 0 is unity. For simplicity we choose $n(\tau,z)=\delta(z)\theta(\tau)$. Now we should minimize the action (47) and (48) with this source. In the momentum representation the saddle-point configuration of the field a (we work in a gauge $\varphi=0$) is:

$$a_\alpha=2\pi\mu\frac{\hat{\epsilon}_{\alpha\beta}k_\beta i n(\omega)}{\epsilon_1\omega^2+\mu k^2} \quad (\hat{\epsilon}_{\alpha\beta}=-\hat{\epsilon}_{\beta\alpha}, \ \ \hat{\epsilon}_{12}=1) \ , \tag{49}$$

and the corresponding action equals

$$S_0=\tfrac{1}{2}\int dk\, d\omega\frac{(2\pi i n\omega)^2\epsilon_1\mu}{\epsilon_1\omega^2+\mu k^2} \ . \tag{50}$$

Since one instanton corresponds to $n_\omega=i\omega^{-1}$ the integrals in (50) for the uniform and flux phases become, respectively,

$$\begin{aligned} S_0&=2\pi^2 N\int\frac{\nu_1\nu_2(d\omega\, dk)}{(\nu_1|\omega|+\nu_2k^3)|\omega|} \quad \text{(uniform)} \\ S_0&=\frac{N\pi^2}{4}\int\frac{(d\omega\, dk)S^2}{(\omega^2+S^2k^2)^{3/2}}\simeq\frac{N}{8}\ln(k_c^{-1}) \quad \text{(flux)} \ . \end{aligned} \tag{51}$$

These integrals diverge at small momenta, thus the action of one instanton is infinite and their density is zero in both phases. The main contribution to the integral (51) comes from the region of small frequencies and momenta which justifies the employed long-wavelength approximation. In the flux phase the divergency of the integral (51) is only logarithmic.

Evaluating in this case the contribution of the one-instanton configuration to the density matrix of the system with volume W we find that it is proportional to

$$W\exp(-S_0)=\exp[(1-N/24)\ln W]$$

and, thus, can be neglected only if $N\geq N_c=24$. At $N=N_c$ the second-order transition takes place, and at $N<N_c$ the main contribution to the density matrix comes from configurations with macroscopically large numbers of instantons. In Euclidean (imaginary time) language, the instanton is a "magnetic" monopole with unit charge. The transition which happens at $N=N_c$ is similar to the vortex unbinding transition discovered by Berezinsky, Kosterlitz, and Thouless[20] in the two-dimensional (2D) XY model. The properties of the present model in the vicinity of the transition point N_c can be described by the renormalization-group theory developed in[20] for the 2D XY model. Since we regard the problem of the physical properties of a model with $N\simeq N_c$ as unphysical, we shall not dwell upon this point here, but discuss the qualitative properties at $N\lesssim N_c$.

We use the Euclidean language. The energy of a monopole configuration can be obtained analogously to the derivation of one monopole energy (50). It comes to a transparent form if we express it through a monopole charge density $q=\partial n/\partial\tau$. (We note here that this formula is not Euclidean invariant, since in the present model there is no lattice spacing in the imaginary time direction; otherwise n would be vector n_α, and this formula would read $q=\mathrm{div}n$). We get

$$E=\frac{N\pi^2}{4}\int\frac{|q_p|^2}{p^3}dp, \quad q_p=\sum_i q_i\exp(-ipR_i) \ , \tag{52}$$

where we use three-dimensional notation: $p=(\omega S^{-1},k)$, $R=(tS,z)$. Each monopole generates a three-dimensional magnetic field $H=2\pi p_\mu/|p|^2$ [where $H=(b,eS^{-1})$]. In the absence of monopoles the correlation function of magnetic fields H_μ follows from (47), and the presence of monopoles modifies it by the correlation function $\langle qq\rangle_p$ of the monopole charge density:

$$\langle H_\mu H_\nu\rangle=\frac{8}{N}|p|\left[\delta_{\mu\nu}-\frac{p_\mu p_\nu}{p^2}\right]+\frac{p_\mu p_\nu}{p^4}4\pi^2\langle qq\rangle_p \ . \tag{53}$$

At $N<N_c$ at large distances ($R\gtrsim R_c$) monopoles become free and fluctuations of the charge density can be described by the usual weight $\exp[-E\{q_p\}]$ in which the charge density q_p can be regarded as an independent variable. Evaluating in this way the correlation function $\langle qq\rangle_p$, we see that it cancels the singular term in (53). Therefore, at $N<N_c$ the transverse part of the gauge field cannot propagate to scales larger than R_c, in other words, at $N<N_c$ the gauge field acquires a mass $m\propto R_c^{-1}$.

VIII. ELECTROMAGNETIC PROPERTIES

We have shown that low-energy fluctuations in the present model are described by Fermi fields and gauge fields in the dielectric state and by Fermi, gauge, and Bose fields in a state close to the dielectric. In this section we discuss their interaction with a real electromagnetic field and the resulting electromagnetic properties of the whole system. We shall consider only the model with large repulsion ($tN/U\ll 1$) and low (or zero) doping. Moreover, since we are interested mainly in the qualitative interplay between superconducting and normal properties we shall suppose some scattering mechanism (e.g., by defects) which results in finite conductivity at $T=0$ in the normal state of the Fermi liquid.

To obtain the effective interaction with an electromagnetic field A we recall that the electromagnetic field results in the appearance of a phase factor $\exp(iAz_{ij})$ in the term $t_{ij}c^\dagger_{i\alpha}c_{j\alpha}$ describing the hopping processes in Hamiltonian (24). Proceeding further analogously to the derivation of (26) and introducing the operators b_i describing the Bose field at low doping, we obtain the effective interactions of Bose and Fermi fields with the gauge field a_{ij} and the electromagnetic field A:

$$L_{\text{int}}=\sum_{i,j}\bar{c}_{i\alpha}c_{j\alpha}\exp(iAz_{ij})[t_{ij}b_ib_j^\dagger-\exp(-ia_{ij})\Delta_{ij}^{(0)}] \ . \tag{54}$$

The action (54) is invariant under gauge transformations $A \to A + \nabla\theta_A$, $c \to c \exp i\theta_A$, and $a_{ij} \to a_{ij} + \theta_{ai} - \theta_{aj}$, $c \to c \exp(-i\theta_a)$, $b \to b \exp(i\theta_a)$, which means that the charges of the Fermi field with respect to gauge fields A and a_{ij} are $+1$ and -1, correspondingly, the charges of the Bose field with respect to gauge fields A and a_{ij} are 0 and -1.

In the absence of holes the term proportional to t_{ij} in the effective action (54) can be omitted, so that the gauge fields are present in the action in the difference $Az_{ij} - a_{ij}$ only. In that case the effective long-range action of the gauge fields A, a_{ij} can be expressed through the fermionic polarization operator $\Pi_{\alpha\beta}$:

$$S\{A,a\} = \frac{T}{2} \int dk \sum_{\omega} \{ [A_\alpha(\omega,k) - a_\alpha(\omega,k)] \Pi_{\alpha\beta}(\omega,k) \times [A_\beta(\omega,k) - a_\beta(\omega,k)] \} . \quad (55)$$

Therefore, in that case the average over a_{ij} results in the action which does not depend on the electromagnetic field at all. This means that in the absence of holes the electromagnetic response is absent as it should be in a dielectric.

In the lowest-order approximation over the hole density the effective action of the gauge fields becomes

$$S\{A,a\} = \frac{T}{2} \int dk \sum_{\omega} \{ [A_\alpha(\omega,k) - a_\alpha(\omega,k)] \times \Pi_{\alpha\beta}(\omega,k) [A_\beta(\omega,k) - a_\beta(\omega,k)] + a_\alpha(\omega,k) \pi_{\alpha\beta}(\omega,k) a_\beta(\omega,k) \} , \quad (56)$$

where the first term is generated by fermions that interact with both fields A, a_{ij} and the second is generated by bosons which interact only with field a_{ij}. It is convenient to single out from $\Pi_{\alpha\beta}$ and $\pi_{\alpha\beta}$ their longitudinal and transverse parts:

$$\Pi_{\alpha\beta} = \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) \Pi_1 + \frac{k_\alpha k_\beta}{k^2} \Pi_2 ,$$
$$\pi_{\alpha\beta} = \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) \pi_1 + \frac{k_\alpha k_\beta}{k^2} \pi_2 . \quad (57)$$

The appearance of a Bose condensate means that π_1 remains nonzero at $\omega, k \to 0$: $\pi(\omega = 0,\ k = 0) = n_c$. At temperatures above the critical point of Bose condensation $\pi_1(k=0) = \sigma_b \omega + O(\omega^2)$. If the repulsion U is strong, then the number of bosons equals the number of holes which implies, in particular, that σ_b becomes zero in the case of a half-filled band. At finite U we should take into account that the rotator should be described by two Bose fields: One field corresponds to creation of excitations with $m > m_0$ and the other to excitations with $m < m_0$. The number of holes equals the difference between the numbers of bosons of these two types. At finite temperature and for a half-filled band the conductivity becomes finite but exponentially small: $\sigma_b \propto \exp(-U/2T)$, since the minimal energy of excitation energy in this state is $U/2$. If the band is half filled the gap in the bosonic spectrum which is nearly $U/2$ at $U \gg tN$ decreases with U and becomes zero at $U = U_c \approx tN$. At $U < U_c$ a Bose condensate is formed which breaks the gauge symmetry.

Now we turn to fermionic excitations. We begin with the uniform phase. The exchange of virtual quanta of the field a_{ij} leads to repulsion between fermions separated by a large distance. The exchange by virtual quanta of the Bose field can lead to attraction at large distances but its strength is proportional to the density of bosons (i.e., the density of holes) and is weak. Thus, in the framework of the simplest one-plane model with equivalent sites the effective interaction of fermions is repulsive at large distances. The exchange of short-wavelength fluctuations of the Δ_{ij} field usually results in short-wavelength attraction between fermions. Thus, in that system a weak superconductivity with a large correlation length is impossible, but the possibility of a strong superconductivity with a small correlation length ξ cannot be excluded on such general grounds and deserves special study. If the system of fermions is normal then $\Pi_1(k=0) = \sigma_f \omega + O(\omega^2)$. In more complicated models the interaction between fermions can become attractive; in this case at low temperature the system of fermions becomes superconductive and $\Pi_1(k=0) = \rho_f + O(\omega)$.

For instance, in the model which consists of many planes coupled by weak tunneling the exchange of virtual bosons results in weak long-range attraction between fermions on adjacent planes. Since the gauge field a_{ij} is purely two-dimensional the long-range repulsion between fermions on different planes is absent. Thus, in this model the effective long-range interaction between fermions on adjacent planes is attractive and leads to superconductivity of the fermion subsystem. This superconductivity is unusual since it originates from off-diagonal pairing of fermions on adjacent planes. This form of pairing was originally proposed for layered materials in Ref. 7. All the aforementioned reasons for existence (or rather nonexistence) of a superconducting state in the uniform phase can be applied as well to the flux state if N is indeed very large ($N > N_c$). However, as we have seen in Sec. VII, at $N < N_c$ the nonperturbative mechanism provided by instantons results in the effective mass of the gauge field, i.e., it favors the superconductivity of the Fermi subsytem. The other important difference of the flux phase is that, since the density of states is zero on the Fermi surface, its conductivity is proportional to the temperature in its normal state.

To get the effective action of the electromagnetic field which describes its interaction with the whole electronic system we should average the effective action (56) over long-wavelength fluctuations of the gauge field a_{ij}. Performing the averaging we get

$$S\{A\} = \frac{T}{2} \int dk \sum_{n} A_\alpha(\omega,k) P_{\alpha\beta} A_\beta(\omega,k) ,$$
$$P_{\alpha\beta} = \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) P_1 + \frac{k_\alpha k_\beta}{k^2} P_2 , \quad (58)$$

where

$$P_1=\frac{\Pi_1\pi_1}{\Pi_1+\pi_1}\ . \tag{59}$$

The current j appearing as a reaction to the external electromagnetic field A is given by $j=PA$. Thus formula (59) means that the physical conductivity of the whole system is determined by the lowest conductivity of the Fermi or Bose subsystems.

This means that if the Bose subsystem is superconducting and the Fermi subsystem is not, then the physical conductivity is finite and equals σ_f. In the opposite case, if the Fermi subsystem is superconducting and the Bose subsystem is not, then the conductivity is also finite and equals σ_f. If both subsystems are superconductive then the superconductive density of the whole system is $\rho=\rho_f\rho_f/(\rho_f+\rho_b)$. If both subsystems have a finite conductivity, then the resistivity of the whole system is the sum of the resistivities of the subsystems: $\sigma^{-1}=\sigma_f^{-1}+\sigma_b^{-1}$. If the band is half filled and $U>U_c$, then σ_b tends to zero at $T\to 0$. Therefore, the conductivity of the whole system is zero in this state at $T=0$ independent of the state of the fermion subsystem.

IX. CONCLUSIONS

The main purpose of this work was the study of a model that provided gapless fermionic excitations in the dielectric state. In the considered model the operators of the fermionic quasiparticles (spinons) are operators of real electrons dressed by a phase factor $\tilde{c}_i=\exp(i\psi_i)c_i$. In the dielectric state the local gauge symmetry is restored, the fluctuations of are large, and the mean value of the phase factor $\langle \exp i\psi\rangle$ is zero. Thus, in the state with excited quasiparticles the mean number of real electrons is zero and spinons cannot carry the electric charge. The transition in the metal state results in symmetry breaking and formation of the order parameter $\langle\exp(i\psi)\rangle\neq 0$, and the spinons become ordinary Fermi excitations.

In this paper we studied a model with a large number of colors N that allowed us to obtain quantitative results using a N^{-1} expansion. Apparently, for real materials, $N=2$ and the obtained quantitative results (such as each phase energy, and the phase diagram) are unreliable, whereas the qualitative properties of each phase probably do not change for $N=2$. We believe that the only way to study the quantitative properties of a more realistic model with $N=2$ is the variational approach in which a wave function for $N\gg 1$ can be used as a trial function. In the present approach we have not considered the problem of phase separation which can occur in the doped systems. For instance, in the usual Hubbard model with low doping with strong repulsion $U\gg t$ the phase separation is energetically favorable: The magnetic state with one electron per site and a small ferromagnetic bubble which contains all holes has lower energy than the uniform state with an even distribution of holes.[21]

The same phenomenon occurs in the present model if $\tilde{t}>I$ [note that in the Hubbard model without modification (3) $I=t^2/U\ll\tilde{t}$]. In that case in the doped system with $n=n_0-\delta$ (n_0 integer) electrons per site the phase separation takes place: In the main part of the system the number of electrons is n_0 and in the other part it is n_0-1. Above we have considered the model in which $\tilde{t}$ and I are independent parameters, in that model the opposite case is also possible: $\tilde{t}<I$, in which case no phase separation occurs even in the doped systems, and δ is equal to the density of holons. Physically the model with independent parameters t and I can be justified in the two-band model.

[1] J. G. Bednorz and K. A. Muller, Z. Phys. B **64**, 188 (1986).
[2] I. Pomeranchuk, Zh. Eksp. Teor. Fiz. **11**, 226 (1941).
[3] P. W. Anderson, Mater. Res. Bull. **8**, 153 (1973).
[4] P. W. Anderson, Science **235**, 1196 (1987).
[5] G. Baskaran and P. W. Anderson, Phys. Rev. B **37**, 580 (1988).
[6] P. W. Anderson, G. Baskaran, Z. Zou, and T. Hsu, Phys. Rev. Lett. **58**, 2790 (1987).
[7] K. B. Efetov and A. I. Larkin, Zh. Eksp. Teor. Fiz. **68**, 155 (1975).
[8] P. W. Anderson, Phys. Rev. **115**, 2 (1959).
[9] L. B. Ioffe and A. I. Larkin, Int. J. Mod. Phys. **B2**, 203 (1988).
[10] I. Affleck and J. B. Marston, Phys. Rev. B **37**, 3774 (1988).
[11] M. Noga, Czech. J. Phys. B **38**, 210 (1988).
[12] G. Kotliar, Phys. Rev. B **37**, 3664 (1988).
[13] A. H. MacDonald, S. M. Girvin, and D. Yoshioka, Phys. Rev. B **37**, 9753 (1988).
[14] T. W. Appelquist, M. Bowick, D. Karabali, and L. C. R. Wiyewardhana, Phys. Rev. D **33**, 3704 (1986).
[15] P. B. Wiegmann and A. I. Larkin, Zh. Eksp. Teor. Fiz. **72**, 857 (1977).
[16] K. B. Efetov, Zh. Eksp. Teor. Fiz. **78**, 2017 (1980).
[17] P. Noziers and D. Sant James, J. Phys. (Paris) **43**, 1133 (1982).
[18] J. M. Wheatly, T. C. Hsu, and P. W. Anderson, Phys. Rev. B **37**, 5897 (1988).
[19] A. M. Polyakov, Phys. Lett. **59B**, 82 (1975); Nucl. Phys. **B120**, 429 (1977).
[20] V. L. Berezinsky, Zh. Eksp. Teor. Fiz. **59**, 907 (1970); J. Kosterlitz and D. Thouless, J. Phys. C **6**, 1181 (1973); J. Kosterlitz, *ibid.* **7**, 1046 (1974).
[21] L. B. Ioffe and A. I. Larkin, Phys. Rev. B **37**, 5730 (1988).