

Antiferromagnetic order and high-temperature superconductivity

Kerson Huang and Efstratios Manousakis*

*Center for Theoretical Physics, Laboratory for Nuclear Science and Department of Physics,
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

(Received 14 July 1987; revised manuscript received 16 September 1987)

The nearly half-filled Hubbard model with strong on-site repulsion is used as a model for the recently discovered copper oxide high-temperature superconductors. The onset of superconductivity upon doping the half-filled lattice of La_2CuO_4 is explained in terms of the Bose-Einstein condensation of effective mass $O(m_e)$. These holes move freely over an antiferromagnetic background. The mechanism is very different from the BCS mechanism. A fit to data is obtained by adjusting the interplane hopping amplitude, which turns out to be extremely small, though necessary.

I. INTRODUCTION

In this paper we study the recently discovered¹ high-temperature superconductors $\text{La}_{2-x}\text{X}_x\text{CuO}_4$, where X is a divalent element, by idealizing the system as a nearly half-filled Hubbard model with large on-site repulsion, as first suggested by Anderson.² The half-filled lattice, with exactly one electron per site, simulates the insulator La_2CuO_4 , in which nearest-neighbor electrons are thought to have strong antiferromagnetic correlations. Doping, i.e., replacing the trivalent La by a divalent element X , will create vacancies (holes). Anderson argues that the vacancies will enable bound singlet pairs of electrons to become current carriers, and lead to superconductivity. This suggestion has gained credence, when recent experiments on single La_2CuO_4 crystals demonstrated the existence of planar antiferromagnetic order³ in the Cu-O plane.

We shall show that the antiferromagnetic order is important for the onset of superconductivity upon doping, in that it provides a template over which the holes move as noninteracting bosons of charge $+2e$. The approximation, which we shall refer to as antiferromagnetic background approximation (ABA) is that the template can be treated classically, and is left undisturbed by the motion of the holes. The ideal Bose gas of holes will undergo Bose-Einstein condensation below T_c , and the superfluid flow of the condensate induced by an external field will give rise to superconductivity.

The effective mass of the holes in ABA turns out to be about two electron masses, which sets a temperature scale of the order of 2000 K. The much lower observed value of T_c reflects the fact that Bose-Einstein condensation can occur only in three or more dimensions, and is considerably reduced by the quasiplanar nature of the lattice. We are able to fit data by adjusting the interplane hopping amplitude, which turns out to be about 5×10^{-7} times the in-plane hopping amplitude. Our model predicts that T_c is proportional to the cube root of the interplane hopping amplitude. It should increase under external pressure, or with a change in lattice structure that brings the planes closer together.

We estimate that ABA is valid only for doping fractions

much less than 60%. For higher doping we use mean-field theory as a guide, although it cannot be made to fit data by varying the parameters of the model. In mean-field theory superconductivity comes about through Cooper-pair formation (CP). The mechanism differs from that in conventional BCS theory, in that the pairs here are bound by nearest-neighbor antiferromagnetic attractions, instead of forces arising from phonon exchange.

We begin by defining the Hubbard model used as a starting point, and then justify the assumptions used in formulating ABA. We then solve the model in ABA, and obtain an analytical expression for the critical temperature, which is used to fit data from $\text{La}_{x-2}\text{Sr}_x\text{CuO}_4$. Results from mean-field theory are then cited. We end with a discussion on the contrast between Bose-Einstein condensation and the CP mechanism.

II. STRONG-COUPLING HUBBARD MODEL

Consider a Hubbard model on a three-dimensional (3D) rectangular lattice, with M sites and N electrons. We are interested in the nearly half-filled case, for which N is slightly less than M . The Hamiltonian of the system is given by

$$H = U \sum_i N_{i\uparrow} N_{i\downarrow} - \sum_{\langle i,j \rangle, s} t_{ij} (C_{j,s}^\dagger C_{i,s} + \text{H.c.}), \quad (2.1)$$

where $C_{i,s}$ is the annihilation operator for an electron at site i with spin s , $N_{i,s}$ the corresponding occupation number, and $\langle i,j \rangle$ denotes a pair of nearest-neighbor sites. The first term in Eq. (2.1) discourages double occupancy of any site, and the second term allows the electrons to hop from site to site. We take

$$t_{ij} = \begin{cases} t, & \text{if } \langle i,j \rangle \text{ lies in Cu-O plane,} \\ t_3, & \text{if } \langle i,j \rangle \text{ is orthogonal to Cu-O plane.} \end{cases} \quad (2.2)$$

Typically $U \gg t$, $t_3/t \ll 1$, $t \approx 0.5$ eV.⁴ In numerical estimates later we take for definiteness $U = 2.5$ eV, or $t/U = 0.2$. Formally we take the large U limit, and treat the second term in Eq. (2.1) by perturbation theory. In this limit the half-filled lattice is populated exactly one electron per site. In the underfilled case there are vacan-

cies (holes), and the fraction of vacant sites $x=1-N/M$ is called the doping fraction.

The unperturbed Hamiltonian has highly degenerate energy levels equally spaced by U . The lowest level corresponds to states with no doubly occupied sites, the first level corresponds to exactly one doubly occupied site, etc. Let S_0 be the Hilbert subspace spanned by the states having the lowest energy. To second order in standard perturbation theory, the linear combination among the states

of S_0 that best approximates the ground state is to be found by diagonalizing the following effective Hamiltonian:

$$H_{\text{eff}} = H' + H' \frac{Q}{E_0 - H_0} H', \quad (2.3)$$

where H_0 and H' are respectively the first and second term in Eq. (2.1), and Q is the projection operator out of S_0 . Working out H_{eff} explicitly, we obtain

$$H_{\text{eff}} = H_1 + H_2 + H_3, \quad (2.4)$$

$$H_1 = - \sum_{\langle i,j \rangle, s} [\xi_i C_{i,s}^\dagger C_{j,s} + (i \leftrightarrow j)], \quad (2.5)$$

$$H_2 = -2U^{-1} \sum_{\langle i,j \rangle, s} t_{ij}^2 (C_{j,s}^\dagger C_{i,s} N_{i,-s}^\dagger C_{i,s}^\dagger C_{j,s} + C_{j,-s}^\dagger C_{i,-s} C_{i,s}^\dagger C_{j,s}), \quad (2.6)$$

$$H_3 = -U^{-1} \sum_{\langle i,j,k \rangle, s} t_{ij} t_{jk} [\xi_k (C_{k,s}^\dagger C_{j,s} N_{j,-s}^\dagger C_{j,s}^\dagger C_{i,s} + C_{k,-s}^\dagger C_{j,-s} C_{j,s}^\dagger C_{i,s}) + (i \leftrightarrow k)], \quad (2.7)$$

where $\langle i,j,k \rangle$ denotes a triplet of sites in which ij and jk are nearest neighbors, and

$$\xi_i = \begin{cases} 1, & \text{if site } i \text{ is empty,} \\ 0, & \text{otherwise.} \end{cases} \quad (2.8)$$

Introducing the operator

$$b_{ij} = 2^{-1/2} (C_{i\uparrow} C_{j\downarrow} - C_{i\downarrow} C_{j\uparrow}), \quad (2.9)$$

we can also write

$$H_2 = -4U^{-1} \sum_{\langle i,j \rangle} t_{ij}^2 b_{ij}^\dagger b_{ij}, \quad (2.10)$$

$$H_3 = -2U^{-1} \sum_{\langle i,j,k \rangle} t_{ij} t_{jk} [\xi_k b_{kj}^\dagger b_{ji} + (i \leftrightarrow k)]. \quad (2.11)$$

Actually the factor ξ_i is unnecessary as long as we only consider matrix elements of H_{eff} in the sub-Hilbert space S_0 . It is included merely as a reminder. The independent parameters in the model are the antiferromagnetic coupling J , the ratio of hopping parameters r , and doping fraction x :

$$J = 4t^2/U, \quad (2.12)$$

$$r = t_3/t, \quad (2.13)$$

$$x = 1 - N/M. \quad (2.14)$$

The case $r=0$ corresponds to a two-dimensional (2D) lattice.

As is evident from Eq. (2.6), the Hamiltonian H_2 describes the virtual process illustrated in Fig. 1 (a), in which an electron hops to a neighboring site occupied by an electron of the opposite spin, making it doubly occupied momentarily, and redresses the situation either by hopping back itself, or having the other electron hop back. The negative sign favors antiferromagnetism, since each nearest-neighbor triplet pair misses the chance to lower the energy by $J/2$. In the scenario we shall describe later, the onset of superconductivity depends on the existence of antiferromagnetic order, for which H_2 is responsible.

In the form (2.10) H_2 is expressed as a sum of terms of

the form $b^\dagger b$, which might lead one to think that H_2 is simply the total number of nearest-neighbor singlet pairs. But appearance is deceiving, for the commutation relations between b and b^\dagger are complicated:

$$[b_{ij}, b_{ij}^\dagger] = 1 - \frac{1}{2} (N_i + N_j), \quad (2.15)$$

$$[b_{ij}, b_{jk}^\dagger] = \frac{1}{2} \sum_s C_{i,s} C_{k,s}^\dagger, \quad (2.16)$$

where $N_i = N_{i\uparrow} + N_{i\downarrow}$. In fact, as one can readily verify, H_2 has the same matrix elements in S_0 as the Hamiltonian of a Heisenberg model.⁵ Thus, in more than one dimension, H_2 is just as insoluble as the Heisenberg model.

From either Eq. (2.7) or Eq. (2.11) we see that H_3 describes the three-site virtual process illustrated in Fig. 1 (b), which enables a singlet pair to move, provided there is an adjacent empty site. (In contrast, a triplet pair cannot move at all.) As we shall see later, this term is crucial for the onset of superconductivity upon doping.

The Hamiltonian H_1 enables an electron to hop singly to an adjacent empty site. It tends to destroy an established antiferromagnetic order, as illustrated in Fig. 2(a),

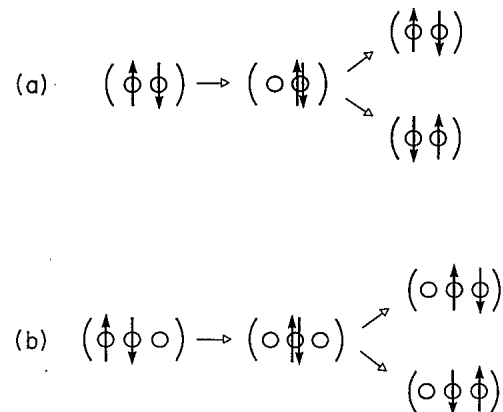


FIG. 1. (a) Virtual process that gives rise to an attraction between nearest-neighbor electrons in a singlet spin state. (b) The three-site interaction enables a singlet spin pair to move.

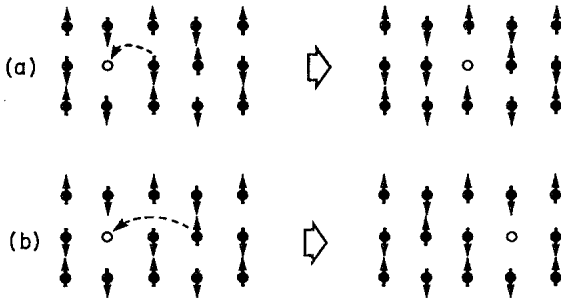


FIG. 2. (a) Moving a hole one space over an antiferromagnetic background ruins it. (b) Moving a hole two spaces, on the other hand, leaves the background undisturbed.

and thus works against any mechanism for superconductivity that is based on antiferromagnetism.

At exactly half filling we have $H_{\text{eff}} = H_2$. [Note, incidentally, that the commutator (2.15) vanishes.] In this case the system is equivalent to a Heisenberg model with antiferromagnetic coupling $-J$ in the Cu-O planes, and $-rJ$ in the orthogonal direction. The ground state is believed to be well described by small quantum fluctuations about the classical Néel state.^{6,7} If the ground state were a pure Néel state, the excited states would be separated from it by an energy gap J . The existence of quantum fluctuations leads to gapless spin-wave excitations. On dimensional grounds we expect that they have rather large velocities, of the order of $c = Ja/\hbar$, where a is the lattice spacing. Taking $J = 0.5$ eV and $a = 1$ Å for an estimate, we have $\hbar c = 0.5$ eV Å, or $c = 5 \times 10^6$ cm/s. Thus the half-filled lattice is hard to excite, and should be an insulator. These expectations, while not rigorously proven in the Hubbard model, are strongly supported by experiments performed on a La_2CuO_4 crystal.³ In particular the estimate for c given above agrees with experiments.

III. ANTIFERROMAGNETIC BACKGROUND APPROXIMATION (ABA)

Let α denote a spin configuration for the half-filled lattice, i.e., it is a list specifying the spin states of the sites. The ground state of the system, which in this case is equivalent to a Heisenberg lattice, is of the form

$$|\text{Heisenberg}\rangle = \sum_{\alpha} K(\alpha) |\alpha\rangle, \quad (3.1)$$

$$|\alpha\rangle \equiv \text{const} \times (C_1^\dagger \cdots C_M^\dagger) |0\rangle, \quad (3.2)$$

where $|0\rangle$ is the vacuum state, the state in which the lattice is entirely unoccupied. The subscript λ on C_λ^\dagger is an abbreviation for the site and spin label $\{i_\lambda, s_\lambda\}$. We must adopt a fixed convention on the ordering of the C_λ^\dagger 's, for example, the one with a higher site label should be written further to the right. We expect $K(\alpha)$ to be peaked at the Néel configuration (that is, at one of the two possible Néel configurations, the other one being decoupled in the ther-

modynamic limit), and falls off rapidly when the number of flipped spins increases.

Now consider very light doping, in which the density of holes is vanishingly small. Define a state with n holes located at $z \equiv \{z_1, \dots, z_n\}$ by

$$|z; \alpha\rangle \equiv |z_1, \dots, z_n; \alpha\rangle \equiv \text{const} \times (C_1^\dagger \cdots C_{M-n}^\dagger) |0\rangle, \quad (3.3)$$

where site labels on the C_λ^\dagger 's refer to the occupied sites, in ascending order to the right. A complete orthonormal basis for the sub-Hilbert space S_0 is generated by allowing α and z to take on all possible values. The order of the z 's in Eq. (3.3) is irrelevant. For example, for a 4-site lattice $\{z_1, z_2\}$ and $\{z_2, z_1\}$ both specify that electrons be created only at sites 3 and 4, and the order of their creation is fixed by convention. Thus, interchanging z_1 and z_2 does not change the state, not even by phase factor. In this sense the holes obey Bose statistics. This is consistent with the usual notion of holes in solids, for a hole here results from removing two electrons from a completely filled site. Since a hole cannot be created on a site where there is already a hole, there is effectively a zero-range hard-core repulsion between holes.

We can trivially rewrite

$$H_{\text{eff}} = \sum_{z, z'} \sum_{\alpha, \alpha'} |z'; \alpha'\rangle \langle z'; \alpha' | H_{\text{eff}} |z; \alpha\rangle \langle z; \alpha|. \quad (3.4)$$

For light doping the background lattice is essentially a Heisenberg lattice with holes. For given z , let us use as a basis the eigenstates of an antiferromagnetic Heisenberg model in which the sites z are taken out, by setting to zero all nearest-neighbor couplings to those sites. We have seen that spin excitations are costly in energy. On the other hand, the holes can move with relatively little effort. Thus, to study low-lying excitations associated with the motion of the holes, we may ignore all but the ground-state spin configuration, and consider a more restrictive sub-Hilbert space spanned by

$$|z; \text{Heisenberg}\rangle = \sum_{\alpha} K(\alpha) |z; \alpha\rangle, \quad (3.5)$$

where $K(\alpha)$ is the same function as in Eq. (3.1).

For our purpose the spin distribution is relevant only in neighborhoods of the holes, of next-nearest-neighbor size. Since there are very few holes, the probability of finding a flipped spin in these neighborhoods is presumed small, and will be neglected. Accordingly, we make the further approximation of replacing $K(\alpha)$ by a δ function peaked at the Néel state, and replace the basis Eq. (3.5) by

$$|z\rangle \equiv |z; \text{Néel}\rangle \equiv |z; \alpha_0\rangle, \quad (3.6)$$

where α_0 denotes the Néel spin configuration. The space spanned by these states is called the *antiferromagnetic subspace*. For a state to remain in this subspace, the motion of a hole must not disturb the antiferromagnetic background. This requires that the site left behind by a moving hole be filled by the spin at that site specified by α_0 . Like a figure in animation, a hole may block out different parts of the background scenery as it moves, but the same part is restored when it passes. To emphasize the point, we term the choice of basis Eq. (3.6) the *anti-*

ferromagnetic background approximation (ABA).

The effective Hamiltonian in ABA is given by

$$H_{\text{eff}} \approx \sum_{z,z'} |z'\rangle \langle z'| H_{\text{eff}} |z\rangle \langle z| . \quad (3.7)$$

First we observe that

$$\langle z' | H_1 | z \rangle = 0 , \quad (3.8)$$

because H_1 takes a state in the antiferromagnetic subspace out of it, as illustrated in Fig. 2(a). The virtual effects of H_1 give an estimate of the range of validity of ABA, which we shall quote later.

Next we note that H_2 is diagonal:

$$\langle z' | H_2 | z \rangle = -(2/U) \delta_{z'z} \sum_{(i,j)} t_{ij}^2 , \quad (3.9)$$

where $\langle ij \rangle$ denotes an actual nearest-neighbor pair, i.e.,

$$\langle z | H_2 | z \rangle = E_0 + 2\Delta E - \begin{cases} 0 & \text{if the holes are not NN,} \\ J/2 & \text{if the holes are in-plane NN,} \\ r^2 J/2 & \text{if the holes are interplane NN.} \end{cases} \quad (3.12)$$

More generally we can write

$$\langle z | H_2 | z \rangle = E_0 + 2\Delta E - \frac{1}{2} J(n_{\parallel} + r^2 n_{\perp}) + \dots , \quad (3.13)$$

where n_{\parallel} and n_{\perp} are, respectively, the number of NN hole pairs in the plane, and orthogonal to the plane. This implies there is an effective two-body potential between holes:

$$V(z_1, z_2) = \begin{cases} \infty & \text{if } z_1 = z_2, \\ -J/2 & \text{if } z_1, z_2 \text{ are in-plane NN,} \\ -r^2 J/2 & \text{if } z_1, z_2 \text{ are interplane NN,} \\ 0 & \text{otherwise.} \end{cases} \quad (3.14)$$

The dots in Eq. (3.13) denote contributions from the simultaneous interaction of three or more holes, which

$$\langle z_1, \dots, z_i + \mu + \nu, \dots, z_n | H_3 | z_1, \dots, z_i, \dots, z_n \rangle = -U^{-1} t(\mu) t(\nu), \quad i=1, \dots, n, \quad \mu + \nu \neq 0 . \quad (3.15)$$

In the antiferromagnetic subspace we can introduce boson operators $h^\dagger(z)$ and $h(z)$, which respectively creates and annihilates a hole at lattice position z , with

$$[h(z), h^\dagger(z')] = \delta_{z,z'} , \quad (3.16)$$

$$\sum_z h^\dagger(z) h(z) = n . \quad (3.17)$$

When a hole is annihilated by $h(z)$, the site at z is refilled by the spin that was removed from the Néel configuration in order to create that hole originally. We can now represent H_3 in the form

$$H_3 = -U^{-1} \sum_z \sum_{\substack{\mu, \nu \\ (\mu + \nu \neq 0)}} t(\mu) t(\nu) h^\dagger(z + \mu + \nu) h(z) , \quad (3.18)$$

neither i nor j can be a hole. Using Eq. (2.2) we rewrite this as

$$\langle z | H_2 | z \rangle = -\frac{1}{2} J(N_{\parallel} + r^2 N_{\perp}) , \quad (3.10)$$

where N_{\parallel} and N_{\perp} are, respectively, the number of actual nearest-neighbor pairs in the Cu—O plane, and the number of interplane pairs. By construction, all such pairs have antiparallel spins. Suppose there is only one hole present. Since it takes away four nearest-neighbor links in the plane, and two in the orthogonal direction, the energy of the system is given by

$$\langle z | H_2 | z \rangle = E_0 + \Delta E , \quad (3.11)$$

$$\Delta E = J(2 + r^2) ,$$

where E_0 is a constant. For two holes we have (using NN as abbreviation for "nearest neighbor"),

will lead to three- and higher-body potentials. All the interactions are negligible when the density of holes is small. Accordingly we neglect them entirely, and take H_2 to be a constant. The effects of the interactions will be commented on later.

We now turn to H_3 , whose action in the antiferromagnetic subspace is illustrated in Fig. 2(b). It enables an electron to "leapfrog" over its nearest neighbor (necessarily of opposite spin), and jump to an initially empty next-nearest-neighbor site. The net result is that a hole skips two lattice spacings at a time, so that it always moves in the same Néel sublattice.

For convenience we denote a lattice vector by μ or ν , and use the notation $t_{ij} = t(\mu)$, where $\{i, j\}$ mark the end-points of μ . The nonvanishing matrix elements of H_3 can then be written in the form

where the z sum extends over all lattice position vectors. For a more formal derivation of Eq. (3.18), note that Eq. (3.15) defines an n -body Hamiltonian acting on the configuration space of holes. It is then straightforward to rewrite the Hamiltonian in second-quantized form, resulting in Eq. (3.18).

We emphasize that the ABA is a kind of semiclassical approximation in which the quantum-mechanical spin fluctuations about the Néel configuration are neglected. In other words, we ignore the coupling between the holes and spin waves. The estimate given at the end of the last section, which yields a high spin-wave velocity, makes the approximation plausible for phenomena below a few hundred degrees Kelvin. An improvement to ABA should

take the spin waves into account, possibly along the lines discussed by Anderson⁸ sometime ago.

IV. THE ONSET OF SUPERCONDUCTIVITY

The Hamiltonian in ABA, which is given by H_3 up to an additive constant, can be diagonalized by Fourier analyzing $\tilde{h}(z)$:

$$\tilde{h}(z) = M^{-1/2} \sum_k e^{ik \cdot z} \tilde{h}(k), \quad (4.1)$$

$$\tilde{h}(k) = M^{-1/2} \sum_z e^{-ik \cdot z} \tilde{h}(z). \quad (4.2)$$

It is easily verified that

$$[\tilde{h}(k), \tilde{h}^\dagger(k')] = \delta_{kk'}, \quad (4.3)$$

$$\sum_k \tilde{h}^\dagger(k) \tilde{h}(k) = n. \quad (4.4)$$

In terms of the Fourier transforms we have

$$H_{\text{eff}} = \text{const} + H_3 = \text{const} + \sum_k E(k) \tilde{h}^\dagger(k) \tilde{h}(k), \quad (4.5)$$

$$E(k) = \frac{1}{2U} \sum_{\substack{\mu, \nu \\ (\mu + \nu \neq 0)}} t(\mu) t(\nu) \{1 - \cos[k \cdot (\mu + \nu)]\}. \quad (4.6)$$

The Hamiltonian is now in diagonal form. For the low-lying excitations we can take the small- k limit of $E(k)$:

$$E(k) = \frac{1}{2U} \sum_{\substack{\mu, \nu \\ (\mu + \nu \neq 0)}} t(\mu) t(\nu) [k \cdot (\mu + \nu)]^2. \quad (4.7)$$

Using Eq. (2.2) we can recast Eq. (4.7) in the form

$$E(k) = \hbar^2 \left(\frac{k_x^2 + k_y^2}{2m^*} + \frac{k_z^2}{2m_3^*} \right), \quad (4.8)$$

$$\frac{\hbar^2}{m^*} = 4Ja^2 \left(1 + \frac{r}{2} \right), \quad (4.9)$$

$$\frac{\hbar^2}{m_3^*} = 4Ja_3^2 r \left(1 + \frac{r}{2} \right), \quad (4.10)$$

where the x - y plane refers to the Cu-O plane, and a and a_3 are, respectively, the in-plane and interplane lattice constants. As expected, the effective mass m_3^* for interplane motion goes to infinity in the 2D limit, when the ratio r of hopping parameters goes to zero.

For light doping, the system is equivalent to an ideal Bose gas of holes with number conservation, and hence exhibits Bose-Einstein condensation below a critical temperature T_B . The superfluid flow of the Bose-Einstein condensate induced by an external magnetic field leads to superconductivity.

For arbitrary masses m_1 , m_2 , and m_3 for motions along the x , y , and z axes, the Bose-Einstein transition temperature (with Boltzmann's constant set to unity) is given by⁹

$$T_B = \frac{2\pi\rho^{2/3}\hbar^2}{[\zeta(3/2)]^{2/3}(m_1 m_2 m_3)^{1/3}}, \quad (4.11)$$

where ρ is the density of holes, and $\zeta(3/2) = 2.612 \dots$. In our case this gives

$$\frac{T_B}{J} = \frac{8\pi}{[\zeta(3/2)]^{2/3}} r^{1/3} \left(1 + \frac{r}{2} \right) x^{2/3}. \quad (4.12)$$

In Fig. 3 we plot T_B against the doping fraction x , for various values of the hopping parameter ratio r , with the data¹⁰ for $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ shown on the same graph. Our curves are computed for $t = 0.5$ eV and the somewhat arbitrary choice $U = 2.5$ eV. This set of parameters gives

$$J = 0.4 \text{ eV} = 4800 \text{ K}, \quad (4.13)$$

$$m^* \approx 2m_e,$$

where m_e is the mass of a free electron. The fact that the effective mass is of electronic magnitude puts the transition temperature on a high-energy scale. (That it comes out to be two electron masses is amusing but accidental.) However, Bose-Einstein condensation can take place only in three or more dimensions, and the phenomenon disappears in the 2D limit ($r \rightarrow 0$). This is reflected in the $r^{1/3}$ dependence in Eq. (4.12). To counteract the large energy scale of 4800 K sufficiently to bring the transition temperature down to the observed range of around 100 K, r has to be as small as $10^{-7} - 10^{-8}$. Thus, the lattice is almost two dimensional, in agreement with experimental findings,³ but it is not known whether these small values of r are quantitatively correct.

The theory predicts that, at least for small doping, one can increase the critical temperature by increasing the interplane hopping. In practice, this may be accomplished through a structural change of the lattice to bring the planes closer, or application of external pressure. The pressure dependence of the critical temperature cannot be predicted within this model, which takes the hopping parameters as given. The conclusion that a more nearly three-dimensional lattice has a higher critical temperature agrees with that of Lee and Read,¹¹ though for an apparently different reason.

Taking the effective potential Eq. (3.14) into account, we have a dilute nonideal Bose gas, whose properties are well understood.¹² The interaction can be summarized by a single parameter, the scattering length D , which may be thought of as an effective hard-core diameter. The small dimensionless parameter in the problem is $\epsilon = \rho^{1/3} D$. The transition temperature changes only by $O(\epsilon)$, which is not significant. The most important effect of the interactions is to alter the single-particle excitation spectrum from being particlelike, as in Eq. (4.7), to being phononlike, with sound velocity proportional to $\epsilon^{1/2}$. This gives a specific heat proportional to T^3 near absolute zero, instead of the $T^{3/2}$ behavior of the ideal Bose gas. The width of the temperature region in which the cubic law applies, however, is $O(\epsilon)$.

The linear phonon spectrum is important in the present application, for any scattering among quasiparticles with a quadratic spectrum will destroy the superfluidity of the condensate. The transition from single-particle behavior to collective-phonon-like behavior, as well as the interplay between them, is well understood.¹³

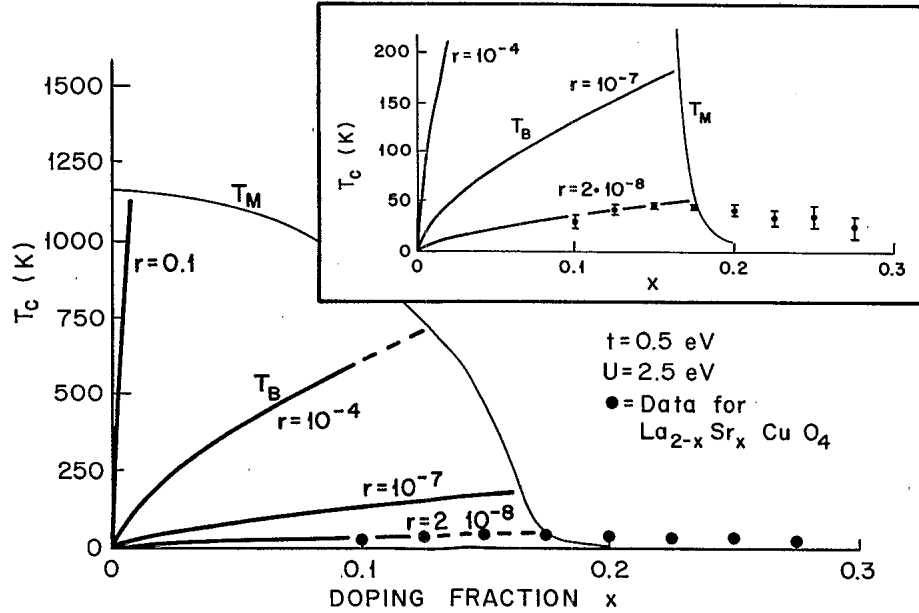


FIG. 3. Critical temperature as a function of doping fraction. The parameter $r = t_3/t$ is the ratio between the hopping parameters normal to and in the Cu—O plane. Different scenarios give the branches T_B (from Bose-Einstein condensation) and T_M (from mean-field approximation). T_M is insensitive to r , but T_B depends on it. The data are from Ref. 10.

When the fraction of holes becomes significant, the antiferromagnetic background loses meaning, and ABA must breakdown. The cause lies in the increasing importance of H_1 , which tends to disrupt the antiferromagnetic order. The second-order energy shift due to H_1 is of the order of nt^2/J , where n is the number of holes, t^2 comes from the squared matrix element of H_1 , and $J = 4t^2/U$ is a typical energy denominator, the price of a flipped spin. Thus the energy shift per lattice site is $Ux/4$, where x is the doping fraction. For ABA to be valid this must be much smaller than the energy per site due to H_2 , which is of order $J = 4t^2/U$. Therefore a criterion for the validity for ABA is

$$x \ll (4t/U)^2. \quad (4.14)$$

For our choice of parameters the condition is $x \ll 0.6$.

V. MEAN-FIELD THEORY

There is yet no reliable method to treat the problem beyond ABA. For qualitative orientation, one can turn to mean-field theory, which has been discussed by Anderson, Baskaran, Zou, and Hsu¹⁴ and by Ruckenstein, Hirschfeld, and Appel.¹⁵ The treatment consists of three drastic approximations: (a) replacing the operator ξ in Eq. (2.8) by the doping fraction x (a poor approximation for small x); (b) neglecting the three-site term H_3 (justifiable in hindsight); (c) linearizing H_2 (a practical necessity).

We shall denote by $\tilde{C}_{k,s}$ the Fourier transform of $C_{i,s}$, and use μ to denote both a lattice vector and the direction of that lattice vector. For example, we shall denote the component of k along μ by k_μ . The mean-field Hamiltonian

is given by

$$H_{MF} = \sum_{k,s} (e_k - \bar{\mu}) \tilde{C}_{k,s}^\dagger \tilde{C}_{k,s} - \sum_k J_k (\tilde{C}_{k1}^\dagger \tilde{C}_{-k1}^\dagger \Delta + \text{H.c.}), \quad (5.1)$$

$$\Delta = (Md)^{-1} \sum_{\langle i,j \rangle} \langle C_{i1} C_{j1} - C_{i1}^\dagger C_{j1}^\dagger \rangle, \quad (5.2)$$

$$e_k = - \sum_\mu (2xt_\mu + PJ_\mu) \cos(k_\mu a_\mu), \quad (5.3)$$

$$P = (Md)^{-1} \sum_{\langle i,j \rangle} \langle C_{i,s}^\dagger C_{j,s} \rangle, \quad (5.4)$$

$$J_k = \sum_\mu J_\mu \cos(k_\mu a_\mu), \quad (5.5)$$

where $\bar{\mu}$ is a constant (the chemical potential). This Hamiltonian is of the BCS type, and can be diagonalized by a Bogoliubov transformation. The gap and chemical potential equations are as follows:

$$1 = \frac{1}{Md} \sum_k \frac{J_k}{E_k} \sum_\mu \cos(k_\mu a_\mu) \tanh(\beta E_k/2), \quad (5.6)$$

$$x = \frac{1}{M} \sum_k \frac{e_k - \bar{\mu}}{E_k} \tanh(\beta E_k/2), \quad (5.7)$$

$$E_k = \sqrt{(e_k - \bar{\mu})^2 + \Delta^2 J_k^2}, \quad (5.8)$$

where β is the inverse temperature. Numerical factors in these equations disagree with the corresponding ones in Refs. 14 and 15. In this respect the quoted references also disagree with each other.

The critical temperature T_M is the temperature at which $\Delta = 0$, and is obtained by solving the above equations numerically, with the result shown in Fig. 3. As a

function of x , it is insensitive to r , and the dimensionality d of the lattice. The important model parameters are t and J . It can be shown analytically that the curve extrapolates to $J/4$ at $x=0$, and vanishes at a value determined by J/t . With t fixed by electronic calculations at 0.5 eV, there is essentially only one adjustable parameter J .

The Cooper pairs here are bound by the attraction between spin-singlet pairs, arising from H_2 . They have zero total momentum, and therefore no translational motion as a whole, and thus the neglect of H_3 is a self-consistent approximation. The stability of the Cooper pairs is undermined by H_1 , which encourages electrons to hop singly to empty sites. When the number of empty sites increases, so does the importance of H_1 . This is why T_M decreases with x , and eventually vanishes.

The work in Ref. 15 actually takes H_3 into account, but linearizes it. The effect of a linearized H_3 is to enhance the stability of the singlet pairs against transition to a triplet state but in the mean-field approximation it has only a small effect on the critical temperature.

Confrontation with the data in Fig. 3 shows that the mean-field approximation is rather poor. No value of J can make the mean-field curve fit the data in any quantitative sense.

VI. DISCUSSION

As we found in ABA, whose basis and results are not in disagreement with experiments, the mechanism for the onset of superconductivity is Bose-Einstein condensation. What sustains superconductivity at moderate doping is

less certain, for the mean-field theory used in this regime is inadequate. For lack of an alternative, however, let us take the suggestion of mean-field theory that the mechanism is Cooper-pair formation (CP). We would like to compare these two different pictures.

Although the ABA and the CP scenarios cannot be bridged by a common set of variables, it is instructive to try to describe them in a common framework, even though we may have to resort to ill-defined concepts. In both schemes there are bosons arising from the binding of electron pairs, and we can think of the system loosely as a superfluid of composite bosons. In the CP case the composite bosons are obviously the Cooper pairs. In ABA they are the holes, whose dynamics originates in the anti-ferromagnetic order. In either case we can associate two energy scales with the system: a binding energy Δ , and a critical temperature T_B . The binding energy measures the integrity of the bosons as particles, and the critical temperature measures the cohesiveness of the Bose-Einstein condensate. In ABA we have $T_B \ll \Delta$. When the system heats up, the Bose-Einstein condensate evaporates long before the bosons dissociate. This is why Δ is irrelevant in this regime. In the CP regime we have just the opposite, i.e., $T_B \gg \Delta$. Upon heating, the bosons quickly dissociate, and T_B never had a chance to enter the picture.

ACKNOWLEDGMENT

This work is supported in part by funds provided by the U.S. Department of Energy under Contract No. DE-AC02-76ER03069.

*Present address: Supercomputer Computations Research Institute, Florida State University, Tallahassee, FL 32306.

¹J. G. Bednorz and K. A. Müller, *Z. Phys. B* **64**, 188 (1986); S. Uchida, H. Takagi, K. Kawaski, and S. Tanaka, *Jpn. J. Appl. Phys. Pt. 2*, **26**, L1 (1987); Z. Zhao *et al.*, *Kexue Tongbao (China)*, March (1987); C. W. Chu, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, and Y. Q. Wang, *Phys. Rev. Lett.* **58**, 405 (1987); R. Cava, R. B. van Dover, B. Batlogg, and E. A. Rietman, *ibid.* **58**, 408 (1987).

²P. W. Anderson, *Science* **235**, 1196 (1987).

³R. J. Birgeneau *et al.*, *Phys. Rev. Lett.* **59**, 1329 (1987). G. Shirane *et al.*, *ibid.* **59**, 1613 (1987).

⁴L. F. Matheiss, *Phys. Rev. Lett.* **58**, 1028 (1987); J. Yu, A. J. Freeman, and J. H. Xu, *ibid.* **58**, 1035 (1987).

⁵V. J. Emery, *Phys. Rev. B* **14**, 2989 (1976).

⁶P. W. Anderson, *Mater. Res. Bull.* **8**, 153 (1973); P. Fazekas and P. W. Anderson, *Philos. Mag.* **30**, 432 (1974); J. Oitmaa

and D. D. Betts, *Can. J. Phys.* **56**, 897 (1978).

⁷J. E. Hirsch, *Phys. Rev. B* **31**, 4403 (1985); *Phys. Rev. Lett.* **54**, 1317 (1985).

⁸P. W. Anderson, *Phys. Rev. B* **86**, 694 (1952).

⁹See, K. Huang, *Statistical Mechanics, 2nd ed.* (Wiley, New York, 1987), Sec. 12.3.

¹⁰J. M. Tarascon, L. H. Greene, W. R. McKinnon, G. W. Hull, and T. H. Geballe, *Science*, **235**, 1373 (1987).

¹¹P. A. Lee and N. Read, *Phys. Rev. Lett.* **58**, 2691 (1987).

¹²See Ref. 9, Secs. 12.4 and 13.8.

¹³See E. Manousakis, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1985 (unpublished).

¹⁴P. W. Anderson, G. Baskaran, Z. Zou, and T. Hsu, *Phys. Rev. Lett.* **58**, 2790 (1987).

¹⁵A. E. Ruckenstein, P. J. Hirschfeld, and J. Appel, *Phys. Rev. B* **36**, 857 (1987).