

momentum operators by applying projection operators and we form an orthogonal set. For the uniform solution,

$$S_3 = i \sum_{\mathbf{k}_s > 0} (L_{\mathbf{k}} \psi_{\mathbf{k}}^\dagger \psi_{-\mathbf{k}}^\dagger + cc).$$

Noting that the method has a variational aspect, $L_{\mathbf{k}}$ need not be taken from the small-oscillation analysis but may be freely chosen to describe average large-amplitude effects. We obtain improved ground-state energy and single- and multiple-excitation spectra.

For the solid-like solution, $f(\mathbf{x})$ is periodic. We expand $\psi = \sum a_{\mathbf{k}}^\alpha \varphi_{\mathbf{k}}^\alpha(\mathbf{x})$, where $[a_{\mathbf{k}}^\alpha, a_{\mathbf{l}}^{\dagger\beta}] = \delta_{\alpha, \beta} \delta_{\mathbf{k}, \mathbf{l}}$. Here $\varphi_{\mathbf{k}}^\alpha$ are a complete set of Bloch tight-binding orbitals for which \mathbf{k} takes on values in the first zone; α labels the zones. For $\mathbf{k} = 0$ the $\varphi_{\mathbf{k}}^\alpha$ are periodic; for $\mathbf{k} \neq 0$ they have a modulating factor. Thus if the linear shift is performed only for the a_0^α , the ground-state expectation values of physical quantities are periodic. If shifts for $\mathbf{k} \neq 0$ are required, the expectation value of the correlation operator ceases to be periodic.

The connection between the two solutions is seen by referring to the quantum problem of a particle in a well with several minima (or stationary points). Because of the tunnel effect, good approximate wave functions are superpositions of functions appropriate to the classical separate regions. By analogy, we take

$$\Psi = \mathcal{P}(N) \mathcal{P}(\mathbf{P}) \int G(R) \exp S_2(R) \cdot \exp S_3(R) \Phi(\dots N_{\mathbf{k}} \dots) dR.$$

The coefficients of the linear and quadratic forms depend on R ; the integral over R includes a discrete sum; $\mathcal{P}(N)$ and $\mathcal{P}(\mathbf{P})$ are projection operators of total number of particles N with total momentum \mathbf{P} . Detailed calculations of properties of liquid and solid helium based on the present approach are in progress.

¹ E. P. Gross, Phys. Rev. **100**, 1571 (1955).

² N. Bogolyubov, J. Phys. (U.S.S.R.) **11**, 23 (1947).

Microscopic Theory of Superconductivity*

J. BARDEEN, L. N. COOPER, AND J. R. SCHRIEFFER

Department of Physics, University of Illinois, Urbana, Illinois

(Received February 18, 1957)

SINCE the discovery of the isotope effect, it has been known that superconductivity arises from the interaction between electrons and lattice vibrations, but it has proved difficult to construct an adequate theory based on this concept. As has been shown by Fröhlich,¹ and in a more complete analysis by Bardeen and Pines² in which Coulomb effects were included, interactions between electrons and the phonon field lead to an interaction between electrons which may be

expressed in the form

$$H_I = \sum_{\mathbf{k}, \mathbf{k}', s, s'} \frac{\hbar\omega |M_{\mathbf{k}}|^2}{(E_{\mathbf{k}} - E_{\mathbf{k}'})^2 - (\hbar\omega)^2} \times c_{\mathbf{k}' - \mathbf{k}, s'}^* c_{\mathbf{k}, s} c_{\mathbf{k}', s'}^* c_{\mathbf{k} + \mathbf{k}, s} + H_{\text{Coul}}, \quad (1)$$

where $|M_{\mathbf{k}}|^2$ is the matrix element for the electron-phonon interaction for the phonon wave vector \mathbf{k} , calculated for the zero-point amplitude of the vibrations, the c 's are creation and destruction operators for the electrons in the Bloch states specified by the wave vector \mathbf{k} and spin s , and H_{Coul} represents the screened Coulomb interaction.

Early attempts³ to construct a theory were based essentially on the self-energy of the electrons, although it was recognized that a true interaction between electrons probably played an essential role. These theories gave the isotope effect, but contained various difficulties, one of which was that the calculated energy difference between what was thought to represent normal and superconducting states was far too large. It is now believed that the self-energy occurs in the normal state, and results in a slight shift of the energies of the Bloch states and a renormalization of the matrix elements.

The present theory is based on the fact that the phonon interaction is negative for $|E_{\mathbf{k}} - E_{\mathbf{k}'}| < \hbar\omega$. We believe that the criterion for superconductivity is essentially that this negative interaction dominate over the matrix element of the Coulomb interaction, which for free electrons in a volume Ω is $2\pi e^2/\Omega \kappa^2$. In the Bohm-Pines⁴ theory, the minimum value of κ is κ_e , somewhat less than the radius of the Fermi surface. This criterion may be expressed in the form

$$-V = \langle -(|M_{\mathbf{k}}|^2/\hbar\omega) + (4\pi e^2/\Omega \kappa^2) \rangle_{av} < 0. \quad (2)$$

Although based on a different principle, this criterion is almost identical with the one given by Fröhlich.^{1,3}

If one has a Hamiltonian matrix with predominantly negative off-diagonal matrix elements, the ground state, $\Psi = \sum \alpha_j \varphi_j$, is a linear combination of the original basic states with coefficients predominantly of one sign. A particularly simple example is one for which the original states are degenerate and each state is connected to n other states by the same matrix element $-V$. The ground state, a sum of the original set with equal coefficients, is lowered in energy by $-nV$. One of the authors made use of this principle to construct a wave function for a single pair of electrons excited above the Fermi surface and found that for a negative interaction a bound state is formed no matter how weak the interaction.⁵

Because of the Fermi-Dirac statistics, difficulties are encountered if one tries to apply this principle directly to (1). Matrix elements of H_I between states specified by occupation numbers (Slater determinants) in general may be of either sign. We want to pick out

a subset of these between which matrix elements are always of the same sign. This may be done by occupying the individual particle states in pairs, such that if one of the pair is occupied, the other is also. The pairs should be chosen so that transitions between them are possible, i.e., they all have the same total momentum. To form the ground state, the best choice is $\mathbf{k}\uparrow, -\mathbf{k}\downarrow$, since exchange terms reduce the matrix elements between states of parallel spin. To form a state with a net current flow, one might take a pairing $\mathbf{k}\uparrow, -\mathbf{k}+\mathbf{q}\downarrow$, where \mathbf{q} is a small wave vector, the same for all \mathbf{k} and such that both states are within the range of energy $\hbar\omega$. The occupation of the pairs may be specified by a single spin-independent occupation number, $m_{\mathbf{k}}=0$ or 1. Nonvanishing matrix elements connect configurations which differ in only one of the occupied pairs.⁶ It is often convenient to specify occupation in terms of electron pairs above the Fermi surface and hole pairs below.

The best wave function of this form will be a linear combination

$$\Psi = \sum_{\mathbf{k}_1 \cdots \mathbf{k}_n} b(\mathbf{k}_1 \cdots \mathbf{k}_n) f(\cdots m_{\mathbf{k}_1} \cdots m_{\mathbf{k}_n} \cdots), \quad (3)$$

where the sum is over all possible configurations. In our calculations, we have made a Hartree-like approximation and replaced b by $b(k_1)b(k_2)\cdots b(k_n)$. We have also assumed an isotropic Fermi surface [so that $b(k)$ depends only on the energy ϵ of the Bloch state involved], and that V is the same for all transitions within a constant energy $\hbar\omega$ of the Fermi surface, $\epsilon=0$. A direct calculation gives for the interaction energy

$$W_I = -4[N(0)]^2 V \int_0^{\hbar\omega} \int_0^{\hbar\omega} \Gamma(\epsilon)\Gamma(\epsilon') d\epsilon d\epsilon', \quad (4)$$

where $N(0)$ is the density of states at the Fermi surface. The kinetic energy measured from the Fermi sea is

$$W_K = 4N(0) \int_0^{\hbar\omega} g(\epsilon) \epsilon d\epsilon, \quad (5)$$

where $g(\epsilon)$ is the probability that a given state of energy ϵ is occupied by a pair, and

$$\Gamma(\epsilon) = \{g(\epsilon)[1-g(\epsilon)]\}^{\frac{1}{2}}. \quad (6)$$

One may interpret the factor $\Gamma(\epsilon)\Gamma(\epsilon')$ as representing the effect of the exclusion principle on restricting the number of configurations which are connected to a given typical configuration. Matrix elements corresponding to $\mathbf{k}\rightarrow\mathbf{k}'$ are possible only if the state \mathbf{k} is occupied and \mathbf{k}' unoccupied in the initial configuration and \mathbf{k}' occupied and \mathbf{k} unoccupied on the final configuration. The probability that this occurs is

$$g(\epsilon)[1-g(\epsilon')]g(\epsilon')[1-g(\epsilon)] = [\Gamma(\epsilon)]^2[\Gamma(\epsilon')]^2. \quad (7)$$

Since matrix elements have probability amplitudes

rather than probabilities, the square root of (7) occurs in (4).

A variational calculation to determine the best $g(\epsilon)$ gives

$$W = W_I + W_K = -\frac{2N(0)(\hbar\omega)^2}{\exp[2/N(0)V]-1}. \quad (8)$$

Thus if there is a net negative interaction, no matter how weak, there is a condensed state in which pairs are virtually excited above the Fermi surface. The product $N(0)V$ is independent of isotopic mass and of volume. The energy W varies as $(\hbar\omega)^2$, in agreement with the isotope effect. It should be noted that (8) cannot be obtained in any finite order of perturbation theory. The energy gain comes from a coherence of the electron wave functions with lattice vibrations of short wavelength, and does not represent a condensation in real space.

Empirically, energies are of the order of magnitude of $N(0)(kT_e)^2$, and of course kT_e is much less than an average phonon energy $\hbar\omega$. According to our theory, this will occur if $N(0)V < 1$, a not unreasonable assumption. In this weak-coupling limit, the energy may be expressed simply in terms of the number of electrons, n_e , virtually excited in coherent pairs above the Fermi surface at $T=0^\circ\text{K}$

$$W = -\frac{1}{2}n_e^2/N(0), \quad (9)$$

where

$$n_e = 2N(0)\hbar\omega \exp[-1/N(0)V]. \quad (10)$$

It is a great advantage energy-wise to include in the ground state wave function only pairs with the same total momentum. Suppose that instead one had chosen a random pairing, $\mathbf{k}_1\uparrow, \mathbf{k}_2\downarrow$, with $\mathbf{k}_1+\mathbf{k}_2=\mathbf{q}$ and consider a typical matrix element $(\mathbf{k}_1, \mathbf{k}_2 | H_I | \mathbf{k}_1', \mathbf{k}_2')$ which vanishes unless $\mathbf{k}_1'+\mathbf{k}_2'=\mathbf{q}'=\mathbf{q}$. We shall assume that the \mathbf{q} 's of all pairs are small so that if \mathbf{k}_1 and \mathbf{k}_2 are both within $\hbar\omega$ of the Fermi surface, so are \mathbf{k}_1' and \mathbf{k}_2' . If we construct a wave function made up of a linear combination of states with such virtual excited pairs and determine the interaction energy, we would find an expression similar to (4) but with (7) replaced by the much smaller quantity:

$$g(\epsilon_1)g(\epsilon_2)[1-g(\epsilon_1)'][1-g(\epsilon_2)']g(\epsilon_1') \times g(\epsilon_2')[1-g(\epsilon_1)][1-g(\epsilon_2)]. \quad (11)$$

The pairing $\mathbf{k}_2 = -\mathbf{k}_1$ corresponds to $\mathbf{q}=0$ for all pairs and insures that if \mathbf{k}_1' is unoccupied, so is \mathbf{k}_2' . This is also true if all pairs have the same \mathbf{q} .

Wave functions corresponding to individual particle excitations may be made of linear combinations of states in which certain occupation numbers, corresponding to real excited electrons or holes, are specified and the rest are used to make all possible combinations of virtual excitations of $\mathbf{k}\uparrow, -\mathbf{k}\downarrow$ pairs. Because of the reduction in phase space available to the pairs, the interaction energy is reduced in magnitude. For small

excitations the consequent increase in total energy is proportional to the number of excited electrons. This means that a finite energy is required to excite an electron from the ground state. The same applies to real excited \mathbf{k} , $-\mathbf{k}$ pairs. If $f(\epsilon)$ is the probability that a Bloch state of energy ϵ is occupied by an excited electron above the Fermi sea, and $1-f(-\epsilon)$ the probability that there is a hole below, one finds for the interaction energy an expression similar to (4) but with $[\Gamma(\epsilon)]^2$ replaced by $g(\epsilon)\{1-[\Gamma(\epsilon)]^2-g(\epsilon)\}$. For small excitations above $T=0^\circ\text{K}$, the total pair energy may be expressed in the weak-coupling limit as

$$W = -\frac{n_c^2}{2N(0)} \left(1 - \frac{4n_e}{n_c}\right), \quad n_e \ll n_c, \quad (12)$$

where n_c is the number of electrons in the virtually excited states at $T=0$ and n_e is the number of actually excited electrons. This leads to an energy gap⁷ (i.e., the energy required to create an electron-hole pair):

$$E_G = \partial W / \partial n_e = 2n_e / N(0) \quad \text{at } T=0^\circ\text{K}. \quad (13)$$

Taking the empirical $W = -H_c^2 / 8\pi$ and estimating $N(0)$ from the electronic specific heat, we find $E_G = k \times 13.8^\circ\text{K}$ for tin. This is to be compared with the experimental value of about $k \times 11.2^\circ\text{K}$. Calculations are under way to determine the thermal properties at higher temperatures.

Advantages of the theory are (1) It leads to an energy-gap model of the sort that may be expected to account for the electromagnetic properties.⁸ (2) It gives the isotope effect. (3) An order parameter, which might be taken as the fraction of electrons above the Fermi surface in virtual pair states, comes in a natural way. (4) An exponential factor in the energy may account for the fact that kT_c is very much smaller than $\hbar\omega$. (5) The theory is simple enough so that it should be possible to make calculations of thermal, transport, and electromagnetic properties of the superconducting state.

* This work was supported in part by the Office of Ordnance Research, U. S. Army. One of us (J.R.S.) wishes to thank the Corning Glass Works Foundation for a grant which aided in the support of this work.

¹ H. Fröhlich, Phys. Rev. **79**, 845 (1950); Proc. Roy. Soc. (London) **A215**, 291 (1952).

² J. Bardeen and D. Pines, Phys. Rev. **99**, 1140 (1955).

³ For reviews of this work, mainly by Fröhlich and by Bardeen, see J. Bardeen, Revs. Modern Phys. **23**, 261 (1951); *Handbuch der Physik* (Springer-Verlag, Berlin, 1956), Vol. 15, p. 274.

⁴ See D. Pines, *Solid State Physics* (Academic Press, Inc., New York, 1955), Vol. 1, p. 367.

⁵ L. N. Cooper, Phys. Rev. **104**, 1189 (1956).

⁶ This looks deceptively like a one-particle problem, but is not, because of the peculiar statistics involved. The pairs cannot be treated as bosons.

⁷ Some references to experimental evidence for an energy gap are R. E. Glover and M. Tinkham, Phys. Rev. **104**, 844 (1956); M. Tinkham, Phys. Rev. **104**, 845 (1956); Blevins, Gordy, and Fairbank, Phys. Rev. **100**, 1215 (1955); Corak, Goodman, Satterthwaite, and Wexler, Phys. Rev. **102**, 656 (1956); W. S. Corak and C. B. Satterthwaite, Phys. Rev. **102**, 662 (1956).

⁸ J. Bardeen, Phys. Rev. **97**, 1724 (1955).

Polarization of Nuclei by Resonance Saturation in Paramagnetic Crystals*†

C. D. JEFFRIES

Physics Department, University of California, Berkeley, California

(Received February 15, 1957)

THE suggestion of Overhauser¹ that the saturation of the spin resonance of the conduction electrons in a metal should give nuclear polarizations of the order of $\beta H / kT$ (β = Bohr magneton, H = magnetic field) catalyzed thinking along these lines; it was soon realized by many² that related dynamic polarization processes are apropos to paramagnetic substances in general. In fact, even earlier, Pound³ had produced enhanced populations by nuclear magnetic resonance saturation in a system with quadrupole splitting. Abragam⁴ discusses the nuclear polarization obtainable by the saturation of the resolved paramagnetic resonance hfs lines in magnetically dilute crystals. In his scheme the strongly allowed electronic magnetic dipole transitions are saturated and the nuclear polarization is induced by suitable relaxation processes through the hfs coupling. We wish to point out that in many cases the saturation of certain so-called forbidden transitions will produce a comparable nuclear polarization directly in the sense that the applied radio-frequency field itself flips the nuclei. Such forbidden transitions are commonly observed in microwave paramagnetic resonance, e.g., in the case of appreciable nuclear quadrupole interactions and in cases where the nondiagonal magnetic hfs terms are not too small. As a specific sample of the latter, consider the following spin Hamiltonian,⁵ appropriate for Co^{++} ions at low temperatures in a magnetically dilute axial crystal in an external magnetic field:

$$\mathcal{H} = \beta [g_{\parallel} H_z S_z + g_{\perp} (H_x S_x + H_y S_y)] + A I_z S_z + B (I_x S_x + I_y S_y) + \mathcal{H}_{\text{relax}} + \mathcal{H}_{\text{rf}}$$

The first two terms are the electronic Zeeman terms and are assumed to be much larger than the magnetic hyperfine terms in A and B . The energy levels are shown schematically in Fig. 1 where we have taken $A \simeq B$, $S = \frac{1}{2}$ and $I = 2$, for illustration purposes. The various states are characterized in zero order by electronic and nuclear magnetic quantum numbers m and M , respectively. However, the term in B mixes the states so that to first order we have for the wave functions $\psi_a = \psi(\frac{1}{2}, 2)$, $\psi_b \simeq \psi(\frac{1}{2}, 1) + (B/H_z)\psi(-\frac{1}{2}, 2)$, \dots , $\psi_j \simeq \psi(-\frac{1}{2}, 2) - (B/H_z)\psi(\frac{1}{2}, 1)$. The $\mathcal{H}_{\text{relax}}$ term gives relaxation transitions between the various states, the dominant ones being those shown for ($\Delta m = \pm 1$, $\Delta M = 0$), corresponding to the electron spin-lattice relaxation. The $\Delta(m+M) = 0$ relaxation transitions are considerably weaker, as are also those for ($\Delta m = 0$, $\Delta M = \pm 1$), not shown. For simplicity the latter transitions are neglected; from reasonable assumptions concerning relaxation mechanisms it can be shown