PHYSICAL REVIEW VOLUME 176, NUMBER 2 10 DECEMBER 1968

Ground-State Energy of the Anderson Model by the Cluster Variation Method

SHYAMALENDU M. BOSE
Department of Physics, The Catholic University of America, Washington, D. C. 20017
AND
TOMOYASU TANAKA
Department of Physics, The Catholic University of America, Washington, D. C. and U. S. Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

(Received 23 July 1968)

The Anderson Hamiltonian has been treated by the application of the cluster variation method for the cooperative phenomena, which was first proposed by Kikuchi and later generalized by Morita. The calculation has been carried out by retaining up to two-particle off-diagonal correlations. The self-consistent equations obtained for the number of d electrons with opposite spins are found to have similar behavior and similar self-consistent solutions to those of Anderson. The ground-state energy of this system has been calculated for the cases of infinite d-d correlation energy and for a d-d correlation energy twice the energy of the d orbital. This ground-state energy is found to be lower than the free-electron energy by an amount $2N(0)D^2\exp[-|\epsilon_d|/N(0)|V|^2]$, where $N(0)$ is the density of conduction states of one spin at the Fermi surface, D is the conduction bandwidth, ϵ_d is the energy of the d orbital measured from the Fermi surface, and V is the mixing strength between the conduction-electron states and the d orbital.

I. INTRODUCTION

THE discovery of the existence of a logarithmic term in the conduction electron-magnetic impurity scattering amplitude by Kondo[1] and the successful explanation of the presence of a resistance minimum at low temperatures led many investigators to study the various physical properties of dilute magnetic alloys. However, since the Kondo treatment is valid only at high temperatures, various attempts have been made to improve and to extend Kondo's theory to lower temperatures. In particular, the analysis of the nature of the ground state of such a system has become a subject of many theoretical investigations. The problem is, therefore, drastically simplified and reduced to find the ground state of a magnetic impurity of spin magnitude $\frac{1}{2}$ that interacts antiferromagnetically with conduction electrons (the s-d Hamiltonian).

Nagaoka[2] treated this problem by the method of two-time Green's function. By introducing an *ad hoc* truncation scheme in his Green's function equations he obtained a set of simultaneous integral equations. An approximate solution of the integral equations for the antiferromagnetic coupling led to the existence of a quasibound state.

Suhl,[3] and Suhl and Wong,[4] applied scattering theory to the problem and showed that no bound state can exist for such a system. Abrikosov[5] employed infinite-order perturbation theory and also reached the conclusion that a bound state should not exist. Hamann,[6] and Falk and Fowler,[7] later solved Nagaoka's truncated Green's function equations rigorously and found solutions exactly identical to Suhl and Wong's solution.

Yosida,[8] Okiji,[9] Heeger and Jensen,[10] Kondo,[11] and others have investigated the problem by the Rayleigh-Ritz perturbation method. In such an approach the ground state is assumed to be singletlike. Yosida constructed the trial wave function with one localized quasiparticle above the Fermi surface together with an undisturbed Fermi sea. This quasiparticle was then allowed to interact with the impurity via the s-d exchange interaction forming a singlet configuration. By using a modified perturbation method he derived a bound state for the case of antiferromagnetic coupling, whereas there was no bound state for the ferromagnetic coupling. The main objection to Yosida's theory is that his trial wave function does not preserve the particle-hole symmetry. In their calculation, Heeger and Jensen incorporate the particle-hole symmetry in a BCS-type wavefunction. Considering only an antiferromagnetic coupling they obtain a bound state that is the same as that obtained by Yosida. A more refined variational perturbation by Okiji has verified Yosida's calculation.

On the other hand, Kondo[12] uses a perturbation scheme to obtain the ground-state energy. He assumes that the ground-state energy has two parts; one is an analytic function of exchange integral J and the other is an infinite series of nonanalytic functions of J. Kondo calculated the first three terms of this series explicitly

---

[*] This work was supported in part by U.S. Air Force Office of Scientific Reseach Grant AF-AFOSR-445-66 and partly by National Aeronautics and Space Administration Research Grant No. NsG 649.

[1] J. Kondo, Progr. Theoret. Phys. (Kyoto) **32**, 37 (1964).

[2] Y. Nagaoka, Phys. Rev. **138**, A1112 (1967).

[3] H. Suhl, Phys. Rev. **138**, 515 (1965); Physics **2**, 39 (1965); Phys. Rev. **141**, 483 (1966).

[4] H. Suhl and D. Wong, Physics **3**, 17 (1967).

[5] A. A. Abrikosov, Physics **2**, 5 (1965).

[6] D. R. Hamann, Phys. Rev. **158**, 570 (1967).

[7] D. Falk and M. Fowler, Phys. Rev. **158**, 567 (1967).

[8] K. Yosida, Phys. Rev. **147**, 223 (1966); Progr. Theoret. Phys. (Kyoto) **36**, 875 (1966).

[9] A. Okiji, Progr. Theoret. Phys. (Kyoto) **36**, 875 (1966).

[10] A. J. Heeger and M. A. Jensen, Phys. Rev. Letters **18**, 485 (1967).

[11] J. Kondo, Progr. Theoret. Phys. (Kyoto) **36**, 429 (1966).

[12] J. Kondo, Phys. Rev. **154**, 644 (1967).

and assumed the remainder to be the terms of a geometric series, and carried out the summation. The ground-state energy obtained by Kondo is lower than those obtained from variational calculations.

Recently, Schrieffer and Wolff[13] have shown that by a canonical transformation of the Anderson Hamiltonian one gets the *s*-*d* exchange Hamiltonian plus some spin-independent terms, and thus the two models are equivalent. However, Dworin[14] has shown that the resistivities calculated from the models are different from each other. Also, more recently, Horwitz, Alexander, and Fibich[15] have claimed from the detailed singlet-triplet formulation of the two models that they are not exactly equivalent.

Efforts have also been directed to the study of the bound state for the Anderson model.[16] Starting with the Anderson Hamiltonian, Appelbaum[17] used the perturbation approach of Kondo[12] to obtain the ground-state energy. He too calculates the first few terms of the infinite series explicitly and then guesses the succeeding terms to be a geometric series and carries out the summation of the series. As in the case of the *s*-*d* model, he finds a negative exponential as the ground-state energy. Anderson,[18] by employing the variational technique, also obtains an energy lowering of the system in the ground state.

In order to investigate the question of the equivalence of the two models in detail we thought it would be useful to investigate the Anderson model by still another method. The approach here will be to treat the problem by the technique of the cluster variation method of cooperative phenomena. The method was first proposed by Kikuchi[19] and later generalized by Morita.[20] The variational free energy of a system is constructed by introducing correlations among different particles, which are then expressed in terms of the Hamiltonian and the trial reduced density matrices. By expressing the trial free energy in such a way, we can introduce approximations in a systematic manner. The number of terms that we retain depends on the nature of the problem. If the trial free energy is truncated after keeping correlations among *n* different particles of the system, we say we are treating clusters of *n* particles. The cluster variation method has been applied to the Heisenberg model with arbitrary spin and range of exchange by Morita and Tanaka,[21] and more recently by Halow, Tanaka, and Morita[22] to the calculation of the ground-state energy of the *s*-*d* Hamiltonian. Here we shall follow the method of the latter paper where the reduced density matrices are expressed in terms of unknown expectation values and constructed such that their normalization and reducibility conditions are automatically satisfied, thus avoiding the problem of calculating the unknown Lagrange multipliers which would otherwise have to be introduced to satisfy these conditions. The variation of the trial free energy is then taken with respect to these unknown elements of the reduced density matrices.

II. CLUSTER VARIATION METHOD APPLIED TO ANDERSON MODEL

The Hamiltonian of our system may be written as

$$H=\sum_{\mathbf{k},\sigma}\epsilon_{\mathbf{k}}c_{\mathbf{k}\sigma}{}^{*}c_{\mathbf{k}\sigma}+\sum_{\sigma}\epsilon_{d}c_{d\sigma}{}^{*}c_{d\sigma}+U c_{d\sigma}{}^{*}c_{d\sigma}c_{d-\sigma}{}^{*}c_{d-\sigma}+V\sum_{\mathbf{k},\sigma}(c_{\mathbf{k}\sigma}{}^{*}c_{d\sigma}+c_{d\sigma}{}^{*}c_{k\sigma}), \quad (1)$$

where $c_{\mathbf{k}\sigma}{}^{*}$ and $c_{\mathbf{k}\sigma}$ are the creation and annihilation operators for the conduction electrons with momentum $\mathbf{k}$ and spin σ. $c_{d\sigma}{}^{*}$ and $c_{d\sigma}$ are the same for the local *d* electrons with spin σ. Hence the first term on the right-hand side represents the total energy of the conduction band whose width we will take as $2D$, the second term is the energy of the *d* orbital, the third term gives the Coulomb repulsion between spin up and spin down electrons in the *d* level, and the last term represents the energy due to mixing between the conduction electron states and the *d* orbital.

As in Refs. 21 and 22, the trial free energy of our system can be written as

$$\begin{aligned}\mathfrak{F}&=\mathrm{tr}\rho_t(H+k_BT\ln\rho_t)\\&=\sum_j \mathrm{tr}_j h^{(1)}(j)\rho_t{}^{(1)}(j)+\sum_{j>k}\mathrm{tr}_{j,k}h^{(2)}(j,k)\rho_t{}^{(2)}(j,k)\\&\quad+k_BT\Big[\sum_j\gamma^{(1)}(j)+\sum_{j>k}\gamma^{(2)}(j,k)+\cdots\Big], \quad (2)\end{aligned}$$

where $h^{(1)}(j)$ is that part of the Hamiltonian which involves one particle operators in state j and $h^{(2)}(i,j)$ is the part which involves two particles representing the interaction between particles in states i and j. $\rho^{(1)}(j)$ and $\rho^{(2)}(j,k)$ are the one- and two-particle reduced density matrices. The entropy term of the free energy has been expanded in terms of the many-particle cumulants $\gamma^{(1)}(j)$, $\gamma^{(2)}(j,k)$, etc., which are given by

$$\begin{aligned}\gamma^{(1)}(j)&=\mathrm{tr}_j\rho_t{}^{(1)}(j)\ln\rho_t{}^{(1)}(j),\\ \gamma^{(2)}(j,k)&=\mathrm{tr}_{j,k}\rho_t{}^{(2)}(j,k)\ln\rho_t{}^{(2)}(j,k)\\&\quad-\mathrm{tr}_j\rho_t{}^{(1)}(j)\ln\rho_t{}^{(1)}(j)-\mathrm{tr}_k\rho_t{}^{(1)}(k)\ln\rho_t{}^{(1)}(k). \quad (3)\end{aligned}$$

These quantities show the effects of the many-particle correlations. If we truncate the series [Eq. (2)] after $\gamma^{(n)}(1,2\cdots n)$ we say that we are considering up to n-particle clusters.

Since we have up to two-particle-interaction terms in our Hamiltonian, it is reasonable that we retain up to $\gamma^{(2)}(j,k)$ terms in the free energy and neglect three-

---

[13] J. R. Schrieffer and P. A. Wolf, Phys. Rev. **149**, 491 (1966).

[14] L. Dworin, Phys. Rev. **164**, 818 (1967).

[15] G. Horwitz, S. Alexander, and M. Fibich, Phys. Rev. **168**, 495 (1968); M. Fibich and G. Horwitz, *ibid.* **168**, 508 (1968).

[16] P. W. Anderson, Phys. Rev. **124**, 41 (1961).

[17] J. A. Appelbaum, Phys. Rev. **165**, 632 (1968).

[18] P. W. Anderson, Phys. Rev. **164**, 352 (1967).

[19] R. Kikuchi, Phys. Rev. **81**, 988 (1951).

[20] T. Morita, J. Phys. Soc. Japan **12**, 1060 (1957).

[21] T. Morita and T. Tanaka, Phys. Rev. **145**, 288 (1966).

[22] J. Halow, T. Tanaka, and T. Morita, Phys. Rev. **175**, 680 (1968). See also J. Halow, Ph.D. thesis, The Catholic University of America, 1968 (unpublished).

and other higher-particle correlations. The variational free energy is then written as

$$\mathfrak{F}=\sum_{\mathbf{k},\sigma}\epsilon_{\mathbf{k}}\,\mathrm{tr}_{\mathbf{k}\sigma}n_{\mathbf{k}\sigma}\rho_t^{(1)}(\mathbf{k},\sigma)+\sum_{\sigma}\epsilon_d\,\mathrm{tr}_{\sigma}n_{d\sigma}\rho_t^{(1)}(d,\sigma)+U\,[\mathrm{tr}_{\sigma}\rho^{(1)}(d,\sigma)n_{d\sigma}]\,\mathrm{tr}_{-\sigma}\rho_t^{(1)}(d,-\sigma)n_{d-\sigma}[+V\sum_{\mathbf{k},\sigma}\mathrm{tr}_{\mathbf{k}\sigma,d\sigma}\rho_t^{(2)}(\mathbf{k}\sigma,d\sigma)$$
$$\times(c_{\mathbf{k}\sigma}{}^*c_{d\sigma}+c_{d\sigma}{}^*c_{\mathbf{k}\sigma})+k_BT\{\sum_{\mathbf{k},\sigma}\mathrm{tr}_{\mathbf{k}\sigma}\rho_t^{(1)}(\mathbf{k},\sigma)\,\ln\rho_t^{(1)}(\mathbf{k},\sigma)+\sum_{\sigma}\mathrm{tr}_{\sigma}\rho_t^{(1)}(d,\sigma)\,\ln\rho_t^{(1)}(d,\sigma)+\sum_{\mathbf{k},\sigma}[\mathrm{tr}_{\mathbf{k}\sigma,d\sigma}\rho_t^{(2)}(\mathbf{k}\sigma,d\sigma)$$
$$\times\ln\rho_t^{(2)}(\mathbf{k}\sigma,d\sigma)-\mathrm{tr}_{\mathbf{k}\sigma}\rho_t^{(1)}(\mathbf{k}\sigma)\,\ln\rho_t^{(1)}(\mathbf{k}\sigma)-\mathrm{tr}_{\sigma}\rho_t^{(1)}(d\sigma)\,\ln\rho_t^{(1)}(d\sigma)]\}].$$

We then write the reduced density matrices in terms of unknown expectation values so that the variation with respect to these elements can be carried out explicitly. As shown in Ref. 22, these can be constructed very simply. The relevant density matrices are given below:

$$\rho_t^{(1)}(\mathbf{k},\sigma)=\begin{pmatrix}\langle n_{\mathbf{k}\sigma}\rangle & 0\\ 0 & 1-\langle n_{\mathbf{k}\sigma}\rangle\end{pmatrix},\qquad \rho_t^{(1)}(d,\sigma)=\begin{pmatrix}\langle n_{d\sigma}\rangle & 0\\ 0 & 1-\langle n_{d\sigma}\rangle\end{pmatrix},$$

$$\rho_t^{(2)}(\mathbf{k}\sigma,d\sigma)=\begin{pmatrix}\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle & 0 & 0 & 0\\ 0 & \langle n_{\mathbf{k}\sigma}\rangle-\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle & \langle c_{\mathbf{k}\sigma}{}^*c_{d\sigma}\rangle & \\ 0 & \langle c_{d\sigma}{}^*c_{\mathbf{k}\sigma}\rangle & \langle n_{d\sigma}\rangle-\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle & 0\\ 0 & 0 & 0 & 1-\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle+\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle\end{pmatrix}.$$

Before we can perform the variation, we must carry out the trace operations in $\mathfrak{F}$. In order to do this we shall have to express the density matrices in a diagonal form. All one-particle density matrices are already diagonal. Upon diagonalization, $\rho_t^{(2)}(\mathbf{k}\sigma,d\sigma)$ assumes the following form:

$$\rho_t^{(2)}(\mathbf{k}\sigma,d\sigma)$$
$$=\begin{pmatrix}\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle & 0 & 0 & 0\\ 0 & \tfrac{1}{2}\{A+B+[(A-B)^2+4\,|\,\xi_{\mathbf{k}d}{}^{\sigma}\,|^2]^{1/2}\} & 0 & 0\\ 0 & 0 & \tfrac{1}{2}\{A+B-[(A-B)^2+4\,|\,\xi_{\mathbf{k}d}{}^{\sigma}\,|^2]^{1/2}\} & 0\\ 0 & 0 & 0 & 1-\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle+\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle\end{pmatrix},$$

where $A=\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle$, $B=\langle n_{d\sigma}\rangle-\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle$, and $\xi_{\mathbf{k}d}{}^{\sigma}=\langle c_{\mathbf{k}\sigma}{}^*c_{d\sigma}\rangle$.

We can then write the trial free energy as

$$\mathfrak{F}=U^{(1)}+U^{(2)}-TS^{(1)}-TS^{(2)},\tag{5a}$$

where

$$U^{(1)}=\sum_{\mathbf{k},\sigma}\epsilon_{\mathbf{k}}\langle n_{\mathbf{k}\sigma}\rangle+\sum_{\sigma}\epsilon_d\langle n_{d\sigma}\rangle,\tag{5b}$$

$$U^{(2)}=U\langle n_{d\sigma}\rangle\langle n_{d-\sigma}\rangle+V\sum_{\mathbf{k},\sigma}(\xi_{\mathbf{k}d}{}^{\sigma}+\xi_{\mathbf{k}d}{}^{\sigma*}),\tag{5c}$$

$$S^{(1)}=-k_B\sum_{\mathbf{k},\sigma}[\langle n_{\mathbf{k}\sigma}\rangle\ln\langle n_{\mathbf{k}\sigma}\rangle+(1-\langle n_{\mathbf{k}\sigma}\rangle)\ln(1-\langle n_{\mathbf{k}\sigma}\rangle)+\langle n_{d\sigma}\rangle\ln\langle n_{d\sigma}\rangle+(1-\langle n_{d\sigma}\rangle)\ln(1-\langle n_{d\sigma}\rangle)],\tag{5d}$$

$$S^{(2)}=-k_B\sum_{\mathbf{k},\sigma}\{\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle\ln\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle+\tfrac{1}{2}[\langle n_{\mathbf{k}\sigma}\rangle+\langle n_{d\sigma}\rangle-2\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle+((\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathbf{k}d}{}^{\sigma}\,|^2)^{1/2}]$$
$$\times\ln\tfrac{1}{2}[\langle n_{\mathbf{k}\sigma}\rangle+\langle n_{d\sigma}\rangle-2\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle+((\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathbf{k}d}{}^{\sigma}\,|^2)^{1/2}]+\tfrac{1}{2}[\langle n_{\mathbf{k}\sigma}\rangle+\langle n_{d\sigma}\rangle-2\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle$$
$$-((\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathbf{k}d}{}^{\sigma}\,|^2)^{1/2}]\ln\tfrac{1}{2}[\langle n_{\mathbf{k}\sigma}\rangle+\langle n_{d\sigma}\rangle-2\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle-((\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathbf{k}d}{}^{\sigma}\,|^2)^{1/2}]$$
$$+[1-\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle+\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle]\ln[1-\langle n_{\mathbf{k}\sigma}\rangle-\langle n_{d\sigma}\rangle+\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle]-\langle n_{\mathbf{k}\sigma}\rangle\ln\langle n_{\mathbf{k}\sigma}\rangle-(1-\langle n_{\mathbf{k}\sigma}\rangle)\ln(1-\langle n_{\mathbf{k}\sigma}\rangle)$$
$$-\langle n_{d\sigma}\rangle\ln\langle n_{d\sigma}\rangle-(1-\langle n_{d\sigma}\rangle)\ln(1-\langle n_{d\sigma}\rangle)\},\tag{5e}$$

where we have introduced the following decouplings: $\langle n_{d\uparrow}n_{d\downarrow}\rangle=\langle n_{d\uparrow}\rangle\langle n_{d\downarrow}\rangle$ and $\langle n_{\mathbf{k}\sigma}n_{d\sigma}\rangle=\langle n_{\mathbf{k}\sigma}\rangle\langle n_{d\sigma}\rangle$, assuming that the correlations between d electrons of opposite spin, and between conduction and d electrons are negligibly small. Introduction of this kind of decoupling in the problem does not imply that we are carrying out the calculation in Hartree-Fock approximation. This is because the decoupling is used only in quantities that have nonvanishing contribution in the Hartree-Fock approximation. When a quantity is a correlation itself, e.g., $\xi_{\mathrm{k}d}{}^{\sigma}=\langle c_{\mathrm{k}\sigma}{}^{*}c_{d\sigma}\rangle$, it is retained explicitly.

The expectation values $\xi_{\mathrm{k}d}{}^{\sigma}$, $\langle n_{\mathrm{k}\sigma}\rangle$, and $\langle n_{d\sigma}\rangle$ are to be determined in such a way that the free energy $\mathfrak{F}$ given by Eq. (5) has a stationary value. Let us first carry out the minimization with respect to $\xi_{\mathrm{k}d}{}^{\sigma}$. This leads to the expression

$$V+\frac{\xi_{\mathrm{k}d}{}^{\sigma *}k_BT}{[(\langle n_{\mathrm{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathrm{k}d}{}^{\sigma}\,|^2]^{1/2}}\ln\left\{\frac{\langle n_{\mathrm{k}\sigma}\rangle+\langle n_{d\sigma}\rangle-2\langle n_{\mathrm{k}\sigma}\rangle\langle n_{d\sigma}\rangle+[(\langle n_{\mathrm{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathrm{k}d}{}^{\sigma}\,|^2]^{1/2}}{\langle n_{\mathrm{k}\sigma}\rangle+\langle n_{d\sigma}\rangle-2\langle n_{\mathrm{k}\sigma}\rangle\langle n_{d\sigma}\rangle-[(\langle n_{\mathrm{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathrm{k}d}{}^{\sigma}\,|^2]^{1/2}}\right\}=0. \qquad (6)$$

To calculate $\xi_{\mathrm{k}d}{}^{\sigma}$ in the ground state, we take the limit that the temperature $T\rightarrow 0$. Since the first term V in Eq. (6) is finite, the coefficient of T in the second term has to diverge. Now $\xi_{\mathrm{k}d}{}^{\sigma *}$ has a finite value and also the denominator is finite since it involves the square root of the sum of positive quantities whose maximum value is unity. This implies that the logarithmic term has to diverge. Since only the denominator of the logarithmic term can go to zero, V and $\xi_{\mathrm{k}d}{}^{\sigma *}$ have to be of opposite signs so that Eq. (6) can be satisfied. This amounts to a unique solution for $\xi_{\mathrm{k}d}{}^{\sigma}$ in terms of other unknown quantities.

We have

$$\langle n_{\mathrm{k}\sigma}\rangle+\langle n_{d\sigma}\rangle-2\langle n_{\mathrm{k}\sigma}\rangle\langle n_{d\sigma}\rangle-(\langle n_{\mathrm{k}\sigma}\rangle-\langle n_{d\sigma}\rangle)^2+4\,|\,\xi_{\mathrm{k}d}{}^{\sigma}\,|^2=0,$$

which gives

$$\xi_{\mathrm{k}d}{}^{\sigma}=\pm[\langle n_{\mathrm{k}\sigma}\rangle\langle n_{d\sigma}\rangle(1-\langle n_{\mathrm{k}\sigma}\rangle)(1-\langle n_{d\sigma}\rangle)]^{1/2}. \qquad (7)$$

As mentioned before, the sign of $\xi_{\mathrm{k}d}{}^{\sigma}$ must be chosen in such a way that the product $\bar{V}\xi_{\mathrm{k}d}{}^{\sigma}$ is a negative real number so that Eq. (6) is satisfied.

Introducing this solution in the trial free energy function at $T=0$ we get

$$\mathfrak{F}_{T=0}=\sum_{\mathrm{k},\sigma}\epsilon_{\mathrm{k}}\langle n_{\mathrm{k}\sigma}\rangle+\sum_{\sigma}\epsilon_d\langle n_{d\sigma}\rangle+U\langle n_{d\sigma}\rangle\langle n_{d-\sigma}\rangle$$
$$-2\,|\,V\,|\sum_{\mathrm{k},\sigma}[\langle n_{\mathrm{k}\sigma}\rangle\langle n_{d\sigma}\rangle(1-\langle n_{\mathrm{k}\sigma}\rangle)(1-\langle n_{d\sigma}\rangle)]^{1/2}. \qquad (8)$$

Minimizing $\mathfrak{F}$ with respect to $\langle n_{\mathrm{k}\sigma}\rangle$ and $\langle n_{d\sigma}\rangle$, we are led to the following two coupled equations:

$$\epsilon_{\mathrm{k}}-|\,V\,|\left(\frac{n_{d\sigma}(1-n_{d\sigma})}{n_{\mathrm{k}\sigma}(1-n_{\mathrm{k}\sigma})}\right)^{1/2}(1-2n_{\mathrm{k}\sigma})=0, \qquad (9)$$

and

$$\epsilon_d+Un_{d-\sigma}-|\,V\,|\sum_{\mathrm{k}}\left(\frac{n_{\mathrm{k}\sigma}(1-n_{\mathrm{k}\sigma})}{n_{d\sigma}(1-n_{d\sigma})}\right)^{1/2}(1-2n_{d\sigma})=0, \qquad (10)$$

where we have omitted the $\langle\cdots\rangle$ notations for the expectation values. Let us now introduce the following definitions:

$$\epsilon_{d\sigma}=\epsilon_d+Un_{d-\sigma}, \qquad (11)$$

$$\eta_{d\sigma}=n_{d\sigma}(1-n_{d\sigma}), \qquad (12)$$

$$\eta_{\mathrm{k}\sigma}=n_{\mathrm{k}\sigma}(1-n_{\mathrm{k}\sigma}), \qquad (13)$$

$$\epsilon_{0\sigma}=2\,|\,V\,|\sum_{\mathrm{k}}[n_{\mathrm{k}\sigma}(1-n_{\mathrm{k}\sigma})]^{1/2}. \qquad (14)$$

Equations (9) and (10) then assume the following forms:

$$\epsilon_{\mathrm{k}}-|\,V\,|\,(\eta_{d\sigma}/\eta_{\mathrm{k}\sigma})^{1/2}(1-2n_{\mathrm{k}\sigma})=0, \qquad (15)$$

$$\epsilon_{d\sigma}-(\epsilon_{0\sigma}/2\eta_{d\sigma}{}^{1/2})(1-2n_{d\sigma})=0. \qquad (16)$$

From Eq. (15) we obtain

$$\eta_{\mathrm{k}\sigma}=\eta_{d\sigma}\,|\,V\,|^2/(\epsilon_{\mathrm{k}}{}^2+4\eta_{d\sigma}\,|\,V\,|^2), \qquad (17)$$

which can be solved for $n_{\mathrm{k}\sigma}$ to get

$$n_{\mathrm{k}\sigma}=\tfrac{1}{2}[1-\epsilon_{\mathrm{k}}/(\epsilon_{\mathrm{k}}{}^2+4\eta_{d\sigma}\,|\,V\,|^2)^{1/2}]. \qquad (18)$$

Equation (18) gives the distribution of states modified from the Fermi distribution by the presence of the *s-d* interaction. If Eqs. (17) and (14) are combined, we obtain

$$\epsilon_{0\sigma}=2\,|\,V\,|^2\,\eta_{d\sigma}{}^{1/2}\sum_{\mathrm{k}}[(\epsilon_{\mathrm{k}}{}^2+4\eta_{d\sigma}\,|\,V\,|^2)^{1/2}]^{-1}.$$

Replacing the sum over **k** by an integral over ϵ_{k} and recalling that the width of the conduction band is $2D$ we get

$$\epsilon_{0\sigma}=2N(0)\,|\,V\,|^2\,\eta_{d\sigma}{}^{1/2}\int_{-D}^{D}[d\epsilon_{\mathrm{k}}/(\epsilon_{\mathrm{k}}{}^2+4\eta_{d\sigma}\,|\,V\,|^2)^{1/2}]$$
$$=4N(0)\,|\,V\,|^2\,\eta_{d\sigma}{}^{1/2}\sinh^{-1}(D/2\eta_{d\sigma}{}^{1/2}\,|\,V\,|), \qquad (19)$$

where the density of states has been assumed to be a constant and is given by $N(0)$, the density of the conduction states of one spin per unit energy at the Fermi surface.

Combining Eqs. (16) and (19), we obtain the following equation:

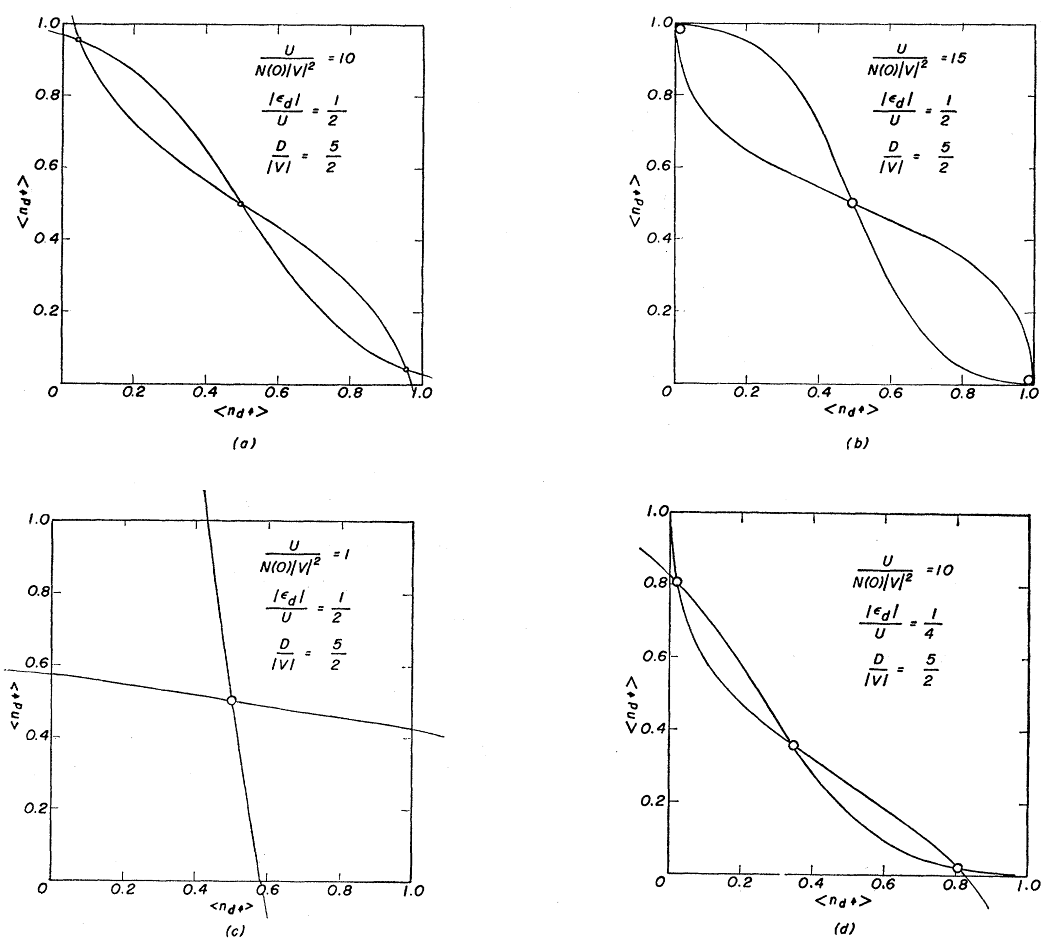
$$\sinh^{-1}(D/2\eta_{d\sigma}{}^{1/2}\,|\,V\,|)=\epsilon_{d\sigma}/2N(0)\,|\,V\,|^2(1-2n_{d\sigma}). \qquad (20)$$

In order to determine the number of d electrons of a given spin, we shall have to solve Eqs. (11) and (20) simultaneously, i.e., the following two equations must be solved self-consistently:

$$\sinh^{-1}\frac{D}{2[n_{d\uparrow}(1-n_{d\uparrow})]^{1/2}\,|\,V\,|}=\frac{\epsilon_d+Un_{d\downarrow}}{2N(0)\,|\,V\,|^2(1-2n_{d\uparrow})},$$

$$\sinh^{-1}\frac{D}{2[n_{d\downarrow}(1-n_{d\downarrow})]^{1/2}\,|\,V\,|}=\frac{\epsilon_d+Un_{d\uparrow}}{2N(0)\,|\,V\,|^2(1-2n_{d\downarrow})}. \qquad (21)$$

To show the meaning of these two equations, we plot for five different cases in Fig. 1. In Figs. 1(a)–1(c), we have chosen $U=2\,|\,\epsilon_d\,|$, and in Figs. 1(d) and 1(e) $U=4\,|\,\epsilon_d\,|$. From Figs. 1(a), 1(b), 1(d), and 1(e) it



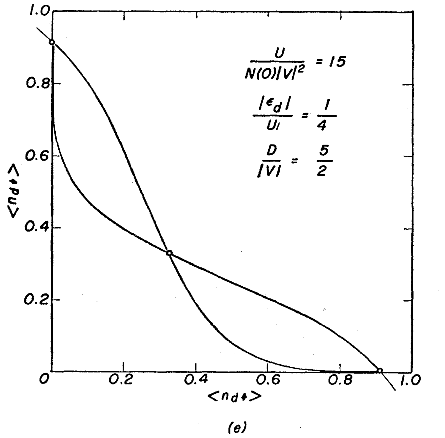


FIG. 1. In (a)–(e) we have plotted $\langle n_{d\uparrow}\rangle$ versus $\langle n_{d\downarrow}\rangle$ for different values of $U/N(0)\,|V|^2$, $|\epsilon_d|/U$ and $D/|V|$. Three solutions are possible when $U/N(0)\,|V|^2\gg1$ [(a), (b), (d), and (e)]. There is only one nonmagnetic solution when $U/N(0)\,|V|^2\sim1$ [(c)].

is clear that when $U/N(0)\,|V|^2$ is rather large there are three sets of solutions for these equations. The first one is at $n_{d\uparrow}=n_1$ and $n_{d\downarrow}=n_2$, the second one is at $n_{d\uparrow}=n_{d\downarrow}$ and the third one is at $n_{d\uparrow}=n_2$ and $n_{d\downarrow}=n_1$. $n_{d\uparrow}=n_{d\downarrow}$ solution is unstable. The other two solutions are the stable ones energetically and they correspond to the up and down configurations of a localized spin. However, when $U/N(0)\,|V|^2$ is rather small, as in Fig. 1(c), the set of equations have only one self-consistent solution $n_{d\uparrow}=n_{d\downarrow}$. In this case the net spin of the system is zero and hence this solution is magnetically uninteresting. Since $U=2\,|\epsilon_d|$ is the most favorable case for localized moment, we find that the condition $n_{d\uparrow}=1-n_{d\downarrow}$ is satisfied for the two magnetic solutions in Figs. 1(a) and 1(b). It is interesting to observe that although the self-consistent equations for $n_{d\uparrow}$ and $n_{d\downarrow}$ derived in this paper look quite different from Anderson equations [Eq. (27) of Ref. 16], they have similar behavior and similar self-consistent solutions.

III. GROUND-STATE ENERGY

We can now proceed to the evaluation of the ground-state energy. The energy lowering of the system due to the presence of a magnetic impurity is given by

$$W=\sum_{\mathbf{k}>k_F,\sigma}\epsilon_{\mathbf{k}}n_{\mathbf{k}\sigma}+\sum_{\mathbf{k}<k_F,\sigma}|\epsilon_{\mathbf{k}}|(1-n_{\mathbf{k}\sigma})+\sum_{\sigma}\epsilon_d n_{d\sigma}+Un_{d\sigma}n_{d-\sigma}-\sum_{\sigma}\epsilon_{0\sigma}\eta_{d\sigma}{}^{1/2}-\epsilon_d. \tag{22}$$

Using Eqs. (18) and (14) in Eq. (22) we get

$$W=\sum_{\mathbf{k}>k_F,\sigma}\epsilon_{\mathbf{k}}\frac{1}{2}\left(1-\frac{\epsilon_{\mathbf{k}}}{(\epsilon_{\mathbf{k}}+4\eta_{d\sigma}|V|^2)^{1/2}}\right)+\sum_{\mathbf{k}<k_F,\sigma}|\epsilon_{\mathbf{k}}|\frac{1}{2}\left(1+\frac{\epsilon_k}{(\epsilon_k+4\eta_{d\sigma}|V|^2)^{1/2}}\right)$$
$$+\epsilon_d(n_{d\uparrow}+n_{d\downarrow}-1)+Un_{d\uparrow}n_{d\downarrow}-[\epsilon_{0\uparrow}\eta_{d\uparrow}{}^{1/2}+\epsilon_{0\downarrow}\eta_{d\downarrow}{}^{1/2}].$$

Converting the sum over **k** to an integral over $\epsilon_{\mathbf{k}}$ and carrying out the spin sum we have

$$W=N(0)\int_0^D\epsilon_{\mathbf{k}}\left(1-\frac{\epsilon_{\mathbf{k}}}{(\epsilon_{\mathbf{k}}+4\eta_{d\uparrow}|V|^2)^{1/2}}\right)d\epsilon_{\mathbf{k}}+N(0)\int_0^D\epsilon_{\mathbf{k}}\left(1-\frac{\epsilon_{\mathbf{k}}}{(\epsilon_{\mathbf{k}}+4\eta_{d\downarrow}|V|^2)^{1/2}}\right)d\epsilon_{\mathbf{k}}$$
$$+\epsilon_d(n_{d\uparrow}+n_{d\downarrow}-1)+Un_{d\uparrow}n_{d\downarrow}-[\epsilon_{0\uparrow}\eta_{d\uparrow}{}^{1/2}+\epsilon_{0\downarrow}\eta_{d\downarrow}{}^{1/2}]$$
$$=N(0)D^2-N(0)\int_0^D\frac{\epsilon^2d\epsilon}{(\epsilon^2+4\eta_{d\uparrow}|V|^2)^{1/2}}-N(0)\int_0^D\frac{\epsilon^2d\epsilon}{(\epsilon^2+4\eta_{d\downarrow}|V|^2)^{1/2}}$$
$$+\epsilon_d(n_{d\uparrow}+n_{d\downarrow}-1)+Un_{d\uparrow}n_{d\downarrow}-[\epsilon_{0\uparrow}\eta_{d\uparrow}{}^{1/2}+\epsilon_{0\downarrow}\eta_{d\downarrow}{}^{1/2}].$$

Carrying out these straightforward integrals and using Eq. (19) for $\epsilon_{0\uparrow}$ and $\epsilon_{0\downarrow}$, we get

$$W=-N(0)|V|^2[\eta_{d\uparrow}[1+2\sinh^{-1}(D/2\eta_{d\uparrow}{}^{1/2}|V|)]]+\eta_{d\downarrow}[1+2\sinh^{-1}(D/2\eta_{d\downarrow}{}^{1/2}|V|)]$$
$$+\epsilon_d(n_{d\uparrow}+n_{d\downarrow}-1)+Un_{d\uparrow}n_{d\downarrow}. \tag{23}$$

We shall treat the ground-state energy in two limiting cases. First let us consider the Appelbaum approximation where $U/N(0)|V|^2$ tends to infinity. In such an approximation the solutions of the self-consistent Eqs. (21) are obtained as $n_{d\uparrow}\approx1$ and $n_{d\downarrow}\approx0$ (cf. Fig. 5 of Ref. 16). Hence we shall have $\eta_{d\uparrow}\approx\eta_{d\downarrow}\approx\eta_d$ and $n_{d\uparrow}+n_{d\downarrow}=1$. Then from Eqs. (21) we obtain

$$U=4N(0)|V|^2\sinh^{-1}(D/2\eta_d{}^{1/2}|V|). \tag{24}$$

Equation (23) then reduces to

$$W=-2N(0)|V|^2\eta_d. \tag{25}$$

Again substituting Eq. (24) into Eq. (21), we have

$$\epsilon_d+2N(0)|V|^2\sinh^{-1}(D/2\eta_d{}^{1/2}|V|)=0,$$

from which we obtain in the weak-coupling limit $(|\epsilon_d|\gg N(0)|V|^2)$

$$\eta_d|V|^2=D^2\exp[-|\epsilon_d|/N(0)|V|^2]. \tag{26}$$

Substituting Eq. (26) into Eq. (25), we have

$$W=-2N(0)|V|^2\exp[-|\epsilon_d|/N(0)|V|^2]. \tag{27}$$

Let us now consider the opposite limit of $U=2|\epsilon_d|$, which is the most favorable case for a localized moment. In this approximation the solutions of the self-consistent equations (21) are such that the condition $n_{d\uparrow}+n_{d\downarrow}=1$ is always satisfied. Hence we have $\eta_{d\uparrow}=\eta_{d\downarrow}=\eta_d$. Using these relations in Eqs. (21) we obtain

$$U=4N(0)|V|^2\sinh^{-1}(D/2\eta_d{}^{1/2}|V|),$$

which is the same as Eq. (24). Hence in this approximation also we obtain

$$W=-2N(0)|V|^2\eta_d.$$

Again, as before, for weak-coupling limit

$$\eta_d|V|^2=D^2\exp[-|\epsilon_d|/N(0)|V|^2],$$

and hence

$$W=-2N(0)D^2\exp[-|\epsilon_d|/N(0)|V|^2], \tag{28}$$

which is the same as Eq. (27). W is the energy difference between a system with a magnetic impurity and the unperturbed system at absolute zero. Thus, Eqs. (27) and (28) show that the ground-state energy of the perturbed state is lower than the ground-state energy of the unperturbed conduction electrons. The expression for the ground-state energy has the same exponential behavior as that obtained by Appelbaum, although he has an extra factor of 2 in the denominator of his exponent. Anderson also obtained a similar energy lowering but his numerical factor in the exponent is quite different from Appelbaum's and our result.