Ground State of a Magnetic Impurity in a Metal

PHILIP W. ANDERSON
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 21 July 1967)

A new type of ground-state wave function for a magnetic impurity in a metal is postulated, and its parameters are determined variationally. The basic idea is that formation of a singlet ground state requires the phase shifts of all electrons near the Fermi surface to be adjusted simultaneously. The method resembles a configuration-interaction treatment of in-out correlation. The result is a phase shift which varies as $(E-E_F)^{1/2}$ near the Fermi surface and a binding energy larger than in previous variational treatments.

THERE has been much recent interest in the behavior of magnetic impurities in metals below the Kondo-Suhl temperature.[1–7] The prevailing opinion at present is that the ground state is in some sense a singlet,[3,8–13] which forms near T_K with binding energy $\sim k_B T_K$. If the ground state is to be a singlet, it seems reasonable that it should be treated by methods and with a model which is not incompatible with the "ordinary" nonmagnetic impurity. For instance, such an impurity obeys the Friedel sum rule,[14] so that the $T=0$ phase shift at the Fermi surface is determined by electrical neutrality, which is difficult within the usual "Kondo model." Such a method and its results are the subject of this paper.

The strong Coulomb interaction, which, in the magnetic impurity, causes magnetism, can in a singlet only induce the "in-out" correlation effect. In-out correlation in nonmagnetic impurities should be dealt with as in atomic systems, by some combination of unrestricted Hartree-Fock (HF) and configuration interaction. The corresponding methods in the magnetic impurity system, however, run into a paradox which is both the difficulty of and the key to the problem. This is that, using the result of a recent paper,[15] the lowest unrestricted HF solutions have no matrix elements to any singlet solution or to each other, because they have phase shifts at the Fermi surface for opposite spins which are greatly different. A singlet state in any real sense—i.e., having no net spin within a finite sphere around the impurity—must have equal phase shifts for the two spins, by the Friedel theorem.

Our technique will be to construct a solution from antisymmetrized products of single-particle scattered wavefunctions, as in HF, but to modify each such determinant so that it can interact with the other configurations chosen: Essentially by fiat, we change the phase shifts in a small region near the Fermi surface so as to make them equal and allow finite matrix elements. The form of the modification is chosen variationally, by minimizing the over-all energy.

As in the usual HF approach,[16] we work with the following Hamiltonian as a model which contains most of the relevant physics:

$$\mathcal{H}=\sum_{k,\sigma}\epsilon_k n_{k\sigma}-E_d\sum_{\sigma}n_{d\sigma}+Un_{d\uparrow}n_{d\downarrow}+\sum_{k,\sigma}V_{dk}(C_{d\sigma}^{\dagger}C_{k\sigma}+C_{k\sigma}^{\dagger}C_{d\sigma}). \quad (1)$$

For the time being let us assume

$$U\gg E_d\gg\rho(0)\langle\tilde{V}^2\rangle_{\mathrm{Av}}\equiv\Delta, \quad (2)$$

where $\rho(0)$ is the density of states at the Fermi surface, which is taken as the zero of energy. V_{dk} becomes small for $|\epsilon_k|\rangle$ a cutoff D, which is taken of order $\sim E_d$ for simplicity.

Our variational assumption about the wave function is the following:

$$\Phi=\alpha(\Phi^{\uparrow}+\Phi^{\downarrow})+\beta\Phi^{0}. \quad (3)$$

Here $\Phi^{\uparrow}$, $\Phi^{\downarrow}$, and Φ^0 are determinantal wave functions made up from two sets of scattered wave functions, $\Psi_{k\sigma}^{+}(r)$ and $\Psi_{k\sigma}^{-}(r)$, and we define Fermion operators $C_{k\sigma}^{+}$ and $C_{k\sigma}^{-}$ annihilating the two types of waves. At a considerable distance from the Fermi surface, Ψ_k^{+} is the same as the scattered wave function of energy ϵ_k with U in (1) set equal to zero and the resonance at E_d—i.e., the up-spin HF solution—and Ψ_k^{-} is the corresponding down-spin solution, with the resonance at E_d+U. These wave functions, of course, are also relatively shifted by the well-known energy shift $\propto\delta/R$.[17]

[1] J. Kondo, Progr. Theoret. Phys. (Kyoto) **32**, 37 (1964).
[2] H. Suhl, Phys. Rev. **138**, A515 (1965).
[3] Y. Nagaoka, Phys. Rev. **138**, A1112 (1965).
[4] A. A. Abrikosov, Physics **2**, 5 (1965).
[5] S. D. Silverstein and C. B. Duke, Phys. Rev. Letters **18**, 695 (1967).
[6] H. Suhl, Phys. Rev. Letters **18**, 743 (1967) (an excellent summary of the present situation).
[7] D. R. Hamann and P. M. Bloomfield, Phys. Rev. (to be published).
[8] J. R. Schrieffer and D. C. Mattis, Phys. Rev. **140**, A1412 (1965).
[9] J. Kondo, Progr. Theoret. Phys. (Kyoto) **34**, 204 (1965).
[10] K. Yosida, Phys. Rev. **147**, 223 (1966); Progr. Theoret. Phys. (Kyoto) **36**, 875 (1966); A. Okiji, *ibid.* **36**, 712 (1966).
[11] A. A. Abrikosov, Physics (to be published).
[12] A. J. Heeger and M. A. Jensen, Phys. Rev. Letters **18**, 488 (1967).
[13] J. Kondo, Progr. Theoret. Phys. (Kyoto) **36**, 429 (1966).
[14] J. S. Langer and V. Ambegaokar, Phys. Rev. **121**, 1091 (1961).
[15] P. W. Anderson, Phys. Rev. Letters **18**, 1049 (1967).
[16] P. W. Anderson, Phys. Rev. **124**, 41 (1961).
[17] We will use repeatedly the essentially complete equivalence of the phase-shift scheme for doing the magnetic-impurity problem [A. Blandin and J. Friedel, J. Phys. Radium **19**, 573 (1958)] and the local orbital scheme of Ref. 16, which equivalence is proved in P. W. Anderson and W. L. McMillan, in *Proceedings of the International School of Physics "Enrico Fermi" Course XXXVII* (Academic Press Inc., New York, 1967) p. 64ff.

Near the Fermi surface, both Ψ's are modified, in a way which may be characterized by the appropriate scattering phase shift as a function of energy $\delta^{+,-}(E)$, which is to be the variational parameter of our method. That this is the case may be seen as follows[18]: By definition of the scattering matrix T and using the unscattered wave functions ϕ_k,

$$T\phi = V\Psi,$$

$$\Psi_k - (E_k - \mathcal{H}_0)^{-1} T\phi_k = (E_k - \mathcal{H}_0)^{-1}(E_k - \mathcal{H}_0 - V)\Psi_k = 0,$$

$$(\phi_{k'}, \Psi_k) = (\epsilon_{k'} - E_k)^{-1} T_{k'k}. \tag{4}$$

E_k is the energy of the scattered wave function Ψ_k. Near the energy shell, this reduces to

$$(\phi_{k'}, \Psi_k) = [\pi\rho(\epsilon)]^{-1} \frac{\sin\delta_k}{\epsilon_{k'} - E_k} = \frac{\sin\delta_k}{\pi(n_k - n_{k'}) + \delta_k}. \tag{5}$$

We have defined $k = \pi n_k / R$, where R is the box radius here, and we observe that the phase factor is automatically eliminated by using real "box" eigenfunctions. As demonstrated in Ref. 15 and 18, the behavior at the pole as given by (5) makes the normalization correct.

It is clear that we may make (5)—which is an expansion of the Ψ in terms of a complete set of states—a definition of the Ψ by fiat. We believe that this is actually an accurate assumption because off-energy-shell deviations would be expected to be on the scale of E_d, since the electrons by our choice of Hamiltonian interact with nothing outside the impurity: Asymptotic wave functions should hold up to the central atom. As in Ref. 15 and Ref. 18, then, the overlap integral of two spherical scattered wave functions $\Psi_k{}^\alpha$ and $\Psi_k{}^\beta$, $(\alpha, \beta = +, -)$ may be taken to be

$$(\Psi_k{}^\alpha, \Psi_{k'}{}^\beta) = \sin(\delta_k{}^\alpha - \delta_{k'}{}^\beta) / [\pi(n_k - n_{k'}) + \delta_k{}^\alpha - \delta_{k'}{}^\beta]. \tag{6}$$

We define the three determinantal components of our wave function as follows:

$$\Phi\uparrow = \prod_k^{k_F} (C_{k\uparrow}{}^+)^\dagger (C_{k\downarrow}{}^-)^\dagger \Phi_{\text{vac}},$$

$$\Phi\downarrow = \prod_k^{k_F} (C_{k\uparrow}{}^-)^\dagger (C_{k\downarrow}{}^+)^\dagger \Phi_{\text{vac}},$$

$$\Phi_0 = \prod_k^{k_F} (C_{k\uparrow}{}^-)^\dagger (C_{\uparrow k}{}^-)^\dagger \Phi_{\text{vac}}. \tag{7}$$

The variations of the phase shifts $\delta_k{}^{+,-}$ and the phase shifts for the wave functions (7) are sketched in Fig. 1.

The energy of our variational function contains two δ-dependent parts which we calculate as follows. The first and simpler is the extra kinetic energy of any of these determinants caused by the anomaly of δ at the Fermi surface. We use (4):

$$\Delta E_{\text{kin}} = \sum_{k<k_F, k'>k_F} |(\Psi_k, \phi_{k'})|^2 (\epsilon_{k'} - \epsilon)_k$$

and with (5)

$$\Delta E = \frac{2}{\pi^2\rho^2} \sum_{k<k_F} \sum_{k'>k_F} \frac{\sin^2\delta_k}{\epsilon_{k'} - \epsilon_k} = \frac{2}{\pi^2} \int d\epsilon_k \sin^2\delta_k \ln(E_d/\epsilon_k) \tag{8}$$

(introducing the cutoff at E_d previously mentioned, and assuming the $\sin^2\delta$ functions for both spins identical). For present purposes we assume the local state resonance structure near E_d to be basically unmodified, and we include only the corrections to δ near k_F in the integral (8).

The second contribution to the energy is the off-diagonal matrix element connecting the two types of singlet wave functions in (3). In the limit U large, at least, these are good approximate eigenstates of all but the V_{dk} terms of the Hamiltonian (1), but the states have been so constructed that the V_{dk} terms connect them with finite matrix elements. This construction is the crucial feature of our wave functions: The theorem of Ref. 15 demonstrates that energy can be gained from forming a linear combination such as (3)—i.e., singlet formation—*if and only if* the phase shifts for the two spins are forced to be identical at the Fermi surface. A way of putting it is that the V_{dk} terms represent a local operator, and a local operator cannot move electrons back and forth to infinity: It must connect only states with the same numbers of electrons in a finite region containing the impurity.

Consider a typical term of such a matrix element, say,

$$M_k{}^{\uparrow 0} = V_{kd}(\Phi_0 \mid C_{k\uparrow}{}^\dagger C_{d\uparrow} \mid \Phi_\uparrow). \tag{9}$$

The expression on the right is the overlap integral of two determinantal wave functions, Φ_0 and $C_k{}^\dagger C_d \Phi_\uparrow$. It will be clearer how this is evaluated if we first do the simple case

$$S_{0\uparrow} = (\Phi_0, \Phi_\uparrow). \tag{10}$$

The spin-down pieces of Φ_0 and $\Phi_\uparrow$ are identical, so that we are left with the overlap integral[18] (we assume N states of angular momentum l below the Fermi surface)

$$S_{0\uparrow} = \iiint \prod_{k=1}^{N} (dr_k)(N!)^{-1}$$

$$\times \sum_{P,P'} (-1)^{P+P'} \prod_k^{k_F} [\Psi_k{}^+(Pr_k)][\Psi_k{}^-(P'r_k)]$$

$$= \operatorname*{Det}_{k,k'=1\cdots N} |(\Psi_k{}^+, \Psi_{k'}{}^-)|$$

$$= \operatorname*{Det}_{nk,nk'=1\cdots N} |\sin(\delta_k{}^+ - \delta_{k'}{}^-)/\pi(n_k - n_{k'}) + \delta_k{}^+ - \delta_{k'}{}^-|. \tag{11}$$

[18] J. Friedel, Phil. Mag. **43**, 1115 (1952)

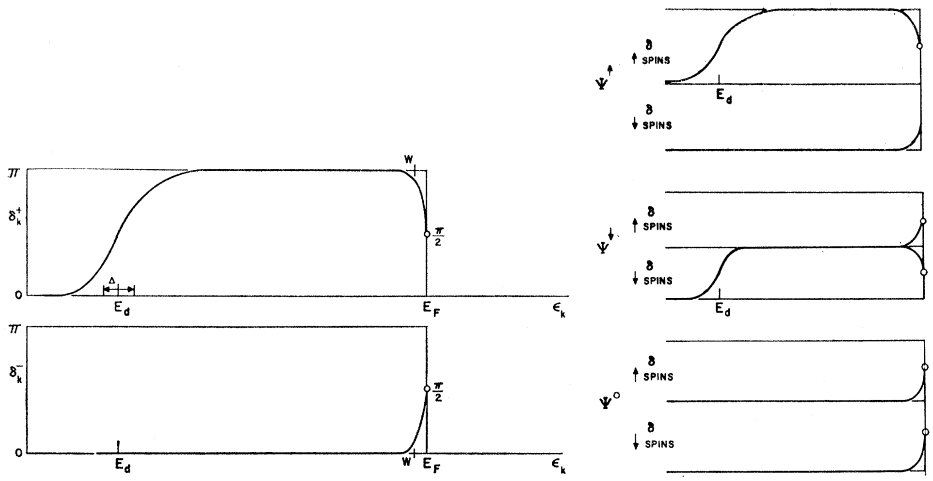

FIG. 1. (a) Schematic dependence of phase shifts δ^+ and δ^- on energy. (b) Phase shifts of states in the three components of the ground-state wave function.

We are assuming $U \gg$ all other energies, so that the Ψ^- do not contain any ϕ_d. Thus the off-energy-shell contributions in Ψ^+ from ϕ_d do not enter. In all cases we want to concentrate our attention on the Fermi-surface contributions, so this kind of approximation is made. We believe that any correction in this would be an irrelevant numerical renormalization of V_{kd}, E_d, etc.

As far as we can see, this determinant cannot in general be evaluated exactly, but a case which can is sufficiently flexible that it probably gives us an excellent indication of the behavior of (11). This is the case in which one of the phase shifts remains constant, so that $\sin(\delta_k{}^+ - \delta_{k'}{}^-)$ depends on k or k' only. Again, we are making an on-energy-shell approximation: We use the numerator at the pole for all values. It has a wider validity: For many ranges of the variables one or the other *is* constant. To make this approximation let us define

$$\delta_n = (\delta_k{}^+ - \delta_k{}^-)/\pi,$$

where $n = n_{kF} - n_k$, and measures the distance from E_F. Then

$$S_{0\uparrow} \simeq \prod_{n=1}^{N} (\sin\pi\delta_n/\pi) \ \mathrm{Det}\,|\, 1/n - n' + \delta_n \,|.$$

This can be evaluated by an algorithm due to Cauchy[19]:

$$\begin{aligned} S_{0\uparrow} &= \prod_n^N \frac{\sin\pi\delta_n}{\pi} \prod_{m<n}^N (n-m+\delta_n-\delta_m)(m-n) \Big/ \prod_{m,n}^N n-m+\delta_n, \\ &= (-1)^N \prod_n^N \frac{\sin\pi\delta_n}{\pi\delta_n} \prod_{m<n}^N \frac{[1+(\delta_n-\delta_m)/(n-m)]}{[1+\delta_n/(n-m)][1-\delta_n/(n-m)]}, \\ &= (-1)^N \exp\left\{ \sum_n^N \ln\left(\frac{\sin\pi\delta_n}{\pi\delta_n}\right) + \sum_{m<n}^N \ln\left(1+\frac{\delta_n-\delta_m}{n-m}\right) - \ln\left(1+\frac{\delta_n}{n-m}\right) - \ln\left(1-\frac{\delta_m}{n-m}\right)\right\}. \end{aligned} \tag{12}$$

It will turn out that in all cases of interest the first few terms in an expansion of the exponent in powers of δ are adequate. The linear term manifestly cancels out. Using the standard product expansion for $(\sin\pi\delta)/\pi\delta$, we obtain for the result taking into account the quadratic terms

$$\begin{aligned} (-1)^N \exp &- \left\{ \sum_n^N \delta_n{}^2 \sum_m \frac{1}{m^2} + \tfrac{1}{2} \sum_{m<n} \left\{ \frac{(\delta_n-\delta_m)^2}{(n-m)^2} - \frac{\delta_n{}^2}{(n-m)^2} - \frac{\delta_m{}^2}{(n-m)^2} \right\} \right. \\ &= (-1)^N \exp\left(-\frac{1}{2} \left\{ \sum_{n=1}^N \left[\sum_{m=n}^\infty \left(\frac{\delta_n}{m}\right)^2 + \sum_{m=N-n+1}^\infty \left(\frac{\delta_n}{m}\right)^2 + \sum_{m=1}^{N-n} \frac{(\delta_{n+m}-\delta_n)^2}{m^2} \right] \right\} \right) \\ &\simeq (-1)^N \exp\left[-\tfrac{1}{2} \sum_{n=1}^N \delta_n{}^2 (n^{-1} + (N-n+1)^{-1}) - \tfrac{1}{2} \sum_{n=1}^N \sum_{m=1}^{N-n} \frac{(\delta_{n+m}-\delta_n)^2}{m^2} \right]. \end{aligned} \tag{13}$$

[19] I am grateful to J. M. Luttinger for suggesting the source of this algorithm: G. Polya and G. Szegö, *Aufgabe und Lehrsatze aus der Analysis* (Dover Publications, Inc., New York, 1945), Vol. 2. p. 98.

Let us note several features of this result. First, if δ is finite as $n\to 0$ (or $\to N$) the overlap is zero as $\exp[-\frac{1}{2}(\delta^2)\ln N]$. This is the exact version of the orthogonality theorem proved as an inequality in Ref. 15. It is interesting that the main difference from the previous result is to replace $\sin^2\delta$ by δ^2. A phase shift of π corresponds to $S\propto N^{-1/2}$, which is precisely the overlap one expects, since then one state contains precisely one localized electron which in the other is free, and a local-free overlap integral is $\sim N^{-1/2}$. We should note that when δ is finite and large it is not valid to retain only the quadratic terms in the expansion of (12), but the change in the result is a numerical factor in the first few factors of the product which can be computed in detail if need be.

For the actual overlap (12) to be finite, δ_n must vanish as $n\to 0$ fast enough to converge the first sum. This gives us a Fermi-surface structure of a width W which we may define by setting

$$\sum_{n=1}^{\infty}\frac{\delta_n^2}{n}=\ln\frac{E_d}{W}. \tag{14}$$

In later sections and the Appendix we handle this structure in more detail, working out as best we can the additive numerical constants in expressions like (14), which is correct to logarithmic accuracy. These constants will be vital to the later evaluation of the energy.

Let us estimate the contribution to the third sum in (13) coming from the region of the resonance at E_d by assigning a width Δ to that resonance as follows:

$$\sum_{m=1}^{N-n}\frac{(\delta_{m+n}-\delta_n)^2}{m^2}\simeq\sum_{m=(E_d+\Delta)\rho-n}^{N-n}(1/m^2)$$
$$=[(E_d+\Delta)\rho-n]^{-1}-(N-n)^{-1}$$

The sum over n of the last term cancels the second sum in (13); the first gives us

$$\sum_{n=1}^{N}\frac{\delta_n^2}{N-n}+\sum_{n=1}^{N}\sum_{m=1}^{N-n}\frac{(\delta_{m+n}-\delta_n)^2}{m^2}=\ln\frac{E_d}{\Delta}. \tag{15}$$

Thus

$$S_{0\uparrow}\simeq(\Delta W)^{1/2}/E_d\ll 1. \tag{16}$$

(The contribution of $n\sim W\rho$ to the second two sums is not logarithmic, though it may, of course, be finite: All calculations are of necessity limited to a kind of logarithmic accuracy.)

The actual evaluation of the desired matrix element (9) requires one more trick of manipulation with these overlap determinants. Acting on $\Phi_\uparrow$ with $C_{d\uparrow}$ is not an operation which is easy to evaluate numerically, but physically it is clear enough what that does: It removes the resonance at E_d and leaves the phase shift at all energies above E_d shifted by π. It is an excellent approximation to consider $\Phi_\uparrow$ as a projected wave function in which the projection operator $1-n_{d\uparrow}=0$, as Φ_0 is projected on $n_d=0$. Then except at the resonance near E_F these may be changed into each other by multiplication with C_d or $C_d^\dagger$.

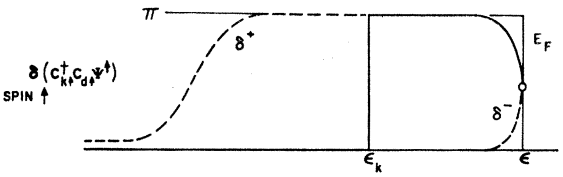

FIG. 2. Phase shifts of the wave function obtained by multiplying by one component V_{kd} of the single-particle potential.

We then wish to compute $(\sum_k C_k^\dagger)C_d\Phi_\uparrow$. Now $C_d\Phi_\uparrow$ is a wave function with one less electron than Φ_0, but because their phases are now shifted by π at the Fermi surface the energy levels match precisely at both $n\gg E_d\rho$ and at $n=0$. We now insert into $C_d\Phi_\uparrow$ a wave function having constant amplitude V_{kd} for each ψ_k in Φ_0, which is to say our overlap determinant now looks like

$$V_{kd}\begin{vmatrix}1 & (\Psi_1^-,\Psi_1^+) & (\Psi_1^-,\Psi_2^+) & \cdots\\ 1 & (\Psi_2^-,\Psi_1^+) & \cdots & \\ 1 & & & \\ 1 & & & \\ 1 & & & \\ 1 & & & \end{vmatrix}. \tag{17}$$

To evaluate this we resort simply to the sum of the minors of the first column. The Mth minor is now an $(N-1)\times(N-1)$ determinant with the Mth wave function Ψ_M^- omitted. The nm element of this Mth minor is (we again assume that only one phase shift varies, for the sake of simplicity only in this case)

$$D_{nm}{}^M=\frac{\sin\pi\delta_n}{\pi[n-m+\delta_n]},$$
$$m=1\cdots N-1,\ n=1,2\cdots M-1,\ M+1\cdots N$$
$$=\frac{\sin(\pi\delta_n{}'^M)}{\pi[n-m+\delta_n{}'^M]}\operatorname{sgn}(M-n),$$
$$n=1\cdots N-1,\ m=1\cdots N-1, \tag{18}$$

with

$$\delta_n{}'^M=\delta_n,\qquad m<M$$
$$\delta_n{}'^M=\delta_n+1,\qquad m\geq M. \tag{19}$$

The effect of the sgn function in (17) is just such as to remove the $(-)^M$ in the minor expansion of (17), so that the net result is that the kth matrix element is just obtained by introducing a saltus of π in the

phase shift at k (very little to our surprise: that implies simply dropping a state at that point). We show in Fig. 2 the phase shifts including the saltus at ϵ_k.

Thus we have

$$\sum_k M_k{}^{\uparrow 0} = \sum_k V_{kd} \exp\left\{\sum_n \ln \frac{\sin\pi\delta_n{}'^M}{(\delta_n{}'^M)} + \sum_{m<n}\left[\ln\left(1+\frac{\delta_n{}'^M-\delta_m{}'^M}{n-m}\right)-\ln\left(1+\frac{\delta_n{}'^M}{n-m}\right) - \ln\left(1-\frac{\delta_m{}'^M}{n-m}\right)\right]\right\}, \quad (20)$$

with $M=\epsilon_k\rho$. Again, keeping terms only to logarithmic accuracy in the exponent, we may limit ourselves to the quadratic part of the exponent, as in (13). The major difference between (20) and (13) is that now, because of the saltus in δ, there remains an actual divergence, which together with the $N^{-1/2}$ dependence of V_{kd} and the sum over k combines to make (20) finite as a whole, while it would not be so if the Fermi-surface structure were omitted.

The most important range for ϵ_k is $E_d \gg \epsilon_k \gg W$. In this range the effect of the saltus is to replace the smooth cutoff at E_d by a sharp one at ϵ_k. Let us compute this element:

$\epsilon_k \gg W$:

$$\ln[M_k{}^{0\uparrow}/V_{kd}] \cong -\tfrac{1}{2}\left(\ln\frac{\epsilon_k}{W} + \sum_{n=1}^{M}\frac{\delta_n{}^2}{N-n} + \sum_{n=W\rho}^{M}\sum_{m=M-n}^{N-n}\frac{1}{m^2}\right)$$

$$\cong -\tfrac{1}{2}\left[\ln\frac{\epsilon_k}{W} + \sum_{n=W\rho}^{M}\frac{1}{N-n} + \sum_{n=W\rho}^{M}\left((M-n)^{-1}-(N-n)^{-1}\right)\right]$$

$$= -\tfrac{1}{2}[\ln\epsilon_k/W + \ln M] \quad (21)$$

(we remember that $M=\epsilon_k\rho$ and ρ is of order N, so that we get $\sim N^{-1/2}$, as stated.) Thus we have, for $\epsilon_k \gg W$:

$$M_k{}^{0\uparrow} \cong e^{1/4}(e/2\pi)^{1/2} W^{1/2} V_{kd}/\rho^{1/2}\epsilon_k. \quad (22)$$

The factor $(2\pi/e)^{-1/2}$ comes from the evaluation of the nonlogarithmic factors at the saltus given in the Appendix. The factor $e^{1/4}$ is the result of an approximate evaluation of the sums involving the Fermi-surface structure.

At this point we have enough information to determine W variationally to logarithmic accuracy. The excitation energy of Φ_0 relative to $(\Phi^{\uparrow}+\Phi^{\downarrow})$ is $|E_d|$ [the kinetic energy (8) is insignificant]. Since also (22) is small on this scale as is $S_{0\uparrow}$ we may take $\alpha=\frac{1}{2}\sqrt{2}$ and

$$\beta = -\sqrt{2}\left(\sum_k M_k{}^{0\uparrow}\right)/|E_d|, \quad (23)$$

and the energy gained by the β admixture is (a factor inside the logarithm is neglected)

$$E_g \cong -2\left(\sum_k M_k{}^{0\uparrow}\right)^2/E_d$$

$$\cong -(e^{3/2}\rho\,|V|^2/\pi E_d) W \ln^2(E_d/W). \quad (24)$$

The energy loss may be calculated from (8), again using the Appendix to get the result approximately to the next order better than logarithmic:

$$E_l = \frac{2}{\pi^2}\int \sin^2\delta \ln\frac{E_d}{\epsilon_k}\, d\epsilon_k$$

$$\simeq 1.215(W/\pi)\ln E_d/W. \quad (25)$$

We now vary with respect to W to minimize the energy. The result is

$$\ln E_d/W \cong 1.215/e^{3/2}(E_d/\rho V^2),$$

$$W \cong E_d \exp(-E_d/4.48\rho V^2) = E_d \exp(-1/4.48\rho J), \quad (26)$$

using the Schrieffer-Wolff exchange integral value. This width is better than the result of most previous variational calculations.[10–12] Our total binding energy is extremely sensitive to details and to higher-order terms since the lowest-order approximation (26) cancels out in the binding energy and the next correction to the ln must be kept. For the record, it is

$$-E_g + E_l = E_b = -1.215(W/\pi). \quad (27)$$

If our estimates are correct, this is larger than that obtained in other variational calculations.[20]

[20] With the numerical factors as we have used them, we can identify V^2/E_d as Kondo's J, and then our variational energy is much lower even than his in Ref. 13, which he believed (but is no longer certain) was already too low with respect to perturbation-theory extrapolations. Excessive weight should not be placed on such comparisons, since all variational theories leave out perturbation-theoretic terms which are algebraic in J: One makes the implicit assumption that perturbation theory starting from the new ground state will not give results differing by an amount greater than $\sim W/D$ because of the Fermi-surface structure relative to the usual perturbation theory, since only that fraction of the electron states are strongly perturbed, while the perturbation-theoretic energy comes from the whole spectrum. It should be noted that Kondo's wave function is rather close in effect to that used here, being much more specialized in its choice of the same set of scattering states for each spin, but allowing for more freedom in the off-energy-shell parts. Nonetheless, the motivation and physical meaning of his calculations are very different from ours.

The cancellation of the biggest terms in E_g has an important physical meaning. This is that the kinetic energy of the Fermi-surface structure (which is of the nature of a bound state) is considerably larger than its binding energy $\sim W$. Thus the uncertainty principle does *not* require that its extension in space be of order $\hbar V_F/W$, but rather of order $\hbar v_F/E_l \sim (\hbar v_F/W)\, E_d/\rho v^2$, a factor $\sim 1/J\rho$ smaller. This makes more plausible the apparently surprising fact that strong interaction effects are not observed in the $\sim 10^{-5}$ concentration range.

Next we attempt to get some indication of the detailed behavior of δ_k for low energies $\epsilon_k \ll W$. Again the basic energy-balance problem is controlled not by the low-energy matrix elements but by the fact that δ for low energies enters critically into the high-energy matrix elements. This enables us to write down a precise equation in the limit $\Delta/E_d \to 0$ (weak coupling: the Kondo limit) for the variation of δ with energy, even though all our other numerical results are approximate even in this limit. That is, there are three ways in which δ_n as a function of energy near E_F affects the multiplicative factors *outside* the logarithms in (24) and (25):

(a) As in (13), the term involving W in (21) is the result of a sum over n:

$$\exp(-\tfrac{1}{2}\sum_n \delta_n^2/n).$$

Given that δ_n is some function

$$\delta_n = f(\epsilon_n/W), \tag{28}$$

we have

$$\sum_n \frac{\delta_n^2}{n} = \int_0^{E_k} f^2\left(\frac{\epsilon_n}{W}\right)\frac{d\epsilon_n}{\epsilon_n} = \ln\frac{E_k}{W} + I_a,$$

$$I_a \cong \int_0^1 \frac{f^2(x)\,dx}{x} + \int_0^\infty [f^2(x)-1]\frac{dx}{x}. \tag{29a}$$

I_a depends on the *form* of f and not on the scale factor W, which may be defined arbitrarily so long as f approaches unity for $\epsilon/W \gg 1$ and 0 for $\epsilon/W \ll 1$. This arbitrariness leads to no difficulties in the variational problem: the scale factor cancels out.

(b) A factor which we ignored in the approximate evaluations (21) and (16) is the factor common to both matrix elements,

$$\exp\left\{-\frac{1}{2}\left[\sum_{n=1}^{N}\sum_{m=1}^{N-n}\frac{(\delta_{n+m}-\delta_n)^2}{m^2}\right]\right\}$$

or at least that portion of it coming from the Fermi-surface structure. This term is vital in the present problem, because (29a) does not prevent $\delta(\epsilon)$ from becoming infinitely steep at some point other than $\epsilon = 0$. This is the term, in fact, which enforces continuity of f. Clearly it is qualitatively proportional to $(d\delta/d\epsilon)^2$ but that is not an adequate approximation to it. The quantity which enters, then, is

$$\exp(-\tfrac{1}{2}I_b),$$

$$I_b \cong \int_0^\infty dx \int_0^\infty \frac{dy}{y^2}[f(x+y)-f(x)]^2. \tag{29b}$$

Clearly if $f \to 0$ as $x \to \infty$, the upper end is highly convergent and may be extended to ∞.

(c) Finally, there is a form-dependent term multiplying the logarithm in (8). Here we have made a special assumption: that in the HF solution both resonances are far from the Fermi surface. Thus $\delta_k^+ + \delta_k^- = \pi$, and

$$\sin^2\delta_k^{\pm} = \cos^2\tfrac{1}{2}\pi\delta(\epsilon_k).$$

This is a specialization for convenience; in general we should have $\delta_k^+ + \delta_k^- = \alpha(k)$, $\delta_k^+ - \delta_k^- = \pi\delta$, where $\alpha(k)$ is a function determined by electrical neutrality at the Fermi surface and by an auxiliary variational procedure elsewhere. Nevertheless, it is easy to show that with given α the energy varies with $\cos^2\frac{1}{2}\pi\delta$. Inserting (28) in (8) we have

$$\Delta E = \frac{2W}{\pi^2}\int_0^\infty dx \cos^2\tfrac{1}{2}\pi f(x)\,(\ln E_d/W + \ln x).$$

Thus we have a third multiplicative constant outside the logarithms, which again in principle is scale-independent.

$$I_c = \int_0^\infty dx \cos^2\tfrac{1}{2}\pi f(x). \tag{29c}$$

Now let us write down the total energy including the form-dependent factors (29);

$$\Delta E = (2/\pi^2) I_c W \ln E_d/K_1 W$$
$$-(\rho\,|\,V\,|^2/\pi E_d) W \ln^2 E_d/K_2 W \exp[-(I_a+I_b)]. \tag{30}$$

K_1 and K_2 are form-dependent *additive* terms, which would be vital in getting the energy exactly, but do not determine W in the limit $E_d \gg W$. In that limit clearly the variational problem controlling the *form* of δ is to minimize I_a and I_b, keeping I_c fixed or vice versa (I_a and I_b *increase* with δ, and I_c decreases). This may be done with a Lagrange multiplier,

$$\delta(I_a + I_b + \lambda I_c) = 0, \tag{31}$$

where from (30) λ is essentially of order unity:

$$\lambda \propto \frac{1}{(\rho V^2/E_d)\,\ln E_d/W}. \tag{31'}$$

Performing the variations in (31) we get

$$\frac{f(x)}{x} + \int_0^x \frac{dy}{y^2}[f(x)-f(x-y)]$$
$$-\int_0^x \frac{dy}{y^2}[f(x+y)-f(x)] - \tfrac{1}{4}(\pi\lambda)\sin\pi f(x) = 0. \tag{32}$$

This nonlinear integral equation is probably not soluble in general, but limiting behaviors can be determined. First let us examine the behavior at the Fermi surface ($x\to 0$). Here the last term is linear in f and is completely dominated by the first two, each of which can be singular as $x\to 0$. First, we observe that f must go to 0 more slowly than x, for at $x=0$ the second term reduces to

$$\int_0^\infty f\frac{dy}{y^2},$$

which is infinite if $f\sim x$, and finite if $f\sim x^{1+\epsilon}$; while $f(x)/x$ goes to a finite nonzero limit if $f\sim x$, and to zero if $f\sim x^{1+\epsilon}$. On the other hand, f must go more rapidly to zero than $(\ln x)^{-1/2}$ in order that the integral I_a be convergent. By trial, we discover that

$$f=Ax^{1/2}, \qquad x\ll 1 \tag{33}$$

gives an exact solution for the first two terms of (32):

$$Ax^{-1/2}+Ax^{-1/2}\int_0^1\frac{dz}{z^2}\left[2-(1-z)^{1/2}-(1+z)^{1/2}\right]$$
$$-Ax^{-1/2}\int_1^\infty\frac{dz}{z^2}\left[(1+z)^{1/2}-1\right]=0.$$

The integrals are barely convergent, but elementary.[21] The solution cannot proceed in series starting from this term because the third term has high-end singularities.

To estimate the large-x behavior where f is smooth, the second and third terms may be dropped, and we get

$$1/x\simeq\tfrac{1}{4}(\pi\lambda)\,\sin\pi(1-f)\cong\tfrac{1}{4}(\pi^2\lambda)\,(1-f)$$
$$f\simeq 4/\pi^2\lambda x, \qquad x\gg 1. \tag{34}$$

Again, the peculiar form of the second term prevents straightforward series solution. An approximate form which fits both ends is

$$f\simeq\left(\frac{x}{(2/\pi^2\lambda)+x}\right)^{1/2}. \tag{35}$$

In conclusion, in this paper we have formulated a new type of wave function for an impurity in a metal. The idea of this wave function is relatively simple: that is, the wave function is to be made up out of several configurations. For each configuration a different set of one-electron wave functions is chosen variationally. The entire complexity comes from the presence of the Fermi surface, which seems inevitably to bring an infinite number of these wave functions into the computation, and leaves one with the serious technical problems, first of parameterizing many-body wave functions with an infinite number of degrees of freedom, and second of evaluating matrix elements between wave functions differing for each of an infinite number of particles. The techniques we have found for these problems may not be unique but are usable. The only serious competitor would be the use of a pseudopotential varying from one component of the state to another, but the matrix-element evaluation would remain a problem.

The basic difficulty seems to be expressible in conventional many-body theoretic terms as a divergent vertex correction

$$\sum_{k,k'}\frac{(T_{kk'})^2}{(\epsilon_k-\epsilon_{k'})^2}$$

to the scattering matrix element. The basic structure of our idea for dealing with this is to introduce two or more different "vacuum" states which are defined so that these corrections can be handled, at the cost of there being matrix elements of $\mathcal{H}$ between the different vacua. (This idea will be clearer when we discuss the calculation of the excitation spectrum in a subsequent paper.) Some of the more standard techniques were attempted but did not seem to work; of course, we cannot guarantee that none of them would.

This technique, and our methods for computing matrix elements between disjoint "ground" states, may have much wider applicability. The basic infinite vertex correction seems to appear in a number of optical-transition problems as well,[22] particularly x-ray emission edges and possibly even in ordinary optical transitions in metals. Our methods of matrix-element evaluation are still effective, but unfortunately in the optical case the divergence due to the finite $\delta(E_F)$ actually occurs and the complicated problem of the full spectral distribution of the matrix elements comes to the fore.

I would like to acknowledge specific suggestions of value from J. M. Luttinger, J. J. Hopfield, W. L. McMillan, J. Kondo, and D. R. Hamann and many valuable discussions with these and many others, including W. Brinkman, L. Dworin, J. Appelbaum, J. M. Rowell, T. Geballe, M. P. Sarachik, and A. M. Clogston.

[21] This form has somewhat indirect verification from several sets of experimental data. A preliminary calculation of the single-particle excitation spectrum indicates that $d\delta/dE\sim E^{-1/2}$ is interpretable as a peak in the density of states. This would give $\chi\propto T^{-1/2}$ at low temperatures, which fits well a number of data: T. H. Geballe, B. T. Matthias, A. M. Clogston, H. J. Williams, R. C. Sherwood, and J. P. Maita [J. Appl. Phys. **37**, 1181 (1966)] mention this as a good fit to data on Fe in Ir; and it also gives an excellent fit to the data of M. D. Daybell and W. A. Steyert [Phys. Rev. Letters **18**, 398 (1967)] on Fe in Cu. D. R. Hamann, Ref. 7, arrives at a similar behavior from the Nagaoka theory. If we may identify $\delta(E)$ with $\delta(T)$, the temperature behavior of the Fermi surface phase shift, in some cases the resistance may be $\propto\delta^2\propto T$, which is a rough fit to data on Fe in Rh of B. R. Coles, [Phys. Letters **8**, 243 (1964)] and on Fe in Ir of M. P. Sarachik (to be published).

[22] This aspect of these methods has been evolved in conversations with J. J. Hopfield and D. R. Hamann. There is some relationship to the results of G. D. Mahan, Phys. Rev. (to be published).

APPENDIX. EVALUATION OF THE NONLOGARITHMIC FACTORS IN THE MATRIX ELEMENTS

In order to calculate the correct width of the resonance structure (which should of course agree at least roughly with the Kondo temperature as calculated from perturbation theory), it is essential to get at least the multiplicative factors correct to better than logarithmic accuracy. We find that the additive constants are more complicated (although in principle they could be done) and that even the multiplicative constants can be at best approximately evaluated.

There are several places where such factors enter. The simplest, if trickiest, comes from the power series expansion in δ which we have made in evaluating $M^{\uparrow 0}$. This breaks down at the saltus in δ because there we have factors $1\pm\delta/m$, where δ is not small and m is not large. Consider the matrix element again in the range $\epsilon_k \gg W$.

The wave functions at energies below ϵ_k are identical and may be ignored. From ϵ_k we start out with a phase shift of π [let us imagine actually that it is $\pi(1-\epsilon)$ and go to the limit $\epsilon\to 0$ to keep everything finite at all times]. From π the phase shift decreases continuously toward zero at the Fermi surface; let us continue on above the Fermi surface to some number N, introducing zero shifts so that the upper-limit terms are negligible. Then, using the algorithm and numbering our levels n from the one above ϵ_k, the matrix element is formally

$$\prod_{n=1}^{N}\frac{\sin\pi\delta_n}{\pi\delta_n}\prod_{m=1}^{N-n}\left(1+\frac{\delta_{n+m}-\delta_n}{m}\right)\Big/\prod_{n=1}^{N}\prod_{m<n}\left(1+\frac{\delta_n}{n-m}\right)\left(1+\frac{\delta_m}{n-m}\right). \tag{A1}$$

We use the product expansion

$$\frac{\sin\pi\delta_n}{\pi\delta_n}=\prod_{m=1}^{\infty}\left(1+\frac{\delta_n}{m}\right)\left(1-\frac{\delta_n}{m}\right)$$

and, renumbering the factors in the denominator of (A1), we find that many of them cancel against the corresponding ones in the sine expansion (including all those which go to zero as $\epsilon\to 0$). We are left with the following product:

$$\prod_{n=1}^{N}\prod_{m=1}^{N-n}\left(1+\frac{\delta_{n+m}-\delta_n}{m}\right)\prod_{m=n}^{\infty}\left(1+\frac{\delta_n}{m}\right)\prod_{m=N-n+1}^{\infty}\left(1-\frac{\delta_n}{m}\right).$$

Since δ is continuous, no large terms appear in the first product; and since δ_N and all those near it are zero, the last product is also highly convergent whenever $m\to 1$. We are left with the middle product. Our approximation to this kind of product (which is independent of any particular numbering of levels) is to exponentiate and expand the logarithm. This is very convergent above some value of n and $m\gg 1$, so let us evaluate it only up to K, first by our approximation (setting, now, $\delta_n=1$):

$$\begin{aligned}\prod_{n=1}^{K}\prod_{m=n}^{K}(1+m^{-1})&\simeq\exp\sum_{n=1}^{K}\sum_{m=n}^{K}\ln(1+m^{-1}),\\ &\simeq\exp\sum_{n=1}^{K}\sum_{m=n}^{K}(m^{-1}-1/2m^2),\\ &\simeq\exp\{K-\tfrac{1}{2}\sum_{n=1}^{K}(n^{-1}-K^{-1})\},\\ &\simeq\exp\{K-\tfrac{1}{2}\ln K+\tfrac{1}{2}\}.\end{aligned} \tag{A2}$$

(The sum of $1/n$ can be done more accurately but our approximation has been to neglect the lower limit of integrals in the quadratic terms, so we maintain consistency here).

The same product may be evaluated exactly. It is

$$\begin{aligned}&(1+1/1)(1+\tfrac{1}{2})\cdots(1+1/K)\\ &\quad\times(1+\tfrac{1}{2})(1+\tfrac{1}{3})\cdots(1+1/K)\times\cdots\\ &=\frac{(K+1)^K}{K!}=\frac{(K+1)^{K+1}}{(K+1)!}=\frac{e^K}{K^{1/2}}\times\frac{e}{(2\pi)^{1/2}}.\end{aligned} \tag{A2'}$$

The ratio of these two evaluations is $(e/2\pi)^{1/2}$, which we have inserted in (22).

The other multiplicative factors come from the three integrals we have called I_a, I_b, and I_c in Eq. (29), which depend on the detailed Fermi-surface structure. These can not be done analytically even with the approximation for $f(x)$ given in Eq. (35), but we believe that an accurate enough estimate is obtained by using the simple function

$$f(\epsilon/W)\cong\epsilon/(\epsilon+W). \tag{A3}$$

Then

$$\sum_n\frac{\delta_n^2}{n}=\ln\frac{\epsilon_k}{W}-1, \tag{A4}$$

and

$$I_b=\int_0^\infty dx\int_0^\infty\frac{dy}{y^2}\left(\frac{x+y}{1+x+y}-\frac{x}{1+x}\right)^2=\frac{1}{2}. \tag{A5}$$

I_c may also be calculated explicitly.

$$\begin{aligned}I_c&=\int_0^\infty dx\cos^2\tfrac{1}{2}\pi\frac{x}{x+1}=\int_1^\infty dy\sin^2\frac{\pi}{2y}\\ &=\tfrac{1}{2}\pi\int_0^{\pi/2}\frac{\sin^2 x}{x^2}dx\\ &=\tfrac{1}{2}\pi\{-2/\pi+\mathrm{Si}(\pi)\}=\tfrac{1}{2}\pi(1.215).\end{aligned} \tag{A6}$$

These values have been inserted into Eqs. (22), (24), and (25) in order to obtain the variational results given in the text.