PHYSICAL REVIEW VOLUME 92, NUMBER 1 OCTOBER 1, 1953

Use of Non-orthogonal Wave Functions in the Treatment of Solids, with Applications to Ferromagnetism*

W. J. CARR, JR.
Westinghouse Research Laboratories, East Pittsburgh, Pennsylvania

(Received March 11, 1953)

It is proved by a rearrangement of Löwdin's solution to the many-body problem that a vector model type energy expression is valid for a solid provided (in the simplest case of one electron per atom) the number of nearest neighbors times the overlap integral between them is small compared with unity.

The approximation is also carried to include third order permutation terms, and the spin coupling via closed shells these "triple exchange" terms can produce is discussed.

General expressions are given for the exchange integrals to be used in a vector model solution. These under the proper conditions reduce to the conventional integrals except for a slight modification, which helps to make the integrals negative.

INTRODUCTION

IN the theory of solids, two separate points of view exist, one emphasizing the "free" aspect of electrons in the solid, the other the "bound" aspect. The first, known as the collective-electron or energy-band theory, makes use of one-electron wave functions having the Bloch form of a free-electron wave multiplied by a function having the periodicity of the crystal lattice. The second places more stress on the bound nature of the electrons, and in its simplest form, the Heisenberg model, regards the solid as just a collection of atoms, with each pair behaving as a hydrogen molecule.

A slightly more refined atomic picture is found in the Dirac-Van Vleck formulation, wherein the energy of the solid is assumed to be

$$E=E_0-\sum_{j>i} J_{ij}(\tfrac{1}{2}+2\langle \mathbf{S}_i\cdot \mathbf{S}_j\rangle), \tag{1}$$

with E_0 including the kinetic energy of all electrons plus the classical Coulomb energies of nuclei and smeared out electronic charge; and with J_{ij} a Heitler-London exchange integral between the ith and jth wave functions and $\mathbf{S}_i$ the vector spin operator (in multiples of $h/2\pi$) for the ith electron. The angular brackets indicate an average over a spin function.

The relative merits of the collective *vs* atomic points of view are, in brief, the following. The collective picture has the advantage that the Bloch wave functions are orthogonal, which leads to greater mathematical simplicity. It has a disadvantage in not containing the physical fact that electrons, irrespective of spin, try to avoid one another due to their repulsive Coulomb interaction. This disadvantage causes states of maximum spin to be unduly favored, because the theory allows two electrons with antiparallel spin to move independently of one another while the Pauli principle enforces its automatic correlation on electrons with parallel spin.

On the atomic model, the electrons are made to avoid one another in a certain sense by being localized around the different nuclei, the degree of separation depending upon the overlap of neighboring wave functions. In addition, the localized point of view has the advantage of physical simplicity, since (1) reduces the problem of a solid to the familiar diatomic molecule problem. In spite of these advantages, use of the Dirac-Van Vleck model has been restricted, because of the following objection. It has been pointed out by several writers[1-4] that in solving for the energy of a linear chain of hydrogen-like atoms, as a simple example, one obtains a result of the form

$$E=\frac{E_0-O(NJ)+O(N^2T^2J)\cdots}{1-O(NT^2)+O(N^2T^4)\cdots}, \tag{2}$$

when an antisymmetric wave function is used. The symbol O denotes the order of magnitude of the term, N the number of atoms in the chain, and T an overlap integral for nearest neighbors, which might reasonably have the value 10^{-1} to 10^{-2}. The exchange integral J also refers to nearest neighbors. Setting the denominator equal to unity and neglecting all terms beyond the second in the numerator gives approximately the result expected by (1), but such procedure seems obviously foolish when N is a very large number. Nevertheless, Van Vleck[3] in special cases was able to show that such a naive approach does give substantially the correct answer, because the remaining terms, in a rather obscure manner, tend to cancel each other. Van Vleck's analysis is convincing in that use of this theory leads to no "catastrophe." Doubts have remained as to whether, in general, the model is actually a good approximation.[4]

The purpose of the present paper is to investigate the problem further in order to determine under what conditions, and with what modifications, a vector model similar to (1) can be used; and to consider some applications to ferromagnetism.

---

* Work partially supported by the U. S. Air Force.

[1] J. C. Slater, Phys. Rev. **35**, 509 (1930).
[2] D. R. Inglis, Phys. Rev. **46**, 135 (1934).
[3] J. H. Van Vleck, Phys. Rev. **49**, 232 (1936).
[4] J. C. Slater, Washington Conference on Magnetism, University of Maryland, 1952 [Revs. Modern Phys. **25**, 199 (1953)].

THE WAVE FUNCTION FOR THE SYSTEM

For any one-electron model of a many-electron system, one of the most satisfactory ways of describing the complete wave function is by means of Slater determinants of one-electron functions ψ_i, each multiplied by a spin function ξ_i. The latter is specified by either α_i or β_i, the eigenfunctions for the component of spin in a particular direction.

Thus

$$\Psi=\begin{vmatrix} \psi_1(1) & \xi_1(1) & \cdots & & \\ \cdot & \psi_2(2) & \xi_2(2) & \cdots & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & & \cdot & \\ & \cdot & & & \cdot \\ & & & \psi_N(N) & \xi_N(N) \end{vmatrix}, \tag{3}$$

where the number in parenthesis labels the electron whose coordinates appear and the subscript designates the wave function.

If a linear combination of determinants such as (3) is taken and a large enough number of terms considered, then, no doubt, a good solution for the energy could be obtained for any type of ψ_i, whether spread out over the entire crystal or localized about one or more nuclei. In practice, the solution of the secular equation to determine the correct linear combination is difficult, if not impossible, to obtain when a large number of determinants with different ψ_i's are involved, and therefore in reality the method is highly restricted. More often, one attempts to approximate the wave function by a single determinant, and it is in this practical sense that one model may lead to correct results and the other incorrect. In fact, for the single determinantal wave function there is an optimum set of ψ_i, namely, that set which leaves Ψ stationary with respect to small variations $\delta\psi_i$ and brings about the lowest energy. It conceivably could happen that this set would contain a mixture of both localized and Bloch functions. In any event, to determine what this best set is, one first must solve the general problem of the energy for nonorthogonal ψ_i, which is essentially what the vector model attempts to do. Viewed in this manner, the energy band method is simply a special case of an equation such as (1), obtained by restricting the ψ_i to a certain type of orthogonal one-electron wave function.

FORM OF THE ENERGY EXPRESSION

According to the prescription of Van Vleck,[3] one can calculate the energy from (3) by naively overlooking—in all terms but the first exchange integral—the fact that the ψ's are not in general orthogonal. It is worth doing a calculation similar to this in order to establish the form to seek when using a more rigorous method. However, the present calculation will be made considerably better than usual in order to obtain a correct expression for the energy up to and including the third order in overlap.[5] There is little point in calculating higher terms since they must be negligible if the method is to prove useful. For simplicity, one may think in terms of one electron per atom, although the proof will apply to any number. With the wave function (3) the energy may be expressed

$$E=\frac{\int\Psi^*H\Psi d\tau}{\int\Psi^*\Psi d\tau} =\frac{N!\int\psi_1^*(1)\cdots\psi_N^*(N)\xi_1(1)\cdots\xi_N(N)H\Psi d\tau}{N!\int\psi_1^*(1)\cdots\psi_N^*(N)\xi_1(1)\cdots\xi_N(N)\Psi d\tau} \tag{4}$$

since Ψ can be expanded into $N!$ permutations of the diagonal element, each of which gives the same result because H is symmetric in all electrons. Unless expressly designated, $d\tau$ will always indicate an integral over all orbital, and sum over all spin, variables involved in the integrand. Expanding Ψ in terms of permutations gives

$$\Psi=\psi_1(1)\cdots\psi_N(N)\xi_1(1)\cdots\xi_N(N) +\textstyle\sum_\nu{}'P_\nu(-1)^p\psi_1(1)\cdots\psi_N(N)\xi_1(1)\cdots\xi_N(N), \tag{5}$$

where P_ν is a permutation operator that interchanges the electrons, p the parity of the permutation, and the $\sum'$ is over all permutations except the identity one. Taking the ψ's to be normalized and shortening the notation to $\varphi_0=\psi_1(1)\cdots\psi_N(N)$, $\varphi_s=\xi_1(1)\cdots\xi_N(N)$, $E_0=\int\varphi_0^*H\varphi_0 d\tau$ (E_0 is independent of spin and is the energy obtained without antisymmetrizing the wave function), one finds

$$E=\frac{E_0+\sum_\nu{}'(-1)^p\int\varphi_0^*\varphi_sHP_\nu\varphi_0\varphi_s d\tau}{1+\sum_\nu{}'(-1)^p\int\varphi_0^*\varphi_sP_\nu\varphi_0\varphi_s d\tau}$$

or

$$E=E_0+\frac{\sum_\nu{}'(-1)^p\int\varphi_0^*\varphi_sHP_\nu\varphi_0\varphi_s d\tau -E_0\sum_\nu{}'(-1)^p\int\varphi_0^*\varphi_sP_\nu\varphi_0\varphi_s d\tau}{1+\sum_\nu{}'(-1)^p\int\varphi_0^*\varphi_sP_\nu\varphi_0\varphi_s d\tau}. \tag{7}$$

In a purely formal way, $P_\nu\varphi_0\varphi_s=P_{\nu0}\varphi_0P_{\nu s}\varphi_s$, where $P_{\nu0}$ permutes the orbital part, and $P_{\nu s}$ the spin part, of the wave function. Then, assuming H to be independent of spin, containing only kinetic and electrostatic energy terms,

$$E=E_0+\frac{\sum_\nu{}'(-1)^p[\int\varphi_0^*HP_{\nu0}\varphi_0 d\tau -E_0\int\varphi_0^*P_{\nu0}\varphi_0 d\tau]\langle P_{\nu s}\rangle}{1+\sum_\nu{}'(-1)^p\int\varphi_0^*P_{\nu0}\varphi_0 d\tau\langle P_{\nu s}\rangle}, \tag{8}$$

where $\langle P_{\nu s}\rangle=\sum\varphi_sP_{\nu s}\varphi_s$ (the sum being taken over the spin coordinates) is the mean value of the permutation operator over the spin function obtained by taking the diagonal elements of (3). $\langle P_{\nu s}\rangle$ in this case can have only the values zero or one.

In summing (8) over the various permutations it is convenient to denote a particular permutation P_ν by $\overset{ij\cdots l}{P}_{km\cdots n}$ meaning $k\rightarrow i$, $m\rightarrow j$, $n\rightarrow l$. The bracketed ex-

---

[5] By the order in overlap, we mean the number of integrals over pairs of wave functions centered about different atoms. For example, in Eq. (2) J is of second order, JT^2 of fourth order, etc.

pression in the numerator of (8) for this particular permutation shall be denoted by $J_{km\cdots n}^{ij\cdots l}$. It is shown in Appendix A that J can be put into a very simple form. The expansion of (8) runs as follows:

$$E=E_0+\frac{(-1)\sum\limits_{j>i} J_{ji}{}^{ij}\langle P_{ji}{}^{ij}\rangle+\sum\limits_{k>j>i}(J_{jki}{}^{ijk}\langle P_{jki}{}^{ijk}\rangle+J_{kij}{}^{ijk}\langle P_{kij}{}^{ijk}\rangle)+\cdots}{1-\sum\limits_{j>i} T_{ij}T_{ji}\langle P_{ji}{}^{ij}\rangle+\cdots}. \tag{9}$$

From the Appendix,

$$\begin{aligned}
J_{ji}{}^{ij}\equiv J_{ij}=&\int\psi_i^*(1)\psi_j^*(2)\mathcal{H}(1,2)\psi_j(1)\psi_i(2)d\tau\\
&-\int\psi_i^*\psi_jd\tau\int\psi_j^*\psi_id\tau\int\psi_i^*(1)\psi_j^*(2)\\
&\qquad\times\mathcal{H}(1,2)\psi_i(1)\psi_j(2)d\tau\\
\equiv&(ij|\mathcal{H}(1,2)|ji)-T_{ij}T_{ji}(ij|\mathcal{H}(1,2)|ij),\\
J_{jki}{}^{ijk}=&\int\psi_i^*(1)\psi_j^*(2)\psi_k^*(3)\mathcal{H}(1,2,3)\psi_j(1)\psi_k(2)\psi_i(3)d\tau\\
&-\int\psi_i^*\psi_jd\tau\int\psi_j^*\psi_kd\tau\int\psi_k^*\psi_id\tau\int\psi_i^*(1)\\
&\qquad\times\psi_j^*(2)\psi_k^*(3)\mathcal{H}(1,2,3)\psi_i(1)\psi_j(2)\psi_k(3)d\tau\\
\equiv&(ijk|\mathcal{H}(1,2,3)|jki)\\
&-T_{ij}T_{jk}T_{ki}(ijk|\mathcal{H}(1,2,3)|ijk),
\end{aligned}\tag{10}$$

$$T_{ij}=\int\psi_i^*\psi_jd\tau,$$

$$\begin{aligned}
\mathcal{H}(1,2)=&\frac{-\hbar^2}{2m}(\nabla_1{}^2+\nabla_2{}^2)+V_1+V_2\\
&+\sum_{l\neq i,j}\left(\int\frac{e^2}{r_{1l}}|\psi_l(l)|^2d\tau_l+\int\frac{e^2}{r_{2l}}|\psi_l(l)|^2d\tau_l\right)+\frac{e^2}{r_{12}},
\end{aligned}$$

$$\begin{aligned}
\mathcal{H}(1,2,3)=&\frac{-\hbar^2}{2m}(\nabla_1{}^2+\nabla_2{}^2+\nabla_3{}^2)+V_1+V_2+V_3\\
&+\sum_{l\neq i,j,k}\left(\int\frac{e^2}{r_{1l}}|\psi_l(l)|^2d\tau_l+\int\frac{e^2}{r_{2l}}|\psi_l(l)|^2d\tau_l+\int\frac{e^2}{r_{3l}}|\psi_l(l)|^2d\tau_l\right)\\
&+\frac{e^2}{r_{12}}+\frac{e^2}{r_{23}}+\frac{e^2}{r_{13}},
\end{aligned}$$

$$E_0=W+\sum_i\left(i\left|\frac{-\hbar^2}{2m}\nabla_1{}^2+V_1\right|i\right)+\sum_{j>i}\left(ij\left|\frac{e^2}{r_{12}}\right|ij\right),$$

where V_1 is the potential between electron 1 and all nuclei, $-(\hbar^2/2m)\nabla_1{}^2$ is the kinetic energy operator for electron 1, e^2/r_{12} is the Coulomb interaction between electrons 1 and 2, and W is the electrostatic energy among nuclei.

It will be recognized that J_{ij} is of second order in overlap and $J_{jki}{}^{ijk}$ of third order. The Hamiltonian $\mathcal{H}(1,2)$ is nothing more than the Hamiltonian for two electrons in the time average potential of all other charges in the lattice (overlap neglected). Similar remarks apply for $\mathcal{H}(1,2,3)$.

By employing the same process of division that was used to remove E_0, one can take the next two groups of terms in the numerator of (9) out of the fraction. Thus

$$\begin{aligned}E=E_0-\sum_{j>i}J_{ij}\langle P_{ji}{}^{ij}\rangle+\sum_{k>j>i}(&J_{jki}{}^{ijk}\langle P_{jki}{}^{ijk}\rangle\\&+J_{kij}{}^{ijk}\langle P_{kij}{}^{ijk}\rangle)+R.\end{aligned}\tag{11}$$

For small overlap the series up to R is well behaved, the aggregate of triple permutation terms being small compared with the J_{ij} terms which, in turn, are small compared with E_0. The remainder R is a fraction containing in the numerator terms involving only fourth and higher orders in overlap. For large N, however, both numerator and denominator are seemingly divergent series similar to those illustrated in Eq. (2).

If R can be neglected the derivation is complete. Before attempting to find the conditions under which the remainder in (11) really can be overlooked, it is of interest to point out that the equation is in a form which lends itself to simple interpretation.

It is desired first to compare J_{ij} with the exchange integral in the Heitler-London description of the hydrogen molecule. For that problem the wave functions are $\psi_1(1)\psi_2(2)\pm\psi_1(2)\psi_2(1)$ multiplied by spin functions. The plus sign goes with the singlet state, the minus with the triplet. The energy,

$$\frac{\int\psi_1(1)\psi_2(2)H\psi_1(1)\psi_2(2)d\tau\pm\int\psi_1(1)\psi_2(2)H\psi_1(2)\psi_2(1)d\tau}{1\pm T_{12}{}^2},$$

may be written as

$$E=(12|H|12)\pm\frac{(12|H|21)-T_{12}{}^2(12|H|12)}{1\pm T_{12}{}^2},$$

and, when the overlap is small,

$$E\simeq(12|H|12)\pm[(12|H|21)-T_{12}{}^2(12|H|12)].\tag{12}$$

The first term on the right in (12) is in exact analogy to the E_0 term of (11); and the bracketed term, the exchange integral,[6] differs from J_{12} only in the following

---

[6] The Hamiltonian is $H=H_1+H_2+e^2/r_{ab}+e^2/r_{12}-e^2/r_{1b}-e^2/r_{2a}$, where r_{ab} is the internuclear distance, r_{12} the interelectronic distance, r_{1b}, r_{2a} the distances between electrons on one atom and nuclei of the other, and H_1, H_2 are the Hamiltonians for the isolated atoms. As normally used, ψ_1 and ψ_2 are eigenfunctions for the hydrogen atoms, so the exchange integral becomes

$$\left(12\left|\frac{e^2}{r_{ab}}+\frac{e^2}{r_{12}}-\frac{e^2}{r_{1b}}-\frac{e^2}{r_{2a}}\right|21\right)-T_{12}{}^2\left(12\left|\frac{e^2}{r_{ab}}+\frac{e^2}{r_{12}}-\frac{e^2}{r_{1b}}-\frac{e^2}{r_{2a}}\right|12\right).$$

The e^2/r_{ab} term, being a constant, would also cancel out, but no manner: the H of (12) is the Hamiltonian for two electrons in the potential of two nuclei, while the $\mathcal{H}(1, 2)$ in the corresponding J_{12} term of (11) is, as stated before, a Hamiltonian for two electrons in the average potential of the rest of the crystal. $J_{kij}{}^{ijk}$ bears a similar relation to the three atom problem.

To summarize, the right side of (11) contains, firstly, the spin independent kinetic and classical Coulomb energies; secondly, the spin dependent energy obtained by assuming the electrons to interact in pairs; thirdly, the additional spin dependent part that comes from assuming the electrons as interacting in groups of three. Had the expansion been carried further the effect of groups of four, etc., would have appeared.

If one takes only the first two terms on the right side of (11) and substitutes Dirac's relation[7] $P_{ji}{}^{ij}=\frac{1}{2}+2\mathbf{S}_i\cdot\mathbf{S}_j$, then

$$E\simeq E_0-\sum_{j>i} J_{ij}(\tfrac{1}{2}+2\langle\mathbf{S}_i\cdot\mathbf{S}_j\rangle), \tag{13}$$

which corresponds to the vector model when the spin function is restricted to a simple product φ_s. The J of (13), however, is somewhat different from the J of Van Vleck, a point that will be discussed later.

CALCULATION OF THE ENERGY BY LÖWDIN'S METHOD

An estimate of the conditions under which one is justified in neglecting the remainder in (11) may be obtained by expressing the determinant (3) in terms of minors, rather than permutations, and then using an expansion due to Löwdin.[8]

If the product $\psi_j\xi_j$ in (3) is denoted by φ_j, the wave function can be written

$$\Psi=\sum_i \varphi_j(i)\Psi_{i,j}=\sum_{\substack{j,k\\ j\neq k,\, i\neq l}} \varphi_j(i)\varphi_k(l)\Psi_{il,jk}, \tag{14}$$

where $\Psi_{i,j}$ is the cofactor of $\varphi_j(i)$ in Ψ, and similarly $\Psi_{il,jk}$ is plus or minus the minor of Ψ formed by deleting the ith and lth rows and jth and kth columns. Let the Hamiltonian be written

$$H=W+\sum_i H(i)+\sum_{\substack{i,l\\ i\neq l}} H(i,l),$$

with $H(i)$ depending only on coordinates of the ith electron, and $H(i, l)$ depending only on those of i and l, while W is independent of electron coordinates. In the present case,

$$H(i)=-(\hbar^2/2m)\nabla_i{}^2+V_i, \quad \text{and} \quad H(i,l)=e^2/2r_{il}.$$

From (4) the energy is

$$E=W+\frac{\int\varphi_1{}^*(1)\cdots\varphi_N{}^*(N)\sum_i H(i)\sum_j \varphi_j(i)\Psi_{i,j}d\tau}{\int\varphi_1{}^*(1)\cdots\varphi_N{}^*(N)\Psi d\tau}$$
$$+\frac{\int\varphi_1{}^*(1)\cdots\varphi_N{}^*(N)\times\sum\limits_{\substack{i,l\\ i\neq l}} H(i,l)\sum\limits_{\substack{j,k\\ j\neq k}} \varphi_j(i)\varphi_k(l)\Psi_{il,jk}d\tau}{\int\varphi_1{}^*(1)\cdots\varphi_N{}^*(N)\Psi d\tau}. \tag{15}$$

Now $\int\varphi_1{}^*(1)\cdots\varphi_N{}^*(N)\Psi d\tau=\det t\equiv\Delta$, where t is a matrix with elements $t_{ij}=\int\psi_i{}^*\xi_i\psi_j\xi_j d\tau$. Further,

$$\int\varphi_1{}^*(1)\cdots\varphi_{i-1}{}^*(i-1)\varphi_{i+1}{}^*(i+1)\cdots \times\varphi_N{}^*(N)\Psi_{i,j}d\tau=\Delta_{i,j},$$

which is the cofactor of t_{ij} in Δ, and

$$\int\varphi_1{}^*(1)\cdots\varphi_{i-1}{}^*(i-1)\varphi_{(i+1)}{}^*(i+1)\cdots \times\varphi_{l-1}{}^*(l-1)\varphi_{l+1}{}^*(l+1)\cdots\varphi_N{}^*(N)\Psi_{il,jk}d\tau=\Delta_{il,jk}.$$

Therefore

$$E=W+\sum_{i,j}\frac{\Delta_{i,j}}{\Delta}\int\varphi_i{}^*(i)H(i)\varphi_j(i)d\tau$$
$$+\sum_{\substack{i,l,j,k\\ j\neq k,\, i\neq l}}\frac{\Delta_{il,jk}}{\Delta}\int\varphi_i{}^*(i)\varphi_l{}^*(l)H(i,l)\varphi_j(i)\varphi_k(l)d\tau. \tag{16}$$

Equation (16) is essentially Löwdin's expression,[9] which, as he pointed out, contains no "catastrophe" since the matrix elements are multiplied only by the ratios $\Delta_{i,j}/\Delta$ or $\Delta_{il,jk}/\Delta$. If t^{-1} is the reciprocal matrix of t $(tt^{-1}=1)$, then from the theory of determinants $t^{-1}{}_{ji}=\Delta_{i,j}/\Delta$ and $t^{-1}{}_{ji}t^{-1}{}_{kl}-t^{-1}{}_{jl}t^{-1}{}_{ki}=\Delta_{il,jk}/\Delta$. The key point in the further development is Löwdin's expansion of $t^{-1}=(1+S)^{-1}$ in the power series

$$t^{-1}{}_{ji}=(1-S+S^2-S^3\cdots)_{ji}$$
$$=\delta_{ji}-S_{ji}+\sum_k S_{jk}S_{ki}-\cdots, \tag{17}$$

where the matrix elements $S_{ji}=t_{ji}-\delta_{ji}$ are the off-diagonal elements of the t matrix. They are zero if i and j have different spins or when ψ_i and ψ_j are orthogonal (for example wave functions on the same atom). The sum over spin coordinates of $\xi_i\xi_j$ may be replaced by $\langle P_{ji}{}^{ij}\rangle$ so that $t_{ji}=T_{ji}\langle P_{ji}{}^{ij}\rangle$.

By substituting (17) into (16) and grouping terms, one can obtain an energy expression up to any order in the overlap. It is sufficient to state that the first three groups of terms so obtained are exactly equal to those of (11) (see Appendix B).

---

harm is done by leaving it in, since to do so makes $(12|e^2/r_{ab}+e^2/r_{12}-e^2/r_{1b}-e^2/r_{2a}|12)$ simply the classical electrostatic energy between two atoms with smeared-out charge. This quantity multiplied by T^2, at least in the hydrogen molecule, can be neglected compared with $(12|e^2/r_{ab}+e^2/r_{12}-e^2/r_{1b}-e^2/r_{2a}|21)$, which is the familiar form of the exchange integral. For present purposes, however, it is not demanded that ψ_1 and ψ_2 be eigenfunctions and consequently the general form of the exchange integral is retained.

[7] P. A. M. Dirac, Proc. Roy. Soc. (London) **A123**, 714 (1929).

[8] Per-Olov Löwdin, J. Chem. Phys. **18**, 365 (1950).

[9] The derivation given here, pointed out to me by C. Zener, is somewhat more straightforward than Löwdin's, which is based upon the premise that certain power series expansions exist.

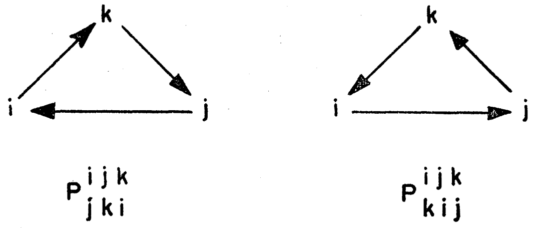

FIG. 1. The two permutation possibilities for three electrons exchanging position.

In the scheme of localized wave functions, if the ψ_i are selected so that only the overlap between nearest neighbors is important, the term $(S^n)_{ji}$ in (17) approximately equals the nth power of the nearest neighbor overlap integral, $S_{ab}=\int\psi_a{}^*\psi_b d\tau$, times the number of paths from j to i via n jumps between nearest neighbors (these paths being through wave functions having the same spin).

It is apparent that in the nearest neighbor approximation $(S^n)_{ji}\leqslant Z^{n-1}(S_{ab})^n$, where Z represents the number of neighboring wave functions which have a nonvanishing overlap with ψ_a. Therefore, a sufficient condition for the convergence of the right side of (17) is $ZS_{ab}<1$.

Thus, having ZS_{ab} small compared with unity is sufficient, though possibly not necessary, for the neglect of R in (11). The proof holds for any number of electrons per atom, since wave functions on the same atom are orthogonal.

If only the criterion for neglect of the three-body compared with the two-body exchange is considered, it is more nearly that $Z'S_{ab}$ must be small compared with unity. The quantity Z' is the number of wave functions of the right spin which overlap both a and b.

EXTENSION TO A LINEAR COMBINATION OF DETERMINANTS WITH THE SAME ψ_i

For a given set of one-electron wave functions ψ_i, an improved total wave function may be obtained by taking a linear combination of determinants,

$$\Psi=\sum_\mu A_\mu\begin{vmatrix}\psi_1(1)\xi_{1\mu}(1) & & \\ & \psi_2(2)\xi_{2\mu}(2) & \\ & \ddots & \\ & & \psi_N(N)\xi_{N\mu}(N)\end{vmatrix}, \qquad (18)$$

where the A's are coefficients, and the spin functions ξ represent a given number of α's and β's which in each determinant are assigned to the ψ_i in a different way. In particular, it is possible to construct eigenfunctions of both the z component and the square of the total spin angular momentum operator in this manner.[10]

The extension of (18) to the energy calculation by the first method is very simple, for Ψ can be written

$$\Psi=\sum_\nu P_\nu(-1)^p\psi_1(1)\cdots\psi_N(N)\sum_\mu A_\mu\xi_{1\mu}(1)\cdots \times\xi_{N\mu}(N)=\sum_\nu P_\nu(-1)^p\varphi_0\varphi_s', \qquad (19)$$

and the calculation thereafter leads to an expression of exactly the form (11), except that $\langle P\rangle$ now means the sum of $\varphi_s'P\varphi_s'$ over the spin coordinates and thus the mean can have values other than zero or one. The spin function φ_s' is a linear combination of product one-electron spins formed by taking the diagonal elements of each determinant multiplied by the coefficient in front of the determinant, i.e.,

$$\varphi_s'=\sum_\mu A_\mu\xi_{1\mu}(1)\cdots\xi_{N\mu}(N). \qquad (20)$$

It has been assumed that φ_s' is normalized.

The remainder in (11) for the wave function (18) will certainly be small under some conditions, and there seems no reason to suspect that the conditions will differ greatly from those calculated for the case of a single determinant.

Thus, for a given set of orbital ψ_i the problem of a many-electron system is reduced to finding the best spin function for the effective Hamiltonian,

$$H_{\text{eff}}=E_0-\sum_{j>i}J_{ij}P_{ji}{}^{ij}+\sum_{k>j>i}(J_{kij}{}^{ijk}P_{kij}{}^{ijk}+J_{jki}{}^{ijk}P_{jki}{}^{ijk}). \qquad (21)$$

Equation (21) corresponds to the usual Dirac-Van Vleck model, extended to include third order permutations and modified as regard to exchange integrals in the manner previously discussed.

SIGN OF THE EXCHANGE INTEGRAL

In the early quantum theory of solids the presence of ferromagnetism in certain elements was attributed by Heisenberg as due to a positive sign for J_{ij}, the exchange integral between d wave functions on neighboring atoms. Of the several attempts[11,12] made to calculate this quantity all have used the standard Heitler-London approximation to the hydrogen molecule exchange integral (see footnote 6). The value of these calculations is to show that the positive and negative terms in the integrand cause cancellation to a high order so that a slight modiffication of the integrand can be of importance.

The correct value of J_{ij} is from (10) $(ij|\mathcal{H}(1,2)|ji)-T_{ij}T_{ji}(ij|\mathcal{H}(1,2)|ij)$. We assume, in analogy with the hydrogen molecule, that $(ij|\mathcal{H}(1,2)|ji)$ is the principal term. The Hamiltonian in this term, as already mentioned, differs from the two-electron Hamiltonian in that it contains the potential of the complete crystal lattice. Far away atoms will, of course, add nothing to this potential, but neighboring atoms might be expected to add a considerable negative term, because the d electron will, to some extent, penetrate the charge cloud around nearby atoms and therefore experience a nuclear attraction.

The remaining point to consider is that, according to experimental evidence, diatomic molecules, with very few exceptions, have a negative exchange integral.

---

[10] For example, see H. Eyring, and G. E. Kimball, J. Chem. Phys. **1**, 239 (1933).

[11] E. P. Wohlfarth, Nature **163**, 57 (1949).

[12] H. Kaplan, Phys. Rev. **85**, 1038 (1952).

Since in a solid the tendency for a negative integral is even greater, as was just shown, Zener's assumption[13] that the exchange between *d* shells in the transition elements is negative becomes a very attractive working model. This assumption makes it necessary for the ferromagnetic coupling in the iron group of elements to come about indirectly, such as via the conduction electrons.

It is to be understood that all of the preceding discussion applies to wave functions *i* and *j* which are nonorthogonal. In the case of accidental orthogonality between wave functions on different atoms, the exchange integral reduces to that of the band theory $(ij|e^2/r_{12}|ji)$ and becomes positive.

THE TRIPLE EXCHANGE

The two groups of triple exchange terms in (21) correspond to the two possibilities for three electrons exchanging position: clockwise and counterclockwise as shown in Fig. 1.

Normally, if the overlapping of wave functions is small, the triple exchange will be relatively unimportant compared with the simple exchange terms. In one common case, however, the triple exchange might be expected to play an important part: namely, the interaction between a wave function ψ_i and a closed shell of electrons, for in this case the simple exchange terms sum to a constant independent of spin. To show this fact one can replace the permutation operator with the spin operators, obtaining

$$\sum_{i,j} J_{ij}\langle P_{ji}{}^{ij}\rangle=\sum_{i,j} J_{ij}(\tfrac{1}{2}+2\langle S_i\cdot S_j\rangle)$$
$$=\text{const}+2\langle \textstyle\sum_i S_i\cdot\sum_j J_{ij}S_j\rangle,$$

and the spin dependent part is zero if *j* is summed over a closed shell.

In many ferro- or antiferromagnetic materials a lattice arrangement exists where atoms possessing a magnetic moment are well separated by atoms having essentially closed shells. Two mechanisms for a spin coupling that can be transmitted through closed shell atoms have been proposed: the double exchange of Zener[14] and Heikes,[15] which takes place in cases of degeneracy, such as in off-stoichiometric compounds; and the superexchange of Kramers,[16] Anderson,[17] and Van Vleck,[18] which can take place in ordinary stoichiometric materials. Both of these mechanisms arise when one considers interactions between different electronic configurations, i.e., between determinants having different sets of ψ_i. The superexchange is just a statement of the fact that the "closed shell" atoms in a second approximation really are not closed, and therefore allow a coupling to take place.

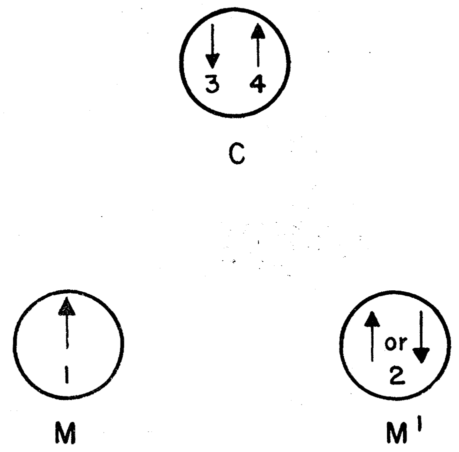

FIG. 2. Arrangements of three atoms.

The triple exchange terms, on the other hand, admit an appreciable coupling in certain instances even in the closed shell approximation. Consider three atoms arranged as in Fig. 2, with *M* and *M'* having one electron each but overlapping each other very little, and *C* having a closed shell of two electrons which overlap both *M* and *M'* considerably.

For the case of *M* and *M'* having parallel spins, the spin function can be written

$$\alpha(1)\alpha(2)[\alpha(3)\beta(4)-\alpha(4)\beta(3)]/\sqrt{2};$$

and for antiparallel spins,

$$[\alpha(1)\beta(2)-\alpha(2)\beta(1)][\alpha(3)\beta(4)-\alpha(4)\beta(3)]/2.$$

Taking account of the fact that $\psi_3=\psi_4$, one easily can show that

$$\sum_{\substack{1\\k>j>i}}^{4} (J_{kij}{}^{ijk}\langle P_{kij}{}^{ijk}\rangle+J_{jki}{}^{ijk}\langle P_{jki}{}^{ijk}\rangle)$$

equals $(J_{312}{}^{123}+J_{231}{}^{123})$ for the triplet and $-(J_{312}{}^{123}+J_{231}{}^{123})$ for the singlet. This result can be summarized by the expression $(J_{312}{}^{123}+J_{231}{}^{123})\langle P_{21}{}^{12}\rangle$, indicating a coupling between 1 and 2 of the same form as the direct coupling. The ratio of this coupling to the direct coupling is, other things being equivalent, of the order $T_{13}T_{12}T_{32}/T_{12}{}^2$, a number that can be quite large if the closed shell atom has a big overlap for *M* and *M'*. If $J_{312}{}^{123}$ is to have appreciable magnitude, however, 1, 2, and 3 must all overlap each other to some extent, so the interaction is probably strongest when the three atoms are approximately at the corners of an equilateral triangle.

A considerable number of materials with configurations such as this exist. Rough estimates indicate that the triple exchange integral can feasibly attain values large enough to account for Curie temperatures observed. If $J_{312}{}^{123}$ is negative it contributes a ferromagnetic coupling, if positive, an antiferromagnetic one.

---

[13] C. Zener, Phys. Rev. **81**, 440 (1951).
[14] C. Zener, Phys. Rev. **82**, 403 (1951).
[15] R. Heikes, thesis, University of Chicago (unpublished).
[16] H. A. Kramers, Physica **1**, 182 (1934).
[17] P. W. Anderson, Phys. Rev. **79**, 350 (1950).
[18] J. H. Van Vleck, J. phys. radium **12**, 262 (1951).

APPENDIX A

The Hamiltonian is

$$H=W-\sum_i \frac{\hbar^2}{2m}\nabla_i^2+\sum_i V_i+\frac{1}{2}\sum_{\substack{i,j\\ i\neq j}} \frac{e^2}{r_{ij}},$$

where W is the interaction among nuclei, $-(\hbar^2/2m)\nabla_i^2$ the kinetic energy operator of the ith electron, V_i the Coulomb interaction between the ith electron and all nuclei, and e^2/r_{ij} the Coulomb interaction between two electrons.

Let H be written

$$H=H(1, 2\cdots n)+\sum_{i=1\cdots n}\ \sum_{j\neq 1\cdots n} \frac{e^2}{r_{ij}}+H', \tag{22}$$

where the first term contains only the coordinates of the electrons 1 to n, the second term is the interaction of these electrons with all other electrons, and H' is the remainder, which does not depend upon the coordinates of electrons 1 to n.

Let

$$\mathcal{H}(1\cdots n)=H(1\cdots n)+\sum_{i=1\cdots n}\ \sum_{j\neq 1\cdots n} \int \frac{e^2}{r_{ij}}|\psi_j(j)|^2 d\tau_j. \tag{23}$$

Now,

$$\int \varphi_0{}^*HP_{m\cdots l}^{1\cdots m\cdots l\cdots n}\varphi_0 d\tau$$

$$=\int \psi_1{}^*(1)\cdots\psi_N{}^*(N)HP_{m\cdots l}^{1\cdots n}\psi_1(1)\cdots\psi_N(N)d\tau$$

$$=\int \varphi_0{}^*(H-H')P_{m\cdots l}^{1\cdots n}\varphi_0 d\tau$$

$$+\int \varphi_0{}^*H'\varphi_0 d\tau\int \varphi_0{}^*P_{m\cdots l}^{1\cdots n}\varphi_0 d\tau, \tag{24}$$

therefore

$$\int \varphi_0{}^*HP_{m\cdots l}^{1\cdots n}\varphi_0 d\tau-\int \varphi_0{}^*H\varphi_0 d\tau\int \varphi_0{}^*P_{m\cdots l}^{1\cdots n}\varphi_0 d\tau$$

$$=\int \varphi_0{}^*(H-H')P_{m\cdots l}^{1\cdots n}\varphi_0 d\tau$$

$$-\int \varphi_0{}^*(H-H')\varphi_0 d\tau\int \varphi_0{}^*P_{m\cdots l}^{1\cdots n}\varphi_0 d\tau, \tag{25}$$

and the right side of (25) is equal to

$$\int \psi_1{}^*(1)\cdots\psi_n{}^*(n)\mathcal{H}(1\cdots n)\psi_m(1)\cdots\psi_l(n)d\tau$$

$$-\int \psi_1{}^*\psi_m d\tau\cdots\int \psi_n{}^*\psi_l d\tau\int \psi_1{}^*(1)\cdots\psi_n{}^*(n)$$

$$\times\mathcal{H}(1\cdots n)\psi_1(1)\cdots\psi_n(n)d\tau. \tag{26}$$

The expression (26) shall be denoted by the usual notation

$$(1\cdots n|\mathcal{H}(1\cdots n)|m\cdots l)-T_{1m}\cdots T_{nl}$$
$$\times(1\cdots n|\mathcal{H}(1\cdots n)|1\cdots n)$$

or simply by $J_{m\cdots l}{}^{1\cdots n}$.

Now the summation over permutations involving a simple interchange gives the terms $-\sum_{j>i} J_{ji}{}^{ij}P_{ji}{}^{ij}$; the summation over permutations involving three electrons is

$$\sum_{k>j>i} (J_{kij}{}^{ijk}P_{kij}{}^{ijk}+J_{jki}{}^{ijk}P_{jki}{}^{ijk}).$$

The integral $\int \varphi_0{}^*H\varphi_0 d\tau$ evaluated in a straightforward manner is E_0 in (10).

APPENDIX B

Equation (16) may be written

$$E=W+\sum_i \frac{\Delta_{i,j}}{\Delta}(i|H(1)|i)+\sum_{\substack{i,j\\ i\neq j}} \frac{\Delta_{i,j}}{\Delta}(i|H(1)|j)$$

$$+\sum_{\substack{i,j\\ i\neq j}} \frac{\Delta_{ij,ij}}{\Delta}(ij|H(1, 2)|ij)$$

$$+\sum_{\substack{i,j\\ i\neq j}} \frac{\Delta_{ij,ji}}{\Delta}(ij|H(1, 2)|ji)$$

$$+2\sum_{\substack{i,j,k\\ i\neq j\neq k}} \frac{\Delta_{ik,jk}}{\Delta}(ik|H(1, 2)|jk)$$

$$+2\sum_{\substack{i,j,k\\ i\neq j\neq k}} \frac{\Delta_{ij,jk}}{\Delta}(ij|H(1, 2)|jk)$$

$$+\sum_{\substack{i,j,l,k\\ i\neq j\neq l\neq k}} \frac{\Delta_{il,jk}}{\Delta}(il|H(1, 2)|jk). \tag{27}$$

By making use of (17),

$$E=W+\textstyle\sum_i(1+\sum_j S_{ij}S_{ji}-\cdots)(i|H(1)|i)$$

$$+\sum_{\substack{i,j\\ i\neq j}}(-S_{ji}+\cdots)(i|H(1)|j)$$

$$+\sum_{\substack{i,j\\ i\neq j}}(1+2\textstyle\sum_k S_{ik}S_{ki}-S_{ij}S_{ji}+\cdots)(ij|H(1,2)|ij)$$

$$+\sum_{\substack{i,j\\ i\neq j}}(-1+\cdots)(ij|H(1, 2)|ji)$$

$$+2\sum_{\substack{i,j,k\\ i\neq j\neq k}}(-S_{ji}+\cdots)(ik|H(1, 2)|jk)$$

$$+2\sum_{\substack{i,j,k\\ i\neq j\neq k}}(S_{ki}+\cdots)(ij|H(1, 2)|jk)$$

$$+\sum_{\substack{i,j.l,k,\\ i\neq j\neq l\neq k}}(S_{ji}S_{kl}-S_{jl}S_{ki}+\cdots)(il|H(1, 2)|jk). \tag{28}$$

Group these terms in the following manner:

$$\begin{aligned}E=\Big[W&+\textstyle\sum_i(i|H(1)|i)+\sum\limits_{\substack{i,j\\ i\neq j}}(ij|H(1,2)|ij)\Big]\\ &+\Big[\sum_{i,j}S_{ij}S_{ji}(i|H(1)|i)\\ &+\sum_{\substack{i,j\\ i\neq j}}\big(2\textstyle\sum_k S_{ik}S_{ki}-S_{ij}S_{ji}\big)(ij|H(1,2)|ij)\\ &-\sum_{\substack{i,j\\ i\neq j}}S_{ji}(i|H(1)|j)-\sum_{\substack{i,j\\ i\neq j}}(ij|H(1,2)|ji)\\ &\qquad-2\sum_{\substack{i,j,k\\ i\neq j\neq k}}S_{ji}(ik|H(1,2)|jk)\Big]+\cdots.\end{aligned}\tag{29}$$

The first group on the right in (29) is E_0 of (11). The second group in (29) is equal to $-\sum_{j>i}J_{ij}\langle P_{ji}{}^{ii}\rangle$ of (11), which one may prove by expanding J_{ij} as defined in (10). In the proof it must be remembered that $S_{ii}=0$ and S_{ij}, $(i|\ \ |j)$, etc. are also zero if i and j have different spins.

We state without further proof, that the next group of terms in (29), which would involve three integrals over pairs of wave functions with different i and j, is equal to the third-order terms of (11).

If the series (17) converges sufficiently, one may neglect the remaining terms in (29) and consequently R in (11).

PHYSICAL REVIEW VOLUME 92, NUMBER 1 OCTOBER 1, 1953

A Modified Theory of Production of Secondary Electrons in Solids*

A. VAN DER ZIEL
Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota

(Received June 11, 1953)

A few difficulties in the previous theories of energy loss and secondary electron production by primary electrons in metals can be removed by replacing the Coulomb interaction between a primary electron and a lattice electron by a screened Coulomb interaction, such as required by the plasma theory. In the new theory Wooldridge's mechanism of secondary electron production seems to have lost most of its importance. The energy loss distribution due to the interaction with the conduction electrons of the metal is found to have a maximum at an energy loss slightly less than $(\hbar^2k_m{}^2/2m)$, where k_m is the value of the wave vector of the conduction electrons at the Fermi level. The theory is also applied to the bound electrons, and it is shown that the screened Coulomb interaction does not change the previous results very strongly. It is also shown that the polarization of the medium has only a small influence upon the previous results for insulators.

I. INTRODUCTION

DEKKER and van der Ziel[1] recently unified the various theories of secondary electron production. This theory, which is essentially a modification of the existing theories of energy loss of fast particles in matter, is based upon a Coulomb interaction between the primary electrons and the lattice electrons. It has the same difficulties for the conduction electrons of the metal as previous theories:

(a) In a single collision between a primary electron and a lattice electron the probability $P(E_{kk'})dE_{kk'}$ of an energy loss between $E_{kk'}$ and $E_{kk'}+dE_{kk'}$ becomes infinite for $E_{kk'}\rightarrow 0$.

(b) The probability $P(k')dk'$ of a transition of a lattice electron to an energy state having an absolute value of the wave vector between k' and $k'+dk'$ becomes infinite at the Fermi level.

(c) The rate of energy loss (dE_p/dx) due to the lattice electrons for a primary electron of energy E_p varies as $E_p{}^{-1}\log(E_p/E_0')$, with a very small value of E_0'.

It is the aim of this paper to investigate whether a screened Coulomb interaction might remedy these defects. It is reasonable to expect so, for, since the difficulties are caused by the interaction of primary electrons and those lattice electrons that are passed by at larger distances, this interaction should be removed by a screened potential function.

Such a screened potential should be expected in metals according to the plasma theory of electron interaction. In that theory the interaction between a primary electron and the electrons of the metal is split into two parts.[2]

(1) An "organized" part, consisting of the interaction with the electron gas as a *whole*, resulting in relatively sharp to very sharp energy losses caused by the excitation of "plasma oscillations." This part of the interaction can explain the discrete energy losses of electrons in metal foils. It is not known how important this process

---

* Supported by U. S. Signal Corps Contract.

[1] A. J. Dekker and A. van der Ziel, Phys. Rev. **86**, 755 (1952).

[2] R. Kronig and J. Korringa, Physica **10**, 406, 800 (1943); H. A. Kramers, Physica **13**, 401 (1947); D. Bohm and E. P. Gross, Phys. Rev. **75**, 1851, 1864 (1949); D. Bohm and D. Pines, Phys. Rev. **80**, 903 (1950); **82**, 625 (1951); D. Pines and D. Bohm, Phys. Rev. **85**, 338 (1952); D. Pines, Phys. Rev. **85**, 931 (1952).