PHYSICAL REVIEW VOLUME 115, NUMBER 1 JULY 1, 1959

New Approach to the Theory of Superexchange Interactions

P. W. ANDERSON
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received February 4, 1959)

The theory of indirect exchange in poor conductors is examined from a new viewpoint in which the *d* (or *f*) shell electrons are placed in wave functions assumed to be exact solutions of the problem of a single *d*-electron in the presence of the full diamagnetic lattice. Inclusion of *d*-electron interactions leads to three spin-dependent effects which, in the usual order of their sizes, we call: superexchange *per se,* which is always antiferromagnetic; direct exchange, always ferromagnetic; and an indirect polarization effect analogous to nuclear indirect exchange. Superexchange itself is shown to be closely related to the poor conductivity, in agreement with experiment. By means of crystal field theory the parameters determining superexchange can be estimated, and in favorable cases (NiO, $LaFeO_3$) the exchange integrals can be evaluated with accuracy of several tens of percent. Qualitative understanding of the whole picture of exchange in iron group oxides and fluorides follows from these ideas.

I. INTRODUCTION

In spite of a considerable literature, a qualitative understanding of the superexchange phenomenon has been lacking. Superexchange may be defined as the exchange mechanism taking place in transition-metal salts in which the ions are fairly well separated by normally diamagnetic groups, and in which conduction is poor at all normal temperatures.[1,2] The experimental fact which is most striking is that in almost every known case this interaction is antiferromagnetic; very few insulators are true ferromagnets.[3] (Ferrimagnetism of course results from antiferromagnetic interactions.) This generalization suggests that there may be a distinct, reasonably universal mechanism for superexchange, and perhaps even that it is related to the Mott mechanism[4] which prevents conduction.

Aside from not convincingly explaining this experimental point, the theories which have been advanced have all suffered from one major difficulty: although it is obvious that a perturbation technique is correct, since the exchange effect is small compared to most energies of the problem, the perturbation problem is very difficult to define satisfactorily. Usually electron interaction problems are solved by starting from some suitable set of single-electron states satisfying an assumed one-electron Hamiltonian, and introducing the interaction as a perturbation; but the appropriate one-electron states here should have the symmetry of the problem, i.e., be running waves, while clearly the outstanding feature of this problem is that the electrons are localized, and actually the localization is caused by their interactions.[4] In addition, another large effect is the overlap of *d*-electrons onto neighboring diamagnetic groups, which is neither small nor easy to take into account, especially because of the problem of orthogonalization. As a result of these difficulties, it is not even clear whether the apparently very different schemes which have been proposed are actually distinct, or what their relationship is.[5] We shall suggest here that many of them are roughly correct in that they represent various ways of looking at parts, larger or smaller, of the same physical mechanism.

The subject divides itself naturally into two parts. First we have to write down the rather abstract, general theory of *d*-electron spins and their interactions. The first and most important part of this job is to understand the concept of a single *d*-electron in the presence of the diamagnetic lattice. Such an electron is actually a quasi-particle like the polaron, because it carries with it a cloud of associated electronic polarization. In particular, it will have an associated spin polarization.[6] The two most important properties of these particles are their kinetic energy, or desire to delocalize themselves, and their repulsion when they are too near each other. Whenever this repulsion predominates and prevents metallic conduction and the formation of bands, we show that the opposite tendency to delocalize causes a necessarily antiferromagnetic interaction which we identify as superexchange.

The physical basis for this is simply that antiparallel electrons can gain energy by spreading into nonorthogonal overlapping orbitals, where parallel electrons cannot. This term is the first in a series of types of interactions, of which we discuss the first few, including all suggestions so far made. Quantitative estimates verify that superexchange dominates when it is present.

The superexchange interaction may be expressed simply in terms of parameters related to other properties of the crystal. In the second part of the paper we choose to estimate the most important from ligand field theory[7] (which may be thought of as the theory of the isolated quasi-spin). We find that we can make at least a semi-quantitative comparison with experiment for oxide and

---

[1] H. A. Kramers, Physica **1**, 182 (1934).

[2] P. W. Anderson, Phys. Rev. **79**, 350 (1950).

[3] R. R. Heikes, Phys. Rev. **99**, 1232 (1955).

[4] N. F. Mott, Proc. Phys. Soc. (London) **A62**, 416 (1949).

[5] A number of these are discussed in J. Yamashita and J. Kanda, Phys. Rev. **109**, 730 (1958); also see R. K. Nesbet, Ann. Phys. **4**, 87 (1958).

[6] It will occasionally be convenient to have a name for these particles. We call them "quasi-spins" or "spin quasi-particles."

[7] For a review see J. S. Griffith and L. E. Orgel, Quart. Revs. (London) **11**, 381 (1957).

fluoride antiferromagnets, and explain most of the qualitative regularities in the data.

II. ISOLATED SPIN QUASI-PARTICLE

As was remarked in the introduction, the only satisfactory single-electron wave functions in a solid-state problem are running waves, and we indeed start the discussion with a single running-wave d-electron isolated in the diamagnetic crystal. We must imagine that the nuclear charges are adjusted in such a way that the d-electron sees a potential not far from what it sees in the real crystal; this is not difficult.

This d-electron in the real crystal is not very accurately simply a superposition of atomic d-electron functions, for three reasons in order of numerical importance:

(1) It must be orthogonalized to all of the wave functions of the electrons already present. In particular, the O^{--} or F^{-} electrons will have formed partially covalent bonds with d-orbitals on the ions, according to the Van Vleck "molecular orbital" scheme as adapted by Stevens and Owen,[8,9] so that the wave function has an antibonding admixture of anion functions. Since the quasi-spin is assumed to be in principle an *exact* solution of the one-electron problem plus the diamagnetic crystal, there is no need to confuse ourselves with any additional one-electron transfer effects,[10] as opposed to the true polarization effects we discuss shortly.

It should be emphasized that the main virtue of the concept is this negative result, that it contains from the start *all* one-electron effects.

We shall soon show that the most important part of the exchange results simply from these altered one-electron functions; what the spin quasi-particle concept does is to isolate these one-electron effects from polarization effects, making it possible to estimate them separately and show the latter are usually small.

(2) The electron, being a charged particle, is surrounded by an accompanying cloud of electronic dielectric polarization. Note that we must not take into account lattice polarization because most of the motions of our particles will be virtual ones with such large perturbation denominators that the motions are too rapid for the nuclei to follow. This electronic dielectric polarization is unimportant except for its numerical effect in reducing somewhat the Coulomb interactions of the particles, because it is not spin-dependent; we shall not therefore even write it down.

(3) More important for spin effects is the spin-dependent polarization caused by the exchange effect. For running d-electrons this takes the form of virtual excitations from filled bands into empty ones, while the d-electron jumps to a different k-state in the d-band; parallel electrons in the full bands go to parallel states, while antiparallel ones exchange spins with the d-electrons. The full expression, in perturbation theory, of this effect is explored in the Appendix.

In second-quantized notation,[11] we can call the many-electron wave function of the diamagnetic lattice Ψ_0, and define a fermion creation operator $s_{\mathbf{k}\sigma}^*$ such that the many electron wave function of a single quasi-particle of momentum $\mathbf{k}$ and spin σ is

$$\Psi(\mathbf{k},\sigma)=s_{\mathbf{k}\sigma}^*\Psi_0. \tag{1}$$

The s's are a set of properly anticommuting fermion operators. We should also note that really there are 5 bands of s's, which may be of complicated degenerate forms, especially in cubic crystals. An s might be expressed in terms of a series of the true one-electron operators c as (see the Appendix)

$$s_{\mathbf{k}\sigma}^*\cong c_{\mathbf{k}\sigma}^{d*}+\sum_{\mathbf{k}',\mathbf{q},\sigma'}\sum_{e,f}\varphi_{\mathbf{k}\sigma}(\mathbf{q},e,f)\times c_{\mathbf{k}+\mathbf{q},\sigma'}^{d*}c_{\mathbf{k}',\sigma'}^{f}c_{\mathbf{k}'-\mathbf{q},\sigma}^{e*}+\text{ordinary polarization}+\cdots. \tag{2}$$

Here the c^*'s are the creation operators for a complete orthonormal set of one-electron wave functions, which are chosen to be in some sense the best possible such set (presumably obtained by a Hartree-Fock method). One might define such a set precisely by the criterion that in (2) only one one-electron term should appear. Here a c^d refers to a one-electron d-band function, a c^e to a function for a nominally empty band in the diamagnet, and a c^f to a nominally full band. In terms of this complete set the electron field operator $\psi_\sigma(\mathbf{r})$ may be written

$$\psi_\sigma(\mathbf{r})=\sum_{\mathbf{k}}\{\sum_{f}c_{\mathbf{k}\sigma}^{f}\varphi_{\mathbf{k}}^{f}(\mathbf{r})+c_{\mathbf{k}\sigma}^{d}\varphi_{\mathbf{k}}^{d}(\mathbf{r})+\sum_{e}c_{\mathbf{k}\sigma}^{e}\varphi_{\mathbf{k}}^{e}(\mathbf{r})\}, \tag{3}$$

---

[8] J. H. Van Vleck, J. Chem. Phys. **3**, 807 (1935).

[9] K. W. H. Stevens, Proc. Roy. Soc. (London) **A219**, 542 (1953); J. Owen, Proc. Roy. Soc. (London) **A227**, 183 (1954).

[10] P. W. Anderson and H. Hasegawa, Phys. Rev. **100**, 675 (1955).

[11] P. Jordan and E. Wigner, Z. Physik **47**, 631 (1928); for a clearer treatment and applications to such problems as the present one see W. Heisenberg, Ann. Physik **10**, 888 (1931); V. Fock, Z. Physik **75**, 622 (1932). A brief review for those somewhat unfamiliar: $c_{\mathbf{k}}^*$ is an operator which is thought of as creating a single electron in the one-electron orbital $\psi_{\mathbf{k}}(r)$, while its Hermitian conjugate $c_{\mathbf{k}}$ destroys this electron. The operator $n_{\mathbf{k}}=c_{\mathbf{k}}^*c_{\mathbf{k}}$ counts the electrons in state k in the sense that a state with exactly n electrons in k diagonalizes it with eigenvalue n. The Pauli principle and Fermi statistics are insured by the anticommutation relations,

$$c_{\mathbf{k}}^*c_{\mathbf{k}'}+c_{\mathbf{k}'}c_{\mathbf{k}}^*=\delta_{\mathbf{k}\mathbf{k}'},$$
$$c_{\mathbf{k}}c_{\mathbf{k}'}+c_{\mathbf{k}'}c_{\mathbf{k}}=0,$$

the $\mathbf{k}=\mathbf{k}'$ relation limiting n to 0 or 1 and the rest enforcing antisymmetry. The assumption we make is that the "clothed," physical quasi-particle operators like the $s_{\mathbf{k}}^*$, which create and destroy the electron together with its accompanying polarization, also exist and obey the same relations. This assumption is, though unproved except in perturbation theory, basic both to usual solid-state theory and quantum field theory. The field operator $\psi^*(\mathbf{r})$ introduced in (3) creates an electron at point r rather than in state $\mathbf{k}$; clearly the relationship is the linear transformation (3). Even this much knowledge of second quantization will not be necessary to understand the physical ideas in this paper.

where the $\varphi(\mathbf{r})$ are the complete orthornormal set of space functions. The second term in (2) is the spin-polarization accompanying the d-electron, and as we indicated the true quasi-particle operator contains also polarization terms (which look like spin-polarization but the operator combination is $c_{\sigma'}{}^{e*}c_{\sigma'}{}^{f}c_{\sigma}{}^{d*}$) and also higher many-electron terms.

The most important property of the s's defined in (2) as single quasi-particles is their kinetic energy

$$E_K=\sum_{\mathbf{k},\sigma}\epsilon(\mathbf{k})s_{\mathbf{k}\sigma}{}^{*}s_{\mathbf{k}\sigma}; \tag{4}$$

again we understand that the sum really contains five bands.

If $\epsilon(\mathbf{k})$ were a constant, an equally satisfactory set of one-electron starting functions would be the localized Wannier functions formed from the s's,

$$s^{*}(\mathbf{R},\sigma)=N^{-\frac{1}{2}}\sum_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{R}}s_{\mathbf{k},\sigma}{}^{*}. \tag{5}$$

Even when the variation of ϵ with $\mathbf{k}$ is simply small relative to other energies in the problem—as it will turn out to be—the localized function represents a good starting point, and the variation of ϵ may often be treated as a perturbation. This variation is small because it does not come from the direct transfer of an electron from the diamagnetic ion to the paramagnetic one, but from the higher order effect, transfer all the way from cation to cation. Here we see one of the advantages of the quasi-spin concept: in starting with these localized functions we have already included all the interactions of the d-electrons with the diamagnetic ions, which may be large; we use as a perturbation only the overlap from one metal ion to another, which is much smaller.

A second useful feature of the localized spin quasi-particles is that, neglecting $\epsilon(\mathbf{k})$ in the large energy denominators involved in the many-electron parts, the expressions for such things as the spin-polarization greatly simplify, as is shown in the Appendix, and take on the form one expects for truly localized functions. This form is almost identical with the corresponding spin-polarization around a nucleus which causes nuclear indirect exchange.[12]

To get a full set of five localized functions the sums in (5) must actually be confined to any appropriate set of five bands in turn. The resulting five localized functions form a reducible representation of the ion's point group and so, in for example the cubic case, can be recombined to form the usual irreducible sets: t, of symmetry like (xy,yz,zx) and $e_g(y^2-x^2, 2z^2-x^2-y^2)$. We shall introduce a new index n and call this set of localized quasi-particle operators $s_n(\mathbf{R},\sigma)$. Explicitly, we can write the running-wave quasi-spin in the jth branch of the d-band in terms of the s_n as

$$s_{\mathbf{k},\sigma}{}^{j*}=N^{-\frac{1}{2}}\sum_{n}a_{nj}\sum_{\mathbf{R}}e^{-i\mathbf{k}\cdot\mathbf{R}}s_n{}^{*}(\mathbf{R},\sigma),$$

and this, inserted in the kinetic energy expression (4), gives the energy in terms of localized functions

$$E_K=\sum_{n,\sigma,\mathbf{R}}\epsilon_{cn}s_n{}^{*}(\mathbf{R},\sigma)s_n(\mathbf{R},\sigma)+\sum_{n,n',\mathbf{R}\neq\mathbf{R}',\sigma}b_{\mathbf{R}-\mathbf{R}'}{}^{nn'}s_n{}^{*}(\mathbf{R},\sigma)s_{n'}(\mathbf{R}',\sigma) \tag{6}$$

(by group theory the first term must clearly be diagonalized by the symmetric functions s_n).

The ϵ_{cn} are the energies in the crystal field, which may be somewhat larger than the constants b which play the part of one-electron "transfer integrals" for the localized functions. In principle both can be obtained by Fourier transformation from the a_{nj}'s and the $\epsilon(\mathbf{k})$, but in actual practice one will compute or measure them from properties of the localized functions. We shall regard the ϵ_c's and the b's, the latter of which are in practice appreciable only for near neighbors, as fundamental constants of the theory. The only purpose then of this whole excursion has been to make the following observation: *that the* $s_n{}^{*}(\mathbf{R},\sigma)\Psi_0$ *are a set of many-electron functions representing spins localized on cations, for which the only off-diagonal matrix elements of the full Hamiltonian are the transfer integrals* $b_{R-R'}$. Thus, aside from these transfer integrals, all the further properties of the spins must follow from true many-electron interaction effects. (We shall find that the b's are of the order of half an electron volt.)

III. INTERACTIONS OF QUASI-PARTICLES

Having set up the picture of individual quasi-spins, we shall now write down their interactions in the order of their numerical importance. The first few are simply the various electrostatic interactions of the purely one-electron parts, which may be written down by starting with the electrostatic energy

$$V=\frac{1}{2}\int d\mathbf{r}\int d\mathbf{r}'\sum_{\sigma,\sigma'}\psi_{\sigma'}{}^{*}(\mathbf{r}')\psi_{\sigma}{}^{*}(\mathbf{r})\times\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\psi_{\sigma}(\mathbf{r})\psi_{\sigma'}(\mathbf{r}'), \tag{7}$$

and expanding $\psi_\sigma(r)$ in terms of the localized functions analogously to (3),

$$\psi_\sigma(\mathbf{r})\cong\sum_{\mathbf{R},n}s_n(\mathbf{R},\sigma)f_n(\mathbf{r}-\mathbf{R})+\sum\nolimits_{\text{other bands}}, \tag{8}$$

$f_n(r)$ being the one-electron Wannier function as normally understood. Actually, (8) is just the first terms of a series which is to be understood as the inverse of (2), an expansion of ψ (or the c's) in powers of the many-electron operators s. Higher many-electron terms in the

---

[12] N. F. Ramsey, Phys. Rev. **91**, 303 (1953); M. A. Ruderman and C. Kittel, Phys. Rev. **96**, 99 (1954); N. Bloembergen and T. J. Rowland, Phys. Rev. **97**, 1679 (1955).

energy, however, are more easily written down by direct perturbation theory than by writing the ψ's in (7) and (8) in terms of many-electron functions.

Inserting (8) into (7), we get as the largest term the Coulomb repulsion of the electrons when they are on the same ion:

$$U_{\mathrm{rep}}=\tfrac{1}{2}\sum_{\mathbf{R},n,n',\sigma,\sigma'} U_{nn'}s_{n'}^{*}(\mathbf{R},\sigma')\times s_{n}^{*}(\mathbf{R},\sigma)s_{n}(\mathbf{R},\sigma)s_{n'}(\mathbf{R},\sigma'), \quad (9)$$

where

$$U_{nn'}=\int d\mathbf{r}\int d\mathbf{r}'|f_{n}(\mathbf{r})|^{2}|f_{n'}(\mathbf{r}')|^{2}e^{2}|\mathbf{r}-\mathbf{r}'|^{-1}. \quad (10)$$

The coefficients $U_{nn'}$ are of the order of 15 electron volts for free-ion functions, but are reduced in the solid by dielectric polarization. To a rough approximation the U's do not depend on n and n'. What dependence there is, is given by the so-called Slater integrals, and will be discussed shortly. For the present $U_{nn'}=U$. The effect of the large magnitude of U_{rep} is that where possible the electrons remain equally shared among the ions. It costs an energy

$$mU-(m-1)U=U, \quad (11)$$

to take an electron from an ion with the correct number, m, of electrons (and thus $m-1$ repulsive interactions with the given one) to one with $m+1$ electrons and thus m repulsive interactions. This in fact is how we estimate U.

Because $U_{nn'}$ is so much bigger than the transfer integrals $b_{\mathbf{R}-\mathbf{R}'}$, the zeroth approximation we use should diagonalize the energy U_{rep} rather than E_{K}: that is, should occupy the localized wave functions $s(\mathbf{R})$ with individual electrons rather than the running-wave functions $s_{\mathbf{k}}$. The lowest states will be those with some exact number m of quasi-spins on each ion; the perturbed virtual states, states with one or more electrons transferred to other ions. These lowest states form a zeroth order degenerate manifold because of the various possibilities of spin σ, and of orbit n.[13]

The next interaction in order of magnitude does not further split this degeneracy. This is the true Coulomb repulsion between electrons on neighboring ions:

$$\tfrac{1}{2}\sum_{R\neq R'}\sum_{n,n'} U(\mathbf{R}-\mathbf{R}')s_{n'}^{*}(\mathbf{R}',\sigma')\times s_{n}^{*}(\mathbf{R},\sigma)s_{n}(\mathbf{R},\sigma)s_{n'}(\mathbf{R}',\sigma'), \quad (12)$$

$$U(\mathbf{R}-\mathbf{R}')=\int\int|f_{n}(\mathbf{r}-\mathbf{R})|^{2}|f_{n'}(\mathbf{r}'-\mathbf{R}')|^{2}\times e^{2}|\mathbf{r}-\mathbf{r}'|^{-1}d\mathbf{r}d\mathbf{r}'. \quad (13)$$

Equation (13) will obviously not depend much on n and n'. Since in all zeroth order states the occupation is uniform, (12) has no effect except to reduce the cost of transferring an electron from $\mathbf{R}$ to $\mathbf{R}'$. Two electrons at a typical ionic distance in oxides, 4 A, repel each other by perhaps 6 electron volts; subtracting this from the isolated-ion 15 electron volts, an upper limit on this cost is about 10 electron volts, further to be reduced by dielectric polarization.[14] (Note that the observed activation energy for d-band conduction will be even smaller because of ionic polarization; this might be 2–4 electron volts, while the effective $U_{nn'}$ might be 5–10 electron volts.)

Next comes the true exchange interaction within the individual cations. This is the Coulomb repulsion of the overlap charge:

$$V_{\mathrm{ex}}{}^{\mathrm{in}}=\tfrac{1}{2}\sum_{\mathbf{R},n\neq n',\sigma,\sigma'} J_{nn'}s_{n'}^{*}(\mathbf{R},\sigma')\times s_{n}^{*}(\mathbf{R},\sigma)s_{n'}(\mathbf{R},\sigma)s_{n}(\mathbf{R},\sigma'), \quad (14)$$

where

$$J_{nn'}=\int d\mathbf{r}\int d\mathbf{r}'\, f_{n}^{*}(\mathbf{r})f_{n'}(\mathbf{r})\times f_{n'}^{*}(\mathbf{r}')f_{n}(\mathbf{r}')e^{2}|\mathbf{r}-\mathbf{r}'|^{-1} \quad (15)$$

is the usual exchange integral. Note that if we confine ourselves to a subspace in which

$$s_{n}^{*}(+)s_{n}(+)+s_{n}^{*}(-)s_{n}(-)\equiv 1,$$
$$n_{n}(+)+n_{n}(-)\equiv 1$$

(i.e., orbital state n is certainly occupied), then

$$\begin{aligned} n_{n}(+)-n_{n}(-)&=2S_{z}{}^{n},\\ s_{n}^{*}(+)s_{n}(-)&=S_{+}{}^{n}=S_{x}{}^{n}+iS_{y}{}^{n},\\ s_{n}^{*}(-)s_{n}(+)&=S_{-}{}^{n}=S_{x}{}^{n}-iS_{y}{}^{n}, \end{aligned} \quad (16)$$

where the capital S^{n}'s are the spin vector associated with state n. Using these relationships, we arrive at the well-known identity between (14) and a spin interaction:

$$V_{\mathrm{ex}}{}^{\mathrm{in}}=-\tfrac{1}{2}\sum_{\mathbf{R},n\neq n'} J_{nn'}(\tfrac{1}{2}+2\mathbf{S}^{n}\cdot\mathbf{S}^{n'}). \quad (17)$$

This interaction will remove the degeneracy of relative spin orientations within the single cations, according to the usual Hund's rule reasoning (see reference 13).

So far no terms have affected the degeneracy of relative spin orientation of spins on neighboring ions; now we shall bring forth the largest terms which do.

By far the largest, when it occurs, is a spin-dependent force resulting from the terms already presented. The mutual repulsion $U_{nn'}$ of electrons on the same ion prevents the permanent occupation of ionized states, and thus makes the d-band substances insulators; but

---

[13] In the actual many-electron ions the $U_{nn'}$'s of course differ slightly from one n, n' to another, by the so-called Slater integrals. These, the $\epsilon_{c}(n)$, and the shortly-to-be described Hund's rule internal exchange integrals give us a single-ion problem whose diagonalization is the main province of ligand field theory. (See reference 7.) The result, however, is, except for the cases of $V^{3+}(d^{2})$ and $Co^{2+}(d^{7})$, rather simply described as an unambiguous occupation of certain t and e orbitals. We shall return to these "single-ion" complications in the experimental section.

[14] There will also be an additional reduction, similar to the reduction of 10–20% in the Slater integrals and spin-orbit coupling in ligand field theory, (see reference 8) caused by the reduction in the amplitude of the wave function on the central ion due to spreading to the ligands.

by virtually occupying, to a small extent, ionized states which are connected to the ground state by the transfer integrals b, the system can gain a certain amount of energy. However, only ionized states in which the electron transfers to an empty state n, σ can contribute; since the transfer integrals b carry the electron without change of spin, this means that the energy is gained only in the presence of an antiparallel neighbor, which means an antiferromagnetic effect. Clearly this part of our theory is a generalization of the "alternant orbitals" idea[15] whereby the wave functions for opposite spins need not be orthogonal to each other and can spread out on to each other's respective sites. What is new is that we have extended the idea to much more general situations and shown that it is closely connected with the insulating property of the transition-metal salts.

First we shall write this interaction down for the simple model of one orbital per spin which has been the province of most previous discussions of superexchange[2,5]; then we shall write down the more general expression.

In such a simple model all the degenerate states in the ground-state manifold have exactly one electron per ion, while all the excited states with one transferred electron have energy U. Between any pair of ions at a distance $\mathbf{R}-\mathbf{R}'$ there is only one $b_{\mathbf{R}-\mathbf{R}'}$; this must act twice to return the state to one of the ground manifold. Thus the second-order perturbation of the energy is

$$\Delta E=-\sum_{\mathbf{R},\mathbf{R}',\sigma,\sigma'}\frac{b_{\mathbf{R}-\mathbf{R}'}{}^2}{U}s^*(\mathbf{R},\sigma)\times s(\mathbf{R}',\sigma)s^*(\mathbf{R}',\sigma')s(\mathbf{R},\sigma'). \qquad (18)$$

The combination of quasi-particle operators is much simplified when we realize that in the ground states one and only one orbital per ion is occupied, so that the identities (16) hold. Then

$$\begin{aligned}s^*(\mathbf{R},+)s(\mathbf{R}',+)s^*(\mathbf{R}',+)s(\mathbf{R},+)&\\+s^*(\mathbf{R},-)s(\mathbf{R}',-)s^*(\mathbf{R}',-)s(\mathbf{R},-)&\\=\tfrac{1}{2}\{n_R(+)[n_{\mathbf{R}'}(-)+1-n_{\mathbf{R}'}(+)]&\\+n_{\mathbf{R}}(-)[n_{\mathbf{R}'}(+)+1-n_{\mathbf{R}'}(-)]\}&\\=\tfrac{1}{2}\{1-[n_{\mathbf{R}}(+)-n_{\mathbf{R}}(-)][n_{\mathbf{R}'}(+)-n_{\mathbf{R}'}(-)]\}&\\=\tfrac{1}{2}-2S_z(\mathbf{R})S_z(\mathbf{R}').&\qquad (19)\end{aligned}$$

[Note that here the constant term has opposite sign to that in (17). This is a real difference and may be observable.] Also,

$$\begin{aligned}s^*(\mathbf{R},+)s(\mathbf{R}',+)s^*(\mathbf{R}',-)s(\mathbf{R},-)&\\+s^*(\mathbf{R},-)s(\mathbf{R}',-)s^*(\mathbf{R}',+)s(\mathbf{R},+)&\\=-[S_x(\mathbf{R})+iS_y(\mathbf{R})][S_x(\mathbf{R}')-iS_y(\mathbf{R}')]&\\-[S_x(\mathbf{R})-iS_y(\mathbf{R})][S_x(\mathbf{R}')+iS_y(\mathbf{R}')]&\\=-2S_x(\mathbf{R})S_x(\mathbf{R}')-2S_y(\mathbf{R})S_y(\mathbf{R}'),&\qquad (20)\end{aligned}$$

so that

$$\Delta E=\text{const}+\sum_{\mathbf{R},\mathbf{R}'}\frac{2|b_{\mathbf{R}-\mathbf{R}'}|^2}{U}\mathbf{S}_{\mathbf{R}}\cdot\mathbf{S}_{\mathbf{R}'}. \qquad (21)$$

This is the antiferromagnetic exchange effect. Our estimates would put the effective exchange integral at 1/50–1/10 electron volts, or 200 to 1000 degrees. As we shall see, all other effects are orders of magnitude smaller, so that it is not surprising that this has been the outstanding effect experimentally.

The above mechanism leads to a nonvanishing result only when $\mathbf{R}$, σ' and $\mathbf{R}'$, σ are occupied states; i.e., when the d-states from and to which transfer takes place are half full. This corresponds to the "both more than half full" case of Anderson[2] and in fact that reference contains, without understanding it, a clumsy but in principle correct calculation, as to a greater or lesser extent do later papers.[5]

In the more complicated 5-orbital d-band, there will be many exchange integrals of this same kind, between each pair of occupied orbitals. There will in addition be off-diagonal exchange terms like

$$\sum_{\sigma,\sigma'}\frac{b_{\mathbf{R}-\mathbf{R}'}{}^{nn'}b_{\mathbf{R}-\mathbf{R}'}{}^{nn''}}{U}s_n{}^*(\mathbf{R},\sigma)\times s_{n'}(\mathbf{R}',\sigma)s_{n'}{}^*(\mathbf{R}',\sigma)s_{n''}(\mathbf{R},\sigma), \qquad (22)$$

and more complicated types. These off-diagonal terms will usually connect to states which are higher than the ground states by crystal field energies, although in the anomalous cases like Co^{++} they can play some role.

We shall discuss the numerical aspect of the above superexchange mechanism in the next part of the paper; now we go on to other effects. A third-order perturbation effect seems to be one of the major contributions when one of the states to which transfer takes place is completely empty. This is the third-order effect of transfer together with the internal exchange effect (14). Suppose that state n on $\mathbf{R}$ is half full, state n' on $\mathbf{R}'$ is completely empty (presumably because of the crystal field effects), and $b_{\mathbf{R}-\mathbf{R}'}{}^{nn'}$ is fairly large. Also suppose that there is a state n'' which *is* partially occupied on atom $\mathbf{R}'$. Then there is a third-order interaction

$$\begin{aligned}\Delta E_3=\sum_{\sigma,\sigma',\sigma'',\sigma'''}\sum_{\mathbf{R},\mathbf{R}'}\frac{|b_{\mathbf{R}-\mathbf{R}'}{}^{nn'}|^2J_{n'n''}}{U^2}&\\\times s_n{}^*(\mathbf{R},\sigma''')s_{n'}(\mathbf{R}',\sigma''')s_{n''}{}^*(\mathbf{R}',\sigma')s_{n'}{}^*(\mathbf{R}',\sigma'')&\\\times s_{n''}(\mathbf{R}',\sigma'')s_{n'}(\mathbf{R}',\sigma')s_{n'}{}^*(\mathbf{R}',\sigma)s_n(\mathbf{R},\sigma).&\end{aligned}$$

This interaction operating on a state with n' completely unoccupied vanishes unless $\sigma'=\sigma$ and unless $\sigma'''=\sigma''$, (that is, J operates only on the transferred electron n', σ) and since state n' is empty the combination $s_{n'}(\mathbf{R}',\sigma)s_{n'}{}^*(\mathbf{R}',\sigma)$ is just unity. Inserting these two

---

[15] P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955).

facts, we get

$$\Delta E_3=\sum_{\sigma,\sigma'}\frac{|b_{\mathbf{R}-\mathbf{R}'}{}^{nn'}|^2 J_{n'n''}}{U^2}s_{n''}{}^*(\mathbf{R}',\sigma)\times s_n{}^*(\mathbf{R},\sigma')s_{n''}(\mathbf{R}',\sigma')s_n(\mathbf{R},\sigma). \quad (23)$$

This is then like a positive, ferromagnetic exchange between n, $\mathbf{R}$ and n'', $\mathbf{R}'$ [we could again show that this is equivalent to $\mathbf{S}_n(\mathbf{R})\cdot\mathbf{S}_{n''}(\mathbf{R}')$]. This is the "less than half full" interaction of previous work[2]; we see it is less by the order of $J_{n'n''}/U$ or about a factor 5–10 than the antiferromagnetic effect, so will only appear when that is absent and under conditions where b is large or U small.

The next effect is the true direct exchange, which comes by taking the Coulomb repulsion of the overlap charge:

$$E_{\text{ex}}=\tfrac{1}{2}\sum_{\mathbf{R},\mathbf{R}',n,n'}\sum_{\sigma,\sigma'}J_{\mathbf{R}-\mathbf{R}'}{}^{nn'}s_{n'}{}^*(\mathbf{R}',\sigma')\times s_n{}^*(\mathbf{R},\sigma)s_{n'}(\mathbf{R}',\sigma)s_n(\mathbf{R},\sigma'), \quad (24)$$

where

$$J_{\mathbf{R}-\mathbf{R}'}{}^{nn'}=\int d\mathbf{r}\int d\mathbf{r}'\, f_n{}^*(\mathbf{r}-\mathbf{R})f_{n'}(\mathbf{r}-\mathbf{R}')\times f_{n'}{}^*(\mathbf{r}'-\mathbf{R}')f_n(\mathbf{r}'-\mathbf{R})e^2|\mathbf{r}-\mathbf{r}'|^{-1}. \quad (25)$$

This is, by an obvious general theorem, always positive (ferromagnetic). We can understand why it is usually less than superexchange by an order-of-magnitude argument. All of these interactions will be estimated by assuming that the overlap of neighboring quasi-particle wave functions really takes place entirely on the intervening diamagnetic ion. Suppose that the amount of overlap is ϵ: i.e.,

$$f_n \cong \psi_d - \epsilon\psi_p(\mathrm{O}^{--}); \quad (26)$$

Then each of the f's in (25) must be of order ϵ because this is the amplitude in the overlap region. Aside from this factor ϵ^4, (25) is just an atomic exchange integral J_{at} on the O^{--} ion, of the order of $\frac{1}{10}$ ry.

The expression $b_{\mathbf{R}-\mathbf{R}'}{}^{nn'}$ contains ϵ^2 and is, aside from that, a one-electron term value T, or energy of interaction between the O^{--} electron and the one-electron potential; as a general rule term values are ~ 1 ry, and are thus an order of magnitude bigger than exchange integrals. U itself is a repulsive Coulomb interaction between single electrons, which is usually intermediate in magnitude between term values and exchange integrals. Thus

$$b^2U^{-1} \simeq \epsilon^4T^2U^{-1} \gg \epsilon^4 J_{\text{at}},$$

the ratio being slightly more than an order of magnitude. On the other hand, the ferromagnetic effect (23) is b^2U^{-1} multiplied by a factor $J_{\text{at}}U^{-1}$, and so may be expected to be at most only slightly larger than direct exchange; both are, however, ferromagnetic so they add. Actually (23) is just the direct exchange with the transferred part of the wave function, so (23) and (24) are two halves of the same effect, direct exchange with the "alternant orbital," nonorthogonally overlapping, function.

We use, here and later, the fact that true ferromagnetic exchange integrals are always small. This is because they are the self-energy of an overlap charge which, because of orthogonality, must be as often positive as negative, and almost always has a node where the wave function itself is largest.

The last effect we take up here is the third in this group of interactions which are only an order or so smaller than true superexchange. Undoubtedly many higher-order interaction effects exist, coupling groups of three spins, etc.; but these will be clearly smaller than the smallest of the effects we consider. This third effect is essentially exchange between one spin and the spin polarization of its neighbor, and is the complete analog of the Ramsey-Kittel-Ruderman-Bloembergen[11] mechanism of indirect exchange interaction between nuclear spins. So far our interactions have really been electron-pair effects depending on the direct action of the exclusion principle: superexchange because the antiparallel spins can transfer over on to each other's ions, while the other two are direct exchange, in the first case of the transferred "superexchange" charge, in the second of the actual overlap of the wave functions (clearly effects of closely the same order of magnitude). In all three cases the sign is unequivocal.

The true indirect exchange interaction is probably best written down directly as an energy perturbation, starting from the one-electron approximation. In the Appendix we show that, in the approximation that the quasi-spins are fairly well localized, the effect of the part of the exchange which causes spin-polarization on a localized spin is

$$Vc_n{}^{d*}(\mathbf{R},\sigma)=\sum_{\mathbf{k}',\kappa,\sigma',e,f}J_n(e,f;\boldsymbol{\kappa},\mathbf{k}')\exp(i\boldsymbol{\kappa}\cdot\mathbf{R})\times c_n{}^{d*}(\mathbf{R},\sigma')c_{\mathbf{k}',\sigma'}{}^f c_{\mathbf{k}'-\kappa,\sigma}{}^{e*}, \quad (27)$$

where

$$J_n(e,f;\boldsymbol{\kappa},\mathbf{k}')=\int d\mathbf{r}d\mathbf{r}'\, e^2|\mathbf{r}-\mathbf{r}'|^{-1}\times f_n{}^*(\mathbf{r}')f_n(\mathbf{r})\varphi_{\mathbf{k}'}{}^f(\mathbf{r}')\varphi_{\mathbf{k}'-\kappa}{}^{*e}(\mathbf{r}). \quad (28)$$

The second-order perturbation comes from applying (27) twice, the second time returning e, $\mathbf{k}+\boldsymbol{\kappa}$ and f, $\mathbf{k}'$ to their proper places (thus maintaining the identity of σ and σ'). The result is

$$E_2=-\sum_{n,\mathbf{R},n',\mathbf{R}';\sigma,\sigma'}c_{n'}{}^{d*}(\mathbf{R}',\sigma')c_{n'}{}^d(\mathbf{R}',\sigma)c_n{}^{d*}(\mathbf{R},\sigma)\times c_n{}^d(\mathbf{R},\sigma')\sum_{\mathbf{k},\mathbf{k}',e,f}\exp[i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{R}-\mathbf{R}')]\times\frac{|J_n(ef;\mathbf{k}-\mathbf{k}';\mathbf{k}')|^2}{\epsilon_e(\mathbf{k})-\epsilon_f(\mathbf{k}')}. \quad (29)$$

The sum over $\mathbf{k}$ and $\mathbf{k}'$ in (29) is rather hard to estimate, but we can get a rough idea of it in the following way. Except for the rather small variation of J with $\mathbf{k}$ and $\mathbf{k}'$, the sum will vanish unless both $\epsilon_e(\mathbf{k})$ and $\epsilon_f(\mathbf{k})$ are functions of $\mathbf{k}$. This reflects the fact that both the electron and the hole must actually travel from $\mathbf{R}$ to $\mathbf{R}'$ in order to recombine there. Thus, taking into account that the electron and hole are probably in fairly deep, narrow bands, we expand

$$\begin{aligned}\epsilon_e(\mathbf{k})&=E_e+\sum_\lambda \exp(i\mathbf{k}\cdot\boldsymbol{\lambda})\beta_\lambda+\cdots,\\ \epsilon_f(\mathbf{k})&=E_f+\sum_\lambda \exp(i\mathbf{k}\cdot\boldsymbol{\lambda})\beta_\lambda'+\cdots,\end{aligned}\tag{30}$$

and expanding the denominator in (29) under the assumption that β_λ and β_λ' are small, we get

$$E_2\sim\frac{|J_n(ef)|^2}{E_e-E_f}\,\frac{\beta_{\mathbf{R}-\mathbf{R}'}\beta_{\mathbf{R}-\mathbf{R}'}'}{(E_e-E_f)^2},$$

where $J_n(ef)$ is an exchange integral with the Wannier functions of bands e and f:

$$J_n=\int d\mathbf{r}d\mathbf{r}'\,f_n^*(\mathbf{r})f_e(\mathbf{r})\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}f_f^*(\mathbf{r}')f_n(\mathbf{r}').\tag{31}$$

f_f we know, since the only core bands which could propagate effectively are the s and p bands on the anions. f_e is probably some combination involving s or p functions on the cation; so J_n will be of order ϵ (for the overlap with f_f) times an atomic exchange integral.

The number β_λ is probably not too small, $\sim\frac{1}{4}(E_e-E_f)$, since the e band may be fairly wide; but β_λ' may well be as little as 1/10 to 1/100 of E_e-E_f, which is of the order 20 electron volts or so. An estimate is $\beta_\lambda/(E_e-E_f)\sim\epsilon$. Thus we get a result

$$\tfrac{1}{4}\left(\frac{\epsilon^3 J_{\mathrm{at}}^2}{E_e-E_f}\right),$$

or slightly smaller than direct exchange. Note, however, that this definitely falls off more slowly with ϵ so that very distant ions or very inner shells will tend to retain this mechanism rather than true superexchange.

It is not clear what the sign of this mechanism will be. Physically, we expect the spin-polarization to draw parallel electrons toward the d shell (being the result of true exchange), so to tend to take the parallel electrons from sp shells of the anions and put them in higher shells of the cation. This leaves antiparallel spin-polarization near the neighbor cation so we expect an antiferromagnetic sign. Note also that the nuclear magnetic resonance shift at the anion will be diamagnetic. Such shifts are indeed observed in $\mathrm{Gd(H_2O)_6}^{+++}$ and in $\mathrm{CrF_3}$,[16] but two orders of magnitude weaker than the usual paramagnetic shift in the iron group, just as we would expect from our estimates. On the other hand, the coefficient in (29) depends in a complicated way on relative phases of wave functions, etc., and actually has been shown to be ferromagnetic in some metal and semiconductor cases.[11] Thus we do not feel that the sign is at all predictable without a more detailed discussion and a clear knowledge of the states e and f.

This is the last interaction we shall discuss. There are obviously many higher terms, but they will all be smaller by one or more orders of magnitude; while there will be essentially no cases in which all four of these interactions are particularly small. That is, superexchange dominates in fairly concentrated or fairly covalent cases except when transfer occurs into empty shells, in which case direct exchange dominates. Dilute or noncovalent cases will always be dominated by the lowest order indirect exchange.[17]

IV. BRIEF DISCUSSION OF THE COMPARISON WITH EXPERIMENT

The two parameters which enter the superexchange energy are the repulsive energy U and the transfer integral b. One could get at U either by knowing the observed activation energy for conduction in the d-band, and adding to that the ionic polarization energy; or by starting from U for a free ion and subtracting the nearest-neighbor repulsion and other corrections. Both of the quantities in the first method are hard to estimate, and in particular experimental values of the d-band energy are few and questionable. As a guess by the first method, the activation energy is 2–3 electron volts, and the ionic polarization energy[18]

$$2\left(\frac{e^2}{K_{\mathrm{el}}\cdot r_0}-\frac{e^2}{K_{\mathrm{tot}}r_0}\right),$$

where K is the dielectric constant, and where r_0 is an effective radius of the ion itself (say of order 1–1.5 A). Taking a value of K_{tot} of 5, this is about 3–4 electron volts, giving an estimate of 6 electron volts, with no particular variation from cation to cation.

More reliable perhaps, especially for comparisons, is to take the values of U for free ions from tables,[19] reduce by a 4-ev nearest-neighbor correction and by perhaps 4 ev for electronic polarization, take off another 10%

---

[16] R. G. Shulman (to be published); Knox, Shulman, and Waite (to be published).

[17] This clear separation suggests a nomenclature which we adopt here and suggest for the future: namely, the "alternant orbital" mechanism = superexchange, the next two = direct exchange, and the perturbation term = indirect exchange (since it is the term analogous to nuclear spin indirect exchange).

[18] Rittner, Hutner, and Dupre, J. Chem. Phys. 17, 198, 204 (1949) discuss the polarization energy in detail, but more exact values than the above "cavity" estimate hardly seem justified.

[19] Charlotte E. Moore, *Atomic Energy Levels*, National Bureau of Standards, Circular No. 467 (U. S. Government Printing Office, Washington, D. C., 1949 and 1952).

TABLE I. U-value estimates, in electron volts.

	Third ionization potential	Second ionization potential	Free-ion U	Corrections	U
Mn^{++}	33.7	13.8	19.9	10.0	9.9
Fe^{++}	30.6	16.0	14.6	9.5	5.1
Co^{++}	33.5	17.0	16.4	9.6	5.8
Ni^{++}	35.2	18.2	17.0	9.7	6.3
Cu^{++}	36.8	20.3	16.5	9.6	5.9
Fe^{+++}	~53.5	30.6	~23	~13	~10

for the covalency correction, and arrive at the values in Table I. Fe^{+++} is only roughly estimated; also a 20% covalency correction for this is used. It is striking that only Mn^{++} and Fe^{+++} show any appreciable variation from constancy; a value of 6 ev, with perhaps 9–10 for the half-filled shells, would make good sense. There is no reason to expect the identity of the anion to change these numbers much.

We may guess that the numbers in Table I are not in error by much more than one electron volt, i.e., $\sim\pm 20\%$. This accuracy comes from the fact that between the experimental value of around 2–3 ev,[20] and the ionic value (less 6 ev for the proximity correction and 10% for covalency) the difference is accounted for entirely by electronic and ionic polarization energy of the surroundings. We cannot change the distribution between these by very much (say at most an electron volt, or 50%); hence our 1-ev confidence in Table I.

In estimating b we are helped by a very fortunate circumstance: that b is closely related to the crystal field parameter Dq. $10Dq$ is the difference in energy of a t from an e orbital, and is in general made up of two contributions[7]: a "covalency" contribution coming from the overlap effects, because the e orbitals point towards the octahedral ligands, the t's away; and a purely electrostatic repulsion. Thus Dq (aside from this electrostatic contribution) is a measure of the extent to which the d-function has spread onto the anions.

One way to see how this "covalency" contribution is related to b is to look at the three-center problem of two d orbitals on opposite sides of an anion. b is essentially the difference in energy between the states in which ψ has the same sign on each cation and that antisymmetric about the anion. The former, by symmetry, can contain no anion p-function, the latter no s-function. Thus this difference is the same as the effect of removing all p-function admixture from the wave function (assuming the s admixture relatively small). But that is also at least an important contribution to the crystal field.

A less rough estimate comes from perturbation theory. Starting with an e orbital at energy 0, a p orbital at energy $-E$, and a (now O^{--}-cation, not cation-cation) transfer integral b', the energy perturbation of the e orbital is $+b'^2/E$. On the other hand, the effective transfer integral to an equivalent e orbital is also b'^2/E.

TABLE II. Transfer integrals.

Orbit on first atom	Direction to first	Orbit on second	Direction to second	b
$e_{x^2-y^2}$	x, y	$e_{x^2-y^2}$	x, y	$(5/2)Dq$
$e_{x^2-y^2}$	z	$e_{x^2-y^2}, e_{z^2}$	x, y, z	0
$e_{x^2-y^2}$	x, y	e_{z^2}	x, y	$(5/2\sqrt{3})Dq$
$e_{x^2-y^2}$	x, y	e_{z^2}	z	$(5/\sqrt{3})Dq$
e_{z^2}	x, y	e_{z^2}	x, y	$(5/6)Dq$
e_{z^2}	x, y	e_{z^2}	z	$(5/3)Dq$
e_{z^2}	z	e_{z^2}	z	$(10/3)Dq$

Under the assumption that the overlap of a t orbital is relatively negligible, this perturbation is the covalent contribution from this anion to $10Dq$. To get the precise relationship, define $e_{x^2-y^2}$ to be the orbital of x^2-y^2 symmetry and e_{z^2} that of $2z^2-x^2-y^2$. The normalized relative amplitudes in the various directions are as follows:

	x	y	z
$e_{x^2-y^2}$	$\frac{1}{2}$	$\frac{1}{2}$	0
e_{z^2}	$\frac{1}{2\sqrt{3}}$	$\frac{1}{2\sqrt{3}}$	$\frac{1}{\sqrt{3}}$

and $10Dq$ is made up of the sum of the contributions from all the anions, i.e., it is equal to the contribution which would result from unity wave function amplitude.

Using this, we can find the transfer integrals b in terms of the covalent part of Dq for all the possible combinations of orbitals. This is shown in Table II.

Thus we may work out the various exchange integrals given the covalent contribution to Dq, using the formula

$$J_{\mathrm{eff}}=2b^2U^{-1}. \tag{32}$$

From Table II it is easily verified that the total interaction in any direction between two half filled e_g-shells (two electrons) is equivalent to exchange between two electrons with an exchange integral

$$J_{\mathrm{eff}}(\text{filled } e_g)=2[(10/3)Dq]^2U^{-1}. \tag{33}$$

Thus, in a substance with a shell of n d-electrons, the exchange integral for the total spin $S=n/2$ is

$$J=2n^{-2}[(10/3)Dq]^2U^{-1}. \tag{34}$$

The simplest substance for a comparison with experiment is NiO. Here we have just the half-filled shell contributing, with no effect of the t-electrons and therefore a minimum of nearest-neighbor interaction. According to Van Vleck,[21] the Curie temperature is given in molecular field theory by

$$T_N=2JzS(S+1)/3k. \tag{35}$$

The crystal field parameters for O^{--} octahedra are about the same as for H_2O octahedra,[20] and in both cases are probably mostly covalent.[7] Thus we can use

---

[20] F. J. Morin, Bell System Tech. J. **37**, 1047 (1958).

[21] J. H. Van Vleck, J. Chem. Phys. **9**, 85 (1941).

TABLE III. Exchange integrals and comparison with experiment.

Cation	$J_{\text{eff calc}}$	$T_{N\,\text{calc}}$	$T_{N\,\text{renorm}}$ to NiO	T_N (oxide) obs	θ obs	$T_N(M^{++}F_2)$	$\theta(M^{++}F_2)$	$T_N(\text{La}M^{+++}O_3)$	$T_N(M^{+++}F_3)$
Cu^{++}	585–290°	1170–585°	(675–340°)	453? 220?					
Ni^{++}	460	920	(520)	520		73.2	~100		
Co^{++}	525	879	495	290(326)[a]	280(530)[a]	37	50		
Fe^{++}	620	930	525	185(194)[a]	? (545)[a]	78.3	117		
Mn^{++}	240	340	190	117	610	66.5	97		
Fe^{+++}	750	1050						750	394
Co^{+++}	2200								460
Mn^{+++}								100	43
Cr^{+++}								320	80

[a] See reference 24.

the H_2O value,

$$(Dq)_{\text{Ni}^{++}} = 850 \text{ cm}^{-1}.$$

Inserting this and U from Table I into (34), we get $J_{\text{eff}} = 460°$, $J = 115°$, and from (35) with $S = 1$, $(T_N)_{\text{NiO}} = 920°$. The observed value is[22] $T_N = 520°$, so that our calculation has come out 80% too high.

The molecular field theory is known to give too high a value for the Néel temperature relative to J by the order of 20–30%; the theory behind the Curie-Weiss constant is more accurate but θ has not been measured for NiO. Thus we consider our error to be about 40–50%.

The U value, as we have said, is probably fairly accurate. The b value, on the other hand, cannot be fixed with any exactness. The first and probably least serious error is the fact that we assume in comparing Dq and b that there is no overlap of the t (xy, yz, etc.) functions with the O^{--}, i.e., we assume no π bonding. The magnitude of actual t-shell exchange integrals suggests that this is perhaps $\frac{1}{3}$ to $\frac{1}{4}$ of the σ contribution, so that b may be *underestimated* by 25% or so because of this.

A source of error which should be about the same, and of opposite sign, is the neglect of the s-function contribution to both Dq and b. Since the s and $p\sigma$ functions change relative signs from one side to the other of the anion, s will add to Dq and subtract from b. We might think of the $p\sigma$ function as being slightly hybridized with s. Because of this compensation of these two errors, the main problem is the third one: the relative amounts of the electrostatic and covalent contributions to Dq.

From Table III, which contains calculations similar to those we made on NiO for other iron group cations, we will see that the calculated T_N is closest for the trivalent ion Fe^{+++}, next best (~ twice too high) for the divalent ions in oxides, and quite a bit too high for fluorides. Although fluorides have much the same Dq as oxides, they are generally expected to be much less covalent, while the triply-charged ion Fe^{+++} in oxides is the most covalent situation. Thus the major uncertainty in our calculations is in this question, and we really must consider our J's as upper limits until some method of more accurate estimation of b is worked out. In reference 7 will be found a discussion of the relative importance of covalent and electrostatic contributions, while in reference 9 we find actual measurements of the electron transfer—unfortunately not of the energy.

We do not expect, however, very much variation in these correction factors among doubly charged ions in O^{--}, so we can predict to a fair extent the T_N's which would be observed for the series MnO—CuO, normalizing to the observed NiO value, and neglecting other complicating effects which certainly exist. In Table III we give a list of actually calculated J's and T_N's for these substances as well as for a $Fe^{+++}-O-Fe^{+++}$ 180° bond, and of T_N's for the divalent oxides renormalized to make NiO fit. The J's in this table are the effective J's of (33) for a single pair of electrons equivalent to all the e interactions. Then the T_N's are obtained from this simply by the quantum correction $S(S+1)/S^2 = (S+1)/S$ (various factors z, $\frac{1}{3}$, etc., dropping out). The Dq values we use are listed in Table IV, and are the most recent, those given by Holmes and McClure.[23]

Table III illustrates that while many qualitative features are explained by our theory, the quantitative picture is as yet not clear and complete. Let us take up the various aspects in order:

(A) Oxides $M^{++}O^{--}$. We have included the three cases of Cu^{++}, Co^{++}, and Fe^{++} not because they are expected to agree quantitatively with theory but to show the general order-of-magnitude correlation with theory, particularly of the Curie-Weiss constant θ. In Cu^{++} the question is one of how the $e_{x^2-y^2}$ orbitals are aligned in the given crystal structure; depending on this, there is an uncertainty by a factor 2, as we indicate in the table.

In Co^{++} and Fe^{++} the problem is the effect of orbital degeneracy, as well as (in Co^{++}) the mixture of e and t occupancies. The problem of exchange in the presence of orbital degeneracy has been discussed by Kanamori,[24] who showed that θ, and to a lesser extent T_N, are depressed in Co^{++} by orbital degeneracy. The actual

---

[22] C. H. La Blanchetais, J. Phys. Radium **12**, 765 (1951).

[23] O. G. Holmes and D. S. McClure, J. Chem. Phys. **26**, 1686 (1957).

[24] J. Kanamori, Progr. Theoret. Phys. Kyoto **17**, 177 (1957).

values he gives are shown in parentheses. It is clear that a serious discrepancy exists only in Fe^{++}. However, it may be that there are some ferromagnetic components to the interactions in both Co^{++} and Fe^{++}, as well as the effect of nearest-neighbor exchange in the latter which we now discuss.

The reliable comparison is MnO *vs* NiO. This is not nearly as bad as it looks, because we expect[25] the nearest-neighbor interactions, large because of the large number of t electrons and π bonds, to depress the Curie point by an unknown factor which can easily be as large as 2. What is striking is the large reduction of T_N in MnO, as compared to the increasing amount of nearest-neighbor interaction in the series NiO$-$MnO, as indicated by the increasing $\theta - T_C$ ratio.[26]

(B) Lanthanates LaM^{+++}O$_3$. These we include mainly to show the excellent agreement for Fe^{3+}$-$O$-$Fe^{3+}. This is not unexpected because there is much less electrostatic correction to b here. Note also that in spite of the Dq's of Mn^{+++} and Cr^{+++} being very much larger even than Fe^{+++}, the t interactions (π bonds) operative here are much weaker. This is our first encounter with the sets of rules suggested by Kanamori and Nagamiya,[27] Wollan, Child, Koehler, and Wilkinson,[28] and by Goodenough[29] on semiempirical grounds. These rules are precisely equivalent to the predictions one would make from the present ideas about superexchange (or actually, even from the primitive notions of reference 2). The rules are (together with the explanation by these ideas):

(a) When filled e orbitals overlap a given anion, the result is strong antiferromagnetism (ordinary superexchange between $e(\sigma)$ orbitals).

(b) When a filled e orbital and an empty one overlap opposite ends of an anion, the result is ferromagnetism (direct exchange between the e orbital and the t shell).

(c) When empty e's overlap an anion, the result is weaker antiferromagnetism (π-bond superexchange of the t shells). In rules (b) and (c) the weakness of the interaction is only relative, because normally these involve trivalent cations with large, covalent Dq's where the comparable e-shell superexchanges involve divalent ions.

(C) Difluorides. These repeat the general pattern of the NaCl structure oxides, except that there is no $\theta - T_c$ difference, and that because of the angle of the bond we expect $\pi - \sigma$ interaction (i.e., $t-e$) as well as $e-e$ and $t-t$.

Again, Co^{++} is small, presumably due to orbital effects, and Fe^{++} is not as large as expected. Mn^{++} does not drop severely, probably because of the $t-e$ interaction. All are much smaller than the oxides, reflecting most probably a much smaller covalency in spite of the Dq's being comparable to those in the oxides.

TABLE IV. Octahedral Dq values for iron group ions.

No. of d electrons	Ion	$10Dq$ (cm^{-1})	
1	Ti^{+++}	20 300	
2	V^{+++}	18 000	
3	Cr^{+++}	17 000	
	V^{++}	11 800	
4	Cr^{++}	13 900	
	Mn^{+++}	21 000	
5	Mn^{++}	7800	
	Fe^{+++}	13 800	
6	Fe^{++}	10 000	
	Co^{+++}	∼18 000	
7	Co^{++}	9800	
8	Ni^{++}	8500	
9	Cu^{++}	12 500	(uncertain: strong tetragonality)

(D) Trifluorides. Again we find the sharp drop below the half-filled shells. It is clear, from the fact that the Mn^{++}$-$Fe^{+++} ratio is not greater than in oxides (the trifluorides have 180° bonds) that the angular factor does not play a very large part in the difluorides. The Co^{3+} case is most interesting because it shows that the special properties of Fe^{3+} are indeed due to its being the only usual trivalent ion with a greater than half-filled shell. We predict actually an even larger value for CoF$_3$; perhaps there are some complications with low-spin states as in the lanthanates of Co^{3+}.[29]

V. CONCLUSION

The main purpose of this paper has been to show from an entirely theoretical point of view how the main motivation of superexchange must be the desire of the electrons to set their spins antiparallel so that they may spread out into not quite orthogonal, overlapping orbitals.

In doing this we find that the exchange effect is expressed in terms of two parameters: the repulsive energy of coincident d-electrons, U, which besides being an empirical parameter (d-band conduction; optical absorption?) is easily estimated theoretically; and the transfer integrals b. Our own estimates of b are at best inaccurate; what is important is that b is a simple parameter, referring to the properties of isolated d-electrons, and thus is not at all beyond computation or measurement. In the present high state of development of ligand field theory, even the direct exchange integrals are not out of reach of computation; it is only at the true indirect exchange effect that we reach a great degree of complexity.

The empirical part of this paper makes no claims to completeness or accuracy, but merely shows that with the rough indications furnished by ligand field theory a good semiquantitative understanding of at least the simplest and largest superexchange effects can be

---

[25] F. Stern, Phys. Rev. **94**, 1412 (1954).

[26] P. W. Anderson, Phys. Rev. **79**, 705 (1950).

[27] J. Kanamori, J. Phys. Chem. Solids (to be published); T. Nagamiya, "State of atoms in magnetic crystals," Conference of Welsh Foundation, December, 1958.

[28] Wollan, Child, Koehler, and Wilkinson, Phys. Rev. **112**, 1132 (1958).

[29] J. B. Goodenough, J. Phys. Chem. Solids **6**, 287 (1958).

reached. It shall again be emphasized that more accurate computations, and a wider comparison with experiment, are possible and desirable.

ACKNOWLEDGMENTS

I have been helped by discussions with C. J. Ballhausen, C. Kittel, F. J. Morin, L. E. Orgel, T. Moriya, and J. C. Phillips. R. G. Shulman has helped immeasurably at several stages of this work. This work was begun during a stay at the University of California, made possible by a grant from the U. S. Carbon Corporation.

APPENDIX

In this Appendix the spin-polarization part of the many-electron wave function of a single spin will be derived to the lowest order of perturbation theory, and transformed to localized wave functions. We also demonstrate the simplifications appropriate when localization is a good approximation.

By starting from an optimized set of one-electron wave functions $\varphi_{\mathbf{k}}{}^{n}(\mathbf{r})$, the electron field $\psi_\sigma(\mathbf{r})$ may be written down as in Eq. (3) of the text,

$$\psi_\sigma(r)=\sum_{\mathbf{k}}\Big[\sum_f c_{\mathbf{k}\sigma}{}^{f}\varphi_{\mathbf{k}}{}^{f}(\mathbf{r})+c_{\mathbf{k}\sigma}{}^{d}\varphi_{\mathbf{k}}{}^{d}(\mathbf{r}) +\sum_e c_{\mathbf{k}\sigma}{}^{e}\varphi_{\mathbf{k}}{}^{e}(\mathbf{r})\Big]. \tag{3}$$

We remember that f denotes normally full bands, e normally empty ones and d the d-band. To get the interaction energy between electrons we must insert this into the Coulomb repulsion (7):

$$V=\tfrac{1}{2}\int d\mathbf{r}\int d\mathbf{r}'\ \psi_\sigma{}^{*}(\mathbf{r})\psi_{\sigma'}{}^{*}(\mathbf{r}') \times e^2|\mathbf{r}-\mathbf{r}'|^{-1}\psi_{\sigma'}(\mathbf{r}')\psi_\sigma(\mathbf{r}). \tag{7}$$

Take as unperturbed state

$$\Psi=c_{\mathbf{k}\sigma}{}^{d*}\Psi_0=c_{\mathbf{k}\sigma}{}^{d*}\Big[\prod_{f,\mathbf{k},\sigma'} c_{\mathbf{k}\sigma'}{}^{f*}\Big]\Psi_{\text{vac}}. \tag{A1}$$

The perturbed state is

$$\Psi'=\Psi+\delta\Psi,$$

where

$$\delta\Psi=[E_0-H_0]^{-1}V\Psi. \tag{A2}$$

The part of this which leads to spin polarization is that in which electron $\mathbf{r}'$ goes from the d-band to an empty one, while electron $\mathbf{r}$ goes from a full band into the d-band; for this we take the c^f component of ψ_σ, the c^d of $\psi_\sigma{}^*$, c^e of $\psi_{\sigma'}{}^*$, and c^d of $\psi_{\sigma'}$. This means that the relevant part of V is

$$V\longrightarrow \sum_{e,f,\sigma',\mathbf{k}',f'',\mathbf{k}'''} c_{\mathbf{k}''\sigma'}{}^{d*}c_{\mathbf{k}'''\sigma}{}^{e*}c_{\mathbf{k}\sigma}{}^{d}c_{\mathbf{k}'\sigma'}{}^{f} \times\int d\mathbf{r}d\mathbf{r}'\ e^2|\mathbf{r}-\mathbf{r}'|^{-1}\varphi_{\mathbf{k}''}{}^{d*}(\mathbf{r}') \times\varphi_{\mathbf{k}'''}{}^{e*}(\mathbf{r})\varphi_{\mathbf{k}}{}^{d}(\mathbf{r})\varphi_{\mathbf{k}'}{}^{f}(\mathbf{r}'). \tag{A3}$$

Let us call the integral in the above expression

$$\int d\mathbf{r}d\mathbf{r}'\cdots=\delta(\mathbf{k}'''+\mathbf{k}''-\mathbf{k}'-\mathbf{k})J_{\mathbf{k}}(e,f;\mathbf{k}',\mathbf{k}''); \tag{A4}$$

then the perturbed function $\delta\Psi$ in (A2) is

$$\delta\Psi=-\sum_{\mathbf{k}',\mathbf{k}'',\sigma',e,f} J_{\mathbf{k}}(e,f;\mathbf{k}',\mathbf{k}'') \times(\epsilon_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''}{}^{e}+\epsilon_{\mathbf{k}''}{}^{d}-\epsilon_{\mathbf{k}'}{}^{f}-\epsilon_{\mathbf{k}}{}^{d})^{-1} \times c_{\mathbf{k}'',\sigma'}{}^{d*}c_{\mathbf{k}',\sigma'}{}^{f}c_{\mathbf{k}+\mathbf{k}'-\mathbf{k}'',\sigma}{}^{e*}\Psi_0. \tag{A5}$$

This is the major part of the spin-polarization. It consists of virtually putting the d-electron with unchanged spin into an empty band e, while replacing it with a spin-independent excitation from band f into band d.

Now, if we introduce Wannier functions

$$c^d(\mathbf{R},\sigma)=N^{-\frac{1}{2}}\sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}}c_{\mathbf{k}\sigma}{}^{d},$$

the localized quasi-particle is approximately

$$s^{d*}(\mathbf{R},\sigma)=c^{d*}(\mathbf{R},\sigma)-\sum_{\mathbf{k}\mathbf{k}'\mathbf{k}''ef\sigma'} J_{\mathbf{k}}(e,f;\mathbf{k}',\mathbf{k}'') \times[\epsilon_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''}{}^{e}+\epsilon_{\mathbf{k}''}{}^{d}-\epsilon_{\mathbf{k}}{}^{d}-\epsilon_{\mathbf{k}'}{}^{f}]^{-1} \times(e^{i\mathbf{k}\cdot\mathbf{R}}c_{\mathbf{k}'',\sigma'}{}^{d*})c_{\mathbf{k}',\sigma'}{}^{f}c_{\mathbf{k}+\mathbf{k}'-\mathbf{k}'',\sigma}{}^{e*}. \tag{A6}$$

This turns out to contain various localized $c^d(R',\sigma')$ functions with $\mathbf{R}'$ near $\mathbf{R}$, so is quite complicated. If it is at all a good approximation to use localized particles, we may simplify (A6) very much by using

(a) $\epsilon_{\mathbf{k}''}{}^{d}-\epsilon_{\mathbf{k}}{}^{d}\simeq 0$,

(b) $J_{\mathbf{k}}(e,f;\mathbf{k}',\mathbf{k}'')$

$$=N^{-1}\sum_{\mathbf{R}',\mathbf{R}''}\int d\mathbf{r}d\mathbf{r}'\ e^2|\mathbf{r}-\mathbf{r}'|^{-1} \times\exp[i(\mathbf{k}''\cdot\mathbf{R}''-\mathbf{k}\cdot\mathbf{R}')] \times f^{d*}(\mathbf{r}-\mathbf{R}'')f^{d}(\mathbf{r}'-\mathbf{R}')\varphi_{\mathbf{k}'+\mathbf{k}-\mathbf{k}''}{}^{e*}(\mathbf{r}')\varphi_{\mathbf{k}'}{}^{f}(\mathbf{r})$$

$$\simeq N^{-1}\sum_{\mathbf{R}'}\int d\mathbf{r}d\mathbf{r}'\ e^2|\mathbf{r}-\mathbf{r}'|^{-1} \times\exp[i(\mathbf{k}''-\mathbf{k})\cdot\mathbf{R}']f^{d*}(\mathbf{r}-\mathbf{R}')f^{d}(\mathbf{r}'-\mathbf{R}') \times\varphi_{\mathbf{k}'+\mathbf{k}-\mathbf{k}''}{}^{e*}(\mathbf{r}')\varphi_{\mathbf{k}'}{}^{f}(\mathbf{r})$$

$$=J(e,f;\mathbf{k}''-\mathbf{k},\mathbf{k}'), \tag{A7}$$

which is not a function of $\mathbf{k}$ except through the combination $\mathbf{k}''-\mathbf{k}$. This allows us to sum (A6) first over $\mathbf{k}$, keeping $\mathbf{k}''-\mathbf{k}$ fixed, which leads to

$$\delta\Psi=\sum_{\mathbf{k}',\mathbf{k}''-\mathbf{k},\sigma',e,f} J(e,f;\mathbf{k}''-\mathbf{k},\mathbf{k}') \times(\epsilon_{\mathbf{k}'+\mathbf{k}-\mathbf{k}''}{}^{e}-\epsilon_{\mathbf{k}'}{}^{f})^{-1}\exp[i(\mathbf{k}''-\mathbf{k})\cdot\mathbf{R}] \times c_{\mathbf{k}'+(\mathbf{k}-\mathbf{k}''),\sigma}{}^{e*}c_{\mathbf{k}',\sigma'}{}^{f}c^{d*}(\mathbf{R},\sigma')\Psi_0. \quad \text{(A8)}$$

This is the expression used in the text, which is entirely analogous to the polarization accompanying nuclear indirect exchange. Clearly there are also both the nonlocalized parts and terms in which, instead of $c^{d*}(\mathbf{r},\sigma')$, we have excitations into empty bands e'. We expect that smaller exchange integrals J, and increasing energy denominators, will make these effects quite negligible.

PHYSICAL REVIEW VOLUME 115, NUMBER 1 JULY 1, 1959

Antiferromagnetic Structures of MnS_2, $MnSe_2$, and $MnTe_2$†

J. M. HASTINGS, N. ELLIOTT, AND L. M. CORLISS
Chemistry Department, Brookhaven National Laboratory, Upton, New York

(Received February 13, 1959)

The antiferromagnetic structures of MnS_2, $MnSe_2$, and $MnTe_2$ have been obtained by neutron diffraction. The disulfide exhibits ordering of the "third" kind, the ditelluride, ordering of the "first" kind, and the diselenide, an arrangement which is intermediate between the two. The structures are discussed from the viewpoint of indirect exchange. Magnetic susceptibilities, obtained by the Gouy method, are reported for all three compounds.

NEUTRON diffraction studies of the three polymorphic forms of MnS have shown[1] that the antiferromagnetic structures developed in these materials may be correlated with crystal structure. More specifically, in α-MnS, second nearest neighbor manganese atoms are joined by sulfur atoms located on octahedral sites, whereas in the two β forms, tetrahedrally coordinated sulfur atoms join nearest neighbor manganese atoms. Furthermore, the separation and spatial arrangement of manganese atoms are the same in both α-MnS and the zinc blende form of MnS. The observed magnetic structures indicate that strong antiferromagnetic correlations exist between second nearest neighbors in the case of α-MnS and between nearest neighbors in the two β forms. Indirect exchange coupling thus appears to involve octahedral sulfur bonds in the first instance and tetrahedral bonds in the second.

In an effort to further elucidate the role of the anion in antiferromagnetism, these studies have been extended to the homologous series consisting of MnS_2, $MnSe_2$, and $MnTe_2$. These compounds crystallize with the pyrite structure, which is a NaCl-like arrangement of M and X_2 groups with the axes of the X_2 groups parallel to the various body diagonals. The structure is shown schematically in Fig. 1. A salient feature is the presence of nearly regular tetrahedra whose corner positions are occupied by three metal atoms and one member of an X_2 group, and whose center is occupied by the other member of the X_2 pair. Magnetic as well as crystal-chemical evidence[2] suggest that these compounds may be regarded as essentially ionic combinations of Mn^{+2} and $X_2^{=}$ groups. From this point of view, the face-centered cubic structure of manganese ions might be expected to exhibit second nearest neighbor antiferromagnetic correlations as in MnO and α-MnS. On the other hand, the existence of nearly tetrahedral Mn-X-Mn linkages between nearest neighbor manganese atoms would suggest the possibility of an ordering scheme based upon nearest neighbor antiferromagnetic correlations as in β-MnS. It would not appear possible at present to predict, either on theoretical grounds or on the basis of empirical knowledge of antiferromagnetic

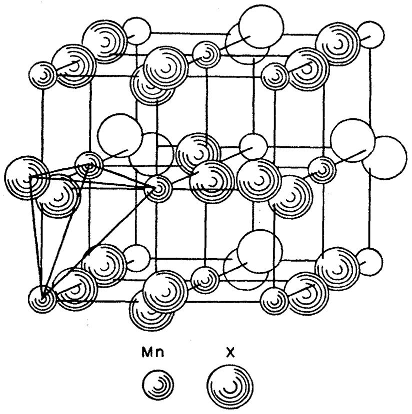

FIG. 1. The pyrite structure.

---

† Research performed under the auspices of the U. S. Atomic Energy Commission. A preliminary account of this work has been published in J. Appl. Phys. **29**, 391 (1958).

[1] Corliss, Elliott, and Hastings, Phys. Rev. **104**, 924 (1956).

[2] N. Elliott, J. Am. Chem. Soc. **59**, 1958 (1937); L. Néel and R. Benoit, Compt. rend. **237**, 444 (1953).