PHYSICAL REVIEW VOLUME 79, NUMBER 2 JULY 15, 1950

Antiferromagnetism. Theory of Superexchange Interaction

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

(Received January 9, 1950)

In this paper the general formalism of Kramers indicating the existence of superexchange interaction has been reduced, under simplifying assumptions, to the point where actual formulas for the interaction can be written down directly in terms of spin operators, with certain exchange and transition integrals as parameters. Two results of physical interest are the following: (a) superexchange must be expected to show the directional properties (as far as directional relations of interacting magnetic ions are concerned) of the orbitals in the outer shell of the non-magnetic connecting ions; and (b) the sign of the effective exchange integral depends upon the sign of the internal exchange coupling of an added electron on the magnetic ion.

I. INTRODUCTION

IN the course of a general study of the phenomenon of antiferromagnetism[1] it became clear that in many substances the atoms which strongly interact magnetically are quite definitely separated from each other by intervening non-magnetic ions. In such substances as MnSe and MnTe this can be verified simply by drawing spheres of ionic radii about the anions, which then separate practically completely even nearest neighbor cations. However, the best illustration of this fact is the substance MnO, in which, by means of neutron diffraction,[2] Shull has been able to actually identify the pattern of magnetic spins below the antiferromagnetic Curie point. This pattern was at first baffling, since it showed that the 12 nearest neighbor spins (MnO is an NaCl structure; the Mn^{++} ions are thus in a face-centered cubic lattice) to a given spin are uncorrelated with this spin; instead, the four separate simple cubic lattices of which the f.c.c. lattice is made up are each lined up in a perfect antiferromagnetic pattern, but entirely uncorrelated one with the other.[3] The pattern is indicated in Fig. 1.[4]

This pattern means that the strongly coupled Mn^{++} ions are separated by a supposedly non-magnetic O^{--} ion directly between them, as we can see from Fig. 1. The Curie point of MnO is 122°K; this means that the exchange coupling between these separated ions is more than a tenth of ordinary exchange integrals (for instance in Fe), and it seems inconceivable that this large magnitude can come from "ordinary" or "direct" exchange. Thus we are led to believe that the primary interaction in MnO, and to a lesser certainty in other similar antiferromagnets (FeO, MnS, etc.) is a superexchange interaction operating through the intermediary of the O^{--} ions.

II. A SIMPLE CASE

Kramers[5] first pointed out that it is possible to have an exchange spin-coupling through the agency of intermediate non-magnetic atoms. The resolution of this paradox lies in Kramers' mechanism: his assumption that "excited states" of the intermediate atoms are present is equivalent to assigning a weight in the total wave function to configurations in which the intermediate atoms have some paramagnetism. As an example we may take such a crystal as MnO; in addition to the totally ionic state $Mn^{++}O^{--}$ it is reasonable from all available evidence to allow a considerable admixture of the state in which at least one *p*-electron from the oxygen has gone into an *s* or *d*-state on Mn^{++}, and thus the oxygen is paramagnetic and can enter into "magnetic" interactions.

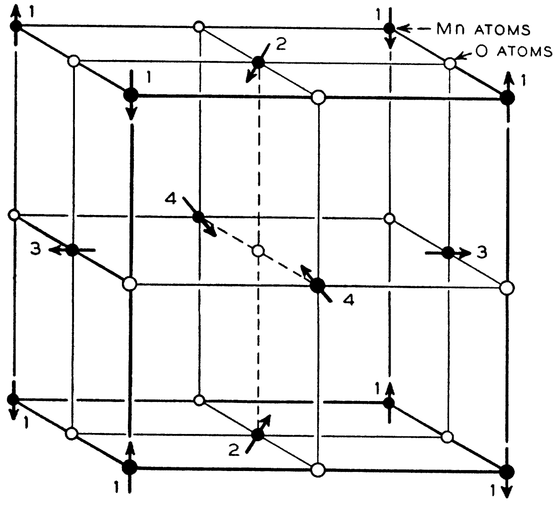

FIG. 1. Pattern for the MnO structure.

---

[1] To be published in a series of papers by G. H. Wannier and the author.

[2] C. G. Shull and J. S. Smart, Phys. Rev. **76**, 1256 (1949). Dr. Shull has kindly given us more recent results as to the most probable actual pattern present.

[3] Dr. Shull informs us that the data are not inconsistent with a pattern suggested by Neél, which has the sublattices not uncorrelated but arranged (presumably by weaker forces of some kind) in such a way as to make all spins parallel in 111 planes, and these planes antiparallel in succession. Such a pattern is not different in essentials from Fig. 1, and still requires strong next nearest neighbor interaction.

[4] In a later paper of this series we shall show that there is theoretical as well as experimental evidence for the pattern found by Dr. Shull. A modified Weiss molecular field treatment indicates that, if T_c is the Curie point and θ the constant in $\chi=c/(T+\theta)$, a value of $\theta/T_c=5$, as in MnO, indicates that the next-nearest neighbor interaction is high and that the pattern is probably that of Fig. 1. Also, quantum mechanical considerations seem to indicate that possibly no *nearest neighbor* antiferromagnetism can exist in the f.c.c. lattice.

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

Let us, to fix our ideas, think of the system of two magnetic ions, such as Mn^{++}, separated by an O^{--} ion. We shall, for simplicity, consider only two of the electrons on the O^{--} ion, assuming them to be in the same p-state; in addition we shall consider the magnetic ions as possessing one electron each, in (for example) a d-state. (We always ignore orbital moments, however, assuming them to be quenched). By hypothesis there is no direct overlap between the two d wave functions; also by hypothesis there is a finite probability that an electron will jump from the O^{--} ion into (perhaps) an s-state on one of the magnetic ions. We must think (as Kramers did not point out) of this s-state as coupled by some strong spin dependent interaction with the d-electron already on the magnetic ion; otherwise there is no superexchange.

Qualitatively it is now easy to see the process involved in superexchange. To the ground state only two spin-states are available, since the p-state spins are necessarily antiparallel: a singlet and a triplet. If, for simplicity, we assume that in the "excited" state the two electrons, which here occupy one magnetic ion, are coupled strongly antiparallel by internal interactions there are in this state also a singlet and a triplet only. The situation is shown schematically in Table I. We see that in the *excited* state there is a triplet-singlet splitting because there can be an appreciable exchange integral between the remaining p-electron on the negative ion and the magnetic electron on the magnetic ion to which an electron has not been transferred. Now, we assume that the true wave function is given by

$$\psi^{s,\,t} = a\psi^{s,\,t}{}_{\text{ground}} + b\psi^{s,\,t}{}_{\text{excited}}. \tag{1}$$

The part of the Hamiltonian causing the ground-excited transition is not such as to mix singlet and triplet states. Thus one or the other superscript in (1) holds and there is a singlet-triplet splitting for the true wave function of approximately

$$\Delta E_{s-t(\text{apparent})} = b^2(\Delta E)_{s-t(\text{excited})}. \tag{2}$$

This is equivalent to a normal exchange coupling, reduced by the quantity b^2. Note that the total number of spins has not changed, and thus that the effective magneton number will not be altered greatly by superexchange coupling.

III. THE KRAMERS FORMULA

In his original paper, Kramers gives a perturbation treatment of the superexchange problem. His ideas are applicable to a system for which we define a "ground" state in which: (a) the electron wave-functions are orthogonal, (b) all "non-magnetic" ions contain only filled shells, (c) the effect of exchange between atoms is neglected, although exchange internal to the ions is included.[6] The unperturbed Hamiltonian H_0 is that which this ground state satisfies. He then shows that one lifts the permutation degeneracy of this ground state neither by including the exchange terms H^{ex} (violating (c)) nor by the transition terms H^{tr} (violating (b)) separately, but that the first terms entering are of third order in these interactions, containing H^{tr} twice and H^{ex} once.[7]

TABLE I. Interaction scheme in superexchange.

	Electron number		1	2	3	4
Ground Orbital State	Electron is on ion		magnetic$_1$	O^{--}	O^{--}	magnetic$_2$
	in orbital state		d_1	p	p	d_2
	Magnetic State	singlet	spin up	(antiparallel)		spin down
			(no splitting because no 1–4 interaction)			
		triplet	spin up	(antiparallel)		spin up
Excited Orbital State	Electron is on ion		magnetic$_1$	magnetic$_1$	O^{--}	magnetic$_2$
	in state		d_1	s_1	p	d_2
	Magnetic State	singlet	(antiparallel)		spin up	spin down
				(splitting if 3–4 interact)		
		triplet	(antiparallel)		spin up	spin up

We shall take Kramers' final expression for the superexchange interaction matrix $H_{tt'}$ as given; it is simply the third-order perturbation under these conditions:

$$H_{tt'} = \sum_{u,\,u'} \frac{(t|H^{tr}|u)(u|H^{ex}|u')(u'|H^{tr}|t')}{(E_u - E_t)(E_{u'} - E_t)}. \tag{3}$$

Here t and t' refer to the various initially degenerate spin states of the ground state, while u and u' are various spin states of the excited orbital state. We have implicitly included the strong internal exchange couplings, such as that postulated in our simple model between the d and s-electrons on Mn^{+}, in the unperturbed Hamiltonian, so that E_u and $E_{u'}$ are not necessarily equal. $H_{tt'}$ is then an effective Hamiltonian due to the third-order perturbation, acting on the unperturbed states.

So far we have been presenting essentially the ideas of Kramers. The purpose of this paper is to find a method of evaluating the sum (3) in such a way that even in complicated cases we can immediately write down $H_{tt'}$ in terms of spin operators referring to the ground state. These can then be evaluated to give the spin-dependence of the superexchange coupling.

---

[6] The assumption (c) will be justified later simply on the grounds that the internal exchange splittings are as large as the ground-excited orbital splittings.

[7] For some applications it is of interest to point out that H^{tr} contains, besides obvious terms giving the non-ionicity of the compound (thus violating assumption (b)), some other terms which result from the physical fact of violation of assumption (a). If, in the simple model of Table I, the wave-functions "d" and "p" are not orthogonal, there is a splitting in the *ground* state. Since we have artificially excluded this possibility by orthogonalizing, there must be terms in H^{tr} restoring the actual physical state of affairs. These terms will have the order of magnitude, (exchange integral)×(overlap integral),[2] which is generally much smaller than the interactions involved in the antiferromagnets. However, such terms might go far in explaining the observed magnitudes of exchange interactions in more magnetically dilute substances.

IV. EVALUATION OF THE KRAMERS EXPRESSION

A short introductory note on the method to be used is in order. The basic idea is similar to one Serber[8] used in a study of configuration interactions in atomic spectra: we shall extend the Dirac vector model of spin-coupling by exchange forces to the problem of the interaction of two configurations, which we shall call the ground (ionic) state and the excited state. If this interaction can be assumed small Kramers' matrix element (3) is the correct form for it, written as a perturbation on the levels of the ground configuration. There is a second-order effect, but this is necessarily the same for all parts of the ground configuration and merely changes the average total energy of the ground state.

We shall treat at first only the four-electron system with the ground configuration (A) consisting of one electron on each of two magnetic ions, (orbitals d_1 and d_2) and two electrons in one orbital (p) of the intervening ion. The excited configuration (B) is that in which one electron from a p-orbital has gone into an orbital on one of the magnetic ions (say "s_1"). Later work will indicate the essentially easy extension to more complicated systems. The Kramers expression (3) involves three matrix elements, two of which connect A with B (M_{AB}) and one of which involves only the excited configuration (M_{BB}). In order to couple the d_1 and d_2 spins, it is necessary to have *two* spin-dependent factors in this expression. One of these is always furnished by M_{BB}, which gives a coupling of p and d_2, but in order also to couple d_1 with p and "complete the chain" between d_1 and d_2 there must be a second spin dependence either in M_{AB} or hidden in the expression (3). There are three ways in which this spin-dependence can be achieved: (a) M_{AB} may be an exchange matrix element causing the transition. We treat this case in Appendix I, although we believe it to be secondary. In the other possibilities we allow M_{AB} to be spin-independent, but assume that the energy or the very existence of the separate multiplets of the configuration B are so spin-dependent that the denominator terms in (3) provide the second spin-dependence, either (b) due to strong internal exchange coupling between s_1 and d_1, or (c) due to the operation of the exclusion principle if s_1 and d_1 happen to be the same orbital, which would give a result like our "simple example."

Cases (b) and (c) may be treated together. To treat these two cases we shall, for the time being, assume that the forces causing the actual orbital transition are of a one-electron type; we might use a Hartree field caused by the unperturbed orbitals to determine the transition probabilities. Since one-electron Hamiltonians are spin-independent we shall show that $(t|H^{tr}|u)$ can be thrown into a very simple form. For example, we may choose as our basic wave-functions for orbitals a, b, c, d, etc. the functions

$$\psi(s_1s_2s_3s_4\cdots)=\psi(s) = A[a(1)\chi_{s1}(1)b(2)\chi_{s2}(2)\cdots\text{etc.}] \quad (4)$$

where the χ's are the basic spin-function α and β, and A is an operator antisymmetrizing these products in electrons, 1, 2, $3\cdots$. Then if the orbital states are

$$\psi_t=a(1)b(2)c(3)d(4),\ \cdots, \quad (5)$$
$$\psi_u=a'(1)b(2)c(3)d(4),\ \cdots,$$

$$(t|H^{tr}|u)=(a|H^{tr}|a')\delta_{tu}, \quad H^{tr}\simeq H_{1\text{-electron}}, \quad (6)$$

where t and u are taken to represent the sets of spins s_1, $s_2\cdots$ belonging to the various orbits. One difficulty is here apparent: if some of the sets $abcd\cdots$, or $a'bcd\cdots$, of orbitals are identical, some of the wave-functions (4) will vanish identically or become linearly dependent. For the time being we shall use Serber's trick of pretending that the orbitals are not identical, and will later remove the non-existent wave-functions by a simple strategem.

In our further development it will be necessary to choose for our upper states, u, not the basic functions (4) but linear combinations of these chosen in such a way that part or all of the spin-dependent exchange Hamiltonian is diagonalized. In this case the new states are obtainable by means of a unitary (canonical) transformation from the old states (4) (which we shall call "s"):

$$\psi_u=\textstyle\sum_s(u|s)\psi(s). \quad (7)$$

Similarly, it will be necessary to pick definite linear combinations, t, for the lower states (here primarily in order to choose states which are allowed by the symmetries in the orbital function); again there is a unitary transforming matrix $(t|s)$. Now we can write down a new form for the transition matrix $(t|H^{tr}|u)$, in terms of any such linear combinations:

$$(t|H^{tr}|u)=(a|H^{tr}|a')\textstyle\sum_{s,\,s'}(t|s)(\delta_{ss'})(s'|u) = (a|H^{tr}|a')(t|u). \quad (8)$$

Thus the matrix causing transitions between the two sets of orbitals is simply a constant times the appropriate canonical transformation of the spin functions. The step of starting always from (4) means that we have a $1-1$ correspondence among spin functions for the two states, in spite of the difference in orbitals; thus the canonical transformation $(t|u)$ exists.

Now it is necessary to consider the spin-dependent, or exchange, part of the Hamiltonian. The part of the perturbation $(u|H^{ex}|u')$ is such a spin-dependent Hamiltonian. In addition, before superexchange can occur we showed that it is necessary that the problem contain a strong spin-dependent coupling (included, as we pointed out, really in the unperturbed Hamiltonian) sufficient to affect appreciably the energy levels E_u. Among such "strong" couplings we include the infinitely strong coupling involved in the use of the same orbital for two electrons.

The spin-dependent part of the Hamiltonian, referred exclusively to upper states u, can be written down[9] in terms of operators referring to the spins of

---

[8] R. Serber, Phys. Rev. **45**, 461 (1934)

[9] P. A. M. Dirac, *Quantum Mechanics* (Clarendon Press, Oxford, 1930), first edition, p. 216.

electrons in the various orbitals. If we have orbitals i, j, etc. the Hamiltonian is

$$V=-\sum_{i<j}\tfrac{1}{2}V_{ij}(1+\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j)+\text{const.} \tag{9}$$

V_{ij} is the matrix element referring to interchange of orbitals i and j. We divide (9) artificially into two parts

$$V=V_1+V_2 \tag{10}$$

where V_1 contains all the terms which are included in E_u. In our simple case of Table I, we would have:

$$V=-\tfrac{1}{2}[V_{d_1s_1}(1+\boldsymbol{\sigma}_{d_1}\cdot\boldsymbol{\sigma}_{s_1})+V_{pd_2}(1+\boldsymbol{\sigma}_p\cdot\boldsymbol{\sigma}_{d_2})], \tag{11}$$

where

$$V_1=\tfrac{1}{2}V_{d_1s_1}(1+\boldsymbol{\sigma}_{d_1}\cdot\boldsymbol{\sigma}_{s_1})$$
$$V_2=\tfrac{1}{2}V_{pd}(1+\boldsymbol{\sigma}_p\cdot\boldsymbol{\sigma}_{d_2}).$$

Suppose that we choose our upper spin-states u in such a way as to diagonalize the operator V. Then the basic Eq. (3) for the coupling becomes, with (8) (upper and lower states differ only in the choice of orbital a or a' for one electron):

$$H_{tt'}=|(a|H|a')|^2\sum\frac{(t|u)(u|V_2|u)(u|t')}{(E_u-E_t)^2}. \tag{12}$$

If V_1 is sufficiently strong that we may say that only those states belonging to the lowest eigenvalue E_u in (12) can occur, the others being completely absent because of large denominators, it is open to us to modify V_2 in such a way as to leave out entirely the high states. Then we are permitted to "pretend" that E_u-E_t is constant for all states.

The method of modifying V_2 in such a way as to eliminate the states belonging to higher values of E_u is well known. We simply multiply by the appropriate projective operator which "projects" V_2 on the subspace belonging only to the lowest eigenvalue of V_1.

$$V_2'=PV_2,$$

or

$$(V_2')_{uu'}=\sum_{u''}P_{uu''}(V_2)_{u''u'}. \tag{13}$$

At this point it is possible also to take into account the effect of having pairs of identical orbitals in either upper or lower states. In the upper states we simply use a slight modification of the above procedure. That is, we ignore the states which are not allowed—states which contain the same pair of orbitals with parallel spins—by applying the appropriate projective operator (such as (16′) below) to eliminate such states. However, we shall show in Appendix II that this is not quite enough to take care of the linear dependences between states when pairs of orbitals are identical (for example, the states $A[a(1)\alpha(1)a(2)\beta(2)]$ and $A[a(1)\beta(1)a(2)\beta(2)]$ are identical but for sign.) There we show that, because the unitary transformation (7) acting on such an "incomplete" set of states can lead to non-normalized wave functions, we must multiply the projective operator by a factor 2 to take care of a re-normalizing factor whenever a pair of orbitals are equal.

If pairs of orbitals in the lower state are identical, we can take care of this fact simply by ignoring such states in our final result. We shall, later, actually take advantage of the fact that there is a basic limitation on the permissible lower states. Again, however, for every pair of identical orbitals we must multiply our final interaction by 2 to renormalize the wave functions, as will be shown in Appendix II. Now that we have a constant denominator in (12), we can apply the rules of matrix multiplication and obtain simply:

$$H_{tt'}=\frac{|(a|H^{tr}|a')|^2}{[(E_u-E_t)_{(\text{lowest})}]^2}\times\sum_u(t|u)(u|V_2'|u)(u|t')=b^2(t|V_2'|t') \tag{14}$$

where

$$b=\frac{(a|H^{tr}|a')}{(E_u-E_t)_{(\text{lowest})}}. \tag{15}$$

We see that we have now written down the superexchange interaction operator $H_{tt'}$ in terms of V_2', which is in general easily obtainable for a particular case. As it stands in (13), V_2' is an operator working only on spin states of the "excited" state; however, it is just the idea of our calculation, and of the simplification (8), that we can establish a precise correspondence among states t and u, and, in the end, simply write V_2' as though it pertained to electrons in the lower states. Physically, this means that since the transition probability is spin-independent, the electrons retain their spins on making the transition.

Our example of Eq. (11) will make this clearer. There are two cases, depending on the sign of the "internal" exchange integral V_{d_1s}:

(A). V_{d_1s} large, >0. Then parallel spins $\boldsymbol{\sigma}_{d_1}$, $\boldsymbol{\sigma}_s$ are favored, and we want to write for P an operator which eliminates all states for which $\boldsymbol{\sigma}_{d_1}$ and $\boldsymbol{\sigma}_s$ are antiparallel. Such an operator is

$$P_1=\tfrac{1}{4}(3+\boldsymbol{\sigma}_{d_1}\cdot\boldsymbol{\sigma}_s). \tag{16}$$

Then in this case (inserting the extra factor 2):

$$H_{tt'}=b^2(t|\tfrac{1}{4}(3+\boldsymbol{\sigma}_{d_1}\cdot\boldsymbol{\sigma}_{p_1})V_{p_2d_2}(1+\boldsymbol{\sigma}_{p_2}\cdot\boldsymbol{\sigma}_{d_2})|t'). \tag{17}$$

We have simply read $\boldsymbol{\sigma}_{d_1}$ for $\boldsymbol{\sigma}_s$, this being the electron making the transition.

(B). V_{ds} large, <0 (or d and s are same orbital). Parallel spins for the d and s orbitals are to be eliminated in this case, and the proper projective operator is[10]

$$P_2=\tfrac{1}{4}(1-\boldsymbol{\sigma}_{d_1}\cdot\boldsymbol{\sigma}_{s_1}) \tag{16'}$$

$$H_{tt}'=\tfrac{1}{4}b^2(t|V_{pd}(1-\boldsymbol{\sigma}_{d_1}\cdot\boldsymbol{\sigma}_{p_1})(1+\boldsymbol{\sigma}_{p_2}\cdot\boldsymbol{\sigma}_{d_2})|t'), \tag{17'}$$

---

[10] We shall hereafter simply understand the factor 2 when P_2 is due to a Pauli principle limitation, i.e. d and s are the same orbital. Since b^2 is unknown anyway, this leads to no important change in the theory.

TABLE II. Sign of the interaction.

	If internal coupling favors spin of added electron	and exchange coupling of p-electron and other atom is	then coupling is
1	parallel	ferromagnetic	antiferromagnetic
2	parallel	antiferromagnetic	ferromagnetic
3	antiparallel	ferromagnetic	ferromagnetic
4	antiparallel	antiferromagnetic	antiferromagnetic

where in both cases

$$b=\frac{(s|H_{1\text{-electron}}|p)}{E_{(\text{bonding})}-E_{(\text{ionic})}}. \tag{18}$$

To compute the effective coupling we observe that $\boldsymbol{\sigma}_d\cdot\boldsymbol{\sigma}_p$ acting on wave-functions in which $\boldsymbol{\sigma}_p$ is the spin of one of a pair of electrons in the p-orbital gives an identically zero result, as it must always make an $S=1$ state out of an $S=0$ state, and the $S=1$ wave functions vanish identically. Then the true coupling is proportional to

$$(\boldsymbol{\sigma}_{d1}\cdot\boldsymbol{\sigma}_p(1))(\boldsymbol{\sigma}_{d2}\cdot\boldsymbol{\sigma}_p(2))$$

$\boldsymbol{\sigma}_p(1)+\boldsymbol{\sigma}_p(2)=0$; thus the coupling is, by a simple expansion into components,

$$\begin{aligned}(\boldsymbol{\sigma}_{d1}\cdot\boldsymbol{\sigma}_p(1))(\boldsymbol{\sigma}_{d2}\cdot\boldsymbol{\sigma}_p(2))&=-(\boldsymbol{\sigma}_{d1}\cdot\boldsymbol{\sigma}_p)(\boldsymbol{\sigma}_{d2}\cdot\boldsymbol{\sigma}_p)\\&=-(\boldsymbol{\sigma}_{d1}\cdot\boldsymbol{\sigma}_{d2}),\end{aligned} \tag{19}$$

where the cross-product term $i\boldsymbol{\sigma}_p\cdot(\boldsymbol{\sigma}_{d1}\times\boldsymbol{\sigma}_{d2})$ has vanished, again because $\boldsymbol{\sigma}_p$ vanishes identically.

The effective coupling is thus

$$\mp\tfrac{1}{4}b^2V_{pd}(\boldsymbol{\sigma}_{d1}\cdot\boldsymbol{\sigma}_{d2})\begin{bmatrix}-\text{Case } A\\ \\+\text{Case } B\end{bmatrix}.$$

There is an equal term also present due to transitions to atom 2 and exchange with atom 1. Note that (17) is correctly isotropic in respect to lattice rotations.

V. GENERALIZATIONS

Various generalizations of this method from the simple case come immediately to mind. The most apparent is the generalization to the case in which V_1 is not infinitely strong, but instead leads simply to a finite splitting of the levels E_u, sufficient for the energy denominators in (12) to be somewhat different. For example, suppose that in Case A

$$\frac{(\Delta E)_{\text{parallel}}}{(\Delta E)_{\text{anti-parallel}}}=x,\ x<1.$$

Then it is necessary to leave in the sum (12) an amount of the antiparallel states proportional to x^2. A "partial projective operator" which will do just this is

$$\begin{aligned}P'&=\tfrac{1}{4}(3+\boldsymbol{\sigma}_d\cdot\boldsymbol{\sigma}_s)+\tfrac{1}{4}x^2(1-\boldsymbol{\sigma}_d\cdot\boldsymbol{\sigma}_s)\\&=\tfrac{1}{4}[(3+x^2)+(1-x^2)\boldsymbol{\sigma}_d\cdot\boldsymbol{\sigma}_s],\end{aligned} \tag{20}$$

then

$$\begin{aligned}H=\tfrac{1}{4}b^2V_{pd}[(3+x^2)+(1-x^2)(\boldsymbol{\sigma}_{d1}\cdot\boldsymbol{\sigma}_{p1})]&\\\times[1+\boldsymbol{\sigma}_{d2}\cdot\boldsymbol{\sigma}_{p2}]&\end{aligned} \tag{21}$$

note that if $x^2=1$, or $V_{\text{strong}}=0$, there is no coupling, since the singlet and triplet averages of H are not then different.

The reader may object that the splittings caused by Hund's rule that internal exchange favors parallel spins are, in general, quite small (of the order of 2 to 3 ev) compared to normal orbital splittings, and thus that x should be nearly equal to unity, and the internal exchange should also be a small perturbation. Then the terms considered in the main body of the paper would be fourth-order, and thus small compared to those due to exchange transitions, which are considered in Appendix I. These latter are easily worked out but do not lead to such simply interpretable results as the ones we are considering. However, it must be remembered that the very fact that superexchange is so strong indicates that the energy level representing ionization of the anion is very low in the antiferromagnetic crystals, as one could also conclude by noting that O$^-$, Te$^-$, etc., have large negative electron affinities. BaO, a compound similar to those considered, has its first exciton level at roughly 3.8 ev which is certainly *not* large compared to the 2 to 3 ev exchange splittings. We therefore believe that the terms here considered are probably the dominant ones.

A second trivial generalization is to the case in which the ions have spins of greater than one-half. In this case we must again guess the direction of coupling of an added electron (perhaps by Hund's rule) and proceed essentially as in the simple case. The projective operators are derived as follows:

$$(\mathbf{S}\cdot\boldsymbol{\sigma})_{(S+\frac{1}{2}\boldsymbol{\sigma})=S+\frac{1}{2}}=S \tag{22}$$

$$(\mathbf{S}\cdot\boldsymbol{\sigma})_{(S+\frac{1}{2}\boldsymbol{\sigma})=S-\frac{1}{2}}=-(S+1) \tag{22$'$}$$

so that in case we wish to eliminate parallel spin $-\mathbf{S}$ cases we use

$$P_{S(\text{anti-parallel})}=\left(1-\frac{\mathbf{S}\cdot\boldsymbol{\sigma}}{S}\right)\cdot\frac{S}{2S+1} \tag{23}$$

and in the opposite case

$$P_{S(\text{parallel})}=\left(1+\frac{\mathbf{S}\cdot\boldsymbol{\sigma}}{S+1}\right)\cdot\frac{S+1}{2S+1}. \tag{23$'$}$$

Here again it is possible to apply the argument of equation (19) and thus the coupling is always proportional to $\mathbf{S}_1\cdot\mathbf{S}_2$.

Table II expresses the available information on the sign of the interaction.

The dependence of the coupling on the sign of the internal coupling can be verified in some actual cases, since by Hund's rule we know that an electron will be added most easily with parallel spin for ions with less than half-filled shells (Cr^{++} for instance) and antiparallel for ions with more than half-filled shells (Mn^{++}). Indeed, it is observed that chromium telluride is ferromagnetic, although manganese telluride, analagous in structure and very similar in many ways, is antiferromagnetic. The dichlorides of these two ions, and also of V and Fe, show, again, this sign change although here the crystal structures are not known.[11] Many other compound pairs could easily be investigated for this feature.

It is easily seen that the presence of other electrons in the same shell as that occupied by the "p" (or transition-making) electron does not change the mathematics. So long as only one electron leaves O^{--} at a time, the electrons in the shell other than the one associated with the absent one are represented in V_1 by infinite terms; thus the wave functions in which they are not antiparallel by pairs are absent. However, these wave functions are not considered in the ground state either; thus these electrons have no effect, since the related projection operator does not enter the true Hamiltonian. This reasoning has an interesting consequence: if superexchange occurs through p-electrons it has a p-type of directionality. That is, if the electron making the transitions is primarily the p-electron, as one suspects, the remaining p-electron in *the same state* will provide the exchange coupling. However, this p-state can only couple with the magnetic ion diametrically opposite to the ion to which the transition is made, and the strongest coupling will thus "look like" a p wave-function, with maxima at 180° from the O^{--} ion. One must realize, however, that the s-electrons of the outer shell will contribute a smaller direction-independent coupling.

This conclusion is interesting in view of the quoted results of Shull[12] on MnO, in which he finds a strong coupling of next nearest neighboring Mn ions rather than of nearest neighbors; this indicates that the superexchange directly through the O^{--} ion may be more powerful than that between locations making angles of 90° with the O^{--} ions.[12] This directionality property is simply a result of the superexchange mechanism, and it should be emphasized as being more general than any of the special assumptions of the theory such as the perturbation treatment or one-electron transitions.

It can be shown that the general form of the superexchange interaction is much the same when the electronic structure of the entire solid is taken into consideration. The easiest procedure in this case is to use the full Hamiltonian for the unperturbed energy for all but a few atoms surrounding a pair of chosen magnetic ions; then the exchange and "state-mixing" Hamiltonians are introduced as simultaneous perturbations on the ground state so defined. This is legitimate as long as there is only a small probability of two electrons "jumping" to or from the same ion at the same time; i.e., as long as the perturbation treatment is valid.

This paper incorporates the results of many helpful discussions with several of my colleagues, particularly Drs. G. H. Wannier, C. Kittel, and W. Shockley. Professor J. H. Van Vleck made several very useful suggestions.

APPENDIX I

Consideration of Transitions Involving Exchange

In the text we have always assumed that the potential causing the transition between orbitals was of the form $\Sigma V(r_i)$ involving only one-electron forces. In case the potential contains important terms of the form $\Sigma V(r_{ij})$ one can carry through the same kind of analysis, but the results are somewhat more complicated and lead to less clarification of the problem.

Let us confine ourselves to the simple case worked out in the text. Then the type of exchange integral which is important is

$$V_{12}=\int s_1(1)d_1(2)p_2(3)d_2(4)V(r_{12})d_1(1)p_1(2)p_2(3)d_2(4). \quad (24)$$

The essential idea is that the proof of Dirac[9] for formula (9) does not break down in this case;[13] the actual transition matrix elements are matrix elements of the operator

$$V=\tfrac{1}{2}V_{12}(1+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2), \quad (25)$$

where we have made the same identification of orbitals in ground and excited states as before:

$$\left[\begin{array}{l}\boldsymbol{\sigma}_1=\text{spin matrix of electron in }\begin{bmatrix}d_1\text{ (ground)}\\ d_1\text{ (excited)}\end{bmatrix}\text{ orbital}\\ \\ \boldsymbol{\sigma}_2=\text{spin matrix of electron in }\begin{bmatrix}p_1\text{ (excited)}\\ s_1\text{ (ground)}\end{bmatrix}\end{array}\right. \quad (26)$$

The proof merely follows the lines of Dirac's proof. First: as an operator acting on products of spin+electron wave functions, $V(r_{ij})$ is equivalent to

$$V_{ij}P_0(ij), \quad (27)$$

where V_{ij} is simply an orbital exchange matrix element of the type (24), while $P_0(ij)$ is the permutation operator interchanging orbitals i and j. Then $P_0(ij)$ on wave-functions forced by the Pauli principle to be antisymmetric is just equivalent to the same permutation acting on the spin wave-functions corresponding to the two orbitals, by the usual argument:

$$P_0\psi(S)=\pm P^{-1}P_0\psi(S)=\pm P_S\psi(S),$$

where P^{-1} is the permutation operator acting on both spins and orbitals. Thus the result of operating with $P_0(ij)$ on our basic wave-functions (4) is the same as that of multiplying by an

---

[11] Foex, *Le Magnetisme* III, (Strasbourg, 1939) p. 222. Also, one should note that the theory is not here as well founded, since the main coupling may not be superexchange.

[12] Still a further experimental confirmation of our general ideas can be found in the Weiss molecular field theory mentioned in a previous footnote applied to a series of compounds MnO, MnS, MnSe, MnTe. Since the electronegativity of the anion decreases along this series, one would expect a decrease in ionic character, and thus an increase in the constant b of Eq. (15). At the same time, the molecular field theory leads us to expect that the θ/T_c ratio should decrease with increasing superexchange coupling of the Mn^{++} ions and thus should decrease along this series. This seems to be the case; the values are (Bizette, *Thesis*, Paris 1946): MnO: 5.0; MnS: 3.2; MnSe: ~3.0.

[13] R. Serber, reference 8, has also proved the theorem used here, that (25) is true (except for factors $\sqrt{2}$) even for configuration transitions.

operator $P_S(ij)$. $P_S(ij)$ can then be proved to be:

$$P_S(ij)=\tfrac{1}{2}(1+\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j). \tag{28}$$

Now the steps by which we made it possible to ignore the energy denominators in (3) can be carried through as before, leaving us with simple matrix multiplication of the three types of matrix elements. The transition matrices $(t|H|u)$ here are simply

$$\tfrac{1}{2}V_{12}(1+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2),$$

which multiply the "modified" exchange matrix $(u|V_2'|u)$.

Let us do an actual example. Suppose that the energy denominators do not enter, and thus that

$$V_2'=\tfrac{1}{2}V_{34}(1+\boldsymbol{\sigma}_3\cdot\boldsymbol{\sigma}_4).$$

Both exchange and one-electron transitions are assumed to be important, and thus the matrix element causing transitions is

$$\tfrac{1}{2}V_{12}(1+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2)+(a|H^{tr}|a').$$

Then Kramers' expression (3) is

$$\begin{aligned}H_{tt'}=\Big\{t\Big|\Big[b+\frac{V_{12}}{E_u-E_t}\Big(\frac{1+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2}{2}\Big)(V_{34})\Big(\frac{1+\boldsymbol{\sigma}_3\cdot\boldsymbol{\sigma}_4}{2}\Big)\Big]&\\ \times\Big[b+\frac{V_{12}}{E_u-E_t}\Big(\frac{1+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2}{2}\Big)\Big]\Big|t'\Big\}.&\end{aligned}$$

The square of $\frac{1}{2}(1+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2)$ is unity, since this is the representative of a simple permutation operator,. and thus, since it commutes with $(1+\boldsymbol{\sigma}_3\cdot\boldsymbol{\sigma}_4)$, the part of this expression involving V_{12}^2 has no spin-dependence. Similarly the b^2 part has no spin-dependence since we have left out the spin-dependence due to the denominator. Only cross-terms between the one-electron and spin transitions remain:

$$\begin{aligned}H_{tt'}&=\frac{bV_{12}V_{34}}{E_u-E_t}\times\tfrac{1}{2}(t|(1+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2)(1+\boldsymbol{\sigma}_3\cdot\boldsymbol{\sigma}_4)1(t')\\ &=-\frac{bV_{12}V_{34}}{E_u-E_t}\times\tfrac{1}{4}(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_4).\end{aligned}$$

We see that here the sign of the coupling cannot be predicted by Hund's rule, although our other conclusions are still valid, and in particular the coupling is still roughly proportional to the square of the non-ionicity, and is still a $(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_4)$ coupling. One might expect that this type of coupling, which would be roughly the same for more than half-filled and less than half-filled shells, in some cases might predominate over the one-electron coupling, and as a result in some cases the sign of the coupling might not change between the two cases, although its magnitude could vary greatly.

APPENDIX II

Justification of the Factor 2 Due to Pairs of Identical Orbitals

In this appendix we shall try to indicate the two sources which lead to an error of a factor of $\sqrt{2}$ in our transition matrix elements computed according to the prescription (8) as compared with the true transition matrix element when identical orbitals are present. The rigorous proof has been given by Serber (reference 8, pp. 463–464).

First we shall compute the transition probability for a very simple case. Suppose our configurations are of two orbitals, and one (A) has two different orbitals (ab), while the second has two identical ones (a^2). Then transitions will only be possible between the singlet states

$$\left.\begin{aligned}\psi_A{}^S&=\tfrac{1}{2}[a(1)b(2)+a(2)b(1)][\alpha(1)\beta(2)-\alpha(2)\beta(1)]\\ \psi_B{}^S&=\tfrac{1}{2}\sqrt{2}a(1)a(2)[\alpha(1)\beta(2)-\alpha(2)\beta(1)].\end{aligned}\right\} \tag{29}$$

These are the rigorous nomalized states, and it is easy to see that if $H=H'(1)+H'(2)$

$$\begin{aligned}(\psi_A{}^S|H|\psi_B{}^S)&=\tfrac{1}{4}\sqrt{2}[\alpha(1)\beta(2)-\alpha(2)\beta(1)]^2\\ &\quad\times[(a(1)b(2)+a(2)b(1))|H'(1)+H'(2)|a(1)a(2)]\\ &=\sqrt{2}(a|H'|b)\end{aligned} \tag{30}$$

The result of our prescription (8), however, is perfectly obvious since both states are singlets, obtained by the same process from the basic wave-functions (4), and thus

$$(t|u)=1 \tag{31}$$

and the matrix element is

$$(a|H'|b). \tag{32}$$

The first source of error is fairly obvious. The basic wave functions (4) in the case of configuration B are not orthogonal; one is

$$(1/\sqrt{2})a(1)a(2)[\alpha(1)\beta(2)-\alpha(2)\beta(1)],$$

and the other is

$$(1/\sqrt{2})a(1)a(2)[\alpha(2)\beta(1)-\alpha(1)\beta(2)],$$

and so the result of the unitary transformation (7) (in this case simply antisymmetrizing) is

$$a(1)a(2)[\alpha(1)\beta(2)-\alpha(2)\beta(1)], \tag{33}$$

which is not normalized, and in fact is too large by a factor of $\sqrt{2}$. Thus we predict that the matrix element (32) is too *large* by a factor $\sqrt{2}$ since it is computed on the basis of a non-normalized wave function. The existence of further orbitals in the configuration does not change this result, and we get the general rule: "For every pair of orbitals in either configuration which are identical in one but not in the other, our transition matrix element must be divided by $\sqrt{2}$ caused by the linear dependence of the basic wave functions."

This factor, unfortunately, is in the wrong direction and makes the error in (32) now 2 rather than $\sqrt{2}$.

The other factor of 2 we can find by looking at the steps in Eq. (30). We notice that the Hamiltonian depends symmetrically on the coordinates of electrons (1) and (2), and in fact is the sum of a term for each electron. Thus it happened that $a(1)a(2)$ combined with *both* $a(1)b(2)$ and $a(2)b(1)$. This is the source of the missing factor 2. What this means is that Eq. (6) of the text is in error when orbitals are identical. Actually, there are cases in which the basic wave-functions combine in case a pair of orbitals are equal. The wave-function

$$A[a(1)\chi_1(1)a(2)\chi_2(2)\cdots] \tag{34}$$

combines (if χ_1 and χ_2 are different, as they must be), both with

$$A[a(1)\chi_1(1)b(2)\chi_2(2)\cdots], \tag{35}$$

and, except for a changed sign,

$$(A7\text{-}b)\quad A[a(1)\chi_2(1)b(2)\chi_1(2)\cdots]. \tag{35$'$}$$

This is due to the fact that (34) is actually identical with the wave-function with χ_1 and χ_2 interchanged.

Fortunately, since only the singlet wave-function in these two orbitals can enter due to the fact that there are identical orbitals in one of the configurations, (35) and (35′) will always enter as an antisymmetrized linear combination, and the only result of the failure of formula (6) of the text is the missing factor of 2, due to the fact that the wave-function with identical orbitals combines with twice too many wave-functions in the other configuration.

Thus, to take account of these two effects the transition matrix element must be multiplied by $2/\sqrt{2}=\sqrt{2}$ for each pair of identical orbitals in either configuration which do not also appear in the other configuration. We can equally well simply multiply the total interaction, or if we like the projective operator, by 2.

The separate matrix elements referring only to upper or to lower configurations are not modified by the presence of identical orbitals, as was shown by Dirac.[9] The procedure of the text is thus shown to be correct.