REVIEWS OF MODERN PHYSICS VOLUME 25, NUMBER 1 JANUARY, 1953

Models of Exchange Coupling in Ferromagnetic Media

J. H. Van Vleck

Harvard University, Cambridge, Massachusetts

INTRODUCTION

IN contradistinction to superconductivity, the mechanism responsible for ferromagnetism is at present fairly well understood in broad outline. It is generally agreed that the exchange forces between electrons provide the coupling between the elementary magnets which is a prerequisite to ferromagnetism and which is so well represented empirically by the Weiss molecular field. The accurate solution of the eigenvalue problems for a complicated solid is, however, so complex that it is necessary to resort to some kind of simplifying model, for which at least approximate calculations can be made. Some of the more important models, each of which has its own fervent supporters, are as follows:

(a) the original Heisenberg theory.[1]

(b) the collective electron ferromagnetism of Stoner and Wohlfarth.[2]

(c) the generalized Heisenberg model, wherein there is a nonintegral number of spins per atom, and the spins are continually being redistributed among the different lattice sites. Model (c) differs from (b) in that the motions of the electrons instead of being uncorrelated, are so constrained that states corresponding to higher degrees of ionization are excluded. Hence we shall sometimes refer to this model as that of minimum polarity. For large interatomic distances or very narrow bands, (c) is approximately the correct physical picture, while for close spacings or wide bands, (b) is more suitable.

(d) Zener's theory.[3]

We shall not attempt to describe the details of (b) and (d), as they are the subject of other papers in this conference, and we shall assume that the reader is familiar with the rudiments of (a), which are presented in practically any book on magnetism. On the other hand, no elaboration of (c) is, as far as we know, published in the literature; the fullest description appears in an unpublished doctor's thesis of H. Hurwitz;[4] it has also been more briefly mentioned in articles by the writer.[5]

Zener's Theory

In Zener's model the exchange effects responsible for ferromagnetism are to be found mainly in the interaction between $3d$ and $4s$ electrons rather than between $3d$ electrons as ordinarily assumed. According to Zener, the $3d-4s$ coupling is so strongly ferromagnetic that it dominates the $4s-4s$ and $3d-3d$ interactions, which he assumes are antiferromagnetic. There are two difficulties which the Zener theory must surmount in order to be acceptable. One is that the $4s$ conduction band may be so wide that the promotion energy required to make $4s$ electrons parallel inhibits ferromagnetism. The other difficulty is a more subtle but less serious one. It can be shown that if the s band is half-full (one conduction electron per atom), it is possible to construct linear combinations of $4s$ wave functions such that the spin of the $4s$ band alternately points north and south, as one passes from one atom to another, and such that the translational kinetic energy is no higher than for the Zener model.[6] With this scheme one can make the spins of the $3d$ and $4s$ electrons parallel when they are on the same atom, and the spins of neighboring $3d$ electrons antiparallel. The implication of this possibility is that with the Zener hypothesis and a half-filled conduction band, the configuration of lowest

---

[1] W. Heisenberg, Z. Physik **49**, 619 (1928).

[2] E. C. Stoner, Phil. Mag. **21**, 145 (1936); Proc. Roy. Soc. (London) **A165**, 372 (1938); **169**, 339 (1939); J. Phys. et radium **12**, 372 (1951); E. P. Wohlfarth, Proc. Roy. Soc. (London) **A195**, 434 (1951); Phil. Mag. **42**, 374 (1951); G. C. Fletcher and E. P. Wohlfarth, Phil. Mag. **42**, 106 (1951).

[3] C. Zener, Phys. Rev. **81**, 440; **82**, 403; **83**, 299 (1951); **85**, 324 (1952).

[4] H. Hurwitz, thesis, Harvard University, 1941.

[5] J. H. Van Vleck, Revs. Modern Phys. **17**, 42 (1945); Physica **15**, 204 (1949).

[6] We shall indicate here briefly how the waves of spin reversal are constructed in two simple cases, *viz.*, the one-dimensional free electron model with periodic boundary conditions, and the Bloch approximation of tight binding. Consider first free electrons in a strip long enough to include $2n+1$ nuclei spaced distances d from each other. Then the $2n+1$ wave functions of lowest energy satisfying the boundary conditions are usually taken as $e^{-in\alpha x}, \cdots, 1, \cdots, e^{in\alpha x}$, with $\alpha=2\pi/(2n+1)d$. However, equivalent linear combinations are (neglecting normalization factors)

$$e^{-in\alpha x}+\cdots+1\cdots+e^{in\alpha x}=\sin[(n+\tfrac{1}{2})\alpha x]/\sin\tfrac{1}{2}\alpha x$$

and $2n$ other functions differing by integral multiples of d in origin. Each of these functions has a maximum at one particular nucleus and nodes at the others. The waves of spin reversal are constructed by assigning alternatively north and south spins to these functions, and, of course, antisymmetricizing. With the tight-binding approximation the construction is even easier, for if each state in the band is occupied once, the energy is the same as for individual atomic functions, and one simply alternates in the spin assignment to successive atoms.

The equivalence of the band energy with these rather unorthodox systems of wave functions with alternating spins to that with a conventional system is contingent on there being one electron for each orbital state or in other words on the s band being half-full. The writer unfortunately did not realize this limitation when he presented his paper at the Washington conference. It should also be mentioned that the alternating arrangement gives as favorable $s-d$ exchange energy as when the s-spins are all parallel only if the radius of the d shell can be regarded as small compared with the atomic volume, for in the staggered model the expectation value of the spin moment varies sinusoidally with position, and is zero mid-way between two atoms.

energy is antiferromagnetic rather than ferromagnetic. However, this difficulty is probably rather academic, as the *s* band is about 30 to 40 percent rather than 50 percent full in most ferromagnetic materials, and then the translational kinetic energy of the *s* electrons is lower for the parallel rather than the alternating arrangement.

The Difference in Polarity between Models (b) and (c)

We shall not pursue these questions further, and shall henceforth discuss the more orthodox theories (a-b-c) in which it is assumed that ferromagnetism results because of, rather than in spite of, the interaction of $3d$ electrons with each other. Some $3d-4s$ exchange, however, doubtless enters, and Zener can well claim that we should not overlook $3d-4s$ exchange completely. I think it can be regarded as agreed that the simple Heisenberg theory is too crude, since the experimental values of the saturation magnetization show clearly that there is not an integral number of spins per atom. We shall center our discussion on nickel, as it is probably the most important and typical case. It is generally recognized that it has 0.6 electrons per atom in the $4s$ band, and 9.4 in $3d$ states, so that symbolically the atomic configuration is $3d^{9.4}\ 4s^{0.6}$. The $4s$ conduction band is so wide that the $4s$ electrons are pretty well represented by the band theory based on itinerant, uncorrelated electrons. The question is how the $3d$ electrons are to be treated. In the Stoner-Wohlfarth model of collective electron ferromagnetism, the motions of the $3d$ electrons through the lattice are completely uncorrelated except for the constraints imposed by the Pauli exclusion principle, or in other words, no allowance is made for the fact that the Coulomb repulsions tend to keep electrons apart. Thus, although in nickel the average configuration of the *d* electrons on an atom is $d^{9.4}$ it is a mixture of configurations

$$d^{10},\ d^{9},\ d^{8},\ d^{7},\ \cdots.$$

The work required to tear off an electron increases rapidly with the degree of ionization. Hence the energy is appreciably lower if, for instance, two atoms with eighteen *d* electrons all told, are each in $3d^9$ states rather than one in $3d^8$ and the other in $3d^{10}$.

In the "forgotten model" (c), the states of higher ionization are arbitrarily ruled out completely, and the configuration $3d^{9.4}$ is considered to be

$$40 \text{ percent } 3d^{10} \qquad 60 \text{ percent } 3d^{9} \qquad \text{(model c).}$$

The lattice sites occupied by $3d^9$ and $3d^{10}$ are continually redistributing themselves. As a result there is a band structure and an excess specific heat, though neither is as large as with model (b). It should be particularly emphasized that the results of neutron diffraction show that the redistribution of the ions among the different lattice sites takes place in a time shorter than the time of passage of a neutron through an atom, so that the scattering pattern is the same as though each atom carried the same mean charge. Otherwise there would be a diffuseness in pattern due to fluctuation effects. Dr. Shull and Dr. Kittel point out to the writer that the passage time is of the order 10^{-13} sec, and consequently the band width cannot be much smaller than about 0.1 volt and still be consistent with the experiments on neutron diffraction.

The true state of affairs is in between (b) and (c). To investigate which of the two is closer to reality we must probe a little more quantitatively and examine the extent to which the collective electron model (b) involves different degrees of ionization. Here the estimates can be made only with certain limiting approximations, but fortunately even at the extremes the results are not widely different. The limiting approximations are obtained by assuming (1) that one particular *d* state or (2) that all five *d* states participate in the "effective band structure," i.e., the portion of the band over which electrons are redistributed in passing from the nonferromagnetic to the ferromagnetic condition or vice versa. According to G. C. Fletcher,[7] for instance, the *d* band has an over-all width of 2.7 volts, of which only the upper part is shown in our Fig. 1, reproduced from his paper. In the nonferromagnetic condition, the band is doubly populated (i.e., each direction of spin represented) out to the line labeled "Paramagnetic Ni," and is vacant thereafter. In the ferromagnetic condition the band is doubly populated out to the line labeled "Ferromagnetic Ni," and the balance of the band is singly populated. What we call the effective structure or width is essentially the interval to the right

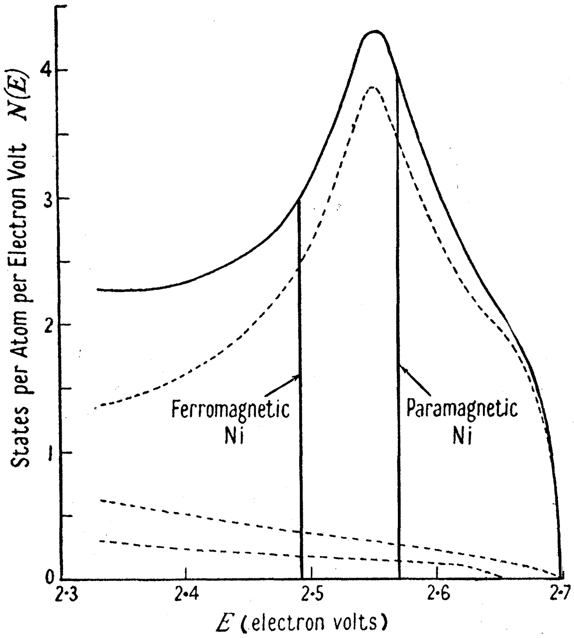

FIG. 1. Density curve for the upper part of the $3d$ band of nickel, according to Fletcher.

---

[7] G. C. Fletcher, Proc. Phys. Soc. (London) **A65**, 192 (1952).

of the line labeled "Ferromagnetic Ni" in Fig. 1. It is obviously much less than the over-all band width. If all the different energy levels in this interval are based on a single d state, than the highest degree of ionization which is obtainable is d^8, as all the four d states other than the given d state are completely filled and are effectively *hors de combat* as far as spin reversals are concerned. On the other hand, the substates of north and south pointing spins for the given d state can be regarded as completely uncorrelated in the unmagnetized condition, and from this fact it is readily shown that the fractions of the time that a nickel atom is in the configurations d^{10}, d^9, d^8 are, respectively, the coefficients of x^0, x, x^2 in

$$(0.7+0.3x)^2. \tag{1}$$

The numbers 0.7 and 0.3 owe their existence to the fact that in nickel the number of d electrons per atom is 9.4, and so for each spin direction the fifth d state, corresponding to the most loosely bound electron, is occupied 70 percent of the time. From (1) we find that the percentages work out as follows:

$$49 \text{ percent } d^{10},\quad 42 \text{ percent } d^9,\quad 9 \text{ percent } d^8. \tag{1$'$}$$

The other limiting case is obtained by assuming that all five d states participate in the effective band structure. The exclusion principle involves restrictions only on electrons in the same d state. Consequently the fractional percentages of atoms in the configurations d^{10}, d^9, $d^8 \cdots$ are, respectively, the coefficients of x^0, $x \cdots$ in the expression

$$(0.94+0.06x)^{10}, \tag{2}$$

or in other words

$$53.8 \text{ percent } d^{10},\quad 34.4 \text{ percent } d^9,\quad 9.9 \text{ percent } d^8,\quad 1.7 \text{ percent } d^7,\quad 0.2 \text{ percent } d^6. \tag{2$'$}$$

The percentages for d^5, d^4, $\cdots$ are negligible.

If, as Fletcher and other writers have done, one uses the approximation of tight binding, there is a secular determinant of degree 5 for each direction of the propagation vector. One particular root depends but little on the direction of propagation, and is responsible for the hump in Fig. 1. If, regardless of the direction of the propagation vector, the significant root involves one and the same combination of d wave functions, the relevant estimates of polarity are those given in (1) and (1′). If, on the other hand, five different orthogonal linear combinations are appropriate, respectively, to one-fifth each of the possible choices of the propagation vector in the critical region, then the correct estimates are those involved in (2) and (2′). Actually, the appropriate linear combinations vary continuously rather than discretely with the choice of the propagation vector, and the actual amount of polarity with the collective electron model is doubtless intermediate between (1) and (2), but somewhat closer to (2) than (1).

Comparison of Energies of Models (b) and (c)

If assumption (1) is used the excess in polar or correlation energy (i.e., energy of interelectronic repulsion) of model (b) over model (c) is

$$0.09[E(d^8)-E(d^9)]-0.09[E(d^9)-E(d^{10})], \tag{3}$$

as is seen in subtracting from (1′) the percentages of polarity (40 percent, d^{10}, 60 percent d^9) for model (c). Here $E(d^8)$ denotes the energy for the configuration d^8 (really $3d^84s^{0.6}$), etc. If (2′) is appropriate, we have instead for the excess

$$\begin{aligned}&+0.099[E(d^8)-(d^9)]+0.017[E(d^7)-E(d^9)]\\&\quad+0.0018[E(d^6)-E(d^9)]-0.140[E(d^9)-E(d^{10})].\end{aligned} \tag{4}$$

We have, for convenience, grouped terms in such a way as to use $E(d^9)$ as essentially an origin of energy. The energies of the higher ionized configurations d^7, d^6, etc., are not known spectroscopically (except for d^7), but surely

$$E(d^{9-n})-E(d^9)>n[E(d^8)-E(d^9)],\quad n=2, 3. \tag{5}$$

We will be underestimating the excess energy, i.e., favoring model (b), if we replace the inequality by an equality sign in (5). Then (4) becomes simply[8]

$$0.138[E(d^9)-E(d^9)]-0.138[E(d^9)-E(d^{10})]. \tag{6}$$

To evaluate (4) or (6) numerically it is not fair to use the spectroscopic energies of the ordinary nickel atom or ion in the configuration $\mathrm{Ni}(d^{10})$, $\mathrm{Ni}^+(d^9)$, $\mathrm{Ni}^{++}(d^8)$, $\cdots$ since the d electrons are partially screened by the 0.6 s electrons. The screening influence of an s electron in the solid state is somewhat greater than in the free ion, as the mean charge per electron is, on the average, compressed into the atomic volume. It is for orders of magnitude, sufficient to assume that in the solid state, the 0.6 s electrons per atom have the same screening action as 1.0 s electrons do in the gas. The relevant energies for our sequence d^{10}, d^9, d^8 are then

$$\mathrm{Ni}^-(3d^{10}4s),\quad \mathrm{Ni}(3d^94s),\quad \mathrm{Ni}^+(3d^84s).$$

Even so, the energy of the negative nickel ion is not known spectroscopically, but can be obtained, though only very crudely, by extrapolation by means of the irregular doublet law, i.e., the assumption that the square root of the binding energy is a linear function of the screening constant. In this fashion one finds that

$$[E(d^8)-E(d^9)]-[E(d^9)-E(d^{10})]=7.6 \text{ volts}. \tag{7}$$

The expressions (3) and (6) are then, respectively, 0.7 and 1.1 volts. Fortunately, for our purposes an error of a volt or two in the estimation of $E(d^{10})$ is of little consequence. If one made no allowance at all for the

---

[8] With the assumptions made, both coefficients in (6) are exactly equal in numerical value, and equal to the excess of the fraction of d^{10} in (2′) over 0.40. This statement is readily shown to be a consequence of the fact that the mean deficit in charge per atom is 0.60e.

screening effect of the 0.6s electron, the value of the expression (7) would instead be 12.5 volts.[8a]

We must now evaluate for comparison how much increase in energy model (c) involves in comparison with (b) because (c) does not have as wide a band structure. If a single d state is assumed to enter it is easy to obtain an approximate estimate for the loss in banding energy. Namely, the degree of polarity of the collective electron model in the ferromagnetically saturated state is then the same as that of model (c), regardless of the degree of magnetization in the latter. Furthermore, the mean square energy dispersion can be shown to be the same in the model (c) as in the ferromagnetic state of model (b). It is therefore natural to assume that model (c) in any state of magnetization has the same band width as model (b) in the ferromagnetic state. This is all on the assumption (1) of a single effective d state, so that the exclusion principle prohibits two co-directed spin-holes on the same atom. From inspection of Fig. 1 one sees that the difference in band energy between the paramagnetic and ferromagnetic states in the collective electron model is about 1.4×0.05 or 0.07 electron volts, only one tenth of the difference (7) in correlation energy.

If instead the assumption involved in (2′) concerning the occupancy of d states is made, the ferromagnetic state in the collective electron model still involves more polarity than the generalized Heitler-London model. The various degrees of polarity correspond to the different coefficients of x in $(0.88+0.12x)^5$, whence the percentages are found to be

$$52.7 \text{ percent } d^{10},\quad 36.0 \text{ percent } d^9,\quad 9.8 \text{ percent } d^8,\quad 1.3 \text{ percent } d^7,\quad 0.1 \text{ percent } d.^6$$

With the approximation of using the equality sign in (5), the excess of polar energy in the paramagnetic as compared with the ferromagnetic collective electron model is

$$[0.138-0.127][\{E(d^8)-E(d^9)\} -\{E(d^9)-E(d^{10})\}]=0.08 \text{ volt}.$$

On the other hand, the decrease in band energy in the paramagnetic as compared with the ferromagnetic state is only 0.07 volt. Thus elimination of excess polarity still pays off, though not as much as with the assumptions (1−1′). Extrapolation to zero excess polarity indicates that model (c) as compared with model (b) has 1.0 volt higher band energy, but 1.1 volts lower energy of interelectronic repulsion, and so wins out, though by a slim margin.

This method of extrapolation is none too reliable, but another approach is available. Let us, for rough purposes of estimation, assume the band shape is rectangular. If one assumed that model (c) had no band energy at all, the decrease in energy, due to band effects, in passing from model (c) to model (b) would be $\frac{1}{2}\times 0.06\times 2.7\times 9.4=0.75$ volt. This figure is obtained by noting that the lowest 94 percent of the band is filled in the paramagnetic state by the 9.4 d electrons, and that in first approximation the center of the band coincides with where it would be in the absence of any banding. We have used Fletcher's value 2.7 volts of the over-all band width. However, these estimates based on the assumption that the generalized Heisenberg model (c), has no band structure at all are unduly unfavorable to this model. Actually this model has an appreciable band structure, because of the fact that the positions of the d^9 and d^{10} sites are continually being interchanged. Unfortunately, the detailed band structure cannot be feasibly calculated, but estimates of orders of magnitude can be made from examination of the root-mean-square dispersion in energy, which can be readily computed, thanks to the invariance of the diagonal sum. One finds that with the assumption of a single effective d state, the r.m.s. band width with model (c) is about 70 percent as wide with model (c) as with model (b). (More precisely, the percentage is 75 with assumption (1′) and 66 with (2′).[9] If the rectangular assumption concerning band shape is applicable to both models (b) and (c), the difference in band energy between the two models amounts to 0.3×0.75 or only 0.2 volt, about a fifth of the difference (7) in correlation energy.

With any of our various ways of estimation, the improvement in correlation energy in (c) *vis-a-vis* (b) more than offsets the contrary change in band energy. Using the Ritz variational principle, we see that the true state of affairs, while undoubtedly intermediate between (b) and (c), is probably closer to (c) than (b). Without relying on detailed numerical estimates, one can see qualitatively that model (b) is more or less between Scylla and Charybdis. If the band width is too

---

[8a] *Note added in proof.*—Dr. C. Herring points out to the writer that in reality the screening by the 4s electrons is progressively greater for the configurations d^{10}, d^9, d^8, since more negative charge is sucked in, the greater the positive charge of the lattice site. This effect tends to diminish the difference (7). A limiting case, doubtless too extreme, is one in which each unit cell is neutral, and there is simply a redistribution of configurations $3d^{10}$, $3d^94s$, $3d^84s^2$ among the different lattice sites. Instead of (7), the relevant energy difference in determining the occurrence of the d^{10}, d^9, d^8 configurations is then $E(3d^84s^2)+E(3d^{10})-2E(3d^94s)$, which in nickel amounts to about 3 volts when one averages over the different multiplets belonging to each configuration. Such a model, wherein the d and s migrations compensate so as to preserve neutrality, cannot be considered as either (b) or (c). It goes even further than our generalized Heitler-London model in ironing out fluctuations in polarity, but may not suppress the d^8 states as completely are in our model (c). Possibilities such as Herring suggests show that too much significance should not be attached to quantitative comparisons of models (b) and (c), as the actual situation is even more complicated, but at the same time essentially support our main contention that in dealing with d shells correlation effects are far more important than in ordinary conductors, and that the ordinary band picture is at best a crude approximation.

[9] The rms band width is proportional to the square root of the probability that an electron on an adjacent lattice-site makes a transition to a hole on some given atom. With assumption 1 (one effective d-state), this probability is proportional to 2×0.40 with model (c) and to $2\times 0.49+0.42$ with model (b) (cf. Eq. (1′)). With assumption 2 (five effective d-states), the probability is proportional to 10×0.40 with model (c) and to $10\times 0.538+9\times 0.344+8\times 0.099\cdots$ with model (b) (cf. Eq. (2′)). In model (c) an electron cannot jump away from a d^9 site, and hence only the d^{10} sites contribute to the transition probability.

wide, ferromagnetism will be suppressed, and the model as actually used is close to this limit. On the other hand, if the band is narrowed, the correlation effects will be so large compared to the band structure that (c) will be closer to reality than (b).

Computational Difficulties with Model (c)

There is, however, one practical difficulty with model (c), and that is the fact that there is apparently, at least as yet, no feasible way of making good numerical calculations with it, and in this respect model (b) is undoubtedly superior. Even the original over-simplified Heisenberg model has computational difficulties, as Slater has pointed out in an earlier paper in this conference. From a practical standpoint, if the overlap integrals are small, a fairly good answer, at least qualitatively, appears to be obtainable by neglecting the nonorthogonality except in the primary exchange integrals, but a rigorist can well find all sorts of objection to this procedure. A long time ago, I pointed out in a semiquantitative way that it was plausible[10] that the nonorthogonality corrections largely canceled out. The following example shows, I think, that the nonorthogonality effects cannot be as bad as the pessimist would make them out. Consider, for example, two hydrogen atoms banded together to form a hydrogen molecule in a large box full of widely separated hydrogen atoms. We know that the binding energy of this molecule can be computed as a two-atom problem. However, the nonorthogonality corrections coming from the remote hydrogen atoms, though infinitesimal, can be made, by taking the box sufficiently large, to spoil the convergence and hence the apparent rigor of the calculation made in the standard way—a rather absurd state of affairs. Consequently, it appears to us probable that the convergence can be greatly improved by using a different grouping of terms or iteration procedure.

Even if one adopts an optimistic point of view as regards the effects of nonorthogonality, calculations with model (c) are apparently not feasible at low or even ordinary temperatures. At very high temperatures, large compared to the Curie point, a series development of the partition function may be utilized. This development has been studied by H. Hurwitz.[4] Model (c), like model (b), involves besides the two-electron exchange integral J, the one-electron "transfer" or "hopping" integral K associated with the transfer of an electron from one lattice site to another:

$$J=\int\cdots\int\psi_A(1)\psi_B(2)H\psi_A(2)\psi_B(1)dv_1dv_2,$$

$$K=\int\cdots\int\psi_A(1)H\psi_B(1)dv_1.$$

Hurwitz shows that K first affects the molecular field in the approximation K^3, whereas J enters already in the first power. Unfortunately, K is doubtless much larger than J, so that development in powers of K/kT has even worse convergence than that in J/kT, which, though slow, is better than once supposed.[11]

One thing stands out clearly. That is that such phenomena as curvature of the $1/\chi-T$ graph above the Curie point, excess specific heat linear in T at low temperatures, nonintegral numbers of magnetons per atom, different magneton numbers deduced from saturation and from the susceptibility at high temperatures, cannot be construed as evidence favoring model (b) over (c), as is often done. It is true that except perhaps for the curvature effect these phenomena are not present with the oversimplified model (a), but they all can exist in (c), because the mean configuration is $3d^{9.4}4s^{0.6}$, and there are continual fluctuations in the way the d^{10} and d^9 states are distributed among the different lattice sites. Hurwitz shows that the specific heat already enters in the approximation K^2/k^2T^2, but, of course, a series development of this character is irrelevant at low temperatures.

Origin of the Exchange Integral

We now turn to the mooted question of the origin of the positive[12] exchange integral necessary for ferromagnetism. Two possible mechanisms can be debated. One is the original Heisenberg hypothesis that the integral has an inter-atomic origin, i.e., arises from the exchange of electrons on different atoms, but neighbors of each other. The other mechanism, suggested by Slater[13] in 1936, is that the positive integral has an intra-atomic origin, and is a reflection of the fact that in the configuration d^8 the triplets are deeper than the singlets. We might, incidentally, mention that in general our views are not so very different from those of Slater in his 1936[13] and 1937[14] papers. In particular, his 1937 article[14] had a calculation of the energies for the states of a crystal of one reversed spin that indicated that the deepest state is essentially one of type (c) rather than (b), but Slater has not pushed this viewpoint subsequently.

Whether interatomic exchange really gives a positive integral, rather than a negative one such as serves as the common basis of valence theory, is a question which has been argued pro and con in the literature and is one into which we do not propose to enter. The numerical calculations necessary to yield a definitive answer are exceedingly difficult. The most recent effort is that of Kaplan,[15] who comes out with a positive integral, but even so the results, though definitely en-

---

[10] J. H. Van Vleck, Phys. Rev. **49**, 232 (1936).

[11] V. Zehler, Z. Naturforsch. **5a**, 349 (1950).

[12] The interatomic exchange integrals of which we speak are those computed with ordinary atomic orbitals. The integrals computed with the orthogonal Wannier functions are always positive, but are not germane to our discussion.

[13] J. C. Slater, Phys. Rev. **49**, 537, 931 (1936).

[14] J. C. Slater, Phys. Rev. **52**, 198 (1937).

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

couraging, are inconclusive because the accurate computations were not made at the internuclear distance appropriate to actual ferromagnetic material.

The idea that the exchange responsible for ferromagnetism has an intra-atomic origin has many elements of attractiveness. In the first place, if correct it makes it unnecessary to invoke interatomic effects unlike those in standard valence theory. In the second place, it provides an easy explanation as to why the transition elements are ferromagnetic, whereas the alkalis and alkaline earths are not. Namely, the intra-atomic mechanism requires degenerate orbital states, as we will explain in more detail, and this degeneracy is missing in s bands. (On the other hand, an explanation of why conductors with p bands are not ferromagnetic is provided only after more detailed quantitative examination, inasmuch as p levels have spatial degeneracy.) We should mention that Zener's theory also has an intra-atomic origin of the exchange integral, but with the exchange between s and d rather than between degenerate d states, such as we are discussing here.

It is not difficult to estimate the size of the molecular field constant (the manifestation of the exchange integral) which is to be expected from the collective electron model if one uses approximations similar to those employed in (1) or (2). With (1) the answer is particularly simple: there can be no ferromagnetic coupling at all due to interatomic exchange, as two electrons can never be in a triplet state on an atom, if only one state is relevant. With (2), the molecular field constant θ' has a value approximately 1750 cm^{-1} or 2500°. (We follow the usage of Stoner and Wohlfarth in denoting the field constant by θ' and in evaluating it in degrees rather than cm^{-1} or volts; $2\theta'$ is the work required to reverse an individual spin against this field.) This estimate is reached as follows: In the collective electron model, the difference ΔE in exchange energy between the unmagnetized state and that of maximum magnetization is $\Delta E=-\frac{1}{2}\eta\theta'$, where $\eta=0.6N$ is the number of holes. On the other hand, if ΔE arises solely from intra-atomic effects, the value of ΔE must also be $\Delta E=-\frac{1}{2}JNp$, inasmuch as two electrons have exchange energy $-J$ if their spins are parallel, and on the average $-\frac{1}{2}J$ if their spins are uncorrelated. Here p denotes the fraction of lattice sites which are in the configuration d^8. The numerical value of θ' follows from comparison of the two expressions for ΔE on taking $p=0.14$ as in Eq. (6), and $J=7500$ cm^{-1}. This value of p is one obtained by making assumptions for the correction of the influence of states of higher polarity d^7, d^6, etc., analogous to those used in reducing (4) to (6). For the effective exchange integral J, Slater obtained the value 7430 cm^{-1} on averaging over the various d states in the configuration $3d^84s^2$. As noted earlier in the paper, we assume that the screening effect of the 0.6 $4s$ electron in the solid state is approximately the same as in the gaseous configuration $3d^94s$. However, the electrostatic intervals separating the various states of $3d^84s$ are little greater than those of $3d^84s^2$, and so we use the round value 7500 cm^{-1} for J. We should mention that Slater long ago[13] estimated the molecular field constant by using considerations somewhat similar to those of the present paragraph, but we believe that our value is a somewhat more refined one, and we are unable to agree with certain quantitative details of his calculation.

A molecular field constant of 2500° which we deduced above for the collective model is more than adequate to give the proper ferromagnetic behavior. It is larger than the value 1230° obtained by Wohlfarth from observed susceptibility data on the assumption of a parabolic band, or the value 1760° deduced by Wohlfarth and Fletcher[2] with rectangular bands. In our opinion, it is probable that the actual molecular field constant is somewhat smaller than either of these estimates. In the first place, the actual band shape in the collective electron model apparently has a sharp peak near the upper end, as shown in Fig. 1, narrowing the effective band width and hence reducing the molecular field necessary to produce ferromagnetism. In the second place, and probably more important, the correlation effects arising from interelectronic repulsion alter the band structure, as we will amplify below.

If the correlation effects are so large that the model (c) of minimum polarity is relevant, then clearly the molecular field should be calculated with this model. At first glance, any attempt to provide an intra-atomic mechanism for this field with (c) appears hopeless, as the essence of (c) is that states of higher polarity are absent. Hence, in the model in its purest form, there are no states at all of structure d^8, d^7, etc., and hence no possibility of favorable intra-atomic exchange energy as a result of electron spins being parallel in unclosed shells. However, model (c) is an ideal limit, and the true wave function, which is a linear combination of everything in sight, doubtless contains some terms of structure d^8, d^7, etc. In other words, model (c) is never completely pure and is always contaminated by some states of higher polarity as, so to speak, a concession to model (b). It is only this contamination which is responsible for intra-atomic exchange. At first one wolud guess that any such contamination would be inadequate, since we have seen that even in model (b) the molecular fields calculated with the intra-atomic mechanism are not a great deal larger than those required by the observed ferromagnetism, and so it would seem that any second-order effect would be insufficient. This is not really the case, for with model (b) there is the familiar band structure which tends to inhibit parallel aligning of the spins and so makes the molecular field constant much larger than it would be otherwise, for it must be sufficient to overpower the unfavorable band structure. We have emphasized that with model (c) there is also a band structure, and, in fact, the so-called excess specific heat is probably of almost the same order of magnitude with (c) as with (b). However, the band structure is of a basically different type with (c) than with (b), being no longer an uncorrelated one-electron affair, and the band energy (i.e., the energy including the transfer integral K but not the exchange integral J) depends much less drastically on spin alignment than in (b). The spin first begins to enter when third powers of K are included, and it is not even clear whether the band structure tends to favor states of high or low spin, though probably the latter in a weak way. Since in (c) the inhibiting effect of the band structure on the spin is to a large extent removed, it is probably a better approximation with (c) to calculate the molecular field as though the band structure were completely absent (essentially the simple Heisenberg theory with $0.6N$ atoms having $s=\frac{1}{2}$) than to try and take over calculations based on model (b). The molecular field constant is then the same as the Curie constant, and hence only 630° in nickel.

It appears likely that the contamination by states of higher polarity can give a molecular field of this order of magnitude in model (c). This is indicated by a skeleton perturbation calculation which we present in the next paragraphs.

Rough Calculation of Perturbation by Higher Polar States in Model (c)

Consider a pair of neighboring atoms in the configuration d^9+d^9. We wish to examine the effect of perturbations by d^8+d^{10} or $d^{10}+d^8$. For brevity, we do not include the s conduction electrons in our notation, and write d^8, d^9, d^{10} for d^8s, d^9s, $d^{10}s$. In dealing with the spatial degeneracy we shall assume that the transfer integral K associated with the transition of an electron from a d state of one atom to the d state of its neighbor is unaffected by which particular spatial substate one selects. In other words we assume that transitions are equally probable from any given substate of atom A to all five substates of B. This is an extreme form of the assumption (2) explained earlier in the paper, but is sufficient for estimating orders of magnitude. A straightforward application of second-order perturbation theory shows that when the disturbing influence of the excited states d^8+d^{10} and $d^{10}+d^8$ is considered, the excess of the energy of singlet ground state over the triplet is

$$ {}^1E-{}^3E=2K^2\left[\frac{4}{h\nu}-\frac{6}{h\nu+\Delta}\right]. \tag{8}$$

Here $h\nu$ denotes the mean excess of energy of the triplet states of d^8+d^{10} over the energy of d^9+d^9 and Δ is the mean separation of the singlet and triplet states of the d^8 ion (including screening correction for the 4s electron). The multiplicity of the configuration d^8+d^{10} is, of course, the same as that of the d^8 ion. The coefficient of the term in $1/h\nu$ is 4 rather than 5 (the degree of spatial degeneracy) because the d^8 ion can be only in a singlet when the two "holes" are in the same spatial substate, and then is incapable of perturbing the triplet ground level. The coefficient of $1/(h\nu+\Delta)$ is 6 because the perturbation of the singlet ground level is correspondingly augmented. (If there were one rather than five effective d substates, the coefficients would be 0, 2 rather than 4, 6, and the expression (8) would necessarily be negative, showing that if assumption (1) rather than (2) were used, the coupling resulting from intra-atomic exchange is antiferromagnetic as we saw was also the case in model (b).) For the coupling to be ferromagnetic it is necessary that the expression (8) be positive. The fact that (8) is negative if the exchange separation Δ in the upper state is zero or small is a manifestation of the fact that the perturbation by polar states is making an incipient collective electron band theory out of the original model starting with "frozen" d^9 lattice sites; and, of course, in a collective electron scheme without exchange, the states of lowest multiplicity are lowest because the band need not be filled so high. Our perturbation calculation does not include correction for the fact that the d^9 lattice sites are migratory even in model (c); this gives a fine structure in the ground state in addition to the one we are computing here. The superposition of the interchange of d^9 and d^{10} sites in model (c) would make a rigorous calculation of the perturbing influence of the upper polar states exceedingly difficult, but for estimating orders of magnitude the present procedure appears adequate.

To evaluate (8), we take $K=0.1$ volt, as this is the order of magnitude indicated, for instance, by Fletcher's paper.[7] The separation Δ is twice the exchange integral for the upper state, which we previously estimated as 7500 cm^{-1}=0.9 volt. Assumptions similar to those used in connection with (6) suggest[16] that $h\nu=7.6$ volts, but there is a correction arising from the fact that the configuration d^9+d^9 has an extra electrostatic energy e^2/r, where r is the separation of neighboring ions. As a result $h\nu$ is reduced to about 1.7 volts and the expression (8) is approximately 0.01 volt.

The existence of a nonvanishing expression (8) is equivalent to the existence of a coupling $-\alpha\mathbf{s}_A\cdot\mathbf{s}_B$ between the two d^9 ions under consideration The expression (8) is in fact, identical with the coefficient α in the vector model. Thus ferromagnetism will result if α is positive. The coefficient θ' of the molecular field is $\theta'=\frac{1}{4}z\alpha$. The number z of neighbors we must not take as 12, the number of nearest neighbors in a face-centered lattice, but rather as $0.6\times12=7.2$ inasmuch as only 60 percent of the lattice sites have the d^9 holes. The value of the constant θ' according to our rough perturbation calculation is thus 200°, almost the right order of magnitude. So the traces of higher polarity in model (c) may be able to yield an adequate molecular

---

[16] The energy interval computed in (7) relates to $h\nu$, and not say $h\nu+\frac{1}{2}\Delta$, for in computing (7) we "lumped" all the states of the configuration d^8 at the ^3F level, the lowest state. This approximation, incidentally, slightly favors model (b) in the comparison of the energies of models (b) and (c) earlier in the paper.

field purely on the basis of intra-atomic exchange coupling. It should, however, be cautioned that this calculation of the molecular field is only a rough one, and should be taken *cum grano salis*. It is much more sensitive to the precise values selected for the energy intervals[17] than the computations earlier in the paper of the relative correlation energies of model (b) and (c), which are much more reliable. The only reason that we came out with a positive field was the drastic lowering of the frequency denominators in (8) because of the corrections to the energy because d^9+d^9 had a higher ionic energy e^2/r than d^8+d^{10}. One at first wonders whether there should not be similar corrections in our comparisons earlier in the paper of the relative energies of models (b) and (c). This, however, is not the case, for the essence of model (b) is that there is no correlation between the motion of the various electrons, so that a positive ion has no extra probability of having a closed shell on neighboring sites even though the electrostatic energy is thereby lowered. Actually there is doubtless some correlation, but this is not pure model (b), but rather a tendency in the direction of (c).

It should also be noted that any comparison of molecular fields calculated on the basis of an intra-atomic mechanism with those found empirically in ferromagnetic media is uncertain, because actually the intra-atomic and interatomic exchange effects are additive. The corrections for interatomic contributions cannot be reliably made for, as we have already emphasized, neither the order of magnitude of the interatomic exchange integrals, nor even their sign, is really known at present.

In summary, it may be said that the results of the present paper are rather discouraging, as the essence of what we have said is that the truth is somewhere between the collective electron model (b) and the generalized Heisenberg (c), probably closer to (c) than (b). Unfortunately, it is much more feasible to make detailed calculations with (b) than with (c). The calculations with the collective electron model have an elegance and clear-cut character which stand in sharp contrast to the crude qualitative estimates which are the best which can be done at present with (c). The gist of this paper is that it would be highly desirable if good methods of computation with (c) could be developed, including some means of circumventing the convergence difficulties associated with nonorthogonality. At present, for definite numerical answers it is certainly necessary to use model (b) rather than (c). However, this fact does not imply that (c) is physically unsound, as some writers seem to have done, and computational difficulties, important as they are, should not obscure the recognition in principle of the situation which conforms closest to physical reality.

---

[17] The configuration d^8 actually has two different triplet and three different singlet levels. In (8), the two triplets were assumed to coincide, with a similar approximation for the singlets, but a calculation as gross as the present one does not warrant the use of additional frequency denominators.

SUMMARY OF THE DISCUSSION

R. Smoluchowski, *Carnegie Institute of Technology, Pittsburgh, Pennsylvania:* It is perhaps advisable to make a brief appraisal of the relative position of the various theories presented in the preceding four papers and to summarize the main points brought up in the discussion. As it often happens in theoretical physics the differences between the theories stem not from basically different concepts of the phenomenon of ferromagnetism. but rather from a different "taste" or philosophy of theoretical approximation.

Slater's method of attacking the full problem with all necessary wave functions orthogonalized and solving it by means of numerical methods such as a digital computer is probably the most complete and quantum-mechanically exact procedure. The difficulties in handling an actual ferromagnetic material in that manner seem, at the present time, to be enormous. On the other hand, the success in the treatment of the paramagnetic oxygen molecule is most promising and it provides strong arguments in favor of the molecular orbitals method. It supports, also, the well-known dependence of the balance between kinetic and exchange energies upon the distances between the 3-*d* shells. The estimate of this dependence was until now based on rather qualitative arguments.

It is natural that the difficulties of the direct procedure make all approximate methods, which use intuitive concepts, of paramount interest and importance. Van Vleck considers the excited states of atoms in the crystal lattice and allows in this way for nonintegral number of spins per atom. In first approximation, the nonorthogonality corrections are considered to be negligible. This generalized form of the Heisenberg approximation appears capable of explaining many facts which were not understandable in terms of the classical Heisenberg model. If successful, it will be a very useful method, and at the same time it will allow the application of intuitive physical reasoning and will give insight into the physical significance of the mathematical procedures.

The collective electron theory developed by Stoner and Wohlfarth is based on a set of assumptions and parameters, one of them being the exchange energy. The calculations are relatively simple and the success in correlating various ferromagnetic and other properties is indeed striking. The effort is now directed towards analyzing the assumptions and the important quantities in terms of other more basic properties. This is particularly true of the exchange energy and of