

On the theory of ferromagnetism

By W. Heisenberg in Leipzig

With 1 Figure. (Received on 20 May 1928)

Translated by D. H. Delphenich

WEISS's molecular forces will be attributed to a quantum-mechanical exchange phenomenon, and indeed, it will be treated as the exchange process that was successfully enlisted in recent times by HEITLER and LONDON in order to interpret homopolar valence forces.

Introduction. Ferromagnetic phenomena have been interpreted in a formally satisfying way by the well-known WEISS theory (*). That theory is based upon the assumption that every atom in a crystal experiences a directed force from the remaining atoms of the lattice that should be proportional to the number of already-directed atoms. By contrast, the origin of these atomic fields was unknown, and several obstacles stood in the way of any interpretation of the WEISS forces on the basis of classical theory: *Magnetic* interactions between atoms are always a few orders of magnitude smaller than the atomic fields that follow from ferromagnetic experiments. Indeed, *electric* interactions lead to the correct order of magnitude; however, one would rather expect that the electrical interactions of two atoms would be proportional to the square of the cosine of their mutual angle of inclination, rather than the cosine, which contradicts the assumptions of WEISS's theory. Other complications were discussed more thoroughly by LENZ (**), and ISING (***) succeeded in showing that the assumption of directed, sufficiently large forces between any two neighboring atoms of a chain did not suffice to generate ferromagnetism.

The ferromagnetic complex of questions has entered a new arena with the UHLENBECK-GOUDSMIT theory of spin electrons. In particular, it follows from the known factor $g = 2$ in the EINSTEIN-DE HAAS effect (which was, in fact, measured for ferromagnetic substances) that in a ferromagnetic crystal only the magnetic eigenmoment of the electrons is oriented, but not, by any means, the atoms. Thus, the possibility of interpreting the WEISS forces as electrical interactions, independent of the relative spin directions of the electrons, goes away, since we know that such forces do not exist. Furthermore, by applying PAULI-FERMI-DIRAC statistics, PAULI (****) has been able to show that paramagnetism or diamagnetism will always result from neglecting the interaction of the electrons in a metal.

(*) P. WEISS, Journ. de phys. (4) **6** (1907), 661 and Phys. Zeit. **9** (1908), 358.

(**) W. LENZ, Phys. Zeit. **21** (1920), 613.

(***) E. ISING, Zeit. Phys. **31** (1925), 253.

(****) W. PAULI, Zeit. Phys. **41** (1927), 81.

§ 1. A model for the foundations of the theory. The basic idea of the theory that we seek here is this: Empirical results exhibit ferromagnetism as an entirely similar state of affairs to what was previously observed in the spectrum of the helium atom. At the time, it seemed to follow from the levels in the helium atoms that a powerful interaction prevailed between the spin directions of two electrons that led to the splitting of the level structure into systems of singlets and triplets. At the time, this difficulty could be resolved by verifying that the apparently large interaction would emerge indirectly from a resonance or exchange phenomenon that would be characteristic of all quantum-mechanical systems of identical particles. This is also closely related to *explaining ferromagnetic phenomena as being implied by this exchange phenomenon*. We will attempt to show that the COULOMB interaction, together with the PAULI principle, will succeed in evoking the same effects as the molecular field that WEISS postulated. It was only in recent times that mathematical methods were developed for the treatment of such a complicated problem in the important investigations of WIGNER (*), HUND (**), HEITLER and LONDON (***)).

Before I go into the actual calculations, I would like to give a brief overview of the methods of approximation that can come to be applied in the treatment of electronic motions in metals.

Method. I. From PAULI (*loc. cit.*) and SOMMERFELD (****), in the first approximation, electrons can be assumed to be completely free. In the second approximation, one might, perhaps, add in the interactions with the lattice points [HOUSTON (†)]. The interaction of electrons with each other is neglected completely.

Method II. As a first approximation, one calculates the motion of an electron in a force field (that, by no means, needs to be small) that is periodic (in three directions). In the next approximation, one might perhaps consider the perturbations that arise from the deviations from periodicity in the lattice. The treatment of the interaction of electrons with each other encounters the same difficulties here as it does in method I.

Method III. In the first approximation, one thinks of the lattice separations as being very large and assumes that every electron thus belongs to its own atom. In the next approximation, one considers the exchange of electrons that move in the unperturbed system with *equal* energies at *different* points, which was first considered by HEITLER and LONDON (*loc. cit.* I). States in which *more* electrons are found in comparison to the number that is found in *one* atom in the unperturbed system *will not be considered* in this approximation.

The difference between these three methods becomes clearer when we explain it by another example, namely, the hydrogen molecule, which was treated rigorously by

(*) E. WIGNER, *Zeit. Phys.* **40** (1927), 883; **43** (1927), 624.

(**) F. HUND, *ibid.*, **43** (1927), 788.

(***) W. HEITLER and F. LONDON, *ibid.* **44** (1927), 455, cited as I in what follows; W. HEITLER, *ibid.* **46** (1927), 47 (cited as II); *ibid.*, **47** (1928), 835 (cited as III); F. LONDON, *ibid.*, **46** (1928), 455.

(****) A. SOMMERFELD, *ibid.*, **47** (1928), 1; cf., also W. V. HOUSTON, *ibid.*, **47** (1928), 33, and C. ECKART, *ibid.*, **47** (1928), 38.

(†) W. V. HOUSTON, *ibid.* **48** (1928), 449.

HEITLER and LONDON (*loc. cit.* I). In method I, the electrons were, once more, first treated as free, which would naturally not yield any suitable starting solution for the calculation. In method II, one starts with the solutions of the two-center problem [cf., HUND (*)]. A level that describes electron 1 as being in a $1s$ state around nucleus a and electron 2 as being a $1s$ state around nucleus b in the limiting case of infinite nuclear separations would split into four levels (1 to 4) that might be characterized by the table:

	Nucleus a	Nucleus b
1	1	2
2	2	1
3	1,2	–
4	–	1,2

The interaction of the two electrons will first be considered in higher-order approximations. Method II will be directly identical to the method that was employed by HEITLER and LONDON. Only levels 1 and 2 included in an unperturbed system. It will be assumed that levels 3 and 4 lead to substantially higher-lying energy values. The diversity of levels in unperturbed systems will then be more meager for method III than it is for methods I or II.

There is, indeed, no argument, *a priori*, for preferring any of the three approximation procedures over the other ones. Method I will be most closely applicable to metals of very large conductivity, while method III is most applicable to metals of very feeble conductivity. Method II is in the middle between these two limiting cases.

I have based the following calculations upon method III, since only it can permit a quantitative treatment of the electron interactions.

§ 2. The distribution of the level values. The following calculations define a simple generalization of the HEITLER-LONDON investigations (*loc. cit.* I) to the case of $2n$ electrons in a state of interaction (the number of electrons is now assumed to be even, upon purely formal grounds). One will then find $2n$ electrons in $2n$ different (indeed, they are not different energetically, but positionally) quantum cells.

We shall first assume only that the quantum numbers of the electrons in their atoms are the same for all atoms. Other stationary states of the unperturbed system will not be considered, since it will be assumed that they would lead to much high energy values.

One is then dealing with the determination of the energy values of the stationary states of the total system, which will belong to the state that was described above when the COULOMB interaction of the charges in an atom with the charges of any other atom is considered to be a perturbation. Due to the great computational complications that have appeared up to now, it will only be possible for us to attempt the perturbation calculations up to the first approximation. Whether this first approximation will actually be successful for the cases that nature presents must remain undecided. We take the eigenfunctions of the unperturbed system to be, say, products of the SCHRÖDINGER eigenfunctions of the hydrogen atom, or better yet, the eigenfunctions that correspond to

(*) F. HUND, Zeit. Phys. **40** (1927), 742.

the rest of the atoms considered, just like in the cited paper of HEITLER and LONDON; it is entirely superfluous to repeat those Ansätze here explicitly. These eigenfunctions are certainly not orthogonal, but the deviation from the usual treatment that is required first differs in the terms of order two, so we can apply the usual method of treating things in the case of orthogonal eigenfunctions. The electrons of an atom can be exchanged with those any other atom as a result of perturbations. As long as one overlooks the perturbing terms of order two, only simple transpositions between two neighboring atoms will occur. If one chooses the simplest case to be the one in which any atom in the unperturbed system possesses *one* valence electron then the “exchange terms” for the perturbing energy will reduce to the expressions that were given by HEITLER and LONDON:

$$J_{(kl)} = \frac{1}{2} \int \psi_k^\kappa \psi_k^\lambda \psi_l^\kappa \psi_l^\lambda \left(\frac{2e^2}{r_{kl}} + \frac{2e^2}{r_{\kappa\lambda}} - \frac{e^2}{r_{\kappa k}} - \frac{e^2}{r_{\kappa l}} - \frac{e^2}{r_{\lambda l}} \right) d\tau_k d\tau_l. \quad (1)$$

Here, k and l mean the numbers of two electrons, while κ and λ mean the numbers of the remaining atoms to which k and l belong in the unperturbed state. A very important constant that enter into the perturbation calculations is the purely “static” interaction:

$$J_E = \int d\tau_1 d\tau_2 \cdots d\tau_{2n} (\psi_1^1)^2 (\psi_2^2)^2 \cdots (\psi_{2n}^{2n})^2 \left\{ \sum_{\substack{k,l \\ k>l}} \frac{e^2}{r_{kl}} + \sum_{\substack{\kappa,\lambda \\ \kappa>\lambda}} \frac{e^2}{r_{\kappa\lambda}} - \sum_{\substack{k,\lambda \\ k \neq \lambda}} \frac{e^2}{r_{k\lambda}} \right\}. \quad (2)$$

Due to their smallness, we can leave the magnetic interactions completely outside of consideration. Nevertheless, the spin moments of all electrons will become partly parallel and partly anti-parallel as a result of the *exchange processes*. If one adds the fundamental Pauli principle to this, viz., that the eigenfunctions of the total system should be anti-symmetric in all electrons, then an entirely well-defined total magnetic moment will belong to each level value of the perturbed system that will be characterized by the rotational moment $sh / 2\pi$ of the system. In all, there will be $(2n)!$ levels in the unperturbed system (if one ignores the Pauli principle and spin). *A statistical treatment of ferromagnetism will be possible when all energy values that belong to a given value of s have been calculated.* This problem is generally not soluble in this form, since $2n$ is a very large number. We can only hope to obtain a general insight into the distribution of the eigenvalues for a given s . In what follows, we will calculate the *number* of levels, the center of mass of the energy (thus, the mean value of energy for a given s), and the *mean-square variance* of the energy about that mean value. We shall then make the generally somewhat arbitrary assumption that, in the first approximation, the energy values are distributed around the mean in a GAUSSIAN error curve, such that the breadth of the error curve is calculated from the mean-square variance.

From the investigations of WIGNER, HUND, and HEITLER (*loc. cit.*) and the assumption of the PAULI principle, every value s of the total spin moment belongs to *one* system of levels (“ σ ”) that are characterized by the well-defined partitioning of $2n$ into summands:

$$2n = \underbrace{2+2+\cdots+2}_{(n-s) \text{ times}} + \underbrace{1+1+\cdots+1}_{2s \text{ times}} . \quad (3)$$

The partition of the “reciprocal” system is then called simply:

$$2n = (n - s) + (n + s). \quad (4)$$

HEITLER (*loc. cit.* II) has given the following formula for the mean value – i.e., the “center of mass of energy” – of the system σ :

$$E_\sigma = \frac{1}{f_\sigma} \sum_P \chi_\sigma^P J_P . \quad (5)$$

In this, χ_σ^P means the group character that belongs to the permutation P , and $f_\sigma = \chi_\sigma^E$ is the *number* of levels in the system. The energy of the unperturbed levels is omitted, as an additive constant. We further calculate the mean-square variance $\overline{\Delta E^2}$ of the energy about the value E_σ : The energy value is given by the square root of an equation of degree f_σ that one obtains when one sets the following determinant equal to zero:

$$\begin{vmatrix} \sum_P b_{11}^P J_P - x & \sum_P b_{12}^P J_P & \cdots & \sum_P b_{1f}^P J_P \\ \sum_P b_{21}^P J_P & \sum_P b_{22}^P J_P - x & \cdots & \vdots \\ \vdots & \cdots & \cdots & \vdots \\ \sum_P b_{f1}^P J_P & \cdots & \cdots & \sum_P b_{ff}^P J_P - x \end{vmatrix} = 0. \quad (6)$$

The sum of the roots of this equation $\sum x_n$ is given by the coefficients of x^{f-1} , so $\sum_{i,P} b_{ii}^P J_P = \sum_P \chi_\sigma^P J_P$, which leads to equation (5). The sum $\sum_{n>m} x_n x_m$ is given by the coefficients of x^{f-2} in (6), and one then has:

$$\begin{aligned} \sum_{n>m} x_n x_m &= \sum_{i>k} \sum_{P,P'} b_{ii}^P b_{kk}^{P'} J_P J_{P'} - \sum_{i>k} \sum_{P,P'} b_{ik}^P b_{ki}^{P'} J_P J_{P'} \\ &= \frac{1}{2} \left(\sum_{i,k} \sum_{P,P'} b_{ii}^P b_{kk}^{P'} J_P J_{P'} - \sum_{i,k} \sum_{P,P'} b_{ik}^P b_{ki}^{P'} J_P J_{P'} \right). \end{aligned} \quad (7)$$

In the last expression, one sums over all independent values of i, k . $\sum_k b_{ik}^P b_{ki}^{P'}$ is now the i^{th} diagonal term of the product matrix $b^P \cdot b^{P'}$. Since the matrices b define a representation of the group, one will have:

$$b^P \cdot b^{P'} = b^{P \cdot P'}.$$

If one then observes that $\chi^P = \sum_i b_i^P$ then it will follow that:

$$\sum_{n>m} x_n x_m = \frac{1}{2} \sum_{P, P'} (\chi^P \chi^{P'} - \chi^{P \cdot P'}) J_P J_{P'}. \quad (8)$$

If we set $x_n = E_\sigma + \Delta E_n$ (here, the index σ properly belongs to x_n , ΔE_n , as well; we shall drop it, for the sake of clarity in our notation) then we will have:

$$\begin{aligned} \sum_{n>m} x_n x_m &= \sum_{n>m} (E_\sigma + \Delta E_n)(E_\sigma + \Delta E_m) && (n, m = 1, \dots, f_\sigma) \\ &= \frac{f_\sigma(f_\sigma - 1)}{2} E_\sigma^2 + \sum_{n>m} \Delta E_n \Delta E_m; \end{aligned} \quad (9)$$

since $\sum \Delta E_n = 0$, one will have:

$$\sum_{n=1}^{f_\sigma} \Delta E_n^2 = -2 \sum_{n>m} \Delta E_n \Delta E_m. \quad (10)$$

It ultimately follows from (5), (8), (9), (10), and the equation $f_\sigma = \chi_\sigma^E$ that:

$$\sum_{n=1}^{f_\sigma} \Delta E_n^2 = \frac{1}{f_\sigma} \sum_{P, P'} (\chi_\sigma^E \chi_\sigma^{P \cdot P'} - \chi_\sigma^P \chi_\sigma^{P'}) J_P J_{P'} \quad (11)$$

and

$$\overline{\Delta E_\sigma^2} = \frac{1}{f_\sigma^2} \sum_{P, P'} (\chi_\sigma^E \chi_\sigma^{P \cdot P'} - \chi_\sigma^P \chi_\sigma^{P'}) J_P J_{P'}. \quad (12)$$

In order to be able to apply this formula, we must calculate the group characters of the permutations of the various classes. Since all J_P vanish, with the exception of the cases $P = E$ and P of the class (12) (viz., transpositions), only the following types will come under consideration for $\chi_\sigma^{P \cdot P'}$:

$$\chi_\sigma^E, \chi_\sigma^{(12)}, \chi_\sigma^{(123)}, \chi_\sigma^{(12)(34)}.$$

These group characters can be calculated by a method of SCHUR that was described by HEITLER (*loc. cit.* III). It next yields for the reciprocal system of levels (*):

(*) The values for f_σ and $\chi_\sigma^{(12)}$ that were given by HEITLER [*loc. cit.*, III, equation (32)] are marred by printing errors.

$$\left. \begin{aligned}
\chi_{n-s,n+s}^E &= \frac{(2n)!(2s+1)}{(n-s)!(n+s+1)!}, \\
\chi_{n-s,n+s}^{(12)} &= \frac{(2n-2)!(2s+1)}{(n-s)!(n+s+1)!} (s^2 + s + n^2 - 2n), \\
\chi_{n-s,n+s}^{(123)} &= \frac{(2n-3)!2(n-1)}{(n-s)!(n+s+1)!} \{6s^2 + 9s^2 + s(2n^2 - 10n + 3) + n(n-5)\}, \\
\chi_{n-s,n+s}^{(12)(34)} &= \frac{4(2n-4)!}{(n-s)!(n+s+1)!} \{2s^2 + 5s^4 + 4s^2(n^2 - 5n + 4) + s^2(6n^2 - 30n + 19) \\
&\quad + 2s(n^4 - 6n^3 + 15n^2 - 14n + 3) \\
&\quad + n^4 - 6n^3 + 14n^2 - 9n\}.
\end{aligned} \right\} \quad (13)$$

The characters of the system of levels that actually is present differ from the characters of their reciprocals that are employed here only by their signs. Indeed, the character of the reciprocal system is equal to (equal and opposite to, resp.) that of the system itself when the permutation P arises from an even (odd, resp.) number of transpositions.

Up to this point, everything is true in complete generality, with no relationship to any special assumptions that we might make about the crystal lattice or the atomic structure of the ferromagnetic substance.

In order to be able to calculate, we must now specialize our assumptions somewhat further. It follows from formula (1) and the calculations of HEITLER and LONDON that $J_{(12)}$ decreases exponentially with increasing distance. For the most part, one can exchange an atom in a lattice only with its “neighbors”; exchanges with atoms that lie further away than the “neighboring atoms” will then be neglected. The number of “neighbors” of an atom is, e.g., 1 in a molecular lattice of diatomic molecules, 2 in a linear chain, 4 in a quadratic surface lattice, 6 in a simple cubic lattice, 8 in a cubic, space-centered lattice, and 12 in a cubic, face-centered lattice.

We shall make only the assumption that all non-vanishing exchange terms J_P should be *equal* (we call that value J_0). That must be case when the remaining atoms are non-magnetic; i.e., centrally-symmetric. We then now calculate E_σ and $\overline{\Delta E_\sigma^2}$ for a lattice in which every atom z has neighbors. Thus, we will consider only the highest powers in n and s and drop the lower terms; that means that we shall overlook the effects on the boundary surface of the crystal.

The number of transpositions that lead to the value J_0 (i.e., the number of atom-pairs of least separation) is $z \cdot 2n / 2 = z \cdot n$. Thus, it follows from (5) and (13) that:

$$E_\sigma = -z \cdot \frac{s^2 + n^2}{2n} J_0 + J_E. \quad (14)$$

In order to calculate the values of $\overline{\Delta E_\sigma^2}$, we next need the value of the expression:

$$A_{P,P'} = \frac{1}{f_\sigma^2} (\chi^E \chi^{P,P'} - \chi^E \chi^{P',P}) \quad (15)$$

for the various possible combinations of P and P' . That yields (up to lower powers of n and s):

$$\left. \begin{aligned} 1. P = P' : & \quad A_{(12)(12)} = \frac{(n^2 - s^2)(3n^2 + s^2)}{4n^4}, \\ 2. P \text{ and } P' \text{ have a common element :} & \quad A_{(12)(13)} = \frac{(n^2 - s^2)s^2}{4n^4}, \\ 3. P \text{ and } P' \text{ have no common element :} & \quad A_{(12)(34)} = \frac{-(n^2 - s^2)s^2}{2n^5}. \end{aligned} \right\} \quad (16)$$

If an atom has z neighbors then type 1 will occur $z \cdot n$ times, type 2, $z(z-1)n$ times, and type 3, $z^2 n^2 / 2$ times.

It finally follows from equation (12) that $\overline{\Delta E_\sigma^2}$ is:

$$\overline{\Delta E_\sigma^2} = J_0^2 \left(zn \cdot A_{(12)(12)} + 2z(z-1)n A_{(12)(13)} + 2 \frac{z^2 n^2}{2} A_{(12)(34)} \right)$$

and

$$\overline{\Delta E_\sigma^2} = J_0^2 z \frac{(n^2 - s^2)(3n^3 - s^2)}{4n^3}. \quad (17)$$

The mean deviation of the energy from the mean (14) then has order of magnitude $\Delta E_\sigma \sim J_0 \sqrt{n}$. In the foregoing formulas, $-\sigma$ means always means the system of levels that belongs to the partition (3) and thus to the total spin moment s .

§ 3. Statistics: connection with WEISS's formulas. The following arguments will be founded upon the aforementioned, generally somewhat arbitrary, assumption that the distribution of energy values about the mean has the approximate form of a GAUSSIAN error curve. Since the total number of levels that belong to the spin moment s amounts to f_σ^* , we shall thus assume that:

$$\frac{f_\sigma}{\sqrt{2\pi\overline{\Delta E_\sigma^2}}} \cdot e^{-\frac{\Delta E^2}{2\overline{\Delta E_\sigma^2}}} d\Delta E$$

levels lie between $E_\sigma + \Delta E$ and $E_\sigma + \Delta E + d\Delta E$.

From the calculations up to now, the direction of the total spin moment s in a crystal will remain completely undetermined when, as we can conclude from the result of the EINSTEIN-DE HAAS effect, the orbital moments of the electrons in the crystal

(*) Without the Pauli principle, due to the f_σ -fold degeneracy of the total number of levels in all systems that belong to the partition would be equal to f_σ^2 . However, that degeneracy will be removed by the Pauli principle.

compensate for each other. From known formulas, one will have for this total energy ($g = 2$ for the spin!):

$$E' = - \frac{e}{\mu c} \frac{h}{2\pi} H \cdot m \quad (s \geq m \geq -s) \quad (18)$$

($m =$ electron mass).

The problem now arises of calculating the most probable value of m for a given temperature and a given value of H . To abbreviate, we then introduce:

$$\left. \begin{aligned} \alpha &= \frac{1}{kT} \cdot \frac{e}{\mu c} H \cdot \frac{h}{2\pi}, \\ \beta &= \frac{z J_0}{kT}. \end{aligned} \right\} \quad (19)$$

One then defines the state sum (up to an inessential factor that is independent of α):

$$\sum_{s=0}^n \sum_{m=-s}^{+s} \int_{-\infty}^{+\infty} d\Delta E \frac{f_\sigma}{\sqrt{2\pi \Delta E_\sigma^2}} \cdot e^{\alpha m + \beta \frac{s^2}{2\pi} \frac{\Delta E}{kT} - \frac{\Delta E^2}{2\Delta E_\sigma^2}} = \sum_{s=0}^n \sum_{m=-s}^{+s} f_\sigma e^{\alpha m + \beta \frac{s^2}{2\pi} + \frac{\Delta E_\sigma^2}{2k^2 T^2}}.$$

Finally, it follows from (17) (again, while dropping unimportant factors) that:

$$S = \sum_{s=0}^n \sum_{m=-s}^{+s} f_\sigma e^{\alpha m + \beta \frac{s^2}{2\pi} - \beta^2 \frac{s^2(4n^2 - s^2)}{8n^3 z}}. \quad (20)$$

If one denotes the expression $e^{\beta \frac{s^2}{2\pi} - \beta^2 \frac{s^2(4n^2 - s^2)}{8n^3 z}}$ by $g(s)$ and considers that $f_\sigma = \binom{2n}{n+s} - \binom{2n}{n+s+1}$ then after substituting the summation series will become:

$$S = \sum_{m=-n}^{+n} \sum_{s=|m|}^n e^{\alpha m} \cdot g(s) \left[\binom{2n}{n+s} - \binom{2n}{n+s+1} \right],$$

and a transformation that is analogous to partial integration will make:

$$\begin{aligned} S &= S_1 + S_2 = \sum_{m=-n}^{+n} e^{\alpha m} \cdot g(m) \binom{2n}{n+s} + \sum_{m=-n}^{+n} \sum_{s=|m|}^n e^{\alpha m} [g(s+1) - g(s)] \binom{2n}{n+s+1}, \\ S_2 &= \sum_{s=0}^n \frac{e^{\alpha(s+1)} - e^{-\alpha s}}{e^\alpha - 1} \cdot [g(s+1) - g(s)] \binom{2n}{n+s+1}. \end{aligned}$$

Since $g(s)$ is an even function of s , one can drop the term in S_2 in $e^{-\alpha s}$ in the approximation that is employed here and extend the sum in it from $s = -n$ to $s = +n$. One will then see that up to factors of order of magnitude 1, S_2 will also assume the form of S_1 (if one replaces the summation symbol s in S_2 with m).

It ultimately follows that:

$$S = F \cdot \sum_{m=-n}^n e^{\alpha m} g(m) \binom{2n}{n+m}, \quad (21)$$

where F is a function of the quantities α and β that has order of magnitude 1 (when α and β are of that order of magnitude).

For our later calculations, it will be assumed that the expression under the summation sign in S exhibits a steep maximum at the location $m = m_0$ ($m_0 \approx$ mean value of m), which shall be verified later. We set $m = m_0 + (m - m_0)$ and develop the exponent of $g(m)$ in powers of $(m - m_0)$. That will yield:

$$g(m) \approx e^{\beta \frac{m_0^2}{2n} + \beta \frac{m_0}{n} (m - m_0) - \frac{\beta^2}{8n^3 z} (4m_0^2 n^2 + m_0^4 + 8m_0 n^2 (m - m_0) - 4m_0^3 (m - m_0))}.$$

Thus, S again becomes, up to inessential factors that do not depend upon α

$$\begin{aligned} S &= F \sum_{m=-n}^n e^{\left(\alpha + \beta \frac{m_0}{n} - \beta^2 \frac{m_0}{nz} + \beta^2 \frac{m_0^3}{2n^3 z} \right) m} \binom{2n}{n+m} \\ &= F \left[2 \cos \left(\frac{\alpha + \beta \frac{m_0}{n} - \beta^2 \frac{m_0}{nz} + \beta^2 \frac{m_0^3}{2n^3 z}}{2} \right) \right]^{2n}. \end{aligned} \quad (22a)$$

We obtain the most probable value m_0 from S using the equation:

$$m_0 = \frac{\partial}{\partial \alpha} \log S = n \tan \frac{\alpha + \beta \frac{m_0}{n} - \beta^2 \frac{m_0}{nz} + \beta^2 \frac{m_0^3}{2n^3 z}}{2}. \quad (22)$$

The term that originates from the differentiation of $\log F$ can be neglected in comparison to n . In this, one also finds the belated justification for the fact that have considered only the highest powers in n and s in the exponent in the state sum, and thus neglected factors whose order of magnitude was 1.

Formula (22) is, in essence, identical with the known WEISS formula for ferromagnetism. The fact that $\tan x$ appears in (22), instead of $\cot x - 1/x$ (as in WEISS) originates in the fact that only two orientations for the spin moment in an external field are possible. If one sets:

$$\frac{m_0}{n} = y; \quad a + \beta \frac{m_0}{n} - \beta^2 \frac{m_0}{nz} + \beta^2 \frac{m_0^3}{2n^3 z} = 2x \quad (23)$$

then one will get the equations:

$$\left. \begin{array}{l} \text{I. } y = \tan x, \\ \text{II. } 2x = \alpha + y \left(\beta - \frac{\beta^2}{z} \right) + \frac{\beta^2}{2z} y^3. \end{array} \right\} \quad (24)$$

This formula differs from the WEISS formulas by the cubic term on (24.II). Once more, one can construct the solutions of (24) most simply graphically. Fig. 1 is drawn for the values $\alpha = 0.2$, $\beta = 2$, $z = 8$.

For small or negative values of the constant β , one will get paramagnetism. Ferromagnetism enters in when the tangent of the curve II for $y = 0$ subtends a smaller angle with the x -axis than the tangent of I; the influence of the cubic terms is first ignored in this. The condition for ferromagnetism then reads:

$$\beta \left(1 - \frac{\beta}{z} \right) \geq 2. \quad (25)$$

This condition can be fulfilled only for high values of z . The maximal value on the left-hand side of (25) is $(\beta_{\max} = z/2) \frac{z}{2} \left(1 - \frac{1}{2} \right)$, and it follows that:

$$z \geq 8. \quad (26)$$

Ferromagnetism is then possible only for lattice types for which an atom has at least eight neighbors. That is the case for Fe, Co, Ni, whose lattices are all cubic, some of which are space-centered ($z = 8$) and some of which are face-centered ($z = 12$). By also considering the terms of third degree in 24.II, it can perhaps even happen for $z = 7$ that curve II indeed (for $\alpha = 0$) increases steeper than I at the origin, but that later on, two intersection points of I and II will occur. Since $z = 7$ does not occur anyway, one should not attach any physical meaning with that possibility.

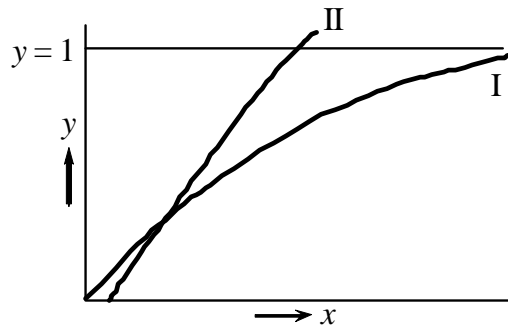


Fig. 1.

When β increases above the value $z/2$ (increasing β corresponds to decreasing temperature), from equation (25), the “strength of the molecular field” will again

decrease, and for $\beta > \frac{z}{2} \left(1 + \sqrt{1 - \frac{8}{z}} \right)$, there will no longer be any paramagnetism.

Mathematically, in fact, this comes about because of the assumed GAUSSIAN distribution for the energy values, which has the consequence that level values will also occur for small values of s that (for positive J_0) lie just as deeply to deeper than the energy that belongs to $s = n$. In reality, there is certainly *no* energy value that lies deeper than the one for $n = s$ (when J_0 is positive). Naturally, the deviations of the actual energy distribution curve from the assumed one will become more and more noticeable for decreasing temperature.

One must improve the provisional theory that was attempted here by calculating the higher variances of the mean $\overline{\Delta E^3}$, $\overline{\Delta E^4}$, etc. and correspondingly construct improved distribution curves for the values of the terms. Corresponding higher powers of β would appear on the left-hand side of equation (25) in this improved theory; the left-hand side of (25) is thus actually a transcendental function of β . The value that we have obtained here for the left-hand side of (25) might correspond to a development of that function in powers of $\frac{\beta}{z} = \frac{J_0}{kT}$ that has been truncated at the second term (our term $\beta - \beta^2/z$ can indeed differ from the first two terms of the true power series of that function only by small amounts). However, it follows from this argument that for higher values of β/z – say, $\frac{\beta}{z} \geq \frac{1}{2}$ – an examination of the higher variances $\overline{\Delta E^n}$ would be imperative for the study of ferromagnetism. Such a more precise examination of the distribution curve would also most likely displace the limiting value (26) of z . However, nothing in our results would change very much qualitatively.

If β is substantially smaller than the limiting value that is given by (25) then, as was mentioned already, equation (24) will yield paramagnetism. For small values of α , calculation gives ($\tan x \approx x - \frac{x^3}{3} \dots$):

$$y = \frac{\alpha}{2 - \beta + \frac{\beta^2}{z}} + \frac{\alpha^3}{\left(2 - \beta + \frac{\beta^2}{z} \right)^4} \left(\frac{\beta^2}{2z} - \frac{2}{3} \right) + \dots \quad (27)$$

The first term of this series gives the CURIE law with a modification that is similar to the one in the WEISS theory:

$$m_0 \propto \frac{1}{T - \Theta} \frac{T}{T \left(1 + \sqrt{1 - \frac{8}{z}} \right) - \Theta \left(1 - \sqrt{1 - \frac{8}{z}} \right)}.$$

The critical temperature in this is Θ :

$$\Theta = \frac{2J_0}{k \left(1 - \sqrt{1 - \frac{8}{z}} \right)}.$$

Before I go on to a discussion of the numerical values of J_0 , I would like to justify the previous assertion that the terms in the state sum (21) exhibit a very steep maximum at the location $m = m_0$. To that end, I shall calculate the mean-square variance $\overline{\Delta m^2}$ of m about m_0 . It is:

$$\overline{\Delta m^2} = \overline{m^2} - m_0^2$$

and

$$\overline{m^2} = \frac{1}{S} \frac{\partial^2 S}{\partial \alpha^2}.$$

From (22a):

$$\frac{\partial S}{\partial \alpha} = n (e^x + e^{-x})^{2n-1} (e^x - e^{-x}) F + \frac{\partial F}{\partial \alpha} (e^x + e^{-x})^{2n},$$

$$\begin{aligned} \frac{\partial^2 S}{\partial \alpha^2} &= \frac{n(2n-1)}{2} (e^x + e^{-x})^{2n-2} (e^x - e^{-x})^2 F + \frac{n}{2} (e^x + e^{-x})^{2n} F \\ &\quad + 2n \frac{\partial F}{\partial \alpha} (e^x + e^{-x})^{2n-1} (e^x - e^{-x}) + \frac{\partial^2 F}{\partial \alpha^2} \frac{n}{2} (e^x + e^{-x})^{2n}. \end{aligned}$$

This implies that, in the first approximation:

$$\overline{m^2} = m_0^2 + \frac{n}{2} (1 - \tan^2 x) + 2n \tan x \cdot \frac{\partial}{\partial \alpha} \log F.$$

One will then have:

$$\overline{\Delta m^2} = \frac{n}{2} \left(1 - \tan^2 x + 4 \tan x \frac{\partial}{\partial \alpha} \log F \right).$$

The most probable deviation of the moment m from the expected value m_0 is then only of order of magnitude \sqrt{n} , so the terms that were neglected in the exponent of $g(m)$ in equation (22) will be order $\Delta m^2 / n$, and thus of order 1, which agrees with the degree of approximation that we have consistently sought. The neglected terms have the same order as the boundary surface effects.

§ 4. Magnitudes and signs of the “molecular field.” The constant β must have order of magnitude 1, in order for ferromagnetism to be possible; one must then have $J_0 \sim kT$, where T will assume values on the order of 10^3 degrees for Fe, Co, Ni. It follows that $J_0 \sim 10^{-13}$ erg $\sim \frac{1}{100}$ the energy of the hydrogen ground state. That is just the order of the energy contribution that one would expect for the exchange term of the form (1) when the

atoms lie close to each other. If the atomic separations become larger then the exchange terms will decay exponentially. That is the basis for the fact that iron or nickel salt solutions are never ferromagnetic.

The question of the sign of J_0 is much more difficult to answer. In their theory of the homopolar bond, HEITLER and LONDON make that assumption that J_0 is negative in complete generality, which would exclude ferromagnetism. For the special case in which the electrons are found to be unperturbed in the $1s$ state, it follows, in fact, from general theorems that the energy values must lie in a way that would correspond to negative values of J_0 . Such an argument is, in turn, applicable to only electrons in the $1s$ state, and one can show that J_0 will generally be positive for high principal quantum numbers. One must then deal with the expression:

$$J_0 = \frac{1}{2} \int \psi_k^\kappa \psi_k^\lambda \psi_l^\kappa \psi_l^\lambda \left(\frac{2e^2}{r_{kl}} + \frac{2e^2}{r_{\kappa\lambda}} - \frac{e^2}{r_{\kappa k}} - \frac{e^2}{r_{\lambda k}} - \frac{e^2}{r_{\lambda l}} \right) d\tau_k d\tau_l, \quad (1)$$

in which κ and λ are the indices for the atomic nuclei, while k and l are the those of the electrons. Initially, ψ will be a hydrogen eigenfunction, but later on, it will be shown that the argument is just as valid for other central fields in the vicinity of the nucleus. One can then say with certainty that J_0 will be positive for very small values of $r_{\kappa\lambda}$, since the term $1 / r_{\kappa\lambda}$ will then outweigh all of the other ones. However, that result does not need to have any physical meaning, since for very small values of $r_{\kappa\lambda}$, even the entire approximation becomes illusory (cf., the case of $1s$ terms!). One then comes to the values of J_0 for very large $r_{\kappa\lambda}$. When J_0 is positive there, one must assume that it remains positive for all values of $r_{\kappa\lambda}$, in general. We thus investigate further how a charge distribution of density $\psi_k^\kappa \psi_k^\lambda$ appears at large distances $r_{\kappa\lambda}$, first for perhaps the higher s terms. The Schrödinger functions contain an e -function as the most important term, and

$\psi_k^\kappa \psi_k^\lambda$ thus contains the factor $e^{-\frac{r_{\kappa\kappa} + r_{k\lambda}}{a_0 n}}$ ($a_0 =$ Bohr hydrogen radius, $n =$ principal quantum number; thus, no confusion with the electron number $2n$ should be created). If one drops the remaining factors then the density will be constant on confocal ellipsoids of rotation around the two nuclei. For increasing distance between the nuclei, the charge ellipsoid degenerates into a cylinder around the connecting line between the nuclei. (This happens for both values of the principal quantum number.) Furthermore, the e -function appears multiplied by a polynomial in $r_{\kappa\kappa}$ ($r_{\kappa\lambda}$, resp.) of degree $n - 1$. The zero locus of this polynomial lies entirely in the neighborhood of the nucleus; at greater distances from it, it will suffice to replace the polynomial with its highest power r^{n-1} . The behavior of the central force at distances of order a_0 from the nucleus is entirely inessential when only $r_{\kappa\lambda}$ is sufficiently large. The density distribution of the charge over the length of the aforementioned cylinder is therefore non-uniform, but otherwise approximately proportional to $r_{\kappa\kappa}^{n-1} r_{k\lambda}^{n-1}$. For small values of n , this distribution is still quite uniform and one can easily see that the negative terms in J_0 can substantially predominate. For increasing n , by contrast, the density distribution assumes an ever steeper maximum at the midpoint between the two nuclei. In the limit of very large values of n , the mean

value of the terms of type $1 / r_{\kappa\kappa}$, when taken over the density distribution that was given above, tends to the value $2 / r_{\kappa\lambda}$:

$$\overline{\frac{1}{r_{\kappa\kappa}}} = \overline{\frac{1}{r_{\kappa\lambda}}} = \overline{\frac{1}{r_{l\kappa}}} = \overline{\frac{1}{r_{l\lambda}}} \rightarrow \frac{2}{r_{\kappa\lambda}}.$$

By contrast, the term with $1 / r_{kl}$ – viz., the “self-potential” of the density distribution – increases beyond all limits with increasing n . J_0 is then certainly positive for sufficiently high principal quantum numbers. One can easily show that nothing will change in this result when one carries out the calculations for p , d , ..., or any other higher state. The limiting value of n for which J_0 can become positive for the first time is difficult to determine exactly. A rough calculation yields $n = 3$. This limiting value will possibly depend upon values of the remaining quantum numbers. The fact that, e.g., the oxygen molecule empirically possesses a magnetic moment of $2 \cdot \frac{1}{2} \frac{h}{2\pi}$ in the ground state seems to show that J_0 can already be positive for $n = 2$. On the other hand, it can follow from the many-times-observed critical temperatures (e.g., for γ -iron) that there are many times J_0 that can also be negative for higher principal quantum numbers.

Concluding remarks. The calculations that were described here lead to two conditions for the appearance of ferromagnetism:

1. The crystal lattice must be a type such that any atom has at least 8 neighbors.
2. The principal quantum number of the electrons that are responsible for magnetism must be $n \geq 3$.

Both conditions together do not reach far enough to single out Fe, Co, Ni from all other materials; however, Fe, Co, Ni do satisfy the conditions. It was certainly to also be expected that the theory that was contrived here can meanwhile serve as only a qualitative schema in which ferromagnetic phenomena will perhaps be classified later. The theory admits an extension for the case of several exchanges per atom; an incisive study of the $J_{(kl)}$ values, as well as the distribution curve of the term values, will be requisite. I hope to be able to go into these questions, as well as a thorough comparison of the theory with the experimental results later.

Leipzig, Institut für theoretische Physik der Universität.
