A Survey of the Theory of Ferromagnetism*

J. H. VAN VLECK

Harvard University, Cambridge, Massachusetts

THE WEISS MOLECULAR FIELD

THE outstanding development in the phenom-enological description of ferromagnetism is undoubtedly Weiss' theory of the molecular field,[1] which he first published in 1907. Shortly before this, Langevin had developed his well-known theory of paramagnetism, based on the fundamental idea that the orientation of a molecular dipole of moment μ in a field H is governed by the Boltzmann distribution law. With this hypothesis it is easy to show[2] that the moment per unit volume is given by the expression

$$M = N\mu L(\mu H/kT), \tag{1}$$

where N is the number of molecules per unit volume, and $L(x)$ is the so-called Langevin function

$$L(x) = \coth x - x^{-1}. \tag{2}$$

The basic idea of the Weiss theory is that the
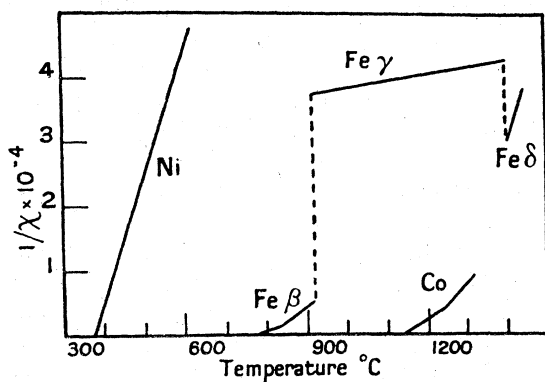

FIG. 1. The reciprocal of the susceptibility as a function of temperature above the Curie point.

---

[1] P. Weiss, J. de Phys. 6, 667 (1907).
[2] P. Langevin, J. de Phys. 4, 678 (1905). For a review of the Langevin theory and properties of the Langevin function, see, for example, Chap. II of the writer's *Electric and Magnetic Susceptibilities* (Oxford University Press, New York, 1932).

effective field acting on an elementary magnet in a ferromagnetic medium is not to be identified with the applied field H, but is, rather, to be taken as $H+qM$, where M is the intensity of magnetization, and q is a proportionality factor independent of temperature. The portion qM is called the "molecular field," and is clearly a manifestation of the cooperative phenomenon in virtue of which the atomic magnets tend to be parallel.

The Weiss theory has the merit of great simplicity, for many phenomena can be explained by taking the argument of the Langevin function as $H+qM$ instead of H. With this modification, Eq. (1) becomes

$$M = N\mu L(\mu[H+qM]/kT). \tag{3}$$

If we can neglect saturation effects, and so make the familiar approximation[2] $L(x) = \frac{1}{3}x$ appropriate to small x, then after solving for M, the relation (3) reduces to

$$\chi = M/H = N\mu^2/3k(T - T_C). \tag{4}$$

Here χ denotes the susceptibility M/H and

$$T_C = N\mu^2 q/3k. \tag{5}$$

Equation (4) obviously fails at $T = T_C$, and so the Weiss theory immediately gives us a critical point or "Curie temperature." Below the latter it is no longer allowable to make the approximation $L(x) = \frac{1}{3}x$. Instead, the moment ceases to be linear in the field strength; enormous magnetization can be obtained without the necessity of corresponding applied fields, and the behavior becomes ferromagnetic.

This simple analysis furnishes a remarkably satisfactory description of the salient experimental facts. We shall allude to what are perhaps the two most important of the many successes of the Weiss theory. One of these is the linear relation which is predicted by Eq. (4) between the reciprocal of the susceptibility and the temperature above the Curie point. As shown by Fig. 1, taken from Stoner's *Magnetism and Molecular Structure*, the linearity is on the whole quite well
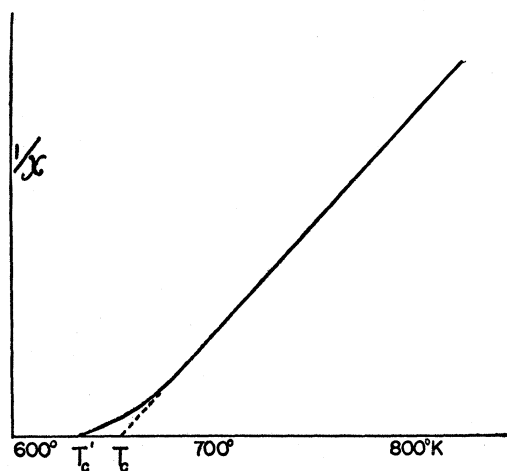
Fig. 2. Enlargement of Fig. 1 near the Curie point for nickel, to illustrate the distinction between the paramagnetic Curie point T_C and the ferromagnetic one T_C'.

confirmed experimentally, except for the discontinuities at the polymorphic transition points, whose effect is particularly pronounced in iron. Figure 1 is drawn to too small a scale to show the fact, illustrated schematically in Fig. 2, that as the temperature is lowered to the vicinity of the Curie point, the experimental curves begin to deviate more from linearity. The intercept of the axis $1/\chi = 0$ obviously corresponds to infinite susceptibility or in other words to ferromagnetism. Usually it occurs at a lower temperature T_C' than the value T_C which is obtained by extrapolation from the linear behavior at higher temperatures. The quantities T_C and T_C' are sometimes called the paramagnetic and ferromagnetic Curie points. The distinction between them is relatively small, representing so to speak a second-order effect. For instance in nickel, the difference $T_C - T_C'$ is about 20°, whereas T_C' itself is 630°K.

The other success of the Weiss theory which we should like to emphasize is its prediction concerning the variation of saturation magnetization with temperature below the Curie point. If $T < T_C$, Eq. (3) can be solved for M in terms of T only numerically. The solution need, however, be made only once and for all, if we employ "reduced units,"

$$\zeta = M/N\mu, \quad \tau = T/T_C,$$

wherein the magnetization is measured relative to the saturation intensity $N\mu$ achieved at the absolute zero, and the temperature is expressed as a multiple of the Curie point. Then, according to the Weiss theory, we have for all substances, a relation of the form $\zeta = f(\tau)$, where f is a "universal function" involving no undetermined constants. The graph of this function is shown in the curve of Fig. 3 labelled $S = \infty$. It is seen that the general trend of the experimental points does indeed follow the classical curve, although they do agree better with the curve $S = \frac{1}{2}$ yielded by quantum-mechanical refinement, which we will discuss later. The universality of the $\zeta - \tau$ graph yielded by the classical theory of Weiss is usually called the law of corresponding states. The analogy to the law of corresponding states in van der Waals equation of state in thermodynamics is obvious; in fact, we may well apply to the Weiss theory the words which Harvey Davis so aptly uses to describe that of van der Waals: namely, that it is "qualitatively right but quantitatively wrong," and based "half on theory, and half on genius at empirical guessing." The empirical part of the Weiss theory is, of course, the ad hoc introduction of the molecular field.

In deriving the universal curve shown in Fig. 3, one approximation is necessary, and that is to neglect the applied field H in comparison with the Weiss field qM in the argument of the Langevin function. The reader may perhaps object that the full intensity of magnetization is never obtained without some applied field, so that the approximation $H = 0$ is not really allowable. However, the values of the necessary fields are always small compared with the resulting induction, and below the Curie point, the intensity of magnetization appearing in Eq. (3) or graphed in Fig. 3 is to be construed as relating to what is called the saturation intensity. It is not the same as the true saturation $N\mu$ which would be obtained were all the elementary dipoles perfectly parallel, and which is reached only at the absolute zero. According to the Weiss model, a ferromagnetic solid is to be thought of as composed of a large number of microcrystals, each of which is spontaneously magnetized to the extent given by (3) with $H = 0$. The directions of magnetization of the microcrystals do not all coincide, but instead are random in the absence of an external applied field. A comparatively small field, of the order 10^2 to 10^3 gauss, however, will suffice to align the

directions of magnetization of the microcrystals, and Eq. (3), with $H=0$, gives the intensity of magnetization when this state is achieved. This transition from random to parallel alignment of the elementary domains gives rise to phenomena such as remanence, hysteresis, the Barkhausen effect, and the like. They are not described by the simple Weiss theory, which does not contain such quantum-mechanical complications as spin-orbit coupling, and so has no mechanism to inhibit the microcrystals from having their directions of spontaneous magnetization perfectly parallel even in infinitesimal fields, instead of fields of the order 10^3 gauss. The transition phenomena associated with the lining-up of the magnetizations of the various micro-crystals belongs to a subject usually referred to as "domain theory," which will not be treated in this article.[3] On the whole this theory is at present much more qualitative than quantitative in nature, although one must not overlook some important Russian work,[4] which has more quantitative features than does the traditional Bloch model.[5]

The great mystery of the Weiss theory was how to explain the large molecular fields. They were supposed to be a manifestation of powerful coupling between elementary magnets. However, at the time, the only known interaction between them was the classical dipole-dipole coupling, whose potential is

$$V_{ij} = \frac{\mathbf{u}_i \cdot \mathbf{u}_j}{r_{ij}^3} - 3\frac{(\mathbf{u}_i \cdot \mathbf{r}_{ij})(\mathbf{u}_j \cdot \mathbf{r}_{ij})}{r_{ij}^5}. \qquad (6)$$

Equation (6) embodies the effect which one demonstrates experimentally when one places a number of compass needles close together. However, the resulting interaction is far too weak to yield the coupling required by the Weiss theory. It gives a maximum value 4π for the constant q in the molecular field qM, whereas the successful application of the Weiss theory requires that q be of the order 10^5. Furthermore, with the classical dipolar mechanism (6), the value of q depends strongly on how the specimen is cut, and on the direction of the field relative to the crystal-

[3] For a discussion of the domain theory, see the paper by W. F. Brown, Jr., Rev. Mod. Phys. **17**, 15 (1945).
[4] L. Landau and E. Lifshitz, Physik. Zeits. Sowjetunion **8.2**, 153 (1935).
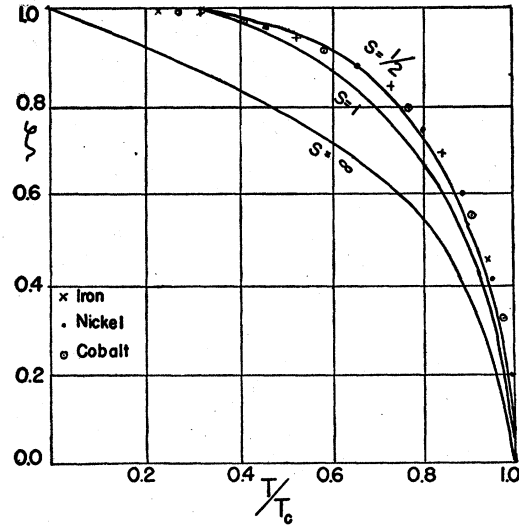[5] F. Bloch, Zeits. f. Physik **74**, 295 (1932).

FIG. 3. Saturation intensity of magnetization as a function of temperature, in reduced units. The abscissa is the ratio of the temperature to the Curie temperature, and the ordinate is the ratio of the saturation intensity of magnetization to the value of this intensity at the absolute zero. The curve $S = \infty$ is for the classical theory of Weiss. The other curves are based on the Brillouin rather than Langevin function, and are the corresponding quantum-theory versions appropriate to a spin quantum number S.

lographic axes. The reason that this is so is that the minimum value of the potential energy (6) is achieved with parallel alignment $\mathbf{u}_i = \mathbf{u}_j$ when the directions of $\mathbf{u}_i$ and $\mathbf{r}_{ij}$ coincide, whereas the stable configuration is the anti-parallel one $\mathbf{u}_i = -\mathbf{u}_j$ when $\mathbf{u}_i$ is perpendicular to $\mathbf{r}_{ij}$. Thus ferromagnetic anisotropy, according to the classical picture if taken literally, should be, so to speak, a first-order effect. Actually (except for the case of pyrrhotite) it is a second-order one, in the sense that the saturation intensities of magnetization are approximately the same in all directions, although the fields required to produce saturation may be different. As by Eq. (5) the Curie temperature is proportional to the constant q of the molecular field, another way of expressing these deficiencies of the classical dipolar model is to say that the latter would give Curie points, comparable with the very low temperatures achieved by adiabatic demagnetization, rather than of the order 10^3 °K. Furthermore the Curie temperature would depend on how the specimen was cut and on the orientation of the applied field relative to the crystallographic axes. Cryogenic experiments on certain paramagnetic

salts do indeed reveal critical temperatures measured in hundredths or thousanths of a degree which are attributable to dipole-dipole interactions, but this is not our subject. We may also remark, parenthetically, that in the electric rather than magnetic case, the dipolar forces can be much larger, essentially because molecular electric dipole moments are much larger than magnetic when expressed in terms of Gaussian units. (E.g., one Debye unit of electric polarity is 10^{-18} whereas one Bohr magneton is $.92 \times 10^{-20}$). In consequence, it is possible for Rochelle salt, and a few other substances, to show the electric analog of ferromagnetism, just from dipolar forces alone, and in such cases there is tremendous anisotropy, as we should expect.

ADVENT OF QUANTUM MECHANICS

The discovery of the true quantum mechanics in 1926 was a great help in understanding ferromagnetic phenomena. In the first place, it was accompanied by the Uhlenbeck-Goudsmit concept of electron spin. The latter has a ratio of magnetic moment to angular momentum equal to e/mc instead of the classical $e/2mc$. This fact enables us to understand why it is that experiments on magnetization by rotation and on rotation by magnetization made by Barnett and others[6] always (again, excepting pyrrhotite) yield a gyromagnetic ratio almost twice the classical value. This behavior is to be expected if most of the orbital angular momentum is largely destroyed by interatomic forces in the solid state, leaving only the spin. The quantum theory of ferromagnetism is usually developed on the basis that the orbital contributions to the magnetic moment are negligible. Actually, they cannot be forgotten entirely, as evidenced by the fact that gyromagnetic ratios of ferromagnetics are usually nearer 1.9 than 2.0. Although quantum mechanics enables us to understand qualitatively that the orbital angular momentum is largely suppressed, remarkably little has been done, on the whole, in the way of quantitative calculations[7]

to show that just enough orbital angular momentum will be left to make the gyromagnetic ratio about 0.1 unit lower than the spin only value 2.0.

One thing which quantum theory has obviously done is to introduce a discrete series of orientations rather than a continuous distribution as in the classical Langevin theory. When this modification is made, Eq. (3) is replaced by

$$M = 2NS\beta B_S\left(\frac{2S\beta(H+qM)}{kT}\right), \qquad (7)$$

where $B_S(y)$ is the so-called Brillouin function

$$B_S(y) = \frac{\sum_{R=-S,\cdots,S} R e^{-Ry/S}}{S \sum_R e^{-Ry/S}}$$

$$= \frac{2S+1}{2S} \coth\left(\frac{2Sy+y}{2S}\right) - \frac{1}{2S} \coth\left(\frac{y}{2S}\right).$$

Here S is the spin quantum number of the atom, and β is the Bohr magneton $he/4\pi mc$; we have assumed complete quenching of the orbital angular momentum, so that the total effective angular momentum of the atom is specified by S, and the g-factor is 2. When the Brillouin function is used, the law of corresponding states of course only applies to atoms with the same spin S. It has already been mentioned that in Fig. 3, better agreement with experiment is achieved by using a low value of S than the classical $S = \infty$. This behavior is reasonable. The incomplete d shell in nickel, for instance, probably has the form d^9, with only one free spin, so that the most relevant curve should be that for $S = \frac{1}{2}$.

The greatest service of quantum mechanics, however, has been in the removal of the mystery of the large Weiss molecular fields. The puzzle was solved by Heisenberg[8] in 1928, who showed that the explanation is provided by the exchange forces characteristic of quantum mechanics. These forces cannot be described in simple intuitive language. In a certain sense these forces are purely orbital in character as they arise from the overlapping of orbital wave functions and are

[6] For the most recent gyromagnetic work, and references to earlier literature, see S. J. Barnett, Proc. Am. Acad. **75**, 109 (1944), **73**, 401 (1940); Rev. Mod. Phys. **7**, 129 (1935).

[7] For discussion of the quantum-mechanical mechanism for the quenching of orbital angular momentum in ferromagnetic solids, which is different from that in paramagnetic salts, see Sommerfeld and Bethe, *Handbuch*

der Physik, second edition, Vol. 24/2, p. 613 (Heisenberg model) and H. Brooks, Phys. Rev. **58**, 909 (1940). (Stoner model). Brooks finds that the orbital residues are of the order of magnitude to be expected theoretically.

[8] W. Heisenberg, Zeits. f. Physik **49** 619 (1928).

associated with the dependence of the energy on the type of orbital symmetry. However, because of the constraints imposed by the Pauli exclusion principle, there is a correlation between orbital symmetry and spin alignment, so that there is very large apparent spin-spin coupling, despite the fact that in a certain sense the exchange forces have nothing to do with the magnetic moment of the electron. (In chemical language, the spin alignment may be regarded as an "indicator" of the orbital symmetry). Dirac[9] showed that, apart from a constant term of no interest to us, the effective coupling between spins due to the exchange effect is equivalent to a potential energy of the form

$$V_{ij} = -2J_{ij}\,\mathbf{S}_i \cdot \mathbf{S}_j, \qquad (8)$$

where J_{ij} is the exchange integral connecting atoms i and j, and $\mathbf{S}_i$ is the spin angular momentum vector of atom i, measured in multiples of the quantum unit $h/2\pi$.

Two facts are immediately to be noted about Eq. (8). One is that exchange integrals can often be fairly large, and so it is no longer difficult to understand the large amount of interaction between the elementary magnets envisaged by the Weiss field. The other is that the potential is isotropic, or in other words, does not depend on how the spin magnets $\mathbf{S}_i$ and $\mathbf{S}_j$ are aligned relative to the radius vector $\mathbf{r}_{ij}$ joining atoms i and j, whereas this alignment was involved in the classical coupling (6) through the term $(\mathbf{\mu}_i \cdot \mathbf{r}_{ij})(\mathbf{\mu}_j \cdot \mathbf{r}_{ij})$. Hence in the first approximation ferromagnetic media should be isotropic, in agreement with the experiment. How then does one explain the fact that actually there is always some anisotropy, though it is not the main effect? One immediately recalls that besides (8), there is superposed the much smaller classical electromagnetic coupling energy (6), which will give some anisotropy. This is true, but the resulting anisotropy is not sufficiently great, and the proper explanation of anisotropy is probably to be found in spin-orbit coupling, wherein the incomplete quenching of the orbital angular momentum interferes with the freedom of the spin. We shall not pursue this

subject further, although there is an appreciable literature on the quantum theory of ferromagnetic anisotropy.[10]

Derivation of Eq. (8)

It is quite easy to derive Eq. (8) if one accepts the fact that for a two-electron system, the permitted values of the energy are

$$W = K_{ij} \pm J_{ij}, \qquad (9)$$

where K_{ij} is the so-called Coulomb integral, which we need not write down, and J_{ij} is the exchange integral

$$J_{ij} = \int \cdots \int \psi_i(1)\psi_j(2)\mathcal{H}\psi_j(1)\psi_i(2)dv_1 dv_2. \qquad (10)$$

Here $\mathcal{H}$ denotes the Hamiltonian operator, and ψ_i is the wave function for state i. Equation (9) is a well-known result of the elementary Heitler-London theory of chemical bonding. The upper and lower signs correspond respectively to singlet and triplet states. In writing (9), it has been assumed that the wave functions are orthogonal. Actually they are not because they relate to different atomic centers, but the resulting error is probably unimportant for our purposes.[11]

Now by the law of cosines we have

$$2\mathbf{s}_i \cdot \mathbf{s}_j = \mathbf{s}_{i+j}{}^2 - \mathbf{s}_i{}^2 - \mathbf{s}_j{}^2,$$

where $\mathbf{s}_{i+j}$ is the vector sum of $\mathbf{s}_i$ and $\mathbf{s}_j$. In quantum mechanics, $\mathbf{s}_i{}^2$, $\mathbf{s}_j{}^2$, and $\mathbf{s}_{i+j}{}^2$ are all diagonal matrices in the two-electron problem, such that

$$\mathbf{s}_i{}^2 = \mathbf{s}_j{}^2 = s(s+1) = \tfrac{3}{4}, \quad \mathbf{s}_{i+j}{}^2 = s_{i+j}(s_{i+j}+1)$$

where the singlet and triplet have respectively $s_{i+j} = 0$ and $s_{i+j} = 1$. Hence the characteristic values of the expression

$$-\tfrac{1}{2} - 2\mathbf{s}_i \cdot \mathbf{s}_j \qquad (11)$$

are ± 1, the upper sign applying to the singlet.

We next note that the exchange energy, whose characteristic values are given by the second member of (9), and the expression (11) are simultaneously diagonal, differing from each other merely by a factor J_{ij}. Since matrices equal in one system of representation are always equal, it follows that the exchange effects are equivalent to multiplying (11) by J_{ij}, and so (8) is proved. In writing (8) we have omitted the term $-\tfrac{1}{2}J_{ij}$ corresponding to the first member of (11) since this term does not depend on spin alignment, and is hence of no interest in the problem of magnetism. It is to be emphasized that because of the matrix identity, this proof, though starting with a two-electron system, is general, and applies regardless of how many electrons may be involved in the characteristic value problem of the whole solid body.

In this derivation of (8) we have supposed that there is only one valence electron per atom, so that s may be identified with S. The extension to the case of n electrons per atom is, however, immediate. We have merely to sum over the various electrons of each atom, noting that the

[9] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, New York, 1935) second edition, Chap. X; for the application to ferromagnetism see Chap. XII of the writer's *Electric and Magnetic Susceptibilities*.

[10] For references see J. H. Van Vleck, Phys. Rev. **52**, 1178 (1938), also Brooks, reference 7.

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

total spin $\mathbf{S}_i$ of atom i is the vector sum of the spins $\mathbf{s}_i$, $\mathbf{s}_{i+1}$, $\cdots \mathbf{s}_{i+n}$ of the various constituent electrons. Thus we have

$$J_{ij}\mathbf{s}_i\cdot\mathbf{s}_j+J_{i+1,\,j}\mathbf{s}_{i+1}\cdot\mathbf{s}_j+\cdots+J_{i+n,\,j}\mathbf{s}_{i+n}\cdot\mathbf{s}_{j+1}$$
$$+J_{i,\,j+1}\mathbf{s}_i\cdot\mathbf{s}_{j+1}+\cdots+J_{i+n,\,j+n}\mathbf{s}_{i+n}\cdot\mathbf{s}_{j+n}$$
$$=J_{ij}(\mathbf{s}_i+\mathbf{s}_{i+1}+\cdots+\mathbf{s}_{i+n})\cdot(\mathbf{s}_j+\cdots+\mathbf{s}_{j+n})$$
$$=J_{ij}\mathbf{S}_i\cdot\mathbf{S}_j. \qquad (12)$$

In (12), we have assumed that all electrons not in closed shells on a given atom have the same exchange integrals, so that $J_{ij}=J_{i+1,\,j}=\cdots=J_{i+n,\,i+n}$. Our simple form of the theory thus does not take account of the fact that, especially in a crystalline field, the d orbits are of diverse types, but without this simplification the calculation would be completely intractable. Actually, only s shells have the property that the exchange integrals are the same for all constituent electrons, and an incomplete s shell can have only one electron. (Complete shells do not contribute to (8) since their total spin is zero.) Hence the generalization to incomplete shells with $n>1$ may seem unwarranted and deceptive. However, we can, for rough qualitative purposes, consider several electrons to be effectively in a common s state if their orbital angular momenta have been quenched by the interatomic forces and if they all overlap neighboring atoms to about the same extent.

CALCULATION OF THE SUSCEPTIBILITY WITH THE HEISENBERG MODEL

The total exchange energy is obtained by summing over all pairs of atoms in the crystal. We shall make the usual assumption that the exchange integral is negligible except between adjacent atoms and has the same value J for all neighboring pairs. Then the total interaction energy is

$$V_{\text{total}}=-2J\sum_{\text{neighbors}}\mathbf{S}_i\cdot\mathbf{S}_j. \qquad (13)$$

We must now proceed to examine the magnetic properties which follow from the coupling (13). We have already intimated that (13) gives a large force between elementary magnets, and so is in at least qualitative agreement with the Weiss molecular field. We must, however, go a step further, and see if the quantitative formulas for magnetic properties resulting from (13) are reasonably close to those of the Weiss theory. The problem of calculating the magnetic susceptibility associated with (13) is far from an easy one. In fact, it can be solved only under limiting conditions, or else with simplifying approximations, usually questionable. Heisenberg[8] showed that the formulas of the Weiss theory are a consequence of the exchange coupling (13) if one rather arbitrarily assumes that all states of

the same resultant spin for the whole crystal (or microcrystal) have the same energy. Then the constant of proportionality q of the Weiss field turns out to have the value

$$q=zJ/2N\beta^2, \qquad (14)$$

in terms of the exchange integral (10). Here z is the number of nearest atoms possessed by any given atom.

An extremely simple method of getting this result has been given by Stoner,[12] and we shall now present essentially his procedure. Heisenberg's original calculations[8] by means of group theory, or the corresponding computations by means of the vector model[9] are probably more rigorous, but are much more intricate, and in any case (14) is only an approximation.

Stoner's assumption is that in considering the action of the other spins of the crystal on a given spin, the instantaneous values of the other spins may be replaced by their averages with respect to the time, which will be denoted by bars. Since all the atoms of a crystal are considered equivalent, these averages will be independent of the atom. So the potential acting on a given atom is taken as

$$V_i=-2J\sum_{\text{neighbors of }i}\mathbf{S}_i\cdot\bar{\mathbf{S}}_j$$
$$=-2zJ(S_{xi}\bar{S}_{xj}+S_{yi}\bar{S}_{yj}+S_{zi}\bar{S}_{zj}).$$

If the intensity of magnetization is directed along the z axis, then

$$\bar{S}_{xj}=\bar{S}_{yj}=0, \quad \bar{S}_{zj}=M/2N\beta,$$

and hence

$$V_i=-(zJM/N\beta)S_{zi}.$$

If we add V_i to the potential $-2\beta H S_{zi}$ due to the applied magnetic field, we see that the total energy of orientation of the spin magnet is the same as though the applied field H were replaced by $H+qM$, where q has the value (14). Since we are dealing with a quantum-mechanical rather than classical theory, we must, of course, employ formulas based on the Brillouin rather than Langevin function, or in other words use (7) rather than (2). If in (7) we use the approximation $B_S(y)=y(S+1)/3S$ appropriate to weak fields, then (7) takes the form $M=CH/(T-T_C)$

[12] E. C. Stoner, *Magnetism and Matter* (Methuen & Company, London, 1934), p. 358, Proc. Leeds Phil. Soc. 2, 56 (1930); Phil. Mag. 10, 43 (1930).

where the Curie temperature T_C and Curie constant C have the values

$$T_C = 2JzS(S+1)/3k, \quad C = 4N\beta^2 S(S+1)/3k. \quad (15)$$

Stoner's substitution of average for instantaneous values, or Heisenberg's ultimately equivalent assumption that all states of the same total crystalline spin have the same value, obviously represents only an approximation, the exact nature of which is so obscure that the theory can scarcely be said to rest on a rigorous or logically convincing basis until more accurate methods of calculation are found. As has been stressed by Néel,[13] the instantaneous molecular field acting on a given atom is by no means the same as the average, and so the Weiss theory is in a certain sense equivalent to supposing that all fluctuations in the molecular field can be neglected. Néel shows, for instance, that the fluctuations in the molecular field are responsible for the differences between the paramagnetic and ferromagnetic Curie points.

It would obviously be desirable to find a more accurate method of calculation with the Heisenberg model, especially a systematic scheme of successive approximations which will enable one to come closer and closer to the true result. Heisenberg[8] himself made an interesting attempt in this direction. He made a presumably more refined calculation on the assumption that the energy levels for a given total crystalline spin have a Gaussian distribution, whose mean square fluctuation can be computed from group theory, or, alternatively, by means of the vector model. Unfortunately the results, whose quantitative formulas we shall omit, are, if anything, less satisfying than those of the first approximation. They lead to far too much curvature in the graphs of $1/\chi$ against T above the Curie point, and to excessively drastic conditions for ferromagnetism, namely that the number of neighbors be eight or greater. Actually, ferromagnetic alloys are known in which the magnetic ions are arranged on a simple cubic lattice, where $z = 6$. On the other hand, the first approximation (15) gives too lenient conditions, allowing even a linear chain to be ferromagnetic, whereas it can

be shown rigorously by other methods that such a chain should be only paramagnetic.[14] It is thus of interest to carry out third and fourth approximations, and see whether they will lead to reasonable results, intermediate between those of the first and second. The higher approximations should be of particular interest, because they involve other details of the crystalline structure than merely the neighbor number z, which is all that enters not only in the first approximation (14–15) but also in the second or Gaussian one. For instance, it is known from another method of calculation, viz. the Bloch spin-wave theory to be discussed later, that the simple cubic grating should be ferromagnetic, whereas the hexagonal plane grating should not, although both involve six nearest neighbors. Clearly the same result should also ultimately follow from a series development which is a continuation of Heisenberg's procedure. Indeed, in the higher approximations, there appear terms which represent groupings of parallel spins in closed chains, the possibilities of which are not enumerable in terms merely of the neighbor number.

Most instructive calculations of the third and fourth approximations have been made by Opechowski.[15] He shows that, in principle, the reciprocal of the susceptibility above the Curie point should be expressible in the form

$$\frac{1}{\chi} = \frac{3k}{4N\beta^2 S(S+1)} \times \left[T - \Delta + \frac{a}{T} + \frac{b}{T^2} + \frac{c}{T^3} + \cdots \right]. \quad (16)$$

Here the first and second approximations, corresponding respectively to (14–15) and to the Gaussian assumption are obtained by terminating the infinite series with the terms $-\Delta$ and a/T respectively. Opechowski[15] has computed for the case $S = \frac{1}{2}$ of one electron per atom the values of the further coefficients b and c respectively, although the determination of the fourth-order moment necessary to evaluate c is quite difficult. If the substance has a Curie point, i.e., a temperature at which the susceptibility becomes infinite, the right side of (16) must have a zero for a real

[13] M. L. Néel, Ann. de physique 18, 5 (1932), 8, 237 (1937).

[14] H. Bethe, Zeits. f. Physik 71, 205 (1931).
[15] W. Opechowski, Physica 4, 181 (1937).

positive value of T. Opechowski finds that this property is indeed restored to the simple cubic lattice if the development is terminated at b/T^2 rather than a/T. When, however, he pushes the calculations a step further, and computes c, absurd results are obtained. Not even a face-centered lattice, where $z=12$, remains ferromagnetic. Thus for practical purposes, attempts to compute the susceptibility by means of an infinite series such as (16) must be regarded as fruitless, for the convergence is so bad that the method is not really usable.

Other methods of calculation than the series one are obviously desirable. One step in this direction has been made by Fay[16] for a simple cubic lattice. He splits the entire crystal into elementary cubes each containing eight atoms, and within each cube the exchange interactions are taken into account rigorously by means of formulas which Serber[17] has developed for the characteristic values of an eight-electron system. The coupling between electrons of different cubes is handled by methods equivalent to the first approximation of Heisenberg's theory. Thus the secular problem connected with (13) can be solved "half-rigorously," i.e., taking into account the fluctuations in half of the interactions, but employing essentially mean values for the remainder, since any given atom has three neighbors within its own cube, and three in other cubes. Fay finds it makes comparatively little difference whether the problem is treated "half-rigorously," or whether the entire problem is handled by the first approximation, where only the average fields are used throughout. Namely, the formula for the Curie point becomes $T_C = 2.55J/k$ instead of $T_C = 3J/k$, and the $(1/\chi) - T$ curves above the Curie point are nearly straight lines. Thus incorporation of half the interactions rigorously does not lead to anything like as much deviation from the simple formulas of the original Weiss theory as Heisenberg's Gaussian calculation would suggest.

Another way in which the Heisenberg model can be simplified is by changing the interaction potential. This is done in the so-called Ising model, which uses a coupling energy of the form $-2Js_{zi}s_{zj}$ rather than $-2J\mathbf{s}_i \cdot \mathbf{s}_j$ as in (13). We

use the notation s rather than S as it is assumed that there is only one free spin per atom. In a certain sense the Ising model is a purely mathematical fiction, as it neglects the interactions $-2J(s_{xi}s_{xj}+s_{yi}s_{yj})$ between the components of spin perpendicular to the direction of the magnetic field, which are often important physically. There is a considerable literature on the calculations of crystalline characteristic values and Curie points with the Ising model.[18] Such treatments have the great merit of being clear-cut and rigorous for the assumed problem. However, they have been confined primarily to one- or two-dimensional rather than three-dimensional lattices. Even were rigorous calculation with the Ising model possible for the actual lattice pattern, the results should not be identified too closely with the actual magnetic behavior of the material simply because of the inadequacy and arbitrariness of the model. Nevertheless, the Ising model is undoubtedly useful as a stepping-stone to further mathematical insight into the characteristic value problems associated with a magnetic crystal.

THE BETHE-PEIERLS-WEISS METHOD

Another and quite different procedure has recently been developed by Peter Weiss,[19] a young American not to be confused with the French physicist Pierre Weiss, the father of the molecular field, who died in 1940. The essential feature of this new calculation, not yet published because of war delays, is that it adapts to the case of ferromagnetism a method which had previously been devised by Bethe and Peierls[20] in the order-disorder problem encountered in the structure of alloys. The method has both power and generality. Unlike Fay's treatment, for instance, it is not confined to the simple cubic lattice. Peierls[20] himself showed how it could be adapted to magnetic susceptibilities by a simple change of notation in the order-disorder problem, provided one employed the Ising rather than Heisenberg

[16] C. H. Fay, Proc. Nat. Acad. Sci. 21, 537 (1935).
[17] R. Serber, J. Chem. Phys. 2, 697 (1934).

[18] E. Ising, Zeits. f. Physik 31, 253 (1935). For the most recent work with the Ising model and references to the literature see the article by Wannier, Rev. Mod. Phys. 17, 50 (1945).
[19] P. R. Weiss and J. H. Van Vleck, Phys. Rev. 55, 673 (1939) (abstract only); P. R. Weiss, unpublished doctor's thesis, Harvard University, 1941.
[20] H. A. Bethe, Proc. Roy. Soc. 150A, 552 (1935); R. Peierls, Proc. Camb. Phil. Soc. 32, 477 (1936).

model. Weiss, on the other hand, uses the true Heisenberg model, and with the latter, a calculation of the Bethe-Peierls type is both more difficult and more interesting than for the Ising potential, since it involves the quantum mechanics of angular momentum vectors, instead of classical scalar quantities.

The central idea of the procedure of Weiss is to consider a given atom i and its z nearest neighbors $i+1 \cdots i+z$ as a structural unit U. The interactions involving atom i are thus taken into account rigorously. The atoms $i+1$, $i+2$, $\cdots$, $i+z$ form a shell around i. Their coupling with each other and with the central atom i is treated rigorously, inasmuch as a quantum-mechanical system of $z+1$ spins is a soluble problem if z is not too large. It is, however, necessary to assume that the spin quantum number of each atom is $\frac{1}{2}$, because the problem becomes too complicated with more than one magnetic electron per atom. The coupling of atoms $i+1$, $\cdots$, $i+z$ with atoms exterior to the shell U is included by means of an approximation. It is assumed that the influence of the outside atoms can be represented by the addition of a potential energy term in the partition function for the system of $z+1$ atoms. The magnitude of this term is determined by the requirement that the mean magnetic moment of the central atom i be the same as that of any atom $i+1$, $\cdots$, $i+z$ in the shell surrounding i, since in the crystal all atoms are strictly equivalent on the average. Insertion of the extra term to allow for coupling of atoms exterior to U is equivalent to supposing that the outer atoms $i+1$, $\cdots$, $i+z$ of U are subject not merely to the applied field H, but rather to the sum $H+H_e$ of H and an effective field H_e. On the other hand, the central atom i is subject only to H, since the interaction with its neighbors has been taken into account rigorously, and it is for the latter reason that the Bethe-Peierls attack is such a good approximation. The effective field H_e is considerably more general than the molecular field qM used in the classical Weiss theory, and is not to be confused with the latter. In the first place H_e need not be linear in M, although this is the case above the Curie point, and even when H_e is a linear function $H_e = \alpha M$ of M, the coefficient of proportionality α is not assumed to be independent of the temperature T. Instead α is a complicated function of T determined by the requirement that atoms i and $i+k$ $(1 \leq k \leq z)$ have the same mean moment. It is to be emphasized that it is not necessary to resort to series development of α, or other functions, in descending powers of T, as in (16), and so the Bethe-Peierls scheme gives closed expressions, instead of arbitrarily terminating a series at a certain point.

We now present a brief outline of the mathematical details of the model. Since atom i interacts equally with its various neighbors $i+1$, $\cdots$, $i+z$, the resultant spin S' of atoms $i+1$, $\cdots$, $i+z$ will be a "good quantum number," which can assume the values 0, 1, $\cdots z$. The total spin of the entire system U will be S_U, where $S_U = S' \pm \frac{1}{2}$ unless $S' = 0$, in which case $S_U = \frac{1}{2}$. Let $W(S')$ be the energy of the system $i+1$, $\cdots$, $i+z$ in the absence of i and of external fields. Then when the coupling with i, and the fields H, H_e explained above are introduced, a quantum-mechanical analysis, similar in many ways to that used in the Paschen-Back effect, shows that the energy levels are

$$W_U = W(S') + 2M_U \beta H' + \frac{1}{2}J \pm [\beta^2 (H' - H)^2 + 2M_U J \beta (H' - H) + (S' + \frac{1}{2})^2 J^2]^{\frac{1}{2}}, \quad (17)$$

with $H' = H_e + H$. When M_U has any of the values $-S'$, $-S'+1$, $\cdots$, S' there are states corresponding to both sign choices in (17), but when $M_U = -S'-1$, $S'+1$ only the branch of the radical should be used for which the coefficient of J is negative. Here M_U is the magnetic quantum number associated with the projection of S_U along H. The partition function is

$$Z = \Sigma \exp (-W_U/kT), \quad (18)$$

where the summation is over the various energy states of our system of $z+1$ atoms, which are given by (17). It is to be emphasized that there are many states $W(S')$ belonging to a given value of S', the number being determined by the so-called "branching rule" for compounding angular momentum vectors. For the simple and body-centered lattices, $W(S')$ is independent of S' and may be taken to have the value zero, since for these arrangements none of the atoms $i+1$, $\cdots$, $i+z$ are nearest neighbors to each other. This statement is not, however, true of the face-centered cubic, or plane hexagonal lattices. In the face-centered case, the exact characteristic values of $W(S')$ are too difficult to determine exactly, and so Weiss is obliged to make certain simplifying approximations, which, however, still take account, at least qualitatively, of the spreading of the energy levels $W(S')$ belonging to the same S'. For the plane hexagonal grating, the determination of $W(S')$ is equivalent to calculation of the benzene ring by means of the Heitler-London model, a problem already solved by Serber[17] and others. The requirement that atoms i and $i+k$ have the same mean moment implies that

$$\partial \log Z / \partial H = (1/z)(\partial \log Z / \partial H'). \quad (19)$$

The intensity of magnetization is the same as NkT times

the expression (19), and in this fashion, the susceptibility can be found.

Weiss has considered the cubic, body-centered and face-centered cubic lattices, also the plane hexagonal grating, and the linear chain. He finds that the various types of cubic lattices should all be ferromagnetic below a certain temperature, or, in other words, possess Curie points. On the other hand, he finds that the linear chain should not be ferromagnetic, which is as it should be, for this particular arrangement can be treated rigorously, as has been done by Bethe,[14] who proved it should be only paramagnetic. It is particularly gratifying that Weiss' analysis yields no Curie point for the plane hexagonal grating, in contrast to the simple cubic lattice, which possesses the same number $z=6$ of neighbors. It is known from other methods, *viz.* Bloch's spin waves, to be discussed later, that the simple cubic grating should be ferromagnetic, and the plane hexagonal not. The different results obtained by Weiss in these two cases show that his method has considerable sensitivity, for difference between the two arrangements begins only with the term b/T^2 in (16).

It is, of course, possible to expand the results obtained by Weiss as a series in descending powers of $1/T$, although, as already emphasized, the great advantage of the Bethe-Peierls method is that it yields closed forms rather than series. However, the expansion is useful in order to let us see how far the model agrees with the first few terms of the exact development (16), which have been evaluated by Opechowski. Weiss finds agreement inclusive of b/T^2 in (16). The Bethe-Peierls model used by Weiss represents a far better approximation than Heisenberg's Gaussian assumption, but at the same time yields results not too far out of line with those of the simple molecular field corresponding to the first approximation of Heisenberg.

Weiss finds that the Curie points for the simple and body-centered lattices are respectively $1.85J/k$, $2.9J/k$. The values furnished by the first approximation (15) of the Heisenberg theory are respectively $3J/k$ and $4J/k$. With the Gaussian model,[8, 9] the simple cubic lattice has no Curie point at all, while the body-centered one has $T_c=2J/k$. It is, of course, to be expected that

the results obtained by Weiss should be intermediate between the first and second approximations of the Heisenberg theory, for the Gaussian distribution overemphasizes the effect of the spread of energy values for a given total crystalline spin, whereas the first approximation represented by (15) neglects the spreading entirely.

The graphs obtained by Weiss with the Bethe-Peierls model for the reciprocal of the susceptibility as a function of the temperature above the Curie point are remarkably close to straight lines, despite the fact that the series (16) converges so slowly that it would be entirely unwarranted to stop with the second term $-\Delta$. The curvature is greatest near the Curie point, and so one can understand, at least qualitatively, why the paramagnetic and ferromagnetic Curie points are different.

Weiss has not yet extended his calculations with the Bethe-Peierls method to the ferromagnetic region below the Curie point, where the computations become considerably more difficult. It is not, however, to be expected that this method will be a particularly accurate one at low temperatures, since there the elementary magnets are nearly parallel, and long distance order, as well as waves of propagation not envisaged in the Bethe-Peierls model, become important.

THE BLOCH SPIN WAVES

Fortunately, another method, particularly adapted to the low temperature region, was developed some time ago by Bloch.[21] The starting point of his attack is Slater's[22] observation that the characteristic values of the Heisenberg exchange coupling can be rigorously determined if the spins of all but one atom are parallel (i.e., $N-1$ atoms have $M_{z_i}=+\frac{1}{2}$, and one has $M_{z_i}=-\frac{1}{2}$). The solutions can be interpreted as representing waves of disturbance in which the reversed spin is propagated through the crystal with various possible wave-lengths. The fundamental hypothesis of the Bloch calculation is that if there are k reversed spins, the solution can be obtained by additive superposition of k solutions in which a single spin is reversed. This will be an allowable approximation only if k is small com-

[21] F. Bloch, Zeits. f. Physik **61**, 206 (1930); Kramers, Leiden Comm. Supplement 83; Moller, Zeits. f. Physik **82**, 559 (1933); W. Opechowski, Physica **4**, 715 (1937).
[22] J. C. Slater, Phys. Rev. **35**, 509 (1930).

pared with N, so that the probability of two or more reversed spins being at the same point of the lattice is small. (When reversed spins coincide, the approximation is meaningless, as one cannot turn a northward spin southward more than once.) Hence the Bloch method of calculation with the Heisenberg model is only satisfactory in the immediate neighborhood of complete saturation $M = N\beta$, and consequently only at temperatures very low compared to the Curie point. Bloch has shown that the plane hexagonal grating should never become ferromagnetic, even at the absolute zero, unlike the various cubic lattices. He shows that for the latter the formula giving the variation of magnetization with temperature is, for small kT/J,

$$M = N\beta[1 - A(kT/J)^{\frac{3}{2}}] \qquad (20)$$

where A is a numerical constant, which has the value 0.1174 for the simple cubic lattice.[23] Equation (20) leads to more rapid deviations of the moment from the limiting value $N\beta$ when the temperature is raised above zero than does the conventional first approximation (7), (14) which is the quantum version of the molecular field of Weiss. The variation of M with temperature predicted by (20) has been confirmed by experiment.[24]

INSUFFICIENCY OF THE HEITLER-LONDON MODEL—STONER'S METHOD

It is to be emphasized that all the calculations with the Heisenberg theory which we have discussed in the preceding pages are based on the Heitler-London model. The latter represents a completely non-polar approximation or, in other words, it supposes that the electrons responsible for ferromagnetism always remain on the same atom and do not participate in electrical conduction. This is an idealization never completely realized in fact. Another limiting case is furnished by the model of Sommerfeld, Wigner, Seitz, and others analogous to that of Hund and Mulliken in the theory of valence.[25] In this model of "itinerant electrons," one supposes that the $3d$

electrons circulate independently and freely from one atom to another. The resulting momentary shortage or surplus of charge on any particular atom makes the crystal instantaneously polar. Theoretical calculations from this point of view have been made by Bloch,[26] by Slater,[26] and especially by Stoner.[27] Evidently, the truth is somewhere between the Heitler-London model and that of itinerant electrons. The actual intermediate case is too difficult to calculate. Hence Stoner's calculations fulfill a very useful role in rounding out the theory, for if we have available the predictions of theory for the approximation of completely free as well as completely bound electrons then we can see qualitatively by interpolation what is to be expected between the two limiting cases. When, as fortunately often proves true, the results are quite similar in the two extremes, we can be reasonably confident that the behavior is not very different in the intermediate region.

The calculations of Stoner are based on a well-defined, clear-cut model, and so have a certain characteristic elegance. It is assumed that the electron energy levels are attributes of the whole crystal rather than the individual atom, and can be handled by the Fermi-Dirac statistics. A molecular field is used to represent the exchange interaction, which is hence treated in a somewhat phenomenological fashion, and no more completely or rigorously than in the first approximation of the Heisenberg theory. The Stoner procedure can hence be characterised as the superposition of the Weiss molecular field on the Sommerfeld theory of electronic conduction, and depicts what is called "collective electron ferromagnetism."

For readers used to thinking in mathematica terms, the content of the Stoner theory is most readily visualized by writing down the formula which he uses for the characteristic function, viz.

$$\Gamma = \int_0^\infty \nu(\epsilon) \log [1 + e^{\eta - (\epsilon - qM\beta - \beta H)/kT}] d\epsilon$$
$$+ \int_0^\infty \nu(\epsilon) \log [1 + e^{\eta - (\epsilon + qM\beta + \beta H)/kT}] d\epsilon, \qquad (21)$$

[23] The value .0587 of A given on p. 338 of the writer's *Electric and Magnetic Susceptibilities* is incorrect, as he copied a numerical error in Bloch's original paper.

[24] M. Fallot, Annales de Physique **6**, 305 (1936).

[25] For comparison of the Heitler-London-Slater-Pauling and Hund-Mulliken theories of valence see, for example, J. H. Van Vleck and A. Sherman, Rev. Mod. Phys. **7**, 167 (1935).

[26] F. Bloch, Zeits. f. Physik **57**, 545 (1939); J. C. Slater, Phys. Rev. **49**, 537, 931 (1936) and **52**, 198 (1937).

[27] E. C. Stoner, Proc. Roy. Soc. **165A**, 372 (1938); **169A**, 339 (1939); also Proc. Leeds Phil. Soc. **3**, 457 (1938); Phil. Mag. **25**, 899 (1988).
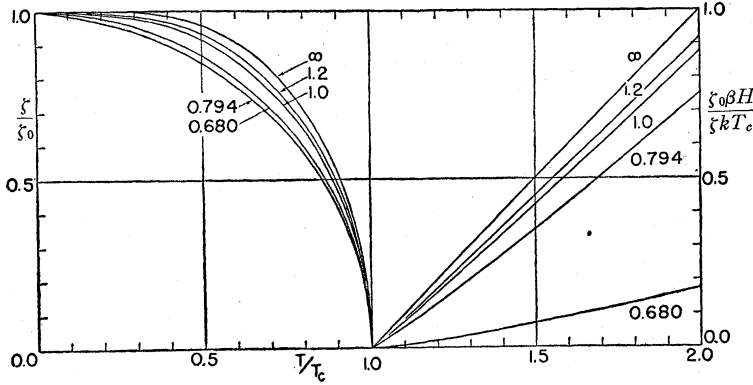
FIG. 4. The left side of the figure gives the saturation intensity of mag-
netization measured relative to the value at the absolute zero, as a function
of the reduced temperature below the Curie point in Stoner's theory. The
ordinate for the right side, which applies above the Curie temperature, is
proportional to the reciprocal of the susceptibility. The various curves are
for different values of a parameter $k\theta'/\epsilon_0$ which is proportional to the ratio of
the exchange integral to the spread in energy (so-called "Fermi energy") due
to electron migration. The limiting case $k\theta'/\epsilon_0 = \infty$ coincides with the
classical Weiss molecular field theory if the latter is applied to a Brillouin
function for $S=\frac{1}{2}$.

where $\nu(\epsilon)$ is the number of orbital states in the
energy interval ϵ, $\epsilon+d\epsilon$ and η is the threshold
constant, which is determined by the require-
ment $N=(\partial\Gamma/\partial\eta)_T$ that the total number of
electrons per unit volume be N. Once Γ is known,
the free energy F, specific heat C_V, and magnetic
moment M can be simply determined as follows:

$$F=NkT-kT\Gamma, \qquad c_V=\frac{\partial}{\partial T}\left[kT^2\left(\frac{\partial\Gamma}{\partial T}\right)_\eta\right]. \quad (22)$$

$$M=kT(\partial\Gamma/\partial H)_\eta. \quad (23)$$

The two integrals in (21) correspond to electrons
with spins respectively parallel and antiparallel
to the field H. The parts of the exponential
coefficients which are proportional to M and to H
represent the energy due to the molecular and to
the applied field, while ϵ is the translational or
"Fermi" energy. If we could neglect the fields
qM and H, the two integrals could be combined,
and we would have the usual quantum theory of
electronic conduction. Stoner assumes that the
density of energy levels $\nu(\epsilon)$ has the same form of
dependence $\nu(\epsilon)=a\epsilon^{\frac{1}{2}}$ on the energy as in the
standard Sommerfeld model, but he does not
make the assumption that the proportionality
constant a has necessarily the same value
$2\pi(2m)^{\frac{3}{2}}h^{-3}V$ as for a free electron gas, or in other
words the "effective mass" for determining the

energy levels need not be the same as the true
electronic mass. Stoner notes that for his as-
sumption $\nu(\epsilon)=a\epsilon^{\frac{1}{2}}$ "to be applicable with fair
approximation, it is merely necessary that the
band form should approximate to this standard
type over part of its range, up to and somewhat
beyond the top of the Fermi distribution. As far
as can be judged from the graphical representa-
tion of the form of the d band in nickel, computed
by the Wigner-Seitz method, the approximation
is here reasonably close."

Even with the specialization $\nu(\epsilon)=a\epsilon^{\frac{1}{2}}$, the
determination of the threshold constant M and
the solution of Eq. (23) for M, which enters on
the right side through the exponential in (21) as
well as explicitly on the left side, is not easy. The
calculations must be made numerically. The
Curie point T_C is clearly the temperature above
which (23) ceases to admit a non-vanishing solu-
tion for M when $H=0$. The situation is not as
simple as in the conventional Weiss theory, where
the law of corresponding states permits the
elimination of all undetermined constants. There
still remains as an undetermined parameter the
ratio of the energy level spacing to the constant
of the molecular field. In order to have a con-
venient dimensionless ratio ζ, $k\theta'/\epsilon_0$ Stoner uses
in place of M, a, q constants ζ, θ', ϵ_0 given by

$$\zeta=M/N\beta, \qquad \epsilon_0=[3N/4a]^{\frac{2}{3}}, \qquad k\theta'=qN\beta^2.$$

The quantity ϵ_0 has the physical significance of being the width of the filled energy band, while θ' plays the role of a sort of characteristic temperature, and ζ is the ratio of the magnetization to the value which would result were all the spins perfectly parallel.

In Fig. 4, taken from Stoner's paper, the left- and right-hand portions correspond roughly to Figs. 3 and 2. The abscissa in Fig. 4 is the reduced temperature, and the left-hand ordinate is the saturation magnetization measured relative to that at $T=0$, while the right-hand one, applicable above the Curie point is proportional to the reciprocal of the susceptibility. In comparing Fig. 4 with Fig. 3, it should be remembered that the structural unit of the Stoner theory is a single electron, and so the curve of Fig. 3 which is relevant is that for $S=\frac{1}{2}$. In fact, the curve labelled ∞ in Fig. 4 is identical with the curve $S=\frac{1}{2}$ of Fig. 3. This agreement is to be expected, for if $k\theta'/\epsilon_0$ is infinite, the spreading of the electronic energy levels is negligible compared to the effect of the molecular field, or in other words of exchange. Then the Fermi-Dirac statistics employed by Stoner reduce to the Boltzmann variety, and if we use spins with Boltzmann statistics subject to no other forces than a given phenomenological molecular field, it is immaterial whether the electrons are bound or migratory. In connection with Fig. 4, it is to be emphasized that the magnetization which is achieved at the absolute zero is not necessarily the same as that corresponding to complete parallel alignment of the spins. As ζ denotes the magnetization measured in multiples of $N\beta$, complete parallelism would demand $\zeta_0=1$. Prior to Stoner's theory, Bloch[26] had pointed out that ferromagnetism is impossible if the spread of translational energy levels is too large compared to the exchange integral. The corresponding critical value of $k\theta'/\epsilon_0$ in the Stoner theory is $\frac{2}{3}$. If $k\theta'/\epsilon_0$ is sufficiently large, the value of ζ_0 is unity, but a particularly interesting result of the theory is that in the interval $\frac{2}{3}<k\theta'/\epsilon_0<2^{-\frac{1}{3}}$ the saturation ζ_0 at $T=0$ is intermediate between 0 and the value unity characteristic of complete parallelism. The behavior in this intermediate region is shown in Fig. 5.

The Stoner theory has many noteworthy results. Among them are the following:

(1) The fact, already alluded to, that the saturation intensity at the absolute zero need not be an integral number of Bohr magnetons. This accords with experiment. At the same time, however, it should be emphasized that this explanation by the Stoner theory is open to the criticism that there is less than 20 percent difference between the upper and lower limits of the inequality for $k\theta'/\epsilon_0$ that delineates the transition zone covered by Fig. 5. Hence it should presumably happen but rarely that the exchange integral has the proper value to yield the intermediate region characterized by non-integral magneton numbers.

(2) In reduced units, the curves of saturation intensity against temperature (left half of Fig. 4) have about the same shape as in the standard molecular field theory. Hence the rough agreement of the latter with experiment does not preclude the existence of a collective energy-band structure.

(3) Above the Curie point, the reciprocal of the susceptibility is roughly a linear function of temperature, but still there is an appreciable curvature, especially near the Curie point.

(4) The slope of the asymptote of the $1/\chi - T$ curves does not have the value $k/N\beta^2$ corresponding to the first term of the ordinary Weiss-Heisenberg theory (cf. Eq. (16) with $S=\frac{1}{2}$). Here we must of course take N to be not the total number of atoms per unit volume, but rather the
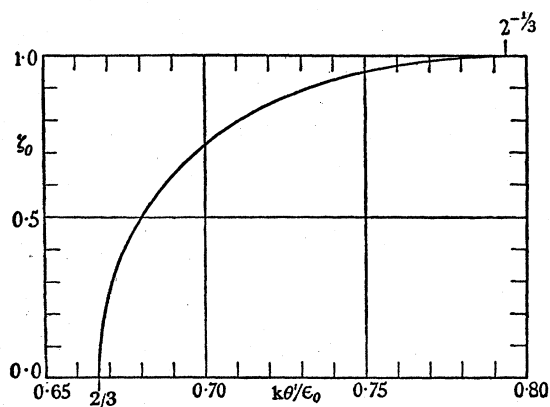

FIG. 5. This figure shows that for certain values of the exchange integral the saturation intensity of magnetization at the absolute zero in the Stoner theory is less than that corresponding to complete parallelism of the spins. The ordinate is the ratio of the saturation moment at $T=0$ to the value which would result were the elementary magnets perfectly parallel. The abscissa is proportional to the ratio of the exchange to Fermi energy.

number with free spins (e.g., 60 percent of the total number of atoms in nickel). The direction of the deviation is in accord with experiment, viz. the asymptotic value of $1/\chi$ is smaller than that furnished by the first term of (16). If, however, one uses in the Stoner theory the ratio of band width to molecular field deduced from the low saturation at the absolute zero, the predicted deviation is considerably larger than the observed. In nickel, for instance, the actual reduction amounts to only about 30 percent.

(5) The law of approach to the magnetization at the absolute zero is approximately $\zeta/\zeta_0 = 1 - bT^2$. In the case of iron or nickel, this formula is in fair accord with Fallot's experimental data,[24] though the agreement is not as good as when the Bloch expression $1 - cT^{\frac{3}{2}}$ based on the spin wave theory is used (cf. Eq. (20)). For certain alloys, however, the differences between the results obtained by using the two formulas are much more pronounced, and Fallot's measurements unequivocally favor the $T^{\frac{3}{2}}$ rather than T^2 law.

(6) There is a specific heat due to the redistribution of the d electrons among the various translational states when temperature changes. The theory shows that it should be linear in the temperature. Such a contribution to the specific heat seems to be confirmed experimentally. It is usually referred to as the "excess specific heat," and is a characteristic of transition elements such as nickel.

(7) From the observed excess specific heat it is possible to determine the ratio $\epsilon_0/k\theta'$, which is a measure of the importance of the "Fermi energy" compared to the effect of the molecular field. In nickel, the value thus obtained for this ratio agrees quite well from that deduced from the reciprocal of the susceptibility above the Curie point. On the other hand it does not check with the value of the ratio deduced from the value $\zeta_0 = .6$ of the saturation magnetization at the absolute zero. If the latter is used to determine $\epsilon_0/k\theta'$ the calculated electronic specific heat per gram atom is $2.4 \times 10^{-3}T$, whereas the observed value is $1.7 \times 10^{-3}T$. The agreement in orders of magnitude is certainly good, and, as Stoner states, the discrepancy in the exact values may be attributed to a difference in the band form in the actual metal from the standard from $\nu(\epsilon) = a\epsilon^{\frac{1}{2}}$ assumed in the computations.

COMPARATIVE MERITS OF THE HEISENBERG AND STONER THEORIES

One is apt to wonder whether from the agreement with experiment or other considerations one can deduce whether the non-polar Heisenberg model or the polar Stoner one comes closest to reality. In our opinion, it is impossible to say anything very definite on this subject. One immediately recognizes that there are a good many features common to both theories. For example, both give a reasonable behavior of the saturation below the Curie point (item (2) above). The mode of approach to the magnetization at the absolute zero is, however, represented somewhat better by the Heisenberg model (cf. item (5)). At first thought, it may seem that a non-polar theory is incapable of offering any explanation of the excess specific heat, or of the fact that the magnetization at the absolute zero is not an integral number of Bohr magnetons per atom (items (1) and (6)). This is no longer true if one uses a slight variant of the Heisenberg theory in which not all atoms have a free spin. As a typical case, let us consider nickel. Here the usual Heisenberg model with one spin per atom is to be identified with a shortage rather than surplus of one electron compared with a closed shell, in other words with the configuration d^9 rather than d. Now, all told, there are 10 electrons per atom in nickel in addition to the electrons in the interior part $1s^2 2s^2 2p^6 3s^2 3p^6$ of the atom which are so tightly bound that they can be omitted from consideration. These 10 electrons have their choice of being housed in $4s$ or $3d$ states. The neutral Ni atom has the configuration $4s^2 3d^8$, indicating that lower energy is obtained by occupancy of the two $4s$ places than by filling out the $3d$ shell to the completed group $3d^{10}$. In the solid however, the $4s$ wave functions of adjacent atoms will overlap considerably, and as a result the totality of $4s$ states for the crystal will be spread into a quite wide band, which is presumably broad compared with the $3d$ band. As $4s$ configurations are generally not ferromagnetic, the spins will be paired in the portion of the $4s$ band which is occupied, and hence will not contribute to the magnetism. Probably the $4s$ band is so wide that the upper portion of it is higher than the $3d$ band, which may in first approximation be treated as infinitely narrow in

comparison. If we assume that 70 percent of the $4s$ band is above the $3d$ band, then the most stable configuration will be that in which on the average 0.6 of an electron per atom is in a $4s$ state, and 9.4 electrons in $3d$, so that symbolically the configuration is $4s^{.6}3d^{9.4}$. The fact that the saturation magnetization at $T=0$ is .6 Bohr magneton per atom is then explained without the need of invoking the Stoner mechanism in the intermediate case shown in Fig. 5. How about the excess specific heat with this picture? The energy is undoubtedly different when two $3d^9$ atoms are adjacent than when a $3d^9$ is next to a $3d^{10}$. There will hence be a specific heat associated with the redistribution of the $3d^9$ and $3d^{10}$ states among the different lattice sites when the temperature is varied. Sixty percent of the total number of atoms are always to be kept in the configuration $3d^9$, and forty percent in $3d^{10}$, but there are an enormous variety of ways in which the two varieties of atoms can be peppered about in the crystal. This effect is rather hard to treat in a quantitative way, although some preliminary calculations have been made by H. Hurwitz.[28] It is not known yet, for instance, whether the resulting specific heat is linear in the temperature, or whether the agreement on the order of magnitude of the parameters as determined by the various criteria is as good as discussed in item (7) for the Stoner theory. Still, Hurwitz shows that it is probable that the specific heat is not vastly different in order of magnitude from that furnished by the Stoner collective electron mechanism. Incidently if we had a pure Heisenberg model with one free spin per atom, there would still be some specific heat, due to the spectrum of energy states for the exchange energy of the crystal, or in other words due to the fluctuations in the molecular field. At low temperatures this specific heat can be treated by the Bloch spin-wave method. It is, however, proportional to $T^{\frac{3}{2}}$ and so not large enough to account for the observed excess specific heat, which is linear in T.

Our suggestion of a lattice built on $3d^8$ and $3d^9$ atoms is in a certain sense a compromise between the simple Heisenberg picture using only one type of atom and the Stoner model of freely migrating

electrons. The latter neglects the interplay between the various electrons, so that it is possible for large surpluses or deficits of charge to pile up on a given atom, which can hence be regarded as momentarily existing in configurations such as $3d^9$, $3d^8$, $3d^7$, etc. In other words the Stoner theory neglects what Wigner and Seitz[29] call the correlation energy. In our opinion the correlation effect is a rather large one, as the indications are that the spread in $3d$ levels due to inter-atomic overlap is smaller than the higher ionization potentials for removal of several electrons which would have to be overcome in order for the electrons to migrate freely through the crystal without much regard to one another. In other words, our generalization of the Heisenberg-Heitler-London model does not involve the high degree of instantaneous polarity found in the unmodified collective electron viewpoint, and so perhaps comes closer to the truth. In the last analysis, it is quibbling to try and say which idea is right, since the addition of appropriate correction terms, *viz.* the states of higher ionization in the former and correlation effects in the latter, ultimately merge the two together. Incidentally, there is, as far as we can see, no reason why in the Stoner model the number of $3d$ electrons per atom, or rather number of holes in the $3d$ shell per atom, need be taken as unity. The unit value assumed by Stoner applies only if the $4s$ shell is half filled, so that the average configuration is $4s3d^9$. This will be the case if the $4s$ band splitting is rather symmetrical, so that about half of the $4s$ band is above, and half below the $3d$ band. However, there is no reason why this should necessarily be true. With two undetermined parameters, *viz.* the total number of d electrons and the ratio $k\theta'/\epsilon_0$, the extended Stoner theory would have so much leeway in the way of undetermined constants that it would be hard to tell whether any agreement obtained with experiment would be significant or fortuitous. The fact that the original Heisenberg theory has no undetermined constants when reduced units are used must be regarded as both an element of strength and

---

[28] H. Hurwitz, unpublished doctor's thesis, Harvard University, 1940.

[29] E. Wigner, Phys. Rev. **46**, 1002 (1934); Trans. Faraday Soc. **34**, 678 (1938). In the latter of these references Wigner shows that the correlation corrections are so large that numerical calculations of specific heat made with the usual collective electron model uncorrected for correlation are open to serious question.

weakness in this theory—strength in that any agreement achieved with experiment is not the result of an *ad hoc* choice of constants, but weakness in that not enough flexibility is included to allow for the complicated structure of a real solid.

It still remains to compare the Heisenberg and Stoner theories in their predictions regarding the reciprocal of the susceptibility as a function of temperature above the Curie point (items (3) and (4) of our preceding discussion of Stoner's work). We believe it is futile to try and decide between the two theories on this ground. We have seen that Opechowski's work shows that the series in (16) converge so badly that it is silly to correlate with the first two terms of (16) the slope of a straight line which happens to represent fairly well the theoretical dependence of $1/\chi$ on T over most of the range of T accessible to observation. Thus even in the Heisenberg theory one should not expect agreement between the magneton numbers deduced from saturation at $T=0$ and from use of the first two terms of (16). More specifically, the calculations of Weiss with the Bethe-Peierls method indicate that above the Curie point, except near the latter, the graph of $1/\chi$ against T does not deviate greatly from a straight line. On the whole, with any point of view, it is easier to understand the existence of curvature in the relation between $1/\chi$ and T, or the discrepancy between the paramagnetic and ferromagnetic Curie points, than it is why the relation is as nearly linear as it is empirically. Namely, with any type of theory one always neglects certain corrections or perturbations the inclusion of which will presumably introduce more curvature. The observed linearity would seem to indicate that there is a fortuitous compensation of the various effects, and that the slope does not correspond to the value obtained by using the first term of some series development but rather that a straight line is simply a rough approximation to a complicated analytical relation. Among the various complications which affect the relation between $1/\chi$ and T are the following:

(a) the fact that exchange coupling cannot be represented rigorously by the usual molecular field. In other words there exist fluctuations in the latter which are not covered by the conventional first approximation of the Heisenberg theory;

(b) distortion due to the remains of spin-orbit coupling which persist despite quenching of most of the orbital angular momentum in a solid, and which interfere with the freedom of spin orientation;

(c) in an ionic model, the migrations of electrons through the crystal;

(d) in a generalized Heitler-London model, the redistribution of the configurations d^9 and d^{10} among the different lattice sites.

The calculations of Weiss with the Bethe-Peierls method allow in principle for (a), and those of Stoner include (c), but not (a). In unpublished work, Hurwitz[28] has investigated qualitatively the effect of (d), having to rely mainly on examination of how the badly convergent series development (16) is affected by including (d). He finds that (d) increases the fluctuations of the molecular field which are present even with all atoms alike, and so tends usually to accentuate curvature. None of the existing calculations have considered (b) in any detail. All told, the observed amount of linearity in the $1/\chi - T$ relations must be regarded as something of a mystery.

In order to have simpler conditions, it is illuminating to have experimental data on metallic gadolinium (not to be confused with the celebrated salt gadolinium sulphate, which is only paramagnetic). Such data are provided by the noteworthy experiments of Trombe.[30] As the $4f$ electrons responsible for the magnetism of the rare earth group are deeply cloistered in the interior of the atom, and as the Gd^{+++} ion is in an S state, effects (b) and (c) are presumably wanting. Trombe finds indeed that the curvature in the relation between $1/\chi$ and T is considerably smaller for gadolinium than for nickel. Also he finds that at $T=0$ the saturation magneton number has, within the experimental error, exactly the right value for a Gd^{+++} ion, indicating that we have here to consider only atoms in a single configuration. What is especially interesting is that above the Curie point, the susceptibility is, except near T_c, well represented by using a Curie constant $35N\beta^2/3k$. This is the

---

[30] M. F. Trombe, Ann. de physique 7, 385 (1937).

value corresponding to the first approximation of the Heisenberg theory, in which we terminate the development (16) with Δ and use $S=7/2$ since the Gd^{+++} ion is in an ^8S state. It is clear that since the $4f$ electrons are deeply buried in the atom, one must use the Heisenberg rather than Stoner theory in gadolinium. However, it is not apparent why the fluctuations in the molecular field, whose effect is represented by the higher members of (16) should be so unimportant. In other words, there is still the complication (a) even if (b), (c), and (d) are absent, and we have indicated that (a) alone is enough to change the apparent value of the Curie constant over the domain in which linearity of $1/\chi$ in T holds. There thus appears to be something of a contradiction between iron or nickel and gadolinium. In other words the simple theory seems to work too well for the latter. Possibly the fact that the gadolinium atom has a rather high spin, *viz.* $S=7/2$ makes the fluctuations smaller than for an atom with a comparatively low spin such as nickel. There is some reason to believe that this is the case, as calculation shows[31] that in (16) the ratio a/Δ^2, which is dimensionless and a measure of the effect of the first member of the fluctuation corrections, is inversely proportional to S^2+S.

QUESTION OF THE SIGN OF THE EXCHANGE INTEGRAL

In order to have ferromagnetism, it is essential that the exchange integral defined by (10) be positive. Otherwise the exchange effects tend to destroy rather than amplify the magnetic moment, or in other words the constant of the molecular field is negative and has a demagnetizing effect. Hitherto in this article we have tacitly omitted examination of whether the exchange integral had the proper sign to produce ferromagnetism. Obviously it would be desirable if one could calculate this sign theoretically, for then one would be able to predict whether a given material might be ferromagnetic. We use the word "might" rather than "is" because the work of Bloch and Stoner shows that even with a favorable sign of the exchange integral, ferromagnetism is suppressed if the spreading of energy levels due to electron migration is too

large. Unfortunately it is so difficult to calculate exchange integrals in ferromagnetic materials that this has not been done. Even for electrons belonging to free hydrogenic atoms the computation of these integrals would be laborious, and undoubtedly interatomic forces distort the d electrons so badly that conclusions based on calculations with free atoms would probably not be reliable, at least without some further analysis.

There is, however, one working rule, first developed by Slater,[32] which seems to furnish a rather convenient qualitative index as to whether ferromagnetism is possible.[33] It states that ferromagnetism can exist only if the atomic radius is small compared to the interatomic distance.

Furthermore, Sommerfeld and Bethe[34] have shown that the Slater rule has a simple, though rather crude quantummechanical explanation, as follows. The bulk of the contribution to the integral (10) presumably comes from regions, in the six-dimensional coordinate space, which are exterior to the outermost node of the wave functions of both electrons, at least provided the azimuthal quantum number is zero or small. The contribution of the region inside the nodes is considered of minor significance by Sommerfeld and Bethe because the wave functions will not overlap sufficiently to penetrate strongly the interior portion of a neighboring atom and also to a certain extent because the volume covered by the interior regions of the atom is small. Thus, though this is perhaps the weakest part of the argument, the dominant part of the integral is considered to arise from regions for which the factor $\psi_i(1)\psi_j(2)\psi_i(2)\psi_j(1)$ is positive, so that the sign behavior of the integral can be guaged by examining the sign of $\mathfrak{IC}$. After allowing for terms that cancel out because ψ_i, ψ_j satisfy respectively the wave equations for free atoms i and j respectively, it is found that the significant parts of $\mathfrak{IC}$ are those corresponding to the potential energy due to the nucleus of the other atom, with allowance for screening by non-valence electrons, and to the mutual potential energy of the two electrons involved in the exchange integral. If the nuclear attractive term predominates over the inter-electronic repulsive term, the factor $\mathfrak{IC}$ will be essentially negative; otherwise it will be positive. Now if the radius of the atom is at all comparable with the interatomic distance, there will be appreciable overlapping of the wave functions, i.e., a significant factor $\psi_i(1)\psi_j(2)\psi_i(2)\psi_j(1)$, at points where the interelectronic repulsion is overshadowed by the nuclear term inasmuch as a nucleus contains many electronic charges. On the other hand, when the atomic radius is very small compared with the interatomic distance, the only appreciable overlapping will be in a more or less void space midway between the two atoms, where the influence of both

---

[31] See J. H. Van Vleck, J. Chem. Phys. **5**, 320 (1937), especially Eq. (56) and footnote 20.

[32] J. C. Slater, Phys. Rev. **35**, 509 (1930); **36**, 57 (1930).
[33] E. C. Stoner, Proc. Leeds Phil. Soc. **2**, 391 (1933).
[34] A. Sommerfeld and H. A. Bethe, *Handbuch der Physik*, second edition, Vol. XXIV/2, p. 596.

atomic centers is unimportant and the important energetic term arises only because of the possibility of the two electrons themselves sometimes being close together in this midway region. Hence one can see qualitatively why in order for the repulsive terms to outweigh the nuclear effects and give a positive exchange integral, one must satisfy the Slater criterion that the atomic radius be small compared to the interatomic distance. The Sommerfeld-Bethe argument, moreover, goes further and indicates why ferromagnetism is found only for d and f electrons and never for s electrons. Namely, the atomic wave functions contain a factor r^l, where l is the azimuthal quantum number. In s-states the absence of a node at the nucleus presumably makes the negative contribution to the exchange integral predominate if the atomic diameter is not completely negligible compared to the inter-atomic distance, so that there is some overlapping. Hence we can understand why s-electrons never give ferromagnetism. On the other hand, in d- or f-states the factor r^l makes the integrand small in the interior of an atom where the nuclear forces are greatest. In order for there to be ferromagnetism it is not sufficient just that the exchange integral be positive, but in addition it must be sufficiently large to predominate over any energy band structure due to electron migration (cf. Fig. 5). The factor r^l is helpful not merely because it enhances the likelihood of a positive sign, but also because it will make the integral larger once it has become positive, by accentuating the outer regions where the interelectronic repulsion is the prime effect. We have not said anything about p electrons, but clearly it would be optimism to expect this very qualitative argument to provide us with any definite predictions regarding this intermediate case.

FEEBLE PARAMAGNETISM

From the preceding discussion it is apparent that the large positive exchange integrals necessary for ferromagnetism will be realized only under rather special favorable conditions. Normally the exchange integral will be negative, and the spins will tend to set themselves anti-parallel. This is the situation ordinarily encountered for the electrons responsible for chemical bonding. In fact, the usual theories of valence, electron-pairing, etc., are founded on the idea that a saturated chemical bond corresponds to the compounding of the individual atomic spins to a zero resultant molecular spin. This is just the reverse of the state of affairs necessary for ferromagnetism, since the latter demands that the configuration of lowest energy be one of high rather than compensated spin. Thus ferromagnetism may be regarded as a sort of exception to the usual valence rules.

For the normal case of negative exchange integrals, the effect of exchange is demagnetizing,

and may roughly be described by taking the Curie temperature T_C to be negative in the Weiss-Curie formula $\chi = C/(T - T_C)$, since T_C has the same sign as the exchange integral. If the exchange coupling is powerful enough so that $-T_C$ is large compared with T, then the susceptibility will be independent of temperature, and much smaller than for a free atom, where $T_C = 0$. Thus a mechanism is provided for understanding why many substances, far outnumbering the strongly paramagnetic, or ferromagnetic ones, have a feeble paramagnetism which is more or less independent of temperature. In a feebly paramagnetic material, the spin of an individual atom can be as large as in a highly magnetic substance, but the exchange effects are demagnetizing and tend to make the resultant spin of the entire crystal small. A typical metal of this type is aluminum, which is only weakly paramagnetic despite the fact that the normal state of the Al atom is ^2P.

Strictly speaking, any discussion of feeble paramagnetism is outside the scope of a report on ferromagnetism. However, weakly paramagnetic materials are so common, and the theoretical explanation of their behavior ties in so closely with that of ferromagnetics, that before closing it has seemed worth while to comment on them briefly. We should mention that there are other effects besides negative exchange coupling which enfeeble paramagnetism. In particular there is the suppression of paramagnetism due to the spreading of energy levels caused by the migration of the electrons through a solid conductor. Pauli[35] first showed that this dispersion of translational energy levels removes most of the spin paramagnetism when there is no exchange coupling, and we have seen that Bloch[26] and Stoner[27] have demonstrated that this can still be true even when the exchange integrals have a sign favorable to ferromagnetism, provided they are not too large. Since the wave functions for s and p electrons usually overlap adjacent atoms considerably, the migration effects are probably sufficient, regardless of the sign of the exchange integrals, to suppress most of the paramagnetism in solids not having d or f electrons. Special men-

[35] W. Pauli, Zeits. f. Physik **41**, 81 (1927); an excellent review of this subject is given in sections 29 and 140 of Seitz' book, *The Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940).
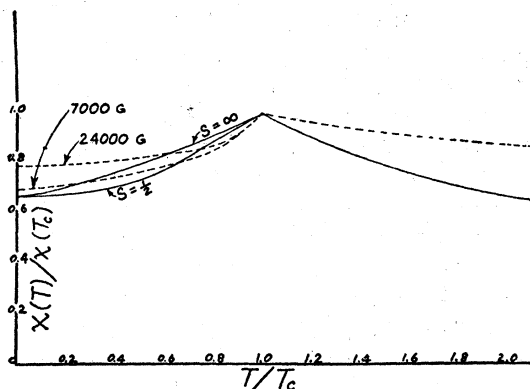
FIG. 6. The susceptibility $\chi(T)$ as a function of the reduced temperature in an anti-ferromagnetic material. The solid curves are theoretical, while the dashed lines are measurements on MnO by Bizette, Squire, and Tsai. The scale of ordinates is such that the susceptibility is unity at the Curie point.

tion should be made of the fact that in the non-ferromagnetic elements of the transition groups, the paramagnetism, though still of what may be called the "feeble" variety, is considerably larger than usual for this category. Precisely this behavior is to be expected theoretically, since the wave functions of d electrons overlap other atoms less than do those of the s or p type, and in consequence there is a narrower band structure, with less suppression of paramagnetism. For a very interesting description of the experimental data on this subject, data on the widths of the energy bands, etc., the reader is referred particularly to the discussion in Mott and Jones' book.[36]

So far we have cited only feeble paramagnetism which arises because of nearly completing quenching of the magnetic moments of atoms which are paramagnetic when free. Sometimes distortion due to interatomic forces can originate a weak paramagnetism in materials whose isolated atoms or ions are diamagnetic. This situation is more likely to arise in an ionic salt than in a conductor, inasmuch as electrons of ions in closed shells usually migrate but little. Examples are probably furnished by Sc_2O_3, TiO_2, CeO_2, and numerous cobaltamines.[37]

---

[36] Mott and Jones, *The Theory of the Properties of Metals and Alloys* (Oxford University Press, New York, 1936), pp. 189–200. Also see F. Seitz, *The Modern Theory of Solids*, section 29.

[37] For further details see p. 302 of the writer's *Electric and Magnetic Susceptibilities*.

ANTI-FERROMAGNETISM

In the preceding discussion we have stressed how difficult it is to know whether feeble paramagnetism is to be attributed to negative exchange integrals rather than to some of the other mechanisms which we have mentioned. There is, however, one class of materials, known as "anti-ferromagnetics" in which it is quite clear that the suppression of paramagnetism is to be identified with exchange coupling. These substances, which are not very common, have a susceptibility which passes through a maximum when the temperature is raised, as shown in Fig. 6. On the other hand, ordinary feeble paramagnetism is characterized by a susceptibility which is more or less independent of temperature, but which usually decreases monotonically and slowly as the temperature is raised. Examples of anti-ferromagnetic media include MnO, MnSe, Cr_2O_3, and CrSb.

The peculiar behavior illustrated in Fig. 6 is nicely explained theoretically in the following way. Suppose that we have a crystal whose constituent atoms can be resolved into two sublattices A and B such that the nearest neighbors of the atoms of A are atoms belonging to B, and *vice versa*. The simple cubic and body-centered cubic lattices are both of this type. In our rudimentary derivation of the Weiss molecular field from the exchange potential (8), we assumed that on the average the spins of all atoms pointed in the same direction. This supposition is admissible if the exchange integral is positive, as then the state of lowest exchange energy is that of completely parallelism. With a negative integral, however, the exchange energy of two atoms is a minimum if their spins are anti-parallel. Hence the configuration of deepest energy for the crystal as a whole is that in which the spins of sublattice A all point northward, and those of B all southward or *vice versa*. Consequently if a method based on mean values is used, the theory should be generalized by assuming that the mean direction of the spin in sublattice A is the opposite from that in B, apart from polarization effects due to an external applied field. With this generalized model, a "staggered" non-vanishing molecular field can exist which tends to direct the spins of A in one direction and those of B in the other. We shall

omit the mathematical analysis, which is due to Néel,[38] Bitter,[39] and the writer.[40] It yields the result that there is a certain critical or Curie temperature above which the spontaneous field of this character disappears. With similarly rather than oppositely directed spins for the two sublattices, no spontaneous field would have been possible at any temperature. If one treats the staggered model by a method of approximation analogous to the usual first-order one of the Heisenberg theory, then it turns out that the formula for the Curie temperature above which the staggered molecular field disappears is precisely the same as that (15) obtained for the ferromagnetic rather than anti-ferromagnetic case, except that now one must use the absolute value of the exchange integral. (Without use of the absolute value, the critical temperature would be negative, and hence meaningless.) As the temperature is lowered, the inner field becomes stronger and stronger, and so it is harder for the field to make the spins all point in one direction when they want to be arranged alternately in opposite directions. Hence, without going through the analysis, one can see qualitatively why the susceptibility diminishes as the temperature is lowered from the Curie temperature to the absolute zero. Above the Curie temperature, on the other hand, the staggering effect is washed out, and the susceptibility decreases with increasing temperature, obeying the familiar formula $C/(T+\Delta)(\Delta>0)$ obtainable even without introducing interlocking lattices. At the absolute zero, the inhibiting effect of the powerful internal fields on any change of alignment due to a weak external field is complete only if the latter has, apart from sense, the same direction as the former. If applied perpendicular to the alternating inner fields, the external field can still give rise to an outstanding moment by twisting slightly the orientations of the elementary magnets. Since the molecular field is presumably directed at random, the parallel and perpendicular cases, considered respectively by Bitter and Néel, should be weighted in the ratio $1:2$, and it is for this reason that the susceptibility χ_0 at the absolute zero for weak fields is theoretically

[38] L. Néel, Ann de physique 17, 64 (1932); 5, 256 (1936).
[39] F. Bitter, Phys. Rev. 54, 79 (1937).
[40] J. H. Van Vleck, J. Chem. Phys. 9, 85 (1941).

$\frac{2}{3}$ that χ_C at the Curie point. The observed values of χ_0/χ_C, instead of being 0.67, range from 0.3 to 0.85, so that the agreement is qualitative rather than quantitative. It is especially to be emphasized that in a truly ferromagnetic medium a comparatively weak field will suffice to orient the moments of the elementary domains and make them parallel to the applied field, producing so-called saturation. On the other hand, in an anti-ferromagnetic material, even a very powerful field will produce but little rotation of the domain alignments, as the low susceptibility makes the energy of the crystal rather insensitive to external fields. One can understand qualitatively why the observations of Bizette, Squire, and Tsai[41] on MnO shown in Fig. 6 reveal a somewhat higher susceptibility, below the Curie point, for a field of 24,000 gauss than for one of 7,000 gauss. Namely, in an anti-ferromagnetic medium a very powerful field will rotate the elementary domains into more favorable alignment, i.e., make the molecular field of more domains perpendicular to the applied field, instead of being random. In the limiting case of complete perpendicularity, the susceptibility below the Curie temperature T_C is independent of temperature and has the same value as at T_C. The rather mild dependence of susceptibility on field strength shown in Fig. 6 is to be contrasted with the behavior of a truly ferromagnetic material where a field of less than 7000 gauss would have produced complete parallelism. When the susceptibility is finite but still is somewhat influenced by field strength, one should expect some weak hysteresis phenomena, and in MnS a feeble hysteresis is indeed observed, wherein the loop is so narrow that it encloses but little area.

In the preceding paragraphs we have perhaps inadvertently given the impression that feebly paramagnetic and anti-ferromagnetic media represent two distinct cases. Actually there is no sharp dividing line between the two, though a convenient criterion for anti-ferromagnetism is a maximum susceptibility at a particular temperature. Also our statement at the beginning of the discussion on feeble paramagnetism that the temperature dependence of susceptibility is of the form $C/(T+|T_C|)$ in the first approximation

[41] H. Bizette, C. F. Squire, and Tsai, Comptes rendus 207, 449 (1938); C. F. Squire, Phys. Rev. 56, 922 (1939).

comparable with the Heisenberg theory is seen to be not really correct below the Curie point if a division into alternating sublattices is possible, for then one has a behavior of the type shown in Fig. 6. In either case, however, the variation with temperature is small, in fact less with the situation covered by Fig. 6, which gives $\chi_0 = \frac{2}{3}\chi_c$ rather than $\chi_0 = 2\chi_c$. The fact that most weakly paramagnetic materials do not show the peculiar maximum in χ may mean that the crystal structure is not such as to permit the requisite division into sublattices, or much more likely, that the effects of electron migration are appreciable. With the Stoner collective electron mechanism, localization on individual atoms is lost, and no molecular field of consequence is possible as long as the mean exchange integral is negative, since staggering requires local order. Here we thus have a case where the results are different with the approximations of bound and itinerant electrons, and the former must be closer to the truth when a behavior such as that shown in Fig. 6 is realized.

CONCLUSION

In closing we should like to stress once more the great parallelism in many respects between the theory of ferromagnetism and the theory of valence. Both topics are readily tractable only for models corresponding to the limiting cases in which the electrons are treated respectively as tightly bound on individual atoms and as wandering freely from atom to atom without cognizance of each others' positions. The original Heisenberg theory of magnetism is the analogue of the Heitler-London-Slater-Pauling theory of valence, while the Stoner treatment resembles the Hund-Mulliken approach. In both subjects the truth is doubtless intermediate between the two extremes, and the results in the limiting cases usually do not differ too much from each other, giving one some confidence in a similar behavior in the intermediate region. In both valence and ferromagnetism, the agreement with experiment is often surprisingly good in the first approxima-

tion and in many cases spoiled by inclusion of the second or third approximation, leaving it something of a mystery why so nearly the first-order result would presumably be restored were it possible to carry through the calculations rigorously. Along with the similarity that we have noted, there are two main differences between the theory of ferromagnetism and that of the chemical bond. In the first place, the former requires a positive rather than negative exchange integral, in order that the energy be lowest for parallel rather than anti-parallel spin alignment. Secondly, in magnetism the entire crystal rather than the molecule is the structural unit, and consequently the potential mathematical and physical complexity is greater than for even the worst organic molecule.

Has the quantum theory of ferromagnetism been a success? The answer depends to a large extent on whether one takes the attitude of the optimist or the pessimist, that is to say, whether one concentrates on the successes of the theory or looks for the flaws and emphasizes all the difficulties of rigor and convergence. Analogous remarks apply to the quantum theory of valence.[25] If one demands a theory that will predict in advance just which alloys will be ferromagnetic, what quantitatively will be the remanence, hysteresis, etc., then one must admit failure, although as time progresses, greater refinement of the theory will perhaps increase its usefulness for qualitative prognostication. Similarly one is doomed to disappointment in the present applications of quantum mechanics to chemistry if one expects advance information, say, on just which substitutions in a benzene ring will give particular physical and chemical properties such as, say, that of photographic developer, etc. If, on the other hand, what one desires of a theory is a qualitative understanding of why nature works as it does, and the satisfying confidence of having a mechanism which would, no doubt, explain even the most complicated phenomena if the difficult mathematical calculations could be made, then the quantum theory of ferromagnetism or of valence is certainly a success.