Magnetic Susceptibility of a Diamagnetic Electron Gas— The Role of Small Effective Electron Mass

E. N. ADAMS II
Institute for Nuclear Studies, University of Chicago, Chicago, Illinois
(Received July 21, 1952)

An exact formula for the "steady" bulk susceptibility of a degenerate electron gas is derived. The deviations from the Landau-Peierls formula are studied for several situations in which the effective electron mass is small. It is concluded that, in any case for which the Landau-Peierls susceptibility differs from the Landau susceptibility, the total susceptibility differs from the Landau-Peierls susceptibility by an amount of the same order of magnitude.

The magnetic susceptibility of solid Bi is discussed on a model developed by H. Jones. It is found that for several reasons, Jones' model requires major modification. A particular modification is suggested as a possible basis for a theory of the Bi susceptibility.

Some implications of our conclusions are discussed in connection with Bardeen's theory of superconductivity and the theory of the de Haas-van Alphen effect.

1. INTRODUCTION

ALTHOUGH some metals are ferromagnetic or strongly paramagnetic, most metals have rather small magnetic susceptibilities of the order of 10^{-6}. The suceptibility of such an "ordinary" metal arises partly from the magnetic moments of electrons bound to lattice ions and partly from the moments of the conduction electrons, the two kinds of susceptibility being of roughly the same order of magnitude. If the conduction electrons behaved as a gas of free electrons, the second contribution would be paramagnetic and one would expect that the sign of the susceptibility would depend on whether the diamagnetism of the ions or the paramagnetism of the conduction electrons predominates. Actually, certain metals, e.g., Bi, exhibit a diamagnetism which, although weak, is conspicuously greater than any which could be accounted for as a contribution of electrons bound to lattice ions. In such cases it is necessary to attribute the additional diamagnetism to the conduction electrons. Clearly, if the large diamagnetism results from the orbital moments of the conduction electrons, however, the electronic motions in the lattice potential must be quite different from the motions in a free electron gas. Thus, in order to understand the susceptibility of a diamagnetic metal, one must know something of the way in which the electrons move in the lattice potential of the metal.

The basic theoretical analysis of the "orbital" susceptibility of conduction electrons is a result of the efforts of Peierls,[1] who worked in the approximation of "tight binding." The same problem has been worked through by Wilson,[2] who has discussed it in the approximation of nearly free electrons, although his analysis is quite general. The only serious assumption in these treatments is that the electron-electron interaction may be ignored except in so far as it can be represented by a potential which has the periodicity of the lattice and which is the same for every electron. That assumption will also be made in this paper. The results of their analyses may be summarized as follows: the calculated susceptibility splits naturally into two parts, the first part, given by the Landau-Peierls formula, being

$$\chi_{LP}=-\tfrac{1}{3}\mu_B{}^2 n(\zeta)\langle\alpha^{xx}\alpha^{yy}-[\alpha^{xy}]^2\rangle. \qquad (1.1)$$

Here μ_B is the Bohr magneton, ζ is the Fermi energy, $n(\zeta)$ is the level density at the Fermi surface, and, for example, α^{xy} is the product of the electron mass by the xy component of the reciprocal mass tensor; the angular bracket denotes an average over the free Fermi surface, which need not lie in a single band. The second contribution to the susceptibility is much more complicated, and will be discussed in detail later. At this point, it is sufficient to remark that it gets its greatest contribution from terms which have the general appearance of second-order energies so it may be thought of as arising because of virtual transitions between bands.

A number of attempts have been made to account for the observed susceptibilities of some of the diamagnetic metals on the basis of the theory referred to above.[3] In the applications it has usually been assumed that the major contribution to the orbital susceptibility is given by Eq. (1.1). On this assumption, it is easy to show that a large diamagnetism cannot arise solely as a consequence of a high level density $n(\zeta)$. For if one replaces the angular bracket in (1.1) by unity, its value for free electrons, one obtains for the susceptibility

$$\chi_L=-\tfrac{1}{3}\mu_B{}^2 n(\zeta). \qquad (1.2)$$

Formula (1.2) was first derived by Landau[4] for the case of a gas of free electrons. The susceptibility χ_L is just one-third as large as the Pauli[5] susceptibility χ_S,

[1] R. Peierls, Z. Physik **80**, 763 (1933).
[2] A. H. Wilson, *The Theory of Metals* (Cambridge University Press, Cambridge, 1936).
[3] See, for example, N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Clarendon Press, Oxford, 1936), for a discussion of some applications.
[4] L. Landau, Z. Physik **64**, 629 (1930).
[5] W. Pauli, Z. Physik **41**, 81 (1926).

given by

$$\chi_S = \mu_B^2 n(\zeta). \tag{1.3}$$

The susceptibility χ_S is derived from the spin magnetic moment of the electron. A comparison of the expressions (1.2) and (1.3) shows clearly that a free electron gas is paramagnetic, and that a mere change in level density does not lead to diamagnetism. Thus, if the electron gas model is adequate for an understanding of the observed susceptibilities of the diamagnetic metals, it must be supposed that for some metals the angular bracket in Eq. (1.1) has a value considerably in excess of three.

The angular bracket in Eq. (1.1) has a simple physical interpretation. Because the electron gas is in a potential, a magnetic field does not cause an electron to precess in the manner of a free electron, but rather in a more complicated manner. Thus in the expression (1.2) for the susceptibility one should not use the Bohr magneton computed from the free electron mass, since the electron precesses as though its mass were quite different from that of a free electron. The angular bracket in Eq. (1.1) tells how the effect of the more complicated precession may be taken account of properly. It consists of a modification of the factor μ_B^2, so that $(1/m)^2$ is replaced by the proper combination of components of the reciprocal effective mass tensor. A large value of the angular bracket is associated with a small effective mass. Thus the argument of the previous paragraph is equivalent to the statement that a degenerate electron gas can be diamagnetic only if the average effective mass of the electrons at the Fermi surface is considerably smaller than the free electron mass.

The possibility that a really large diamagnetic susceptibility might arise as a consequence of the activities of a few electrons of very small effective mass has been contemplated by Bardeen,[6] who has envisaged a situation in which the susceptibility would be great enough to drive out the magnetic field. In Bardeen's problem the interaction of electrons with lattice vibrations splits the group of electron states at the Fermi surface into two groups. Bardeen treated the electrons in the energetically lower groups as electrons in an ordinary conduction band and used the formula Eq. (1.1) to calculate the magnetic susceptibility. He considered the possibility that the $\alpha^{\mu\nu}$ of Eq. (1.1) are of the order 10^4 corresponding to energy gaps in the electron spectrum of 10^{-3} ev. It seemed to the writer that the existence of such small energy gaps made the application of the ordinary susceptibility formula very doubtful, and it is in this doubt that the present work had its origin. However, the writer has not attempted a detailed application to Bardeen's theory, because he does not understand in what sense Bardeen can apply ordinary band theory in his particular problem. Instead, therefore, a detailed investigation was made of another theory which accounts for a large diamagnetism in terms of a small number of electrons with small effective masses, to wit, the theory given by Jones[7] to account for the susceptibility of Bi. In examining Jones' theory, we have been particularly interested in estimating the order of magnitude of other terms in the susceptibility than those leading to Eq. (1.1) because it is not obvious that such terms are really negligible when the effective masses of some electrons are very small. In fact, we will see immediately that these terms ought to be large when the effective masses are small.

An electron state which has a very small effective mass is located at a point in wave-number space where the curvature of the energy surface is large. Such regions of high curvature are small and are associated with energy gaps corresponding to planes of the reciprocal lattice. If the effective mass is very small, its magnitude is directly proportional to the size of the energy gap with which the high curvature is associated, and some or all of the α's which occur in Eq. (1.1) are large with magnitudes proportional to the reciprocal energy gap. Thus the large contributions to the Landau-Peierls susceptibility, Eq. (1.1), come from regions on the Fermi surface where electron states in two or more bands are almost degenerate. Now two states which are nearly degenerate can be appreciably admixed by what would otherwise be a very weak perturbation, and the second-order energy resulting from this admixing will have a magnitude proportional to the reciprocal energy gap. The magnetic perturbation is a weak perturbation which can cause such an admixture of states in different bands, and the second-order energy associated with this admixing makes a contribution to the susceptibility. Such contributions are not included in the susceptibility given by Eq. (1.1), and have usually been neglected in the theoretical studies of the observed susceptibilities of the diamagnetic metals. Wilson[2] has given a brief argument to the effect that one is justified in ignoring these contributions under most circumstances. It will be shown below that one should not neglect these contributions in the case of the diamagnetic metals, if it is true that most of the susceptibility is contributed by a few states. In fact, we will show that in such a case the Landau-Peierls formula, Eq. (1.1), does not give any reasonable approximation to the susceptibility.

In subsequent sections we derive an expression for the "steady" diamagnetic susceptibility, i.e., that part of the susceptibility which is independent of temperature and field strength, and use this expression to study the contribution to the susceptibility from the kinds of configurations of states which are believed to occur in the diamagnetic metals. The starting point in the derivation is the Hamiltonian of the generalized Wannier theory.[8] In Sec. 2 this Hamiltonian is diagonalized to order $\mathcal{H}^2$ by a unitary transformation. In Sec. 3 the diagonalized Hamiltonian is used in a statistical treatment to deduce a general expression for the steady susceptibility. The formula found here cannot well be compared with

[6] J. Bardeen, Phys. Rev. **80**, 567 (1950); **81**, 829 (1951).

[7] H. Jones, Proc. Roy. Soc. (London) **147**, 395 (1934).

[8] E. N. Adams, Phys. Rev. **85**, 41 (1952); **86**, 427 (1952).

Wilson's,[2] since some integrals which occur in his final formulas do not occur in the Wannier theory. In Sec. 4 a discussion is given of the effective mass tensor. In particular, we examine a special case in which the off-diagonal matrix elements can be expressed in terms of the effective mass. In Sec. 5 a calculation is made of the contribution to the susceptibility made by the simple configuration which occurs when the Fermi surface just overlaps the face of a Brillouin zone. This configuration is of interest because all terms in the susceptibility can be calculated almost exactly. The discussion of Sec. 5 shows that under certain circumstances a filled zone can make large contributions to the susceptibility. In Sec. 6 a discussion is given of the configuration which Jones[7] hypothesized to exist in Bi. While the complete calculation cannot be carried through for this configuration, it is shown that the terms which Jones neglected are of the same order as those which he kept. It is argued that in general the susceptibility contributions which are usually neglected are of the same order as the difference between the Landau susceptibility, Eq. (1.2), and the Landau-Peierls susceptibility, Eq. (1.1). It is concluded that Jones' theory of Bi is probably wrong; a modification of his model is suggested, but not investigated.

2. THE HAMILTONIAN

The Hamiltonian which will be used in calculating the susceptibility is that of the Wannier theory.[8] While the validity of this theory has been established only approximately, the applications to be made here do not depend on the approximations of the theory. The same results could be obtained from the Schrödinger theory by more tedious means. For an electron in a uniform constant magnetic field, the Schrödinger Hamiltonian will be written:

$$H=(p^2/2m)+V_L(x)+\tfrac{1}{2}\omega\delta_{\mu\nu\lambda}u^\mu\{x^\nu, p^\lambda\} +\tfrac{1}{8}m\omega^2\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}u^\mu u^\alpha\{x^\nu, \{x^\beta, \delta_{\gamma\lambda}\}\}. \quad (2.1)$$

Here $V_L(x)$ is the periodic potential of the unperturbed lattice, ω is the Larmor frequency $(e\mathcal{H}/2mc)$, u^μ is a unit vector in the direction of the magnetic field, $\delta_{\mu\nu\lambda}$ is the Levi-Civita δ symbol which is antisymmetric on all pairs of indices, and the curly brackets signify anticommutators. It is, of course, unnecessary to introduce the explicit symmetrization, but it will prove convenient subsequently because of the manifestly Hermitian character of the symmetrized terms. From the Schrödinger Hamiltonian, the Wannier Hamiltonian can be written down immediately.[8] First a slight change of notation from that used in reference 8 will be made. The notational change is simply to use Latin capitals instead of Greek letters to denote the operators for the electron momentum and coordinate. Thus $\mathbf{P}^\mu$ denotes the Wannier operator for the μ component of electron momentum, $P_n{}^\mu$ denotes that part of the operator which is diagonal in bands, P^μ that part which is nondiagonal in bands. Similarly, $\mathbf{X}^\mu$ denotes the operator for the μ-component of the electron coordinate, $x^\mu+X_n{}^\mu$ the part which is diagonal in bands, X^μ the part which is nondiagonal in bands. As previously, p^μ denotes the μ-component of the "crystal momentum" and x^μ the μ-component of what we might call the "crystal coordinate." $P_n{}^\mu$, P^μ, and X^μ are functions of the p^ν but not not of the x^ν. $X_n{}^\mu$ is independent[8] of p^ν and x^ν. The explicit forms of $\mathbf{X}^\mu$, $\mathbf{P}^\mu$ are given in reference 8. The Wannier Hamiltonian is now obtained from the Schrödinger Hamiltonian by the following rules of substitution:

$$p^2/2m+V_L(x)\rightarrow\mathbf{E}(p), \quad x^\mu\rightarrow\mathbf{X}^\mu, \quad p^\mu\rightarrow\mathbf{P}^\mu. \quad (2.2)$$

Thus the Wannier Hamiltonian corresponding to (2.1) is:

$$\mathbf{H}=\mathbf{E}(p)+\tfrac{1}{2}\omega\delta_{\mu\nu\lambda}u^\mu\{\mathbf{X}^\nu, \mathbf{P}^\lambda\} +\tfrac{1}{8}m\omega^2\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}\{\mathbf{X}^\nu, \{\mathbf{X}^\beta, \delta_{\gamma\lambda}\}\}. \quad (2.3)$$

The operator $\mathbf{E}(p)$ is diagonal in the bands and its diagonal element in the jth band $\langle jp|E(p)|pj\rangle$ is just the energy of an unperturbed state of the electron in the jth band and with wave number p. However, the other terms are not diagonal, so before doing any statistical treatments it is convenient to diagonalize the Hamiltonian (2.3) to the second order in ω. As a preliminary step we transform $\mathbf{H}$ by the operator:

$$\mathbf{u}_1=\sum_n \tau_{nn}\exp\{-iX_n{}^\mu p^\mu\}. \quad (2.4)$$

The operator $\mathbf{u}_1$ is *not* periodic with the reciprocal lattice periodicity, so the transformed operators are no longer periodic functions of p. It can be seen, however, that the diagonal parts of the transformed operators remain periodic in p. The principal results of the transformation (2.4) is to remove the constants $X_n{}^\mu$ from $\mathbf{X}^\mu$.

$$\mathbf{u}_1(x^\nu+X_n{}^\nu)\mathbf{u}_1{}^{-1}=x^\nu. \quad (2.5)$$

The transformation (2.4) considerably simplifies the calculation of the formula for the susceptibility. In addition, it shows that the $X_n{}^\mu$ need not appear in the final formula. Because we make this transformation, however, we cannot easily compare our final formula with Wilson's,[2] which still contains the $X_n{}^\mu$ in the form of the integral $\int(u_{pn}{}^*\partial u_{pn}/\partial p^\mu)d\tau_0$. The existence of the transformation (2.4) shows that Wilson's formula could be simplied by the use of sum rules and relations among matrix elements.

Even after the transformation (2.4) we can continue to write the Hamiltonian in the form (2.3), but with the understanding that the transformed $\mathbf{X}$, $\mathbf{P}$ now occur therein. The next step is to transform $\mathbf{H}$ by a unitary transformation which removes all parts of $\mathbf{H}$ which are nondiagonal in the bands and are of order ω. The transformation which does this will be called $\mathbf{u}_2$. Writing $\mathbf{u}_2=\exp\{i\mathbf{S}_2/\hbar\}$, we find that a possible choice of $\mathbf{S}_2$ is

$$\mathbf{S}_2=\tfrac{1}{2}m\omega\delta_{\alpha\beta\gamma}u^\alpha[\{X^\beta, x^\gamma\} +\sum_{jj'}{}'\tau_{jj'}1(\hbar/miE_{jj'})(2\{P_n{}^\gamma, X^\beta\}_{jj'} +\{P^\gamma, X^\beta\}_{jj'}]. \quad (2.6)$$

The transformed Hamiltonian takes the form:

$$\mathbf{u}_2\mathbf{H}\mathbf{u}_2^{-1}=[\mathbf{E}(p)+\tfrac{1}{2}\omega\delta_{\mu\nu\lambda}u^\mu\{x^\nu, P_n{}^\lambda\}+\tfrac{1}{8}m\omega^2\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}u^\mu u^\alpha\{x^\nu, \{x^\beta, \partial P_n{}^\lambda/\partial p^\nu\}\}]+\omega h_1+\tfrac{1}{8}\omega^2\delta_{\mu\nu\lambda}u^\mu\{x^\nu, h_{21}{}^\lambda\}+\tfrac{1}{8}\omega^2 h_2+O(\omega_0{}^3)+\cdots. \quad (2.7)$$

h_1, $h_{21}{}^\lambda$, and h_2 are Hermitian functions of p; nondiagonal terms of order ω_2 have been neglected.

The bracketed quantities are just the development[9] of $\mathbf{E}(p^\mu-m\omega\delta_{\mu\nu\lambda}x^\nu u^\lambda)$ to order ω^2. If this were the complete Hamiltonian, the electrons in a given band would move about in a complicated way which could be described by attributing to the electron a velocity-dependent mass tensor. Transitions between bands would not occur as a result of the magnetic forces, and wave packets would move on constant energy surfaces. The complete expression for the steady susceptibility would be the Landau-Peierls susceptibility, which is given by (1.1).

The chief concern of this paper will be with the remainder of the susceptibility which arises from the terms h_1, $h_{21}{}^\lambda$, and h_2 in the Hamiltonian. These terms are of a rather complicated form, even after the maximum simplification has been made. In order to express these terms, it is convenient to define four real dimensionless tensor operators $l^{\mu\nu}$, $d^{\mu\nu}$, $r^{\mu\nu}$, and $s^{\mu\nu}$.

$$\begin{aligned} l_{jj'}{}^{\mu\nu}&\equiv(1/\hbar)\textstyle\sum_{j''}'(X_{jj''}{}^\mu P_{j''j'}{}^\nu+P_{jj''}{}^\nu X_{j''j'}{}^\mu),\\ d_{jj'}{}^{\mu\nu}&\equiv(i/\hbar)\textstyle\sum_{j''}'(X_{jj''}{}^\mu P_{j''j'}{}^\nu-P_{jj''}{}^\nu X_{j''j'}{}^\mu),\\ r_{jj'}{}^{\mu\nu}&\equiv(1/\hbar)\{X^\mu, P_n{}^\nu\}_{jj'},\\ s_{jj'}{}^{\mu\nu}&\equiv(i/\hbar)[X^\mu, P_n{}^\nu]_{jj'}. \end{aligned} \quad (2.8)$$

The straight brackets denote commutators.

In terms of the operators (2.8) we may write

$$\begin{aligned} h_1&=(\hbar/2)\delta_{\mu\nu\lambda}u^\mu l_n{}^{\nu\lambda},\\ h_{21}{}^\lambda&=m\delta_{\alpha\beta\gamma}u^\alpha[\{\partial X^\lambda/\partial p^\beta, P^\gamma\}+\{X^\beta, s^{\gamma\lambda}+d^{\gamma\lambda}\}+3\{X^\beta, s^{\lambda\gamma}\}+2i[X^\lambda, l^{\beta\gamma}]]_n,\\ h_2&=\textstyle\sum\tau_{nn}\hbar^2\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}u^\mu u^\alpha[\{X^\nu, \{X^\beta, \delta_{\gamma\lambda}+3\alpha_n{}^{\gamma\lambda}\}\}_n\\ &\quad\times(m/\hbar^2)-(2m/\hbar)\{X^\lambda, \partial l^{\beta\gamma}/\partial p^\nu\}\\ &\quad+\textstyle\sum_{n'}'(1/E_{nn'})\{-d_{nn'}{}^{\beta\nu}d_{n'n}{}^{\lambda\gamma}\\ &\quad-d_{nn'}{}^{\beta\nu}(2s_{n'n}{}^{\lambda\gamma}+s_{n'n}{}^{\gamma\lambda})+l_{nn'}{}^{\beta\gamma}l_{n'n}{}^{\nu\lambda}\\ &\quad+3l_{nn'}{}^{\beta\gamma}r_{n'n}{}^{\nu\lambda}+\text{c.c.}\}\\ &\quad+\textstyle\sum_{n'}'(1/iE_{nn'})\{3r_{nn'}{}^{\beta\nu}d_{n'n}{}^{\gamma\lambda}-s_{nn'}{}^{\gamma\beta}l_{n'n}{}^{\nu\lambda}-\text{c.c.}\}\\ &\quad+\textstyle\sum_{n'}'(1/E_{nn'})\{-s_{nn'}{}^{\nu\beta}s_{n'n}{}^{\gamma\lambda}\\ &\quad+2r_{nn'}{}^{\nu\lambda}r_{n'n}{}^{\beta\gamma}+\text{c.c.}\}\\ &\quad+\textstyle\sum_{m'}'(1/iE_{nn'})\{r_{nn'}{}^{\beta\nu}(3s_{n'n}{}^{\gamma\lambda}+5s_{n'n}{}^{\lambda\gamma})-\text{c.c.}\}]. \end{aligned} \quad (2.9)$$

Here $E_{nn'}$ denotes $E_n-E_{n'}$ and the subscript $_n$ on an operator means the diagonal part of its matrix. Although the operator h_2 is complicated, it has several characteristics which can be described qualitatively. In the first place, h_2 vanishes for free electrons, since each of its terms involves off-diagonal matrix elements of the electron coordinate. In the second place, there are qualitatively different kinds of terms in h_2. For example, the first term involved $\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}u^\mu u^\alpha\{X^\nu, \{X^\beta, \delta_{\gamma\lambda}\}\}_n$ is always positive, while some of the other terms may have either sign according as the predominant energy denominators are positive or negative. This remark is made to emphasize the fact that h_2 is not simply a second-order energy correction, since the sign of a second-order energy must depend on whether the important intermediate states lie above or below the state corrected. What such behavior means will be discussed further when we obtain a formula for the susceptibility.

3. THE SUSCEPTIBILITY

The steady magnetic susceptibility may be defined in terms of the free energy F by

$$\chi=-\frac{\partial^2 F}{\partial\mathfrak{IC}^2}\bigg|_{\mathfrak{IC}=0}. \quad (3.1)$$

For a Fermi gas,

$$F=N\zeta-kT\textstyle\sum_i\ln[1+\exp\{-(E_i-\zeta)/kT\}]. \quad (3.2)$$

Since ζ depends on $\mathfrak{IC}$ quadratically, and since $\partial F/\partial\zeta=0$, ζ may be treated as a constant. In evaluating the susceptibility one makes use of the fact that the expression for χ may be converted into a trace which can then be evaluated in any convenient representation. It is easily seen because of the constancy of ζ and N, that the required form of the susceptibility is

$$\chi=-\frac{\partial^2}{\partial\mathfrak{IC}^2}\operatorname{Tr}[g(\mathbf{H})], \quad (3.3)$$

in which

$$g(\mathbf{H})\equiv-kT\ln[1+\exp\{-(\mathbf{H}-\zeta)/kT\}]. \quad (3.4)$$

Introducing the Larmor frequency here, (3.3) can be written as

$$\begin{aligned}\chi&=-\left(\frac{\mu_B}{\hbar}\right)^2\frac{\partial^2}{\partial\omega^2}\operatorname{Tr}[g(\mathbf{H})]\bigg|_{\mathfrak{IC}=0}\\&=-\left(\frac{\mu_B}{\hbar}\right)^2\operatorname{Tr}\left[\frac{\partial^2}{\partial\omega^2}g(\mathbf{H})\bigg|_{\mathfrak{IC}=0}\right].\end{aligned} \quad (3.5)$$

Since the value of a trace is invariant against changes of representation, we may use the transformed Hamiltonian (2.7) in $g(\mathbf{H})$. In addition, in getting (3.5) we have assumed a choice of representation for which the wave functions do not depend on ω. It is clearly most advantageous to take the trace using the representation in terms of eigenfunctions of the unperturbed Hamil-

[9] J. M. Luttinger, Phys. Rev. 84, 814 (1951).

tonian. At this point one can see why the Wannier formalism is especially suitable for the susceptibility calculation. Whereas with the Schrödinger Hamiltonian the unperturbed eigenfunctions are the complicated Bloch functions, with the Wannier formalism the unperturbed eigenfunctions are plane waves. Thus in setting up the Wannier formalism all of the complications of the problem have been put into the Hamiltonian, leaving the wave functions as simple as possible. Since we have already diagonalized $g(\mathbf{H})$ in the band indices, the trace will just involve a sum over bands of integrals on p of the diagonal parts of $g(\mathbf{H})$. Because the eigenfunctions are plane waves these integrals are very much like phase space integrals in classical statistical mechanics. The density of states in p space is just $2/(2\pi\hbar)^3$, provided that the integrals over p are restricted to the first Brillouin zone.

To facilitate calculating the trace, imagine that $g(\mathbf{H})$ has been developed as a formal power series in $\mathbf{H}$. It is not difficult to take the trace of the general term, and after doing so to sum the series. Thus we need first to calculate

$$T^{(N)}\equiv-\left(\frac{\mu_B}{\hbar}\right)^2\mathrm{Tr}\left[\frac{d^2\mathbf{H}^N}{d\omega^2}\right]_{\mathfrak{H}=0}$$

$$=-\left(\frac{\mu_B}{\hbar}\right)^2\frac{2}{(2\pi)^3}\sum_n\int_{B,Z}d^3k e^{-ik\cdot x}\left[\frac{d^2H_n{}^N}{d\omega^2}\right]_{\mathfrak{H}=0}e^{ik\cdot x}. \tag{3.6}$$

$T^{(N)}$ naturally breaks into two parts, one the result of first-order terms, the other the result of second-order terms. Call these $T_a{}^{(N)}$ and $T_b{}^{(N)}$, respectively. Since $\partial\mathbf{H}/\partial\omega$ commutes with $\mathbf{E}$, it is readily found that

$$T_a{}^{(N)}=-(\mu_B/\hbar)^2\,\mathrm{Tr}[(\tfrac{1}{2}\delta_{\mu\nu\gamma}u^\mu\{x^\nu,P_n{}^\lambda\}+h_1)^2\times N(N-1)\mathbf{E}_n{}^{N-2}]. \tag{3.7}$$

Using the fact that only terms which are real and even in x can survive, this is found to be

$$T_a{}^{(N)}=(\mu_B/\hbar)^2\,\mathrm{Tr}[\{\tfrac{1}{4}(\delta_{\mu\nu\lambda}u^\mu\{x^\nu,P_n{}^\lambda\})^2+h_1{}^2\}\times\partial^2\mathbf{E}^N/\partial\mathbf{E}^2]. \tag{3.8}$$

The calculation of $T^{b(N)}$ will now be described. It is found that the term in $(h_{21}{}^\lambda)^2$ can give nothing. The term in h_2 leads to a contribution which will be called $T_{b_2}{}^{(N)}$ and is

$$T_{b_2}{}^{(N)}=-(\mu_B/\hbar)^2\,\mathrm{Tr}[\tfrac{1}{4}h_2\partial\mathbf{E}^N/\partial\mathbf{E}]. \tag{3.9}$$

$T_{b_1}{}^{(N)}$, the remaining trace, can be written as

$$T_{b_1}{}^{(N)}=-(\mu_B/\hbar)^2\,\mathrm{Tr}[N\langle\mathbf{E}^{N-1},(m/4)\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}\times u^\mu u^\alpha\{x^\nu,\{x^\beta,\partial P_n{}^\gamma/\partial p^\lambda\}\rangle_S]. \tag{3.10}$$

The notation $\langle\mathbf{E}^{N-1},o\rangle_S$, means that the N factors are to be completely symmetrized as to order, so that o occurs once in each of the N ordinal positions. Now the operator in the face bracket (3.10) is to be commuted through all the $\mathbf{E}$'s until it stands on the left. After the factors have been ordered in this way the plane wave eigenfunctions can be dropped and p treated merely as an integration variable. The surviving part of $T_{b_1}{}^{(N)}$ is:

$$\begin{aligned}T_{b_1}{}^{(N)}=&-(\mu_B/\hbar)^2\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}u^\mu u^\alpha\\&\times\mathrm{Tr}[(m/4)\{x^\nu,\{x^\beta,\partial P_n{}^\gamma/\partial p^\lambda\}\}(d\mathbf{E}^N/d\mathbf{E})\\&-(\hbar^2/2)(\partial P_n{}^\beta/\partial p^\nu)(\partial P_n{}^\gamma/\partial p^\lambda)(d^2\mathbf{E}^N/d\mathbf{E}^2)\\&-(\hbar^2/3m)P_n{}^\nu P_n{}^\beta(\partial P_n{}^\gamma/\partial p^\lambda)(d^3\mathbf{E}^N/d\mathbf{E}^3)].\end{aligned} \tag{3.11}$$

The last term in $T_{b_1}{}^{(N)}$ may be integrated by parts using $(1/m)P_n{}^\nu d^3\mathbf{E}^N/\partial\mathbf{E}^3\equiv\partial(\partial^2\mathbf{E}^N/\partial\mathbf{E}^2)\partial p^\nu$. Finally, we get

$$\begin{aligned}T_{b_1}{}^{(N)}=&-(\mu_B/\hbar)^2\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}u^\mu u^\alpha\\&\times\mathrm{Tr}[(m/4)\{x^\nu,\{x^\beta,\partial P_n{}^\gamma/\partial p^\lambda\}\}d\mathbf{E}^N/d\mathbf{E}\\&-(\hbar^2/6)(\partial P_n{}^\beta/\partial p^\nu)(\partial P_n{}^\gamma/\partial p^\lambda)d^2\mathbf{E}^N/d\mathbf{E}^2.\end{aligned} \tag{3.12}$$

Now, before collecting terms, it can be shown that adding the first term of $T_{b_2}{}^{(N)}$ to the first term of $T_a{}^{(N)}$ will give something which eventually contributes no susceptibility. The proof consists in observing that these terms are just the ones which would have survived if we had assumed from the outset that $[p,x]=0$ and the Hamiltonian had the form $\mathbf{E}[p+(e\mathfrak{A}/c)]$. Thus these terms lead to the susceptibility which one gets on "nonquantum" mechanics for a Hamiltonian of the form $\mathbf{E}[p+(e\mathfrak{A}/c)]$. Van Leeuwen[10] has given an argument which shows that for such a Hamiltonian the susceptibility is precisely zero, since the energy depends only on the velocity. The argument is quite independent of statistics and can be used here to show that the two terms must cancel. There remains of $T^{(N)}$

$$\begin{aligned}T^{(N)}=&-(\mu_B/\hbar)^2\,\mathrm{Tr}[h_1{}^2(d^2\mathbf{E}^N/d\mathbf{E}^2)\\&+\tfrac{1}{4}h_2(d\mathbf{E}^N/d\mathbf{E})]+\mu_B{}^2\tfrac{1}{6}\delta_{\mu\nu\lambda}\delta_{\alpha\beta\gamma}u^\mu u^\alpha\\&\times\mathrm{Tr}[(\partial P_n{}^\beta/\partial p^\nu)(\partial P_n{}^\gamma/\partial p^\lambda)d^2\mathbf{E}^N/\partial\mathbf{E}^2].\end{aligned} \tag{3.13}$$

$T^{(N)}$ was a typical term in the power series development of the trace (3.5). It is a trivial matter to sum the series again. While doing so, we will split the susceptibility into several parts for convenience of subsequent discussion. In re-expressing the last term, we introduce the usual notation:

$$\partial P_n{}^\gamma/\partial p_\lambda\equiv\alpha_n{}^{\gamma\lambda}, \tag{3.14}$$

and set the $u^\alpha=\delta_{\alpha 3}$, thus choosing axes so that the Z direction is along the magnetic field. Then from the last term we get

$$\chi_1=+\tfrac{1}{6}\mu_B{}^2\delta_{\nu\lambda 3}\delta_{\beta\gamma 3}\,\mathrm{Tr}[\alpha_n{}^{\beta\nu}\alpha_n{}^{\gamma\lambda}d^2g(\mathbf{E})/d\mathbf{E}^2]. \tag{3.15}$$

Denoting as x and y two directions perpendicular to the field, we may write (3.15) as

$$\chi_1=+[2\mu_B{}^2/3(2\pi\hbar)^3]\textstyle\sum_n\displaystyle\times\int d^3p(\alpha_n{}^{xx}\alpha_n{}^{yy}-[\alpha_n{}^{xy}]^2)d^2g(E_n)/dE_n{}^2. \tag{3.16}$$

[10] For a discussion see J. H. Van Vleck, *Electric and Magnetic Susceptibilities* (Oxford University Press, London, 1933), p. 94.

Now if $f(E)$ denotes the Fermi function, we have $dg(E)/dE = f(E)$. The integral in (3.16) can be performed using the δ function character of df/dE. In order to do this, one introduces the energy as a variable of integration to replace the component of p normal to the energy surface. Then (3.16) becomes

$$\chi_1 = -\tfrac{2}{3}\mu_B^2/(2\pi\hbar)^3 \int_{SF} dS_P(\alpha^{xx}\alpha^{yy} - [\alpha^{xy}]^2)/|\nabla_p E|. \tag{3.17}$$

Equation (3.17) just gives the Landau-Peierls susceptibility, Eq. (1.1):

$$\chi_1 = -\tfrac{1}{3}\mu_B^2 n(\zeta)\langle \alpha^{xx}\alpha^{yy} - [\alpha^{xy}]^2\rangle. \tag{1.1}$$

Next we get the contributions to the susceptibility from the first term of expression (3.13), using the expression for h_1 given in Eq. (2.9). By arguments analogous to those used above it is easily found that

$$\chi_2 = \tfrac{1}{4}\mu_B^2 n(\zeta)\langle (l^{xy} - l^{yx})^2\rangle. \tag{3.18}$$

Here as in Eq. (1.1) the angular bracket denotes an average over the Fermi surface which need not lie in a single band.

Finally, from the second term in the expression (3.13) we get a contribution

$$\chi_3 = -(\mu_B^2/4\hbar^2)[2/(2\pi\hbar)^3] \int_{VF} d^3p h_2(p). \tag{3.19}$$

The integral in Eq. (3.19) goes over all occupied states, so over the occupied regions in every zone.

The complete susceptibility is given as the sum of χ_1, χ_2, and χ_3. However, in most cases the bulk of the contribution is made by χ_1, since χ_2 and χ_3 contain energy denominators which make their contributions negligible. We will subsequently investigate special cases in which such is not the case. At this point, however, we want merely to characterize these three susceptibilities. χ_1 is the susceptibility which one derives from quantum statistical mechanics using a Hamiltonian $\mathbf{E}[p + (e\mathfrak{A}/c)]$. χ_1 includes certain effects caused by interaction between states in different bands, *viz.*, those which can be described in terms of an effective mass which is velocity dependent. χ_1 may be of either sign since the angular bracket may be either positive or negative. χ_2 is always positive. Its form suggests that it results from something like the orbital moment of the electron about the center of whatever unit cell the electron is momentarily in. If we set $l \sim 1$, we see that χ_2 is of the same order of magnitude as χ_1. χ_3 consists of two kinds of terms, *viz.*, terms which are always diamagnetic, and terms which have a sign determined by the distribution of neighboring states in the energy spectrum. It is probable that χ_3 could only be evaluated for the simplest cases, since it contains terms corresponding to complicated "roundabout" transitions. The condition that χ_3 be important is that the energy gaps between some occupied bands be much smaller than the Fermi energy. We will see in a later section that small energy gaps between bands can result in a quite large value of χ_3.

4. THE EFFECTIVE MASS

We have earlier introduced the effective mass ratio tensor $\alpha^{\mu\nu}$ defined by

$$\alpha_n{}^{\mu\nu} \equiv m\partial^2 E_n(p)/\partial p^\mu \partial p^\nu. \tag{4.1}$$

In this section we will discuss the effective mass for a case in which it is possible to evaluate the susceptibility completely by expressing the off-diagonal matrix elements which occur in the expression (3.18) in terms of the effective mass ratio tensor. We will first write down the derivation of another expression for $\alpha_n{}^{\mu\nu}$. Observe that the operators $\mathbf{P}^\mu$ and $\mathbf{X}^\nu$ are operators for the momentum and coordinate of the electron, so they have the commutation relation

$$[\mathbf{P}^\mu, \mathbf{X}^\nu] = (\hbar/i)\delta_{\mu\nu}. \tag{4.2}$$

In the Wannier theory these operators have the following matrix forms:[8]

$$\begin{aligned} \mathbf{P}^\mu &= \textstyle\sum_n \tau_{nn} P_n{}^\mu + \sum'_{nn'} \tau_{nn'} P^\mu{}_{nn'}, \\ \mathbf{X}^\nu &= x^\nu + \textstyle\sum'_{nn'} \tau_{nn'}(\hbar/miE_{nn'}) P^\nu{}_{nn'}. \end{aligned} \tag{4.3}$$

The expressions (4.3) give the correct forms after the transformation $\mathbf{u}_1$ given by Eq. (2.4). Calculating the commutator and making use of the fact that $\alpha_n{}^{\mu\nu} = \partial P_n{}^\mu/\partial p^\nu$, one gets

$$\alpha_n{}^{\mu\nu} = \delta_{\mu\nu} - \textstyle\sum'_{n'}(1/mE_{n'n})(P_{nn'}{}^\mu P_{n'n}{}^\nu + P_{nn'}{}^\nu P_{n'n}{}^\mu). \tag{4.4}$$

We want to discuss the size and sign of $\alpha_n{}^{\mu\nu}$ for several cases of interest. In the first place, the $\alpha^{\mu\nu}$ should ordinarily be smaller than unity for the valence band of a metal. This is so because the few intermediate states which are of lower energy and which could, therefore, lead to larger values of $\alpha^{\mu\nu}$ actually lie so much lower in energy that their contributions are small. On the other hand, there are many excited states in higher bands, all of whose contributions have such a sign as to lower the value of $\alpha^{\mu\nu}$. There can quite easily be such states which are not much higher in energy, so it is reasonable to expect that for some values of p in the band the reduction of $\alpha^{\mu\nu}$ should become appreciable. Thus for a valence electron band $\alpha^{\mu\nu}$ is probably of the order of unity or somewhat smaller, being perhaps considerably smaller for a fraction of the states. If the excited band should lie very close for some values of p, the values of $\alpha^{\mu\nu}$ in that part of the band could be very much smaller than unity, but of absolute value very much greater than unity. In the case of a conduction band which lies not too far above a valence band, we ex-

pect large positive values of the $\alpha^{\mu\nu}$ for states in the neighborhood where the bands are very close together. Where a number of bands are close in energy, we simply expect $|\alpha^{\mu\nu}|\gg 1$, but we do not known the sign except for the highest and lowest bands.

In order to calculate the $\alpha_n{}^{\mu\nu}$ from Eq. (4.4), one would need to know the off-diagonal matrix elements of the electron momentum. Such knowledge would in turn involve a more detailed knowledge of the wave functions than one can ordinarily hope to attain. It is, in fact, more likely that one could get information about $\alpha^{\mu\nu}$ directly from the energy spectrum. In this section we will actually consider the $\alpha^{\mu\nu}$ as known, and express the off-diagonal elements in terms of the $\alpha^{\mu\nu}$. Of course, we can do so only in an especially simple situation. As will be seen, however, in this situation we can still learn something about the $\alpha^{\mu\nu}$ from Eq. (4.4). The simple case to be considered is one which arises when, in a certain region of p space, two bands and only two lie very close together. What we are interested in here is a configuration for which the energy gap is extremely small so that values $|\alpha|\gg 1$ occur. In such a case a very good approximation may be had by entirely neglecting the influence of other bands. In Fig. 1 we describe the configuration using the unreduced zone description of the bands. The near degenerate states are distributed along the faces of a plane in the momentum space, the plane being represented in cross section by a solid line. Another solid line represents the cross section of an energy surface which overlaps into the second band, while a dotted line indicates how the cross section might look in the absence of the energy discontinuity across the zone face. The large curvature of the energy surface near the zone face is proportional to one of the $\alpha^{\mu\nu}$. Because the curvature is high, the energy separation between bands grows rapidly with p in the direction normal to the face. Taking an origin of momentum coordinates at the point 0 and designating the normal direction as then x direction, one can say that for p_x about equal to p_m, the energy gap has already doubled.

$$p_m=(m\Delta E_0/|\alpha^{xx}|)^{\frac{1}{2}}. \tag{4.5}$$

The significance of p_m will be discussed again later. The dashed lines of Fig. 1 are intended to be at a distance p_m from the zone face and thus to show what states in the momentum space have large α values. The calculation to be made next applies only to states lying between the planes represented by the dashed lines.

ΔE will denote E_{21}, hence will be a positive number. Let us calculate the commutator $[\mathbf{X}^\mu, \mathbf{X}^\nu]$ using the form (4.3) and noting that because of the energy denominators only one intermediate state is important. All matrix elements of this commutator vanish identically because the $\mathbf{X}^\mu$ are electron coordinate operators. The vanishing of the diagonal element of the commutator gives

$$P_{12}{}^\mu P_{21}{}^\nu - P_{12}{}^\nu P_{21}{}^\mu = 0, \tag{4.6}$$

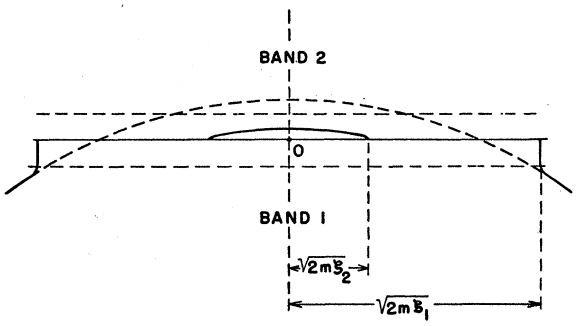

FIG. 1. Profile of an energy surface which just overlaps the face of a zone. $\zeta_1=\zeta_2+\Delta E_0$, where ΔE_0 is the energy discontinuity at 0. $pm=(m\Delta E_0/\alpha)^{\frac{1}{2}}$ gives the separation of the dashed lines from the plane.

Using Eq. (4.6) and the one intermediate state approximation, one gets from Eq. (4.4)

$$\begin{aligned}\alpha_1{}^{\mu\nu}&=\delta_{\mu\nu}-2(P_{12}{}^\mu P_{21}{}^\nu/m\Delta E),\\ \alpha_2{}^{\mu\nu}&=\delta_{\mu\nu}+2(P_{12}{}^\mu P_{21}{}^\nu/m\Delta E),\end{aligned} \tag{4.7}$$

or

$$2P_{12}{}^\mu P_{21}{}^\nu=(\delta_{\mu\nu}-\alpha_1{}^{\mu\nu})m\Delta E. \tag{4.8}$$

It will be seen later that Eqs. (4.7) and (4.8) and several equations deducible therefrom are all that we need to calculate the complete susceptibility when the Fermi surface overlaps a zone face.

Some equations which follow from (4.7) will be useful later. Adding the two Eqs. (4.7), one gets

$$\alpha_1{}^{\mu\nu}+\alpha_2{}^{\mu\nu}=2\delta_{\mu\nu}. \tag{4.9}$$

Equation (4.9) can be integrated near $p=0$ to give

$$P_1{}^\mu+P_2{}^\mu=2p^\mu. \tag{4.10}$$

Also one can write from (4.9),

$$(\alpha_2{}^{\mu\nu}-\alpha_1{}^{\mu\nu})=2(\delta_{\mu\nu}-\alpha_1{}^{\mu\nu}),$$

which integrates to

$$(P_2{}^\mu-P_1{}^\mu)=2(p^\mu-\alpha_1{}^{\mu\nu}p^\nu). \tag{4.11}$$

Finally, we will prove that if $\alpha^{\mu\nu}$ is diagonalized only one component can be large. In a weak potential theory with the potential represented by a single cosine this is obvious. However, the point here is to show that there is no conceivable set of circumstances under which we get more than one large α with only one excited state, no matter how many Fourier components are important. The proof follows immediately if one squares the two sides of Eq. (4.8) for $\mu\neq\nu$ and then uses Eq. (4.8) to rewrite the left-hand side. One gets, for example,

$$(\alpha_1{}^{\mu\nu})^2=(1-\alpha^{\mu\mu})(1-\alpha^{\nu\nu}),\quad(\mu\neq\nu). \tag{4.12}$$

Equations (4.12) must hold when axes are chosen so

that $\alpha^{\mu\nu}$ is diagonal, so one gets

$$\begin{aligned}(1-{}^{xx}\alpha_1)(1-\alpha_1{}^{yy})&=0,\\(1-\alpha_1{}^{xx})(1-\alpha_1{}^{zz})&=0,\\(1-\alpha_1{}^{yy})(1-\alpha_1{}^{zz})&=0.\end{aligned}\tag{4.13}$$

Choosing $1-\alpha_1{}^{xx}\neq 0$, it follows that $\alpha_1{}^{yy}=\alpha_1{}^{zz}=1$.

5. THE COMPLETE SUSCEPTIBILITY FOR A SPECIAL CASE

If the Fermi surface slightly overlaps the face of a zone boundary across which there is a small gap, the electrons in the neighborhood of the gap will, under certain circumstances, make an unusually large contribution to the susceptibility. Equation (1.1) indicates that this will be the case when the field is oriented perpendicular to the direction in which $\alpha^{\mu\mu}$ is large. Clearly, $\alpha^{\mu\mu}$ is largest when μ refers to the direction normal to the zone face. This direction we denote as the x direction. The solid curve of Fig. 1 describes a cross section of the Fermi surface perpendicular to the magnetic field. Figure 2 shows a yz cross section, the circles being the intersections of the Fermi surface with the zone face.

We wish to calculate the contribution of this overlap configuration to the magnetic susceptibility for the case that the magnetic field direction lies in the zone face. What we will actually calculate is the excess above the Landau susceptibility as given by Eq. (1.2). In order to get the excess, we merely replace α^{xx} by $(\alpha^{xx}-1)$ in the susceptibility χ_1. Equation (4.8) shows that $(\alpha^{xx}-1)$ is a particularly simple expression which has a sharp maximum (as a function of p_x) at the zone face. We make one more elaboration in that together with the contribution of the overlap configuration discussed above we will take the identical contribution which comes from overlapping the opposite zone face. Then we can say that for small overlap the occupied states in the second band fill the ellipsoidal region

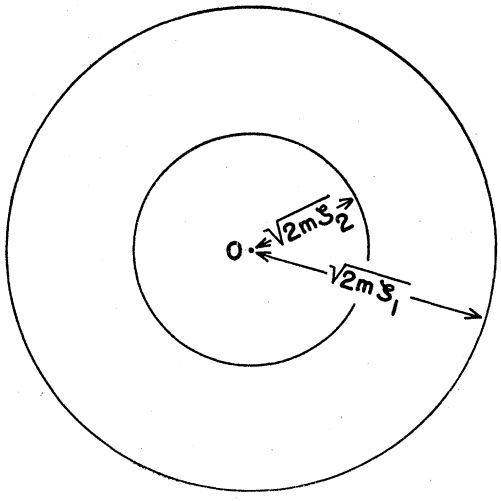

FIG. 2. Intersection of Fermi surface with plane discontinuity.

inside of an energy surface corresponding to an overlap energy ζ_2. The ellipsoid, because of the large value of $\alpha_2{}^{xx}$, is greatly flattened along the x axis.

In order to evaluate χ_2 and χ_3, we must evaluate certain sums over intermediate states. As we saw in the previous section, this becomes possible when there is but a single important intermediate state. Introducing the notation,

$$(\alpha_1{}^{xx}-1)\equiv-\alpha,\tag{5.1}$$

we can see from Eq. (4.9) that

$$(\alpha_2{}^{xx}-1)=\alpha.\tag{5.2}$$

In terms of α we can rewrite Eq. (4.8) as

$$2P_{12}{}^{\mu}P_{21}{}^{\nu}=m\alpha\Delta E\delta_{\mu x}\delta_{\nu x}.\tag{5.3}$$

We can now use Eq. (5.3) to rewrite the four expressions (2.8) so they involve only diagonal matrix elements. In our approximation,

$$\begin{aligned}l_{jj'}{}^{\mu\nu}&=0,\quad d_{jj'}{}^{\mu\nu}=0,\\r_{jj'}{}^{\mu\nu}&=(1/miE_{jj'})P_{jj'}{}^{\mu}(P_{j'}{}^{\nu}+P_j{}^{\nu}),\\s_{jj'}{}^{\mu\nu}&=(1/mE_{jj'})P_{jj'}{}^{\mu}(P_{j'}{}^{\nu}-P_j{}^{\nu}).\end{aligned}\tag{5.4}$$

Since $l_{jj'}{}^{\mu\nu}$ vanishes, the susceptibility χ_2 vanishes for our problem. We must simplify χ_3 which is given by (3.18). Inserting Eq. (5.4) into h_2 of Eq. (2.9), one gets for χ_3,

$$\begin{aligned}\chi_3=&(\mu_B{}^2/4)(2/(2\pi\hbar)^3)\delta_{\nu\lambda z}\delta_{\beta\gamma z}\delta_{\lambda x}\delta_{\gamma x}\\&\times\Big[\int_{(1)}\alpha d^3p\{-(1/\Delta E)(2\delta_{\beta\nu}+3\alpha_1{}^{\beta\nu}+3\alpha_2{}^{\beta\nu})\\&+(1/m\Delta E^2)[(P_2{}^{\beta}-P_1{}^{\beta})(P_2{}^{\nu}-P_1{}^{\nu})\\&+2(P_2{}^{\beta}+P_1{}^{\beta})(P_2{}^{\nu}+P_1{}^{\nu})\\&+5(P_2{}^{\beta}+P_1{}^{\beta})(P_2{}^{\nu}-P_1{}^{\nu})]\}\\&+\int_{(2)}\alpha d^3p\{-(1/\Delta E)(2\delta_{\beta\nu}+3\alpha_1{}^{\beta\nu}+3\alpha_2{}^{\beta\nu})\\&+(1/m\Delta E^2)[-(P_2{}^{\beta}-P_1{}^{\beta})(P_2{}^{\nu}-P_1{}^{\nu})\\&-2(P_2{}^{\beta}+P_1{}^{\beta})(P_2{}^{\nu}+P_1{}^{\nu})\\&+5(P_2{}^{\beta}+P_1{}^{\beta})(P_2{}^{\nu}-P_1{}^{\nu})]\}\Big].\end{aligned}\tag{5.5}$$

The integrals $\int_{(1)}d^3p$, $\int_{(2)}d^3p$ are over the occupied states in the first and second bands, respectively. Already in the form (5.5), it can be seen that χ_3 has two different kinds of terms. The first and last integrals in each bracket have the same sign in either band while the others are of opposite signs in the two bands. The terms of the second kind will contribute only if there is a region of p space for which the states are occupied in one band but not the other.

We can simplify Eq. (5.5) still further. Using Eqs. (4.9), (4.10), and (5.3), we reduce (5.5) to

$$\chi_3=(2\mu_B{}^2/(2\pi\hbar)^3)\Big[\int_{(1)} d^3p\alpha\{-(2/\Delta E)+(2/m\Delta E^2)p_y{}^2\}$$

$$+\int_{(2)} d^3p\alpha\{-(2/\Delta E)-(2/m\Delta E^2)p_y{}^2\}\Big]. \quad (5.6)$$

The total susceptibility is the sum of χ_3 and the surface integrals χ_1:

$$\chi_1=+(2\mu_B{}^2/3(2\pi\hbar)_3)\int_{(1)} dS_p\alpha/|\nabla_p E|$$

$$-(2\mu_B{}^2/3(2\pi\hbar)^3\int_{(2)} dS_p\alpha/|\nabla_p E|. \quad (5.7)$$

Note that in both Eqs. (5.6) and (5.7) α is a function of p which vanishes strongly when p_x is large. The strong energy dependence of α comes about partly because α contains an energy denominator and partly because the nondiagonal matrix elements fall off away from the zone face. It is difficult to estimate the energy dependence of the matrix elements without some kind of model of the potential so we will merely note that there is such energy dependence; later in the section we describe the results obtained on a specific model. In any case we can say that when p_x gets large enough α drops sharply to zero. We can estimate the value of p_x for which the drop-off occurs by noting that when the kinetic energy $\alpha p_x{}^2/2m$ equals half the gap energy $(\Delta E/2)$ it is no longer energetically profitable for the electrons to try to be in standing wave type states, so α will fall off. The critical value of p_x we have denoted as p_m; it was given in Eq. (4.5). We will do the integrals χ_1 and χ_3 on the assumption that α has the constant value α_0 for $p_x<p_m$ and the value 0 for $p_x>p_m$. This assumption cannot introduce any large errors. However, because of the different energy dependences of the different terms, we may get relative values of various terms which are wrong by as much as a factor 2. It will become clear that such errors will not affect the qualitative description of the dependence of the susceptibility on overlap.

We calculate the excess susceptibility due to overlap for two cases (a) and (b). In case (a) the overlap is small so the occupied states in the upper band occupy a flattened ellipsoid. A central cross section of the Fermi surface looks like the solid curve of Fig. 1. In case (b) the overlap is so large that the Fermi surface in the second band is more like a spherical cap sitting on a narrow collar cut from an ellipsoid. A central cross section looks like the solid curve of Fig. 3.

(a) The integral $\chi_1{}^{(2)}$ is easily found to be

$$\chi_1{}^{(2)}=-\tfrac{1}{3}\mu_B{}^2N(\zeta_2)\alpha_0. \quad (5.8)$$

Here

$$N(\zeta_2)\equiv 8\pi m(2m\zeta_2)^{\frac{1}{2}}(2\pi\hbar)^{-3}\alpha_0{}^{-\frac{1}{2}}, \quad (5.9)$$

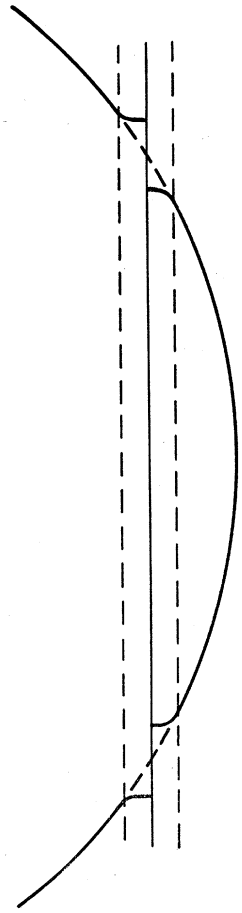

FIG. 3. Cross section of energy surface for large overlap. The dotted line shows the corresponding energy surface in the absence of interaction.

and the superscript (2) denotes the contributions from the second band. In order to evaluate $\chi_1{}^{(1)}$, we notice that the bulk of the integral comes from the "collar" on the surface (1). This collar is of radius $(2m\zeta_1)^{\frac{1}{2}}$ and width $2p_m$ when we take a pair of zone faces together. Carrying out the surface integral, one gets

$$\chi_1{}^{(1)}=\tfrac{1}{3}\mu_B{}^2\alpha_0N(\Delta E_0/2). \quad (5.10)$$

Thus

$$\chi_1=\tfrac{1}{3}\mu_B{}^2\alpha_0[N(\Delta E_0/2)-N(\zeta_2)]. \quad (5.11)$$

Equation (5.11) is valid for $\zeta_2\lesssim\Delta E_0/2$. χ_3 is easily calculated to be

$$\chi_3{}^{(1)}=\mu_B{}^2N(\Delta E_0/2)\alpha_0[(\zeta_1/\Delta E_0)^2-(2\zeta_1/\Delta E_0)], \quad (5.12)$$

$$\chi_3{}^{(2)}=-\tfrac{1}{3}\mu_B{}^2N(\zeta_2)\alpha_0[(4\zeta_2/\Delta E_0)$$
$$+(8/5)(\zeta_2/\Delta E_0)^2]. \quad (5.13)$$

In case (b) the overlap is very large, so the upper band integrals must be recalculated. $\chi_1{}^{(2)}$ now comes from a "collar" integral also, and is found to be

$$\chi_1{}^{(2)}=-\tfrac{1}{3}\mu_B{}^2\alpha_0N(\Delta E_0/2). \quad (5.14)$$

$\chi_3{}^{(2)}$ is an integral over all occupied states inside the

upper collar region.

$$\chi_3^{(2)}=-\mu_B^2N(\Delta E_0/2)\alpha_0[(\zeta 2/\Delta E_0)^2+(2\zeta_2/\Delta E_0)]. \quad (5.15)$$

The lower band integrals are still given by Eqs. (5.10) and (5.12). In case (b) we have

$$\chi_1=\chi_1^{(1)}+\chi_1^{(2)}\sim 0. \quad (5.16)$$

We can now survey the trend of the various contributions to χ, beginning with small overlap and proceeding to large overlap. For small overlap the Landau-Peierls susceptibility is paramagnetic as shown by Eq. (5.11). As the number of overlapping electrons increases, the susceptibility χ_1 decreases monotonically until for large overlap χ_1 is practically zero as shown by Eq. (5.16). Thus, surprisingly, the excess Landau-Peierls susceptibility is always paramagnetic. The susceptibility χ_3 is negative as overlap begins. The contribution from the lower band becomes less negative as the overlap proceeds since the term $(\zeta_1/\Delta E_0)^2$ in Eq. (5.12) is always greater than 1. At the same time the diamagnetic contribution from the upper band states is steadily growing. As the overlap becomes large, the susceptibility χ_3 finally takes on the simpler form

$$\chi_3=-\mu_B^2N(\Delta E/2)\alpha_0[1+(2\zeta_2/\Delta E_0)]. \quad (5.17)$$

This susceptibility is diamagnetic and grows linearly with the maximum overlap energy. In Fig. 4 the general behavior is shown. Figure 4 is not plotted directly from the formulas above, since the method of integration used there permits relative errors between terms of a factor 2. As noted above such errors arise because of the energy dependence of the matrix elements. Therefore, we have again carried through the susceptibility calculation using the explicit form of the matrix elements obtained for the well-known case of a weak cosine potential. Thus in Fig. 4 the relative magnitudes of the various terms are correct. The formulas for this special case show that the integrations above are not bad. We label these formulas to correspond to those above.

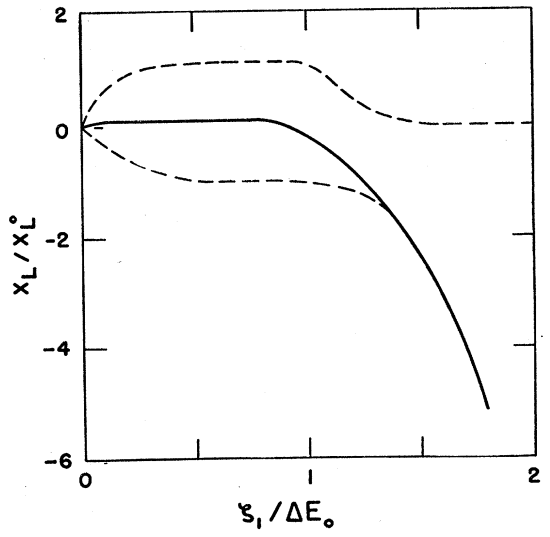

FIG. 4. Susceptibility as a function of ζ_1. Overlap begins for $\zeta_1=\Delta E$. The upper dashed curve represents χ_1, the lower dashed curve χ_3. The solid curve is the total susceptibility. All susceptibilities are measured in units of $\chi_L^0=\frac{1}{3}\mu_B^2N(\Delta E/2)\alpha_0$.

(a) $$\chi_1=\tfrac{1}{3}\mu_B^2\alpha_0[N(\Delta E_0/2)-N(\zeta_2)], \quad (5.11)\text{cp}$$

$$\chi_3^{(1)}=\mu_B^2N(\Delta E_0/2)\alpha_0[(\sqrt{2}/3)(\zeta_1/\Delta E_0)^2-(3\pi\sqrt{2}/16)(\zeta_1/\Delta E), \quad (5.12)\text{cp}$$

$$\chi_3^{(2)}=-\tfrac{1}{3}\mu_B^2N(\zeta_2)\alpha_0[(3\zeta_2/\Delta E_0)+(8/5)(\zeta_2/\Delta E_0)^2]; \quad (5.13)\text{cp}$$

(b) $$\chi_1^{(2)}=-\tfrac{1}{3}\mu_B^2\alpha_0N(\Delta E_0/2), \quad (5.14)\text{cp}$$

$$\chi_3^{(2)}=-\mu_B^2N(\Delta E_0/2)\alpha_0[(\sqrt{2}/3)(\zeta_2/\Delta E_0)^2+(3\pi\sqrt{2}/16)(\zeta_2/\Delta E_0)]. \quad (5.15)\text{cp}$$

The asymptotic behavior for large overlap is surprising, because it indicates that under certain circumstances one might get a very large contribution to the susceptibility from states which lie far below the Fermi surface. We will now see how large the effects are when an entire zone face is overlapped. We will assume that α satisfies

$$\alpha\sim 4E_K/\Delta E. \quad (5.18)$$

Relation (5.18) is true for a weak cosine potential[3] if E_K is the energy corresponding to the center of the face. Then $\Delta E/2\sim 2E_K/\alpha$, and by (5.8)

$$\alpha_0N(\Delta E_0/2)=8\pi m(2mE_K)^{\frac{1}{2}}(2\pi\hbar)^{-3}. \quad (5.19)$$

The right-hand side of (5.19) is just the density of energy levels for a degenerate free electron gas with states occupied for all energies below E_K, so the susceptibility described by Eq. (5.17) is

$$\chi_3\sim\chi_L^{(0)}(6\zeta_2/\Delta E), \quad (5.20)$$

in which $\chi_L^{(0)}$ is the Landau susceptibility for a free electron gas with Fermi energy E_K. If the complete face of a lower zone were thus overlapped, we would attain a value $\zeta\sim E_K/2$ at least, and find

$$\chi_3\sim\chi_L(3E_K/\Delta E)\sim\alpha\chi_L{}^0. \quad (5.21)$$

Equation (5.21) describes the contribution to the total susceptibility which would be made by a pair of zone faces which are completely overlapped by the Fermi surface, no matter how far below that surface. For comparison it can be noted that the maximum value of χ_1 as given by Eq. (5.11) is about $\chi_L^{(0)}$, using the same approximation (5.18) for α.

The calculation as given in this paper does not give much insight into the physical process which leads to the large contribution to χ_3. However, an examination of the detailed deduction of the results in Sec. 3 shows that $\frac{1}{4}$ of the contribution comes from the $\mathfrak{A}^2$ term in the original Hamiltonian and $\frac{3}{4}$ from the $P\cdot\mathfrak{A}$ terms

taken in second order. While it is not easy to say what sign the latter contribution should have, it is clear that the result found here is qualitatively correct for the $\mathcal{C}^2$ term which is positive for all states and hence must lead to a diamagnetic contribution to the susceptibility from all bands. That the order of magnitude of the contribution arising from a filled band is correct can be verified by noting that in the face of a zone there are of the order of $N\alpha^{-\frac{1}{2}}$ states whose magnetic energies are each of the order

$$\epsilon \sim \tfrac{1}{2}m\omega^2 X_{nn'}{}^2 \sim \tfrac{1}{2}\hbar^2\omega^2\alpha/\Delta E, \qquad (5.22)$$

where N is the number of electrons per unit volume on the metal. Thus the magnetic energy of this group of states is $U \sim \mu_B{}^2 \mathcal{H}^2 \alpha N(E)$ which gives $\chi \sim \alpha\chi_L$ as found above. This large susceptibility contribution from a filled band corresponds to a physical situation in which there is a large ionic susceptibility because the ion has a low-lying excited state.

If in a metal there were a large susceptibility arising from overlap of weak gaps, a large magnetic anisotropy effect could easily result. For, as a result of a 90° rotation of the magnetic field direction, the susceptibility contribution of that face could be made to go from an original large value to zero. It is thus conceivable that a number of weak faces oriented in a simple pattern could rise to a rather complicated pattern of magnetic anisotropy in a single crystal specimen. We think it possible that such a situation exists in Bi, which we will discuss in detail in the next section.

6. APPLICATION TO BISMUTH

In the last section it was shown for a certain simple configuration of the Fermi surface in which the effective electron masses are very small that the Landau-Peierls formula, Eq. (1.1), does not give a good approximation to the total susceptibility. In this section we will argue that for any configuration of electron states with small effective masses the susceptibility contributions χ_2 and χ_3 are as important as the Landau-Peierls susceptibility. Naturally, we cannot calculate these susceptibilities for any very complicated configuration of states, so what we will do is to discuss an example of a more complicated configuration and show that the order of magnitude of the terms in χ_3 is the same as that of the Landau-Peierls susceptibility. The example which we will discuss comprises two configurations which are of general interest, *viz.*, the top of a band which is almost full and the bottom of a band which is almost empty.

The complicated situation we chose to discuss is the Fermi surface of Bi. For several reasons Bi seemed especially suitable for our purpose. A detailed analysis of the Bi susceptibility had already been made by Jones.[7] He had found that some of the electrical and magnetic properties could be understood in a simple manner as resulting from the anomalous behavior of very small groups of electrons with very small effective masses. Jones had developed a specific set of hypotheses giving a correlation of the various anomalous electric and magnetic properties with specific sites on the Fermi surface and had thereby reached quantitative conclusions concerning the values of the $\alpha^{\mu\nu}$ which would be required to explain the observations. Finally, Jones had apparently been able to reach reasonable agreement with experiment taking account of only the Landau-Peierls part of the susceptibility. Thus Bi was very satisfactory for our purpose since practically important configurations seemed to occur in Bi, small electron masses play an important role, and the effects which we are concerned with had not previously been studied for Bi.

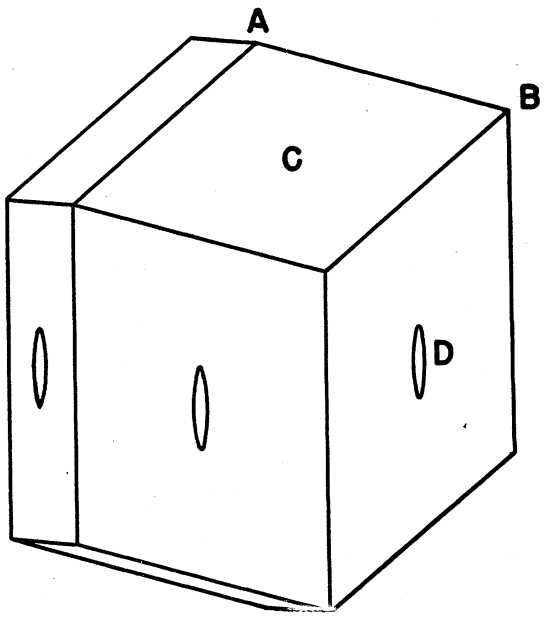

FIG. 5. Brillouin zone of bismuth containing five electrons: $E_A = 12.5$ ev; $E_B = 15$ ev; $E_0 = 9.2$ ev; $E_D = 7.5$; $E_F = 10$ ev.

Jones believed that the smallness of the electrical conductivity of Bi was evidence that most of the electrons in Bi occupy a nearly full Brillouin zone, so that the free Fermi surface is very small. The Brillouin zone which Jones suggested is shown in Fig. 5. If full, it would contain exactly the five valence electrons per atom which are present in pure Bi. Jones supposed that the zone is not quite full, but that there is a small number of holes in the inner zone at A and a small number of electrons overlapping into the outer zone at D. The number of overlapping electrons was estimated to be of the order of 10^{-4} per atom.[11] The overlapping electrons were supposed to occupy greatly elongated ellipsoids as indicated in Fig. 5. The large diamagnetic susceptibility which is observed when a magnetic field is applied along the principal axis of a single crystal of Bi was attributed to the diamagnetic precession of the

[11] See reference 3 for this estimate which is different from that of the original paper reference 6. Jones' theory is discussed in some detail in this reference.

overlapping electrons at the sites D. Jones decided that these electrons would have to have effective mass ratios $\alpha^{xx}\sim\alpha^{yy}\sim 200$ for motion perpendicular to the principal axis and $\alpha^{zz}\sim 1$ for motion along the principal axis. The large diamagnetic susceptibility which is observed when a magnetic field is applied perpendicular to the principal axis was attributed to the holes at A. Jones did not attempt a quantitative discussion of the holes or the susceptibility perpendicular to the principal axis.

A number of experiments give information which is in at least qualitative agreement with Jones' model.[10] The susceptibility of crystalline Bi is diamagnetic and about 25 times the value of χ_L which one calculates assuming that the 5 valence electrons per atom constitute a free electron gas having the electron density of Bi. It is difficult to account for such a large diamagnetism except on the assumption that some of the valence electrons have small effective masses. Furthermore, this large diamagnetism disappears almost completely when the Bi is melted so it is hard to avoid the conclusion that the diamagnetism is intimately associated with the lattice structure and it is reasonable to suppose it to be associated with the Brillouin zone structure. The ratio of susceptibilities parallel and perpendicular to the principal axis can be changed by a factor 2 by alloying with the pure Bi a small fraction of an atomic percent of Pb, which has one less valence electron per atom than Bi. This observation is compatible with the hypothesis that the principal axis susceptibility arises from the motions of a very small number of overlapping electrons, and that this number is substantially reduced percentage-wise by the substitution of the small number of Pb atoms. Finally, the field dependence of the Bi susceptibility is very spectacular (the de Haas-van Alphen effect);[12] the susceptibility oscillates in high fields and the oscillatory behavior persists to rather high temperatures. The reason for the oscillation was first recognized by Peierls and has since been studied in great detail.[13,14] Existing theories of the effect are not really applicable to metals because in the theories the electrons are treated as free. However, these theories show that a free electron gas could not have such an oscillation of the susceptibility at the field strengths used unless some electrons had extremely small masses. The totality of the evidence strongly suggests, therefore, that the anomalous magnetic properties of Bi originate from some configuration or configurations of conduction electrons which have small effective masses, probably as a consequence of the structure of the Brillouin zones.

In the following paragraphs we will discuss the two configurations which occur in Jones model of the Fermi surface from the point of view of the "neglected" susceptibilities. We will begin by discussing the hypothetical configuration of the electrons overlapping at the sites D. Figure 5 shows that for these electrons states all components of the diagonal reciprocal effective mass tensor are necessarily positive, and Jones found that agreement with experiment is only possible if precisely two components are large. Now it follows from Eq. (4.4) that if $\alpha^{\mu\mu}$ is large and positive for some state, then there exists a nearby state for which $\alpha^{\mu\mu}$ is large and negative. For suppose there are M states which are close together and interact strongly so that the sum is Eq. (4.4) gets its chief contribution from these states. Summing Eq. (4.4) over those states we find

$$\sum_{n=1}^{M}\alpha_n{}^{\mu\nu}=M\delta_{\mu\nu}. \tag{6.1}$$

M is always a small number, so it is clear that the positive and negative contributions must approximately cancel. The form of the right-hand side of (4.4) indicates that the lower-lying states will have negative $\alpha^{\mu\mu}$ while the higher-lying states will have positive $\alpha^{\mu\mu}$. Finally, the $\alpha^{\mu\mu}$ of the lowest state must be either smaller than one or negative. It would seem likely then, that on Jones' model the overlapping states at D are not in the lowest state of the set of near degenerate states responsible for the great size of the $\alpha^{\mu\nu}$. In that case, it must be that the lower band states on the other side of the zone face also have large (but negative) $\alpha^{\mu\nu}$, hence that there is only a small energy gap across the zone face. We can estimate the gap energy by noting that the free electron kinetic energy at the D site is about 7.5 volts and the α corresponding to the direction normal to the zone face is 200 for the upper band states. The corresponding α in the lower band must be also of the order 200, so we could conclude by Eq. (5.18) that the energy discontinuity across the zone face was of the order of 0.15 ev. An examination of the kinetic energies at various points in the zone of Fig. 5 shows that such a small gap would be insufficient to restrain the electrons within the rather oblong Brillouin zone. A further point is that, since more than one of the $\alpha^{\mu\nu}$ was assumed to be large, at least one additional excited band would have to be supposed to lie within a small fraction of a volt of the overlapping states at the sites D. Mr. Joel McClure and the writer have constructed all of the reciprocal lattice planes which are relevant to the D sites and have found that no other planes come at all close. It would seem most improbable, then, that the situation with regard to the D sites could be quite as Jones has supposed.

Even though we do not believe that Jones description of the overlap configuration is quantitatively correct, we want to discuss the order of magnitude of the terms χ_2 and χ_3 which would have been expected for such a configuration. We will show that for any configuration like that at the D sites χ_1 would be a poor approximation to the total susceptibility. In order to demonstrate that the terms in χ_2 and χ_3 could not be ignored, we will

[12] de Haas-van Alphen, Leiden Comm. **212a** (1930).

[13] E. H. Sondheimer and A. H. Wilson, Proc. Roy. Soc. (London) **A210**, 173 (1952). In this paper can be found references to much of the work beginning with Peierls original paper.

[14] R. B. Dingle, Proc. Roy. Soc. (London) **211**, 500 (1952); **212**, 38 (1952).

simply select one of the terms from χ_3' and show that it alone could give a susceptibility as large as χ_1. From the expression (2.9) for we can select the term

$$h\equiv-\sum_{n'}(\hbar^2/E_{nn'})[s_{nn'}{}^{\nu\beta}s_{n'n}{}^{\gamma\lambda}+\text{c.c.}]\delta_{\nu\lambda z}\delta_{\beta\gamma z}. \quad (6.2)$$

Using the definition of $s_{jj'}{}^{\mu\nu}$ given in (2.8), we can write (6.2) as

$$h=-\sum_{n'}(2\hbar^2/m^2E_{n'n}{}^3)(P_{n'}{}^{\beta}-P_n{}^{\beta})(P_{n'}{}^{\nu}-P_n{}^{\nu}) \times P_{nn'}{}^{\gamma}P_{n'n}{}^{\lambda}\delta_{\nu\lambda z}\delta_{\beta\gamma z}. \quad (6.3)$$

Since we want to see the order of magnitude, we will ignore off-diagonal components of the $\alpha^{\mu\nu}$ and write

$$P_n{}^{\beta}\doteqdot p^{\beta}\alpha_n{}^{\beta\beta}. \quad (6.4)$$

Then

$$h=-\sum_{n'}(2\hbar^2/m^2E_{nn'}{}^3)[p_x{}^2(\alpha_{n'}{}^{xx}-\alpha_n{}^{xx})^2|P_{nn'}{}^{y}|^2 + p_y{}^2(\alpha_{n'}{}^{yy}-\alpha_n{}^{yy})^2|P_{nn'}{}^{x}|^2]. \quad (6.5)$$

For simplicity, let us set $|\alpha^{xx}|\sim|\alpha_{yy}|\sim\alpha$. Then (6.5) is of the order

$$h\sim-4\hbar^2(p_x{}^2+p_y{}^2)\alpha^3/m\Delta E^2, \quad (6.6)$$

in which ΔE denotes some average excitation energy which is of the order of the excitation from the first to second band. Now if ζ_2 is the energy to which an ellipsoid in zone 2 is filled, there will be occupied in the lower band a volume of about

$$V_{K1}\sim(2m)^{\frac{3}{2}}\Delta E\alpha^{-1}(\Delta E+\zeta_1)^{\frac{1}{2}}. \quad (6.7)$$

Inserting h as given by (6.6) into Eq. (3.18) and using the volume given by Eq. (6.7), we get an estimate of the susceptibility contribution of the term Eq. (6.2) from the low band. This is

$$\chi\sim\mu_B{}^2[2/(2\pi\hbar)^3](2m)^{\frac{3}{2}}\Delta E\alpha^2(\Delta E+\zeta_1)^{\frac{1}{2}}(p_x{}^2+p_y{}^2/m\Delta E^2);$$

or, setting $(p_x{}^2+p_y{}^2)_{\max}\sim m\Delta E/\alpha$, $\Delta E\sim E_F/\alpha$, we find

$$\chi\sim\chi_L\alpha^{\frac{1}{2}}. \quad (6.8)$$

We can compare the susceptibility Eq. (6.8) with the surface integral χ_1 over the ellipsoid in the second zone. This is

$$\chi_1=-(\mu_B{}^2/3)(8\pi/(2\pi\hbar)^3)m(2m\zeta_2)^{\frac{1}{2}} \times|\alpha^{xx}\alpha^{yy}|^{\frac{1}{2}}|\alpha^{zz}|^{-\frac{1}{2}}. \quad (6.9)$$

Now as we saw in the last section, $\Delta E/2=\zeta_2$ in (6.9) gives the approximate maximum value which χ_1 will attain for a single ellipsoid, regardless of the value of ζ_2, in view of the fact that $\alpha^{\mu\mu}$ approach unity for very large ζ_2. Putting $\zeta_L\sim\Delta E/2$ into (6.9), we get

$$\chi_{1\,\max}\sim-\chi_L\alpha^{\frac{1}{2}}. \quad (6.10)$$

We see that for the kind of overlap configuration which Jones considered, the susceptibility χ_3 is of the same order as the susceptibility χ_1. As it is, we considered only those occupied states in the lower band which lay in a "collar" around the central point of overlapping. If we had assumed that the gap was small all the way across a zone face and that the lower band states were occupied across the face, we would find for χ_3 something larger in powers of α than what we wrote in Eq. (6.8); such a result emphasizes the incompatibility of the assumption of large α and of small overlap.

We have shown that it is doubtful that for a configuration of the type at the D sites one can get a useful interpretation of the susceptibility merely on the basis of the Landau-Peierls formula in the case that the gap across the zone face is small. We have also shown the assumption that the zone of Fig. 5 *is* about full to be incompatible with the assumption that any small groups of overlapping electrons at D sites would have very small effective masses. We will now discuss the holes in a zone corner like A or B. We need not go into great detail in such a discussion since the arguments are very much the same as we have used before. In the first place, if the $\alpha^{\mu\nu}$ are really large for the corner states, it means that the energy of excitation to the next band is very small. In an unreduced zone picture this must mean that there is a very small gap across one or more of the faces which make up the corner. Since a corner is always energetically higher than the points on the faces which form it, we would find that because of the smallness of the gap across one face the Brillouin zone would not be "strong" enough to keep the electrons in. This again is a special argument applying only to the model of Bi in which high corners and small gaps are both assumed. The argument that if the $\alpha^{\mu\nu}$ are large χ_3 has terms of the same order as χ_1 applies to this configuration also. While we can hardly discuss χ_2, it is a fair assumption that it too may be appreciably large in both configurations. For if two or more of the $\alpha^{\mu\mu}$ are large, there must be at least two important intermediate states. With two intermediate states, "roundabout" transitions need not be small so χ_2 need not be small.

The discussion which we have given for the special configurations which occur in Jones' model of Bi can easily be recast to apply to other configurations. It will always be found that when the excitation energy from a group of occupied states to other states is much smaller than the Fermi energy, the Landau-Peierls formula does not contain a good approximation to the total susceptibility contribution from that group of states. So goes the general conclusion. In the course of this investigation the writer has come to doubt that Jones' theory of Bi is even qualitatively correct. One may even question whether a zone picture of energy bands is appropriate in Bi, since Bi shows some tendency toward covalent bonding. Without attempting to settle this question the writer would like to suggest that there is a modification of Jones' scheme which appears to offer the possibility of accounting in a rough way for both the parallel and perpendicular susceptibilities. The modified model might well retain some degree of validity, even

if there is a tendency towards forming a covalent bond between nearest neighbors, since the electrons which provide the diamagnetism are distinct from those which provide the covalent bond. Suppose that we use the Brillouin zone description of the bands, but admit that the gaps across some of the zone faces may be rather large. Let us suppose that the energy gaps across the $(1\bar{1}0)$ faces at D amount to several volts so that indeed there is little or no overlap at D. In such a case any overlapping electrons at D would not have especially small masses so would not contribute any great diamagnetism. In fact, if the density of energy levels were quite high, such electrons might give an appreciable *paramagnetic* contribution on account of their spin moments. The source of the observed large diamagnetism of Bi would then have to be elsewhere, and we suggest that the probable source is a group of regions, one of which we have labeled C in Fig. 5. The details of the Bi zone structure indicate that if the overlap at D is not too great there will be appreciable overlap at C. A number of "strong" planes at about the Fermi radius intersect the energetically high corners of the zone shown in Fig. 5 and effectively prevent any electrons from getting near the corners. With the corners of the zone thus cut off there is not room in the zone for five electrons per atom so overlap must occur somewhere. Assuming that the $(1\bar{1}0)$ faces are strong enough to restrict the amount of overlap at D, then the overlap could only occur at C. Let us suppose that it happens so, and consider what the energy level structure might be like in the neighborhood of C. The top faces of the zone are formed by the (221) planes. One of these planes would be expected to have a gap across it which is smaller than that across the $(1\bar{1}0)$ planes but still substantial. The gap would not be small enough in itself to explain any very small effective masses. However, there is another feature of the zone structure which could provide the required complexity of energy levels at the C sites. The strong (211) planes intersect in lines which pass obliquely through the zone faces formed by the (221) planes, and the point at which the line passes through the (221) face is very close to the point where we would have expected overlap to occur most easily. It would not be strange if there are electron states of very small mass in the neighborhood of the intersection point. Thus it would appear that the C sites are naturally endowed with the kind of energy structure which could contribute to a large diamagnetism.

No attempt has been made to work out the consequences of such a model in detail, but some of them can be foreseen without calculation. One of the most striking is that the angle which the (221) faces make with the principal axis is favorable for explaining the observed anisotropy effect of the diamagnetism, without bringing in two separate sites as sources of large diamagnetism. In this respect the model proposed here is even simpler than Jones' original model. On the other hand, other results of the present work make it clear that the diamagnetism could not be simply characterized by a set of effective mass values and a number of overlapping electrons, since off-diagonal matrix elements would play an important role. It would probably be quite difficult to establish the proposed model on a quantitative basis just because of the complications which come into the calculation of the susceptibility when the effective mass is very small. There would be analogous anomalies in a calculation of the electrical properties, so it is not clear whether or not the electrical properties would follow naturally from the model. At the moment all we claim for the model is that it is not palpably absurd, it involves only the most natural hypothesis about the energy structure, and it probably leads to magnetic anomalies of the kind observed for Bi.

7. SUMMARY AND CONCLUSIONS*

In this paper we have derived an expression for the orbital magnetic susceptibility of an electron gas which moves in a periodic potential. The derivation makes use of the Wannier formalism. The final formula for the susceptibility is somewhat simpler than previous formulas, because the Wannier formalism contains implicitly certain sum rules which were not previously noticed.

Using our formulas for the susceptibility we have studied the question as to whether the Landau-Peierls formula, Eq. (1.1), gives a good approximation to the total orbital susceptibility. What we have found is that if any portion of the Landau-Peierls susceptibility is the result of electrons with small effective masses, then the total orbital suceptibility differs from the Landau-Peierls susceptibility by an amount of the same order as that portion. More precisely, if the energy gaps between bands are appreciably smaller than the Fermi energy, the total susceptibility differs from the Landau-Peierls susceptibility by an amount which is of the same order of magnitude as the difference between the Landau-Peierls susceptibility and the Landau susceptibility, Eq. (1.2).

We have treated in detail the case of a weak plane in the reciprocal lattice and shown that such a plane can give large contributions to the orbital diamagnetism when it is overlapped by the Fermi surface. The large diamagnetism resulting from such a plane was shown to be not identical with the Landau-Peierls diamagnetism[6] in fact, for large overlap this diamagnetism was much larger than the Landau-Peierls diamagnetism. If one or more such planes formed the boundary of a filled zone, the zone could make a large contribution to the susceptibility, even though there is no Fermi surface in the zone. This part of the susceptibility could be expected to be anisotropic in a manner determined by the arrangement of the weak planes. It was pointed out that the occurrence of such a susceptibility would cor-

* See also "Note added in proof" at end of this section.

respond to some electrons being weakly bound to the lattice ions forming larger ions with very low-lying excited states.

We have examined Jones' theory of the Bi diamagnetism and concluded that his hypothesis concerning the location of the Fermi surface is incompatible with his assumptions concerning the effective masses. We have suggested a modification of Jones' hypothesis which we believe to be more in accord with the detailed structure of the Brillouin zones and the evidences of strong binding in Bi. It was indicated how the modified model may give a simpler explanation of the over-all features of the Bi susceptibility.

In conclusion, we would like to briefly discuss two theories on which we believe our discussion has some bearing but to which we cannot make a detailed application. The first in Bardeen's theory[6] of superconductivity which was referred to earlier. Bardeen has discussed a case in which the α's were supposed to be so large that the diamagnetism, Eq. (1.1) would be more than sufficient to drive out the field. While we cannot discuss Bardeen's theory itself, it should be clear from the examples discussed in this paper that (Eq. (1.1) is unreliable when the electron effective mass is very small. Another feature of Bardeen's theory which we can comment on in this connection is his "self-consistent field" argument showing that electrons with small effective mass can keep the magnetic field out of a metal because the penetration depth is considerably smaller than the electron wave function. In making this argument, Bardeen set up a Hamiltonian which was the same as the ordinary free electron Hamiltonian except that the effective mass occurred in the place of the free electron mass. The approximation involved in writing such a Hamiltonian is known,[8,9] and the terms neglected are just the terms which give the additional contributions to the susceptibility which are the subject of this paper. As we have seen when the effective mass is very small, the neglected terms are just as important as those which are kept. Thus we can conclude that Bardeen's calculation concerning the penetration of the magnetic field is based on an incomplete Hamiltonian, and that the terms which have been omitted are so large as to bring into question the outcome of a calculation made with the correct Hamiltonian.

There is another matter to which our results cannot be applied and yet about which they indicate something significant, namely the interpretation of the de Haas-van Alphen effect in the diamagnetic metals. At attainable field strengths the effect cannot be observed except in those substances for which there exist electrons of very small electron mass. Dingle[13] has done a careful analysis of a number of factors which cause the magnetic susceptibility of a real material to differ from the ideal crystal-bulk susceptibility considered here. He has included, for example, the effects of finite level width and of finite crystal size, and has done the calculation so as to get the complete temperature dependence and field dependence of the susceptibility. The apparent purpose of his work is to lay the foundation for a study of the de Haas-van Alphen effect in a real metal. Dingle worked in the approximation that the electrons are completely free, and in this approximation obtained the result quoted above that electrons of very small mass must be responsible for the observed de Haas-van Alphen effects. The part of his work which has as yet been published indicates that he considers the effective mass approximation adequate even for very refined calculations. The results of our work, on the other hand, make it certain that the effective mass approximation should be completely inadequate for an interpretation of observed de Haas-van Alphen effects. We have shown that if the effective masses are small there is an important additional term in the steady susceptibility arising from the term in the Hamiltonian which is neglected in the effective mass approximation. There will be similar contributions to each term in the development of the susceptibility in powers of the field, these contributions corresponding to higher order correction effects of the "neglected" term in the Hamiltonian (2.7) which explicitly involves the magnetic field rather than the vector potential. Accordingly, we must conclude that the effective mass approximation is not adequate for interpreting the de Haas-van Alphen effect in real metals. It unfortunately appears that an adequate calculation which takes into account multiple transitions between bands is almost outside the range of possibility.

The writer expresses his appreciation of enlightening discussions with a number of persons in the Institute for the Study of Metals and in particular the assistance of Mr. Joel McClure in constructing the Fermi surface of Bi.

Note added in proof.—After a number of conversations on the subject of this paper, I think it well to restate explicitly just what is claimed to be the validity of these conclusions. In the first place it is not claimed that the example which is carried through to the end is necessarily directly analogous to the situation obtaining in any of the diamagnetic metals. This example is given to show that in the one case for which the off diagonal terms can be readily computed, they are really as big as the Landau-Peierls term. The calculation cannot readily be made for more interesting configurations, but the estimates given show that sizes of matrix elements and numbers of states of large matrix element are such as to make it probable that in these cases, too, the off diagonal terms are really as large as the Landau-Peierls term. I think these last arguments extremely suggestive of the conclusions but it cannot be claimed that they quite prove the conclusions. However, irrespective of the detailed arguments presented, it can be seen that the key fact in the case of interest is that there exist off diagonal terms which, before integration over k, are large like the inverse cube of the energy gap.

For ordinary energy gaps such terms are negligible but for very small energy gaps they become dominant, unless, of course, they cancel one another. I have looked for and failed to find a reason for such a cancellation.

Finally, it is certainly not concluded that observed diamagnetic susceptibilities are not to be interpreted in terms of effective masses. However, it is concluded that the relation of an empirically determined effective mass to the reciprocal curvature of an energy surface will not be simple when the mass is very small.

PHYSICAL REVIEW VOLUME 89, NUMBER 3 FEBRUARY 1, 1953

The Proton-Proton Reaction and Energy Production in the Sun*

E. A. FRIEMAN,† *Polytechnic Institute of Brooklyn, Brooklyn, New York*
AND
L. MOTZ, *Rutherfurd Observatory, Columbia University, New York, New York*

(Received September 9, 1952)

The proton-proton capture cross section is recalculated using the recently computed deuteron wave functions of Feshbach and Schwinger and the most recent value of the Fermi G factor. The energy production of the p-p cycle is redetermined and applied to the sun, using the luminosity calculations of Epstein. It is shown that the p-p cycle outweighs the carbon cycle by a considerable factor under the assumed conditions of temperature and density.

I. INTRODUCTION

IN view of recent results which affect the conclusions of Bethe and Critchfield[1] on the proton-proton reaction,

$$H^1+H^1\rightarrow H^2+\beta^++\nu,$$

this process is recalculated and its astrophysical implications discussed for the sun. Epstein[2] showed recently that the energy production in the sun arising from the p-p cycle outweighs that arising from the carbon cycle by at least a factor of 12. This calculation utilized the new determination of the cross sections for the carbon cycle by Hall and Fowler[3] extrapolated from their measurements at 100 kev. In the present calculation the most recent value of the Fermi constant as well as new deuteron wave functions were used in the calculation of the cross section. The energy production per process also differs from that described in reference 1 because the He^3 reaction replaces the sequence used there.

II. THE CROSS SECTION FOR THE $P-P$ REACTION[4]

The cross section for the process, expressed as a function of the relative velocity v, is

$$\sigma(v)=(G/v)|\bar{M}|^2 f(W), \tag{1}$$

where $\bar{M}$ is the nuclear matrix element,

$$\bar{M}=\int \Psi_f \bar{\Gamma} \Psi_i d\tau, \tag{2}$$

and $\bar{\Gamma}$ is the Gamow-Teller interaction operator,

$$\bar{\Gamma}=(3)^{-\frac{1}{2}}\sum_i \tau^i\bar{\sigma}^i. \tag{3}$$

The proton space wave functions are those given by Yost, Wheeler, and Breit[5] for a square well and are given in concise form in reference 1. We can represent them as

$$\Psi_p(r,\bar{\sigma},\tau)=\Psi_p(r)\,{}^1(\sigma)_0\,{}^3(\tau)_{-1}, \tag{4}$$

where we have used Rosenfeld's[6] notation to represent the spin and isotopic spin wave functions.

The deuteron wave functions used are those computed by Feshbach and Schwinger[7] for "Yukawa well" interactions of different ranges for both the central and tensor portions of the interaction potential. These eigenfunctions can be represented by

$$\Psi_d=(4\pi)^{-\frac{1}{2}}\left[\frac{u(r)}{r}+(2)^{-\frac{1}{2}}S_{12}\frac{w(r)}{r}\right]{}^3(\sigma)_{m_S}\,{}^1(\tau)_0. \tag{5}$$

Using the well-known property of the tensor operator S_{12}, namely, that its average over all angles vanishes, it is seen that there is no D wave contribution to the matrix element. The spin summation is then easily performed using the Hermitian character of the G-T operator, since the sum over the final states of $|\bar{\Gamma}|^2$ is just the expectation value in the original state. The contribution to the cross section from the sum is simply a factor of two. The square of the matrix element can then be written

$$|\bar{M}|^2=8\pi\left|\int_0^\infty \psi_p(r)u(r)rdr\right|^2. \tag{6}$$

* This paper is based on a Ph.D. thesis submitted to the Graduate School of the Polytechnic Institute of Brooklyn. A preliminary report of this work was presented at the May, 1951, meeting of the American Physical Society.

† Now at the James Forrestal Research Center, Princeton University, Princeton, New Jersey.

[1] H. A. Bethe and C. Critchfield, Phys. Rev. **54**, 248 (1938).

[2] I. Epstein, Astrophys. J. **112**, 207 (1950).

[3] R. N. Hall and W. A. Fowler, Phys. Rev. **77**, 197 (1950).

[4] The symbols used here are the same as those given in reference 1 unless otherwise stated.

[5] Yost, Wheeler, and Breit, Phys. Rev. **49**, 174 (1936).

[6] L. Rosenfeld, *Nuclear Forces* (Interscience Publishers, Inc., New York, 1949).

[7] H. Feshbach and J. Schwinger, Phys. Rev. **84**, 194 (1951). We wish to thank Professor Feshbach for a prepublication copy of his paper.