Saturation Hall Constant of Semiconductors

JOHN A. SWANSON
International Business Machines Corporation, Poughkeepsie, New York
(Received January 20, 1955)

The asymptotic value of the Hall constant of semiconductors as the magnetic field becomes very large is shown on general grounds to be $1/(p-n)ec$, where p and n are the concentrations of holes and of electrons respectively; this result is independent of the band shapes in the neighborhood of energy maxima or minima. The unscattered motion of electrons, which determines the saturation Hall constant, is examined by means of the usual crystal momentum force equation, and also through direct investigation of the properties of stationary states in crossed electric and magnetic fields. The saturation magnetoresistance is also discussed, within the framework of the crystal momentum approximation.

INTRODUCTION

THE Hall constant of semiconductors may exhibit considerable variation with magnetic field strength. At vanishingly small fields, the Hall constant of extrinsic semiconductors takes the form

$$R_0=\pm\frac{\mu_H}{\mu}\frac{1}{n_\pm ec}, \tag{1}$$

where $n_\pm$ is the concentration of holes (+) or of electrons (−), μ is the drift mobility, μ_H is the Hall mobility, and e and c have their usual significance. The ratio μ_H/μ in general differs from unity, since electrons in different states may be deflected through different individual Hall angles.

As the magnetic field is increased the Hall constant changes, and may eventually approach a saturation value, R_∞. Theoretical calculations of Hall constant *versus* magnetic field strength have been carried out for certain special types of energy band—those having spherical[1] or ellipsoidal[2] surfaces of constant energy in crystal momentum space. In these instances the saturation Hall constant is

$$R_\infty=\pm 1/n_\pm ec. \tag{2}$$

In this note we shall examine specifically the theory of the saturation Hall constant, and shall attempt to generalize somewhat the conditions under which formula (2) is valid.

I. PHYSICAL INTERPRETATION OF THE SATURATION CONSTANT

At saturation the longitudinal current can be calculated exactly as though the particles were unscattered, that is, as though they followed their natural motion in the electric and magnetic fields.

The reasoning leading to this conclusion, and to the value of the saturation constant itself, is especially straightforward if all the particles have a single effective mass, m^*. The orbits of such particles in crossed electric and magnetic fields are well known; they are cycloids with angular frequency

$$\omega=eH/m^*c \tag{3}$$

and translational velocity

$$v_0=(E/H)c \tag{4}$$

in the direction $(\mathbf{E}\times\mathbf{H})$. A simple proof of this consists in transforming to a coordinate system moving with velocity $\mathbf{v}_0$ relative to the initial system. In this new coordinate system there is no electric field.

Let the mean free time of a particle in a given cycloidal orbit be τ. Then on the average a particle will traverse the angle

$$\theta'=\omega\tau, \tag{5}$$

before being scattered. If θ' is small, the Hall angle θ between current and electric field is given by

$$\theta=\omega\tau_H, \tag{6}$$

where τ_H is a weighted average of the mean free times for different orbits, in general not identical to the average used in calculating the drift mobility. The definition of the Hall mobility,

$$\mu_H=c\lim_{H\to 0}\frac{\theta}{H}, \tag{7}$$

leads to the equation

$$\mu_H=(e/m^*)\tau_H. \tag{8}$$

This abbreviated conventional analysis shows that the Hall constant at low magnetic fields depends on the distribution of mean free times.

When the magnetic field is sufficiently large, θ' for all particles becomes greater than 2π, while θ cannot exceed $\pi/2$. The particles follow their natural motion for one or more revolutions, and hence between collisions possess average velocity v_0. Since θ approaches $\pi/2$, this velocity has the direction of the normal longitudinal current. It is reasonable to suppose that collisions merely carry a particle from one orbit to another without adding any characteristic displacement in the longitudinal direction. In this event the longi-

[1] A. H. Wilson, *Theory of Metals* (Cambridge University Press, Cambridge, 1953), Chap. VIII.
[2] Motoichi Shibuya, Phys. Rev. **95**, 1385 (1954).

tudinal current i_l is

$$i_l = \pm n_\pm e v_0, \tag{9}$$

and the saturation Hall constant is clearly given by Eq. (2).

The fact that scattering may be ignored at high fields leads us to make a more general analysis of the motion of electrons in crystals under crossed electric and magnetic fields. Proof that scattering does not affect the saturation Hall constant in the general case will be deferred until Sec. III, since it will be desirable to have the theory of the unscattered motion already in mind when considering this and related questions.

II. MOTION IN CROSSED FIELDS

Orbits in Momentum Space

The particle orbits in momentum space are determined by the force equation

$$dP/dt = -e[\mathbf{E} + (\mathbf{v}\times\mathbf{H})/c], \tag{10}$$

and the associated velocity equation

$$v_i = \partial W/\partial P_i, \tag{11}$$

where $\mathbf{P}$ is the crystal momentum, $\mathbf{v}$ the velocity, and W the band energy. This description of the motion does not include quantum effects specific to the oscillation of the particle in its orbit. We shall discuss this question in Sec. IV.

In order that $\dot{P}_i$ may be continuous we shall not employ the reduced zone scheme initially, but instead allow P_i to take on all real values. The energy in one band then becomes a periodic function of crystal momentum, i.e., three primitive $\mathbf{P}$-vectors may be defined such that any displacement which is a sum of primitive displacements leaves the energy function unchanged. The images of the orbits in the reduced Brillouin zone will be discussed later.

Let us set up a right-handed coordinate system with orthogonal unit vectors $\mathbf{u}_x$, $\mathbf{u}_y$, and $\mathbf{u}_z$ pointing in the direction of the x, y, and z axes respectively. The directions of the electric and magnetic fields are specified as follows:

$$\mathbf{E} = E\mathbf{u}_x,$$
$$\mathbf{H} = -H\mathbf{u}_z, \tag{12}$$

so that the vector $(\mathbf{E}\times\mathbf{H})$ points in the direction of the y-axis.

The motion is confined to a plane in momentum space

$$P_z = \text{const}. \tag{13}$$

The other component equations of (10),

$$\dot{P}_x = -e[E - v_y H/c], \tag{14}$$

$$\dot{P}_y = -e v_x H/c, \tag{15}$$

can be combined to yield

$$dP_x/dP_y = (cE - v_y H)/v_x H, \tag{16}$$

which in conjunction with Eq. (11) determines the orbit. A more informative equation is obtained by using the well-known relation

$$\dot{W} = \mathbf{v}\cdot d\mathbf{P}/dt, \tag{17}$$

which follows from Eq. (11). Substituting the expression for $d\mathbf{P}/dt$ given by Eq. (10), we have

$$\dot{W} = -eEv_x, \tag{18}$$

and from Eqs. (15) and (18) it follows that

$$dW/dP_y = v_0, \tag{19}$$

where v_0 has the value cE/H, as in Eq. (4). We shall base our further discussion on this relation.

The orbits in P-space, as given by Eqs. (13) and (19), may be visualized as follows: Confining ourselves to the plane $P_z = \text{const}$, let the energy as a function of P_x and P_y be plotted in what was formerly the P_z direction. Then the orbits in P-space are the projections on the plane $P_z = \text{const}$ of the intersection of planes with slope $dW/dP_y = v_0$ with that energy surface.

It is clear that all of the orbits in P-space are either: (a) infinite in extent with average direction along the electric field, or (b) reentrant. For if an orbit were confined to a bounded region of the plane $P_z = \text{const}$ and were not re-entrant, there would have to be an infinite number of distinct intersections of a particular plane of slope (19) with the energy surface, and this is of course impossible in cases of physical interest.

If v_0 is different from zero and not too large, orbits of both type (a) and type (b) exist. The orbits are obviously reentrant near a maximum or minimum of the energy surface (i.e., the surface which represents energies in the $P_z = \text{const}$ plane, and conceived in the preceding as erected above that plane), providing v_0 is small enough. It will be shown later that orbits of type (a) must then also exist. The special case $v_0 = 0$ corresponds to an infinite magnetic field or zero electric field. In this case all orbits are constant energy contours. For $v_0 = \infty$ all the orbits are infinite straight lines, corresponding to motion in a pure electric field.

We are interested primarily in the time average of the velocity of particles in these momentum orbits. Using the angular parentheses to designate the time average of the included quantity, we find from Eqs. (14) and (15) that

$$\langle v_y\rangle = v_0 + \langle\dot{P}_x\rangle c/eH, \tag{20}$$

$$\langle v_x\rangle = -\langle\dot{P}_y\rangle c/eH = 0, \tag{21}$$

the last equation following from the fact that P_y is bounded (since the energy in one zone is bounded). Furthermore, for any re-entrant orbit $\langle\dot{P}_x\rangle$ is zero, and thus

$$\langle v_y\rangle = v_0 \text{ (all reentrant orbits)}. \tag{22}$$

It is Eq. (22) which allows us to realize the more general validity of Eq. (2).

We next consider the images of these orbits in the

reduced Brillouin zone. Let us select one plane $P_z=$ const and one energy plane $W=v_0P_y+$const. Then if there exists no vector in the direction of the electric field which can be expressed as a sum of the primitive momentum displacement vectors, the image of the orbits so defined will be everywhere dense in the fundamental zone or on a surface in the fundamental zone. Let us for simplicity choose a direction for E which contains such a vector. Then the orbits in the topological space of the reduced zone are "closed" curves. We may generate a complete set of such closed orbits by imagining the reduced zone to be completely populated at one instant of time by the array of points representing the stationary states. The orbits subsequently traced out by these points give us our complete set.

Having a complete set of orbits in the reduced zone we can attach a more definite meaning to the terms of the following statement: the sum of the average velocities associated with all orbits is zero. (If several orbits coincide, the contribution of each must, of course, be counted.) This must be so since the total current contributed by a full Brillouin zone is zero. One should recall here that the flow of points in P-space is "incompressible." Since all the reentrant orbits have the average velocity v_0, those orbits which are not reentrant in the infinite P-space have on the average an opposite velocity. Both reentrant and non-re-entrant orbits must exist in general in order that the sum of all average velocities be zero.

Application to the Hall Effect

Let us suppose that the great majority of electrons in a given band have energies less than a certain energy W_0. Consider the volume in extended P-space composed of points associated with energies less than or equal to W_0. If this volume is composed of subvolumes which are finite in extent and disjoint, then the current carried by the corresponding filled orbits in the fundamental zone is $-n_-ev_0$, where n_- is the number of electrons in the band. This follows from the fact that the electrons are confined to these disjoint regions if the velocity v_0 is small enough. If there are several partially filled bands, each satisfying the above criterion, the same formula for the current will apply, with n_- equal to the total number of electrons in such partially filled bands. The image of the disjoint regions on the reduced zone may be either connected or disconnected.

The same result clearly holds for holes. An empty orbit can be considered to carry the current $+ev_0$. If the regions in P-space corresponding to all relevant unfilled states are finite in extent and disjoint, the current is n_+ev_0, where n_+ is the number of unoccupied states.

Experimentally, saturation of the Hall constant is obtainable only if the mean free time is sufficiently large that at available fields most of the particles average at least one traversal of their orbits between scattering collisions. If the electrons are inhomogeneous, i.e., are spread among several bands or different energy minima, the saturation condition may be realized by different groups of electrons at different field strengths. As a result the Hall constant may still change materially after the first apparent approach to saturation.

We conclude this section by pointing out that when both holes and electrons are present the longitudinal current density is

$$i_l=ev_0(n_+-n_-), \tag{23}$$

and the Hall constant is

$$R_\infty=1/(n_+-n_-)ec. \tag{24}$$

This is a well-known formula, now validated more generally. For a near-intrinsic semiconductor the criterion for saturation is somewhat different, as explained in Sec. III.

III. STATISTICAL TREATMENT

In this section, we shall apply the Boltzmann transport equation to the problem of determining the distribution function at high fields. This analysis enables us to show that the high field Hall constant is determined by the free motion of the carriers. The discussion also leads naturally to a study of the saturation magnetoresistance.

Let f be the probability of occupation of states of a band in the reduced zone of P-space. If the concentration of carriers is uniform in space, the Boltzmann equation for this band is

$$\frac{d\mathbf{P}}{dt}\cdot\nabla_P f=-e[\mathbf{E}+(\mathbf{v}\times\mathbf{H})/c]\cdot\nabla_P f=\left[\frac{\partial f}{\partial t}\right]_{\text{coll}}, \tag{25}$$

where the symbol ∇_P denotes the gradient with respect to crystal momentum. The right member of Eq. (25) represents the net rate of change of occupation probability due to scattering collisions.

From Eq. (25) one can deduce in a general manner the behavior of the electric field as the magnetic field increases without limit while the longitudinal current density remains constant. Since the right member of Eq. (25) remains finite while the force $d\mathbf{P}/dt$ becomes infinite, the gradient of the distribution function along the direction of the force must approach zero. Thus the distribution function tends to become constant along the unscattered orbits. It follows that the electric field must become infinite; otherwise, the limiting distribution function would have zero current associated with it. However, the component of the electric field in the direction of $\mathbf{H}$ will remain finite, since a finite electric field is sufficient to cause a current in that direction. In the limit the electric field therefore becomes perpendicular to the magnetic field.

It follows from the above that the current density at

saturation in the plane perpendicular to H is determined by the unscattered motion of the carriers in the magnetic and transverse electric fields. This current, as we have seen, is orthogonal to both fields. Should the magnetic field be applied in a direction not perpendicular to that of the longitudinal current density, a finite electric field in the same direction as the magnetic field is required to produce a balancing current, such that the resultant current lies in the longitudinal direction. The current density due to the unscattered motion of the particles in the magnetic and Hall electric fields is thus the projection on the axis perpendicular to E and to H of the longitudinal current density. It follows that the saturation Hall constant is determined by this unscattered motion, and has the same value for all directions of H, provided that the definition of the Hall constant is generalized in the usual way:

$$\mathbf{E}_H = -R(\mathbf{i}\times\mathbf{H}), \tag{26}$$

in which $\mathbf{E}_H$ is the Hall field, and $\mathbf{i}$ the vector current density.

In the following we shall treat the saturation Hall constant and also the magnetoresistance in a more rigorous fashion for the case in which the magnetic field is perpendicular to the longitudinal current density. We will adopt the same notation and direction for each field as in the previous section.

It is convenient for the analysis to constrain one field to be a constant multiple of the other, i.e.,

$$E = Hv_0/c. \tag{27}$$

The constant v_0 determines a set of unscattered orbits in the crossed fields. Let us define for every such orbit a specific orbit time t' such that t' is the time (measured with respect to an arbitrary fixed reference moment) at which the particle occupies a particular point of its orbit in a magnetic field of unit strength. In a general field H, the corresponding time t is related linearly to t'

$$t' = Ht, \tag{28}$$

and Eq. (25) becomes in terms of t'

$$Hdf/dt' = [\partial f/\partial t]_{\text{coll}}. \tag{29}$$

In Eq. (29) the right member is a partial derivative with respect to time at a fixed point in momentum space, while the term df/dt' is a total derivative with respect to specific time along the orbit passing through the point in question.

Let us assume that at sufficiently high fields the following expansions are valid:

$$f - f_0 = h_0 + h_1H^{-1} + h_2H^{-2} + \cdots, \tag{30}$$

$$[\partial f/\partial t]_{\text{coll}} = s_0 + s_1H^{-1} + s_2H^{-2} + \cdots. \tag{31}$$

It will prove useful to define the additional quantities

$$\bar{h}_0 = f_0 + h_0, \tag{32a}$$

$$\bar{h}_i = h_i \quad (i \neq 0). \tag{32b}$$

The quantity $\bar{h}_0$ is the saturation probability distribution function for the particular value of v_0 selected.

Applying Eq. (28) to the expansions (29) and (30) we find that

$$d\bar{h}_{i+1}/dt' = s_i. \tag{33}$$

The first equation of the set is

$$d\bar{h}_0/dt' = 0 \tag{34}$$

showing that $\bar{h}_0$ is constant along each of the orbits determined by v_0. Thus, provided that the series (30) converges, the saturation Hall constant is rigorously shown to be determined by the unscattered motion. It should be noted, however, that Eq. (34) does not determine the distribution of electrons among the orbits. As one would expect, this distribution is determined by the law of scattering.

Let us in fact note that every s_i must satisfy the condition

$$\oint_{v_0} s_i dt' = 0. \tag{35}$$

The loop integral sign indicates that the scattering function s_i is to be integrated around one complete revolution of the particle in one of the orbits determined by the constant v_0. The proof of Eq. (35) consists in integrating Eq. (33) with respect to the time around one of the closed orbits. The integral of the left hand member is the change of h_i around the orbit, which must be zero.

Equation (35) constitutes a condition on h_i, since h_i determines s_i through the scattering law. This condition is needed to specify h_i uniquely, since to any h_i satisfying (33) we can add a function constant along the orbits but otherwise arbitrary, thereby obtaining a new solution. In particular, the value of $\bar{h}_0$ along a given orbit is determined by Eq. (35) as applied to s_0.

Rather than discuss the general case further, we shall suppose that a mean free time exists. The problem is then greatly simplified, since the functions s_i are given in terms of h_i by the relation

$$s_i = -h_i/\tau. \tag{36}$$

The set of equations resulting from substitution of this expression for s_i into Eqs. (33) is easily soluble, and moreover the solution may be found along each orbit independently. For suppose h_i has been determined; then we may determine h_{i+1} by integrating part way around an orbit:

$$h_{i+1}(P) - h_{i+1}(P_0) = -\int_{t_0'}^{t'} h_i dt'/\tau. \tag{37}$$

Here P_0 and t_0' are the coordinates of an arbitrary reference point on the orbit, and P and t' are the coordinates of a general point. Equation (36) determines h_{i+1} along every orbit to within an arbitrary constant. This constant is determined from the con-

dition

$$\oint_{v_0} h_{i+1}dt'/\tau=0, \tag{38}$$

following from Eq. (33).

We next show that the series (30) obtained in this manner converges if H is sufficiently large. We note that the maximum absolute value of h_{i+1} along a particular orbit is less than the maximum absolute value of h_i integrated around the complete orbit. This follows from Eq. (37). Indicating the maximum absolute value along an orbit by the subscript m, we have

$$h_{i+1,m}<h_{i,m}\oint_{v_0} dt'/\tau. \tag{39}$$

Since h_i must take both positive and negative values, the inequality (39) is in general rather strong. It is useful for our purposes to express this inequality in a different form. Taking the line integral with respect to the actual orbit time t we have

$$\frac{h_{i+1,m}}{h_{i,m}H}<\oint_{v_0} dt/\tau. \tag{40}$$

The left member of Eq. (40) is the ratio of the maximum absolute values of adjacent terms in the series (30). Thus the series converges for a given orbit if

$$\oint_{v_0} dt/\tau<1. \tag{41}$$

This condition may be assured for all relevant (occupied) orbits by choosing a sufficiently large magnetic field.

Thus if a mean free time exists, the convergence of series (30) for sufficiently high magnetic fields can be proven rigorously. We note, however, that the series is not convergent for all magnetic fields. In fact, it is not difficult to show that for the condition

$$\oint_{v_0} dt/\tau>2\pi \tag{42}$$

the series diverges.

When the left member of inequality (41) is very small compared to unity, it is approximately equal to the probability of the occurrence of a scattering collision during the time of one revolution of the particle. We may consider the maximum absolute value (along an orbit) of the ith term in the series (30) as a rough measure of the ith order current contributed by a particle in that orbit. Applying inequality (40) to the first and second terms of the series (30), we see that the major part of the current contributed by particles occupying a given band and region of momentum space is due to the unscattered motion of these particles if the probability per revolution of scattering in that band and region of momentum space is small.

It would be wrong, however, to conclude that the saturation Hall constant is effectively reached when the preceding sentence is true for all relevant bands and regions of momentum space. The condition for the saturation of the Hall constant is that the current density due to the first term of (30) is much larger than the current density due to the second term. But if both nearly empty and nearly filled bands must be considered, the contribution to the current of the holes balances the contribution of an equal number of electrons as far as the current due to the unscattered motion is concerned. The magnitude of the current due to h_0 is therefore reduced, and the current associated with h_1 may continue to be dominant or appreciable at fields much higher than those required to reduce the probability of scattering per revolution to a very small quantity for all orbits. This enables us to understand the implausible discontinuity exhibited by formula (24) at the intrinsic point. At any given finite field the Hall constant is continuous as a function of the net impurity concentration, but in a region of concentrations close to intrinsic it will not have reached its saturation value. Only in the limit of infinite magnetic fields will the Hall constant have reached saturation for all impurity contents, and therefore it is only in this unrealizable case that the discontinuity exists.

To conclude this section we discuss briefly the possibility of expanding the distribution function in the following manner:

$$f=f_0+E(g_1H^{-1}+g_2H^{-2}+\cdots). \tag{43}$$

This expansion is incomplete in the sense that terms of order higher than the first in the electric field are omitted; nevertheless the differential equation governing the g's can be solved in a manner similar to that employed above, with the integrals in this case taken along constant energy contours. The previous expansion is complete since it contains all powers of E by virtue of the constraint $Ec=v_0H$. This constraint does not detract from the generality of the solution, since v_0 is arbitrary.

Saturation Magnetoresistance

Let the Hall angle θ be expressed as

$$\theta=\pi/2-\phi. \tag{44}$$

Then near the saturation of the Hall constant

$$\phi\doteq i_E/i_l, \tag{45}$$

where i_l is the current due to the free motion of the particles, and i_E is the current in the direction of **E**. The electric field E_l in the direction of the longitudinal current can be expressed in terms of ϕ:

$$E_l\doteq E\phi. \tag{46}$$

The resistivity at infinite magnetic field, ρ_∞, which

we shall call the "saturation resistivity," is given by

$$\rho_\infty = E_l/i_l = \phi E/i_l = R_\infty H i_E/i_l. \tag{47}$$

In general the difference in resistivities

$$\Delta\rho = \rho_\infty - \rho_0,$$

where ρ_0 is the resistivity at zero magnetic field, does not vanish, and a "saturation magnetoresistance" exists, proportional to this difference. Since h_0 does not give rise to current in the direction of the electric field, the saturation magnetoresistance is determined by h_1.

The form of the functions $h_0(P)$ and $h_1(P)$ depends on the value of v_0. Thus it may be that the saturation magnetoresistance is current-dependent. In the following we shall be concerned with the limiting value of the magnetoresistance as the longitudinal current density is reduced to zero. The saturation magnetoresistance should not deviate appreciably from this limiting value if the longitudinal drift velocity is much less than the thermal velocity associated with each energy minimum.

We wish to obtain a formula for h_1, but in order to do so it is necessary first to determine h_0. Let us concentrate our attention upon a particular orbit and a particular constant energy contour (energy W_0) which intersects it at a point P_0. Then, if v_0 is sufficiently small,

$$f_0(W) \doteq f_0(W_0) + (\partial f_0/\partial W)(W - W_0), \tag{48}$$

where W is the energy along the orbit. Now consider a point on the energy contour and the adjacent point on the orbit which has the same P_x coordinate. Then, in view of (19),

$$f_0(W) - f_0(W_0) = (\partial f_0/\partial W) v_0 (P_y - P_{0y}). \tag{49}$$

Since P_{0y} is arbitrary we may set $\bar{h}_0(W)$ equal to $f_0(W_0)$. Then Eq. (37) (for $i=-1$) constitutes a condition on P_{0y}:

$$\oint_{v_0} (P_y - P_{0y}) dt'/\tau = 0. \tag{50}$$

Since we are specializing to cases in which v_0 is small, the condition (50) may be approximated by

$$\oint_W (P_y - P_{0y}) dt'/\tau = 0, \tag{51}$$

an integral around a constant energy contour. Rewriting Eq. (49) we have

$$h_0(W) = f_0(W) - (\partial f_0/\partial W) v_0 (P_y - P_{0y}). \tag{52}$$

The two terms on the right-hand side of this equation must correspond to the first two terms of the expansion (43). In fact, had we solved the latter expansion we should have found that

$$g_1 = -(\partial f_0/\partial W)(P_y - P_{0y}), \tag{53}$$

where P_{0y} is determined by Eq. (51). One may employ (52) as an adequate approximation of the correct function near each energy minimum when v_0 is small.

The resultant formula for h_1 is

$$h_1(P) - h_1(P_0) = \frac{\partial f_0}{\partial W} v_0 \int_{t_0'}^{t'} (P_y - P_{0y}) dt'/\tau. \tag{54}$$

To Eq. (54) must be added the usual condition

$$\oint_W h_1 dt' = 0. \tag{55}$$

Equations (50), (53), and (54) determine the saturation magnetoresistance. Should the contours of constant energy in a given plane $P_z = \text{const}$ have a center of inversion in the point of maximum or minimum energy, and should the mean free time possess the same symmetry properties, the problem is further simplified. For then the constant P_{0y} is the same for all the contours in the plane $P_z = \text{const}$, and in fact is equal to the P_y-coordinate of the center of inversion.[3]

It should be noted that this method of finding the magnetoresistance is related to Shockley's "tube integral" method[4] discussed further in the Appendix.

The quantity

$$g_2 = h_1 c/v_0 \tag{56}$$

is independent of v_0. Let I be the operation (consisting of the usual integration over P-space) which, acting on a distribution function, yields the value of the current density in the x-direction (direction of $\mathbf{E}$). Then it is easy to see by virtue of Eq. (47) that

$$\rho_\infty = R_\infty^2 I g_2. \tag{57}$$

IV. STATIONARY STATES IN CROSSED FIELDS

Thus far we have described the unscattered motion of an electron in a periodic potential under crossed electric and magnetic fields in terms of the crystal momentum, i.e., in terms of the stationary states which exist when these fields are absent. In this approximation the electric and magnetic fields are treated as perturbations, a procedure which loses its justification when these fields become excessively large. It is then more appropriate to discuss the motion of the electron in terms of the true stationary states which exist in the presence of the crossed fields. We shall investigate the properties of such stationary states which are relevant to the saturation Hall constant, but not those relevant specifically to the saturation magnetoresistance. That is, we shall attempt to calculate the expectation value

[3] With constant P_{0y}, formula (53) represents a displacement of the equilibrium distribution in P-space, and can be obtained by simpler arguments if one assumes that the effect of electric and magnetic fields is to produce such a displacement. See the paper of M. Kohler, in which the saturation magnetoresistance of metals is regarded from a point of view similar to ours [M. Kohler, Ann. Physik **38**, 283 (1940)].

[4] W. Shockley, Phys. Rev. **79**, 191 (1950).

of velocity of an electron in such a state, but not the probabilities of scattering into other states.

In the following we shall find it convenient to discriminate between "restricted band states" and general states. The term "restricted band state" shall refer to states resulting from a superposition of states primarily from one band or group of degenerate bands. Though not, in general, stationary, such states may be "quasi-stationary" in the sense that the probability density is essentially constant over ordinary time periods. Since the probability of field-induced transitions to higher bands is usually small, most electrons will in fact occupy such states. From the familiar result that a full band carries no current it follows that not all restricted band states can have the same expectation value of velocity. However, this conclusion does not apply to the true stationary states.

Before beginning a direct calculation of the velocity expectation value we take cognizance of an exact solution previously obtained for a special case. The problem of charged particles in crossed electric and magnetic fields in free space has already been treated quantum mechanically by Titeica,[5] whose work was concerned with the magnetoresistance of metals. He found that stationary states exist having the form

$$\psi(x,y,z)=f_n(x-x_0)e^{(i/\hbar)m(v_0-\omega x_0)y}e^{ikz x}, \tag{58}$$

in which $f_n(x)$ is the harmonic oscillator wave function corresponding to the angular frequency $\omega=eH/mc$ and quantum number n. The average position of the electron, x_0, is arbitrary. It is easily seen by direct application of the velocity operator that the average velocity is v_0.

In this case we see that each wave function belongs to a class, the members of which differ only by relative translations along the x-axis and a corresponding phase factor. Such classes exist in general, a fact which enables us to deduce the expectation value of velocity for stationary states when the electron is contained in a crystal, as shown below.

The Hamiltonian for our problem is

$$H=\frac{1}{2m}\left[p_x^2+\left(p_y-\frac{eH}{c}x\right)^2 p_z^2\right]+eEx+V_p(x,y,z), \tag{59}$$

where V_p is the periodic potential. We have chosen the gauge which gives to the Hamiltonian the maximum translational symmetry in the y-direction. We note in passing that if the y- and z-directions are collinear with lattice displacement vectors the solution of Schrödinger's equation can have a form very similar to (58):

$$\psi(x,y,z)=\phi_n(x,y,z)e^{iky y}e^{ikz z}, \tag{60}$$

where ϕ_n is periodic in the y- and z-directions.

It will now be shown that, given any solution of Eq. (59) with expectation value of the x-coordinate $\langle x\rangle=x_0$ and energy W_0, infinitely many solutions exist of identical form except for a phase factor, having each a value of $\langle x\rangle$ and a related energy $W_0+\delta W$ such that

$$eE(\langle x\rangle-x_0)=\delta W. \tag{61}$$

Let $\mathbf{a}$ be a general lattice displacement vector. Then, simultaneously with the coordinate transformation

$$\mathbf{r}=\mathbf{r}'+\mathbf{a}, \tag{62}$$

let us introduce the gauge transformation

$$A_y=A_y'-(eH/c)a_x. \tag{63}$$

In the new coordinate system and gauge the original Hamiltonian is recovered except for the additive term eEa_x. Thus if $\psi(\mathbf{r})$ is a solution in the old gauge, $\psi(\mathbf{r}')$ is a solution in the new gauge.

It follows from gauge transformation theory and may easily be verified directly, that the new wave functions in the original coordinate system and gauge have the form

$$\psi(\mathbf{r})=\psi(\mathbf{r}-\mathbf{a})e^{(i/\hbar)\delta P_y y}, \tag{64}$$

where $\delta P_y=eHa_x/c$.

Let us suppose that a wave packet has been constructed by superposition of states of the form (64). Then the group velocity of this packet is

$$v_g=\delta W/\delta P_y=v_0, \tag{65}$$

which is consequently the expectation value of velocity for the component states.

This result has been shown to apply to all true stationary states. We will show below that it applies also to all quasi-stationary restricted band states satisfying a certain condition. This condition is that the restricted band states in question can be generated from stationary states in the magnetic field alone by regarding the electric field as a perturbation. It is clear that the unperturbed states must be localized with respect to the x-coordinate and must have essentially uniform probability density in the y-direction in order that the electric field constitute a perturbation; but it is probable that in general a complete set of states in the magnetic field can be found having this property. The unperturbed states are automatically restricted band states. Consequently, the perturbed states, if they exist, are also restricted band states.

The expectation value of velocity for the perturbed states will now be calculated directly. Letting the symbol δ indicate the change in the conjoined quantity induced by the perturbation, we have

$$\delta\langle v_y\rangle=\frac{1}{m}\left(\delta\langle p_y\rangle-\frac{eH}{c}\delta\langle x\rangle\right). \tag{66}$$

But to first order in electric field we have

$$\delta\langle p_y\rangle=\sum_n{}'\frac{H_{0n}'p_{yn0}+p_{yn0}H_{n0}'}{W_0-W_n}, \tag{67}$$

[5] S. Titeica, Ann. Physik **22**, 129 (1935).

and

$$\delta\langle x\rangle=\sum_n{}' \frac{H_{0n}''x_{n0}+x_{0n}H_{n0}''}{W_0-W_n}, \tag{68}$$

where W_n is the energy of the nth unperturbed state and $H_{0n}'=H_{0n}''$ are the matrix elements of the perturbation. The sum is taken over all values of n except $n=0$, this value indicating the initial unperturbed state, $\psi_0(x,y,z)$.

The perturbation H' is eEx. However, on transforming to a new coordinate system defined by

$$x=x'-mc^2E/eH^2, \tag{69a}$$

$$\mathbf{r}=\mathbf{r}'+\mathbf{a}, \tag{69b}$$

where $\mathbf{a}$ is a lattice displacement vector, the term eEx drops out of the Hamiltonian (59), and is replaced by the term v_0p_y. In this new coordinate system one can therefore treat v_0p_y as the perturbation. Out of all the states having the same form except for a phase factor and a displacement let us choose as the unperturbed state in this new coordinate system $\psi_0(x',y',z')$. Then the same perturbed state will be obtained in both cases. This is easily seen if ψ_0 has the form (60). For neither perturbation changes this form nor the value of k_y. But for any given orientation of the lattice relative to the magnetic field we may choose another orientation arbitrarily close to the first such that the wave functions may take the form (60) in the new orientation. [It should also be noted that a succeeding infinitesimal rotation about the y-axis insures that the discrete set of values of E allowed by Eqs. (69) is everywhere dense. The restriction to a discrete set is therefore not significant.]

Now if we compute $\delta\langle p_y\rangle$ in the original coordinate system ($H'=eEx$) and $\delta\langle x\rangle$ in the second coordinate system ($H''=v_0p_y$), the resultant expressions when substituted into Eq. (66) cancel. However, we must correct $\delta\langle x\rangle$ by the amount of the coordinate shift $-mc^2E/eH^2$, and this gives us finally the desired result

$$\delta\langle v_y\rangle=v_0. \tag{70}$$

Relation to Orbits in Momentum Space

Let us consider the classical ensemble of particles satisfying the equations of motion (10) and (11), all occupying a given orbit in crystal momentum space and all having the same average value of x-position, but having random y- and z-coordinates. The probability of finding a particle at a given location is stationary for such an ensemble. We shall say that a quasi-stationary restricted band state "corresponds" to an orbit in momentum space if the expectation values of energy and of position are the same for the wave function as for an ensemble of particles in the given orbit. This notion of correspondence allows us to connect the results of the present section with those of Part II.

Not every orbit will possess a corresponding wave function. The localization of the electron by the magnetic field, which leads to the displacement property expressed by Eq. (64), leads in an intimately related way also to the discrete quantization of the energy of an electron having a given expectation value of position. Thus only orbits belonging to a discrete set possess corresponding wave functions.

This quantum effect, regarded from the viewpoint of the crystal momentum approximation, may be called "orbital quantization." The orbital motion of the electron in space results in quantum mechanical interferences which act to select those particular "orbits" for which the interferences are constructive. A first approximation to the resultant energy levels may be found by regarding P_x and x as conjugate oscillatory variables to which the Bohr-Wilson-Sommerfeld phase integral condition may be applied. Through use of Eq. (15) this condition is easily transformed to

$$-\oint P_y dP_x=\oint P_x dP_y=\frac{ch}{eH}(n+\tfrac{1}{2}). \tag{71}$$

The integrals are equal to the area contained by the orbit. It is interesting to note that this condition yields the correct energy levels for the free particle and for the "pseudoparticle" with different longitudinal and transverse masses.

In cases in which bands overlap, the unperturbed wave functions, as stated earlier, must be assumed to involve superpositions of states from each of the degenerate bands. If there should be degeneracy among the unperturbed states, the correct zeroth-order wave functions are to be selected for use in the perturbation calculation. The resultant quasi-stationary states correspond simultaneously to several orbits of equal energy in different bands.

It is reasonable to suppose that as the electric field approaches zero, perturbation theory should become applicable to almost all quasi-stationary states. However, we should like to know to what states in particular the theory is applicable at given values of field strength. We shall not attempt to answer this question rigorously, but instead shall make use of a plausible argument suggested by the correspondence developed here. It is natural to suppose that if the motion of the electron in a magnetic field, as it is described by means of the crystal momentum approximation, is merely perturbed by the electric field, the actual motion is correspondingly perturbed by the electric field, and not altered essentially. But the reentrant orbits are perturbed constant energy contours, while the nonreentrant orbits are altered in essential form. Thus we should expect reentrant orbits to correspond to perturbed states, and nonreentrant orbits to correspond to those "extraordinary" restricted band states which cannot be generated by means of perturbation theory from states stationary in the magnetic field alone.

Such extraordinary states, since they do not occur near the top or bottom of a band, are not likely to be of interest in connection with the semiconductor Hall effect. However, they do play an interesting role in the one-electron theory of the full band. By setting the expectation value of acceleration equal to zero for a quasi-stationary state, it is easily seen that the velocity expectation value is v_0 if and only if the average force exerted by the lattice on the electron is zero; the corresponding statement in the crystal momentum approximation is contained in Eq. (20). In a full band the electrons must, however, exert a force on the lattice ions equal and opposite to the direct force due to the electric field. This force is exerted by electrons in the extraordinary states. Correspondingly, the condition that formula (24) hold is that the relevant electrons and holes exert no average force on the lattice.

In conclusion we point out that at sufficiently high magnetic fields, given constant longitudinal current, a significant number of electrons will make transitions to higher bands. This phenomenon, which is essentially Zener breakdown due to the transverse electric field, could in principle be observed, since it would not be preceded by avalanche breakdown.

APPENDIX. INTEGRAL SOLUTIONS OF THE BOLTZMANN EQUATION

If a mean free time exists, the solution of Eq. (25) can be written in integral form for arbitrary values of H. We have used a series expansion in the preceding work since it simplifies the discussion of saturation effects. The integral solution is given here for the sake of completeness.

The basic idea of this solution is due to Shockley,[4] though his treatment is somewhat generalized here. Let the magnetic field, $\mathbf{H}$, have an arbitrary magnitude, and an arbitrary orientation with respect to the longitudinal current density. The electric field, $\mathbf{E}$, will have a component $\mathbf{E}_1$ parallel to $\mathbf{H}$, and a component $\mathbf{E}_2$ perpendicular to $\mathbf{H}$. The set of orbits in momentum space of which we shall speak are those defined by $\mathbf{E}_2$ and $\mathbf{H}$. Generalizing notation employed by Shockley in treating constant energy contours, we define an angle variable θ for each orbit such that $\theta=2\pi t/T$, where t is the time corresponding to a point of the orbit, and T is the period of one revolution. We further define

$$\omega=2\pi/T,$$
$$g=f-f_0.$$

Then Eq. (25) can be written in the form:

$$\frac{dg}{d\theta}+\frac{g}{\omega\tau}=\frac{\partial f_0}{\partial W}\frac{e}{\omega}\mathbf{E}\cdot\mathbf{v},$$

in which a term $e\mathbf{E}_1\cdot\nabla_P g$ has been dropped, since it is small compared to $e\mathbf{E}_1\cdot\nabla_P f_0$, a quantity included in the right member. This linear equation must be solved subject to the boundary condition $g(\theta)=g(\theta+2\pi)$. Such a solution is given by the expression

$$g(\theta)=A\int_{\theta-2\pi}^{\theta} f_0(1-f_0)\mathbf{E}\cdot\mathbf{v}(\theta')\exp\left[-\int_{\theta'}^{\theta}d\theta''/\omega\tau\right]d\theta',$$

where

$$A=-e(kT\omega)^{-1}\left[1-\exp\left(-\oint d\theta''/\omega\tau\right)\right]^{-1},$$

all integrals being taken along the orbit. The equivalent quantity $-f_0(1-f_0)/kT$ has been substituted for $\partial f_0/\partial W$. Shockley obtained a solution correct to first order in the transverse electric field by integrating instead along constant energy contours. This approximation should be adequate for normal longitudinal drift velocities. In this case θ is the angle for zero electric field, and the factor $f_0(1-f_0)$ may be placed outside the integral sign. One may integrate $g(\theta)v$ around the contour in order to obtain the contribution to the current of states along the contour. The resultant multiple integral, though of somewhat different form, is equivalent to Shockley's integral.

By expanding the exponentials in power series one effectively achieves an expansion in terms of $1/H$. However, the simple results regarding saturation quantities obtained by the direct power series expansion in conjunction with a discussion of the orbits are not apparent from these integrals.