PHYSICAL REVIEW VOLUME 109, NUMBER 4 FEBRUARY 15, 1958

Electron-Hole Recombination Statistics in Semiconductors through Flaws with Many Charge Conditions

CHIH-TANG SAH AND W. SHOCKLEY
Shockley Semiconductor Laboratory, Mountain View, California

(Received July 23, 1957)

A flaw with s electronic units of negative charge makes transitions to charge $s+1$ by hole emission at rate $e(p,s)$ or by electron capture at rate $nc(n,s)$ and returns to charge s at rates $e(n, s+1)$ and $pc(p, s+1)$. Here n is the electron density in the conduction band and p is the hole density in the valence band. The steady-state ratio of populations N_{s+1} to N_s is given by

$$c(n,s)[n+n^*(s+\tfrac{1}{2})]/c(p, s+1)[p+p^*(s+\tfrac{1}{2})],$$

where $n^*(s+\frac{1}{2})=e(p,s)/c(n,s)$ and $p^*(s\frac{1}{2})=e(n, s+1)/c(p, s+1)$. This distribution corresponds to an effective Fermi level for the flaws only for the condition of thermal equilibrium. Expressions for the recombination rate based on the steady-state distribution are derived. For a given transition $s\rightleftarrows s+1$ the following special cases are defined: (1) denuded: $n<n^*$, $p<p^*$; (2) n-dominated: $n>n^*$, $p<p^*$; (3) p-dominated: $n<n^*$, $p>p^*$; (4) flooded: $n>n^*$, $p>p^*$. Diagrams which aid in visualizing the relative importance of the various transitions are presented. Some speculations on the nature of trapping centers are given.

1. INTRODUCTION

THE nonequilibrium recombination statistics for holes and electrons through single-level flaws have been considered in detail by Shockley and Read[1] (hereafter referred to as S-R), Hall,[2] and others.[3] Recently, experimental results[4] have shown that gold, copper, selenium, tellurium, and many other impurities not in the third or the fifth column of the periodic table, have multiple energy levels in the energy gap of germanium and silicon, and their presence influences greatly the recombination and generation of the carriers. The equilibrium statistics for an impurity center or a flaw with multiple energy levels or charge conditions have also been considered recently by Shockley and Last (hereafter referred to as S-L) and others.[5] However, a general and correct nonequilibrium statistics for a multiple-charge-condition flaw in semiconductors has not appeared in the literature.[6]

This paper is concerned with the nonequilbrium but steady-state statistics for holes and electrons in semiconductors with one type of multiple-charge-condition flaws. The complexity of this problem is simplified by the use of two diagrams: the charge-distribution diagram (the R_N diagram) and the recombination-rate diagram (the R_U diagram). Through these diagrams, the most populated charge states of the flaws and most important charge states of the flaws which provide the major part of the carrier recombination and generation traffic can be visualized easily under any injection conditions or carrier densities n and p.

2. DEFINITIONS AND DERIVATION OF EQUILIBRIUM CONDITIONS

The treatment required is essentially similar to that given by Shockley and Read for a single-level center, the essential difference being that sets of values replace single values. For example, in place of a single capture cross section $\langle c_n\rangle$ for an electron in the conduction band (as introduced in S-R), there are a set

$$c(n, -r),\ c(n, -r+1),\ \cdots,\ c(n,s),\ \cdots,\ c(n,t), \quad (2.1)$$

where $-r$ represents the most positive charge condition considered for the center, corresponding to net charge $+r$ times the absolute value of the electronic charge, and $+t$ is similarly the most negative condition. Each electron-capture cross section involves a transition of plus unity increase of s in the sequence and is associated with an electron-emission probability, denoted by $e(n, s+1)$. We shall use the following notation (see Table I for definitions):

Rate of electron capture by flaws initially in charge condition s,

$$s\rightarrow s+1 \qquad nN_sc(n,s). \quad (2.2)$$

Rate of electron emission by flaws initially in charge condition $s+1$,

$$s+1\rightarrow s \qquad N_{s+1}e(n, s+1). \quad (2.3)$$

---

[1] W. Shockley and W. T. Read, Jr., Phys. Rev. **87**, 853 (1953).

[2] R. N. Hall, Phys. Rev. **83**, 288 (1951); **87**, 387 (1952).

[3] P. T. Landsberg, Proc. Phys. Soc. (London) **A65**, 604 (1952); P. T. Landsberg and T. S. Moss, Proc. Phys. Soc. (London) **B69**, 661 (1956); S. G. Kalashnikov, J. Tech. Phys. USSR **26**, 241 (1956) [translation: Soviet Phys. (Tech. Phys.) **1**, 237 (1957)]. H. Y. Fan, Phys. Rev. **92**, 1424 (1953); J. Nishizawa and Y. Watanabe, Sci. Repts. Research Insts., Tôhoku Univ. **7**, 149 (1955).

[4] C. B. Collins *et al.* Phys. Rev. **105**, 1168 (1957), and references therein; H. H. Woodbury and W. W. Tyler, Phys. Rev. **105**, 84 (1957), and references therein; J. A. Burton, Physica **20**, 845 (1954); J. A. Armstrong *et al.*, Bull. Am. Phys. Soc. Ser. II, **2**, 265 (1957).

[5] W. Shockley and J. T. Last, Phys. Rev. **107**, 392 (1957); P. T. Landsberg, Proc. Phys. Soc. (London) **B69**, 1056 (1956); B. H. Champness, Proc. Phys. Soc. (London) **B69**, 1335 (1956); H. Brooks in *Advances in Electronics and Electron Physics*, edited by L. Marton (Academic Press, Inc., New York, 1955), Vol. 1, p. 127.

[6] T. Landsberg, Proc. Phys. Soc. **B70**, 282 (1957). This author considered the case of very small deviation from the thermal equilibrium condition so that a quasi-Fermi level (or steady-state Fermi level) can be used to describe the charge distribution of the flaws.

TABLE I. List of symbols.

$c(n,s)$	average capture probability of electron by flaw in charge condition s
$c(p,s)$	average capture probability of hole by flaw in charge condition s
$e(n,s)$	average emission probability of electron by flaw in charge condition s
$e(p,s)$	average emission probability of hole by flaw in charge condition s
n	electron density in the conduction band
n_i	equilibrium electron or hole density in an intrinsic specimen
$n(s+\frac{1}{2})$	density of electron in the conduction band when the Fermi level falls at $E(s+\frac{1}{2})$
$n^*(s+\frac{1}{2})$	density of electron in the conduction band when the Fermi level falls at $E^*(s+\frac{1}{2})$
$n_m(s+\frac{1}{2})$	see Eq. (7.2a)
p	hole density in the valence band
$p(s+\frac{1}{2})$	density of hole in the valence band when the Fermi level falls at $E(s+\frac{1}{2})$
$p^*(s+\frac{1}{2})$	density of hole in the valence band when the Fermi level falls at $E^*(s+\frac{1}{2})$
s	index for the sth charge state of the flaw with s electronic charges
E_i	intrinsic Fermi level
$E(s+\frac{1}{2})$	energy associated with the transitions of flaws between charge conditions s and $s+1$
$E^*(s+\frac{1}{2})$	see Eq. (3.9)
$E_m(s+\frac{1}{2})$	see Eq. (7.2)
E_{nd}	the nth donor energy level defined by Shockley and Last, or $E(-n+\frac{1}{2})$
E_{na}	the nth acceptor energy level defined by Shockley and Last, or $E(n-\frac{1}{2})$
F	Fermi level
F_p	quasi-Fermi level for holes
F_n	quasi-Fermi level for electrons
F_s	Fermi level at the surface
$M(s+\frac{1}{2})$	density of flaw with energy $E(s+\frac{1}{2})$
N_s	density of flaw in charge state s
$R_N(s+\frac{1}{2}, p,n)$	ratio of flaw densities in the two adjacent charge conditions s and $s+1$
$R_U(s)$	ratio of the steady-state recombination rates involving energies $E(s+\frac{1}{2})$ and $E(s-\frac{1}{2})$
$S(s+\frac{1}{2}, F_s)$	surface recombination velocity due to transitions $s\rightleftarrows s+1$
U_{2a}	an activation energy associated with the second acceptor trapping level
$U(n, s+\frac{1}{2})$	net electron capture rate for transition of flaw from charge s to $s+1$
$U(p, s+\frac{1}{2})$	net hole capture rate for transition of flaw from charge condition $s+1$ to s
$U(s+\frac{1}{2})$	the net steady-state electron or hole capture rate
$\delta n, \delta p$	added electron or hole density
$\tau(n,s)$	average electron lifetime for unit flaw density $=1/c(n,s)$
$\tau(p,s)$	average hole lifetime for unit flaw density $=1/c(p,s)$

Similarly, the rates for transitions involving holes are written as

$$s\rightarrow s+1 \qquad N_s e(p,s), \tag{2.4}$$

$$s+1\rightarrow s \qquad pN_{s+1}c(p, s+1). \tag{2.5}$$

The assumptions on the basis of which rate expressions like (2.1) to (2.5) may be expected to be valid are:

(1) The electron and hole gases are nondegenerate and are in thermal equilibrium [see S-R, Eq. (3.1)].

(2) The relaxation time of a flaw, after it has undergone a change in charge, is short compared to the average time between changes of charge.

(3) Processes involving three imperfections are unimportant, such as the capture of one electron being enchanced by the presence of another electron or hole to carry off energy (Auger process).

From the foregoing it is evident that transitions between charge conditions s and $s+1$ are characterized by the four emission-capture constants: $c(n,s)$, $e(n,s+1)$, $c(p, s+1)$, $e(p,s)$. As we shall show below, there is one relationship between these four constants and the intrinsic carrier density n_i so that three constants suffice to describe the transition. The exact choice of which three are most convenient depends upon the application being made of the model. Certain derived quantities based on the four emission-capture constants and n_i prove to be particularly convenient for some applications. We shall give the equations for these derived quantities in this section.

Net electron capture ($s\rightarrow s+1$):

$$U(n, s+\tfrac{1}{2})=nN_s c(n,s)-N_{s+1}e(n, s+1), \tag{2.6}$$

Net hole capture ($s+1\rightarrow s$):

$$U(p, s+\tfrac{1}{2})=pN_{s+1}c(p, s+1)-N_s e(p,s). \tag{2.7}$$

The principle of detailed balance requires that $U(n, s+\frac{1}{2})$ and $U(p, s+\frac{1}{2})$ each vanish under thermal-equilibrium conditions so that

$$N_{s+1}/N_s=nc(n,s)/e(n, s+1), \tag{2.8}$$

$$N_{s+1}/N_s=e(p,s)/pc(p, s+1). \tag{2.9}$$

The equality of the two right-hand terms of (2.8) and (2.9) leads to the relationship mentioned above between the emission and capture constants:

$$n_i^2 c(n,s)c(p, s+1)=e(n, s+1)e(p,s), \tag{2.10}$$

where the mass-action law $np=n_i^2$ has been used.

The first derived constant we shall introduce is an energy level $E(s+\frac{1}{2})$ which is identical with the energy levels introduced in describing the equilibrium statistics of flaws in the Shockley-Last paper.[5] The factor $\frac{1}{2}$ in the index for the energy level is used in order to emphasize that the energy is associated with the transitions between charge conditions s and $s+1$. The equation for $E(s+\frac{1}{2})$ is

$$\begin{aligned} E(s+\tfrac{1}{2})&\equiv E_i+kT\ln[e(n, s+1)/n_i c(n,s)] \\ &=E_i+kT\ln[n_i c(p, s+1)/e(p,s)], \end{aligned} \tag{2.11}$$

where E_i is the intrinsic energy level such that when the Fermi level F equals E_i then n and p are each equal to n_i.

The notation involving s, $s+1$ and $s+\frac{1}{2}$ forms a consistent scheme in the following sense: A constant describing a transition is identified by the *initial charge condition.* In each such case the final charge condition is obvious; for example, $c(n, s+1)$ clearly concerns a transition from $s+1$ to $s+2$. A quantity associated with the transition between two conditions but *not specifically with either direction of transition* is represented by the average value of the charge; i.e., $E(s+\frac{1}{2})$ corresponds to average of s and $s+1$.

In terms of $E(s+\frac{1}{2})$ two carrier densities are introduced in analogy with n_1 and p_1 of S-R. These are the densities obtained by setting F equal to $E(s+\frac{1}{2})$ in the expression for electron density under equilibrium conditions. When these densities are present, equal numbers of flaws are in conditions s and $s+1$.

$$\begin{aligned} n(s+\tfrac{1}{2}) &= n_i \exp[(F-E_i)/kT] \\ &= n_i \exp\{[E(s+\tfrac{1}{2})-E_i]/kT\} \\ &= e(n,\, s+1)/c(n,s), \end{aligned} \tag{2.12}$$

$$\begin{aligned} p(s+\tfrac{1}{2}) &= n_i \exp[(E_i-F)/kT] \\ &= n_i \exp\{[E_i-E(s+\tfrac{1}{2})]/kT\} \\ &= e(p,s)/c(p,\, s+1). \end{aligned} \tag{2.13}$$

In terms of $E(s+\frac{1}{2})$, $n(s+\frac{1}{2})$ and $p(s+\frac{1}{2})$ one readily obtains from (2.11) and (2.12):

$$\begin{aligned} N_{s+1}/N_s &= n/n(s+\tfrac{1}{2}) = p(s+\tfrac{1}{2})/p \\ &= \exp\{[F-E(s+\tfrac{1}{2})]/kT\}. \end{aligned} \tag{2.14}$$

This equation is identical with that defined in S-L for equilibrium statistics and thus shows that $E(s+\frac{1}{2})$ is identical with the energy level defined in terms of state-sums in that reference.[5]

For completeness in this section we shall introduce three other derived constants:

$$n^*(s+\tfrac{1}{2}) = e(p,s)/c(n,s), \tag{2.15}$$

$$p^*(s+\tfrac{1}{2}) = e(n,\, s+1)/c(p,\, s+1), \tag{2.16}$$

$$\begin{aligned} E^*(s+\tfrac{1}{2}) &= 2E_i - E(s+\tfrac{1}{2}) \\ &\quad - kT \ln[c(n,s)/c(p,\, s+1)]. \end{aligned} \tag{2.17}$$

The two *starred densities* correspond to equilibrium with F lying at $E^*(s+\frac{1}{2})$. The usefulness of these quantities will be made clear in the next section.

As discussed above, a single transiton is described by three independent parameters which may be chosen in a wide variety of ways from the larger set of ten quantities introduced so far:

$$\begin{aligned} &E(s+\tfrac{1}{2}),\, E^*(s+\tfrac{1}{2}),\, n(s+\tfrac{1}{2}),\, p(s+\tfrac{1}{2}),\, n^*(s+\tfrac{1}{2}),\, p^*(s+\tfrac{1}{2}); \\ &e(n,\, s+1),\, c(n,s),\, e(p,s),\, c(p,\, s+1). \end{aligned} \tag{2.18}$$

The first set of six quantities are related to ratios (homogeneous functions of order zero) of the capture and emission constants. In the next section we shall find that the steady-state distribution is determined by a two-parameter system chosen from the first set of six quantities. One additional constant, proportional to the rate and chosen from the second set of four constants, must be introduced to calculate the rate of recombination or generation.

3. STEADY-STATE CONDITIONS

In a nonequilibrium steady-state situation, in which net recombination or net generation of carriers occurs, the distribution

$$N_{-r}, \cdots, N_s, \cdots, N_{t-1}, N_t, \tag{3.1}$$

does not vary in time. Hence the net rate of transitions to the condition t must vanish so that the effect of $c(n,\, t-1)$ and $e(p,\, t-1)$ must be exactly balanced by transitions back due to $e(n,t)$ and $c(p,t)$. Similarly, since there is no net departure of flaws from condition $t-1$ to condition t, there can be no net arrival of flaws to condition $t-1$ from condition $t-2$. Thus, in general there must be balance between transitions in the two directions between any two conditions s and $s+1$.

This steady-state balance requirement is equivalent to the principle of conservation of electrical charge since it says that the flaw does not accumulate any net charge in the transitions between s and $s+1$.

The net rate of recombination may be calculated either from the net rate of electron capture or the net rate of hole capture. That the equality of these two net rates for any transition $s \rightarrow s+1$ and $s+1 \rightarrow s$ is, in fact, a consequence of the steady-state condition just described can be seen as follows:

$$\begin{aligned} O &= (\text{over-all net rate of transitions } s \rightarrow s+1) \\ &= (\text{net excess of electron capture over electron emission}) \\ &\quad - (\text{net excess of hole capture over hole emission}). \end{aligned} \tag{3.2}$$

Thus, the recombination rate due to transitions between s and $s+1$ may be written in keeping with (2.6) and (2.7) as

$$\begin{aligned} U(s+\tfrac{1}{2}) &= N_s n c(n,s) - N_{s+1} e(n,\, s+1) \\ &= N_{s+1} p c(p,\, s+1) - N_s e(p,s). \end{aligned} \tag{3.3}$$

In order to evaluate $U(s+\frac{1}{2})$, which is a function of the steady-state nonequilibrium densities n and p, the values of N_s and N_{s+1} must be determined.

We shall determine first the ratio N_{s+1}/N_s by equating the total rates of transition $s \rightarrow s+1$ to that of $s+1 \rightarrow s$. The total transition rate from s to $s+1$, resulting from electron capture and hole emission for the N_s flaws in charge condition s may be written as

$$\begin{aligned} \text{Total } s \rightarrow s+1\colon\ & N_s[nc(n,s)+e(p,s)] \\ &= N_s c(n,s)[n+n^*(s+\tfrac{1}{2})], \end{aligned} \tag{3.4}$$

where the density $n^*(s+\frac{1}{2})$ is given by (2.15) which we repeat here for convenience

$$n^*(s+\tfrac{1}{2}) = e(p,s)/c(n,s). \tag{3.5}$$

Thus $n^*(s+\frac{1}{2})$ is the density at which electron capture produces a contribution to $s \rightarrow s+1$ transitions just equal to that of hole emission. It is evident that electrons in the conduction band have a relatively important effect on transitions $s \rightarrow s+1$ only if n is greater $n^*(s+\frac{1}{2})$.

Similarly, the total rate of transitions from $s+1$ to s may be written as

$$\text{Total } s+1 \rightarrow s = N_{s+1} c(p,\, s+1)[p+p^*(s+\tfrac{1}{2})], \tag{3.6}$$

where $p^*(s+\frac{1}{2})$ is given by (2.16) which we also repeat for convenience:

$$p^*(s+\tfrac{1}{2})=e(n,s+1)/c(p,s+1). \tag{3.7}$$

The density $p^*(s+\frac{1}{2})$ makes hole capture equal electron emission for $s+1 \rightarrow s$ transitions.

From the relationship (2.10) between the four emission-capture constants, it follows that

$$n^*(s+\tfrac{1}{2})p^*(s+\tfrac{1}{2})=n_i^2, \tag{3.8}$$

so the starred densities correspond to an equilibrium situation.

Since $n^*(s+\frac{1}{2})$ and $p^*(s+\frac{1}{2})$ satisfy the mass-action law for np, they define an energy level. This level is

$$\begin{aligned} E^*(s+\tfrac{1}{2})&=2E_i-E(s+\tfrac{1}{2})-kT\ln[c(n,s)/c(p,s+1)].\\ &=E(s+\tfrac{1}{2})-kT\ln[e(n,s+1)/e(p,s)]. \end{aligned} \tag{3.9}$$

If the two capture cross sections are equal, an especially simple condition prevails: the energy level $E^*(s+\frac{1}{2})$ is obtained by reflecting $E(s+\frac{1}{2})$ through E_i. If the emission constants are equal, $E^*(s+\frac{1}{2})$ and $E(s+\frac{1}{2})$ are equal. The relationship of the two E's is further discussed in Sec. 6.

That a pair of starred densities corresponding to thermal equilibrium must, in general, exist may be seen by the following argument. If we imagine the Fermi level to be gradually raised from the valence band to the conduction band, the rate of electron capture for a given transition $s \rightarrow s+1$ will continually increase compared to hole emission producing the same transition. When $n=n^*(s+\frac{1}{2})$ the two rates become equal. By the principle of detailed balance, each of these rates is equal respectively to the corresponding reverse process and, therefore, the two reverse rates are also equal. The reverse rates are, of course, $s+1 \rightarrow s$ by electron emission and $s+1 \rightarrow s$ by hole capture. The value $p^*(s+\frac{1}{2})$, which makes these rates equal, must therefore satisfy the mass action law when multiplied by $n^*(s+\frac{1}{2})$. Thus when $F=E^*(s+\frac{1}{2})$ all four of the transition processes proceed at identical rates.

The physical argument of the preceding paragraph also leads at once to the conclusion that when $n=n^*(s+\frac{1}{2})$ and $p=p^*(s+\frac{1}{2})$ the ratio N_{s+1}/N_s must be the inverse of the emission ratio, i.e., the ratio must be $e(n,s+1)/e(p,s)$ so as to have the two emission rates balance. This ratio may also be written as $c(n,s)n^*(s+\frac{1}{2})/c(p,s+1)p^*(s+\frac{1}{2})$. It is helpful to have this relationship in mind in considering the meaning of the complete expression for the steady-state ratio of N_{s+1}/N_s.

Equating the two total rates of transition (3.4) and (3.6) leads to the ratio, denoted by $R_N(s+\frac{1}{2},n,p)$, for N_{s+1}/N_s:

$$\begin{aligned} R_N(s+\tfrac{1}{2},p,n)\equiv N_{s+1}/N_s=c(n,s)[n+n^*(s+\tfrac{1}{2})]/&\\ c(p,s+1)[p+p^*(s+\tfrac{1}{2})].& \end{aligned} \tag{3.10}$$

For a given nonequilibrium situation, this ratio may be considered a known function of n and p, the nonequilibrium carrier densities. As discussed in connection with (2.20), the steady-state distribution function (3.10) has two parameters dependent on the flaw; these two may be taken as $c(n,s)/c(p,s+1)$ and $n^*(s+\frac{1}{2})$ with $p^*(s+\frac{1}{2})$ given by $n_i^2/n^*(s+\frac{1}{2})$. In Sec. 4 [see (4.21)] the ratio of capture constants is expressed in terms of $n(s+\frac{1}{2})$ and $n^*(s+\frac{1}{2})$ and these latter two are used as the two parameters.

The distribution represented by the R_N values cannot be represented by an effective Fermi level for the flaws. This conclusion is reached by noting that if an effective Fermi level F_f does exist, then the ratio of any two successive R_N's is independent of F_f and is given according to (2.14) by

$$\begin{aligned} &R_N(s+\tfrac{1}{2},p,n)/R_N(s-\tfrac{1}{2},p,n)\\ &\qquad=\exp\{[E(s-\tfrac{1}{2})-E(s+\tfrac{1}{2})]/kT\}. \end{aligned} \tag{3.11}$$

If under nonequilibrium conditions

$$np=an_i^2, \tag{3.12}$$

then the ratio of the two R_N's given by (3.10) can be reduced to

$$\begin{aligned} &R_N(s+\tfrac{1}{2},p,n)/R_N(s-\tfrac{1}{2},p,n)\\ &\qquad=\frac{[n+n^*(s+\frac{1}{2})][an+n^*(s-\frac{1}{2})]}{[an+n^*(s+\frac{1}{2})][n+n^*(s-\frac{1}{2})]}\\ &\qquad\quad\times\exp\{[E(s-\tfrac{1}{2})-E(s+\tfrac{1}{2})]/kT\}. \end{aligned} \tag{3.13}$$

It is evident that the fraction multiplying the exponential will be independent of n only if either a equals unity (the equilibrium case) or $n^*(s+\frac{1}{2})$ and $n^*(s-\frac{1}{2})$ are equal (an accidental situation). Thus, it follows that no effective value F_f can in general be found unless equilibrium prevails so that a is unity.

If we assume that the values of the R_N factors are known and the total number of flaws is N_f, then the distribution N_s can be formally expressed as follows:

$$\begin{gathered} N_{-r+1}=R_N(-r+\tfrac{1}{2},n,p)N_{-r}\equiv f_{-r+1}N_{-r},\\ N_s=R_N(s-\tfrac{1}{2},n,p)N_{s-1}\equiv f_sN_{-r},\\ N_t=R_N(t-\tfrac{1}{2},n,p)N_{t-1}\equiv f_tN_{-r}. \end{gathered} \tag{3.14}$$

The factors f_s, defined above, are functions of n and p and are successive products of the form

$$\begin{aligned} f_s\equiv{}&R_N(s-\tfrac{1}{2},n,p)\\ &\times R_N(s-1-\tfrac{1}{2},n,p)\cdots R_N(-r+\tfrac{1}{2},n,p). \end{aligned} \tag{3.15}$$

In terms of the f_s, the value N_{-r} is

$$N_{-r}=N_f/(1+\sum f_s), \tag{3.16}$$

and

$$N_s=N_ff_s/(1+\sum f_s), \tag{3.17}$$

where the sum ranges from $-r+1$ up to t inclusive. (3.17) expresses N_s as a function of N_f, n, p with the capture and emission constants as parameters in the relationship. In the next section we shall consider graphically possible forms for the relationship.

For small disturbances from equilibrium, the R_N's have the values given in the S-L treatment of the equilibrium conditions. As shown in that treatment, under most conditions all of the flaws are substantially in one condition of charge. As we shall show below, under this circumstance, the flaws act as two sets of S-R recombination centers.

The net steady-state rate of electron or hole capture in the transitions between s and $s+1$ can be expressed in terms of known functions of n and p in various ways. From (3.3), (3.5), (3.7), (3.8), and (3.10), we find

$$\begin{aligned}U(s+\tfrac{1}{2})&\\&=N_s(pn-n_i^2)c(n,s)/[p+p^*(s+\tfrac{1}{2})]\\&=N_{s+1}(pn-n_i^2)c(p,s+1)/[n+n^*(s+\tfrac{1}{2})].\end{aligned}\tag{3.18}$$

This net rate can be written as

$$U(s+\tfrac{1}{2})=(N_s+N_{s+1})\times\frac{pn-n_i^2}{\tau(n,s)[p+p(s+\frac{1}{2})]+\tau(p,s+1)[n+n(s+\frac{1}{2})]},\tag{3.19}$$

where the mean life quantities τ are defined as

$$\tau(n,s)\equiv 1/c(n,s),\tag{3.20}$$

$$\tau(p,s+1)\equiv 1/c(p,s+1).\tag{3.21}$$

The interpretation of Eq. (3.19) is as follows: The fraction is simply the S-R formula for the net recombination for the case of unit density of flaws characterized by capture cross sections $c(n,s)$ for electrons and $c(p,s+1)$ for holes. The mean life of an electron in the heavily acceptor doped p-type material is $\tau(n,s)$ for unit density of flaws. The number of flaws which act like such centers is N_s+N_{s+1}. This reduces to N_f if the population of all conditions save N_s and N_{s+1} is negligible.

If one particular condition of the flaw predominates, so that

$$N_s\doteq N_f,\tag{3.22}$$

then the flaws act like one set of N_f centers of the S-R type with constant $c(n,s)$ etc., plus another set of N_f centers of the S-R type with constants $c(n,s-1)$ etc.

The relative importance of the net steady-state recombinations involving the two adjacent energy levels $E(s-\frac{1}{2})$ and $E(s+\frac{1}{2})$ can be studied by considering the ratio of the recombination rates, which can readily be obtained from Eq. (3.18):

$$R_U(s)=U(s+\tfrac{1}{2})/U(s-\tfrac{1}{2})=\frac{c(n,s)}{c(p,s)}\,\frac{n+n^*(s-\frac{1}{2})}{p+p^*(s+\frac{1}{2})}.\tag{3.23}$$

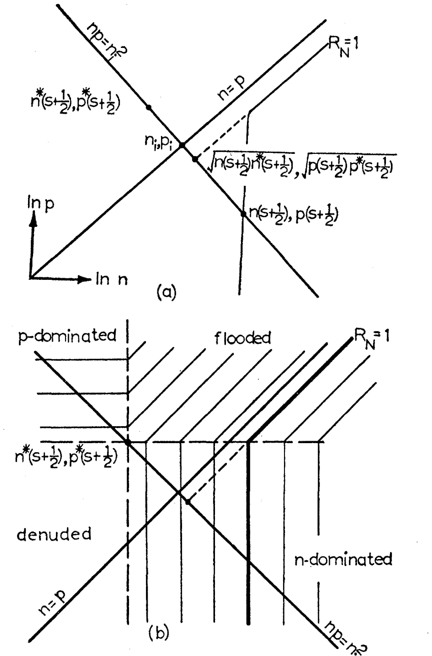

FIG. 1. Principle features of the R_N-diagram: (a) The line $R_N=1$; (b) the R_N=constant family.

Although this expression is formally similar to (3.10), it is a three-parameter rather than a two-parameter expression because $n^*(s-\frac{1}{2})$ and $p^*(s+\frac{1}{2})$ do not satisfy the mass-action law. We shall briefly consider two limiting cases:

(a) Denuded: $n=p=0$,

$$R_U(s)=e(p,s)/e(n,s),\tag{3.24}$$

(b) Flooded: $n^*(s-\frac{1}{2})\ll n=p\gg p^*(s+\frac{1}{2})$,

$$R_U(s)=c(n,s)/c(p,s).\tag{3.25}$$

For the denuded case, the relative importance is determined by the emission processes from flaws in the charge condition s. The flooded case, on the other hand, depends on the rate of entrance by capture by flaws in the charge condition s. A more detailed discussion is given in a subsequent section.

4. CHARGE DISTRIBUTION DIAGRAM, R_N DIAGRAM

In order to visualize how the distribution of the flaws over the possible conditions of charge depends upon n and p for equilibrium and nonequilibrium conditions, it is helpful to represent the dependence of R_N upon n and p on a diagram. Figure 1 shows a typical case, the axes being $\ln n$ and $\ln p$ with intrinsic point at the center of the diagram. In order to understand Fig. 1 in detail, we note that the ratio of capture constants in (3.10) for R_N may be re-expressed using (2.12), (2.13), (2.15), and (2.16) as follows:

$$\begin{aligned}c(n,s)/c(p,s+1)&\\&=[p(s+\tfrac{1}{2})p^*(s+\tfrac{1}{2})]^{\frac{1}{2}}/[n(s+\tfrac{1}{2})n^*(s+\tfrac{1}{2})]^{\frac{1}{2}}.\end{aligned}\tag{4.1}$$

The geometric mean densities $[n(s+\frac{1}{2})n^*(s+\frac{1}{2})]^{\frac{1}{2}}$ appear on Fig. 1 midway between $n(s+\frac{1}{2})$, $p(s+\frac{1}{2})$ and $n^*(s+\frac{1}{2})$, $p^*(s+\frac{1}{2})$.

In terms of these expressions $R_N(s+\frac{1}{2}, p,n)$ can be expressed in terms of the four parameters n, p, n^*, p^*:

$$R_N(s+\tfrac{1}{2}, p,n)=\left[\frac{p(s+\frac{1}{2})p^*(s+\frac{1}{2})}{n(s+\frac{1}{2})n^*(s+\frac{1}{2})}\right]^{\frac{1}{2}}\frac{n+n^*(s+\frac{1}{2})}{p+p^*(s+\frac{1}{2})}. \tag{4.2}$$

As discussed following (3.10), only two parameters are independent because of the mass-action law. In order to see that $R_N=1$ at the point $n(s+\frac{1}{2})$, $p(s+\frac{1}{2})$ it is helpful to note that

$$p(s+\tfrac{1}{2})+p^*(s+\tfrac{1}{2}) \\ =n_i^2[n(s+\tfrac{1}{2})+n^*(s+\tfrac{1}{2})]/n(s+\tfrac{1}{2})n^*(s+\tfrac{1}{2}). \tag{4.3}$$

The remaining features of Fig. 1 are largely self-explanatory. The rounding of the corners of lines of constant R_N, which occur over the regions where

$$|\ln[n/n^*(s+\tfrac{1}{2})]|<2 \text{ and } |\ln[p/p^*(s+\tfrac{1}{2})]|<2, \tag{4.4}$$

is not represented.

Features to be noted in Fig. 1(a) are the fact that the three points with index $(s+\frac{1}{2})$ lie on the thermal equilibrium line with the geometric mean point halfway between $n(s+\frac{1}{2})$, $p(s+\frac{1}{2})$ and $n^*(s+\frac{1}{2})$, $p^*(s+\frac{1}{2})$. Where n and p are large, the ratio R_N is proportional to n/p so that lines of constant R_N are parallel to $n=p$.

Figure 1(b) illustrates the general features of the lines of constant R_N. The lines are spaced at equal intervals of $\ln R_N$. It should be noted that the values of R_N on the equilibrium line are independent fo the value selected for $n^*(s+\frac{1}{2})$, $p^*(s+\frac{1}{2})$ and depend only on the distance from the $R_N=1$ point at $n(s+\frac{1}{2})$, $p(s+\frac{1}{2})$ determined by $E(s+\frac{1}{2})$ as discussed in connection with (2.14). If we imagine that $n^*(s+\frac{1}{2})$, $p^*(s+\frac{1}{2})$ slides along the $np=n_i^2$ line while $n(s+\frac{1}{2})$, $p(s+\frac{1}{2})$ remains fixed, then each $R_N=$constant line rotates through 90° as $n^*(s+\frac{1}{2})$, $p^*(s+\frac{1}{2})$ passes it but does not change its value.

In order to interpret certain portions of Fig. 1(b), we shall introduce the quasi-Fermi levels (q.f.l. or imrefs) F_n and F_p. In terms of these, the intrinsic density n_i and the intrinsic level E_i, the hole and electron densities become

$$n=n_i \exp[(F_n-E_i)/kT], \tag{4.5}$$

$$p=n_i \exp[(E_i-F_p)/kT], \tag{4.6}$$

where, for thermal equilibrium

$$F_n=F_p=F. \tag{4.7}$$

The characteristics of the four limiting regions bounded by dashed lines on Fig. 1(b) are as follows:

Denuded.—$n\ll n^*(s+\frac{1}{2})$, $p\ll p^*(s+\frac{1}{2})$. The steady state is governed by emission with

$$R_N(s+\tfrac{1}{2})=e(p,s)/e(n, s+1). \tag{4.8}$$

This ratio also holds exactly at all points on the line having the ratio $n:p$ equal to $n^*(s+\frac{1}{2}):p^*(s+\frac{1}{2})$. At the point $n^*(s+\frac{1}{2})$, $p^*(s+\frac{1}{2})$ electron capture equals hole emission and electron emission equals hole capture. [See paragraphs following Eq. (3.9).]

Flooded.—$n\gg n^*(s+\frac{1}{2})$, $p\gg p^*(s+\frac{1}{2})$. The steady state is governed by capture and R_N reduces to

$$R_N(s+\tfrac{1}{2})=c(n,s)n/c(p, s+1)p, \tag{4.9}$$

so that the two rates of capture are equal:

$$N_{s+1}c(p, s+1)p=N_s c(n,s)n. \tag{4.10}$$

This is satisfied by the above value of R_N.

n-dominated.—$n\gg n^*(s+\frac{1}{2})$, $p\ll p^*(s+\frac{1}{2})$. For this case, electron capture $nc(n,s)$ dominates hole emission $e(p,s)$ for the transitions $s\rightarrow s+1$ and the ratio of importance is $nc(n,s)/e(p,s)=n/n^*(s+\frac{1}{2})$. Similarly, electron emission dominates hole capture in the ratio $p^*(s+\frac{1}{2})/p$. Thus the steady state results from balance between the electron processes alone and R_N can be written as

$$\begin{aligned} R_N(s+\tfrac{1}{2})&=c(n,s)n/c(p, s+1)p^*(s+\tfrac{1}{2}) \\ &=c(n,s)n/e(n, s+1) \\ &=n/n(s+\tfrac{1}{2})=\exp\{[F_n-E(s+\tfrac{1}{2})]/kT\}. \end{aligned} \tag{4.11}$$

Comparison with Eq. (2.14) shows that the ratio of population N_{s+1}/N_s is that expected for equilibrium with the electron quasi-Fermi level. In other words, in the n-dominated region, the transition is in equilibrium with the electrons.

p-dominated.—$n\ll n^*(s+\frac{1}{2})$, $p\gg p^*(s+\frac{1}{2})$. This case is similar to the n-dominated case. The transition gets into equilibrium with the holes and

$$\begin{aligned} R_N(s+\tfrac{1}{2})&=p(s+\tfrac{1}{2})/p \\ &=\exp\{[F_p-E(s+\tfrac{1}{2})]/kT\}. \end{aligned} \tag{4.12}$$

Small-signal.—$n=n_e+\delta n$, $p=p_e+\delta p$, and δn, $\delta p\ll n_e$, p_e where the subscript e stands for equilibrium. When the electron and hole densities are very near the equilibrium values, the steady-state distribution approaches the thermal equilibrium distribution given by

$$R_N(s+\tfrac{1}{2})=\exp\{[F-E(s+\tfrac{1}{2})]/kT\}. \tag{4.13}$$

This situation corresponds to the region near the line $np=n_i^2$ in Fig. 1.

As shown in connection with Eq. (3.10), the R_N diagram for a given transition depends on two parameters. On diagrams like Fig. 1, these parameters may be chosen as $n(s+\frac{1}{2})$ and $n^*(s+\frac{1}{2})$ or the corresponding hole densities related to them by the mass-action law. Geometrically, it suffices to draw only the $R_N=1$ line; the intersection with $np=n_i^2$ then gives $n(s+\frac{1}{2})$ and the bend point gives respectively $p^*(s+\frac{1}{2})$ or $n^*(s+\frac{1}{2})$ if $n(s+\frac{1}{2})$ is, respectively, greater or less than $n^*(s+\frac{1}{2})$. This geometrical approach is presented in Fig. 2.

Figure 2 illustrates a set of possible $R_N=1$ lines for the case of a transition with an energy $E(s+\frac{1}{2})$ greater than E_i. Eight possible cases corresponding to different values of $n^*(s+\frac{1}{2})$ or alternatively to $c(n,s)/c(p,\ s+1)$ as related by (4.1) are shown. For case (a), the four limiting regions are shown. Similar diagrams for the other cases have been omitted to simplify the figure.

In terms of the ratio $c(n,s)/c(p,\ s+1)$, the various lines have a fairly simple interpretation:

Case (b): for this case $R_N=1$ for flooding with $n=p$, so that $c(n,s)$ and $c(p,\ s+1)$ are equal. This leads to perfect reflection of $n(s+\frac{1}{2})$, $p(s+\frac{1}{2})$ into $n^*(s+\frac{1}{2})$, $p^*(s+\frac{1}{2})$ through n_i, p_i as discussed in connection with Eq. (3.9).

Case (a) corresponds to $c(n,s)>c(p,\ s+1)$ as can be seen from the fact that $R_N=1$ only when $p>n$ in the flooded condition. It is probable that an ordinary donor level (a column-five substitutional impurity) is of this type. Such a level is attractive with charge plus one for electron capture but neutral for hole capture. Furthermore, the energy to be dissipated is smaller for electron capture because $E(s+\frac{1}{2})$ lies above E_i and is near to E_c than E_v. Hence, the value of $c(n,s)$ is probably much larger than $c(p,\ s+1)$ for such levels.

Cases (c), (d), and (e) correspond to hole capture more easily than electron capture. These cases, as we shall discuss in a later section may correspond to a high-lying acceptor level with a multiple negative charge. For such a case, $c(n,s)$ may be small due to Coulomb repulsion.

Case (f) represents the improbable situation in which, when $F=E(s+\frac{1}{2})$ each of the four transitions proceeds at exactly the same rate as the others. It corresponds to

$$n(s+\tfrac{1}{2})c(n,s)=p(s+\tfrac{1}{2})c(p,\ s+1)=e(p,s)=e(n,\ s+1), \quad (4.14)$$

so that the ratio of capture cross sections offsets the ratio of carrier densities. The line f thus cuts through the center of the denuded region as shown. All points in the denuded region correspond to R_N values very nearly equal to unity.

Cases (g) and (h) are hole-trapping levels. (See Sec. 6 for fuller discussion.) The $R_N=1$ line lies in the p-dominated region. This is a consequence of the fact that even when $F=E(s+\frac{1}{2})$ and $p(s+\frac{1}{2})\ll n(s+\frac{1}{2})$, the large ratio of $c(p,\ s+1)/c(n,s)$ causes the transition to be in equilibrium with holes.

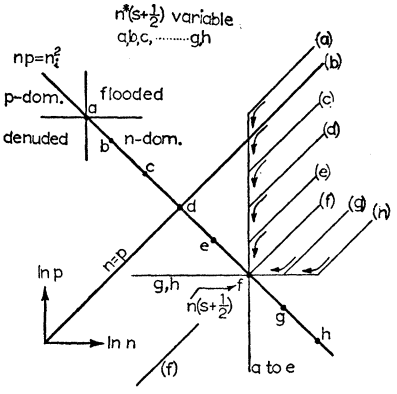

FIG. 2. Possible $R_N=1$ lines for a transition with $E(s+\frac{1}{2})>E_i$.

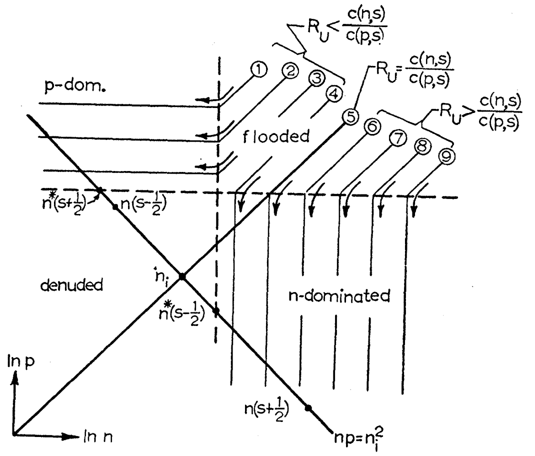

FIG. 3. Principle features of the R_U-diagram, the R_U=constant family, $E(s+\frac{1}{2})>E_i$; $E(s-\frac{1}{2})<E_i$.

5. RECOMBINATION RATE DIAGRAM: R_U DIAGRAM

The relative importance of the net steady-state recombination rate of two adjacent energy levels can be studied by comparing the recombination rates as indicated at the end of Sec. 3. This ratio R_U can also be represented graphically, but the situation is more complicated than R_N because the family of lines are specified by three rather than two parameters.

In the graphical representation of R_N, the distribution of charge, such as that shown in Fig. 1, two points, $n(s+\frac{1}{2})$ and $n^*(s+\frac{1}{2})$, uniquely determine R_N at any given densities p and n. However, to consider $R_U(s,p,n)=U(s+\frac{1}{2})/U(s-\frac{1}{2})$ of (3.23) we need to know $n^*(s+\frac{1}{2})$, $n^*(s-\frac{1}{2})$ and the ratio $c(n,s)/c(p,s)$. The recombination rate ratio,

$$R_U(s,p,n)=\frac{c(n,s)}{c(p,s)}\frac{n+n^*(s-\frac{1}{2})}{p+p^*(s+\frac{1}{2})}, \quad (5.1)$$

is shown in Fig. 3 as a family of R_U=constant lines, with equal $\ln R_U$ spacings. The line $R_U=1$ is not shown explicitly but can be obtained easily for a given $c(n,s)/c(p,s)$. The rounding of the corners near the regions $\ln[n/n^*(s-\frac{1}{2})]<2$ or $\ln[p/p^*(s+\frac{1}{2})]<2$ is also not shown.

Several important features of this diagram shall be discussed. The upper level $E(s+\frac{1}{2})$ or the transitions $s\rightleftarrows s+1$ provides the main contribution to the recombination in the n-dominated region and in the lower right of the flooded region if $c(n,s)/c(p,s)$ is nearly unity. The lower level $E(s-\frac{1}{2})$ becomes important in the p-dominated region and the upper left of the flooded region.

Now suppose that the s condition of the flaw has a negative electronic charge different from zero. Then

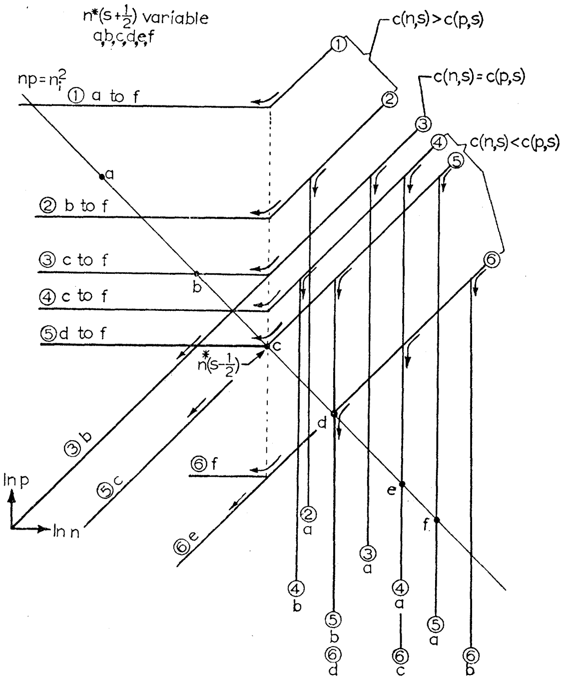

FIG. 4. Possible $R_U=1$ lines for a given $n^*(s-\frac{1}{2})$, variable $p^*(s+\frac{1}{2})$ and $c(n,s)/c(p,s)$.

due to Coulomb repulsion and attraction $c(n,s)/c(p,s)$ would be small compared with unity. Under this situation the $R_U=1$ line would be a line with the numeral label higher than 5 in Fig. 3. Thus, unless the semiconductor is sufficiently n-type, the lower level or the transition $s-1 \rightleftarrows s$ would be the dominant one.

A second general feature of the R_U diagram is that, under the extreme flooded condition, the line $n=p$ coincides with the line of $R_U=c(n,s)/c(p,s)$. This result states that the relative importance of the two levels under the flooded condition is equal to the ratio of the capture probabilities of the flaws in the s charge condition.

In the denuded region expression (5.1) reduces to

$$R_U(s,p,n)=e(p,s)/e(n,s), \tag{5.2}$$

so that the relative importance of the two levels is equal to the ratio of emission probabilities in the s charge condition.

The effect of different values of $p^*(s+\frac{1}{2})$ and the ratio $c(n,s)/c(p,s)$ with fixed $n^*(s-\frac{1}{2})$ is shown in Fig. 4. Six possible groups of $R_U=1$ lines are shown. The dependence of the $R_U=1$ on the ratio $c(n,s)/c(p,s)$ is similar to that shown in Fig. 3 in the flooded region. The position of $p^*(s+\frac{1}{2})$ determines whether the $R_U=1$ line becomes horizontal or vertical in the small densities region.

The situation for line (3)b of Fig. 4 is extremely unlikely. This represents a completely symmetrical case for which $n^*(s-\frac{1}{2})=p^*(s+\frac{1}{2})$ and $c(n,s)=c(p,s)$. The situation for line (5)c is also not likely for which $n^*(s+\frac{1}{2})=n^*(s-\frac{1}{2})$.

The line (1) in Fig. 4 corresponds to a situation where the steady-state recombination comes mainly from the transitions between states s and $s+1$. The lower level or the $s-1$ charge condition becomes important only for highly n-type semiconductors; for ordinary semiconductors, it is an electron-trapping state, or an electron recombination center filled with holes so that it is rather inactive.

A similar situation exists for line (6)b in Fig. 4 for holes.

6. HYPOTHETICAL FOUR-LEVEL FLAW WITH HOLE-TRAPPING TRANSITIONS

In terms of $E(s+\frac{1}{2})$ and $E^*(s+\frac{1}{2})$ relatively simple criteria can be given for the condition in which a transition $s\rightarrow s+1$ may act as a trap for minority carriers. Consider the situation

$$E(s+\tfrac{1}{2})<F<E(s+\tfrac{3}{2}), \tag{6.1}$$

$$E_i<F<E^*(s+\tfrac{1}{2}). \tag{6.2}$$

Inequality (6.1) requires that most of the flaws are in charge condition s under equilibrium conditions. Inequality (6.2) requires that the semiconductor be n-type and furthermore that it be in a p-dominated condition so far as the transition $s\rightarrow s+1$ is concerned. These conditions clearly represent a transition which acts as a trap for minority carriers, holes, in an n-type semiconductor.

For conditions in which the quasi-Fermi level F_p for holes is several kT above $E(s+\frac{1}{2})$, most of the flaws will be in charge condition $s+1$ and in accordance with (4.12) the number in charge condition s will be approximately

$$N_s=N_fp/p(s+\tfrac{1}{2})=N_f\exp[E(s+\tfrac{1}{2})-F_p]/kT, \tag{6.3}$$

provided $p<p(s+\frac{1}{2})$ or equivalently $F_p>E(s+\frac{1}{2})$. For greater hole densities than $p(s+\frac{1}{2})$, the traps will be filled in the sense that the more negative condition $s+1$ will be eliminated and N_s will be practically equal to N_f. This corresponds to each flaw trapping one plus charge or one hole. For still larger values of p, the condition $s-1$ may need to be considered.

The criterion that the traps be highly effective in trapping holes is that a majority of the holes should be in the traps so that $N_s>p$. This condition can be expressed in terms of an energy level $E'(s+\frac{1}{2})$ defined by[7]

$$E'(s+\tfrac{1}{2})=E_i-kT\ln(N_f/n_i). \tag{6.4}$$

When F falls on $E'(s+\frac{1}{2})$, p becomes equal to N_f.

If $c(n,s)$ were zero, the trap would be a perfect trap in the sense that it would play no role in electron-hole recombination. Small values of $c(n,s)$ or large values of $c(p,s+1)/c(n,s)$ lead to condition (6.2) which is re-

---

[7] See reference 1, Eq. (A.11).

quired to make the transition be dominated by holes. The ratio of $c(p,s+1)/c(n,s)$ needed to have (6.2) apply is found from (3.9) to be given by

$$kT\ln[c(p,s+1)/c(n,s)] = E(s+\tfrac{1}{2})+E^*(s+\tfrac{1}{2})-2E_i. \tag{6.5}$$

Thus, for a transition deeper than half the energy gap for hole capture, so that $E(s+\frac{1}{2})>E_i$, large ratios of $c(p,s+1)/c(n,s)$ are required to fulfill condition (6.1) with $E^*(s+\frac{1}{2})>E(s+\frac{1}{2})$. Examples of this situation are shown by lines (g) and (h) of Fig. 2 corresponding to cases with $n_s^*(s+\frac{1}{2})>n_s(s+\frac{1}{2})$ and thus, to $E^*(s+\frac{1}{2})>E(s+\frac{1}{2})$. For ratios of $c(p,s+1)/c(n,s)$ near unity, trapping levels for holes can occur only if $E(s+\frac{1}{2})$ is substantially less than E_i so that $E^*(s+\frac{1}{2})$ can lie above F for an n-type semiconductor.

Similar considerations apply for electron trapping in a p-type semiconductor.

We shall illustrate how trapping levels may occur by considering a hypothetical model of a four-level flow having five conditions of charge with

$$s=-1,\,0,\,1,\,2,\,3. \tag{6.6}$$

In the notation of S-L as shown on Fig. 5, the energy levels are

$$E(-\tfrac{1}{2})=E_{1d},\quad E(\tfrac{1}{2})=E_{1a},\quad E(1\tfrac{1}{2})=E_{2a},\quad E(2\tfrac{1}{2})=E_{3a}. \tag{6.7}$$

The energy level scheme resembles that of Au in germanium, for example.

A point which we wish to illustrate in this section is that when the flaw becomes multiply charged, the starred value $E^*(s-\frac{1}{2})$ obtained from $E(s-\frac{1}{2})$ may lie on the same side of E_i as does $E(s-\frac{1}{2})$ and indeed may lie farther from E_i than does E_s. Under such circumstances the second acceptor state is in effect a trap for holes and the third acceptor state might well be a very-long-time-constant deep trap.

In the case of a localized flaw, such as a gold atom, the neutral condition may well be one in which there are a number of relatively deep-lying unoccupied energy levels. Because of the shielding of the dielectric constant, the "electron affinity" may well be positive even for the second or third electron added. However, when an electron is added to an already negatively charged flaw, it will be necessary to overcome a substantial electrostatic repulsion. This situation is represented in Fig. 6.[8] The top curve shows the potential energy seen by the second added electron which ends up bound to the center with an effective energy denoted by E_{2a}. On the other hand, hole capture involving the same energy level is represented by a potential energy corresponding to charge of -2 on the flaw. Thus, on a semiclassical model, we see that $c(p,2)$ for hole capture may be larger than $c(n,1)$ by a factor larger than the activation energy factor $\exp(U_{2a}/kT)$.

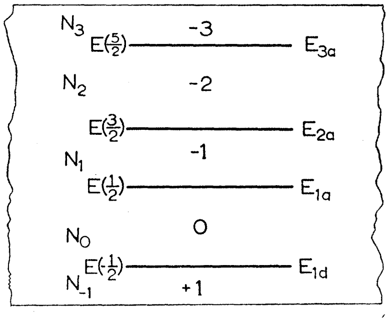

FIG. 5. Energy level scheme for a many-level flaw.

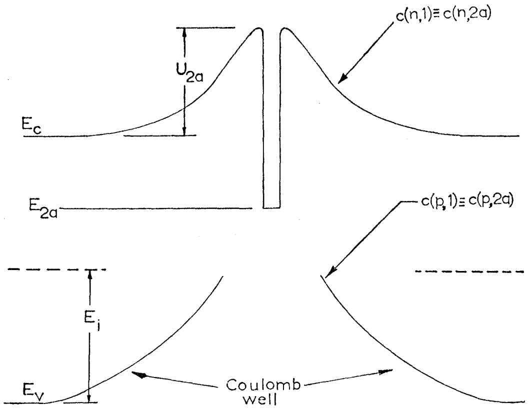

FIG. 6. The capture-emission situation for a second acceptor level showing the different effects of Coulomb charge on electrons and holes.

From (6.5) it is evident that the amount ΔE_{2a} by which $E_{2a}{}^*$ lies above the reflected value $2E_i-E_{2a}$ corresponding to equal capture constants is given by

$$\Delta E_{2a}\equiv kT\ln[c(p,2)/c(n,1)]>U_{2a}. \tag{6.8}$$

If the bound electron is several tenths of an electron volt below E_c, then its wave function will be closely restricted to the neighborhood of the flaw and thus will be largely restricted inside of most of the Coulomb rise of the potential U_{2a}, which may also be as large as several tenths of an electron volt at one atomic diameter from the center. Thus, ΔE_{2a} may well be larger than $2(E_{2a}-E_i)$ with the result that $E_{2a}{}^*$ lies above E_{2a}. This situation is represeted in Fig. 7(c).

As a preliminary to considering high injection level situations for the four-level flaw, the thermal equilibrium situation is reviewed in Fig. 8. As has been shown in S-L, the centers are largely in one condition unless the fermi level falls on an energy level. In this special case, the flaws are half in each of the adjacent conditions. The fraction of the flaws in a given condition fall off exponentially as with slope $1/kT$ or $2/kT$ or $3/kT$ on a logarithmic plot as shown in Fig. 8(a).

In Eq. (3.19) it was shown that the flaw acts as if it

---

[8] W. C. Dunlap, Jr., in *Proceedings of the Conference on Photoconductivity, Atlantic City, November, 1954* (John Wiley and Sons, Inc., New York, 1956), p. 539.

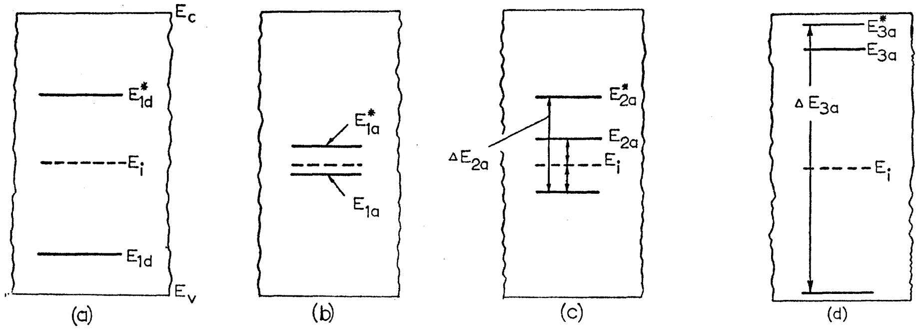

FIG. 7. Energy levels and starred energy levels for a flaw with four levels: (a) first donor level; (b) first acceptor level; (c) second acceptor level and possible hole trap; (d) third acceptor level and possible very-long-time-constant hole trap.

were several sets of simple S-R recombination centers, the number M_s of centers with energies $E(s+\frac{1}{2})$ being given according to the following scheme:

$$E_{1d} \qquad M_{1d}=N_{1d}+N_0, \tag{6.9a}$$

$$E_{1a} \qquad M_{1a}=N_0+N_{1a}, \tag{6.9b}$$

$$E_{2a} \qquad M_{2a}=N_{1a}+N_{2a}, \tag{6.9c}$$

$$E_{3a} \qquad M_{3a}=N_{2a}+N_{3a}. \tag{6.9d}$$

These numbers are represented in Fig. 8(b). Thus, in general the flaws mimic double their number of recombination centers. The decay constant for small injected carrier densities (the reciprocal of the lifetime) is approximately equal to the sum of the decay constants calculated by S-R summed over the two levels associated with the prevalent condition of the flaw.

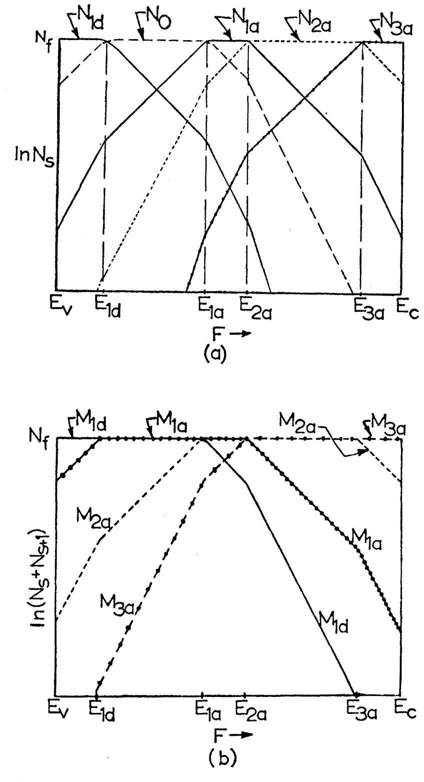

FIG. 8. Occupany numbers and effective number of recombination centers. (a) Occupancy numbers; (b) numbers of recombination centers.

The equations for any particular case may readily be worked out on this basis.

In order to explain what may happen at large injection levels, Fig. 9 has been constructed to correspond to the energy level scheme of Fig. 7. In Fig. 9(a) the lines for $R_N=1$ are shown for these energy levels. Lines E_{1d} and E_{1a} are normal cases in which the distribution in the two conditions adjoining the level $E(s-\frac{1}{2})$ is controlled by the majority carriers when the Fermi level falls on $E(s-\frac{1}{2})$. The cases of E_{2a} and E_{3a} are like Fig. 2(g) or (h) and correspond to a trap for minority carriers since the occupancy depends on minority carriers when F falls on E_{2a} or E_{3a}.

The lines for E_{1a} and E_{1d} in the flooded region have been drawn as if $c(n,1d)/c(p,0)$ were less than $c(n,0)/c(p,1a)$. This may be reasonable if the difficulty of an electron dissipating energy in falling from the conduction band to E_{1d} makes $c(n,1d)$ substantially smaller than the other capture cross sections involved. The writers do not contend that this will inevitably be the case and introduce it chiefly to illustrate a general situation for which the charge condition zero discussed in the next paragraph does not exist in this region.

The new feature is shown in Fig. 9(a). Owing to the crossing of the E_{1d} and E_{1a} lines, there is a region where condition $s=0$ is unlikely to occur for any ratio of n to p. In the region between the lines E_{1a} and E_{1d}, the neutral condition is always less favored than either the 1a or 1d condition. The dashed line divides this region shown in Fig. 9 into two regions of either the condition 1d predominant or 1a predominant. The crossing of E_{1d} and E_{1a} corresponds to equal densities of flaws in the three charge states 1d, 0, and 1a. The lines of $R_U(s)$ are not shown in Fig. 9 for the sake of clarity.

Figure 9(b) shows the variation of n and p under conditions of injection, neglecting changes in chemical charge. It indicates that the majority carrier density is unaffected until the minority carrier density approaches the majority density. Thereafter, the line $n=p$ is followed.

It is evident that very complicate dependences of recombination rate upon injected carrier density can occur for a situation like Fig. 9. For example, if line L_1

of Fig. 9(b) is followed, the flaws, initially all in condition $2a$, transfer to condition $1a$, and later to condition $1d$ for the hypothetical case of gold in germanium shown in Fig. 9(a). If the lifetimes associated with E_{2a} are relatively long compared to those of E_{1a}, there will be a rapid increase in recombination as L_1 crosses E_{2a}. Such changes should be observable experimentally and should furnish a basis for determining location of the constant R_N lines and the R_U lines on these diagrams.

Figure 10 shows another case of four-level flaws for which both an electron trapping state and a hole trapping state are present. In Fig. 10(a) both the $N_s=N_{s+1}$ and $N_s=N_{s+2}$ lines are shown. In general, the $N_s=N_{s+2}$ lines change direction twice. The special feature of this case is evident in Fig. 10(a): In the denuded region the neutral charge condition predominates and the charge conditions -1 and $+1$ cannot be important for any ratio of n to p in this region.

Figure 10(b) shows the $R_U(s)=1$ lines. The regions are labeled by the dominating transitions such as $s\rightleftarrows s+1$. The dashed-line boundary corresponds to $R_U(1)R_U(0)=1$ or equal transition traffic for $-1\rightleftarrows 0$ and $1\rightleftarrows 2$ transitions.

Special features of this diagram are: (1) In the region of small p and $n \div n_i$, the transition $0\rightleftarrows 1$ cannot be important for any ratio of n to p; (2) in the denuded region the transition $-1\rightleftarrows 0$ predominates and (3) for nondegenerate p-type crystals the transition $-1\rightleftarrows 0$ predominates.

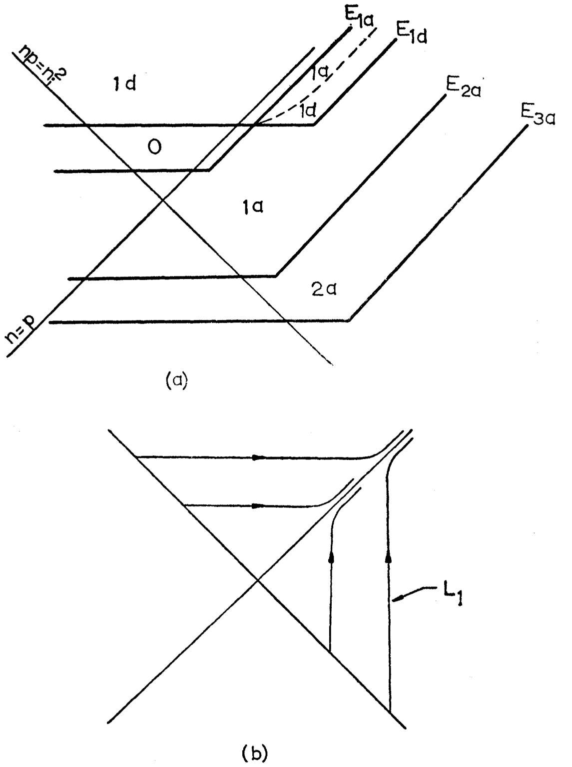

FIG. 9. A hypothetical four-level flaw. (a) The $R_N=1$ diagram; (b) the effect of heavy injection upon n and p assuming negligible changes in chemical and flaw charge densities compared to majority carrier densities.

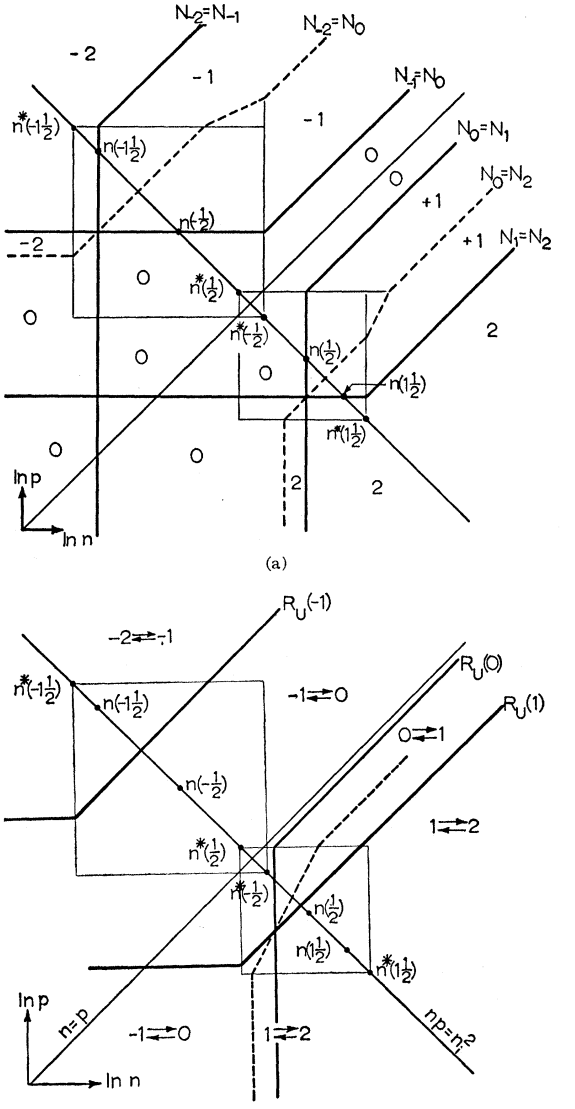

FIG. 10. A hypothetical case with both an electron and a hole trapping state. (a) The lines for $R_N=1$; (b) the lines for $R_U=1$.

From the discussion made for the case shown in Fig. 10, it is evident that if an n^+p junction is made in a semiconductor with this type of four-level flaw, the electrical characteristics, under either forward or reverse bias conditions can be predicted essentially by a single-level, recombination generation center corresponding to the transition $-1\rightleftarrows 0$ and energy $E(-\frac{1}{2})$.

7. DEPENDENCE OF LIFETIME OR SURFACE RECOMBINATION VELOCITY UPON FERMI LEVEL

The expression for recombination rate can be put into a form which may be helpful in interpreting data on lifetime or surface recombination velocity. Equation

(3.19) for the recombination rate $U(s+\frac{1}{2})$ may be rewritten as follows

$$\begin{aligned} U(s+\tfrac{1}{2}) = {} & M(s+\tfrac{1}{2})[e(n,\, s+1)e(p,s)]^{\frac{1}{2}} \\ & \times\{\exp[(F_n-F_p)/kT]-1\} \\ & \div\{\exp[(E_m(s+\tfrac{1}{2})-F_p)/kT] \\ & +\exp[(F_n-E_m(s+\tfrac{1}{2}))/kT] \\ & +2\cosh[(E(s+\tfrac{1}{2})-E_m(s+\tfrac{1}{2}))/kT]\}. \qquad (7.1) \end{aligned}$$

In this expression $E_m(s+\frac{1}{2})$ is the energy corresponding to the geometric mean of $n(s+\frac{1}{2})$ and $n^*(s+\frac{1}{2})$ shown on Fig. 1. Its value is thus

$$\begin{aligned} E_m(s+\tfrac{1}{2}) &= [E^*(s+\tfrac{1}{2})+E(s+\tfrac{1}{2})]/2 \\ &= E_i+(kT/2)\ln[c(p,\, s+1)/c(n,s)]. \qquad (7.2) \end{aligned}$$

The above expressions may be obtained by rewriting (3.19) in terms of the mean densities corresponding to $E_m(s+\frac{1}{2})$. These are given by

$$n_m(s+\tfrac{1}{2})=n_i^2/p_m(s+\tfrac{1}{2})=n_i[c(p,\, s+1)/c(n,s)]^{\frac{1}{2}}. \qquad (7.2\text{a})$$

In terms of them (3.19) becomes

$$\begin{aligned} U(s+\tfrac{1}{2}) = {} & M(s+\tfrac{1}{2})[e(n,\, s+1)e(p,s)]^{\frac{1}{2}} \\ & \times\exp[(F_n-F_p)/kT-1] \\ & \div[p/p_m(s+\tfrac{1}{2})+n/n_m(s+\tfrac{1}{2})+p(s+\tfrac{1}{2})/p_m(s+\tfrac{1}{2}) \\ & +n(s+\tfrac{1}{2})/n_m(s+\tfrac{1}{2})]. \qquad (7.3) \end{aligned}$$

The last two fractions are reciprocals and are equal to the hyperbolic cosine term in (7.1). Because the mean density is a geometric mean of n and n^* terms, the last two terms can be replaced by ratios of n^* to n_m equally well.

We shall next consider the case of small deviations from thermal equilibrium so that F_p and F_n differ little from the equilibrium F. Under these conditions (7.1) reduces to

$$\begin{aligned} U(s+\tfrac{1}{2}) = {} & M(s+\tfrac{1}{2})[e(n,\, s+1)e(p,s)]^{\frac{1}{2}}(F_n-F_p) \\ & \div 2kT\{\cosh[(F-E_m(s+\tfrac{1}{2}))/kT] \\ & +\cosh[(E(s+\tfrac{1}{2})-E_m(s+\tfrac{1}{2}))/kT]\}. \qquad (7.4) \end{aligned}$$

This expression is completely symmetrical in $E^*(s+\frac{1}{2})$ and $E(s+\frac{1}{2})$ because cosh is an even function and differs by equal magnitudes from $E(s+\frac{1}{2})$ and $E^*(s+\frac{1}{2})$. The denominator is symmetrical about $F=E_m(s+\frac{1}{2})$ and is substantially constant if F lies between $E(s+\frac{1}{2})$ and $E^*(s+\frac{1}{2})$. Outside of that range its value is smaller than the constant value by a factor

$$\exp(-|\Delta F|/kT), \qquad (7.5)$$

where ΔF is the amount by which F falls outside the constant range.

Expression (7.4), in a somewhat different form has been used by Many, Harnik, and Margoninski[9] to interpret variations in surface recombination velocity with variations in surface potential. Their treatment is more complete than earlier ones due to Brattain and Bardeen[10] and to Stevenson and Keyes[11] who assume that the capture cross sections are equal and hence that $E_m(s+\frac{1}{2})=E_i$. These treatments, of course, deal with single-level flaws.

If we assume[8] that the quasi-Fermi levels at the surface equal those in the interior and that the energy bands bend going towards the surface so that at the surface the Fermi level lies at F_s above E_i, and assume further that in the bulk of the specimen the hole and electron densities are increased by $\delta p=\delta n$ above p_b and n_b, then

$$n_i^2(F_n-F_p)/kT=(p_b+n_b)\delta p, \qquad (7.6)$$

and the rate of recombination per unit area per unit disturbance in carrier density is

$$\begin{aligned} S(s+\tfrac{1}{2},\, F_s) \equiv {} & M_s(s+\tfrac{1}{2})(p_b+n_b)[e(n,\, s+1)e(p,s)/n_i^2]^{\frac{1}{2}} \\ & \div 2n_i\{\cosh[(F_s-E_m(s+\tfrac{1}{2}))/kT] \\ & +\cosh[(E(s+\tfrac{1}{2})-E_m(s+\tfrac{1}{2}))/kT]\}, \qquad (7.7) \end{aligned}$$

where $M_s(s+\frac{1}{2})$ is the density of flaws in charge conditions s and $s+1$ per unit area of the surface and $S(s+\frac{1}{2},\, F_s)$ defined thus by (7.7) is the contribution of transitions between s and $s+1$ to the surface recombination velocity S.

Equation (7.7) transforms to equation (5) of Many *et al.*[9] by noting that their $q\phi_0$ is $E_m(s+\frac{1}{2})-E_i$.

As pointed out in connection with (7.5), as the potential ϕ_s of the surface, taken as

$$q\varphi_s=F_s-E_i, \qquad (7.8)$$

varies, this gives a plateau for $S(s+\frac{1}{2},\, F_s)$ followed by decreasing regions.

In previous treatments, the surface traps have been

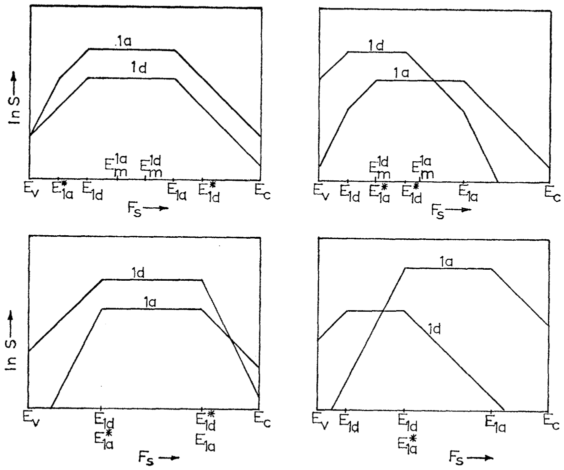

FIG. 11. Possible variations of surface recombination velocity with surface Fermi level for a two-level flaw.

---

[9] Many, Harnik, and Margoninski, in *Semiconductor Surface Physics*, edited by R. H. Kingston (University of Pennsylvania Press, Philadelphia, 1957), p. 85.

[10] W. H. Brattain and J. Bardeen, Bell System Tech. J. **82**, 1 (1953).

[11] D. T. Stevenson and R. J. Keyes, Physica **20**, 1041 (1954).

considered as single-level traps. For such cases the total S is simply the sum of the individual contributions. In the case of flaws, there is an additional possible variation because the effective number of traps may alter as F_s crosses one of the levels $E(s+\frac{1}{2})$.

If we assume a flaw with three charge conditions and two levels, say E_{1d} and E_{1a} for example, then if $E_{1d} < E_{1a} < E_{1d}^*$, the contribution of $M_s(1d)$ to S starts to drop when F_s exceeds E_{1a} because of decreasing $M_s(1d)$ and drops still faster when F_s also exceeds E_{1d}^*, the slope being $-2/kT$ on a $\ln S(1d,F_s)$ *versus* F_s plot.

In Fig. 11 several cases are illustrated to show examples of $\ln S$ *versus* F_s that may arise for a two-level flaw. Only the individual S curves are shown. The total S is their sum. It should be noted that in general if there is overlap of the two plateaus, they will have at least one common limit. It should also be noted that slopes of $\pm 2/kT$ can occur for flaws whereas they cannot for single-level recombination centers. No significance should be attached to the relative heights represented.

Diagrams similar to Fig. 11 can also be constructed for the bulk-recombination constant $(1/\tau)$. The plateaus in this case are concave upward due to the (p_b+n_b) term which has a minimum for $F=F_i$. However, the relative effects of competition of the levels of the flaw causes the same sort of changes compared to several single-level traps for bulk lifetimes as for surface recombination velocity.

PHYSICAL REVIEW VOLUME 109, NUMBER 4 FEBRUARY 15, 1958

Quantum Theory of Galvanomagnetic Effects

PETROS N. ARGYRES
Westinghouse Research Laboratories, Pittsburgh, Pennsylvania

(Received May 21, 1957)

A study is made of the effect of the quantization of the electron orbits in a magnetic field on the galvanomagnetic properties of an isotropic semiconductor or semimetal in the phonon-scattering range. The conductivity tensor is calculated using the quantum-mechanical density operator. A generalized Boltzmann equation is derived, in which the scattering for the density matrix elements of interest in the calculation of the current can be described by a relaxation time $\tau(\epsilon)$, that is for most energies inversely proportional to the density of electronic states in the magnetic field.

The Hall coefficient and transverse resistivity are studied for a number of different sets of conditions for the electron density, the magnetic field, and the temperature. It is found that in all cases the resistivity at very high fields is an increasing function of field. Further the Hall coefficient at high fields does not approach the classical limit. Both Hall coefficient and resistivity under some conditions exhibit an oscillatory dependence on magnetic field. The deviation of the resistivity from classical behavior is easily observable. The deviation of the Hall coefficient is in most cases quite small; however, it may be possible to observe it.

1. INTRODUCTION

THE various effects of a magnetic field on the transport properties of a metal or semiconductor provide us with useful information about the electronic structure of these substances. Transverse and longitudinal magnetoresistance measurements give strong indications about their energy-band structures, Hall-effect measurements determine the type and density of charge carriers, and cyclotron resonance absorption affords a direct measurement of the effective mass of the carriers.

Theories for all these effects have been based in general on the one-electron band model of solids, and, ordinarily, on the Boltzmann transport equation. Such theories with various degrees of generality have been given by many authors.[1]

We shall concern ourselves here with the observation that the treatment of the magnetic field in conjunction with the Boltzmann transport equation is only approximately correct. There are two approximations involved and both limit the applicability of the results of these theories to small magnetic fields. The usual Boltzmann transport equation as applied in the band theory of solids[2] is an integro-differential equation for the distribution function $f(\mathbf{k})$ which gives the probability for a Bloch state with wave vector $\mathbf{k}$ to be occupied. It states that the total time rate of change of the distribution function (for homogeneous media) is equal to its rate of change due to the accelerating fields of force, like electric and magnetic fields, plus that due to the scattering by imperfections in the crystal. The approximations we mentioned above are the following. First, in these treatments the rate of change of $f(\mathbf{k})$ due to the magnetic field $\mathbf{H}$ is taken to be $(e/\hbar c)\mathbf{v}\times\mathbf{H}\cdot\nabla f(\mathbf{k})$, where $\mathbf{v}$ is the average velocity of an electron in the Bloch state with wave vector $\mathbf{k}$. This was first derived for a wave packet by Jones and Zener[3] by neglecting

---

[1] J. McClure, Phys. Rev. **101**, 1642 (1956). This reference contains references to other relevant work.

[2] See, e.g., A. H. Wilson, *The Theory of Metals* (Cambridge University Press, New York, 1953).

[3] H. Jones and C. Zener, Proc. Roy. Soc. (London) **A144**, 101 (1934).