Double Injection in Insulators

MURRAY A. LAMPERT
RCA Laboratories, Princeton, New Jersey
(Received August 10, 1961)

Double injection in insulators is analyzed taking into account that the lifetimes for the injected electrons and holes are different and vary with injection level. Assuming charge neutrality, a detailed solution is obtained for the simple model of an insulator with a single set of recombination centers filled with electrons in thermal equilibrium. The major results are: (i) There is a threshold voltage $V_{\rm th}$ below which the double-injection current is negligible and at which this current rises steeply with voltage. At this threshold voltage the hole transit time, $t_{p,\rm th}=L^2/\mu_p V_{\rm th}$, is comparable to the hole lifetime $\tau_{p,\rm low}$. The subscript "low" refers to the lifetime at low injection levels. The lifetime at high injection levels in this model will generally be longer. (ii) For an electron capture cross section σ_n much smaller than the hole capture cross section σ_p, there is a negative resistance between $V_{\rm th}$ and $V_M\approx(\sigma_n/\sigma_p)V_{\rm th}$. With increasing current, the voltage decreases from $V_{\rm th}$ to V_M. This negative resistance has its origin in an increasing hole lifetime with increasing injection level, owing to electron depopulation of the recombination centers by hole capture. (iii) At still higher currents the double-injection current-voltage characteristic is similar to that for a semiconductor at high injection levels. At sufficiently low currents the neutrality-based, double-injection solution is no longer self-consistent with respect to the neglect of space charge, and the true current is a one-carrier SCL (space-charge-limited) current which, for the simple model analyzed, is the electron SCL current for a trap-free insulator. In real insulators the one-carrier SCL current may mask the voltage threshold effect, (i) above, depending on the physical parameters of the crystal. On the other hand, under the condition specified in (ii) above, the negative resistance will always be observed. Experimentally, the negative resistance should produce either current oscillations or a hysteresis in the voltage vs current for dc applied voltages. Both effects have been widely observed in insulators and high-resistivity semiconductors. It is shown how the theory can be extended to more complicated models.

I. INTRODUCTION

IN this paper the problem of double injection, that is, the simultaneous injection of electrons from a negative contact and holes from a positive contact, into an insulator is discussed. In a previous paper,[1] hereafter referred to as (I), the author, in collaboration with Rose, analyzed field-driven, double-injection current flow in semiconductors and insulators for the injected-plasma case, in which most of the injected electrons and holes are not trapped but remain free to carry current. The theory developed in (I) is applicable when the injected carrier densities are large compared with the density of discrete states in the forbidden energy gap. Under this condition there is a single, common lifetime for electrons and holes. In real insulators, at low injection levels, the lifetimes of electrons and holes are generally significantly different and, accordingly, so are the free carrier densities. Charge neutrality is maintained through absorption of the difference in the free carrier densities in the deep-lying states in the forbidden gap.

In this paper we analyze in detail a particularly simple model of an insulator with the characteristic feature of differing lifetimes for electrons and holes at low injection levels. This model is illustrated in the schematic energy-band diagram of Fig. 1. There is a single, discrete set of recombination centers located in the forbidden gap well below the Fermi level and therefore completely filled with electrons at thermal equilibrium (no applied voltage). We emphasize that this model was chosen for study because of the relative ease of mathematical analysis and because of its relatively simple behavior under double injection. The same mathematical formalism will also yield the solution for far more complicated models.

[1] M. A. Lampert and A. Rose, Phys. Rev. **121**, 26 (1961).

Before embarking on a mathematical analysis of the model of Fig. 1 it is perhaps instructive to see what significant questions concerning its double-injection behavior are suggested by an examination of the problem on physical grounds. Referring to Fig. 1 we see that the presence of the filled recombination centers will interfere with the free transit of holes across the insulator, in that the centers will act as a sink for the injected holes. In this sense the recombination centers present a kind of recombination barrier to the transit of holes. A question which then suggests itself is whether significant double injection, i.e., a two-carrier current substantially in excess of the one-carrier current that can be supported at the same voltage, can take place before sufficient voltage is applied to make the hole drift transit time t_p from anode to cathode comparable to the low-injection-level hole lifetime $\tau_{p,\rm low}$. Here this question is answered negatively, contrary to previous speculation.[2] We show that there is a threshold voltage $V_{\rm th}$ for the onset of double-injection currents, such that $t_{p,\rm th}\simeq 2\tau_{p,\rm low}$, where $t_{p,\rm th}$ is the hole transit time at voltage $V_{\rm th}$.

A second interesting feature of our problem derives from the variation of hole lifetime with injection level. If the capture cross section for holes σ_p of an electron-occupied recombination center is much larger than the capture cross section for electrons σ_n of an unoccupied center (e.g. σ_p corresponding to a Coulomb-attractive capture center, and σ_n to a neutral capture center),

[2] I. Broser and R. Broser-Warminsky, J. Phys. Chem. Solids **6**, 386 (1958). It is argued in this paper that even at low voltages, such that $t_p\gg\tau_{p,\rm low}$ and therefore with only a small fraction of the density of holes injected at the anode reaching the cathode, the double-injection current can be large compared with Ohmic or single-injection, space-charge-limited currents. Their argument leans, incorrectly, on an analog with a sensitive photoconductor, in which indeed a small density of free holes is associated with a much larger density of free electrons.

then with increasing injection level the recombination centers will tend to be depopulated by electrons in favor of holes. Consequently the hole lifetime must increase with increasing injection level; the recombination barrier to hole transit decreases as the double-injection current increases. We are led to inquire whether this behavior of the hole lifetime can produce a negative resistance in the current-voltage characteristic. Our calculations yield an affirmative answer; increasing current is accompanied by decreasing voltage over a voltage range from $V_{\rm th}$ to $V_M \approx (\sigma_n/\sigma_p)V_{\rm th}$.[3]

The electron depopulation of the recombination centers in favor of holes involves a transfer of electrons from the recombination centers to the conduction band. Thus, at the higher injection levels it is as if the recombination centers of density N_R were also shallow donors, thermally ionized at the ambient temperature. One might therefore suspect a close relationship between double injection into our model insulator at high injection levels and double injection, at comparable levels, into the corresponding n-type semiconductor, with thermal free-carrier density $n_0=N_R$, previously studied in (I). This suspicion is borne out by our calculations; the current-voltage characteristics for the two situations are identical at high-injection levels. Thus, our model insulator, in a manner of speaking, is electronically converted into an n-type semiconductor through double injection, and we describe the corresponding portion of the double-injection current-voltage characteristic as the "semiconductor regime." The hole lifetime in this semiconductor regime, $\tau_{\rm high}$ (the electrons and holes have a common lifetime in this regime, hence subscript p is omitted) is much longer than its low-injection-level lifetime $\tau_{p,\rm low}$: $\tau_{\rm high} \simeq (\sigma_p/\sigma_n)\tau_{p,\rm low}$. Consequently the semiconductor regime persists down to voltages much lower than $V_{\rm th}$, since the voltage requirement for validity of this regime is, from (I): $t_p = L^2/\mu_p V < \tau_{\rm high}$. Accepting the fact that voltage varies continuously with current, proved in Appendix A, we now see, perhaps more clearly, why a negative resistance is inherent in our model, for $\sigma_p \gg \sigma_n$.

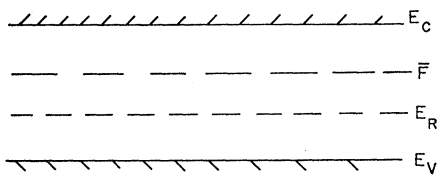

FIG. 1. Schematic energy-band diagram for a simplified mode of an insulator. There is a single set of defect centers, of density N_R, located at energy E_R well below the thermal-equilibrium Fermi level F. These centers function as recombination centers under double injection. E_C and E_V denote the conduction and valence band energy extrema, respectively.

[3] The fact that an increase in free carrier lifetime with injection level can lead to negative resistance has been previously pointed out by Stafeev, who discussed the question qualitatively and with application to diffusive current flow problems: V. I. Stafeev, J. Solid State Phys. Acad. Sci. U.S.S.R. **1**, 841 (1959) [translation: Soviet Phys.—Solid State **1**, 763 (1959)].

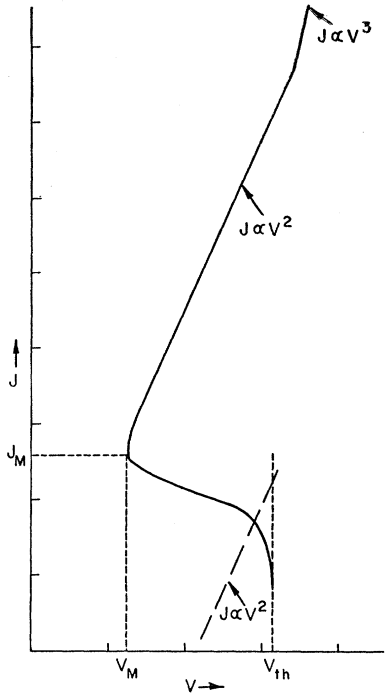

FIG. 2. Schematic current (J)-voltage (V) characteristic (the solid curve), on a log-log plot, for double injection in the model of Fig. 1, for the case that $\langle v\sigma_p\rangle \gg \langle v\sigma_n\rangle$. The lower, dashed curve represents the one-carrier (electron), space-charge-limited current-voltage characteristic for the trap-free insulator.

These various features of double injection for the model of Fig. 1, with $\sigma_p \gg \sigma_n$, are illustrated in the schematic current-voltage characteristic shown on a log-log plot (the solid line) in Fig. 2. They constitute the main results of this paper.

One feature of the present work is of particular interest because it represents a departure from (I) and other related literature[4,5] on field-driven, double-injection currents in solids; the above results have been derived under the assumption of local charge neutrality throughout the insulator. The *double-injection currents studied here are in no way space-charge-limited, but are purely recombination-limited.* [In precisely the same sense, the so-called "Ohmic relaxation regime" of the injected plasma problem studied in (I) is also purely recombination controlled; it is in no way space-charge controlled, as was implied by the derivation in (I). Thus, while the assumption of charge neutrality was not made in (I) in discussing this regime, it could have been made without altering the result. We show this later in this paper and further provide a correction to (I) bearing on this point. In this paper we shall hereafter refer to the "Ohmic relaxation regime" simply as the "semiconductor regime," thereby highlighting the role of the thermally generated free carriers in determining the current-voltage characteristic.]

Interestingly, both at sufficiently low and sufficiently high currents the assumption of charge neutrality is necessarily incorrect and space charge plays an important role in the current flow. At low currents, the

[4] R. H. Parmenter and W. Ruppel, J. Appl. Phys. **30**, 1548 (1959).

[5] M. A. Lampert, RCA Rev. **20**, 682 (1959).

actual current flow in our model insulator is the purely space-charge-limited, one-carrier electron current corresponding to a trap-free insulator. Current varies as the square of the voltage as indicated by the dashed line in Fig. 2. The recombination centers, being filled with electrons in this regime, cannot serve as electron traps. At very high currents, both recombination and space charge play a role in limiting the current. This leads to a cube law dependence of current on voltage,[5] shown as the highest branch of the current-voltage characteristic (the solid curve) in Fig. 2.

The scheme of this paper is as follows. We first discuss, in Sec. II, the assumptions underlying the theory and some consequent limitations on its applicability. In Sec. III are given simplified derivations, for the model of Fig. 1, of the threshold voltage, of the semiconductor (square-law) regime, and of the negative-resistance regime which is present when $\sigma_p \gg \sigma_n$. In the concluding Sec. IV some implications of the theory for experiment are discussed. It is pointed out there how the present theoretical results may be related to some well-known phenomena observed in insulators and in high-resistivity semiconductors.

The mathematical details pertinent to the rigorous solution are given in the Appendices. Because the mathematical technique employed in the analysis of the model of Fig. 1 is applicable to much more general models of insulators, the theory is developed in some detail. In Appendix A the formal solution is obtained for the model of Fig. 1. In Appendix B this solution is studied for the case that $\sigma_p \gg \sigma_n$, the condition leading to a marked negative resistance. In Appendix C the domain of self-consistency of the solution is examined with regard to the assumption of charge neutrality. In Appendix D it is shown that the mathematical technique of the previous appendices can be extended to analyze far more complicated models of insulators. Finally, in Appendix E the space charge in the semiconductor regime is derived and a correction therein made to (I).

The discussion in this paper is confined throughout to steady-state, one dimensional current flow. Further, mks units are employed throughout except where otherwise specified.

II. ASSUMPTIONS

The assumptions on which the theory is based are as follows:

(i) Charge neutrality holds throughout the insulator.

For the model of Fig. 1 this assumption is equivalent to the specific requirement that the injected space-charge density be small compared to the difference between the injected free-electron and free-hole densities, n and p respectively. In Appendix E it is shown that this requirement is indeed met in the semiconductor regime of our problem, which sets in after the common, high-injection level lifetime is attained. This is the square-law branch of the solid curve in Fig. 2. At still higher voltages finally the requirement is no longer met, and space charge must be carefully treated to obtain the correct solution. This is the so-called insulator regime of (I), which is also a regime of the present problem, namely the cube-law branch of the solid curve of Fig. 2. The solution for this regime is already given in (I). The new results in the present problem arise in the domain of currents below the semiconductor regime, and therefore it is in this domain that we must examine the assumption of charge neutrality. Since $p \ll n$ in this domain, for our model insulator, the requirement is simply that the space-charge density be small compared to n. Clearly this is also the condition characterizing significant double-injection current flow as opposed to one-carrier, space-charge-limited current flow. Thus, at currents below the insulator regime, the domain of validity of the neutrality assumption is simultaneously the domain of phenomena peculiar to double-injection current flow.

It is clear that the double-injection current flow described by our theory cannot hold down to arbitrarily low currents. For the theory predicts a threshold voltage for double injection, and corresponding to this voltage there is a space charge which must be exceeded by n if the solution is to be self-consistent with the neutrality assumption. Thus at "low" current levels the flow must be a one-carrier, space-charge-limited current; for the model of Fig. 1 this is the electron current for a trap-free insulator, given by[6] $J = 9\epsilon\mu_n V^2/8L^3$, with ϵ the static dielectric constant. This square-law current-voltage characteristic is shown as the dashed line in the log-log plot of Fig. 2. The transition from one-carrier, space-charge-limited flow to two-carrier, neutralized flow takes place where the respective current-voltage characteristics intersect. This transition is discussed in Appendix C for the case of greatest interest $\sigma_p \gg \sigma_n$. There it is shown that over a substantial range of practical conditions our double-injection solution is self-consistent, and therefore the negative resistance should be exhibited as calculated.

(ii) The current is volume-controlled, i.e., the contacts impose no significant constraints on the currents which are entering or leaving the crystal.

(iii) Diffusion currents are negligible.

This pair of assumptions is identical to the pair of assumptions (ii) and (iii), respectively, of (I). The discussion of these assumptions in (I) is equally valid here and therefore need not be repeated at length. Both assumptions are violated in the immediate neighborhood of an injecting contact. Therefore the solutions obtained by our theory will give the correct over-all description of double-injection current flow only if the separation between injecting contacts is sufficiently large. Since diffusion-type, rather than

[6] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, New York, 1940), 1st ed., p. 172.

field-driven type, solutions dominate near the contacts, "sufficiently large" means $L/L_a \gg 1$, where L is the separation between contacts and L_a the ambipolar, high-lifetime diffusion length: $L_a^2 = 2D_nD_p\tau_{\text{high}}/(D_n+D_p)$, with D_n and D_p the electron and hole diffusion constants, respectively. For $L/L_a \gg 1$, corrections to the current-voltage characteristic over most of the negative-resistance regime of Fig. 2 are of order L_a/L. This is shown in Sec. III C. Only at voltages extremely close to the threshold voltage V_{th} will the corrections be more substantial, and at such voltages and currents the corrections due to space charge will also likely be quite substantial, as shown in Appendix C. Corrections to the current-voltage characteristic in the semiconductor regime, due to the contact constraints, will be small provided $L/L_a \gg 2[1+\ln(J/J_M)]$, where J_M is the current at the voltage minimum V_M in Fig. 2 and J is the actual current.[7]

(iv) Low-field, field-independent mobility conditions obtain.

This assumption is necessary to make the problem analytically tractable. It will frequently be realized in insulator studies. It is most likely to fail in studies of high-mobility semiconductors.

(v) Thermal re-emission of carriers from the recombination centers is negligible.

Specifically, the rate of thermal re-emission of carriers from the recombination centers to a band must be small compared to their capture rate from the band. The thermal re-emission rate per electron becomes negligible compared with the capture rate when the localized states involved lie more than a few kT below the electron demarcation level.[8] This situation ordinarily characterizes a recombination center in insulator studies. In those unusual situations where net recombination is a relatively small difference between large capture and thermal re-emission rates there are substantial calculational difficulties not only with double-injection problems but with photoexcitation problems also.[8]

(vi) The thermally-generated free carriers can be neglected.

First we require that the densities of injected free carriers substantially exceed the densities of thermally generated carriers for both electrons and holes. A second requirement is, referring specifically to the model of Fig. 1, that the hole transit time at the threshold voltage V_{th}, hence the low-injection-level hole lifetime, be shorter than the electron dielectric relaxation time. The converse situation, where the free-carrier lifetime is longer than the dielectric relaxation time, was studied in (I) under the assumption that the lifetime did not change with injection level. Both requirements will commonly be met in insulator studies.

[7] D. O. North (private communication).

[8] A. Rose, *Progress in Semiconductor* (Heywood and Company, Ltd., London, 1957), Vol. 2, p. 109.

III. SIMPLE, APPROXIMATE ANALYTICAL ARGUMENTS

In this section we present simplified derivations, for the model of Fig. 1, of the threshold voltage, of the semiconductor regime, and of the negative resistance regime for the case $\sigma_p \gg \sigma_n$. The rigorous analytical solution is given in Appendices A and B.

The equations characterizing the problem are the current-flow equation:

$$J = en\mu_n\mathcal{E} + ep\mu_p\mathcal{E} = \text{constant}, \tag{1}$$

the expression of neutrality:

$$n - p - p_R = 0, \tag{2}$$

the particle-conservation equations:

$$\mu_p \frac{d}{dx}(p\mathcal{E}) = r = -\mu_n \frac{d}{dx}(n\mathcal{E}), \tag{3}$$

and the recombination-kinetic expressions:

$$r = p\langle v\sigma_p\rangle n_R = n\langle v\sigma_n\rangle p_R, \quad n_R + p_R = N_R. \tag{4}$$

In the above equations J is the total current density, e the magnitude of the electronic charge, μ_n and μ_p the electron and hole mobilities, respectively, $\mathcal{E}$ the electric field intensity, x the position variable, r the recombination rate density, σ_p (σ_n) the cross section for capture of a hole (electron) by a filled (empty) recombination center, and v the free-carrier velocity. Angular brackets $\langle\ \rangle$ about a quantity denote the average of that quantity over the velocity distribution of the free carrier in question. Further, n and p are the injected free-electron and hole densities, respectively, and we have neglected the thermally generated free carriers. Finally, N_R is the density of recombination centers, p_R and n_R the densities of holes and electrons, respectively, in these centers, and in thermal equilibrium $n_R = N_R$.

The boundary condition appropriate to our problem is that the electric field intensity vanish at the hole injecting contact, $x = L$:

$$\mathcal{E} = 0 \quad \text{at} \quad x = L. \tag{5}$$

This boundary condition is discussed in Appendix A.

The two equations in (3) may be conveniently combined to give:

$$\frac{d}{dx}[(p-n)\mathcal{E}] = \frac{(a+1)r}{\mu_p} = \frac{(a+1)n}{\mu_p\tau_n} = \frac{(a+1)p}{\mu_p\tau_p}, \tag{6}$$

with $a = \mu_p/\mu_n$, $1/\tau_n = \langle v\sigma_n\rangle p_R$, and $1/\tau_p = \langle v\sigma_p\rangle n_R$. Equation (4), solved for p_R, gives

$$p_R = \frac{1}{1 + (n/p)\langle v\sigma_n\rangle/\langle v\sigma_p\rangle} N_R. \tag{7}$$

We discuss the different regimes separately.

A. Voltage Threshold

This is a low-injection-level regime in which say $n<N_R/4$. Throughout this regime the recombination centers are still largely filled with electrons, $n_R\approx N_R$, and therefore most of the injected holes are captured by the recombination centers, $p\ll n$. It then follows from (2) and (4) that

$$p\simeq n^2\langle v\sigma_n\rangle/N_R\langle v\sigma_p\rangle. \tag{8}$$

Here we have assumed that $n<N_R\langle v\sigma_p\rangle/\langle v\sigma_n\rangle$ as well as $n<N_R/4$.

Equation (1) is solved for $\mathcal{E}$, giving

$$\mathcal{E}=\frac{J}{en\mu_n}\frac{1}{(1+\theta)},\quad \theta=\frac{p}{n}a\simeq\frac{n\langle v\sigma_n\rangle}{N_R\langle v\sigma_p\rangle}a, \tag{9}$$

where we have also used Eq. (8).

Computing $d(n\mathcal{E})/dx$ from (9), and taking $p_R\simeq n$ in (4), we obtain from Eq. (3) the result:

$$\frac{dn}{n^2}=\frac{e\langle v\sigma_p\rangle N_R}{aJ}dx. \tag{10}$$

From Eq. (9) it follows that $dn/n^2=-(e\mu_n/J)d\mathcal{E}$, with a correction term of order θ^2 which we neglect, assuming $\theta\ll 1$. These results combine to give

$$d\mathcal{E}=-\frac{\langle v\sigma_p\rangle N_R}{\mu_p}dx. \tag{11}$$

Integrating this simple equation, subject to boundary condition (5), we obtain

$$\mathcal{E}=\frac{\langle v\sigma_p\rangle N_R}{\mu_p}(L-x). \tag{12}$$

Since the field does not depend on current J as a parameter, neither will the voltage. Consequently there must be a threshold voltage for double-injection current flow, with the current rising vertically at low currents. This threshold voltage $V_{\rm th}$ is given by

$$V_{\rm th}=\int_0^L\mathcal{E}dx=\frac{\langle v\sigma_p\rangle N_R L^2}{2\mu_p}=\frac{L^2}{2\mu_p\tau_{p,\rm low}},$$
$$\tau_{p,\rm low}=\frac{1}{\langle v\sigma_p\rangle N_R}. \tag{13}$$

Defining $t_{p,\rm th}=L^2/\mu_p V_{\rm th}$, the usual transit-time expression, (13) is equivalent to the result:

$$t_{p,\rm th}=2\tau_{p,\rm low}. \tag{14}$$

This is the result quoted in the Introduction.

Actually the assumption underlying the above derivation, namely $n<N_R/4$, must be invalid near the anode, since the boundary condition (5) requires that n be infinite at $x=L$. When this is taken into account the current rises steeply, but not vertically, near threshold. More precise expressions are obtained in Sec. III C and in Appendix A. The effects of the constraint at the injecting hole contact on this regime of current are discussed in Sec. III C; the effects of space charge are discussed in Appendix C.

B. High-Current, Semiconductor Regime

We discuss this regime for the case of greatest interest, $\langle v\sigma_p\rangle\gg\langle v\sigma_n\rangle$. Under this condition, at high-injection levels the recombination centers fill up with holes, $p_R\simeq N_R$. Consequently, to a high degree of approximation, the electron lifetime $\tau_{n,\rm high}$ is a constant independent of injection level: $\tau_{n,\rm high}\simeq 1/\langle v\sigma_n\rangle N_R$.

Since $p-n=-p_R$, from (2), and $p_R\simeq N_R$ independent of injection level and therefore of position, we may replace p-n in Eq. (6) by $-N_R$:

$$-N_R\frac{d\mathcal{E}}{dx}=\frac{(a+1)n}{\mu_p\tau_{n,\rm high}}. \tag{15}$$

At this point it may be noted that Eq. (15), after replacing N_R by n_T-p_T, is identical with Eq. (14) of (I), which characterizes the semiconductor (ohmic relaxation) regime of the injected plasma problem. The properties of that regime are worked out in detail in (I). These properties, replacing n_T-p_T by N_R, apply equally well to the present case. Nevertheless it is convenient, for purpose of a simplified analysis of the negative resistance regime in the following section, to derive here some simple properties of the semiconductor regime which were not explicitly spelled out in (I). We consider this regime in the limit $n\gg N_R$, in which case $p\approx n$ and the electrons and holes have a common lifetime $\tau_{\rm high}=\tau_{n,\rm high}$ as given above. In this limit n is related to $\mathcal{E}$, through (1), by $(a+1)n=J/e\mu_n\mathcal{E}$. Substitution of this result into Eq. (15) yields the equation

$$\mathcal{E}d\mathcal{E}=-Kdx,\quad K=\frac{J}{e\tau_{\rm high}\mu_n\mu_p N_R}. \tag{16}$$

Integration of this simple equation, subject to boundary condition (5), yields the results:

$$\mathcal{E}=[2K(L-x)]^{\frac{1}{2}},\quad n=\frac{J}{e\mu_n(a+1)[2K(L-x)]^{\frac{1}{2}}}. \tag{17}$$

The applied voltage V is given by $V=\int_0^L\mathcal{E}dx=(\frac{2}{3})(2KL^3)^{\frac{1}{2}}$. Squaring both sides of this equation, we obtain

$$K=(9/8)V^2/L^3\rightarrow J=(9/8)e\tau_{\rm high}\mu_n\mu_p N_R V^2/L^3. \tag{18}$$

This is precisely the result (A33b) of (I), except for the replacement of (n_T-p_T) by N_R. This square-law regime is a branch of the current-voltage characteristic (the solid curve) shown in Fig. 2. Note that this

derivation of the semiconductor regime in no way involves the space charge; indeed we have explicitly assumed neutrality. In Appendix E we show that, consistent with the above derivation, the space charge associated with the double injection is indeed small in this regime.

The correspondence between the semiconductor problem of (I) and the present problem is not a fortuitous accident. The transfer of electrons from the recombination centers to the conduction band that accompanies the capture of injected holes by the centers, constitutes an electronic conversion of the insulator into a semiconductor, as pointed out in the Introduction. With completion of the transfer, the electronic behavior of the insulator is thereafter precisely the same as the equivalent semiconductor. This includes the electronic behavior at still higher injection levels, namely at voltages V such that $Q=CV=\epsilon V/L>eN_RL$. This is the insulator regime of (I). In this regime the neutrality assumption *i*) is no longer a good approximation; space charge now plays an important role, changing the square-law current-voltage dependence to the cube-law dependence:

$$J=(125/18)\epsilon\tau_{\text{high}}\mu_n\mu_pV^3/L^5. \tag{19}$$

Equation (19) is just Eq. (A38b) of (I). The static dielectric constant ϵ makes its appearance in (19) via the Poisson equation which must be used in place of the neutrality relation (2). The cube-law regime is a branch of the current-voltage characteristic (the solid curve) shown in Fig. 2.

C. The Negative-Resistance Regime

The semiconductor regime discussed above is one in which $n>N_R$ throughout the solid. On the other hand, at low double-injection currents in the neighborhood of the voltage threshold, $n<N_R$ throughout the solid, in the simplified picture of Sec. A above. The negative resistance regime is one in which $n<N_R$ over part of the solid, extending out from the cathode, and $n>N_R$ over the remainder of the solid, up to the anode, with $n=N_R$ at a plane x_1, where $0<x_1<L$. (Strictly speaking, this also includes the low-current regime near the voltage threshold, but it proved convenient in Sec. A to simplify the analysis by taking $n<N_R$ everywhere.) The location of x_1 will, of course, depend on the current J.

A considerable simplification in the derivation of the properties of the negative-resistance regime is achieved if the "$n<N_R$" analysis of Sec. A is used up to the plane x_1, i.e., for $0<x<x_1$, and the "$n>N_R$" analysis of Sec. B beyond the plane x_1, i.e., for $x_1<x<L$. The corresponding solutions for n are joined continuously at the plane x_1. Since neither analysis is valid in the immediate vicinity of x_1, the entire derivation is only an approximate one. Accuracy of the results is discussed below.

In the "$n<N_R$" domain, $0<x<x_1$, Eq. (10) is integrated to give

$$\frac{1}{n}=\frac{1}{N_R}+\frac{eN_R\langle v\sigma_p\rangle}{aJ}(x_1-x). \tag{20}$$

The constant of integration in Eq. (20) is chosen to satisfy the boundary condition at $x=x_1$, namely $n=n_1=N_R$. From Eq. (9), taking $\theta=0$,

$$\mathcal{E}=\frac{J}{e\mu_nN_R}+\frac{N_R\langle v\sigma_p\rangle}{\mu_p}(x_1-x). \tag{21}$$

In the "$n>N_R$" domain, $x_1<x<L$, Eq. (16) is integrated to give the same result as previously, namely the relations (17), since the boundary condition (5) still holds. The location of the plane x_1 is determined by (17), taking $x=x_1$ and $n=N_R$. There is a discontinuity in $\mathcal{E}=\mathcal{E}_1$ at x_1, as calculated from the two domains: $\mathcal{E}_1=J/e\mu_nN_R$ from the "$n<N_R$" domain and $\mathcal{E}_1=J/e\mu_nN_R(1+a)$ from the "$n>N_R$" domain. For $a<1$ the discontinuity is not appreciable and accurate results are anticipated, whereas for $a>1$ the discontinuity is substantial, and a correspondingly large error is expected. These expectations are verified below.

The applied voltage V is given by: $V=\int_0^L\mathcal{E}dx=\int_0^{x_1}\mathcal{E}_<dx+\int_{x_1}^L\mathcal{E}_>dx$, where $\mathcal{E}_<$ and $\mathcal{E}_>$ denote $\mathcal{E}$ in the "$n<N_R$" and "$n>N_R$" domains, respectively. Substituting in these integrals for $\mathcal{E}_<$ from (21) and for $\mathcal{E}_>$ from (17), we obtain

$$V=\frac{J}{e\mu_nN_R}x_1+\frac{N_R\langle v\sigma_p\rangle}{2\mu_p}x_1^2+\left(\frac{8K}{9}\right)^{\frac{1}{2}}(L-x_1)^{\frac{3}{2}}. \tag{22}$$

With increasing current, x_1 moves toward the cathode, $x=0$. Finally, the negative-resistance regime terminates at that current J_M' where $x_1=0$. (The prime superscript is used to distinguish certain approximate results of this section from the more rigorous results of Appendices A and B.) Taking $n=N_R$, $x=x_1=0$ and $J=J_M'$ in (17), and using (16), we obtain

$$J_M'=\frac{2(a+1)^2}{a}\frac{eLN_R}{\tau_{\text{high}}}=\frac{2(a+1)^2}{a}eLN_R^2\langle v\sigma_n\rangle. \tag{23}$$

The voltage minimum V_M', i.e., the voltage at the high-current end, $J=J_M'$, of the negative-resistance region, is determined from (22), setting $x_1=0$ and using (16) and (23):

$$V_M'=\frac{4(a+1)}{3a}\frac{L^2N_R\langle v\sigma_n\rangle}{\mu_n}. \tag{24}$$

From (17) and (23) it follows that $x_1=L\{1-(J/J_M')\}$. Substituting this result into Eq. (22) and using

(23) and (24), Eq. (22) can finally be re-written as

$$V=V_M'\left[\tfrac{3}{2}(a+1)\left(\frac{J}{J_M'}\right)\left(1-\frac{J}{J_M'}\right)+\left(\frac{J}{J_M'}\right)^2\right]+V_{\rm th}\left(1-\frac{J}{J_M'}\right)^2, \quad (25)$$

with $V_{\rm th}$ given by (13). This simple, approximate result replaces the much more complicated, implicit relation between V and J given by the rigorous solution, Eqs. (A14) and (A15) of Appendix A. Of course, by the very nature of the derivation, the use of Eq. (25) is restricted to currents $J\leq J_M'$.

Equation (25) clearly exhibits the negative resistance, V, decreasing monotonically as J increases from 0 to approximately J_M'. Further, the derivation clarifies the physical origin of the negative resistance. The "$n>N_R$" domain, $x>x_1$, is a region of low resistance, the "$n<N_R$" domain, $x<x_1$, one of high resistance. Therefore the bulk of the applied voltage appears across the "$n<N_R$" domain. As current increases, the "$n>N_R$" domain grows at the expense of the "$n<N_R$" domain, i.e., x_1 moves toward the cathode: $x_1=L\{1-(J/J_M')\}$. The resulting small increase in the resistance of the "$n>N_R$" domain, due to its increased width, is more than offset by the resulting decrease in the resistance of the "$n<N_R$" domain due to its decreased width—hence the negative resistance. Note that the voltage threshold is also contained in this analysis; as $J\rightarrow 0$ in Eq. (25), $V\rightarrow V_{\rm th}$ as given by (13).

The accuracy of Eq. (25) can be assessed by comparing J_M', V_M', as given by (23) and (24), respectively, to the rigorously derived J_M, V_M as given by (B10) and (A6), and (B9) and (A8), respectively, in the Appendices:

$$\frac{J_M'}{J_M}=\frac{2(a+1)^2[a-\ln(1+a)]}{a^2}, \qquad \frac{V_M'}{V_M}=\frac{4(a+1)[a-\ln(1+a)]^2}{3a[\frac{1}{2}a^2-a+\ln(1+a)]}. \quad (26)$$

For $a\ll 1$, both ratios in (26) approach unity, and the approximate result (25) is very good indeed. For $a=1$, $J_M'/J_M=2.4$ and $V_M'/V_M=1.3$. For $a\gg 1$, $J_M'/J_M=2a$ and $V_M'/V_M=2.7$; the approximate result is here in serious error, as anticipated earlier.

An important feature of the negative-resistance domain is the relatively small change of current with voltage over most of the domain. We first note, from (13) and (24), that $V_{\rm th}/V_M'=3\langle v\sigma_p\rangle/8(a+1)\langle v\sigma_n\rangle$, for $\langle v\sigma_p\rangle\gg\langle v\sigma_n\rangle$, so that $V_{\rm th}\gg V_M'$. Therefore the dependence of V on J, Eq. (25), is completely dominated by the term $V_{\rm th}[1-(J/J_M')]^2$, except for J very close to J_M'. For example, for $J=0.2J_M'$, $V\simeq 0.64V_{\rm th}$. Thus, with current varying by a factor of only 5 from $0.2J_M'$ to J_M', the voltage varies over the much larger range from $0.64V_{\rm th}$ down to V_M'. The rigorous solution of Appendix B yields about the same corresponding voltage swing, namely from $0.67V_{\rm th}$ down to V_M.

If a more realistic model of the hole injecting contact is employed, in which the electron current is blocked at this contact and the hole current is predominantly a diffusion current in the immediate vicinity of the contact, then it may be shown,[7] via an approximate analysis along the above lines, that Eq. (25) is modified as follows:

$$V=\left(1+\frac{L_a}{L}\right)^2 V_{(25)}-\left(\frac{L_a}{L}\right)^{\frac{3}{2}}\left(1+\frac{L_a}{L}\right)^{\frac{1}{2}} V_M'\left(\frac{J}{J_M'}\right)^{\frac{1}{2}}, \quad (27)$$

where $V_{(25)}$ denotes the right-hand side of Eq. (25) and L_a is the high-lifetime, ambipolar diffusion length, $L_a^2=2D_nD_p\tau_{\rm high}/(D_n+D_p)$. In Eq. (27), L is no longer the total spacing between cathode and anode; rather it is the length of solid over which the field-driven currents are dominant. Likewise V is the potential drop over the latter length. Clearly, for $L\gg L_a$, $V\simeq V_{(25)}$; further, there is then small error in taking L as the total crystal length and V as the total applied voltage. Equation (27) is valid, approximately, down to current $J''=J_M'/(2L/L_a+1)$. For $J<J''$, there is no "$n>N_R$" region over the distance L, whereas the existence of such a domain was assumed in the derivation. For $a>1$, Eq. (27) is in error in the same way that Eq. (25) is, and for the same reason. Nonetheless, this analysis establishes the important result that, for $L\gg L_a$, the negative-resistance and threshold regimes are adequately described through use of assumptions (ii) and (iii) of Sec. II, at least down to currents of order $J'\approx L_aJ_M/2L$. Indeed, the breakdown of our double-injection theory, for $L\gg L_a$, is likely to occur, going *down* in current, at a higher current than J' due to space-charge effects, i.e., due to the failure of the neutrality assumption (i). This is discussed in Appendix C.

IV. RELATIONSHIP TO EXPERIMENTS

The theory presented in this paper has been worked out in detail only for the simplest model, that in which there is a single set of defect states in the forbidden gap, completely filled with electrons in thermal equilibrium and functioning as recombination centers under double-injection excitation. The negative resistance in the predicted current-voltage characteristic (Fig. 2) has its origin in the unequal capture cross sections of the defect states, $\sigma_p\gg\sigma_n$. This condition corresponds to acceptor-like behavior of the defect state, where the state is negatively charged when occupied by an electron. We may envisage charge neutrality maintained in this case through the presence in the insulator of an equal number of shallow donors which play no further role in the dynamical behavior of the insulator. In the

current state of the art there is no semiconductor or insulator available whose electronic structure is dominated by such a set of *simple* (single), deep-lying, acceptor levels. On the other hand there are a number of known deep-lying, *double* acceptors in Ge, e.g., Ni, Fe, Mn, etc., and Au is a known triple acceptor. Highly purified Ge, doped with one of these elements and suitably compensated with shallow donors, is a high-resistance insulator at liquid nitrogen temperature. Holes injected into such a sample of Ge would have a low lifetime at low-injection levels and a higher lifetime at higher injection levels where the acceptor levels had been depopulated with electrons in favor of holes. The magnitude of the change in hole lifetime would depend directly on the initial electron occupancy of the deep-lying acceptor states and this occupancy, in turn, is determined by, and therefore controllable through, shallow-donor compensation during crystal growth. Since suitable injecting contacts for both electrons and holes are available, it would appear that Ge and perhaps also Si, for similar reasons, are promising materials with which to look for a quantitative check of our theory. Such experimental studies remain to be carried out. However there do exist reports in the literature, describing phenomena observed in high-resistance Ge at liquid nitrogen temperature, which appear to confirm the broad picture of double injection presented in this paper.

Before giving these references we note that a current-voltage characteristic such as that of Fig. 2, exhibiting a current-controlled negative resistance, might ordinarily be revealed experimentally through either of two striking effects: spontaneous oscillations under application of an appropriate dc voltage, or an apparent breakdown at some critical voltage followed by a marked hysteresis in the current going down in voltage after the "breakdown." The oscillations would be due to the dominance of the negative resistance over the external circuit resistance. If the oscillations are suppressed by the external circuit, then the "breakdown" would likely be observed when, at some critical voltage the current jumps from the lower branch of the characteristic in Fig. 2 to some point on the upper branch, the semiconductor regime. Following this "breakdown," the current will follow this upper branch as voltage is lowered down to some other critical voltage, where the current will drop sharply down to the lower branch, extinguishing the "breakdown."

Both types of phenomena have been observed in high-resistivity Ge at liquid nitrogen temperature, spontaneous oscillations in Au-doped Ge,[9] and "breakdown" and hysteresis in Fe-doped Ge.[10] In neither case are sufficient data furnished by the authors to make a conclusive, quantitative analysis of the observed effects in the light of our present theory. It is interesting to note, however, that in both cases the authors[3,10] attribute the observed phenomena to double injection with hole capture by the deep-lying acceptors. In the Fe-doped Ge study, for example, the "breakdown" was quenched by infrared excitation correlating well with the known infrared quenching of bandgap-excited photoconductivity in n-type Fe-doped Ge. This optical quenching is due to optical release of the holes trapped at the Fe sites into the valence band.

Hysteresis in the dc current-voltage characteristic is a common observation in insulator studies. This effect has been reported by Nicoll[11] for CdSe powders, and further data and analysis were published more recently by Bube.[12] He attributes the effect to double injection accompanied by changing carrier lifetime; however for this material his model is necessarily more complicated than the simple one of Fig. 1 of this paper. Smith[13] has frequently observed a hysteresis associated with steeply rising, and falling, double-injection currents in single crystals of CdS. In his experiments the double injection is verified directly by observation of the bandgap light emitted through the radiative recombination of electrons and holes.

Finally we may note that there are also extensive observations of spontaneous oscillations in photocurrent, with dc applied voltage, in single crystals of insulators: CdS,[14] ZnSe,[15] and high-resistivity GaAs.[16]

ACKNOWLEDGMENTS

The author is pleased to express his indebtedness to Dr. A. Rose for extensive conversations which have helped to clarify many aspects of this work. The simple picture of the negative-resistance regime underlying the argument of Sec. III C is due to Dr. D. O. North. Dr. R. H. Bube supplied the many references to his own and other work on insulators bearing on the subject matter of this paper.

APPENDIX A

Formal Theory for a Single Set of Recombination Centers, Completely Occupied by Electrons in Thermal Equilibrium

The equations defining the mathematical problem are Eqs. (1)–(4) of Sec. III. The key choice of independent variable for our solution is the free-carrier density ratio n/p, which we denote by u.

---

[9] A. A. Lebedev, V. I. Stafeev, and V. M. Tuchkevich, J. Tech. Phys. (U.S.S.R.) **26**, 1419 (1956) [translation: Soviet Phys.—Tech. Phys. **1**, 2071 (1957)].

[10] W. W. Tyler, Phys. Rev. **96**, 226 (1954).

[11] F. H. Nicoll, RCA Rev. **19**, 77 (1958).

[12] R. H. Bube, J. Appl. Phys. **31**, 2239 (1960).

[13] R. W. Smith, Phys. Rev. **105**, 900 (1957), and private communication.

[14] S. H. Liebson, J. Electrochem. Soc. **102**, 529 (1955); R. H. Bube and L. Barton, RCA Rev. **20**, 564 (1959).

[15] R. H. Bube and E. L. Lind, Phys. Rev. **110**, 1040 (1958).

[16] J. Blanc, R. H. Bube, and H. E. MacDonald, J. Appl. Phys. (to be published).

As a function of u, p_R as given by (7) is re-written as

$$p_R=aN_R/(\beta u+a), \quad u=n/p, \quad a=\mu_p/\mu_n, \quad \beta=a\langle v\sigma_n\rangle/\langle v\sigma_p\rangle. \tag{A1}$$

Equations (2) and (A1) give for p and n:

$$p=aN_R/(u-1)(\beta u+a), \quad n=pu=aN_Ru/(u-1)(\beta u+a). \tag{A2}$$

Equation (1) gives for $\mathcal{E}$,

$$\mathcal{E}=J/e\mu_n p(u+a). \tag{A3a}$$

For the model of Fig. 1, p is given by (A2), so that (A3a) becomes

$$\mathcal{E}=\frac{J}{ea\mu_nN_R}h(u), \quad h(u)=\frac{(u-1)(\beta u+a)}{(u+a)}. \tag{A3b}$$

The recombination equation most useful for the present analysis is Eq. (6). Substituting for p_R, p, n, and $\mathcal{E}$ in this equation the corresponding expressions in u as given by (A1), (A2), and (A3b), and carrying out the indicated differentiation, (6) reduces to

$$f(u)du=-\frac{aeN_R^2\langle v\sigma_n\rangle}{J}dx, \quad f(u)=\frac{(u-1)(\beta u+a)^2}{u(u+a)^2}. \tag{A4}$$

The boundary condition used in solving the problem is that stated in (5), namely, $\mathcal{E}=0$ at the hole-injecting contact, $x=L$. This choice of boundary condition requires some discussion. In the more rigorous formulation of the problem, in which the space charge is correctly accounted for by the Poisson equation, $n-p-p_R=(\epsilon/e)d\mathcal{E}/dx$ in place of Eq. (2), there are two boundary conditions on the problem. For a theory of field-driven, double-injection currents with non-constraining contacts the appropriate boundary conditions are those employed in (I), namely the vanishing of the electric field intensity at both injection contacts: $\mathcal{E}=0$ at $x=0$ and at $x=L$. In assuming neutrality Eq. (2), and thereby dropping the $d\mathcal{E}/dx$ term, the order of the differential equation determining $\mathcal{E}$, not written down here, is reduced by one, and therefore it is necessary to drop one boundary condition. We know from the discussion in Sec. III B that at high injection levels the present problem is equivalent to that of double injection into an n-type semiconductor. The analytical solution to this problem in (I) shows that in the semiconductor (Ohmic relaxation) regime the field intensity rises smoothly from zero at the hole-injecting contact, the anode, to a maximum extremely close to the cathode. This is also the solution one obtains by assuming charge neutrality and dropping the condition $\mathcal{E}=0$ at the cathode. (In this case the field is a maximum right at the cathode instead of very close to the cathode.) Indeed, following this prescription, in Sec. III B we explicitly re-derived the $J\propto V^2$ current-voltage characteristic for the semiconductor regime. Actually this same procedure was implicitly followed in the cruder argument of (I), Eqs. (14) and (15).

The fact that the boundary condition (5) is known to be appropriate for the semiconductor regime of our problem does not automatically establish that it is proper at lower currents. Here we appeal to a physical argument. The reason that, in the semiconductor regime of the problem of the plasma injected into an n-type semiconductor, the field rises to a maximum near the electron-injecting contact, the cathode, is that the holes are the difficult carriers to transport across the solid. Volume-distributed, thermally generated electrons are already available for charge neutralization by Ohmic relaxation. It is obvious that, in the present problem also, the holes are the difficult carriers to transport across the solid, at all current levels below the insulator regime. For this reason we expect that the field intensity must grow with increasing distance from the anode.

Returning to the mathematical development, it is evident from (A3b) that $\mathcal{E}=0$ corresponds to $u=1$. Integrating Eq. (A4) from $x=0$ to $x=L$, and letting u_0 denote the value of u at $x=0$, we obtain:

$$\mathcal{J}=J/J_1=1/F(u_0), \quad J_1=aeN_R^2L\langle v\sigma_n\rangle, \quad F(u_0)=\int_1^{u_0}f(u)du. \tag{A5}$$

More generally, at any plane x between cathode and anode, the relationship between x and u is

$$x/L=1-\mathcal{J}F(u)=1-F(u)/F(u_0). \tag{A6}$$

From (A3b) and (A4) it follows that the applied voltage, $V=\int_0^L\mathcal{E}dx$, can be written as

$$V=\frac{J^2}{a^2e^2\mu_nN_R^3\langle v\sigma_n\rangle}\int_1^{u_0}h(u)f(u)du. \tag{A7}$$

Combining (A7) with (A5), we obtain

$$\mathcal{V}=V/V_1=\mathcal{J}^2G(u_0)=G(u_0)/F^2(u_0), \quad V_1=L^2N_R\langle v\sigma_n\rangle/\mu_n, \quad G(u_0)=\int_1^{u_0}g(u)du, \quad g(u)=h(u)f(u)=\frac{(u-1)^2(\beta u+a)^3}{u(u+a)^3}. \tag{A8}$$

More generally, the potential $V(x)$ at any plane x between cathode and anode is given by

$$\frac{V(x)}{V}=1-\frac{\mathcal{J}^2V_1}{V}G(u)=1-\frac{G(u)}{G(u_0)}. \tag{A9}$$

The essential equations for the determination of the current-voltage characteristic are (A5) and (A8). The relationship between current and voltage is an implicit one, via the parameter $u_0=(n/p)_{x=0}$. Except over certain limiting ranges of voltage it is generally impossible to eliminate the parameter u_0 in terms of known, elementary functions and thereby obtain J directly as an analytically describable function of voltage. This same situation characterizes the solutions to space-charge-limited current flow problems, for both one-carrier flow[17] and two-carrier flow [the Appendix of (I)].

The integrals in (A5), with $f(u)$ given by (A4), and (A8), are evaluated by the well-known technique of expanding the integrands in partial fractions:

$$f(u)=A+\frac{B}{u}+\frac{C}{u+a}+\frac{D}{(u+a)^2}, \tag{A10}$$

$$A=\beta^2,\quad B=-1,\quad C=(1-\beta)\{1+(2a+1)\beta\},\quad D=a(a+1)(1-\beta)^2, \tag{A11}$$

$$g(u)=Mu+N+\frac{P}{u}+\frac{Q}{u+a}+\frac{R}{(u+a)^2}+\frac{S}{(u+a)^3}, \tag{A12}$$

$$M=\beta^3,\quad N=\beta^2\{3a-(3a+2)\beta\},\quad P=1,\quad Q=-(1-\beta)\{1-(3a^2-1)\beta+(6a^2+6a+1)\beta^2\},\quad R=a(a+1)(1-\beta)^2\{(a-1)-2(2a+1)\beta\},\quad S=-a^2(a+1)^2(1-\beta)^3. \tag{A13}$$

Substitution of (A10) into (A5) and (A12) into (A8) give the final results:

$$\frac{1}{\mathcal{J}}=F(u_0)=A(u_0-1)+B\ln u_0+C\frac{u_0+a}{1+a}+D\left\{\frac{1}{1+a}-\frac{1}{u_0+a}\right\}, \tag{A14}$$

$$\frac{\mathcal{V}}{\mathcal{J}^2}=G(u_0)=\tfrac{1}{2}M(u_0^2-1)+N(u_0-1)+P\ln u_0+Q\ln\frac{u_0+a}{1+a}+R\left\{\frac{1}{1+a}-\frac{1}{u_0+a}\right\}+\tfrac{1}{2}S\left\{\frac{1}{(1+a)^2}-\frac{1}{(u_0+a)^2}\right\}. \tag{A15}$$

Strictly speaking, the results obtained up to this point are all that we need to obtain a complete picture, within our mathematical model, of the current flow, including potential, field, and density distributions. For given a and β we need only plot numerically all quantities of interest versus u using the expressions derived above. Indeed it is planned to carry out this numerical program in the future with the aid of a digital computer. However, it is obviously desirable to pursue further the analytical study of the problem in order to gain further insight into the solution. The simplified arguments of Sec. III are useful in illuminating the physics of double injection. It remains to put these results on a firm footing, to establish their limitations [as for example in (26)], and to provide a more powerful analytical technique which will be applicable to a more general class of double-injection problems. The following discussion is aimed toward these ends.

First we note that for $u>1$, $f(u)$, $g(u)$, and $h(u)$ are all positive, as is obvious from (A4), (A8), and (A3b), respectively, by inspection. From this it follows from their definitions as integrals, (A5) and (A8), respectively, that both $F(u_0)$ and $G(u_0)$ increase monotonically with increasing u_0, from zero at $u_0=1$ to infinity at $u_0=\infty$. This last property is verified by inspection of the leading (largest) term in the expressions for $F(u_0)$ and $G(u_0)$, (A14) and (A15), respectively, at large u. It is now obvious from (A5) that:

$\mathcal{J}$ is a monotonic decreasing function of u_0 in the interval

$$1\leq u_0\leq\infty,\quad \text{with}\quad \mathcal{J}=\infty\quad \text{at}\quad u_0=1\quad \text{and}\quad \mathcal{J}=0\quad \text{at}\quad u_0=\infty. \tag{A16}$$

$\mathcal{V}$, being the ratio of two monotonic functions, $\mathcal{V}=G(u_0)/F^2(u_0)$ from (A8), need not have such simple behavior as a function of u_0; in particular $\mathcal{V}$ need not be monotonic and, further, it need not approach zero as u_0 approaches infinity. Mathematically, this is the source of the most interesting physical features in the solution to our two-carrier problem.

We first study the behavior of $\mathcal{V}$ at very large u_0: $u_0\gg1$, $u_0\gg a$, $u_0\gg\beta$. The asymptotic expansions (A14), (A15), respectively, in this large-u_0 domain are:

$$F(u_0)\sim Au_0+(B+C)\ln u_0+\left\{\frac{D}{1+a}-C\ln(1+a)-A\right\}_F+(aC-D)\frac{1}{u_0}+a(D-\tfrac{1}{2}aC)\frac{1}{u_0^2}+\cdots, \tag{A17}$$

$$G(u_0)\sim\tfrac{1}{2}Mu_0^2+Nu_0+(P+Q)\ln u_0+\left\{\frac{S}{2(1+a)^2}+\frac{R}{1+a}-Q\ln(1+a)-N-\tfrac{1}{2}M\right\}_G+(aQ-R)\frac{1}{u_0}+(aR-\tfrac{1}{2}S-\tfrac{1}{2}a^2Q)\frac{1}{u_0^2}+\cdots. \tag{A18}$$

The asymptotic expansion for the voltage $\mathcal{V}$ is,

[17] M. A. Lampert, Phys. Rev. 103, 1648 (1956).

from (A8), (A17), and (A18),

$$\mathcal{U}\sim\frac{1}{2\beta}\left\{1-\frac{1}{Au_0}\left[2(B+C)\ln u_0 +2\{\ \}_F-\frac{2NA}{M}\right]+\cdots\right\}. \quad \text{(A19)}$$

In writing (A19) we have replaced $M/2A^2$ by $1/2\beta$, using (A11) and (A13). The bracket $\{\ \}_F$ appearing in (A19) is given in (A17). The expression (A19) directly yields several major results:

(i) There is a threshold voltage,
$$\mathcal{U}_{\rm th}=V_{\rm th}/V_1=1/2\beta \quad \text{or} \quad V_{\rm th}=V_1/2\beta,$$
for two-carrier injection. (A20)

That is, as current $\mathcal{J}$ approaches zero (u_0 approaches infinity) voltage $\mathcal{U}$ approaches $1/2\beta$ rather than zero.

(ii) The two-carrier current $\mathcal{J}$ rises steeply with voltage near threshold $\mathcal{U}_{\rm th}$. (A21)

Over the entire range of currents, $\mathcal{J}=\mathcal{J}(u_0)$, for which the term $[\ ]/Au_0$ in (A19) is small compared to unity, the corresponding voltage change is relatively small.

The results (i) and (ii) confirm the conclusions of Sec. III A, obtained by a simplified analysis, the expression $V_1/2\beta$ of (A20) being identical with the expression $L^2/2\mu_p\tau_{p,\rm low}$ of (13). Both results rest on the neutrality approximation, which breaks down at low currents as discussed in Appendix C.

(iii) For appropriate values of a and β, e.g., for $B+C>0$, i.e., for $\beta<2a/(2a+1)$, there is a negative-resistance region in the two-carrier current-voltage characteristic. (A22)

This follows from (A19) noting that for u_0 sufficiently large the term $2(B+C)\ln u_0$ is the dominant term inside the square brackets, and that for $B+C>0$, $\mathcal{U}$ decreases as u_0 decreases. Also we have used the relation $B+C=\beta\{2a-(2a+1)\beta\}$ from (A11).

Of major interest is the voltage range over which there is negative resistance. In this connection it is to be noted that for $\beta\ll 1$, $\beta\ll a$ the term $2\{\ \}_F$ inside the square brackets in (A19) becomes dominant, with decreasing u_0, over the term $2(B+C)\ln u_0$ at very large u_0. However, under the same conditions, $2\{\ \}_F$ is also positive and the net result is that there is a large voltage range of negative resistance for β very small, as shown in Appendix B.

Returning to the general theory, we expect from (A19) that, under conditions where there is a negative resistance, the plot of $\mathcal{U}$ versus u_0 has the form shown schematically in Fig. 3. (Since $\mathcal{J}$ simply decreases monotonically from ∞ at $u_0=1$ to 0 at $u_0=\infty$, we do not include a plot of $\mathcal{J}$ in Fig. 3.) The negative resistance region in u_0 extends from u_M, that value of u_0 at which $\mathcal{U}$ has a minimum, to $u_0=\infty$. The voltage region of

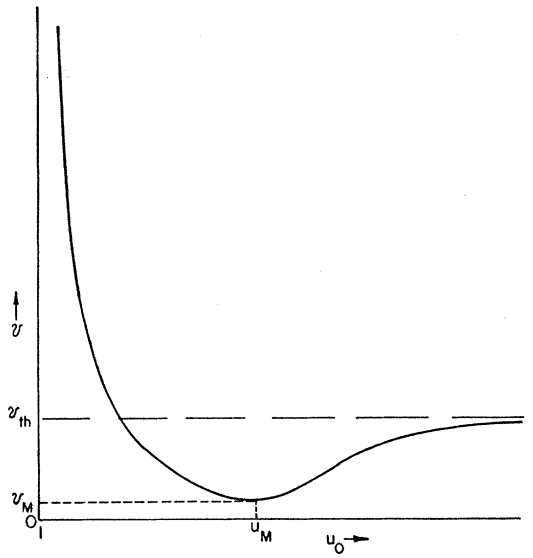

FIG. 3. Schematic plot of the dimensionless voltage variable $\mathcal{U}$ versus the parameter $u_0=n_0/p_0$ for the case that $\langle v\sigma_p\rangle\gg\langle v\sigma_n\rangle$. $\mathcal{U}$ asymptotically approaches $\mathcal{U}_{\rm th}$ as $u_0\to\infty$. The negative-resistance region in voltage lies between $\mathcal{U}_{\rm th}$ and $\mathcal{U}_M$, where $\mathcal{U}_M$ is the minimum value of $\mathcal{U}$.

negative resistance extends from $\mathcal{U}_{\rm th}$ to $\mathcal{U}_M$, the minimum value of $\mathcal{U}$. From (A8), $\mathcal{U}$ has a minimum, i.e., $d\mathcal{U}/du_0=0$, at that u_0 for which

$$F(u_0)\frac{dG(u_0)}{du_0}-2G(u_0)\frac{dF(u_0)}{du_0}=0. \quad \text{(A23)}$$

Noting from (A5) and (A8) that $dF(u_0)/du_0=f(u_0)$ and $dG(u_0)/du_0=g(u_0)$, and $g(u_0)/f(u_0)=h(u_0)$, Eq. (A23) can be written:

$$\tfrac{1}{2}h(u_0)F(u_0)-G(u_0)=0. \quad \text{(A24)}$$

This equation relates three variables a, β, and u_0. From the experimental point of view, a and β are determined by the physical properties of the solid, and therefore Eq. (A24) is to be solved for u_0. From the mathematical point of view, Eq. (A24) is a transcendental equation in the variable u_0 which is difficult to analyze. On the other hand, Eq. (A24) is "only" a cubic equation in the variable β. We therefore find it convenient here to regard a and u_0 as given, and to solve Eq. (A24) for β. After lengthy algebraic manipulation can rewrite Eq. (A24) as

$$\mathcal{B}(\beta)=B_3\beta^3+B_2\beta^2+B_1\beta+B_0=0, \quad \text{(A25)}$$

$$B_3=\frac{u_0-1}{u_0+a}\{(2a+1)(u_0+3a)+au_0\} -\left\{\frac{u_0(u_0-1)(2a+1)}{2(u_0+a)} +6a^2+6a+1\right\}\ln\frac{u_0+a}{1+a}, \quad \text{(A26a)}$$

$$B_2=-\frac{a(u_0-1)(7u_0+18a+5)}{2(u_0+a)}+\left\{\frac{a(u_0-1)(2u_0-2a-1)}{2(u_0+a)}+3a(3a+2)\right\}\ln\frac{u_0+a}{1+a}, \quad \text{(A26b)}$$

$$B_1=\frac{a(u_0-1)(u_0+6a-1)}{2(u_0+a)}-\frac{2a^2(u_0+2a)}{u_0+a}\ln\frac{u_0+a}{1+a}+\frac{u_0(u_0-1)}{2(u_0+a)}\ln\frac{u_0+a}{u_0(1+a)}, \quad \text{(A26c)}$$

$$B_0=\frac{a(u_0-1)}{u_0+a}+\frac{u_0(a+2)+a}{2(u_0+a)}\ln\frac{u_0+a}{u_0(1+a)}. \quad \text{(A26d)}$$

We shall not attempt here a general study of the roots of Eq. (25). The case of greatest interest to us, that in which there is a prominent negative resistance, corresponds to $\beta\ll1$, as already noted. The important root of Eq. (A25) for this case is obtained in Appendix B.

We complete the general analysis with a brief study of the high-current region, corresponding to u_0 near unity, or $u_0=1+\epsilon$ with $\epsilon\ll1$. Systematic expansions of (A17) and (A18), respectively, give for this domain:

$$F(u_0)=\mathfrak{A}_2\epsilon^2-\mathfrak{A}_3\epsilon^3+\mathfrak{A}_4\epsilon^4-\cdots, \quad \text{(A27)}$$

$$G(u_0)=\mathfrak{B}_3\epsilon^3-\mathfrak{B}_4\epsilon^4+\mathfrak{B}_5\epsilon^5-\cdots, \quad \text{(A28)}$$

$$\mathfrak{A}_2=\frac{(a+\beta)^3}{2(1+a)^2}, \quad \text{(A29a)}$$

$$\mathfrak{A}_3=\frac{1}{3(1+a)^3}[a^2(3+a)+4a\beta-(a-1)\beta^2], \quad \text{(A29b)}$$

$$\mathfrak{A}_4=\frac{1}{4(1+a)^4}[a^2(a^2+4a+6)+6a\beta-2a\beta^2+\beta^2], \quad \text{(A29c)}$$

$$\mathfrak{B}_3=\frac{(a+\beta)^3}{3(1+a)^3}, \quad \text{(A30a)}$$

$$\mathfrak{B}_4=\frac{1}{4(1+a)^4}[a^3(a+4)+a\beta(3a+2\beta)(3-\beta)+\beta^3], \quad \text{(A30b)}$$

$$\mathfrak{B}_5=\frac{1}{5(1+a)^5}[a^3(a^2+5a+10)+18a^2\beta-9a(a-1)\beta^2+(a^2-4a+1)\beta^3]. \quad \text{(A30c)}$$

Retaining only the lowest order terms in (A27) and (A28), respectively, and eliminating ϵ, the final result is:

$$\mathcal{J}=(9/8)\mathcal{V}^2\rightarrow J=(9/8)e\tau\mu_n\mu_pN_RV^2/L^3,\qquad \tau=1/(N_R\langle v\sigma_n\rangle)=\tau_{\text{high}}. \quad \text{(A31)}$$

This confirms the solution (18) of Sec. III B. From (A2), $n_0-p_0=p_0(u_0-1)=aN_R/(\beta u_0+a)$; thus, for $\beta u_0\ll a$ or $\langle v\sigma_n\rangle u_0/\langle v\sigma_p\rangle\approx\langle v\sigma_n/\langle v\sigma_p\rangle\ll1$, $n_0-p_0\simeq N_R$. This was a simplifying relation used in deriving (18).

APPENDIX B

Solution for Small β

The coefficients A, etc., of (A11), and M, etc., of (A13) reduce to

$$A=\beta^2,\quad B=-1,\quad C=1+2a\beta,\quad D=a(a+1), \quad \text{(B1)}$$

$$M=\beta^3,\quad N=3a\beta^2,\quad P=1,\quad Q=-1+3a^2\beta,\quad R=a(a+1)(a-1),\quad S=-a^2(a+1)^2. \quad \text{(B2)}$$

We first find the appropriate root of Eq. (A25) which yields the voltage minimum. As will be shown, this root corresponds to large u_0. Taking $u_0\gg1$, $u_0\gg a$, the coefficients in the cubic equation (A25), namely the B_i given in (A26a)–(A26d), reduce to

$$B_3\sim-u_0[\tfrac{1}{2}(2a+1)\ln u_0-(3a+1)], \quad \text{(B3a)}$$

$$B_2\sim au_0(\ln u_0-\tfrac{7}{2}), \quad \text{(B3b)}$$

$$B_1\sim\tfrac{1}{2}u_0[a-\ln(1+a)], \quad \text{(B3c)}$$

$$B_0\sim a-\tfrac{1}{2}(a+2)\ln(1+a). \quad \text{(B3d)}$$

We observe that B_0 is negative for all $a>0$, B_1 is positive for all $a>0$, and $-B_0\ll B_1$. Under these conditions it is readily seen that a physically significant solution, $\beta>0$, to Eq. (A25) is

$$\beta\simeq|B_0|/B_1\simeq\mathfrak{A}(a)/u_0\rightarrow u_0\simeq\mathfrak{A}(a)/\beta,\qquad u_0\gg a,\ 1 \quad \text{(B4)}$$

$$\mathfrak{A}(a)=\frac{(a+2)\ln(1+a)-2a}{a-\ln(1+a)}.$$

For small a, $a\ll1$, $\mathfrak{A}(a)\simeq a/3$; for $a=1$, $\mathfrak{A}(a)=0.257$; and for large a, $a\gg1$, $\mathfrak{A}(a)\simeq\ln a-2$.

More accurately, β is given by (B4) multiplied by the correction factor, $1-B_2|B_0|/B_1^2$. The second term in the correction factor is equal to $2(\ln a)(\ln u_0-\frac{7}{2})/u_0$, and we are clearly justified in neglecting it.

Returning to the physical problem, a and β given, we see that, for $\beta\ll1$, there will be a root u_M of Eq. (A24), and correspondingly a voltage minimum $\mathcal{V}_M$, at $u_M\simeq\mathfrak{A}(a)/\beta$, as given in (B4).

It remains to compute $\mathcal{V}_M$. Using (B1) and (B2), we first rewrite the asymptotic expansions (A17) and (A18):

$$F(u_0)\sim\beta^2u_0+2a\beta\ln u_0+[a-\ln(1+a)]-a^2/u_0+a^2(2a+1)/u_0^2+\cdots, \quad \text{(B5)}$$

$$G(u_0)\sim\tfrac{1}{2}\beta^3u_0^2+3a\beta^2u_0+3a^2\beta\ln u_0+\tfrac{1}{2}[a^2-2a+2\ln(1+a)]-a^3/u_0+a^3(3a+2)/2u_0^2+\cdots. \quad \text{(B6)}$$

For u_0 given by (B4), the constant (u_0-independent) terms in (B5) and (B6) are dominant, so that

$$F(u_M)\sim a-\ln(1+a), \tag{B7}$$

$$G(u_M)\sim\tfrac{1}{2}[a^2-2a+2\ln(1+a)]. \tag{B8}$$

From (A8), (A6), (B7), and (B8) we have finally:

$$\mathcal{U}_M=\frac{G(u_M)}{F^2(u_M)}=\mathcal{U}_M(a)=\frac{\frac{1}{2}a^2-a+\ln(1+a)}{[a-\ln(1+a)]^2}, \tag{B9}$$

$$\mathcal{J}_M=\frac{1}{F(u_M)}=\mathcal{J}_M(a)=\frac{1}{a-\ln(1+a)}. \tag{B10}$$

The negative-resistance region lies between $\mathcal{U}_{\rm th}$, given by (A20), and $\mathcal{U}_M$ given by (B9). For large a, $a\gg1$ but $\beta a\ll1$, $\mathcal{U}_M(a)\simeq\frac{1}{2}$ and $\mathcal{J}_M(a)\simeq1/a$; for $a=1$ $\mathcal{U}_M=2.0$ and $\mathcal{J}_M=3.3$; and for small a, $a\ll1$, $\mathcal{U}_M(a)\simeq4/3a$ and $\mathcal{J}_M(a)\simeq2/a^2$. We see that for $\beta\ll1$, $\beta\ll a$ the voltage range of negative resistance is quite large. Over most of this range the corresponding change of current is relatively much smaller. To see this we first note, from (B5), that $F(u_0)$ is appreciably larger than $F(u_M)$ only if β^2u_0 exceeds $a-\ln(1+a)$. For convenience define u_1 by: $u_1=4\{a-\ln(1+a)\}/\beta^2$. For $u_0=u_1$, it is easily seen that $G(u_1)\simeq\beta^3u_1^2/2$, so that $\mathcal{U}_1=G(u_1)/F^2(u_1)\simeq2\mathcal{U}_{\rm th}/3$. Further, $\mathcal{J}_1=\mathcal{J}(u_1)=1/5\{a-\ln(1+a)\}=0.2\mathcal{J}_M$ from (B10). Thus, as u_0 varies from u_1 down to u_M, the voltage decreases from $2\mathcal{U}_{\rm th}/3$ down to $\mathcal{U}_M$, whereas the current increase is only from $0.2\mathcal{J}_M$ to $\mathcal{J}_M$. Thus, with the exception of the mobility-factor correction required for $a\geq1$, discussed in connection with (26), the results of Sec. III C are confirmed.

For the electron density at the cathode we have, from (A2),

$$u_0\gg1: \qquad n_0\sim\frac{a}{\beta u_0+a}N_R. \tag{B11}$$

At $u_0=u_1$, defined above, $n_0=n_{01}\simeq\beta aN_R/4[a-\ln(1+a)]\ll N_R$. At $u_0=u_M$, $n_0=n_{0M}\simeq aN_R/[\mathcal{C}(a)+a]$, with u_M and $\mathcal{C}(a)$ given by (B4). For $a\gg1$, $n_{0M}\simeq N_R$; for $a=1$, $n_{0M}\simeq0.8N_R$; and for $a\ll1$, $n_{0M}\simeq0.75N_R$.

For the electric field intensity at the cathode, we have from (A3b) and (A5),

$$\mathcal{E}_0=\frac{J}{e\mu_pN_R}\frac{(u_0-1)(\beta u_0+a)}{(u_0+a)}=\frac{J_1}{e\mu_pN_R}\frac{(u_0-1)(\beta u_0+a)}{(u_0+a)F(u_0)}$$

$$\sim\frac{J}{e\mu_pN_R}(\beta u_0+a)\sim\frac{J_1}{e\mu_pN_R}\frac{(\beta u_0+a)}{F(u_0)}, \tag{B12}$$

with $F(u_0)$ given by (A14) or its asymptotic expansion (B5), whichever is appropriate. We first note that for $u_0>u_1$, $F(u_0)\sim\beta^2u_0$, so that $\mathcal{E}_0\sim J_1/e\mu_pN_R\beta=2V_{\rm th}/L$, as seen from (A5), (A20), and (A8). The factor of two comes from the linear variation of $\mathcal{E}$ with x in this range of u, as exhibited in Eq. (12).

As u_0 decreases below u_1 the current moves into the flatter portion of the negative-resistance region, and the applied voltage is concentrated increasingly toward the cathode, as discussed in Sec. III C. A convenient measure of this concentration of applied field near the cathode is the ratio of the true field $\mathcal{E}_0$ at the cathode to the nominal, ohmically distributed field V/L. From (A8) and (B12) we have, noting that $V_1/(J_1/e\mu_pN_R)=L$,

$$\mathcal{E}_0L/V\sim(\beta u_0+a)F(u_0)/G(u_0). \tag{B13}$$

Detailed analysis shows that $\mathcal{E}_0L/V$ has its maximum value at $u_0=u_2=[a^2-2a+2\ln(1+a)]^{\frac{1}{2}}/\beta^{\frac{3}{2}}$. Letting $\mathcal{E}_0=\mathcal{E}_{02}$ and $V=V_2$ at $u_0=u_2$,

$$(\mathcal{E}_0L/V)_{\max}=\mathcal{E}_{02}L/V_2\simeq(3a/8\beta)^{\frac{1}{2}}. \tag{B14}$$

For $\beta\ll1$, this represents a considerable concentration of applied potential near the cathode.

At $u_0=u_M$, $\mathcal{E}_0L/V=\mathcal{E}_{0M}L/V_M\simeq\frac{1}{2}$, so that the potential again varies gradually between anode and cathode.

APPENDIX C

Domain of Validity of the Neutrality Solution

The domain of self-consistency of the neutrality-based solution of Appendices A and B can be obtained directly by calculating the ratio $|(\epsilon/e)(d\mathcal{E}/dx)/(n-p)|$ evaluated, say, at the cathode, $x=0$. The numerator in this ratio is the magnitude of space-charge density accompanying the field distribution calculated using the neutrality assumption. Where this ratio is small compared to unity the neutrality-based assumption is self-consistent with respect to the neglect of space charge. This calculation has been carried out with results that can also be obtained by the simpler argument which we now present.

In the discussion of the neutrality assumption in Sec. II we saw that, because there is a threshold voltage for double injection, the theory is necessarily inconsistent with respect to the neglect of space charge at currents below some critical current $J_{\rm cr}$, and that below $J_{\rm cr}$ the flow is a one-carrier, SCL (space-charge-limited) current. The transition current $J_{\rm cr}$ from one-carrier, SCL flow to two-carrier, neutralized flow is located where the respective current-voltage characteristics intersect. For the model of Fig. 1, the appropriate one-carrier SCL flow is the electron current given by $J\simeq\epsilon\mu_nV^2/L^3$, as discussed in Sec. II. If $J_{\rm cr}$ is written as $J_{\rm cr}=\alpha J_M$, with J_M given by (B10) and (A5), then at the intersection of the two characteristics we have

$$J=\epsilon\mu_nV^2/L^3=\alpha J_M=\alpha f(a), \quad f(a)=\frac{a}{a-\ln(1+a)}. \tag{C1}$$

Using (13), it is convenient to rewrite (C1) as

$$\left(\frac{V}{V_{\rm th}}\right)^2\simeq\frac{4\alpha eaf(a)\mu_p\langle v\sigma_n\rangle}{\kappa\epsilon_0\langle v\sigma_p\rangle^2}\simeq10^{-5}\frac{\alpha af(a)}{\kappa}\left\{\frac{\mu_p\langle v\sigma_n\rangle}{\langle v\sigma_p\rangle^2}\right\}_{\rm pr}, \tag{C2}$$

where κ is the relative dielectric constant and subscript "pr" denotes that the quantities involved are to be evaluated in practical units.

The steep portion of the neutrality-based, two-carrier characteristic in Fig. 2 corresponds to currents $J<0.2J_M$ and voltages $V\simeq V_{\rm th}$. If the one-carrier SCL curve is to cross the two-carrier curve somewhere on this steep portion, then $(V/V_{\rm th})^2$ as given by (C2) with $\alpha=0.2$, must exceed unity.

For pure Ge at liquid nitrogen temperature, $a\simeq 1$, $f(a)\simeq 4$, $\mu_p\simeq 4\times 10^4$ cm^2/volt sec, and $\kappa=16$. Then $(V/V_{\rm th})^2\simeq 10^{-9}\{\bar\sigma_n/\bar\sigma_p{}^2\}_{\rm pr}$, where we have taken $\langle v\sigma_n\rangle = 10^7\bar\sigma_n$ and $\langle v\sigma_p\rangle = 10^7\bar\sigma_p$. Taking $\bar\sigma_n\approx 10^{-14}$ cm^2 and $\bar\sigma_p\approx 10^{-12}$ cm^2 for a singly-charged acceptor state, $(V/V_{\rm th})^2\simeq 10$, whence $J_{\rm cr}\simeq 0.1\alpha J_M$. Thus, about one decade of the steep double-injection characteristic would be observed if the sample were prepared so that the deep-lying acceptor states were only singly-negatively charged. On the other hand, if $\bar\sigma_n\simeq 10^{-15}$ cm^2 and $\bar\sigma_p\simeq 10^{-12}$ cm^2, $(V/V_{\rm th})^2\simeq 1$, and the intersection of the two characteristics would occur at the end of the steep portion of the two-carrier characteristic. In any case, the entire negative-resistance regime is a valid domain.

For CdS at room temperature, $a<1$, $af(a)\simeq 2$, $\mu_p\approx 10^2$ cm^2/volt sec, and $\kappa\simeq 10$. Thus $(V/V_{\rm th})^2\approx 3\times 10^{-12}\{\bar\sigma_n/\bar\sigma_p{}^2\}_{\rm pr}$. In this case any reasonable values of $\bar\sigma_n$ and $\bar\sigma_p$ for a singly charged acceptor state give $(V/V_{\rm th})^2<1$, and there is no prospect of observing the steep portion of the double-injection characteristic. On the other hand, $(V/V_M)^2=(V_{\rm th}/V_M)^2(V/V_{\rm th})^2\approx 5\times 10^{-13}/\bar\sigma_n$, where V_M is taken from (B9) and (A8). Since $\bar\sigma_n$ corresponds to a neutral cross section, $(V/V_M)^2$ certainly exceeds unity, and a substantial portion of the negative-resistance region is a valid domain.

At higher currents, beyond the negative-resistance region, namely in the semiconductor regime, Eq. (18), the neutrality assumption is still valid, as shown in Appendix E. At still higher currents, namely in the insulator regime, the neutrality assumption is once again not valid. However, for this regime we already have the correct solution, including space-charge effects, from (I), namely Eq. (19).

APPENDIX D

Generality of the Mathematical Technique

For the sake of simplicity we chose for detailed study in this paper the simplest model, that of Fig. 1, exhibiting the characteristic behavior of insulators under double injection, namely carrier lifetimes strongly dependent on injection level. Actually the mathematical technique employed is sufficiently general to handle successfully, under the same assumptions, (i) through (vi) of Sec. II, considerably more complicated models, in particular the model with an arbitrary number of sets of recombination centers arbitrarily located in the forbidden gap with respect to the Fermi level. On the other hand, with the inclusion of traps (centers such that the carriers bound in them are in quasi-thermal equilibrium with the free carriers of the same species), the technique of this paper cannot be used to obtain a rigorous solution. However, in the two extremes where the traps are largely empty or largely filled, the technique is once again successful in yielding a solution. The gaps in the solution so obtained occur where either of the two steady-state Fermi levels moves through a trap level. Due to the nature of Fermi-Dirac statistics, these gaps accordingly occupy relatively narrow ranges of voltage. There should be no difficulty in filling in the gaps by simple interpolation between the analytically derived pieces of the solution.

For a general model, including both recombination centers and traps, the current-flow equation (1) and particle-conservation equations (3) are valid as written, whereas the neutrality equation (2) and recombination expressions (4) must be generalized as follows:

$$n+\textstyle\sum_j n_{tj}-p-\sum_k p_{tk}-\sum_i \Delta p_{Ri}=0, \qquad \text{(D1)}$$

$$r=\textstyle\sum_i r_i,$$

$$r_i=p\langle v\sigma_{pi}\rangle(\bar n_{Ri}-\Delta p_{Ri})=n\langle v\sigma_{ni}\rangle(\bar p_{Ri}+\Delta p_{Ri}), \qquad \text{(D2)}$$

$$\bar n_{Ri}+\bar p_{Ri}=N_{Ri}.$$

Here n_{tj} is the density of electrons in electron traps of type j, total density N_{tnj} and located at energy level E_{tnj}; p_{tk} is the density of holes in hole traps of type k, density N_{tpk} and located at energy level E_{tpk}. Since we are dealing with insulators we neglect here the initial, thermal occupancies of the traps, consistent with our neglect of the thermal densities of free carriers. Where a set of traps is nearly filled in thermal equilibrium, this is taken into account very simply, as shown below. The Δp_{Ri} are the increases in the densities of holes in the recombination centers of the set i, $i=1, 2, \cdots$, of total density N_{Ri} and initial hole occupancy $\bar p_{Ri}$. Because a large change in free-carrier density does not necessarily imply a large change in recombination-center occupancy, we must deal with changes in the populations of the recombination centers. Other symbols are defined as previously. Solving Eq. (D2) for Δp_{Ri},

$$\Delta p_{Ri}=a\theta_i N_{Ri}\frac{1-\delta_i u}{a+\beta_i u}, \quad \beta_i=a\frac{\langle v\sigma_{ni}\rangle}{\langle v\sigma_{pi}\rangle}, \quad \theta_i=\frac{\bar n_{Ri}}{N_{Ri}}, \quad \delta_i=\frac{\beta_i(1-\theta_i)}{a\theta_i}. \qquad \text{(D3)}$$

Substituting for Δp_{Ri} from (D3) back into (D2), we have for r_i:

$$r_i=aN_{Ri}\langle v\sigma_{ni}\rangle pu/(a+\beta_i u). \qquad \text{(D4)}$$

Thermal equilibrium between free and trapped carriers is expressed for electrons and holes, respec-

tively, as

$$n_{tj}=\frac{N_{tnj}}{1+N_{nj}/g_{nj}n}, \quad N_{nj}=N_c\exp\left(\frac{E_{tnj}-E_c}{kT}\right), \quad \text{(D5a)}$$

$$p_{tk}=\frac{N_{tpk}}{1+N_{pk}/g_{pk}p}, \quad N_{pk}=N_v\exp\left(\frac{E_v-E_{tpk}}{kT}\right), \quad \text{(D5b)}$$

The g_{nj}, g_{pk} are statistical weight factors for the trap levels in question. In the simplest situation $g=2$. N_c and N_v are the effective state densities for the conduction and valence bands, respectively. We have assumed that each trap has but one level and holds but one carrier.

From (A3a), $(n-p)\mathcal{E}=(J/e\mu_n)[(u-1)/(u+a)]$, so that Eq. (6), with r given by (D2), can be written as

$$\frac{1}{1+a}d\left(\frac{u-1}{u+a}\right)=\frac{1}{(u+a)^2}du=-\frac{e}{aJ}rdx. \quad \text{(D6)}$$

The procedure to demonstrate analytic tractability of the problem is now straightforward. We must show that r, expressed as a function of u, is such that $1/r(u+a)^2$ is integrable over u in known functions. The actual integration is carried out via a partial-fraction expansion of $1/r(u+a)^2$ as in Appendix A. We have also to show that the integration for V,

$$V=\int_0^L \mathcal{E}dx=-\int_1^{u_0}\mathcal{E}\frac{dx}{du}du, \quad \text{(D7)}$$

with $\mathcal{E}$ given by (A3a) and dx/du by (D6), can be carried out in terms of known functions.

We now consider in detail a representative, general model involving only recombination centers.

Case I. Two Sets of Recombination Centers, Arbitrarily Located in Energy; No Traps ($n_{tj}\equiv p_{tk}\equiv 0$).

From (D2) and (D4),

$$r=r_1+r_2=apu\left(\frac{N_{R1}\langle v\sigma_{n1}\rangle}{\beta_1u+a}+\frac{N_{R2}\langle v\sigma_{n2}\rangle}{\beta_2u+a}\right)$$
$$=\frac{apu(A+Bu)}{(\beta_1u+a)(\beta_2u+a)}, \quad \text{(D8)}$$

$$A=a[N_{R1}\langle v\sigma_1\rangle+N_{R2}\langle v\sigma_{n2}\rangle],$$

$$B=\beta_2N_{R1}\langle v\sigma_{n1}\rangle+\beta_1N_{R2}\langle v\sigma_{n2}\rangle.$$

It remains to express p as a function of u. In (D1), replacing Δp_{R1}, Δp_{R2} from (D3), we obtain

$$p=\frac{a}{u-1}\left(\theta_1N_{R1}\frac{1-\delta_1u}{a+\beta_1u}+\theta_2N_{R2}\frac{1-\delta_2u}{a+\beta_2u}\right)$$
$$=\frac{a(C+Du-Eu^2)}{(u-1)(\beta_1u+a)(\beta_2u+a)}, \quad \text{(D9)}$$

$$C=a(\theta_1N_{R1}+\theta_2N_{R2}),$$

$$D=\theta_1N_{R1}(\beta_2-a\delta_1)+\theta_2N_{R2}(\beta_1-a\delta_2),$$

$$E=\theta_1N_{R1}\beta_2\delta_1+\theta_2N_{R2}\beta_1\delta_2.$$

Substituting for p from (D9) into (D8), we have

$$r=\frac{a^2u(A+Bu)(C+Du-Eu^2)}{(u-1)(\beta_1u+a)^2(\beta_2u+a)^2}. \quad \text{(D10)}$$

Substituting for r from (D10) into (D6), we get

$$f_2(u)du=-\frac{ea}{J}dx,$$
$$f_2(u)=\frac{(u-1)(\beta_1u+a)^2(\beta_2u+a)^2}{u(u+a)^2(A+Bu)(C+Du-Eu^2)}. \quad \text{(D11)}$$

This equation can be integrated after a partial-fraction expansion of $f_2(u)$. Substituting for p from (D9) into (A3a), we obtain

$$\mathcal{E}=\frac{J}{ea\mu_n}h_2(u), \quad h_2(u)=\frac{(u-1)(\beta_1u+a)(\beta_2u+a)}{(u+a)(C+Du-Eu^2)}. \quad \text{(D12)}$$

Substituting for $\mathcal{E}$ from (D12) and for dx/du from (D11), (D7) can be written:

$$V=\frac{J^2}{a^2e^2\mu_n}\int_1^{u_0}h_2(u)f_2(u)du. \quad \text{(D13)}$$

The integration in (D13) is carried out via a partial-fraction expansion of $h_2(u)f_2(u)$.

It is clear that this procedure will also work for more than two sets of recombination centers. The only additional complication will then be the presence of polynomials of degree higher than two in the denominators of the integrands. In order to carry out the integrations in these cases, via the partial-fraction expansions, one must first find the roots of these higher degree polynomials.

The mathematical difficulties encountered with the inclusion of traps in the problem is illustrated by the following example.

Case II. A Single Set of Recombination Centers, Arbitrarily Located in Energy, and a Single Set of Electron Traps, Initially Empty.

Dropping subscript j, (D5a) can be rewritten as

$$n_t=N_{tn}pu/(pu+N_n'), \quad N_n'=N_n/g_n. \quad \text{(D14)}$$

Substituting for Δp_R from (D3), dropping subscript i, and for n_t from (D14), the neutrality condition (D1) can be written as

$$p(u-1)+\frac{N_{tn}pu}{pu+N_n'}-a\theta N_R\frac{1-\delta u}{\beta u+a}=0. \quad \text{(D15)}$$

This yields a quadratic equation in p:

$$p^2u(u-1)+p\left(N_n'(u-1)+N_{tn}u-a\theta N_R\frac{u(1-\delta u)}{\beta u+a}\right)-a\theta N_RN_n'\frac{1-\delta u}{\beta u+a}=0. \quad \text{(D16)}$$

It is obvious that, solving this equation for p as a function of u and substituting into (D6) and (D7), it will be impossible to carry out the indicated integrations. We conclude therefore that our mathematical technique breaks down when trapping is included in the problem, if a rigorous solution is sought.

On the other hand, in the two limits $N_n'\gg pu$ and $N_n'\ll pu$ the method is applicable, as we see simply by rewriting (D15) in these limits:

$N_n'\gg pu$, $n_t\simeq\dfrac{N_{tn}}{N_n'}pu$:

$$p\left[u\left(1+\frac{N_{tn}}{N_n'}\right)-1\right]-a\theta N_R\frac{1-\delta u}{\beta u+a}=0, \quad \text{(D17)}$$

$N_n'\ll pu$, $n_t\simeq N_{tn}$:

$$p(u-1)+N_{tn}-a\theta N_R\frac{1-\delta u}{\beta u+a}=0. \quad \text{(D18)}$$

As discussed earlier, a reasonable program is to solve the problem analytically in the two limits (D17) and (D18), and to interpolate between these solutions for the transition region, $N_n'\sim pu$.

Note that if the traps are initially almost completely filled, say with occupancy $\bar{n}_t<N_{tn}$, then we need only be concerned with the trap-filled limit, Eq. (D18), replacing N_{tn} by $N_{tn}-\bar{n}_t$.

APPENDIX E

Space Charge in the Semiconductor Regime. A Correction to (I)

Since $n-p\simeq N_R$ characterizes the semiconductor regime, $Q(V)<\int_0^L e(n-p)dx\simeq eN_RL$, with $Q(V)$ the total injected space charge per unit area at voltage V, so long as $V<V_I$ where $Q(V_I)=eN_RL$. Since the plot of $\mathcal{E}$ vs x with $\mathcal{E}$ given by (17) in the semiconductor regime, is convex, $Q\simeq CV=\epsilon V/L$ as shown in Appendix B of (I). Thus $V_I\simeq eN_RL^2/\epsilon$ and $Q(V)/eN_RL\simeq V/V_I$. Since V_I marks the transition from the semiconductor to the insulator regime of the double-injection problem, as shown rigorously in (I), $Q(V)/eN_RL<1$ in the semiconductor regime, and the solution in this regime is self-consistent with respect to the neutrality assumption.

We take the opportunity here to correct an error in (I). In the physical argument of Sec. III of (I), Q was guessed to be given by $Q=e\bar{n}Lt_{\Omega n}/\tau_{\text{high}}=\epsilon\bar{n}L/N_R\mu_n\tau_{\text{high}}$ [Eq. (5) of (I), and we have replaced $n_T\simeq n_T-p_T$ by N_R, as discussed in Sec. II B]. Here $\bar{n}$ is the average, injected free-electron density. To see that this guess was incorrect, by virtue of a missing factor involving the mobility ratio, we compute the correct $Q=Q_{\text{corr}}$, in similar form, from Eq. (15): $Q_{\text{corr}}=\epsilon L|d\mathcal{E}/dx|_{\text{av}}=(a+1)\epsilon\bar{n}L/aN_R\mu_n\tau_{\text{high}}$. This differs from Eq. (5) of (I) through the factor $(a+1)/a=(\mu_n+\mu_p)/\mu_p$. In retrospect we see that it would have been quite difficult to guess the precisely correct Q without any mathematical analysis at all. If now we combine the relation $Q_{\text{corr}}=CV=\epsilon V/L$ with the correct current expression: $J=e\bar{n}L[(1/t_p)+(1/t_n)]=e\bar{n}V(\mu_n+\mu_p)/L$ in place of Eq. (1) of (I), we obtain directly the semiconductor regime, Eq. (18), except for the factor 9/8.

From this discussion we see that merely guessing the correct space charge and then using the simple approach of Sec. III of (I), which will indeed then give the correct current-voltage relationship, does not in itself establish that the double-injection current flow is actually space-charge-limited, rather than purely recombination-limited. The Poisson equation must be shown to be essential for the derivation before such a conclusion can be drawn. We have shown in this paper, Sec. III B, that the Poisson equation is indeed not essential for derivation of the semiconductor regime. On the other hand, it was shown in (I) that it is essential for the derivation of the insulator regime.