Electronics Research Laboratory
University of California
Berkeley, California

PROPOSAL FOR A TWO-STAGE SEMICONDUCTOR LASER THROUGH TUNNELING AND INJECTION

by

Shyh Wang

This work was supported by the Air Force Office of Scientific Research of the Office of Aerospace Research; the Department of the Army, Army Research Office; and the Department of the Navy, Office of Naval Research under Grant AF-AFOSR-62-340.

July 26, 1963

ABSTRACT

In the present paper, an analysis of laser action in semiconductors is made, and a condition for population inversion is established. The radiation output is found to be proportional to the deviation from the equilbrium value of the occupancy of the allowed states including the states in the conduction and valence band and also the donor and acceptor states. Possible laser levels are discussed.

Following the general analysis, a scheme is proposed for laser action in covalent semiconductors where the top of the valence band and the bottom of the oonduction band do not occur at the same value of k. The scheme consists of simultaneously tunneling electrons into the (000) valley of the conduction band and injecting holes into the valence band. The outstanding feature of the proposed scheme is that it offers a definite possibility for separate amplitude and frequency modulation of the laser output. Various losses associated with the proposed scheme are also discussed.

I. INTRODUCTION

Recently, many workers[1-4] have reported laser action in GaAs diodes biased in the forward direction. In GaAs, the bottom of the conduction band and the top of the valence band occur at the same value of k, (000), that is the center of the Brillouin zone.[5] Therefore, electrons and holes under equilibrium conditions are located near the (000) valley of the conduction and valence bands, respectively. More important, if excess electrons and holes are injected into the base region of a GaAs diode, the excess carriers will have the same value of k.

In an optical transition, it is required by wave mechanics that the total momenta should be conserved, the momentum being equal to $\hbar$ multiplied by the wave number k. Since the wave number k associated with photons is much smaller than that of electrons in a crystalline lattice, the momentum conservation law simply reduces to $k_i = k_f$ where k_i and k_f are the initial and final wave number of the electrons involved in a direct optical transition. Therefore, in semiconductors like GaAs where the electrons and holes have the same value of k, generation or recombination of excess carriers may result in direct absorption or emission of radiation. Such processes are illustrated in Fig. 1a.

In semiconductors such as Ge, however, the bottom of the conduction band occurs at a different value of k.[5, 6] To make an optical transition across the band, electrons and holes must be scattered by lattice phonons to a proper k value as shown in Fig. 1b. Therefore, ordinary absorption or emission of photons in Ge-like semiconductors is indirect or phonon-assisted. This fundamental difference in the band structure of GaAs and Ge explains why intense recombination radiation[7] has been observed in GaAs but not yet in Ge. It is this intense recombination radiation that led to the observation of laser action in GaAs.

II. POPULATION INVERSION IN SEMICONDUCTORS

Before we can discuss laser action, we must first know what constitutes the condition for a population inversion in semiconductors. For clarity and ease of discussion, we shall start with donor and acceptor states. Let N_a and N_d be the density and n_a and n_d be the occupied density of donor and acceptor states, respectively. Under equilibrium conditions, we have

$$n_a = n_{ao} = \frac{N_a}{1+F_a}, \quad n_d = n_{do} = \frac{N_d}{1+F_d} \tag{1}$$

where $F_a = 2 \exp(E_a - E_f)/kT$ and $F_d = \frac{1}{2} \exp(E_d - E_f)/kT$, E_a, E_d and E_f being the energies of the acceptor states, the donor states and the Fermi level, respectively. The rate of transition from the donor to the acceptor states is given by

$$R(d \rightarrow a) = n_d \, P(d \rightarrow a) \, (N_a - n_a) \tag{2}$$

where $P(d \rightarrow a)$ denotes the transition probability from the donor to the acceptor states induced by the photons. Similarly, the rate of transition from the acceptor to the donor states is equal to

$$R(a \rightarrow d) = n_a \, P(a \rightarrow d) \, (N_d - n_d) \tag{3}$$

These transitions are illustrated in Fig. 2. It is well known from quantum mechanics that $P(d \rightarrow a) = P(a \rightarrow d)$. To obtain a net emission of radiation, we must have $R(d \rightarrow a) > R(a \rightarrow d)$ since the donor state is the higher energy state. Thus, we have

$$n_d(N_a - n_a) > n_a(N_d - n_d) \quad \text{or} \quad n_d/N_d > n_a/N_a \tag{4}$$

The above equation constitutes the condition for population inversion between the donor and acceptor states. Physically, Eq. (4)

means that the percentage of occupied donor states must be greater than that of occupied acceptor state to achieve a population inversion between these states.

Note that under equilibrium conditions, we have from Eq. (1),

$$\frac{n_d(N_a-n_a)}{n_a(N_d-n_d)} = F_a/F_d = 4 \exp (E_a-E_d)/kT \ll 1. \qquad (5)$$

Equation (5) is based on Fermi statistics, and it is equivalent to the Boltzmann population ratio, $(N_1/N_2)_o = \exp - (E_1 - E_2)/kT$ used in a two-level maser, the factor 4 in Eq. (5) being a result of spin degeneracy. Under equilibrium conditions, though we may have more electrons in the donor states than in the acceptor states, we do not have a population inversion. In order to have a net emission of radiation, we must fulfill the condition stated in Eq. (4).

Now we can extend our discussion to electrons and holes in the conduction and valence bands, respectively. Let $N_c(E)$ and $N_v(E)$ be the density of states and $n_c(E)$ be the density of occupied states having energy between E and $E + \Delta E$ in the conduction and valence bands, respectively. Let us consider transitions limited to states within energy interval ΔE as shown in Fig. 2. Following exactly the same procedure as we did for the donor and acceptor states, we obtain the following condition for population inversion:

$$n_c(N_v - n_v) > n_v(N_c - n_c). \qquad (6)$$

Under equilibrium condition, by using the Fermi distribution function, we find

$$\frac{n_c(N_v - n_v)}{n_v(N_c - n_c)} = \exp - (E_1 - E_2)/kT \ll 1 \qquad (7)$$

where E_1 and E_2 are the energy of the states under consideration as shown in Fig. 2. Again, Eq. (6) has meaning only when the states under consideration have a non-vanishing transition probability.

III. NON-EQUILIBRIUM SITUATION

From the previous discussion, we see that means must be devised to achieve a non-equilibrium situation in order to get a population inversion. Let us suppose that we have a scheme to inject electrons into the donor states and holes into the acceptor states and that Δn_d and Δp_a be the injected electron and hole concentration in the donor and acceptor states, respectively. Since a hole is equivalent to a missing electron, we have

$$n_d = n_{do} + \Delta n_d, \qquad n_a = n_{ao} - \Delta p_a \tag{8}$$

where n_{do} and n_{ao} are the equilibrium densities of the occupied donor and acceptor states, respectively. Thus, the condition for population inversion becomes

$$(n_{do} + \Delta n_d) N_a > (n_{ao} - \Delta p_a) N_d \tag{9}$$

Using Eq. (5), we can rewrite Eq. (9) as

$$\frac{\Delta n_d}{N_d} + \frac{\Delta p_a}{N_a} > \frac{F_d - F_a}{(1+F_d)(1+F_a)} \tag{10}$$

To see definitely the possibility of a population inversion, we take the extreme case that all the donor states are occupied and all the acceptor states are empty under the non-equilibrium situation. In other words, we take Δn_d and Δp_a to be their respective maximum possible values, that is

$$\Delta n_d = N_d - n_{do}, \quad \Delta p_a = n_{ao}. \tag{11}$$

Substituting Eq. (11) into Eq. (10), we find

$$1 - \frac{1}{1+F_d} + \frac{1}{1+F_a} > \frac{F_d - F_a}{(1+F_a)(1+F_d)} \quad \text{or} \quad 1 > 0. \tag{12}$$

Thus, under this extreme situation, a population inversion is definitely established. Now let us relax the condition in a practical situation. Let

$$\Delta n_d = \alpha_d N_d - n_{do}, \quad \Delta p_a = n_{ao} - \beta_a N_a. \tag{13}$$

In other words,

$$n_d = \alpha_d N_d, \qquad n_a = \beta_a N_a. \tag{14}$$

Substituting Eq. (13) into Eq. (12) and using Eq. (1), we find the condition as

$$\alpha_d - \beta_a > 0. \tag{15}$$

Note that the inequality condition given by Eq. (15) is identical to that given by Eq. (4).

Similarly, we can extend our discussion to electrons in the conduction and valence band, and again for clarity we shall concentrate our attention on electrons within a narrow energy band ΔE as shown in Fig. 2c. If Δn_c and Δp_v are the injected electron and hole density within the given energy interval ΔE in the conduction and valence band, respectively, then we have

$$n_c = n_{co} + \Delta n_c, \qquad n_v = n_{vo} - \Delta p_v. \tag{16}$$

From Eq. (6), the condition for a population inversion becomes

$$\frac{\Delta n_c}{N_c} + \frac{\Delta p_v}{N_v} > \frac{F_c - F_v}{(1+F_c)(1+F_v)} \tag{17}$$

where $F_c = \exp (E_1 - E_f)/kT$ and $F_v = \exp (E_2 - E_f)/kT$. If we define α_c and β_v as the respective percentage of occupied states in the conduction and valence band under the non-equilibrium situation, we find,

$$\Delta n_c = \alpha_c N_c - n_{co}, \qquad \Delta p_v = n_{vo} - \beta_v N_v \tag{18}$$

and consequently, we have for the condition of a population inversion the following

$$\alpha_c > \beta_v \tag{19}$$

IV. CONSIDERATION OF POSSIBLE LASER LEVELS

Once the condition for population inversion is established, we can discuss the possible choice of laser levels, that is the levels between which the laser action takes place. In Fig. 3, a schematic representation of two possible sets of laser levels is shown: (1) the states in the conduction and valence band and (2) the donor and acceptor states. We believe that the latter scheme is much superior. Our reasoning is as follows.

Let us take the first scheme. Suppose that electrons and holes are injected into the base region as in the case of a GaAs diode biased in the forward direction. The important question is how these injected electrons and holes will distribute themselves among the energy states in the conduction and valence band, respectively. For maximum efficiency, we want them in the respective energy interval ΔE as shown in Fig. 3a. In order to achieve this, we must have (1) E_1 and E_2 very close to the band edge and (2) a very low temperature. The condition of operating the laser at a very low temperature is imperative because otherwise the injected carriers will spread over a wide energy range.

If we do operate a semiconductor laser at a very low temperature, we may have an even more efficient scheme, that is the second scheme. Let Δn be the total density of injected electrons. Obviously,

$$\Delta n = \Delta n_c + \Delta n_d \tag{20}$$

Since the shallow donor states are tightly coupled to the electrons in the conduction band according to the effective mass theory,[8] it is reasonable to assume that they are able to maintain a quasi-equilibrium situation among them. Therefore, we expect

$$\frac{\Delta n_c}{\Delta n_d} = \frac{n_{co}}{n_{do}} = \frac{N_c}{N_d} \frac{1 + F_d}{1 + F_c} \tag{21}$$

If we have a highly compensated sample so that the Fermi level lies between the donor and acceptor levels, then the ratio $\Delta n_c / \Delta n_d$ at very low temperatures approaches

$$\frac{\Delta n_c}{\Delta n_d} \sim \frac{N_c}{N_d} \exp - (E_c - E_d)/kT \tag{22}$$

and this ratio can be very small compared to unity. In this way, we are sure that practically all the injected electrons go to the level at which the laser action takes place. The same argument applies to the injected holes. In the second scheme, the conduction and valence bands act merely as transitory storage places for electrons and holes, respectively.

There is another important advantage for the second scheme from the viewpoint of design. For example, we can vary the concentration of N_a and N_d in order to achieve a maximum degree of population inversion at a given temperature. This consideration and other considerations such as the magnitude of the transition probability, the question of reabsorption, etc., will be discussed in later sections when appropriate.

V. TWO STAGE LASER SCHEME FOR Ge-LIKE SEMICONDUCTORS

It has been pointed out in the introductory section that direct radiative recombination can take place in semiconductors like GaAs where the bottom of the conduction band has the same value of k as the top of the valence band. Therefore, for laser action, only a diode structure is needed. The diode when biased in the forward direction supplies to the base region the excess electrons and holes necessary to establish a population inversion as we have discussed in the previous section.

In semiconductors like Ge, however, injection alone is not enough for laser action. To be specific, let us restrict our discussion to Ge. The minimum of the conduction band edge is at the (111) edge in k space.[6] Ordinarily, practically all the free electrons in Ge are located near this conduction band minimum, and practically all the holes are located near the center (000) of the Brillouin zone of the valence band. Injected electrons and holes also behave the same way. Consequently, the band-to-band recombination process has to be phonon-assisted or indirect as shown in Fig. 1b. The probability for such a process is extremely small as pointed out by Van Roosbroeck and Shockley.[9] In order for direct band-to-band recombination to take place in Ge, means must be devised to have excess electrons and holes simultaneously and respectively present (population inversion) at the (000) valley of the conduction and valence bands. Following is the proposed scheme.

Take a pnp structure as shown in Fig. 4. The left-hand-side pn junction is heavily doped on both sides so that tunneling is possible while the right-hand-side np junction is reasonably doped as an ordinary diode. The heavily doped junction is reversely biased so that direct tunneling is possible from the (000) valley of the valence band on the p-side into the (000) valley of the conduction band on the n-side. Such a direct tunneling has been observed experimentally by Morgan and Kane[10] at 77°K and by Hall, and Racette[11] at 42°K.

Once we have electrons (the majority carrier) in the (000) valley of the conduction band, we need to supply holes through minority carrier injection. This is the purpose of the right-hand-side np junction. At very low temperatures, the valence band in the n-region is almost fully occupied, and so are the acceptor states. In Eqs. (15) and (19), $\beta_a = 1$ and $\beta_v = 1$. Obviously population inversion will not occur even though we may have $\alpha_d = 1$ and $\alpha_c = 1$, that is fully occupied donor states and (000) valley states. The idea behind the proposed scheme is illustrated and summarized in Fig. 4. The electrons tunneled into the (000) valley combine with the injected holes presumably near the tunnel junction, and such direct recombination will result in the emission of radiation. It should be pointed out, however, that by direct radiative recombination, the process indicated in Fig. 4, we also include the recombination process through the shallow donor and acceptor states.

A discussion of the essential requirements for the success of the proposed scheme is in order. First, the question arises as to how far the (000) valley electrons can travel before they are scattered into the (111) valley. Up to the present, no direct measurement of such inter-valley scattering has been made. However, we can estimate the magnitude of the mean free path due to phonon-scattering of the (000) valley electrons from optical absorption data. In the indirect transition region, one process (Fig. 1b) involves the optical excitation of an electron from the valence band into the (000) valley of the conduction band and the scattering of the (000) valley electron into the (111) valley accompanied by a phonon emission.

The optical absorption coefficient in the direct and indirect transition regions has been worked out in detail by Bardeen et al[12] and by Roth et al.[13] The important quantities which we want to know are the magnitudes of the matrix element of the momentum operator, P_{if}, and the matrix element of the electron-phonon interaction, M_s, the notations being the same as in reference 12. Using the optical absorption data of Dash and Newman[14] and Macfarlane et al,[15] we find $|P_{if}|^2 = 15(\Delta E)(m)$ where ΔE is the energy of the direct gap at k(000) and m is the free electron mass and $M_s^2 = 3.5 \times 10^{-49}$ erg^{2}cm^3. These values are in good agreement with the values used by Dumke[16] and by Kane[17] in their work. In the calculation, we have used a reduced effective mass of 0.031 m obtained from an effective mass of 0.037m for the (000) valley electrons and an effective hole mass of 0.20m. We also have assumed that the operating temperature is low so only the phonon emission process is important. Following a similar procedure as in the evaluation of electron mobility and summing M_s^2 over all possible final states, we find the transition probability, w, per unit time due to phonon-scattering of (000) valley electrons into various (111)-like valleys to be

$$w = \frac{2\sqrt{2}}{\pi\hbar^4} M_s^2 \, m_c^{3/2} \, E^{1/2} \qquad (23a)$$

where $m_c = 0.55$m is the density of state effective mass for the (111) valley electrons and E is the kinetic energy of the electron after it is scattered into the (111) valley. The mean free path, ℓ, of (000) valley electrons is equal to v_{000}/w where v_{000} is the velocity of the electron. In evaluating ℓ, we note that the momenta must be conserved; that is, $k_i = k_f \pm k_p$ where k_i, k_f and k_p multiplied by $\hbar$ are the momenta associated with the initial and final states of the electron and the momentum of a lattice phonon respectively. Since optical absorption data give M_s^2 for scattering from the (000) to the (111) minima, it is implied that $\pm k_p = k(111) - k(000)$. In other words, $2m_c' E = m_{000}^2 v_{000}^2$ where $m_{000} = 0.037m$ is the effective mass of (000) valley electrons and $m_c' = 0.12m$ is the mobility effective mass of (111) valley electrons. Thus, we obtain

$$\ell = \frac{\pi\hbar^4}{2} \frac{m_c'^{1/2}}{M_s^2 m_c^{3/2} m_{000}} \tag{23}$$

From Eq. (23), we find $\ell = 1.5 \times 10^{-4}$ cm. which corresponds to a mean free time of 6×10^{-12} sec. at 77°K. A high frequency (1000 mc) transistor has a base width of 10^{-4} cm or less; therefore, the problem of phonon scattering can be overcome. Zwerdling et al[18] have estimated for (000) valley electrons a scattering time about 5×10^{-12} sec. from the line width of magneto-optical absorption curves. Since intra-valley scattering should also contribute to the line width, we feel reasonably safe to use the calculated value of 6×10^{-12} sec. as the inter-valley phonon scattering time.

Next comes the question of impurity scattering. Weinreich, et al,[19] in their investigation of the acousto-electric effect in n-type germanium doped with arsenic, have measured the inter-valley scattering between the various conduction band minima, that is the (111)-like valleys. Holonyak et al[20] and Morgan and Kane,[10] in the early work on germanium tunnel diodes have also shown the importance of impurity scattering and its variation with the doping substance. The scattering process is attributed to an impurity admixture of the (000) and the various (111)-like valley states due to the impurity cell potentials. Among the three substances, Sb, P and As, which they have investigated, antimony has a much smaller scattering cross-section than the other two. It is clear from Fig. 5a that direct tunneling to the (000) valley of the conduction band sets in only after the bias voltage E_{app} is greater than $E(000)-E_{fn}$. Below this voltage, indirect tunneling is possible with the assistance of phonon or impurity scattering; therefore, a detailed analysis of the current in the indirect tunneling region gives a measure of the relative importance of phonon and impurity scattering. Recent work of Fritzsche and Tiemann[21] clearly shows that phonon scattering is much more important than impurity scattering in germanium doped with antimony. For the proposed pnp structure for laser action, it is imperative that antimony be used as the dopant in the n-region.

Now we can turn our attention to the injected holes. What we are concerned with is that the injected holes may have a very short lifetime due to high concentration of non-radiative recombination centers accompanied with high concentration of donor and acceptor impurities. Take $D_p = 300\,cm^2/sec.$ at $77^{o}K$ and $\bar{\bar{w}}$= base width = 10^{-4} cm; a simple calculation shows that the lifetime has to be smaller than 10^{-11} sec. for the said recombination process to be significant. This is consistent with the experimental observation that a high frequency transistor with heavily doped base still has a current amplification factor greater than 0.95 at room temperature.

From the above discussion, we can be reasonably sure that if the width of the proposed pnp structure is 10^{-4}cm. or less, a steady state distribution of tunneled (000) valley electrons and injected holes can be maintained in the base region. With diffusion technique available to us, the base width requirement can certainly be met. As pointed out earlier, the simultaneous presence of the (000) valley electrons and injected holes constitutes the working condition of the proposed scheme. Once we have established the realizability condition, we can discuss how we can enhance the direct, radiative recombination of these electrons and holes.

Dumke,[16] using the method of detailed balancing, has calculated the rate of direct radiative recombination from the optical absorption data. In his calculation, he has assumed a thermal equilibrium distribution of electrons among the (111) and (000) valleys and an equilibrium carrier concentration of $n_o = p_o = 2.4 \times 10^{13}/\text{c.c.}$, and he has concluded a direct recombination lifetime of the order of 0.5sec. at 300°K. In the present case, we supply the (000) valley electrons directly through tunneling; hence, we gain a factor of $(m_c/m_{000})^{3/2} \exp(E_{000}-E_{111})/kT$ where the mass ratio takes care of the difference in density of states, while the exponential dependence is the usual Boltzmann factor. Assuming a carrier concentration of $10^{17}/\text{c.c.}$ through injection and tunneling and taking $E_{000}-E_{111} = 0.154$ev and $kT = 0.024$ev, we find $\tau(\text{direct}) = 5 \times 10^{-9}$sec. As the carrier concentration is further increased, the semiconductor eventually becomes degenerate and the lifetime of carriers approaches the ultimate spontaneous lifetime, τ, which is the reciprocal of Einstein's coefficient of spontaneous emission.

$$\tau = \frac{c^3 m^2}{4n^3 h\nu e^2} \frac{\hbar^2}{|P_{if}|^2} \tag{24}$$

where c is the velocity of light, n the index of refraction, ν the frequency, and e the electronic charge. Using the value of P_{if} given earlier and $\Delta E = 0.88$ ev for Ge, we find $\tau = 1.5 \times 10^{-11}$sec. This value of τ should be compared with the inter-valley scattering time of 6×10^{-12}sec. That means, in the spontaneous emission region, about 30 per cent of the tunneled electrons will participate in the photon emission process. Of course, in the stimulated emission region, the lifetime of an electron for the photon emission process decreases drastically and hence the efficiency of the photon emission process approaches unity as the light intensity increases. In order to get into the stimulated emission region, however, the proposed device needs more pumping current by a factor of about three than a comparable GaAs injection diode as the present calculation indicates.

From the above discussion, it is evident that the lifetime of carriers changes with the current density. The sooner the levels between which direct recombination radiation takes place become totally occupied or totally empty, the lower is the threshold current density at which efficient radiation begins. We have argued earlier that donor and acceptor states are more efficient laser levels than conduction and valence band states, and our present discussion further strengthens the argument. Nathan and Burns[22] and Nelson et al,[23] in their work with GaAs diodes have demonstrated that shallow donor and acceptor states play a dominant role in the direct recombination process.

To see how shallow donor and acceptor states can play a part in the direct recombination process, a short review of the structure of these states[24] is in order. Since the effective mass theory has been very successful in explaining the energies of these states and the results of optical absorption and spin resonance experiments, we shall quote the relevant conclusions derived from that theory. First, the electron wave function for the shallow states extends over several thousand crystal cells. As a matter of fact, Conwell[25] has shown that for an impurity concentration greater than 5×10^{16}/c.c., electrons can jump from impurity to impurity, and the impurity states begin to form a band which overlaps the conduction or valence band. Furthermore, for a doping concentration of impurities greater than 10^{18}/c.c., there is a considerable overlap in the electron wave function of the shallow donor and acceptor states. Secondly, the wave function of these shallow states can be expressed as a linear combination of the electron wave function near the respective band minima. Thus, the electrons in these shallow states have the same value of k as the electrons in their respective valley of the conduction and valence bands.

From the above discussion we conclude that if the doping concentration is reasonably high, we can have donor to acceptor state transition as well as conduction to valence band transition. The donor to acceptor state transition may be the dominant one simply because the shallow donor and acceptor states, respectively, become fully occupied and empty before the conductor and the valence band states. In the following discussion, we shall concentrate on the donor to acceptor type of transition to save repetition. Furthermore, we shall assume that the impurity bands have the same structure as the respective conduction and valence bands so that data from band-to-band transition may be used to estimate the performance of the proposed device.

VI. OPERATION FEATURES

Let us take the donor to acceptor type of transition. The net output of radiation power, $\overline{W}$, is equal to

$$\overline{W} = [R(d \rightarrow a) - R(a \rightarrow d)] (E_d - E_a) \tag{25}$$

Using Eqs. (2), (3) and (16), we find

$$\overline{W} = P(d \rightarrow a) [N_a \Delta n_d + N_d \Delta p_a + N_a n_{do} - N_d n_{ao}] (E_d - E_a) \tag{26}$$

Suppose that we completely saturate the donor states through tunneling, then $\Delta n_d + n_{do} = N_d$. Further, we assume that at low temperatures, the acceptor states are completely occupied under equilibrium condition, that is $n_{ao} = N_a$. Thus, Eq. (26) reduces to

$$\overline{W} = P(d \rightarrow a) N_d \Delta p_a (E_d - E_a) \tag{27}$$

Equation (27) simply means that under such circumstances, the output is controlled by the availability of holes. Equation (27) also predicts a linear relationship between the light output and the injection current. Obviously, the bias on the injection junction can be used to modulate the amplitude of the laser output.

The tunnel junction also can serve another function besides tunneling. It has been pointed out by Kohn[24] that the Stark effect on the energies of shallow impurity states may be measurable. For a field of 3000 volts/cm, he has estimated a second-order shift of 5×10^{-5} ev. In a reversely biased tunnel diode, an electric field as high as 10^5 volts/cm may be obtained. It may be worthwhile to give an order of magnitude estimate of the effect. The second-order Stark effect is proportional to the square of the electric field. Hence, for a change of electric field of 10^3 volts/cm at a field of 10^5 volts/cm, we obtain a shift in the energy of the shallow state of $5 \times 10^{-5} \times 2 \times 10^5 \times 10^3 / (3000)^2 = 10^{-3}$ ev. To get a one-tenth of one per cent change in the frequency of the emitted radiation, we need only to change the bias voltage by $10^3/10^5$ or one per cent. Such a small change in the bias voltage should not produce any significant change in the amplitude of the radiation output. Thus, it seems quite plausible to modulate the frequency of laser output by a microwave signal acting as a part of the reverse bias. We should add that the frequency of a laser is determined by the resonant mode, that is dimension, of a resonator. Therefore, the amount of frequency modulation allowed without suffering severe amplitude deterioration is actually limited by the quality factor, Q, of the resonant structure. We should mention that a change in the absorption edge of GaAs under a high electric field has been observed experimentally.[26, 27] This indicates that such effect in covalent semiconductors is measurable.

The proposed two-stage (tunneling and injection) scheme for laser action can be applied to other Ge-like semiconductors where the conduction band minimum and the valence band maximum do not occur at the same value of k. We should also add that should the direct recombination process take place near the injection junction instead of the tunnel junction, then the role of injection and tunnel junction should be interchanged in the above discussion. It is also evident from the discussion that the proposed pnp structure will have a greater mode stability than the diode structure of an injection laser simply because the amplitude of the radiation output can now be changed without changing the condition of the other junction near which radiative recombination takes place. The only drawback of the present scheme is the added loss directly or indirectly due to tunneling. In the following, we shall discuss the various loss mechanisms associated with the present scheme.

First, let us consider the tunnel diode, alone. The conductance of a Ge tunnel diode is illustrated in Fig. 5b according to the experimental work of Hall and Racette, and the sharp increase corresponds to the onset of direct tunneling.[11] It is beyond this voltage where we want to operate the tunnel diode. For Ge, this is around 0.10 ev. However, we notice that before direct tunneling sets in, there is a current due to indirect tunneling (through phonon or impurity scattering) into the (111) valley as shown in Fig. 5a. This component of current constitutes a real loss. To cut down the indirect tunneling process, it is necessary to operate the proposed scheme at low temperatures and to use the right kind of impurity which gives the least admixture of conduction band states. From experimental work of Fritzsche and Tiemann,[21] the ratio of direct to indirect tunneling current is much bigger than a factor of ten for Sb-doped Ge diode; hence, the loss due to indirect tunneling is insignificant. However, if the difference E(000)-E(conduction band minimum) is relatively large, say beyond 1 ev, the loss due to indirect tunneling may become prohibitively large before direct tunneling sets in. Such may be the case with Si.

Now we shall discuss the loss associated with the injection junction. The component of current due to the injection of electrons into the right-hand-side p-region is entirely wasteful. To reduce this wasteful component, we need to have a graded impurity concentration in the base region (n-region) so that the tunnel junction side is heavily doped while the injection junction side is relatively lightly doped. Of course, the base width should be small compared with the mean free path of the (000) valley electrons and recombination of the injected holes with electrons in the (111) valley of the conduction band should be negligible. With diffusion techniques now available to us, the above requirements can surely be met in the laboratory.

The most important loss mechanism is the reabsorption of emitted radiation. During the process, an electron in the valence band may be excited back either to the (000) valley or the (111) valley of the conduction band. We have already included the first process in our analysis of the net radiated power so only the second process constitutes a real loss. We can calculate this loss from optical absorption data.[14,15] Let α be the absorption coefficient due to the phonon-assisted process and d be the linear dimension of the resonant cavity along the direction in which the radiation undergoes multiple reflection. For each passage, the amplitude of the radiation is reduced by a factor $\exp -\alpha d$, and this loss should be added to the loss at the cavity wall. Of course, α depends upon the the wavelength of the radiation, but for the present estimate, we can take a reasonable value of 20cm^{-1} at 77°K. The linear dimension of a semiconductor laser is around 5×10^{-3} cm. Thus $\alpha d = 0.1$ and $\exp - \alpha d = 1 - 0.1$. Now we can follow readily the analysis given by Schawlow and Townes[28] and establish the condition for laser oscillation. We shall start with Eq. (26). In an optical transition, the initial and final states must obey the laws of conservation of energy and momentum, and this restriction should be applied to the product of $N_a N_d P(d \rightarrow a)$ in Eq. (26). If we assume that the impurity bands have the same structure as the conduction and valence bands, then only pairs of states in the impurity bands with matched values of k can participate in the radiation process. For simplicity, we shall further assume that $N_a = N_d$ and that there are N such matched pairs in the frequency range between ν and $\nu + \Delta\nu$. Then, the output of the radiation power is given by

$$\bar{W} = \left[\frac{e^2}{2\pi m^2 \nu^2 h^2} \; I(\nu) \; |P_{if}|^2 \right] N(\alpha_d - \beta_a) h\nu \qquad (28)$$

where $I(\nu)$ is the intensity of radiation at frequency ν, and the factor in the bracket is the rate of transition per sec. Suppose that we have a parallel-plate resonant structure and that the reflection coefficient at the wall is R. Then, for each passage, the loss in radiation energy is equal to $I(\nu)\Delta\nu(1 - R + \alpha d)$. The condition for laser oscillation is reached when the power gain is greater than the power loss. Using Eqs. (24) and (28), we have

$$N(\alpha_d - \beta_a) \geq \frac{\Delta\nu}{\nu} \; \frac{8\pi^2 (1 - R + \alpha d) n^2 \nu^3}{c^2 d} \; \tau \qquad (29)$$

where τ is the spontaneous lifetime of carriers and $\Delta\nu$ is the half-width expressed in frequency of the emitted radiation. For Ge, 1 - R is about 0.4. Therefore, reabsorption of carriers only raises the population inversion requirement for laser oscillation by about 25 per cent.

Finally, we want to mention that recent experimental work[29] on Ge mesa transistors (2N700) strongly suggests that the proposed mechanism for direct recombination is operative. The observed radiation is much stronger than the usual forward-type or avalanche-type radiation, and it is a function of both the tunnel junction voltage and the injection current. The experiment further shows that the observed radiation is stronger for units with lower Zener breakdown voltage. Using the calculation by Kane,[17] we find that for a Ge junction with an impurity concentration of 10^{19}/c. c., the eletrons in tunneling across the junction have a transmission coefficient of e^{-12} at a junction barrier of 1 ev. To avoid heating of the junction due to intra-valley scattering, it is desirable to keep the tunneled electrons within a small energy range above E(000). In other words, an impurity concentration about 10^{19}/c. c. or greater is needed for the tunnel junction. We also like to add that if we do not need separate controls of the tunnel and injection junctions, we can simply connect the pnp structure as a two-terminal device with the base floating. This may offer a great simplification in fabricating the device. However, in this case, impurity concentrations should be so chosen that the injected hole concentration should roughly match the tunneled electron concentration for maximum efficiency. A variation of the present scheme, probably a less efficient one, is to operate a junction diode in a region where both tunneling and avalanche mechanisms are significant and equally important, the former supplying the right kind of electrons while the latter furnishing the necessary holes for direct radiative recombination process.

In conclusion, the proposed scheme offers a definite possibility of obtaining laser action in semiconductors where the (000) valley of the conduction band is slightly above the conduction band minimum. The scheme also provides means for separate modulation of the amplitude and the frequency of the laser output. So far as the laser levels are concerned, we believe there is a definite advantage to choose the donor and acceptor states. To insure success of the proposed scheme, it is necessary to operate the device at very low temperatures and to choose the right kind of doping impurities so that scattering of electrons by lattice phonons and by impurity cell potentials from the (000) valley into the conduction band minimum is not prohibitively fast.

VII. REFERENCES

1. R. N. Hall, *et al.*, *Phys. Rev. Letters*, Vol. 9, 366 (1962).
2. M. I. Nathan, *et al.*, *Appl. Phys. Letters*, Vol. 1, 62 (1962).
3. N. Holonyak, Jr. and S. F. Bevacqua, *Appl. Phys. Letters*, Vol. 1, 82 (1962).
4. T. M. Quist, *et al.*, *Appl. Phys. Letters*, Vol. 1, 91 (1962).
5. H. Ehrenreich, *Jour. Appl. Phys.*, Vol. 32, 2155 (1961).
6. F. Herman, *Proc. IRE*, Vol. 43, 1703 (1955).
7. J. I. Pankove and M. Massoulie, *Bull. Am. Phys. Soc.*, Vol. 7, 88 (1962).
8. C. Kittel and A. H. Mitchell, *Phys. Rev.*, Vol. 96, 1488 (1954).
9. W. Van Roosbroeck and W. Shockley, *Phys. Rev.*, Vol. 94, 1558 (1954).
10. J. V. Morgan and E. O. Kane, *Phys. Rev. Letters*, Vol. 3, 466 (1959).
11. R. N. Hall and J. H. Racette, *Jour. Appl. Phys.*, Vol. 32, 2078 (1961).
12. J. Bardeen, F. J. Blatt and L. H. Hall, *Proc. Atlantic City Photoconductivity Conf.*, 1954, John Wiley & Sons, Inc., New York, 1955.
13. L. M. Roth, B. Lax and S. Zwerdling *Phys. Rev.*, Vol. 114, 90 (1959).
14. W. C. Dash and R. Newman, *Phys. Rev.*, Vol 99, 1151 (1955).
15. G. G. Macfarlane, T. P. McLean, J. E. Quarrington, and V. Roberts, *Phys. Rev.*, Vol. 108, 1377 (1957).
16. W. P. Dumke, *Phys. Rev.*, Vol. 105, 139 (1957).
17. E. O. Kane, *Jour. Appl. Phys.*, Vol. 32, 83 (1961).

18. S. Zwerdling, B. Lax, L. M. Roth and K. J. Button, *Phys. Rev.*, Vol. 114, 80 (1959).

19. G. Weinreich, T. M. Sanders, Jr., and H. G. White, *Phys. Rev.* Vol. 114, 33 (1959).

20. N. Holonyak, L. A. Lesk, R. N. Hall, J. J. Tiemann and H. Ehrenreich, *Phys. Rev. Letters*, Vol. 3, 167 (1959).

21. H. Fritzsche and J. J. Tiemann, *Phys. Rev.*, Vol. 130, 617 (1963).

22. M. I. Nathan and G. Burns, *Appl. Phys. Letters,* Vol. 1, 89 (1962).

23. D. F. Nelson, M. Gershenzon, A. Ashkin, L. A. D'Asaro and J. C. Sarace, *Appl. Phys. Letters*, Vol. 2, 9, (1963).

24. W. Kohn, *Advances in Solid State Physics,* Vol 5, 258 (1957).

25. E. M. Conwell, *Phys. Rev.*, Vol. 103, 51 (1956).

26. A. A. Gutkin and D. N. Nasledor, *Fig. Tver. Tela.*, Vol. 4, 1360 (1962).

27. T. S. Moss, *Jour. Appl. Phys.*, Vol 32, 2136 (1961).

28. A. L. Schawlow and C. H. Townes, *Phys. Rev.*, Vol. 112, 1940 (1958).

29. S. Wang, S. Yee and K. Nosaka, *Appl. Phys. Letters*, Vol. 2, 149, (1963).

VIII. FIGURE CAPTIONS

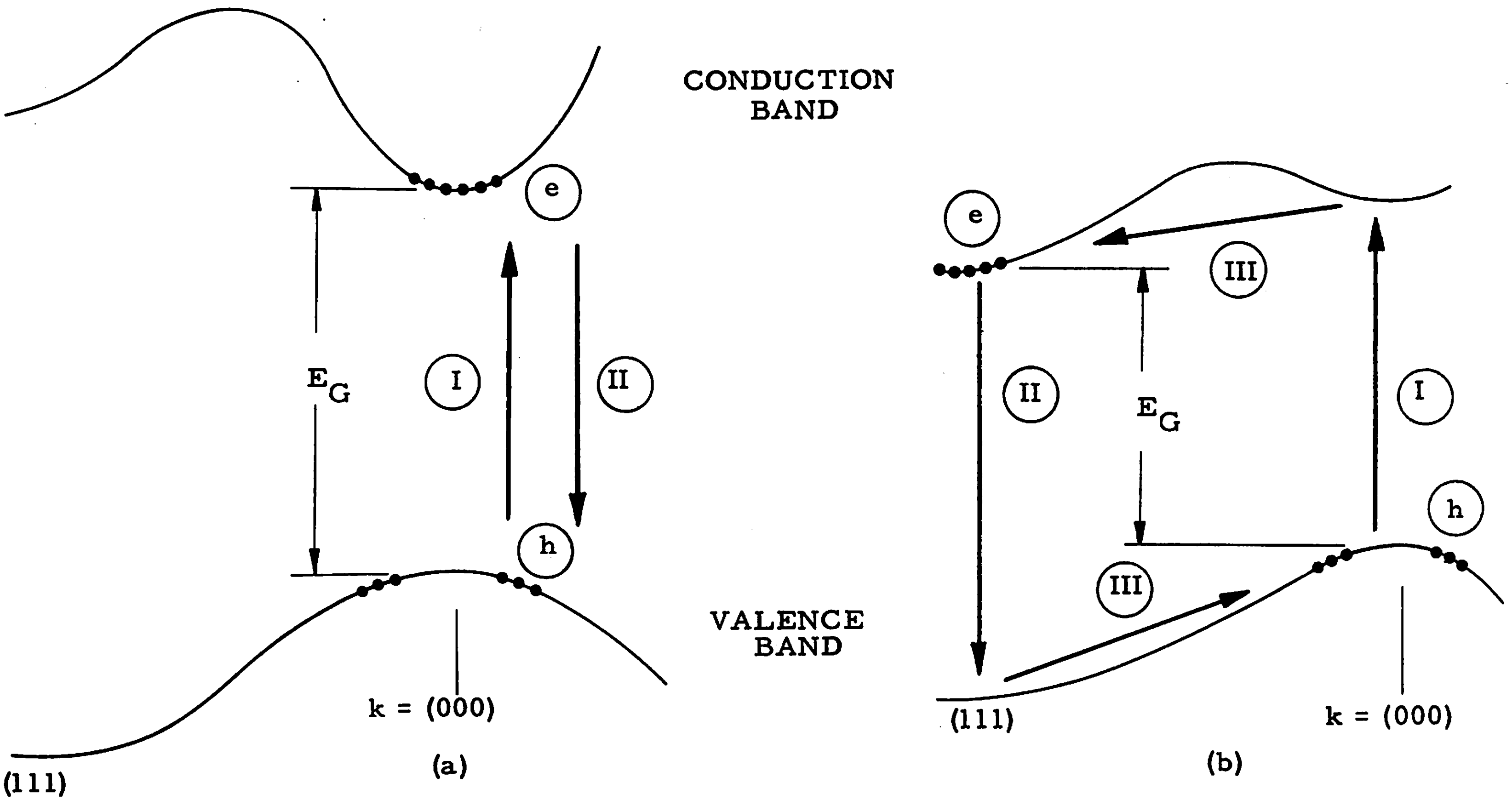
Fig. 1. Sketch of band structures showing the conduction band minimum and the valence band maximum for two types of semiconductors: (a) In GaAs-like semiconductors, direct transition between the conduction and valence band is possible; (b) In Ge-like simiconductors, ordinary absorption or emission of optical radiation must involve phonons.

Fig. 2. (a) Energy level diagram of a semiconductor where the subscripts c, d, f, a, and v denote the conduction band edge, the donor state, the Fermi level, the acceptor state and the valence band edge, respectively. (b) Transitions between donor and acceptor states. (c) Transitions between the conduction band and valence band states, E_1 and E_2 being the energy of the states under consideration.

Fig. 3. Schematic presentation of possible laser levels (a) between states in the conduction and valence band, (b) between the donor and acceptor states. In case (b), the conduction and valence band merely serve as a reservoir to replenish electrons and holes to the donor and acceptor states, respectively.

Fig. 4. Proposed scheme for laser action in Ge. Electrons are tunneled from the p-side (left-hand side) into the (000) valley of the conduction band of the n-region. In the meantime, holes are injected from the right-hand side into the valence band of the n-region. Then, direct, radiative, band-to-band transition may take place.

Fig. 5. (a) Schematic diagram showing the direct and phonon-assisted tunneling process. (b) The conductance curve of a tunnel diode. The break which occurs at $E_{app} = E(000) - E_{fn}$ indicates the onset of direct tunneling, E_{fn} being the Fermi level in the n-region and E(000) being the minimum energy of the (000) valley.



(a)

(b)

Symbols Used:

(e)	Electrons
(h)	Holes
E_G	Energy Gap

Processes Involved:

(I)	Absorption of Photons
(II)	Emission of Photons
(III)	Absorption or Emission of Phonons

Fig. 1

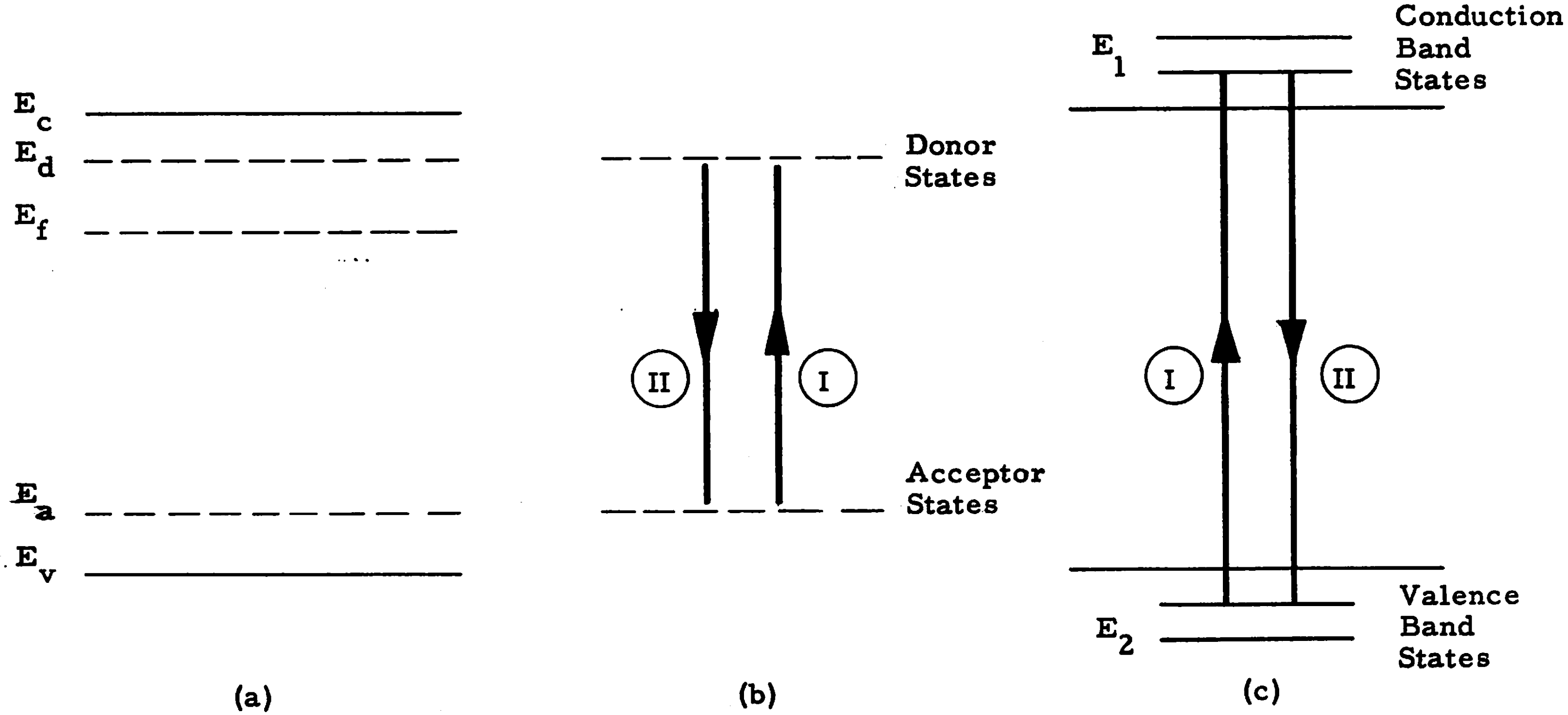

E_c
E_d
E_f
E_a
E_v
(a)
Donor States
II
I
Acceptor States
(b)
E_1
Conduction Band States
I
II
E_2
Valence Band States
(c)

Fig. 2

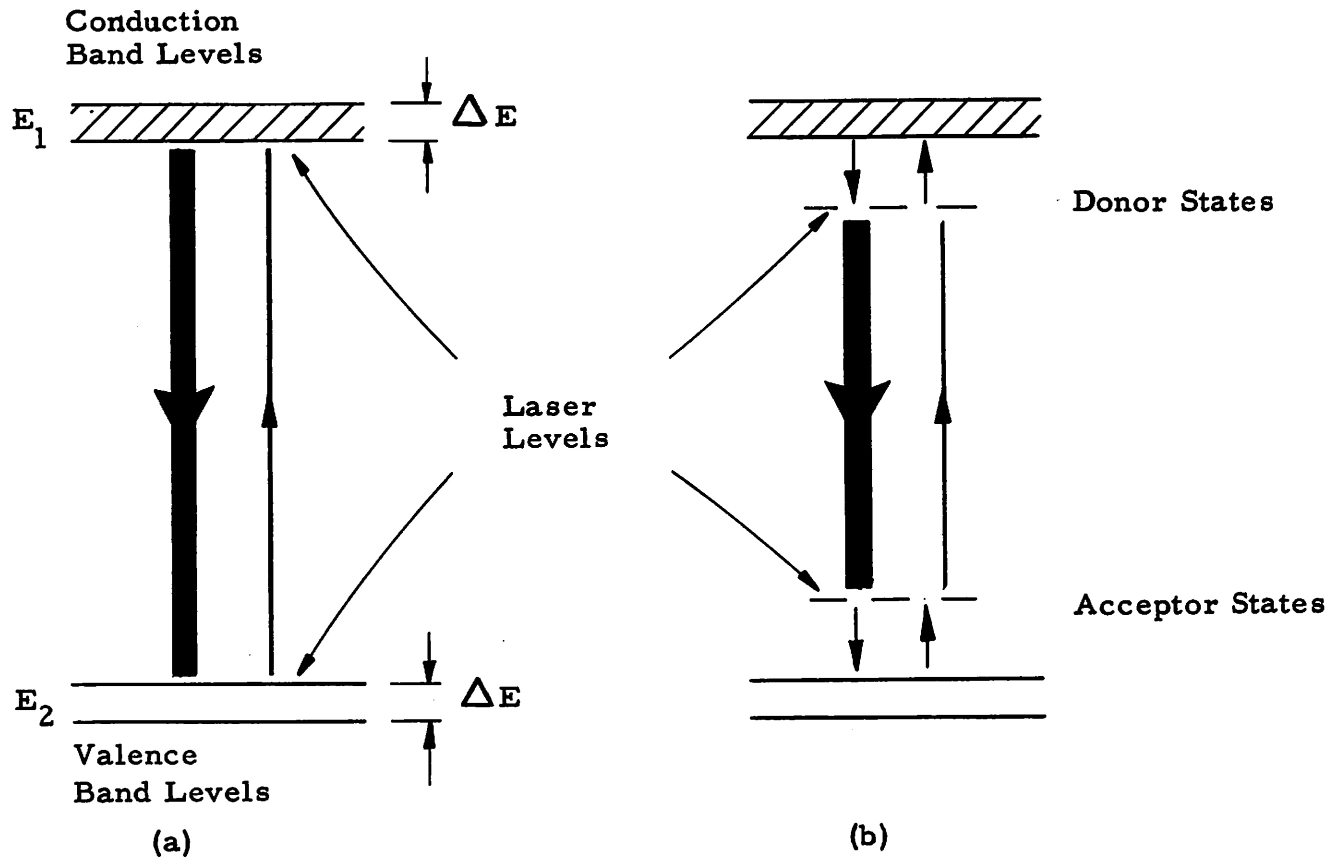
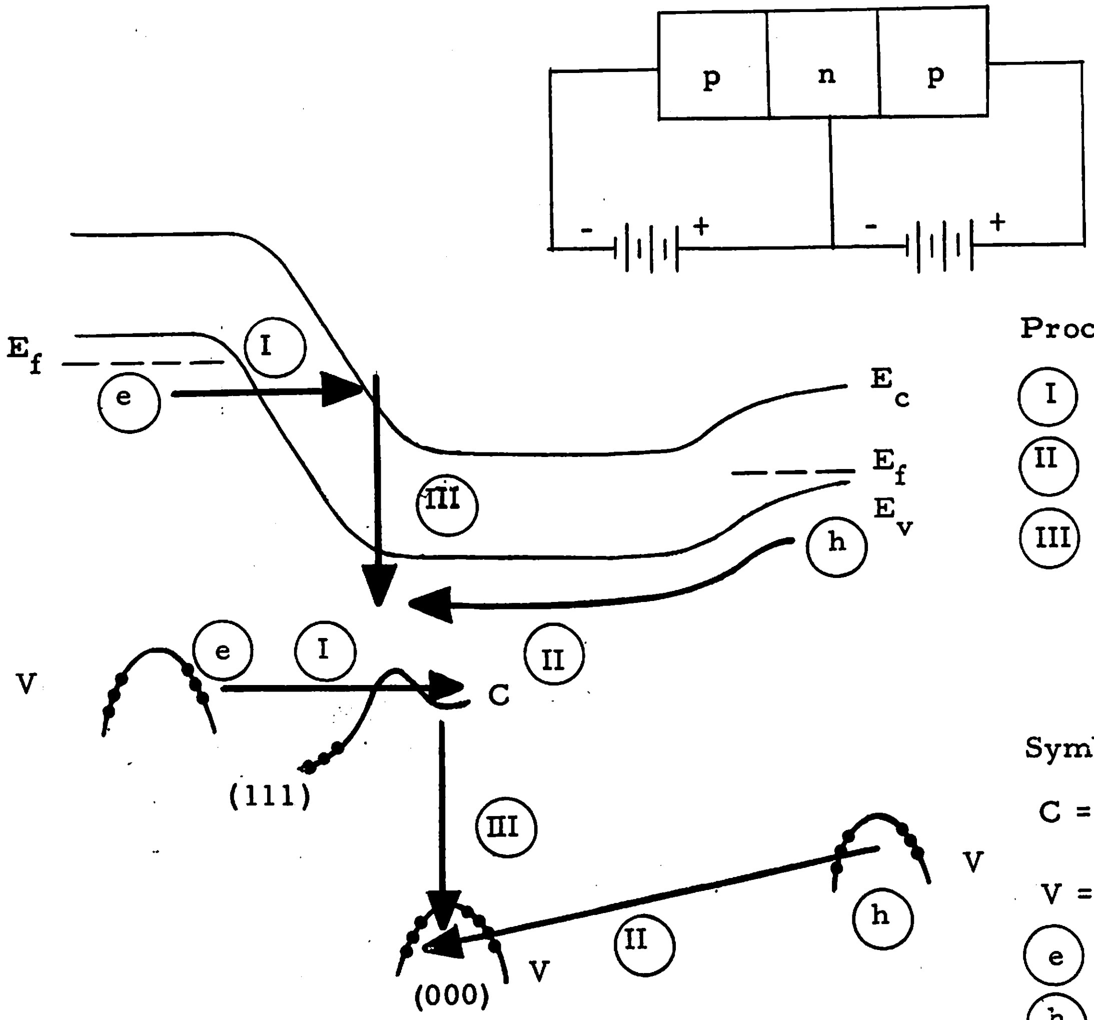

Conduction
Band Levels
E_1
ΔE
Laser
Levels
Donor States
Acceptor States
E_2
ΔE
Valence
Band Levels
(a)
(b)

Fig. 3



Processes Involved:

I	Electron Tunneling
II	Hole Injection
III	Direct Radiative Recombination

Symbols Used:

C = Conduction Band

V = Valence Band

ⓔ	Electrons
ⓗ	Holes

Fig. 4

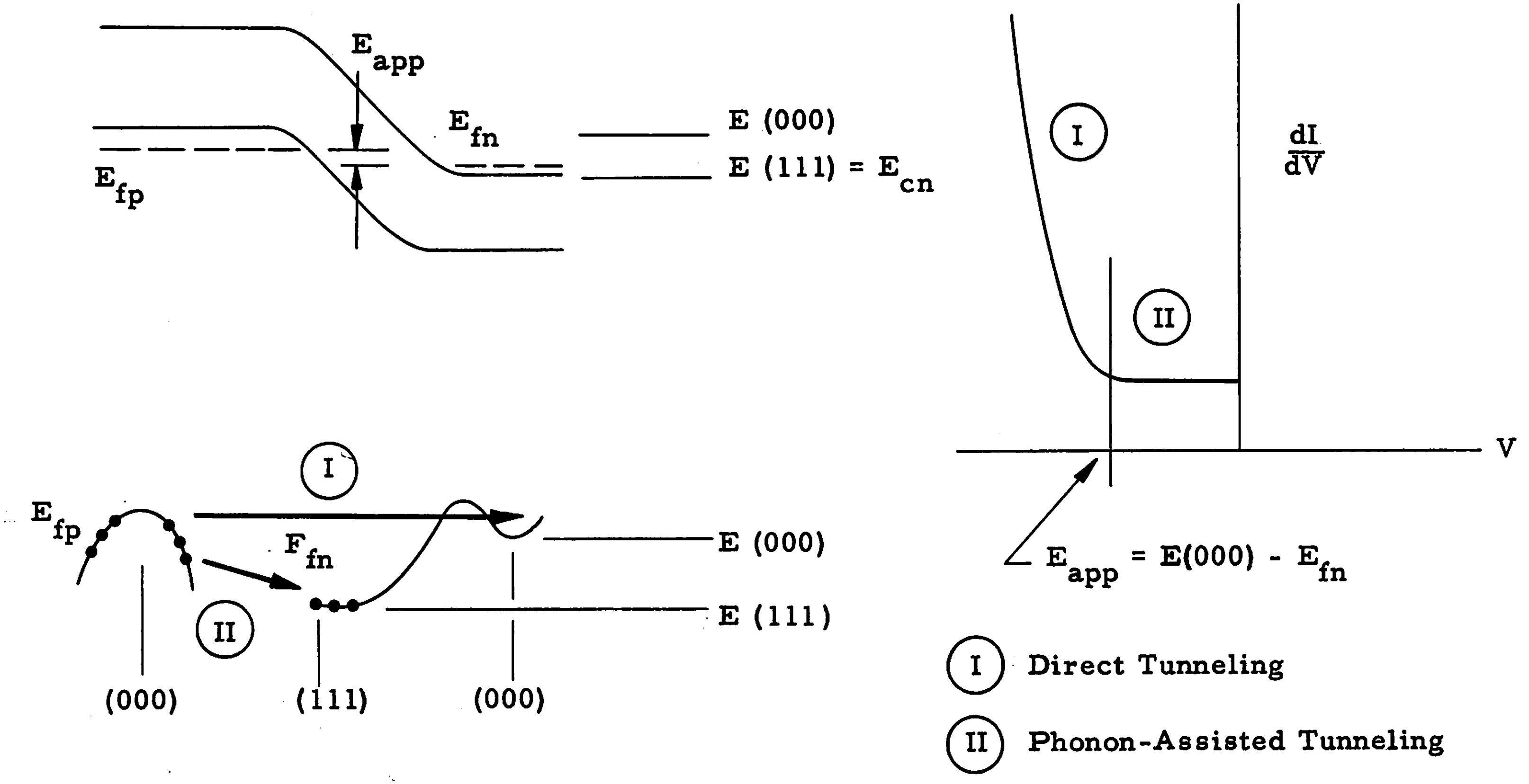

E_{app}
E_{fn}
E_{fp}
E (000)
E (111) = E_{cn}
I
II
E_{fp}
F_{fn}
E (000)
E (111)
(000)
(111)
(000)
$\frac{dI}{dV}$
V
E_{app} = E(000) - E_{fn}
I Direct Tunneling
II Phonon-Assisted Tunneling

Fig. 5