Electron-hole—pair generation and hysteresis in semiconductors

A. Elçi and M. O. Scully
Max-Planck-Institut für Quantenoptik, D-8046 Garching bei München, West Germany and Institute for Modern Optics, Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131

J. M. O'Hare
Department of Physics, University of Dayton, Dayton, Ohio 45469
(Received 22 July 1982; revised manuscript received 17 December 1982)

We consider the electron-hole—pair generation and the resulting free-carrier-density buildup in a semiconductor irradiated with laser beams whose frequencies are below the band gap. Using a simple model, we show that the free-carrier density may exhibit hysteresis and that when placed in a cavity, the semiconductor causes optical multistability at relatively low input cw intensities. For InSb the model predicts hysteresis at or above ~ 100 W/cm^2.

I. INTRODUCTION

Sometime ago it was observed that electron-hole—pair generation plays an important role in the nonlinear absorption properties of a semiconductor irradiated with a laser beam which has a frequency less than the band gap.[1] Specifically, it was seen that even for relatively low-intensity pulsed beams, pair generation can cause the formation of an avalanche in the density of free carriers and thus alter the optical response of the semiconductor. In the present paper we consider this phenomenon in the context of optical multistability. In particular, we show that when the semiconductor is irradiated by a laser beam whose frequency is less than the band gap, the density of electrons and holes can exhibit hysteresis as a result of rapid pair generation, and that if the semiconductor is in a ring cavity, the optical output of the cavity displays multistability.

There is usually a population of mobile carriers even in relatively pure crystals at low temperatures. These carriers, even though small in number, play an important role in the response of a semiconductor to a laser field whose frequency does not allow single-photon interband transitions. For the sake of definiteness, let us consider a slightly n-type material whose band structure is something like the one shown in Fig. 1. There is then a small population of mobile electrons at the bottom of the conduction band. These electrons can absorb photons via emission or absorption of phonons, which is free-carrier absorption and does not depend on the light frequency in a significant way. When the sample is irradiated, free-carrier absorption causes the temperature of electrons to rise and form a hot electron population in the conduction band. Energetic electrons normally relax by phonon emission. However, an electron which has kinetic energy greater than the band gap can preferentially decay to the bottom of the conduction band by producing a pair, that is, by exciting a valence electron to the conduction band.[2,3] The mechanism responsible for this pair generation is the binary Coulomb collisions between conduction and valence electrons. Such a relaxation has a strong resemblance to Auger processes in that it is the inverse of the Auger recombination. Thus, via the heating of the original free carriers, more free carriers are created by pair generation, and they in

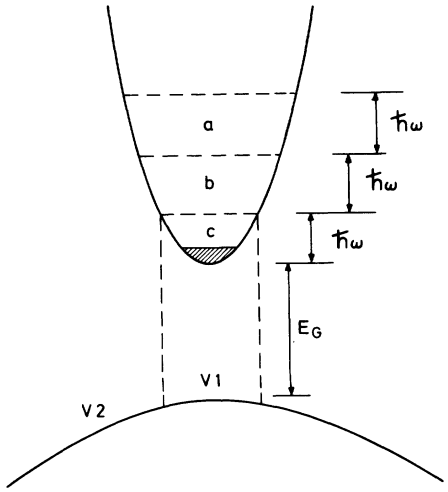

FIG. 1. Bucket model for a two-band semiconductor.

turn can absorb more photons via phonon emission or absorption and thereby create more hot carriers. Unless balanced by electron-hole recombinations, this process can lead to an avalanche growth in the free-carrier density.

In Ref. 3, a simple "bucket" model was developed to describe the behavior of InSb (at $T=4$ K) under intense and pulsed CO_2-laser beams. In this problem the band gap is larger than twice the laser frequency and neither single-photon nor two-photon interband transitions are allowed due to energy conservation. To determine the optical response of InSb to CO_2-laser beams, the bucket model partitions the conduction and valence bands into various regions according to photon energy, and uses a simple set of equations to obtain the particle transfer rates among these regions, or "buckets." Despite its gross simplification of the problem, one may have confidence in this model on the basis of two facts. First, the model leads to results which are in reasonable agreement with experiments in which the beam is pulsed.[3] Second, more precise analysis of the same problem with appropriate electron and hole distributions, which are continuous and energy dependent throughout the bands, has generally confirmed the results of the model.[4] A further advantage of the bucket model is that it can easily be adapted to other semiconductors with different band structures. In the present paper we extend the use of the bucket model to the case in which the laser beam is cw and the semiconductor is placed in a ring cavity. For the sake of definiteness, we also consider InSb, with the simplified band structure as shown in Fig. 1. However, the semiconductor is now irradiated by cw CO_2-laser beams and steady-state conditions prevail.

In Sec. II we study the bucket model for $I<I_{c2}$, where I is the average light intensity inside the semiconductor and I_{c2} is some critical intensity. I_{c2} is on the order of 500 W/cm^2 for InSb. In the problem there appears another critical intensity I_{c1}, such that the model gives two possible values for the free-carrier density if $I_{c1}<I<I_{c2}$. This result is illustrated in Fig. 3 and demonstrates that the free-carrier density can exhibit optically induced hysteresis due to the pair generation. I_{c1} is on the order of 100 W/cm^2 for InSb. For $0<I<I_{c1}$, the model gives a single value for the free-carrier density and no hysteresis is observed. The high-intensity regime corresponding to $I>I_{c2}$ is not investigated in the present paper. In Sec. III we consider the situation in which the sample is placed in a ring-cavity configuration as shown in Fig. 4. We use the Drude theory to take into account the dispersive and absorptive effects of the sample on the field inside the cavity. One can distinguish two types of behavior in the response of the semiconductor. For large detunings from a cavity frequency, the nonlinearity is mostly dispersive in nature, and input-output intensities display hysteresis. When the laser frequency is resonant with a cavity frequency, the nonlinearity is absorptive, and the input-output intensity curve corresponding to the higher density displays bistability by itself. As one varies the frequency detuning, one observes both types of behavior in conjunction.

A final remark here concerns possible practial applications of this hysteresis of the free-carrier density. As the discussion above makes clear, none of the processes which it involves is a resonance phenomenon. The hysteresis is therefore relatively insensitive to light frequency. This is especially important from a practical point of view, since it may lead to a broad-band bistable system.

II. HYSTERESIS OF FREE-CARRIER DENSITY

In this section we briefly describe the bucket model for InSb and solve the relevant rate equations in the steady state. The bucket model for InSb has only two bands, as shown in Fig. 1: a heavy hole band and a conduction band. The model divides the conduction band into three regions: a, b, and c. These regions have equal energy spread, each roughly $\hbar\omega$, where ω is the laser frequency. Starting from c, electrons can go to b and a by free-carrier absorption. Electrons reaching region a can create electron-hole pairs and go to c. Electrons that are in c can recombine with holes in the valence band. The model also divides the valence band into two separate regions, v_1 and v_2. The size of v_1 is supposed to be roughly the same as that of region a. Region v_1 does not contribute to the pair excitation, whereas v_2 does. On the other hand, v_1 is involved in electron-hole recombinations, but v_2 is not. In all regions, whether in the conduction or valence band, electrons relax by phonon emission and go to lower-energy states.

The bucket model uses parabolic bands. If the capital letters N_a, N_b, N_c, N_{v1}, and N_{v2} designate the number of states per unit volume in the various regions described above, then

$$N_c=\int_{E_G}^{E_G+\hbar\omega} dE(2\pi^2)^{-1}(2m_c^*/\hbar^2)^{3/2}(E-E_G)^{1/2}$$
$$=(3\pi^2)^{-1}(2m_c^*\omega/\hbar)^{3/2}, \quad (1)$$

where m_c^* is the effective mass for the conduction band. Let the sample be at $T=4$ K. Then $m_c^*\simeq 0.014m_0$ and $N_c\simeq 3.4\times 10^{17}$ cm^{-3}. Similarly,

$$N_b=(3\pi^2)^{-1}(2m_c^*\omega/\hbar)^{3/2}[(2\omega)^{3/2}-(\omega)^{3/2}]$$
$$\cong 1.8N_c \tag{2}$$

and

$$N_a=(3\pi^2)^{-1}(2m_c^*\omega/\hbar)^{3/2}[(3\omega)^{3/2}-(2\omega)^{3/2}]$$
$$\cong 2.4N_c\ . \tag{3}$$

For the valence-band electrons, from Fig. 1 it follows that

$$N_{v1}=(3\pi^2)^{-1}(2m_h^*E_0/\hbar^2)^{3/2}\cong N_c\ , \tag{4}$$

where m_h^* is the effective hole mass, $m_h^*\simeq 0.4m_0$, and $E_0=m_c^*\hbar\omega/m_h^*$. N_{v2} is taken to be the same as N_a.

Let the lower case letters n_a, n_b, n_c, n_{v1}, and n_{v2} designate the actual electron densities in their respective regions. There are five rate equations for these quantities, which in the steady state become

$$0=\dot{n}_a=\gamma_l[(N_a-n_a)n_b-(N_b-n_b)n_a]$$
$$-\gamma_c n_a(N_b-n_b)$$
$$-\gamma_p n_a n_{v2}(N_c-n_c)^2\ , \tag{5}$$

$$0=\dot{n}_b=-\gamma_l[(N_a-n_a)n_b-(N_b-n_b)n_a]$$
$$+\gamma_l[(N_b-n_b)n_c-(N_c-n_c)n_b]$$
$$+\gamma_c n_a(N_b-n_b)-\gamma_c n_b(N_c-n_c)\ , \tag{6}$$

$$0=\dot{n}_c=-\gamma_l[(N_b-n_b)n_c-(N_c-n_c)n_b]$$
$$+2\gamma_p n_a n_{v2}(N_c-n_c)^2$$
$$+\gamma_c n_b(N_c-n_c)-\gamma_{cv}(N_{v1}-n_{v1})\ , \tag{7}$$

$$0=\dot{n}_{v1}=\gamma_{cv}(N_{v1}-n_{v1})-\gamma_v n_{v1}(N_{v2}-n_{v2})\ , \tag{8}$$

and

$$0=\dot{n}_{v2}=-\gamma_p n_a n_{v2}(N_c-n_c)^2$$
$$+\gamma_v n_{v1}(N_{v2}-n_{v2})\ . \tag{9}$$

The various terms here have the following meanings: $\gamma_{cv}(N_{v1}-n_{v1})$ represents electron-hole recombinations. Recombinations are assumed to affect only regions c and v_1. The terms $\gamma_c n_a(N_b-n_b)$ and $\gamma_c n_b(N_c-n_c)$ represent thermal relaxation in the conduction band. $\gamma_v n_{v1}(N_{v2}-n_{v2})$ describes thermal relaxation in the valence band. $\gamma_p n_a n_{v2}(N_c-n_c)^2$ describes the pair excitation process which is assumed to involve only the a, c, and v_2 regions. The terms that are proportional to γ_l, such as

$$\gamma_l[(N_a-n_a)n_b-(N_b-n_b)n_a]\ ,$$

represent free-carrier absorption. The model assumes that free-carrier absorption takes place in the conduction band but not in the valence band. γ_c, γ_v, γ_p, and γ_{cv} designate the rate coefficients for relaxation of conduction electrons, relaxation of valence electrons, pair excitation, and recombination, respectively. γ_l is a compound rate for free-carrier absorption and is proportional to the laser intensity. Equations (5)–(9) are five equations for five unknowns. One can substitute the expression for particle conservation, which is much easier to deal with, for one of these equations. Summing all five rate equations for $\dot{n}_a$, $\dot{n}_b$, etc., one finds that

$$n_a+n_b+n_c+n_{v1}+n_{v2}=N_{v1}+N_{v2}+n_0 \tag{10}$$

at any time t, and in particular for the steady state at $t\to+\infty$. The right-hand side of (10) follows from the fact that before the light beam is turned on (which corresponds to $t\to-\infty$), $n_a=n_b=0$, $n_{v1}=N_{v1}$, $n_{v2}=N_{v2}$, and $n_c=n_0$. We assume that n_0 is on the order of 10^{16} cm^{-3}. We can also obtain two other equations from (5)–(7) that are more convenient to work with than (5) and (7). Summing (5)–(7),

$$\gamma_p n_a n_{v2}(N_c-n_c)^2=\gamma_{cv}(N_{v1}-n_{v1})\ . \tag{11}$$

This equation states that, in the steady state, the number of electron-hole pairs lost due to recombination must equal the number of pairs created by hot electrons. Now using (11) in (7), we find

$$\gamma_l(N_b n_c-N_c n_b)=\gamma_c n_b(N_c-n_c)$$
$$+\gamma_{cv}(N_{v1}-n_{v1})\ . \tag{12}$$

Keeping (6) and (8) as they are, we now have a new set of five equations, (10)–(12), (6), and (8), for five unknowns.

At this juncture we will make an approximation that follows from the extreme rapidity of the pair excitation process.[3] It turns out that as hot electrons move into region a, they are quickly scattered back to region c via pair excitation. It turns out that the rate for pair generation for InSb is nearly 5 to 6 orders of magnitude larger than the rate for recombination.[3] Thus generally, $n_a \ll n_b, n_c$ even for relatively large laser intensities. We therefore neglect n_a in the various sums that involve n_b, n_c, n_{v1}, and n_{v2}, and the working equations become

$$n_b+n_c+n_{v1}+n_{v2}=N_{v1}+N_{v2}+n_0\ , \tag{13}$$

$$\gamma_l(N_b n_c-N_c n_b)=\gamma_c n_b(N_c-n_c)$$
$$+\gamma_{cv}(N_{v1}-n_{v1})\ , \tag{14}$$

$$\gamma_l(N_b n_c-N_c n_b)=\gamma_l N_a n_b$$
$$+\gamma_c n_b(N_c-n_c)\ , \tag{15}$$

$$\gamma_{cv}(N_{v1}-n_{v1})=\gamma_v n_{v1}(N_{v2}-n_{v2})\ , \tag{16}$$

and

$$\gamma_p n_a n_{v2}(N_c-n_c)^2=\gamma_{cv}(N_{v1}-n_{v1}) . \quad (17)$$

Equation (17), which determines n_a, will not be needed as long as $n_a \ll n_b, n_c$, since the total conduction-band electron density is essentially given by n_b+n_c.

To solve the system of equations (13)–(16), we start from (16), which gives

$$N_{v2}-n_{v2}=(\gamma_{cv}/\gamma_v)(N_{v1}-n_{v1})n_{v1}^{-1} . \quad (18)$$

Substituting (18) and (13), we obtain a second-order equation for n_{v1} in terms of n_b and n_c:

$$n_{v1}^2+n_{v1}(n_b+n_c-n_0-N_{v1}+\gamma_{cv}/\gamma_v) -(\gamma_{cv}/\gamma_v)N_{v1}=0 . \quad (19)$$

This equation has one positive root and one negative root. The positive root gives n_{v1}:

$$n_{v1}=-\tfrac{1}{2}(n_b+n_c-n_0-N_{v1}+\gamma_{cv}/\gamma_v) +\tfrac{1}{2}[(n_c+n_b-n_0-N_{v1}+\gamma_{cv}/\gamma_v)^2 +4\gamma_{cv}N_{v1}/\gamma_v]^{1/2} . \quad (20)$$

From (14) and (15) we find

$$N_{v1}-n_{v1}=\gamma_l N_a n_b/\gamma_{cv} . \quad (21)$$

Equations (20) and (21), when combined, yield an equation for n_b in terms of n_c alone:

$$n_b^2-n_b\frac{[\gamma_{cv}^2N_{v1}-\gamma_l\gamma_{cv}N_a(N_{v1}+\gamma_{cv}/\gamma_v-n_0+n_c)]}{\gamma_l N_a(\gamma_{cv}-\gamma_l N_a)}-\frac{\gamma_{cv}^2N_{v1}(n_c-n_0)}{\gamma_l N_a(\gamma_{cv}-\gamma_l N_a)}=0 . \quad (22)$$

Let the roots of (22) be $n^{\pm}$ (which may be complex):

$$n^{\pm}=[\gamma_{cv}^2N_{v1}-\gamma_l\gamma_{cv}N_a(N_{v1}+\gamma_{cv}\gamma_v^{-1}+n_c-n_0)][2\gamma_l N_a(\gamma_{cv}-\gamma_l N_a)]^{-1} \pm\{\tfrac{1}{4}[\gamma_{cv}^2N_{v1}-\gamma_l\gamma_{cv}N_a(N_{v1}+\gamma_{cv}\gamma_v^{-1}+n_c-n_0)]^2\gamma_l^{-2}N_a^{-2}(\gamma_{cv}-\gamma_l N_a)^{-2} +\gamma_{cv}^2N_{v1}(n_c-n_0)\gamma_l^{-1}N_a^{-1}(\gamma_{cv}-\gamma_l N_a)^{-1}\}^{1/2} . \quad (23)$$

Only real and positive roots are acceptable for n_b. We distinguish two regimes for which $n^{\pm}$ are real: First,

$$\gamma_{cv}>\gamma_l N_a .$$

In this "low-intensity" regime, $n^-<0<n^+$, and therefore n^+ is the only acceptable root. Second,

$$\gamma_{cv}<\gamma_l N_a$$

and

$$[\gamma_{cv}N_{v1}-\gamma_l N_a(N_{v1}+\gamma_{cv}\gamma_v^{-1}+n_c-n_0)]^2>4\gamma_{cv}\gamma_l N_a N_{v1}(n_c-n_0)(\gamma_{cv}-\gamma_l N_a) .$$

In this regime both $n^{\pm}>0$, and either one of the roots, are possible values for n_b, provided that it is less than N_b (and also if the resulting equation for n_c gives an acceptable root).

The parameters for InSb are such that the internal light intensity which corresponds to $\gamma_{cv}=\gamma_l N_a$ is on the order of 10^5 MW/cm^2, which is much larger than $I_{c2}\sim 500$ W/cm^2. As we stated earlier, the discussion of the present paper is restricted to intensities less than I_{c2}, and therefore

$$n_b=n^+ . \quad (24)$$

Note that

$$\gamma_{cv}>\gamma_l N_a \Rightarrow \gamma_{cv}^2N_{v1}>\gamma_l N_a\gamma_{cv}(N_{v1}+\gamma_{cv}\gamma_v^{-1}+n_c-n_0) , \quad (25)$$

since n_c is supposed to be such that $n_c \geq n_0$.

Going back to our system of equations, we note that from Eq. (15) one can obtain another expression for n_b in terms of n_c:

$$n_b=\gamma_l N_b n_c\gamma_c^{-1}[N_c-n_c+\gamma_l(N_a+N_c)/\gamma_c]^{-1} . \quad (26)$$

Finally, using Eqs. (26) and (23), one finds a cubic equation for n_c alone. Let $x=n_c/n_0$. Then,

$$x^3+r_1x^2+r_2x+r_3=0\ , \tag{27}$$

where

$$\begin{aligned} r_1=d^{-1}\{&-\gamma_{cv}\gamma_c N_{v1}-\gamma_I^3 N_a N_b^2(\gamma_{cv}-\gamma_I N_a)(\gamma_c\gamma_{cv}n_0)^{-1}\\ &-\gamma_{cv}\gamma_I N_b N_{v1}n_0^{-1}+\gamma_I^2 N_a N_b n_0^{-1}(N_{v1}+\gamma_{cv}\gamma_v^{-1}-n_0)\\ &-(d+\gamma_c\gamma_{cv}N_{v1})[N_c n_0^{-1}+\gamma_I(N_a+N_c)\gamma_c^{-1}n_0^{-1}]\}\ , \end{aligned} \tag{28a}$$

$$\begin{aligned} r_2=&d^{-1}[N_c n_0^{-1}+\gamma_I(N_a+N_c)\gamma_c^{-1}n_0^{-1}]\\ &\times\{2\gamma_{cv}\gamma_c N_{v1}+\gamma_{cv}\gamma_I N_b N_{v1}n_0^{-1}\\ &+\gamma_{cv}\gamma_c N_{v1}[N_c n_0^{-1}+\gamma_I(N_a+N_c)\gamma_c^{-1}n_0^{-1}]-\gamma_I^2 N_a N_b n_0^{-1}(N_{v1}+\gamma_{cv}\gamma_v^{-1}-n_0)\}\ , \end{aligned} \tag{28b}$$

$$r_3=-d^{-1}\gamma_c\gamma_{cv}N_{v1}[N_c n_0^{-1}+\gamma_I(N_a+N_c)n_0^{-1}\gamma_c^{-1}]^2\ , \tag{28c}$$

and

$$d=\gamma_c\gamma_{cv}N_{v1}+\gamma_I^2 N_a N_b\ . \tag{28d}$$

An acceptable root of (27) must be real and equal or less than N_c/n_0, since by definition n_c cannot be larger than N_c. It must also reduce to 1.0 at exactly $I=0$, since $n_c=n_0$ before the field is applied [see Eq. (10)]. In Ref. 3 it is found that for InSb at $T=4$ K and for $n_0=10^{16}$ cm^{-3}, the parametric values $\gamma_v=10^{-10}$ cm^{3}sec^{-1}, $\gamma_c=2.2\times10^{-8}$ cm^{3}sec^{-1}, $\gamma_{cv}=3.0\times10^8$ sec^{-1}, and $\gamma_I=(1.1\times10^{-21}I)$ cm^{3}sec^{-1}, where I is the average field intensity inside the sample and is given in units of W/cm^2, give the best agreement with pulsed beam experiments.[5] For these values of parameters one finds that there exists a critical intensity I_{c1} such that for $I<I_{c1}$, the cubic equation may either be reducible or irreducible,[6] but there is only one acceptable root. For $I_{c1}<I<I_{c2}$, the cubic equation is irreducible throughout the range of I and there are two acceptable roots. Let

$$p=r_2-r_1^2/3\ , \tag{29a}$$

$$q=r_3-r_1r_2/3+2r_1^3/27\ , \tag{29b}$$

$$\Delta=-(4p^3+27q^2)/27\ , \tag{29c}$$

$$\phi_0=\tan^{-1}(-q^{-1}\Delta^{1/2})+(\pi/2)(1+\mathrm{sgn}q)\ , \tag{29d}$$

and

$$u=-q/2+(-27\Delta/108)^{1/2}\ . \tag{29e}$$

For $I<I_{c1}$, the acceptable root is given by

$$x_0=\begin{cases}2(-p/3)^{1/3}\cos(\phi_0/3+2\pi/3)-r_1/3 & \\ \qquad\qquad\text{if } \Delta>0 & \\ u^{1/3}-pu^{-1/3}/3-r_1/3 \quad \text{if } \Delta<0\ . & \end{cases} \tag{30a}$$

For $I_{c1}<I<I_{c2}$, the acceptable roots are given by

$$x_1=2(-p/3)^{1/2}\cos(\phi_0/3+2\pi/3)-r_1/3 \tag{30b}$$

and

$$x_2=2(-p/3)^{1/2}\cos(\phi_0/3+4\pi/3)-r_1/3\ . \tag{30c}$$

The critical intensity I_{c1} is determined from the transcendental equation

$$2(-p/3)^{1/2}\cos(\phi_0+4\pi/3)-r_1/3=N_c/n_0\ . \tag{31}$$

On the other hand, I_{c2} is determined from the equality of x_1 and x_2, which means

$$\sin(\phi_0/3)=0\ . \tag{32}$$

Figure 2 shows the allowed solutions for n_c/n_0 as

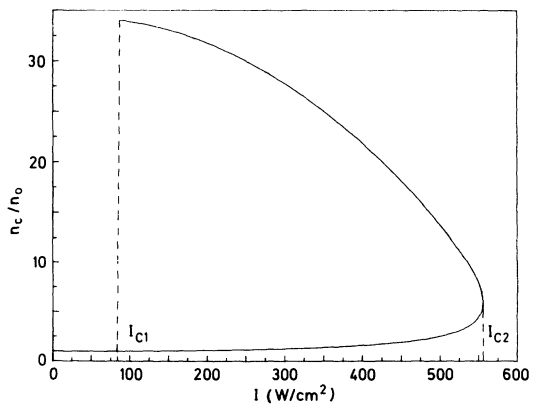

FIG. 2. Electron density in region c as a function of internal light intensity.

functions of I. One solution starts off from 1.0 and monotonically increases as I increases. It corresponds to x_0 for $I<I_{c1}$ and goes over smoothly to x_1 for $I>I_{c1}$. For $I<I_{c1}$, it is the only allowed solution for n_c/n_0. The second solution, which corresponds to x_2, is defined only for $I>I_{c1}$. It starts off from N_c/n_0 and decreases monotonically as I approaches I_{c2}. The two solutions meet at $I=I_{c2}$. Figure 3 shows the corresponding total electron density $n_e \simeq n_b + n_c$ solutions.

An immediate question concerns the stability of these solutions. One can perform a stability analysis using the rate equations (5)—(9) and assuming small fluctuations $n_b \rightarrow n_b + \delta n_b(t)$, $n_c \rightarrow n_c + \delta n_c(t)$, etc., where $\delta n_b \ll n_b$, $\delta n_c \ll n_c$, etc. One then finds that all fluctuations are damped, and that the solutions discussed above are all stable.

The solutions depict the following physical picture. For $I<I_{c1}$, the pair generation, although rapid, is sufficiently compensated by the recombination so that the free-carrier density does not increase significantly. Such a compensation also persists for $I>I_{c1}$, and the slow rise of the lower branch of the density curve in Fig. 3 illustrates this situation. However, as one approaches I_{c2}, the pair generation becomes more efficient and can lead to a density buildup. Furthermore, if I is then decreased, starting from I_{c2}, the pair generation protects the density buildup, and the system moves along the upper curve in Fig. 3. Finally, when I returns to I_{c1}, recombinations take over and there is a sharp decrease (which may be considered as an "avalanche recombination") in the density of free carriers. The system returns to the original lower-density branch. Note that hole relaxation plays an important role in the hysteresis. The rate for hole relaxation is considerably less than the rate for recombination, and therefore hole relaxation acts as a bottleneck in the transfer of electrons from v_1 to v_2. This reduces the efficiency of the entire recombination process and allows a buildup in the density of free carriers. Consistent with this picture, the value of I_{c1} is sensitive to changes in γ_v, and also to changes in the size of N_{v1} relative to the other densities of states. Doubling γ_v increases I_{c1} by about 70% and halving γ_v decreases I_{c1} by about 50%, if all the other parameters are kept unchanged. On the other hand, doubling γ_{cv} increases I_{c1} by less than 6%, and halving it decreases I_{c1} by less than 5%. When N_{v1}/N_c is varied instead of γ_v, the changes induced in I_{c1} are similar to those when γ_v is varied. Finally, we note that free-carrier absorption can retard the relaxation of holes to the band edge more effectively at higher intensities. This somewhat abates the efficiency of the pair generation and causes the density of electrons to decrease along the high-density branch in Fig. 3 as I increases from I_{c1} to I_{c2}.

A brief remark here concerns $I>I_{c2}$. When $I>I_{c2}$, one generally does not get an acceptable solution for n_c from the cubic equation, in that all three roots are real but are larger than N_c. We believe that this is due to the breakdown of the approximation $n_a \ll n_b, n_c$. As we mentioned above, free-carrier absorption becomes more efficient in retarding relaxation of free carriers at high intensities. This leads to a population buildup even in region a. It follows from (10) that the fractional density alotted to region c is decreased if n_a is not neglected (both n_a and n_{v1} increase at high intensities). Thus one should actually work with a biquadratic equation near and above I_{c2}.

III. RING-CAVITY CONFIGURATION

In this section we assume that the semiconductor is placed in a ring cavity as shown in Fig. 4, and use

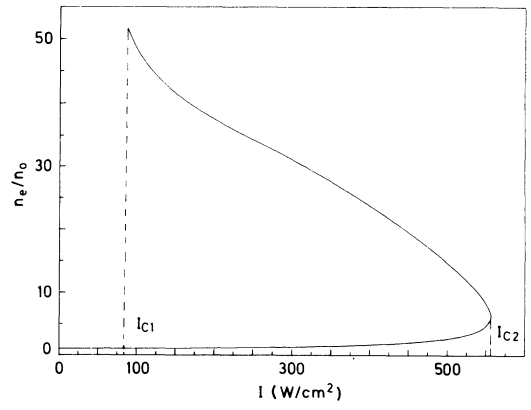

FIG. 3. Total electron density in the conduction band vs I.

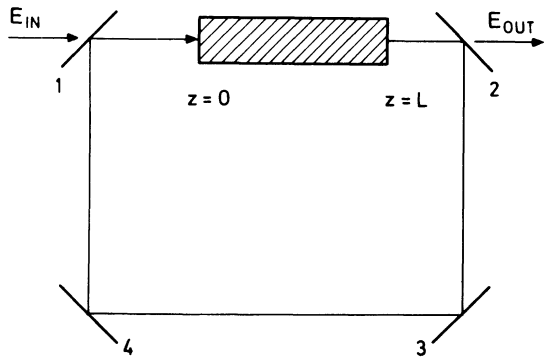

FIG. 4. Ring-cavity configuration. The mirrors at 3 and 4 have 100% reflectivity.

the Drude theory and the mean-field approximation to take into account the dispersive and absorptive effects of the sample on the cavity fields. According to the Drude theory, the dielectric function of the semiconductor can be written as

$$\epsilon(\omega)=\epsilon_1+i\epsilon_2 =\epsilon_\infty\left[1-\frac{\omega_p^2\tau^2}{\omega^2\tau^2+1}+i\frac{\omega_p^2\tau}{\omega(\omega^2\tau^2+1)}\right], \quad (33)$$

where ϵ_∞ is the high-frequency dielectric constant ($\epsilon_\infty=16.0$ for InSb), τ is an effective scattering time, ω is the light frequency, and ω_p is the plasma frequency of the electron-hole plasma of the free carriers, given by

$$\omega_p^2=4\pi e^2\epsilon_\infty^{-1}\left[\frac{n_e}{m_c^*}+\frac{n_h}{m_h^*}\right]\cong\frac{4\pi e^2 n_e}{\epsilon_\infty m_c^*} . \quad (34)$$

Here $n_e\simeq n_b+n_c$, and n_h is the density of the holes, i.e.,

$$n_h=N_{v1}+N_{v2}-n_{v1}-n_{v2} .$$

The last equality in (34) follows from the fact that $m_c^*\ll m_h^*$ and $n_e\simeq n_h$. Thus ϵ in the Drude form depends on the density of the free carriers through the plasma frequency. From the preceding discussion it follows that in the bucket model, ϵ may depend on I in two distinct functional forms. Apart from its functional relation to I via ω_p, ϵ may further depend on I via τ, whose numerical value is adjusted to give the actual losses in the crystal. In Ref. 3 it was found that the best agreement with the pulsed beam experiments is obtained for

$$\tau=2000(n_e/n_0)^{-1/3}\omega^{-1} . \quad (35)$$

We also use (35) for τ.

Let us assume that the semiconductor is placed in a ring cavity as shown in Fig. 4. To simplify the problem we assume that the sample fills the cavity entirely and that the fields are linearly polarized plane waves. In order to obtain the equation of propagation for the field amplitudes, we closely follow the method of Ref. 7, and write the fields in the form

$$\vec{\mathrm{E}}=\hat{x}E(z,t)e^{ik_c z}e^{-i\omega t} \quad (36a)$$

and

$$\vec{\mathrm{D}}=\hat{x}D(z,t)e^{ik_c z}e^{-i\omega t} . \quad (36b)$$

Here $E(z,t)$ and $D(z,t)$ are the slowly varying amplitudes of the electric and displacement fields. k_c is the wave vector related to a cavity frequency Ω_c, such that

$$\Omega_c\cong c\epsilon_\infty^{-1/2}k_c\equiv c'k_c , \quad (37)$$

where c' is the speed of light in the semiconductor. Substituting (36a) and (36b) into the Maxwell equation of propagation,

$$\nabla^2\vec{\mathrm{E}}=c^{-2}\partial_t^2\vec{\mathrm{D}} , \quad (38)$$

and using the fact that $E(z,t)$ and $D(z,t)$ are slowly varying, one finds

$$c'\partial_z E+\partial_t E=-i(\Omega_c-\omega)E +i(\omega/2)(\epsilon-\epsilon_\infty)\epsilon_\infty^{-1}E . \quad (39)$$

To obtain (39), we neglected all second derivatives of E and D, and used

$$D=\epsilon E , \quad (40a)$$

$$\partial_t D\cong\epsilon_\infty\partial_t E , \quad (40b)$$

$$(\omega/c')\partial_t E\cong k_c\partial_t E , \quad (40c)$$

$$\omega^2(\epsilon-\epsilon_\infty)E\cong\omega\Omega_c(\epsilon-\epsilon_\infty)E . \quad (40d)$$

With the use of (33), the equation of propagation for the steady state becomes

$$c'\frac{d}{dz}E=-i\omega(S-iA)E , \quad (41)$$

where

$$S=\Omega_c\omega^{-1}-1+(\omega_p^2\tau^2/2)(\omega^2\tau^2+1)^{-1} \quad (42a)$$

and

$$A=\tau\omega_p^2(2\omega)^{-1}(\omega^2\tau^2+1)^{-1} . \quad (42b)$$

In the ring-cavity configuration, the boundary conditions are[8]

$$E(z=0)=T^{1/2}E_{\mathrm{in}}+RE(z=L) , \quad (43a)$$

$$T^{1/2}E(z=L)=E_{\mathrm{out}} , \quad (43b)$$

where the mirrors at $z=0$ and $z=L$ have the same transmission coefficient T. R is the reflection coefficient, $R=1-T$. E_{in} and E_{out} are the incident and the transmitted fields, respectively. Since the sample fills the cavity for $z=0$ and $z=L$, the mirrors at 1 and 2 are somewhat symbolic, referring actually to the surfaces of the semiconductor. T therefore depends on the free-carrier density, and hence on the internal field intensity. From the Drude model one finds that

$$T\cong 4\epsilon_1^{1/2}(1+\epsilon_1^{1/2})^{-2} , \quad (44a)$$

where

$$\epsilon_1\cong\epsilon_\infty(1-\omega_p^2/\omega^2) . \quad (44b)$$

We now use the mean-field approximation[8] to in-

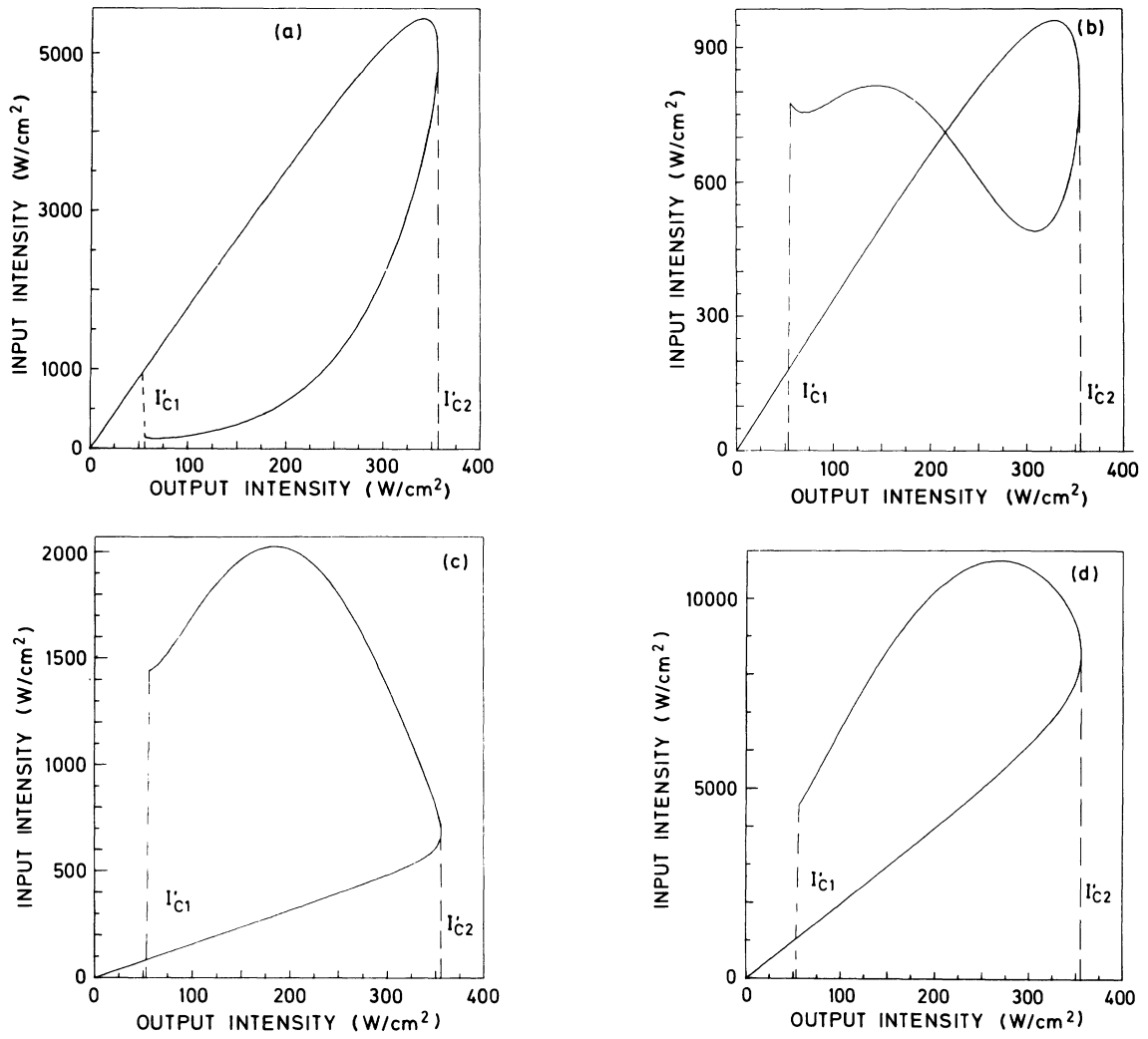

FIG. 5. Input vs output intensities for various detunings: (a) $\Omega_c/\omega=0.9$; (b) $\Omega_c/\omega=0.965$; (c) $\Omega_c/\omega=1.0$; (d) $\Omega_c/\omega=1.1$. For these curves $L=10\,\mu$m. I'_{c1} and I'_{c2} are the critical output intensities corresponding to I_{c1} and I_{c2}.

tegrate (41):

$$E(0)-E(L)\cong -i(\omega L/c')\times[S(E(L))-iA(E(L))]E(L)\ . \tag{45}$$

Clearly, the amplitude $E(L)$ corresponds to I, the average intensity inside the sample. From (45), and (43a) and (43b), we obtain

$$I_{\text{in}}=I_{\text{out}}\{(\omega L/c'T)^2S^2(I_{\text{out}}/T)+[1+(\omega L/c'T)A(I_{\text{out}}/T)]^2\}\ . \tag{46}$$

Here we have converted the solution for the amplitudes into one that involves the corresponding intensities. The use of mean-field approximation restricts this result to thin films.[9] Figures 5(a)–5(d) show the output versus input intensities. As we stated earlier, when Ω_c/ω is varied from 0.9 to 1.1, the nonlinearity goes from being mostly dispersive in nature to being mostly absorptive on $\Omega_c=\omega$, then again back to being dispersive.

ACKNOWLEDGMENTS

We thank Dr. T. W. Nee for useful discussions on the problem, and Mr. Shü-Chuan Cha for the numerical work.

[1]J. F. Figueira, C. D. Cantrell, J. P. Rink, and P. R. Forman, Appl. Phys. Lett. 28, 398 (1976).

[2]E. O. Kane, Phys. Rev. 159, 624 (1967).

[3]T. W. Nee, C. D. Cantrell, J. F. Scott, and M. O. Scully, Phys. Rev. B 17, 3936 (1978).

[4]T. W. Nee and M. O. Scully, J. Appl. Phys. 52, 102 (1981).

[5]These parameter values differ somewhat from the calculated ones. For a comparison of theoretical with experimental best-fit values, see Ref. 4.

[6]J. V. Uspenski, *Theory of Equations* (McGraw-Hill, New York, 1948).

[7]M. Sargent III, M. O. Scully, and W. E. Lamb, Jr., *Laser Physics* (Addison-Wesley, Reading, Mass., 1977), Chap. VIII.

[8]R. Bonifacio and L. A. Lugiato, Lett. Nuovo Cimento 21, 505 (1978).

[9]One also needs to have $|\delta| \ll 1$, where $\delta = m - 2\epsilon_\infty^{1/2} L \lambda_0^{-1}$, λ_0 is the wavelength in vacuum, and m is the nearest integer to $2\epsilon_\infty^{1/2} L \lambda_0^{-1}$.