ADVANCES IN PHYSICS, 1981, VOL. 30, NO. 6, 697–800

Dynamic non-linear optical processes in semiconductors

By ALAN MILLER†
North Texas State University, Texas, U.S.A.

DAVID A. B. MILLER‡
Bell Laboratories, Holmdel, New Jersey, U.S.A.

and S. DESMOND SMITH‡
Optical Sciences Center, University of Arizona, Arizona, U.S.A.

[Received 24 July 1981]

ABSTRACT

In this article *active* non-linear optical effects in semiconductors are reviewed. These processes arise because of the dynamics of excited populations of charge carriers temporarily created in otherwise empty states when a beam of laser radiation is incident on the material. A large number of different effects have been observed in different semiconductor materials recently and the non-linearities cover a wide range of magnitudes both in time-scale and size of non-linearity. The theory and physical concepts relevant to these processes are described and experimental observations using both high power pulsed and low power c.w. lasers are reviewed. Applications in optical bistability, phase conjugation, optical gating and optoelectronic gating are discussed.

CONTENTS

	PAGE
§ 1. INTRODUCTION.	698
1.1. Scope of the review.	698
1.2. Experimental techniques.	701
§ 2. THEORY	704
2.1. Absorption and refraction (linear optical properties).	704
2.2. Saturation, two-level model, density matrix approach.	712
2.3. Scattering, diffusion and recombination of carriers.	721
2.4. Free carrier non-linear properties.	724
2.5. Many-body effects.	725
§ 3. HIGH EXCITATION STUDIES.	727
3.1. Ge.	728
3.2. GaAs.	734
3.3. $InAs_xP_{1-x}$ and $Ga_{1-y}In_yAs$.	742
3.4. CdS and CdSe.	742
3.5. InSb, InAs $Cd_xHg_{1-x}Te$ and PbTe.	743
3.6. Intervalence band saturation in *p*-type semiconductors.	747
§ 4. TRANSIENT GRATINGS	752
4.1. Si.	753
4.2. Ge.	754
4.3. GaAs, InP and InAs.	755
4.4. CdS, CdSe, CdTe and ZnSe.	756

† Permanent address : Royal Signals and Radar Establishment, Malvern, England.
‡ Permanent address : Heriot-Watt University, Edinburgh, Scotland.

0001-8732/81/3006 0697 $02·00 © 1981 Taylor & Francis Ltd

§ 5. Weak excitation studies. 758
5.1. InSb. 759
5.2. GaAs. 769

§ 6. Applications. 771
6.1. Optical gating. 771
6.2. Mode-locking. 775
6.3. Optoelectronic gating. 777
6.4. Phase conjugation. 780
6.5. Optical bistability and the optical transistor. 786

§ 7. Conclusions. 793

Acknowledgments. 793

References. 794

§ 1. Introduction

1.1. *Scope of the review*

It is well known that the electrical properties of semiconductors are greatly influenced by the introduction of charge carriers into otherwise empty bands of energy levels. These carriers are usually thermally excited, either across an energy gap or from impurity states, giving the peculiar sensitivity of electrical properties as well as the microengineering opportunities to construct devices now common in the practice of microelectronics.

Semiconducting materials also have interesting and usable optical properties. Until fairly recently most attention has been focused on the classic ' passive ' effects. Non-linear optical effects leading to frequency mixing, harmonic generation, parametric amplification and the like are relatively large in semiconductors and, in general, larger for smaller energy gaps. In this review we will address a new subject which has arisen in the last few years and concerns ' active ' non-linear processes. This subject arises because of the possibility of exciting the material under the influence of a beam of laser radiation in such a way that real populations of charge carriers are temporarily created in otherwise empty states. It transpires that a relatively small number of such optically excited charge carriers can have profound effects on some non-linear optical properties on timescales ranging between microseconds and picoseconds. An early discussion of these concepts is given by Smith and Miller (1979).

A study of the dynamics of these excited populations shows that at least three characteristic times are intimately involved with the processes. The longest of these is the *carrier lifetime* or *recombination time*. This defines the time for which the excited carriers remain in the otherwise vacant band. It can be as long as hundreds of microseconds under some conditions and shorter than nanoseconds under others. The second longest timescale is imposed by *intraband processes* which tend to *thermalize* an initially non-thermal distribution of optically excited carriers. In magnitude it tends to lie between a few and a few hundred picoseconds depending upon the detailed conditions. The shortest time is known as the *dephasing* or *dipole lifetime*. This time describes how long a dipole transition between two quantum states of the system can maintain its phase coherence. By the uncertainty principle it also describes the effective width of the states. With this explanation of timescales it is clear that the description of the subject as ' *dynamic non-linear optics* ' is justified

whether the observations are made slowly, for example with c.w. lasers, or rapidly with the aid of short-pulse techniques. At the time of writing a fair number of observations of different effects in different materials can be recorded. Many of these investigations, using the new experimental facilities provided by various lasers, began independently and for a variety of different reasons. It can now be seen that there are many underlying common concepts which we attempt to unify in this review.

In optical terms, the extent of the excitation depends upon the balance between the rate of absorption of energy and the appropriate relaxation mechanisms so that some energy states become temporarily *saturated* to a greater or lesser extent. Surprisingly, a large degree of saturation can be obtained in some cases for intensities as low as 1 W/cm^2. At the other extreme, intensities as high as 1 GW/cm^2 can be incident on semiconducting materials before laser damage sets in. In practice, then, we are concerned with induced carrier densities which may range between 10^{14} and 10^{22} carriers per cm^3.

A major division can be made between absorptive and refractive effects ; however, both can be described in terms of complex electric susceptibilities. In our case it is useful to consider the expansion of the component of polarization P_i of the medium in powers of the electric field according to the equation

$$P_i = \chi_{ij}{}^{(1)} E_j{}^{\omega_1} + \chi_{ijk}{}^{(2)} E_j{}^{\omega_1} E_k{}^{\omega_1} + \chi_{ijkl}{}^{(3)} E_j{}^{\omega_1} E_k{}^{\omega_2} E_l{}^{\omega_3}. \quad (1.1)$$

For reasons that will be explained, we shall be mainly concerned with third-order effects, but it is useful to list effects of both active and passive nature in terms of the susceptibilities and the *frequencies* contained therein (Wherrett 1977). This is shown in tables 1 and 2.

We note from tables 1 and 2 that second-order effects are always passive and therefore not of interest in this review. With three terms making up the algebraic sum of frequencies in third-order effects the occurrence of ω_1 and $-\omega_1$ together allows excitation in the sense that we have discussed above.

In the field of atomic laser spectroscopy the ' system ' is sufficiently dilute that individual energy states remain discrete and the phenomenon of saturation is well known and described by the ' two-level atom ' (see, for example, Sargent *et al.* 1974, Yariv 1975). Although strong interactions take place between individual states in the case of solids many of the simple principles concerning the phenomena of saturation, population inversion and power broadening, can be transferred to the more dense situation. Coherent phenomena are difficult to observe due to the very short (picosecond) scattering time out of individual states. Nevertheless ' state filling ' effects *can* be observed and saturation takes place readily through ' *band filling* '.

The macroscopic description of eqn. (1.1) and tables 1 and 2 enables us to make comparisons concerning the order of magnitude of non-linear effects. The range is very large. For example, in second-order effects $\chi^{(2)}$ varies between 10^{-9} and 10^{-15} e.s.u. depending upon the energy gap of the material. In SI units this quantity is 10^{-12}–10^{-18} m/V, the conversion factor being $\chi^{(2)}$ (e.s.u.) $= 2{\cdot}387 \times 10^3$ $\chi^{(2)}$ (SI). In the third-order case, of interest in this article, $\chi^{(3)}$ defined in terms of the fourth rank tensor by the expression

$$P_i{}^{(3)} = \chi_{ijkl}{}^{(3)} E_j{}^{\omega_1} E_k{}^{\omega_2} E_l{}^{\omega_3} \quad (1.2)$$

varies between 10^{-11} e.s.u. for GaAs, away from resonance to 1·0 e.s.u. for

Table 1. Passive processes : real parts of $\chi^{(1)}$, $\chi^{(2)}$ and $\chi^{(3)}$.

	Frequencies of incident fields	Frequencies of fields generated by the polarization of the medium	Susceptibility	Process
1st order	ω_1	ω_1	$\chi^{(1)}(\omega_1 : \omega_1)$	Linear dispersion
2nd order	ω_1, ω_2	$\omega_3[\omega_3 = \omega_1 + \omega_2]$	$\chi^{(2)}(\omega_3 : \omega_1, \omega_2)$	Sum mixing
	ω_1	$\omega_3[\omega_3 = 2\omega_1]$	$\chi^{(2)}(\omega_3 : \omega_1, \omega_1)$	Second harmonic generation
	$\omega_1, 0$	ω_1	$\chi^{(2)}(\omega_1 : \omega_1, 0)$	Electro-optic linear Kerr effect
	ω_1	$\omega_2, \omega_3[\omega_1 = \omega_2 + \omega_3]$	$\chi^{(2)}(\omega_2 : -\omega_3, \omega_1)$	Difference-frequency mixing
	ω_1	$\omega_2[\omega_1 = 2\omega_2]$	$\chi^{(2)}(\omega_2 : -\omega_2, \omega_1)$	Degenerate difference-frequency
3rd order	$\omega_1, \omega_2, \omega_3$	$\omega_4[\omega_4 = \omega_1 + \omega_2 + \omega_3]$	$\chi^{(3)}(\omega_4 : \omega_1, \omega_1, \omega_1)$	Third harmonic generation
	ω_1, ω_2	$\omega_3, \omega_4[\omega_1 + \omega_2 = \omega_3 + \omega_4]$	$\chi^{(3)}(\omega_3 : -\omega_4, \omega_1, \omega_2)$ $\chi^{(3)}(\omega_4 : -\omega_3, \omega_1, \omega_2)$	Four-wave difference-frequency mixing processes
	ω_1	$\omega_2, \omega_3, \omega_4[\omega_1 = \omega_2 + \omega_3 + \omega_4]$	$\chi^{(3)}(\omega_2 : -\omega_3, -\omega_4, \omega_1)$	
	ω_1	ω_1	$\chi^{(3)}(\omega_1 : \omega_1, -\omega_1, \omega_1)$	Intensity-dependent refractive index
	$\omega_1, 0$	ω_1	$\chi^{(3)}(\omega_1 : 0, 0, \omega_1)$	Quadratic Kerr effect

Table 2. Active processes: imaginary parts of $\chi^{(1)}$ and $\chi^{(3)}$.

	Susceptibility	Process
1st order	$\chi^{(1)}(\omega_1;\ \omega_1)$	Linear absorption
3rd order	$\chi^{(3)}(\omega_2;\ \omega_1, -\omega_1, \omega_2)$	Raman scattering
	$\chi^{(3)}(\omega_1;\ \omega_1, -\omega_1, \omega_1)$	Intensity dependent absorption e.g. two-photon absorption or saturable absorption and refraction

InSb at near-resonance to the energy gap at 77 K. (In the SI system these magnitudes are $\chi^{(3)} = 10^{-19}\ m^2/V^2$ and $10^{-8}\ m^2/V^2$ respectively, converted by $\chi^{(3)}$ (e.s.u.) $= 7{\cdot}162 \times 10^7\ \chi^{(3)}$ (SI).) This enormous difference from the 'passive' non-resonant situation is accounted for by band-gap resonant saturation effects which can be excited by relatively weak laser fields in comparison with the usual statement that the electric fields must be comparable to interatomic fields to observe (passive) non-linearity. One should mention that although *refractive* effects are described by the real part of $\chi^{(3)}$, they become 'active' near resonance and so are included in table 2.

Finally, we should comment that the effects can be so large as to invalidate the sense of an expansion in terms of powers of the electric fields. In terms of the two-level model, however, an expression for χ is readily obtained containing in principle *all* orders of the electric field. The various orders of susceptibility in the expansion remain useful to compare orders of magnitude in various materials.

The organization of this review is to some extent arbitrary. Following brief mention of experimental techniques in § 1.2 we describe the theory and underlying concepts in § 2 and then divide description of the experiments in terms of the use of pulsed lasers and relatively high intensities (§ 3) from those of relatively weak excitation (§ 5). The presence of the laser beam can induce a transient grating in a material and the method is important enough to merit a section of its own (§ 4). Finally, we review a variety of applications.

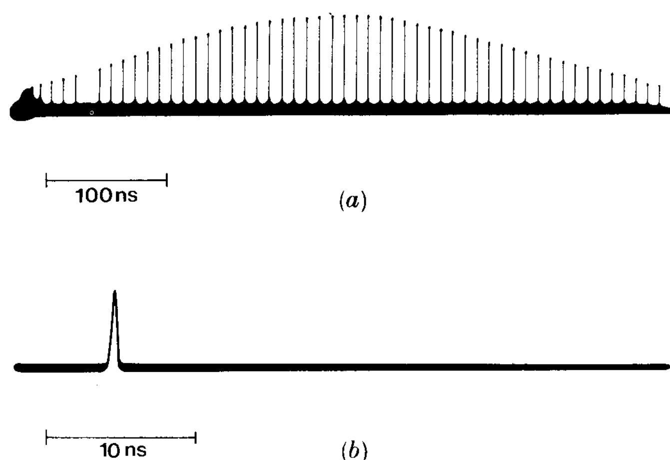
Some topics are excluded: see the work on biexcitons covered by Chemla (1980); we also omit photoconductivity and luminescence since these are adequately covered elsewhere.

1.2. *Experimental techniques*

1.2.1. *The laser sources*

The generation and analysis of mode-locked picosecond pulses is now standard practice using both dye lasers and Nd : Yag or glass (see, for example, Kaiser and Laubereau 1977). Extension of the wavelength range using parametric oscillators is also possible. Peak power values of 100 MW are readily obtained and multiple amplifier configurations can provide powers as high as 10^{12} W. The mode-locked train of picosecond pulses, fig. 1.2.1, occurring every cavity round trip time, provides the basic short-time resolution for time resolved experiments and usually one pulse is switched out by electro-optic methods. Most solid-state experiments are conducted using excite-probe

Fig. 1.2.1



Oscilloscope display of mode locked pulse train (*a*) and of the selected pulse (*b*) rise time of the detection system is not sufficient to resolve the picosecond pulses. (After Kaiser and Laubereau 1977.)

techniques. An optical delay line whose length can be adjusted then provides the picosecond timescale (the pulse travels 30 cm in 1 ns). Typical properties of a 1 ps pulse switched from a pulse train are given in table 3 and fig. 1.2.2 shows schematically a typical layout with excite and probe beams.

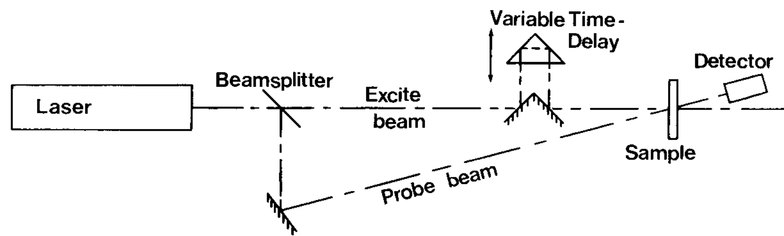
More recently, synchronously mode-locked operation of such systems as argon or krypton-ion pumped dye lasers allows the use of a *regular* sequence of picosecond pulses (up to 10^8 pulses per second) with consequent signal averaging possibilities. Peak powers ~ 10 kW are obtainable with typical pulse durations of a few picoseconds. Optical pulses as short as a tenth of a picosecond have been generated by passive mode-locking of dye lasers. These sub-picosecond pulses have been amplified to GW powers for non-linear measurements (Ippen and Shank 1978).

For infrared wavelengths, notably 10·6 μm CO_2 lasers, single mode ' hybrid-TEA CO_2 ' or injection locked CO_2 lasers giving good reproducible single mode outputs are necessary for high-quality pulsed measurements.

Table 3.

Pulse duration	$t_p = 6$ ps
Frequency width	$\Delta\tilde{\nu} = 2{\cdot}9$ cm^{-1}
$t_p \times \Delta\nu$	0·6
Peak to background ratio	$\sim 10^4$
Pulse intensity	5×10^8 W/cm^{-1}
Pulse energy after amplification	4×10^{-3} J

Fig. 1.2.2



Typical excite-probe layout.

1.2.2. *Measurements*

Absorption measurements are made fairly readily but individual pulse amplitudes are by no means constant. It is therefore quite common to computerize the analysis and make averages over large numbers of pulses.

Detectors are usually working on a nanosecond timescale compatible with reasonable electronics : on shorter timescales time delay is converted into distance in the optical delay path as in fig. 1.2.2.

Non-linear refractive effects can be monitored by a number of different methods :

(1) Reflectivity change—this can be quite large and therefore measurable directly.
(2) Transient ellipsometry.
(3) Transient gratings and four wave mixing.
(4) Beam profile distortion (self-focusing or de-focusing).
(5) Non-linear Fabry–Perot action.
(6) Non-linear magneto-optic effects such as Faraday rotation.

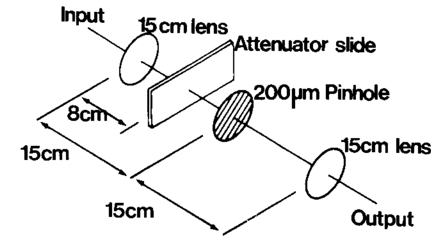
Often only the first order coefficient, n_2, of non-linear refractive effects is measured and this can be defined through

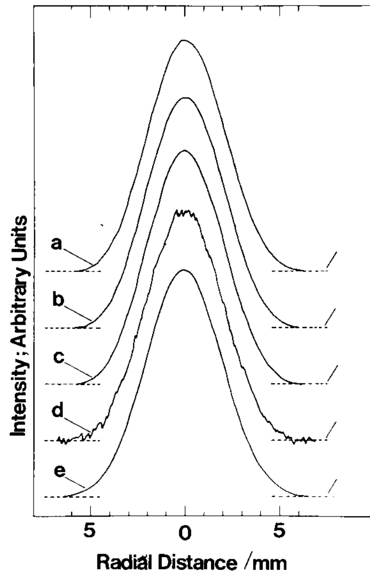
$$n = n_0 + n_2 I \qquad (1.2.1)$$

where I, the intensity, is proportional to the mean square of the electric field. n_2 is measured in practice in units of cm^2/W although calculations (e.g. eqns. (2.4.6) and (5.1.1)) are normally in cgs units ($cm^2 s/erg$). A recent review of general results for n_2 in many materials is given by Chang (1981), although many of the developments reported in the present review are subsequent to Chang's review.

Most experiments require good control of laser intensity. This is far from a trivial problem : one solution that has been effective is the use of a wedged multilayer attenuator in a spatial filter which is capable of continuously changing the intensity over four orders of magnitude whilst retaining an accurately gaussian beam profile. This is described by Miller and Smith (1978) ; fig. 1.2.3 shows the unchanged beam quality over this large range of intensity.

Fig. 1.2.3





Attenuator with beam intensity profiles taken at the position of the output lens (*a*) at 650 mW output power, (*b*) 65 mW, (*c*) 6·5 mW, (*d*) 650 μW (the resolution here is limited by detector noise). Curve (e) is a mathematical gaussian of the same width. (The heights of all traces have been adjusted for comparison.) (After Miller and Smith 1978.)

§ 2. Theory

Many physical processes are involved in dynamic non-linear optics in semiconductors, and to give a full theoretical treatment of all of these would be too extensive for this article, particularly as many of these processes are individually reported elsewhere. However, since many processes are common to the diverse investigations reviewed in this article, it is appropriate to summarize the relevant mechanisms and also to remove some of the barriers of terminology that can exist in relating work from different areas of study. Consequently in this section we try to describe briefly the processes which are invoked frequently in this review.

2.1. *Absorption and refraction* (*linear optical properties*)

The theory of linear absorption and refraction in semiconductors has been amply described in many texts (see for example, Greenaway and Harbeke,

1968, Moss *et al.* 1973, Bassani and Parravicini 1975, Pankove, 1975). Here we will emphasize those aspects which are relevant to non-linear optical processes so that we can establish some theoretical basis for the non-linear effects discussed in this article.

Any isolated atom has a characteristic set of electronic levels with associated wavefunctions and energy eigenvalues. The absorption spectrum of the atom thus consists of a series of lines whose frequencies are given by

$$\hbar\omega = E_j - E_i \quad (E_j > E_i), \tag{2.1.1}$$

where E_j and E_i are a pair of energy eigenvalues. We also know that the spontaneous lifetime, τ, from any excited state to a lower state sets a natural width of order $\hbar/\tau$ on the energy levels due to the uncertainty principle. Perturbation theory to first order gives the probability per unit time that a perturbation of the form $H(t) = H_p \exp(i\omega t)$ induces a transition from the initial state, i, of energy E_i to the final state, f, of energy E_f,

$$P_{if} = \frac{2\pi}{\hbar} |\langle f|H_p|i\rangle|^2 \delta(E_f - E_i - \hbar\omega), \tag{2.1.2}$$

where $|i\rangle$ and $|f\rangle$ are the wavefunctions of the initial and final states. This is often referred to as Fermi's Golden Rule.

If we imagine N similar atoms brought together to form a crystal, each non-degenerate energy level of an atom will spread into a band of N levels. If N is large, these levels can be treated as a continuum of energy states. In a pure semiconductor at low temperature the electrons will completely fill the valence bands while the allowed states above these, the conduction bands, will be empty. The optical properties of semiconductors are dominated by the form of the valence and conduction bands and the energy gap separating them. The wavefunctions and electron energies of these bands can be calculated by various approximate methods. Within the one electron and adiabatic assumptions, each electron moves in the periodic potential field of the lattice leading to the Schrödinger equation for a single particle wavefunction, neglecting spin,

$$\left(\frac{p^2}{2m_0} + V(\mathbf{r})\right) \psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{2.1.3}$$

The solution can be shown to have the form,

$$\psi(\mathbf{r}) = \exp(i\mathbf{k}\cdot\mathbf{r}) \cdot u(\mathbf{k}, \mathbf{r}), \tag{2.1.4}$$

where $\mathbf{k}$ is the electron wave-vector and $u(\mathbf{k}, \mathbf{r})$ has the periodicity of the lattice. The wave-vector, $\mathbf{k}$, is not uniquely defined by means of the wavefunction ; the energy is a periodic function of $\mathbf{k}$, allowing us to restrict the value of $\mathbf{k}$ to the first Brillouin zone in reciprocal space.

The majority of commonly met semiconductors have structures closely related to that of diamond with sp^3 hybridized tetrahedral bonding. In general, this results in three upper valence bands which have retained the p-like character of the atomic states (odd parity) and a spherically symmetric s-like lowest conduction band (even parity) for small values of $\mathbf{k}$ (as an example see germanium, fig. 3.1.1). For larger wavevectors, these s- and p-like states mix and the parities are not well defined. A solution of the Schrödinger

equation using for instance the $\mathbf{k}\cdot\mathbf{p}$ approximation (Kane 1957) including spin-orbit coupling results in the familiar E versus $\mathbf{k}$ electronic dispersion curves in which the energy bands are approximately parabolic near their energy maxima or minima and each band may be characterized by an effective mass parameter, m^*,

$$E(\mathbf{k}) = E(\mathbf{k}_0) + \frac{\hbar^2|\mathbf{k}-\mathbf{k}_0|^2}{2m^*}. \tag{2.1.5}$$

2.1.1. *Direct absorption*

Optical absorption can occur for photon energies in excess of the band-gap energy, transitions occurring between the full valence band and the empty conduction states. The first-order perturbation of the single electron hamiltonian by the radiation is given by,

$$H(t) = \frac{e}{mc}\,\mathbf{A}\cdot\mathbf{p}. \tag{2.1.6}$$

$\mathbf{A}$ is the vector potential of the radiation,

$$\mathbf{A}(\mathbf{r}, t) = A_0\boldsymbol{\xi}\cdot\exp\left[i(\boldsymbol{\varkappa}\cdot\mathbf{r}-\omega t)\right] + \text{c.c.}, \tag{2.1.7}$$

where $\boldsymbol{\varkappa}$ is the wave-vector of the radiation and $\boldsymbol{\xi}$ is the polarization vector in the direction of the electric field. (Here we use gaussian units consistent with the bulk of published papers on the electrodynamics of charged particles.) Using Fermi's Golden Rule, eqn. (2.1.2), the transition probability per unit time between a pair of valence and conduction band states will be given by

$$P_{\mathrm{vc}} = \frac{2\pi}{\hbar}\left(\frac{eA_0}{mc}\right)^2 |\langle\psi_{\mathrm{c}}|\boldsymbol{\xi}\cdot\mathbf{p}|\psi_{\mathrm{v}}\rangle|^2\delta(E_{\mathrm{c}}-E_{\mathrm{v}}-\hbar\omega). \tag{2.1.8}$$

There will be a change in the electron wave-vector after the transition due to momentum conservation, but this is normally neglected because of the small photon wave-vector. Thus, we assume vertical transitions in k-space (the electric dipole approximation). The total transition rate per unit volume, $W(\omega)$, is obtained by integrating over all possible vertical transitions in the first Brillouin zone taking account of all contributing bands,

$$W(\omega) = \frac{2\pi}{\hbar}\left(\frac{2\pi e^2 I}{ncm^2\omega^2}\right)\sum_{v,c}\int\frac{d\mathbf{k}}{(2\pi)^3}|\boldsymbol{\xi}\cdot\mathbf{M}_{\mathrm{cv}}(\mathbf{k})|^2\delta(E_{\mathrm{c}}(\mathbf{k})-E_{\mathrm{v}}(\mathbf{k})-\hbar\omega). \tag{2.1.9}$$

Here, we have replaced the vector potential with the intensity of the radiation through the relation,

$$A_0{}^2 = \frac{2\pi c}{n\omega^2}I, \tag{2.1.10}$$

where n is the refractive index and c is the velocity of light. $\mathbf{M}_{\mathrm{cv}}(\mathbf{k})$ is referred to as the momentum matrix element for a given transition, often assumed to be slowly varying in $\mathbf{k}$, so that

$$\begin{aligned}\boldsymbol{\xi}\cdot\mathbf{M}_{\mathrm{cv}} &= \boldsymbol{\xi}\cdot\langle\psi_{\mathrm{c0}}|\mathbf{p}|\psi_{\mathrm{v0}}\rangle\\ &= \boldsymbol{\xi}\cdot\int_{\substack{\text{crystal}\\ \text{volume}}}\psi_{\mathrm{c0}}{}^*(-i\hbar\nabla)\psi_{\mathrm{v0}}\,d\mathbf{r},\end{aligned} \tag{2.1.11}$$

where ψ_{c0} and ψ_{v0} are the zone centre wavefunctions. The momentum matrix element may be evaluated from $\mathbf{k}\cdot\mathbf{p}$ theory for simple band structures which relates it to the band-gap, E_g, the spin-orbit splitting of the valence bands, Δ_{s0}, and the electron effective mass, m_e^*,

$$|\mathbf{M}_{cv}|^2=\frac{3m_0^2\, E_g(E_g+\Delta_{s0})}{2(3E_g+2\Delta_{s0})}\left(\frac{1}{m_e^*}-\frac{1}{m_e}\right). \quad (2.1.12)$$

Finally, the absorption coefficient, α, is defined through the Beer's Law dependence of intensity on propagation distance, z.

$$\alpha=I^{-1}\frac{dI}{dz}=\frac{\hbar\omega}{I}W. \quad (2.1.13)$$

2.1.2 *Intervalence band absorption*

Although vertical transitions between the valence bands near $\mathbf{k}=0$ in diamond-like semiconductors are forbidden to a first approximation because these states have the same parity, transitions can occur at finite values of $\mathbf{k}$ due to the mixed parity of these states away from the zone centre. Doping a semiconductor p-type can result in a heavy hole valence band depleted of electrons up to some $\mathbf{k}$-value while the light hole band remains filled. This results in absorption in the infrared; in this case the momentum matrix element is directly proportional to the electron wave-vector, $\mathbf{k}$. Intervalence band absorption may also be observed when large hole densities are created by optical excitation.

2.1.3. *Two-photon absorption*

Two-photon absorption (TPA) has been employed as a means of generating large densities of carrier in semiconductors for dynamic non-linear optical studies (see, for example, Fossum *et al.* 1973, Von der Linde and Lambrich 1978, Miller, Johnston *et al.* 1979, Johnston *et al.* 1980). As an excitation method, it has the advantage that a much more homogeneous carrier distribution can be created in the crystal than for single-photon absorption because of the larger penetration depth for the radiation, but it must be remembered that the two-photon process is itself non-linear.

TPA can be treated by second-order perturbation theory. Although two photons of different frequency can be employed, electron–hole pair generation requires only one laser with a photon energy larger than half the band-gap. In this case, the transition probability per unit time between a pair of valence and conduction band states is,

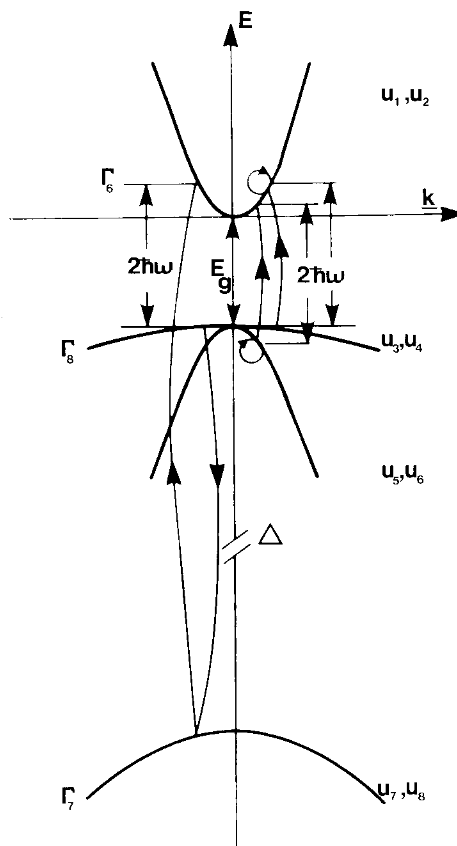
$$P_{vc}=\frac{2\pi}{\hbar}\left(\frac{eA_0}{mc}\right)^4\left|\sum_t\frac{\langle\psi_f|\boldsymbol{\xi}\cdot\mathbf{p}|\psi_t\rangle\langle\psi_t|\boldsymbol{\xi}\cdot\mathbf{p}|\psi_i\rangle}{E_t-E_i-\hbar\omega}\right|^2\delta(E_f-E_i-2\hbar\omega), \quad (2.1.14)$$

where the summation spans all intermediate states, t. The interaction can be regarded as a result of two successive processes. An electron first makes a virtual transition from the initial state, i, to an intermediate level, t, by absorption of one photon. Energy is not conserved at this stage so that absorption of the second photon must take place within a time set by the Uncertainty

Principle resulting in the electron reaching its final state, f, satisfying the condition, $E_f - E_i = 2\hbar\omega$. The intermediate state can be any valence or conduction band state such that momentum is conserved for each individual transition, (i.e. vertical transitions within the electric dipole approximation) but states with energies far from the lowest conduction band and upper valence bands will only have a weak contribution because of the resonant denominator.

Lee and Fan (1974) calculated TPA coefficients for several semiconductors with zinc blende structures assuming parabolic bands and including exciton effects. More recently, Pidgeon *et al.* (1979) have included the non-parabolicity of the bands to determine the frequency dependence. Figure 2.1.1 shows typical two-photon transitions in a model in which only the lowest conduction band and the three highest valence bands are included (three-band model). There are two types of process to be considered. One involves a valence to conduction band transition plus either an intraconduction or intravalence band transition. The second type involves a valence to conduction band transition plus a second interband transition. In all, there are twelve

Fig. 2.1.1



Conduction and valence bands of InSb near $k=0$. Typical two-photon absorption transitions are shown. (After Pidgeon *et al.* 1979.)

final states to be considered, each involving summation over intermediate transitions of these two types. The electron transition rate is,

$$W(\omega)=\frac{2\pi}{\hbar}\frac{4\pi^2e^4I^2}{n^2c^2m^4\omega^4}\sum_f \frac{d\mathbf{k}}{(2\pi)^3}\left|\sum_t \frac{\boldsymbol{\xi}\cdot\mathbf{M}_{\rm ft}(\mathbf{k})\boldsymbol{\xi}\cdot\mathbf{M}_{\rm ti}(\mathbf{k})}{E_{\rm t}-E_{\rm i}-\hbar\omega}\right|\delta(E_{\rm f}-E_{\rm i}-2\hbar\omega),\quad (2.1.15)$$

where the momentum matrix elements, $\mathbf{M}_{\rm ft}$ and $\mathbf{M}_{\rm ti}$ were defined in eqn. (2.1.11).

The two photon absorption coefficient, β, is defined by the relation,

$$-\frac{dI}{dz}=\alpha I+\beta I^2 \qquad (2.1.16)$$

so that

$$\beta=\frac{2\hbar\omega}{I^2}W. \qquad (2.1.17)$$

The two-photon absorption coefficient has proved very difficult to measure accurately, particularly since the generation of electron–hole pairs causes a free carrier absorption which varies during the laser pulse. Measured values in various semiconductors have ranged from 10^{-3} to 16 cm/MW.

2.1.4. *Indirect absorption*

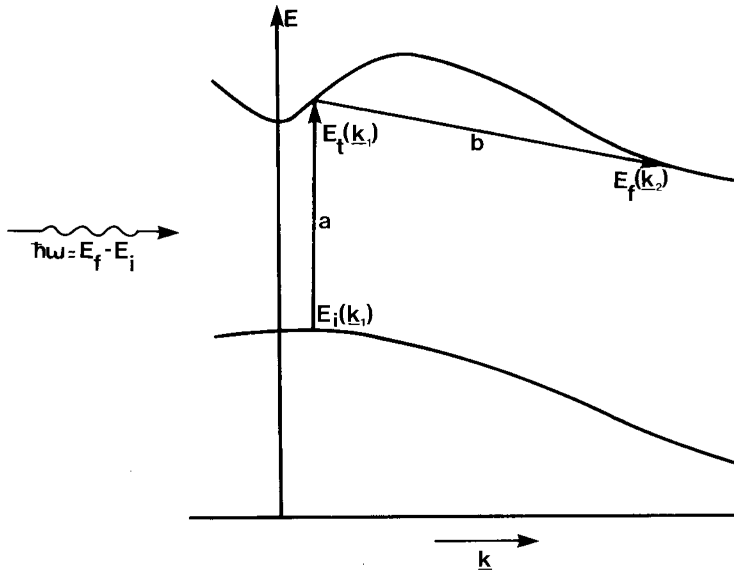
For several semiconductors, e.g. Si, Ge and GaP, the energy minimum of the conduction band and the maximum of the valence band occur at different **k**-values. In this case, optical absorption can result from above band-gap radiation through the additional interaction of a long wave-vector phonon allowing momentum conservation to be satisfied. In the process shown in fig. 2.1.2, an electron is excited by a photon in a virtual transition which violates energy conservation, (a), and a phonon scatters this electron to take up the energy mismatch, (b). The reverse ordering of phonon scattering followed by photon absorption can also take place; phonons can be either emitted or absorbed. Like two-photon absorption, this is an interaction in which two particles contribute to the electron's final energy and is therefore much weaker than direct interband absorption involving only a photon. It can therefore be treated by second-order perturbation theory similar to that described for TPA. The change in the potential experienced by the electron due to the vibration of the lattice is treated as the additional perturbation in the hamiltonian and in its most general form is,

$$H_{\rm L}=\sum_j [V(\mathbf{r}-\mathbf{R}_j-\delta\mathbf{R}_j)-V(\mathbf{r}-\mathbf{R}_j)], \qquad (2.1.18)$$

where $\mathbf{r}$ and $\mathbf{R}_j$ are the position coordinates of the electron and a given atom, j, respectively and $\delta\mathbf{R}_j$ is the displacement of that atom. This can be simplified due to the transitional symmetry of the lattice so that the contributions of the individual phonon branches may be considered separately and the concentration of phonons in each particular mode accounted for. The absorption is difficult to evaluate because all possible combinations of intermediate and final states must be summed and the electron–phonon interaction changes considerably with **k**-value.

Phonon-assisted transitions may also take place in direct gap semiconductors when the photon energy is within an optical phonon energy of the band-gap. This has been suggested as the source of the Urbach absorption tail in some low band-gap semiconductors.

Fig. 2.1.2



Indirect absorption process in a semiconductor.

2.1.5. *Exciton absorption*

The absorption processes discussed above do not take into account Coulomb attraction between the excited electron and the hole state left behind in the valence band when a quantum of radiation is absorbed in the semiconductor. This attraction can lead to the formation of an excited state in which the electron and hole are bound to each other in a hydrogen-like state (Knox 1963, Dimmock 1967, Reynolds and Collins 1981). The binding energies of excitons are typically a few meV and can be seen in the absorption spectrum of some semiconductors as a series of sharp lines just below the fundamental absorption edge. The energy of formation of an exciton is given by,

$$E = E_g + \frac{\hbar^2 |\mathbf{K}|^2}{2(m_e^* + m_h^*)} - \frac{R}{l^2}, \tag{2.1.19}$$

where $\mathbf{K}$ is the exciton wave-vector, R is the exciton Rydberg and l is a quantum integer. Absorption peaks result because transitions can only occur close to $\mathbf{k} = 0$ because after any direct transition, the electron and hole will have equal and opposite $\mathbf{k}$-values. However, since the electron and hole are bound to-

gether in an exciton, this can only be satisfied for $\mathbf{K}=0$. The series of lines converge to the absorption edge and overlap to form a continuum.

Excitons may also be found as a result of indirect transitions with emission or absorption of a phonon. The optical absorption above the band gap is often enhanced by Coulomb interactions between the main bands.

Exciton absorption structure can be difficult to observe in semiconductors. Atomic vibrations and lattice imperfections can broaden the absorption lines into the continuum. Also, excitons can be effectively screened. In particular, low band-gap semiconductors whose excitons have small binding energies and hence large orbital radii, loose their structure at even moderate doping levels. In general, therefore, low temperatures are required for the observation of exciton absorption lines.

2.1.6. *Free carrier absorption*

At frequencies below the threshold for interband transitions and exciton absorption, semiconductors have a continuous absorption spectrum due to intraband transitions (Fan 1956, 1967, Stern 1963). The absorption is approximately proportional to the number of electrons (or holes) in the band and therefore it can be varied by doping. When a free carrier absorbs a quantum of energy, $\hbar\omega$, it must change its wave-vector such that it can make a transition to an empty state at higher energy within the same band. As the photon wave-vector is small, this is only possible if the lattice takes up the momentum in a similar way to indirect absorption. Hence a phonon must be emitted or absorbed in the free carrier transition. The absorption thus depends on the phonon scattering process involved but most materials follow a λ^2 variation in absorption coefficient at long wavelength. This type of free carrier absorption can be masked in p-type semiconductors by intervalence band transitions mentioned earlier which do not require the aid of phonons.

2.1.7. *Refraction*

The macroscopic interaction of electromagnetic radiation with matter is normally expressed in terms of the frequency dependent complex dielectric constant,

$$\epsilon(\omega)=\epsilon'(\omega)+i\epsilon''(\omega), \qquad (2.1.20)$$

where ϵ' describes the refractive properties of a medium while ϵ'' describes its absorption. The latter is related to the absorption coefficient, α, by,

$$\alpha=\frac{\omega}{nc}\epsilon''. \qquad (2.1.21)$$

The real and imaginary parts of the dielectric constant are not independent but are related through dispersion relations. This is because the induced polarization of a medium is linked by causality to the electric field—there is no response before a field is applied. This is manifested in the frequency dependence of the real and imaginary parts of the dielectric constant. The Kramers–Kronig relations give a convenient method of predicting the refractive

properties of a semiconductor in the vicinity of the type of absorption processes discussed here. The real part of the dielectric constant is given by,

$$\epsilon'(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega' \epsilon''(\omega')}{\omega'^2 - \omega^2} d\omega', \tag{2.1.22}$$

where P denotes the principal value of the integral. Conversely,

$$\epsilon''(\omega) = -\frac{2\omega}{\pi} P \int_0^\infty \frac{\epsilon'(\omega')}{\omega'^2 - \omega^2} d\omega'. \tag{2.1.23}$$

We will see that the relationship between the absorptive and refractive properties of semiconductors plays an important part in the understanding of the dynamic non-linear processes discussed in this article.

2.2. *Saturation, two-level model, density matrix approach*

The normal method for calculating linear absorption, as discussed above, is through conventional time-dependent perturbation theory. This is also true for absorption mechanisms which are intrinsically non-linear, such as two-photon absorption. A tacit assumption of such perturbation methods is that the induced population in the upper state is negligible ; when the optically-induced population of the upper state (and/or the associated depletion of the lower state) is not negligible, the absorption will alter due to this finite population. This intensity-dependent process is called saturation since in the high intensity limit in the steady state the net absorption coefficient tends to saturate to a specific level ; however, the effects of saturation can often be observed long before this limit is reached. It is also worth remarking here that the populations may be induced directly from a lower to an upper state or indirectly by, for example, non-radiative scattering from some other optically-excited state ; additionally the change in absorption due to saturation is also generally accompanied by a change in the refractive properties.

In the presence of finite populations, which may be due to doping of the semiconductor, thermal excitation or optical excitation, it is straightforward to amend the perturbation calculations to include the finite populations in lower and upper states. If P is the probability of an absorbing transition per unit time from full state 'a' to empty state 'b' then, if the probability of occupation of state 'a' ('b') is f_a (f_b), the new probability of absorption is $Pf_a(1-f_b)$. With f_a and f_b finite the probability of a transition (now stimulated emission) from state 'b' to state 'a' is also now finite. By Einstein's A and B coefficient argument P is also the probability of stimulated emission transition from full state 'b' to empty state 'a' and so the probability of stimulated emission is $Pf_b(1-f_a)$. So the net probability of an absorbing transition per unit time is now $Pf_a(1-f_b) - Pf_b(1-f_a)$, i.e. $P(f_a - f_b)$. Hence the absorption probability is weighted by the factor $(f_a - f_b)$. (We have neglected degeneracy for simplicity in the above argument.)

Therefore to describe absorption saturation, where the degree of induced population which determines the absorption depends at least to some extent on that absorption, it is common to construct rate equations for the populations (f_a and f_b) and solve self-consistently for the optical absorption. To take a

trivial example if we consider a two-level system where the absorption coefficient for $f_a = 1$, $f_b = 0$ is α_0 and the population difference $(f_a - f_b)$ relaxes to its equilibrium value $(f_{a0} - f_{b0})$ with time constant τ, then the rate equations become

$$\frac{df_a}{dt} = -\frac{\alpha_0(f_a - f_b)I}{\hbar\omega} - \frac{(f_a - f_{a0})}{\tau}, \tag{2.2.1}$$

$$\frac{df_b}{dt} = +\frac{\alpha_0(f_a - f_b)I}{\hbar\omega} - \frac{(f_b - f_{b0})}{\tau}, \tag{2.2.2}$$

since $\alpha_0(f_a - f_b)I/\hbar\omega$ is the net optical transition rate from a to b where I is the intensity.

Solving for the steady state $[d(f_a - f_b)/dt = 0]$ with 'saturation intensity' $I_s = \hbar\omega/2\alpha_0\tau$ and effective absorption coefficient $\alpha = \alpha_0(f_a - f_b)$ gives

$$\alpha = \frac{\alpha'}{1 + I/I_s}, \tag{2.2.3}$$

where $\alpha' = \alpha_0(f_{a0} - f_{b0})$ is the low intensity absorption coefficient. This demonstrates saturation at its simplest with the absorption vanishing altogether at high intensity. In a semiconductor the physical situation may be much more complicated than the trivial one solved here. In general not only will there be direct optical generation rates in the equations but also scattering from and to other states; the relaxation towards equilibrium can also take much more complicated forms (for example the concept of a constant relaxation time may be invalid). Nevertheless, with these extensions rate equations can handle successfully many of the absorption saturation phenomena in semiconductors.

However, rate equation methods belie many of the subtleties of the optical interaction with matter. In particular, when populations are created optically, the quantum-mechanical phase of the resulting excited state wavefunction is at least initially intimately related to the phase of the exciting radiation. In populations created thermally, by doping or, for example, electrical injection, there will of course be no such mutual coherence. The consequences of coherence, such as the 'coherent transients' and 'Rabi oscillations' often seen in atomic systems, will not be considered further in this article; if the scattering processes in the system (e.g. carrier–phonon and carrier–carrier scattering) are strong enough that they randomize the quantum-mechanical phase inside one Rabi period, such effects will be damped out. The method often used to solve the problem more rigorously employs the density matrix, in which coherence is readily included, as also are relaxation processes. We discuss the density matrix formalism itself in the next section. The simplest system handled this way is again the two-level system, giving the optical Bloch equations. The solution for the steady state is standard (see, for example, Yariv 1975) and is actually ultimately equivalent to the simple rate equation solution derived above for the absorption, although now the 'dephasing' is included explicitly. It is useful to quote the solution here primarily to help define much of the terminology which arises from it, which has more general application throughout the subject.

For two energy levels separated by energy $\hbar\omega_0$ and driven by an optical field of frequency ω the problem is solved without perturbation approximations to give for the susceptibility $\chi = \chi' - i\chi''$ (in CGS units).

$$\chi'' = \frac{\mu^2 T_2 \Delta N_0}{\hbar} \frac{1}{1 + (\omega - \omega_0)^2 T_2{}^2 + 4\Omega^2 T_2 T_1} \tag{2.2.4}$$

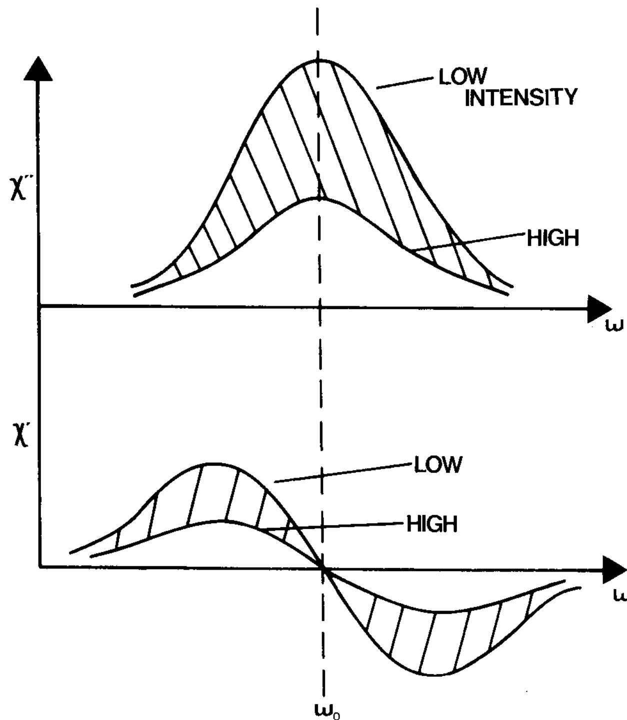
$$\chi' = \frac{\mu^2 T_2 \Delta N_0}{\hbar} \frac{(\omega_0 - \omega) T_2}{1 + (\omega - \omega_0)^2 T_2{}^2 + 4\Omega^2 T_2 T_1}, \tag{2.2.5}$$

where μ is the electric dipole matrix element and ΔN_0 is the equilibrium population difference in the absence of radiation. $\Omega = \mu E_0/\hbar$, (where E_0 is the electric field amplitude) is the Rabi frequency. T_1 and T_2 are often introduced semi-empirically ; if the radiation field were suddenly turned off, T_1 would be the time constant for the population difference to decay to its equilibrium value whereas T_2 would be the time constant for the decay of the oscillating electric dipole field of the system. T_2 can obviously never be longer than T_1, but it can be shorter since collisions may randomize the phase without returning the system to its ground state. T_1 is known variously as the energy, population or ' longitudinal ' relaxation time and T_2 as the phase, dipole or ' transverse ' relaxation time or the dephasing time. (The terms ' longitudinal ' and ' transverse ' arose from magnetic resonance where this formalism was first applied.) χ'' and χ' are of course simply related to absorption and dispersion respectively. The parameter T_2 therefore defines the width (for zero field) of the lorentzian absorption profile described by eqn. (2.2.4). (This width is also known as the ' homogeneous linewidth ', and the lorentzian profile is often called the ' homogeneous lineshape '). With increasing field, it is readily verified that the absorption line shape in eqn. (2.2.4) actually remains lorentzian in form but with increased width ; this phenomenon is called ' power broadening ' although it should be remembered that the absorption is decreasing *everywhere* in the profile with increasing power and the phenomenon might be more accurately defined as ' power-flattening ', the centre of the profile being more flattened than the sides, simply because its absorption is stronger and hence is more easily saturated. The denominator in (2.2.4) and (2.2.5) on resonance (i.e. $\omega = \omega_0$) is of the form $1 + I/I_s$ where the constant I_s is the saturation intensity. However, even off resonance, the denominator can still be arranged in the form $1 + I/I'_s$ (where I'_s is again a constant but is different from I_s). Thus an absorption which appears to saturate in the form $1/(1 + I/I'_s)$, regardless of its cause, is sometimes said to be saturating ' homogeneously '. The form of the absorption saturation described by (2.2.4) is *exactly* the same as that derived from the simple rate equation (see eqn. (2.2.3) above) with T_1 as the relaxation time τ. An absorption resonance behaving like that described by eqn. (2.2.4) is also sometimes referred to as ' Bloch-like '.

Equation (2.2.5) describes the behaviour of the refractive or dispersive part of the susceptibility, showing the classic ' anomalous dispersion ' (i.e. rising positive contribution far below resonance, diminishing negative contribution far above), with the singularity on resonance ' damped ' by T_2. With

increasing field, this too saturates, sometimes called ' saturation of the anomalous dispersion ', leading to an intensity dependence of refractive index which is negative below and positive above resonance. These saturation effects are illustrated in fig. 2.2.1.

Fig. 2.2.1



Effect of χ' and χ'' on saturating a homogeneously broadened system.

The solution of this two-level problem also illustrates another important aspect of saturation in dispersion. If the complex function χ at constant I is examined in the complex plane it can be seen that in general it has poles in the lower and upper halves of the complex plane. This invalidates the Kramers–Kronig relations (eqns. (2.1.22) and (2.1.23)), which are therefore not generally valid for saturation, and eqn. (2.2.5) cannot be derived directly from eqn. (2.2.4) using them. (If $E_0=0$ (the linear case) (i.e. $\Omega=0$), χ only has one pole on the real axis and the relations are of course valid.) One way of interpreting this breakdown physically is that for a specific intensity the induced population depends not only on the intensity of illumination but also the frequency of illumination simply because the absorption depends upon frequency. Hence in the Kramers–Kronig integral over frequency the integral is also over varying induced populations ; in other words, the integral is not really over one constant set of material conditions. But if instead we consider a specific

induced population, the real and imaginary parts of the susceptibility of the above two-level system are again rigorously related through the Kramers–Kronig relations. This can be verified using the relation for the population difference in the two-level system

$$\Delta N = \Delta N_0 \frac{1+(\omega-\omega_0)^2 T_2^{\,2}}{1+(\omega-\omega_0)^2 T_2^{\,2}+4\Omega^2 T_2 T_1} \tag{2.2.6}$$

and rearranging (2.2.4) and (2.2.5) in the form

$$\chi''(\omega) = \frac{\mu^2 T_2}{\hbar} \quad \frac{\Delta N}{1+(\omega-\omega_0)^2 T_2^{\,2}} \tag{2.2.7}$$

$$\chi'(\omega) = \frac{\mu^2 T_2}{\hbar} \quad \frac{\Delta N(\omega_0-\omega) T_2}{1+(\omega-\omega_0)^2 T_2^{\,2}}, \tag{2.2.8}$$

which also emphasizes the fact that the change in both the real and imaginary parts of this susceptibility are linearly proportional to population difference, regardless of cause. (See also, for example, Javan and Kelly 1966.)

The first extension of the theory of the homogeneous two-level system (or equivalently a set of identical two-level systems) is to a set of two-level systems of different separation energies (and possibly different matrix elements and relaxation times), which is therefore an 'inhomogeneous' set (as opposed to the 'homogeneous' set of identical systems). If there is a large number of such systems with a smooth distribution of the parameters (e.g. separation energy), it is possible to integrate the effect of the individual two-level systems to obtain the overall response, the classic case being two-level systems with a uniform distribution of separation energies. For a given radiation frequency at low intensity, those with separation energies within approximately h/T_2 of $\hbar\omega$ can interact with the radiation, but with increasing intensity more two-level systems become involved due to 'power broadening' and so this saturation (often known simply as 'inhomogeneous saturation') is slower than homogeneous, taking the form $1/(1+I/I'_s)^{1/2}$. If the absorption spectrum, in the presence of a strong intensity I at a particular frequency ω_0, is probed with a weak beam of variable frequency there will be a 'dip' in the absorption spectrum of width approximately $1/T_2$ around ω_0, this phenomenon is called 'hole burning'. This contrasts with the situation for 'homogeneous' absorbers in which the entire absorption spectrum is saturated by pumping at any wavelength. It is worth noting also that for 'inhomogeneous saturation' with uniform distribution of separation energies and no variation of any other parameter, there is no associated change in refractive index on saturation at the pumping frequency ω as the changes from systems above ω are exactly cancelled by changes from systems below ω.

2.2.1. *Density matrix formulation*

The density matrix, ρ (or 'density operator' as it is sometimes called), was originally used to describe the density of points in *phase-space* (p_x, x) so that the dynamics of sets of particles could be described by a time-variation of the form (Liouville 1838)

$$\frac{d\rho}{dt} = \left(\frac{\partial\rho}{\partial x}\,x + \frac{\partial\rho}{\partial p_x}\,p_x\right)$$

In quantum mechanics, density matrix formalism is a method of computing expectation values of operators in cases where the precise wavefunction is unknown. Suppose the wavefunction is of the form

$$\psi(\mathbf{r}, t) = \sum_n c_n(t) u_n(\mathbf{r}).$$

If A is an operator corresponding to a physical observable of the system, the expectation value is

$$\langle A \rangle = [\psi(\mathbf{r}, t), A\psi(\mathbf{r}, t)].$$

A particular matrix representation of A is obtained from

$$A_{km} = \int u_k^*(\mathbf{r}) A u_m(\mathbf{r})\, dv$$

where $u_m(\mathbf{r})$ is an arbitrary complete orthonormal set of functions. Whence

$$\langle A \rangle = \sum_{m,n} c_m^* A_{mn} c_n.$$

We assume we have enough information to calculate an *ensemble average* for $c_m^* c_n$, denoted by a bar over that quantity. Thus

$$\overline{\langle A \rangle} = \sum_{m,n} \overline{c_m^* c_n} A_{mn}.$$

We define $\rho_{mn} = \overline{c_m^* c_n}$ as the *density matrix*. We can then write

$$\overline{\langle A \rangle} = \sum_n (\rho A)_{nn} = \mathrm{tr}\,(\rho A).$$

This sum is independent of the choice of functions $u_n(\mathbf{r})$. The type of averaging can be envisaged by the sum over N systems as follows

$$\rho_{nm}(t) = \overline{c_m^*(t) c_n(t)} = \frac{1}{N} \sum_{s=1}^{N} c_m^{(s)*}(t) c_n^{(s)}(t).$$

The diagonal term ρ_{nn} is the probability of finding one of the systems in state $u_n(\mathbf{r})$: the off-diagonal $\rho_{nm}(t)$ is related to the radiating dipole of the ensemble.

The formalism thus contains both *population* and *phase* information and, importantly, can be used with *averaged relaxation times*, which characterize the decay of populations or coherence. To illustrate this we revert to a description of the two-level system.

Each wavefunction of the system satisfies (the time dependent) Schrödinger equation, whence

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} [\rho, H] = (H\rho - \rho H) \ldots, \qquad (2.2.9)$$

where the hamiltonian, H, contains the dipole interaction term μE coupling states of energy $\hbar\omega_1$ and $\hbar\omega_2$ and we can write

$$H = \begin{pmatrix} \hbar\omega_2 & -\mu E \\ -\mu E & \hbar\omega_1 \end{pmatrix}.$$

Using (2.2.9) we obtain the relations :

$$\text{(i)} \qquad \frac{d}{dt}\rho_{21} = -(\omega_2-\omega_1)\rho_{21} + \frac{i\mu E(t)}{\hbar}(\rho_{11}-\rho_{22}) - \frac{\rho_{21}}{T_2}, \qquad (2.2.10)$$

where we have introduced the term ρ_{21}/T_2 to account for the loss of phase coherence between the N eigenfunctions via ' collisions ' since

$$\rho_{21} = \overline{|\rho_2||\rho_1| \exp(i\phi)}$$

and ρ_{21} represents the dipole moment from

$$\langle\mu\rangle = tr(\rho\mu) = \mu(\rho_{12}+\rho_{21}).$$

T_2 is known as the ' dephasing-time ' or ' dipole lifetime '.

Similarly we obtain

$$\text{(ii)} \qquad \frac{d\rho_{22}}{dt} = -i\frac{\mu E(t)}{\hbar}(\rho_{21}-\rho_{21}^*) \qquad (2.2.11)$$

and

$$\text{(iii)} \qquad \frac{d}{dt}(\rho_{11}-\rho_{22}) = 2i\frac{\mu E(t)}{\hbar}(\rho_{21}-\rho_{21}^*), \qquad (2.2.12)$$

the latter following from $\rho_{11}+\rho_{22}=1$.

If N is the density of systems, $N(\rho_{11}-\rho_{22}) = \Delta N$ is the average density of *population difference* between the two levels. If ΔN relaxes to its equilibrium value with time constant T_1 (or τ sometimes) it can be included by modifying (2.2.12) as follows :

$$\frac{d}{dt}(\rho_{11}-\rho_{22}) = \frac{2i\mu E(t)}{\hbar}(\rho_{21}-\rho_{21}^*) - \frac{(\rho_{11}-\rho_{22})-(\rho_{11}-\rho_{22})_0}{T_1}. \qquad (2.2.13)$$

Thus T_1 is the population lifetime ' or recombination time '.

The quantity, $\mu E/\hbar$, is usually known as the Rabi frequency (Rabi 1937), Ω. Its reciprocal $1/\Omega$ has the physical sense of being the time required to invert the populations of the two levels and together with T_1 and T_2 controls saturation processes.

Using this formalism, susceptibilities can be calculated from the polarization P

$$P = N\langle\mu\rangle = \chi E \qquad (2.2.14)$$

The density matrix formulation was introduced into non-linear optics by Bloembergen (1965), and co-workers, with particular relevance to solids and the state of excitation we emphasize in this review. The *first-order* polarization is given by

$$\langle P^{(1)}(\omega)\rangle = \sum_g \sum_n [ex_{gn}\rho_{ng}{}^{(1)}(\omega)], \qquad (2.2.15)$$

where ex_{gn} is the dipole moment operator, whence

$$\langle P^{(1)}(\omega)\rangle = \sum_n \sum_g \frac{|ex_{gn}|^2 2\omega_{ng}}{\omega_{ng}{}^2-(\omega+i\Gamma_{ng})^2}\rho_{gg}{}^{(0)}[E_x \exp(-i\omega t)] \qquad (2.2.16)$$

and the line broadening Γ_{ng} has the physical sense of $1/T_2$.

For the next lowest order, $\rho_{nn}^{(2)}$ corresponds to saturation and with these diagonal terms assumed to be zero, the expressions in terms of the off-diagonal terms are of minor interest in the present context.

Third-order polarization, $P^{(3)}(\omega_4)$ determining the third-order susceptibility $\chi^{(3)}(\omega_4 : \omega_1, \omega_2, \omega_3)$ is written in density matrix form

$$P^{(3)}(\omega_4) = \sum_{ik} \frac{ie}{m\omega_4} [P_{ik}\rho_{ki}(\omega_4 - \omega_{ki}) + \rho_{ik}(\omega_4 - \omega_{ik})P_{ki}] \qquad (2.2.17)$$

and P_{ik} are momentum matrix elements representing dipole transitions.

There can be as many as 48 terms ; if diagonal terms are assumed zero, i.e. there are no population effects, we obtain the expressions applied to off-resonant situations, for example, by Flytzanis (1975).

Saturation is included through non-radiative population rates $1/T_1$ and for the degenerate case $\chi^{(3)}(\omega : \omega, -\omega, \omega)$, Miller *et al.* (1980) give

$$P^{(3)}(\omega) = \sum_{ik} \frac{ie}{m\omega} (P_{ik}\rho_{ki}^{(\omega)} + \rho_{ik}^{(\omega)}P_{ki}) \qquad (2.2.18)$$

$$P^{(3)} = \sum_{i} \frac{|P_{ij}|^2}{(\omega_{ji} - \omega - i\Gamma_2)} \frac{(\omega_{ji} - \omega)^2 + \Gamma_2^2}{((\omega_{ji} - \omega)^2 + \Gamma_2^2 + 4\Omega^2\Gamma_2 T_1)}. \qquad (2.2.19)$$

Large effects can then arise due to the possibilities of multiple resonance.

2.2.2. *Application to semiconductors*

Having summarized elementary saturation models, with the ‘ two-level system ’ models in particular derived initially for the atomic case, we will now examine how the various concepts are applied in semiconductors which are clearly far from being ‘ two-level systems ’. For saturation of optical absorption between bands, two limiting views are common, namely ‘ state-filling ’ and ‘ band-filling.

For the strong direct optical transitions between bands the states coupled optically are effectively those of the same **k**-value in the different bands as the photon momentum is negligible ; thus as far as the direct optical interaction is concerned the absorption looks like that from a collection of ‘ two-level ’ systems, one for each **k**-value. The energy separation varies from **k**-value to **k**-value and so the system appears similar to an ‘ inhomogeneous ’ absorber although it should be noted that the states involved are markedly different from the atomic case where the states are localized on a single atom ; in the semiconductor the states are extended throughout the whole crystal. An implicit assumption in the atomic case is that excitation is not transferred between atoms, the generation and energy relaxation taking place entirely within one atomic system : this will not generally be true for **k**-states in semiconductors where there is a large number of ways of transferring excitation between **k**-states, particularly those in the same band (e.g. carrier–carrier, carrier–impurity and carrier–phonon scattering). But if there is no transfer of excitation between the various optically-coupled states, the semiconductor will behave like the inhomogeneous atomic system ; this can be achieved if the energy relaxation is also direct (e.g. radiative recombination) or takes place entirely

through other non-saturating states not involved in the absorption. An example of such an approach is the intervalence band saturation in germanium discussed below (§ 3.5) where the energy relaxation is by fast optical phonon scattering. The phonons have sufficient energy to remove the carriers entirely from states involved in the optical interaction ; furthermore this relaxation process is not strongly ' bottle-necked ' so that the population relaxation rate is independent of intensity. The resultant saturation shows both ' hole-burning ' and absorption saturation approximately of the form $1/(1+I/I_s)^{1/2}$. This kind of saturation is known as ' state-filling ' saturation as only a small fraction of the states in a band is saturated, namely those directly excited optically. For an example of the detailed theoretical approach to state-filling saturation, the reader is referred to James and Smith (1980 a).

The other extreme case involves strong transfer of excitation between states resulting in the progressive filling of the entire band, hence its name, ' band-filling ' saturation. This occurs when the scattering within the band occurs on a much faster timescale than the recombination between bands. Thus populations created in bands are scattered rapidly usually towards quasi-equilibrium thermal distributions of excited carriers within each band, and these distributions relax more slowly through interband recombination. Thus, for example, in InSb (see § 5.1 (Lavallard *et al.* 1977)) and GaAs (reviewed by Ulbrich 1978) excited above the band-gap energy, the excited electrons and holes appear to thermalize rapidly through carrier–carrier scattering to give the same effective temperature for both the electrons and the holes (this temperature being generally different from the lattice temperature because the carriers are created with some excess energy), and the absorption saturation is then that due to the resulting thermal occupations of the states taking part in the absorption. In this limit, the problem is also relatively soluble through rate equations although now the carrier temperature has to be deduced self-consistently which may involve knowledge of the energy loss rates from the carriers to the lattice. Also spatial diffusion may become important and it is more appropriate to talk of the Boltzmann transport equation than rate equations although the two are essentially similar.

At low temperatures or high concentrations, the thermal carrier distribution will have quasi-Fermi energies inside one or more bands. When the Fermi energy is moved into the band by heavy doping, the optical absorption edge is shifted to higher energies because the transitions to states below the Fermi energy are blocked, these states being (nearly) full ; this shift is known as the Burstein–Moss shift. Consequently optical band-filling saturation, particularly at high intensities and/or low temperatures, is also known as dynamic Burstein–Moss shifting.

Refractive or dispersive effects in band-filling saturation are normally calculated by the Kramers–Kronig relations from the change in absorption spectrum due to the saturation. Of course this refraction is not only sensitive to changes in the absorption at the pumping frequency. For example, in band-filling saturation, pumping significantly above the gap might lead to band filling of states near the bottom of the bands with consequent saturation for photon energies near the gap but with relatively little saturation of the absorption at the pumping frequency, resulting in quite significant change in the refractive properties at the pumping frequency. In general, the degree of absorption

saturation at the pumping frequency is a poor guide to the strength of refractive effects at that frequency due to saturation. (We have seen above that inhomogeneous saturation can have no refractive effect at all.)

Normally when working above the band-gap, it is common to neglect any broadening of the states (such as T_2 broadening) in the linear absorption spectrum and this approximation is usually retained for band-filling saturation. For an example of such an approach see Auston *et al.* (1978). Such an approximation would however be totally inadequate for state-filling saturation: if state broadening were neglected, only those states within a radiation linewidth would be saturated whereas for laser radiation in particular the linewidth is usually very much less than the state broadening found in practice. As discussed above, it is the number of states within a state-broadening width which is saturated in this case. (Broadening may also be important in the linear (and hence also the non-linear) spectrum near the band-gap energy (see, for example, the discussions on InSb and GaAs in §§ 5.1 and 5.2.)

Interband saturation effects which do not correspond to one of these two limiting cases are in general much more difficult to solve, although fortunately these two limits cover many of the observed effects. It is worth noting that in band-filling saturation there are, however, transient effects associated with the relaxation to thermal distributions and also relaxation of the temperature of these distributions. The latter has been observed in, for example, GaAs (see § 3.2, Shank *et al.* 1979 a) although in this case the former was too fast (< 1 ps) to be resolved.

2.3. *Scattering, diffusion and recombination of carriers*

Vital to the understanding of saturation in any material is a knowledge of the time constants associated with electronic scattering, diffusion and recombination. The dynamics of optically excited carriers in semiconductors is a very large and complicated subject. Here we do no more than point out some of the interactions which influence the temporal and spatial evolution of excess carriers created by the absorption of laser radiation in a semiconductor. We refer the reader to reviews of this topic by Ulbrich (1978, 1980) and Shah (1978). Later in this article, we will see that recent experiments which monitor the time dependence of the absorptive and refractive properties have been giving a great deal of insight into this topic. To illustrate the possible mechanisms involved, let us consider electrons which have been excited to some excess energy, ΔE, above the bottom of the conduction band by one of the absorption mechanisms which we have discussed in § 2.1.

2.3.1. *Scattering*

After excitation the electrons will be scattered very rapidly to some other conduction band states. In general, scattering times in semiconductors (and hence the dephasing time, T_2) are short with the result that very high intensities are needed to observe *state-filling* saturation. We can distinguish between those scattering mechanisms which: (a) can take energy from the carrier distribution allowing relaxation of the electrons to the bottom of the conduction band: and (b) those which do not contribute to the overall carrier

cooling. One of three scattering processes may be dominant in this case—impurity, carrier or phonon scattering.

Compensation of donors and acceptors in semiconductors gives rise to ionized impurity centres. The Coulomb interaction between these centres and carriers is an elastic scattering which does not contribute to energy relaxation or redistribution of energy among the carriers. Ionized impurity scattering is the dominant momentum relaxation process which determines the mobility of carriers at temperatures below 100 K in most semiconductors. For low impurity levels and low temperatures, the scattering rate is on the order of 10^{11}–10^{12} s^{-1}.

Coulomb interactions between carriers (e–e, h–h and e–h) cause a rapid redistribution of energy between the carriers particularly at high excitation densities. Thus a thermalized distribution can result within a very short period of time after excitation and is characterized by quasi-Fermi energies for electrons and holes and a carrier temperature different from that of the lattice. The total energy of the carrier distribution remains unchanged in this process. A rate in excess of 10^{13} s^{-1} might be expected for carrier–carrier scattering events at moderate carrier densities of 10^{15}–10^{16} cm^{-3}. Elci *et al.* (1977) have estimated a rate of at least 10^{14} s^{-1} for germanium at carrier densities of 10^{20} cm^{-3}. The carrier–carrier scattering is effective in redistributing momentum amongst the carriers and hence can influence other mechanisms of momentum loss.

The excess energy of the hot-carrier distribution is dissipated mainly by interaction with vibrations of the lattice. There are two mechanisms which we need to consider that result in electron–phonon coupling. The first, optical mode scattering, is due to the electric fields generated by the motion of charged ions which can create large electric dipole moments for optical modes in ionic crystals. The coupling with longitudinal modes is particularly strong and the scattering time for carriers with large excess energies in GaAs for instance by this mechanism is approximately 10^{-12} s. The second coupling method is through the deformation potential. In this case, compression and rarefaction of the crystal due to lattice vibrations modulates the energy band structure. This coupling is weaker but is important for acoustic modes and the optic modes in non-ionic crystals such as silicon and germanium. The acoustic deformation potential scattering rate is typically 10^{8}–10^{9} s^{-1} for GaAs. For indirect gap semiconductors such as silicon and germanium and direct gap semiconductors in which electrons have been excited to energies above the minimum of a side valley, intervalley scattering by large wavevector phonons must also be considered. From measured electron–phonon coupling constants for intervalley scattering in germanium, for instance, this scattering rate is expected to be approximately 10^{14} s^{-1}. Thus an electron excited into the central conduction band minimum of an indirect gap semiconductor such as germanium will be rapidly scattered to the side valleys and subsequently relax to the bottom of these bands by intravalley phonon scattering.

2.3.2. *Diffusion*

The spatial distribution of optically generated carriers through a crystal will be inhomogeneous unless the thickness of the crystal is much less than an absorption length (reciprocal of absorption coefficient). In particular, direct

interband absorption above the fundamental edge will create carriers to the depth of only a few microns. The transverse spatial distribution is also important in certain cases such as self-focusing and transient gratings. Thus, the temporal dependence of the carrier diffusion should be taken into account. Diffusion is characterized in terms of a coefficient, D, which relates the number of carriers, n_d crossing a unit area per unit time to the concentration gradient.

$$n_d = -D\frac{\partial n}{\partial x}. \tag{2.3.1}$$

The values of the diffusion coefficients for the electrons and holes are determined by the mobilities of the carriers through the Einstein relation,

$$eD = \mu kT, \tag{2.3.2}$$

where the magnitude of the mobility, μ, will depend on the dominant scattering process discussed above. When the numbers of electrons and holes are approximately equal as in the case of laser generated electron–hole plasmas, the diffusion of both carrier types is described by an ambipolar diffusion coefficient, D_a, which is related to the coefficients for the electrons and holes, D_e and D_h respectively by,

$$D_a = \frac{2D_eD_h}{D_e + D_h}. \tag{2.3.3}$$

Since any separation of charge would create a Coulomb field between the electrons and holes, the two carriers must diffuse at the same rate.

2.3.3. *Recombination*

Excess electron–hole pairs will normally recombine on larger timescales than the average period between scattering events. The lifetime of optically generated excess carriers may be set by one of several competing recombination routes depending on such parameters as the band structure, lattice temperature, crystal perfection and carrier concentration (Blakemore 1962, Pankove 1975, Smith 1978). Because so many factors can influence the recombination rate, the relative importance of the different mechanisms must be assessed for any given set of experimental conditions. The energy of an electron–hole pair can be transferred to electromagnetic radiation, phonons or other carriers.

Electrons can recombine with holes by making a single transition between the conduction band or exciton level and the valence band. This can be accomplished by the emission of a photon with energy approximately that of the band-gap. Although light emission can often be observed, the recombination rate is usually dominated by other mechanisms. For indirect gap semiconductors, the transition would include the simultaneous emission or absorption of a phonon but this process is very weak. Recombination may result from the emission of a number of phonons instead of a photon, however, this is also very inefficient because of the large number of phonons required to take up the energy of the electron.

Low band-gap semiconductors at room temperature and materials with large generated carrier concentrations are usually dominated by Auger recombination which involves the transfer of energy between the carriers (Landsberg and Beattie 1960). A conduction band electron may recombine with a hole in

the valence band in a collision whereby a third carrier takes up the required energy. The Auger recombination time is thus very dependent on carrier density because it is a three particle interaction.

The lifetime of electron–hole pairs in many semiconductors is dominated by the existence of states near the middle of the forbidden gap due to impurities or lattice flaws. The electron de-excites in a two stage process via these deep traps. The recombination is completed when an occupied trap captures a hole, however the capture rate for holes may be quite different from that of electrons. Again, energy may be given up in the form of radiation, phonons or carrier heating, however the most efficient process is through the interaction with the lattice in most cases. A process in which the electron cascades down through the excited states of the trap emitting phonons between levels is most likely. Special account must be taken of localized flaws near the surface of a semiconductor. Electrons and holes may diffuse to the surface and recombine at these traps ; the surface recombination time can often be higher than in the bulk necessitating special surface treatment to control this.

2.4. *Free carrier non-linear properties*

Free carrier linear optical properties mentioned in § 2.1 can be summarized in the simplest case of free carriers of mass m^* by the standard expression for the complex dielectric constant ϵ, i.e.

$$\epsilon = 1 + \frac{4\pi N e^2}{m^*} \frac{1}{(-\omega^2 - i\Gamma\omega)}, \tag{2.4.1}$$

where Γ represents the line-broadening effects of scattering processes on the 'zero frequency resonance' of free carriers. Separating real and imaginary parts of ϵ to expose the refractive and absorptive effects we obtain

$$\epsilon' = n^2 - k^2 = 1 - \frac{4\pi N^2 e/m^*}{\omega^2 + \Gamma^2} \tag{2.4.2}$$

and

$$\epsilon'' = 2nk = \frac{4\pi N e^2 \Gamma / m^* \omega}{\omega^2 + \Gamma^2}. \tag{2.4.3}$$

If both electrons and holes are excited by an incident laser beam, either by interband transitions if $\hbar\omega_{\text{laser}} > E_g$ or by some other process (e.g. two-photon absorption or assistance of phonons) if $\hbar\omega_{\text{laser}} < E_g$, we refer to the 'electron–hole plasma' and a 'plasma frequency', $\omega_p{}^2 = 4\pi N e^2/m^*$. The average number of carriers excited, N, is given by

$$N = \frac{I}{\hbar\omega} \alpha_{\text{effective}} \tau_R \tag{2.4.4}$$

for intensity I, where $\alpha_{\text{effective}}$ is that part of the absorption coefficient responsible for carrier generation and the *lifetime*, or recombination time is τ_R. Substitution of (2.4.4) into (2.4.2) and (2.4.3) then gives expressions for the *non-linear* (i.e. intensity dependent) free carrier effects.

For free carrier refraction, the frequency is usually sufficiently high that $\omega^2 \gg \Gamma^2$ and thence we obtain

$$\Delta n = \frac{1}{2n_0} \cdot \frac{4\pi e^2}{m_e \omega^2} \sum_j \frac{N_j}{m_j^*/m_e} . \tag{2.4.5}$$

where m_e is the free mass of the electron, m_j^*/m_e is the appropriate mass ratio for the band extremum in question. Whence from (1.2.1),

$$n_2 \simeq \frac{\Delta n}{I} = \frac{2\pi e^2 \alpha_{\text{eff}} \tau_R}{\hbar \omega^3 m^* n_0} \tag{2.4.6}$$

for the case of one dominant carrier of mass m^*.

Absorption coefficients are similarly calculated from

$$\beta = \frac{2\omega k}{c} \tag{2.4.7}$$

by substituting for N from (2.4.4) leading to intensity-dependent absorption. Since free carrier absorption depends explicitly on the scattering processes through Γ in eqn. (2.4.3), the frequency dependence can be more complicated than the expression suggest as Γ is frequency dependent and can vary considerably in differing semiconductor bands. The scattering is related to the *carrier mobility* and is dominated in an electron hole plasma by those with the *smallest* mobility, e.g. the heavy holes. Theoretically appropriate summations can be readily formulated. The intraband scattering processes occur on a picosecond timescale and the redistribution of electrons must also be considered when studying the dynamics on such a timescale.

Free carrier non-linear effects therefore arise from the low (or zero) frequency dipole moment introduced by intensity controlled creation of carriers. A second effect—that of blocking interband transitions—also occurs resonant with the band-gap and can dominate near the band-gap. This is discussed in § 5. For a simple case the relative size of effect has been calculated by Miller, Smith and Wherrett (1980)

$$\frac{n_2 \text{ (Free carrier)}}{n_2 \text{ (Blocking)}} = \frac{4(\omega_g - \omega)}{\omega_g} \tag{2.4.8}$$

for direct gap III–V compound semiconductors. Thus at $\hbar\omega_{\text{laser}} > \frac{3}{4} E_g$ such 'blocking effects' begin to dominate.

Finally, the creation of free holes at the top of a valence band allows transitions between valence bands. Because of the form of valence band structure in semiconductors this effect can be as large as simple intraband hole absorption in some cases. The active free carrier non-linearities discussed here should not be confused with the passive non-linearity seen with a fixed number of carriers in the conduction band due to the non-parabolicity of the band (Wolf and Pearson 1966) although this non-linearity would itself be altered by optically induced populations.

2.5. *Many-body effects*

The semiconductor is clearly a many-body system, but in order to calculate its behaviour simplifying approximations have to be made of which the Hartree and Hartree–Fock self-consistent field approximations are best known : these can reduce the N-body problem to N one-body problems and the results

of such calculations are the so-called ' one-electron ' band structures commonly seen. To include the interaction between the various particles in the semiconductor to a more sophisticated extent is difficult, although it can at least be approached systematically using the Green function method (see, for example, Hedin and Lundqvist 1969 and references therein). The results of many-body effects on band structure calculations in general will not concern us further here. The aspect which is of most interest to us is the consequences of the many-body interaction of optically-excited particles on the absorption and dispersion spectra ; this will give rise to non-linear effects simply because the density of excited particles and hence the optical consequences of their interactions will depend on the exciting radiation intensity.

The physics of even these processes is not simple as it must be treated by similar methods to those used for static band structure calculations and is still the subject of much theoretical work. However, it has received considerable attention as a result of interest in electron–hole liquids and drops in semiconductors and the reader is referred to the many review articles on this subject for a more detailed discussion than is appropriate here (Hensel *et al.* 1977, Rice 1977, Muller and Zimmermann 1979, Rogachev 1980). We have also excluded most of the experimental work on electron–hole drops and liquids in this review as it is already well covered by these previous articles although the creation of electron–hole drops and liquids does cause changes in absorption and dispersion and is thus a dynamic non-linear optical process. We include, however, the work which impinges directly on other investigations reviewed here.

The principal many-body effects which are mentioned in the work covered by this review article are band-gap renormalization and exciton screening effects. All these effects can be regarded as arising from scattering between the various ' particles ', i.e. e (electron)–e, h (hole)–h, e–h, e–ex (exciton), h–ex, ex–ex. The detailed consequences of ex–ex interactions (which can take place through, for example, a van der Waals' force at low densities) with increasing exciton density are not currently clear, although many large non-linear phenomena have been associated with the bi-exciton (a two-exciton ' molecule '). The optical effects related to the bi-exciton have been reviewed elsewhere (Chemla 1980) and will not be considered further here. At high densities, excitons cease to exist, whether due to progressive ' screening ' of the electron–hole Coulombic attraction by increasing density of free carriers (leading to a Mott transition as the excitons are themselves broken into ' free ' carriers) or overlap of excitonic orbits, giving in either case an electron–hole plasma. The precise physics of this process, which leads to disappearance of excitonic absorption features, is still under investigation and will be discussed in reference to particular experiments below (§§ 3.2, 5.2).

Scattering leads not only to transitions between states but also to perturbation of the energies of the states themselves, sometimes called ' self-energy '. In an electron–hole plasma, each electron and hole sees itself in the screened Coulomb potential of the other electrons and holes ; this potential is further modified by exchange potential. Whether the electrons and holes are regarded as gaining self-energy from scattering, screening or from being ' dressed ' by a ' cloud ' of other charged carriers, the physics is the same and one effect is to reduce the energy required to create an electron–hole pair in a plasma, i.e. to reduce the effective band-gap, an effect known as band-gap renormalization.

In principle, this renormalization is different for each **k**-state, which is equivalent to renormalizing the effective mass of the (quasi) particle, although this effect is frequently neglected in discussing band-gap renormaliztion to a first approximation. (Band-gap renormalization is a vital effect in determining the existence of electron–hole drops because if the band-gap renormalization is strong enough, it will make the electron–hole plasma a lower energy state than the exciton gas for certain density ranges, leading to the local ' condensation ' of excitons into ' drops ' of electron–hole ' liquid ').

The formalism for calculating the band-gap renormalization has been summarized by Schmitt–Rink *et al.* (1980). One consequence of this Green function method is that it is also possible to calculate the state-broadening due to the carrier–carrier scattering which appears as the imaginary part of a complex self-energy. The precise manner in which such many-body effects alter the absorption and refraction spectrum is not therefore simple, but the qualitative observation of band-gap renormalization is common. In GaAs, the band-gap is reduced by about 10 meV (approximately 1%) for moderate induced carrier densities at low temperatures and this may be less for higher temperatures (see for example the calculations of Brinkman and Rice (1973) and § 3.2).

§ 3. High excitation studies

Although high-power lasers have been used to produce many non-linear effects in semiconductors (e.g. frequency mixing, parametric amplification, saturation, two-photon absorption, self-focusing and stimulated Raman scattering) we will concentrate on reviewing experiments which involve non-linearities of a dynamic nature. In particular, saturation measurements using pulsed lasers have yielded a great deal of insight into the relaxation times and lifetimes of optically excited carriers. It is not the intention here to review hot-electron effects in semiconductors ; this has been amply covered elsewhere, for instance, see Conwell (1967), Bauer (1974) and Ulbrich (1978, 1980). However, the creation of large densities of carriers using lasers results in dynamic changes in optical properties and this has been increasingly successful in recent years in monitoring the effects of carrier–carrier interactions, carrier–phonon processes, plasmons, non-thermal distributions and screening. Non-linear optical techniques are particularly powerful when the changes in optical properties can be temporally resolved using picosecond pulses with the excite-probe method or when tunable lasers can be used to determine the frequency width of the effect. Interpretation of high excitation results can become extremely complicated, however, because one can rarely separate and study a single mechanism in isolation. For example, the relaxation of highly excited electrons in the conduction band may be simultaneously influenced by exchange effects between the carriers, phonon screening and density dependent relaxation mechanisms.

Besides interest in the fundamental physics of fast electron-dynamics, the value of the experiments conducted so far may be judged by the applications in optical shutters and limiters, laser mode-locking and ultra-fast electrical gating which will be reviewed in § 6. Optical techniques also offer methods of studying electronic materials for devices when all-electronic instruments such as amplifiers and oscilloscopes are response time limited. It should be remembered

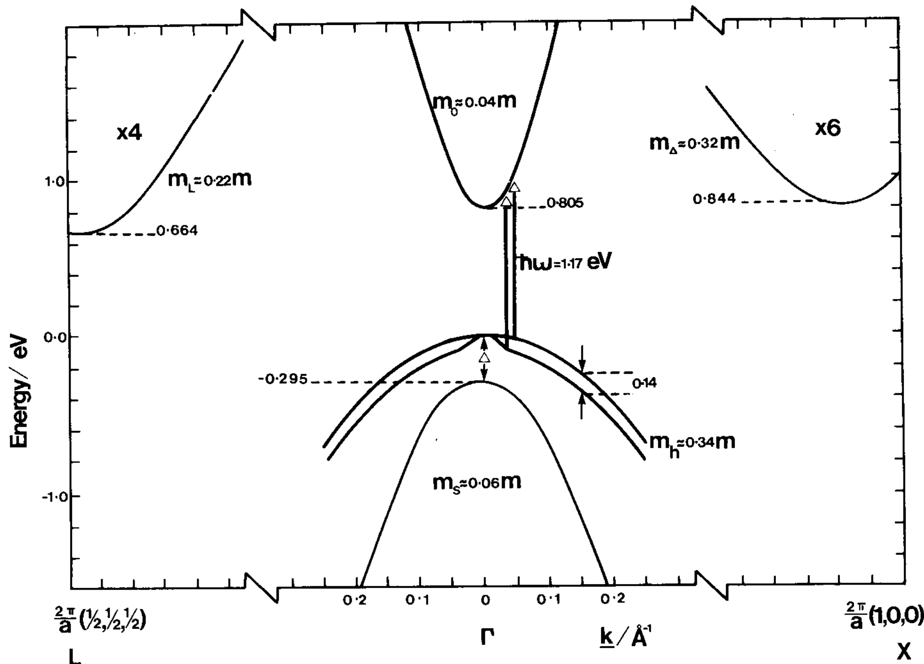
too that laser annealing and optical damage occurs in semiconductors under conditions when a highly excited electron–hole plasma is present in the material. There remains plenty of scope for new and novel experiments to clarify our understanding of the intricacies of how the optical properties of a semiconductor are altered during the time between initial absorption of energy from the laser beam until that energy is finally distributed in the lattice.

3.1. *Ge*

Laser induced saturation of interband transitions in germanium has been studied extensively since the initial measurements on this material by Kennedy *et al.* (1974). For a recent detailed discussion of picosecond photoexcitation in germanium see the reviews by Smirl (1980). The popularity of germanium is partly because it is one of the best characterized semiconductors but mainly it is due to the availability of high-power pulses of a few picoseconds duration at 1·06 μm (1·17 eV) from the mode-locked neodymium glass and YAG lasers.

The principal features of the germanium energy band structure are shown in fig. 3.1.1. It is an indirect gap semiconductor with the minimum conduction band energy consisting of four equivalent valleys at the edge of the Brillouin zone in the [1, 1, 1] direction at L, 0·664 eV (0·73 eV at 77 K) above the valence band extremum while the zone centre gap is 0·15 eV larger. Another six equivalent minima, located near X are at higher energy again, 0·18 eV above the L minima. Thus, using 1·17 eV photons, electrons can be excited to conduction band energies above all of these minima with considerable

Fig. 3.1.1

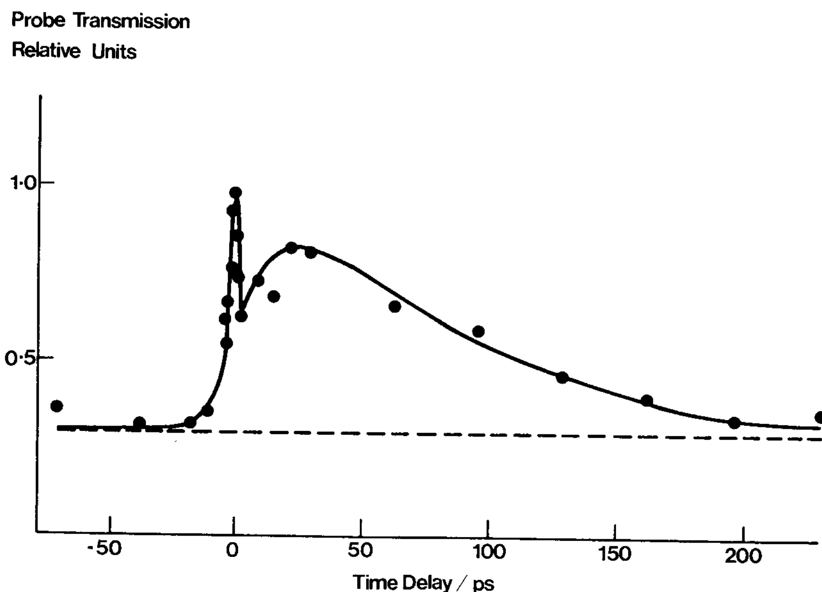


Energy band structure in germanium.

excess energies above the bottom of the conduction band. The absorption coefficient is high ($> 10^4$ cm^{-1}) at 1·06 μm due to direct transitions near the zone centre so that high intensity pulses ($I \sim 10^9$ W/cm^2) can generate very large densities of electron–hole pairs ($\Delta N \sim 10^{20}$ cm^{-3}). At these high excitation levels, electron relaxation is a complicated time dependent combination of various carrier–carrier and carrier–phonon scattering processes both within bands and between different bands.

Saturation of the absorption has been investigated in thin samples (~ 5 μm thickness) by single beam experiments and by the excite-probe technique. Single beam transmission of 10 ps pulses show just over an order of magnitude change in transmission at room and liquid nitrogen temperatures before damage is encountered (Kennedy *et al.* 1974, Smirl *et al.* 1976). In two-beam measurements, an excitation pulse intense enough to exhibit non-linear absorption causes creation of a large density of electron–hole pairs, while a weaker probe pulse which arrives at an angle to and at some time delay after the excitation pulse is used to monitor the change in transmission with time due to the generated carriers. Figure 3.1.2. shows a room temperature result by Shank and Auston (1975) in which a sharp spike is observed at zero delay of the probe followed by an increase and subsequent fall in the transmission over a period of several hundred picoseconds. The spike was originally observed by Kennedy *et al.* (1974) and was thought to arise from the saturation of the coupled valence and conduction band states (state filling); however, this was re-interpreted by Shank and Auston as parametric scattering in an optically generated electron–

Fig. 3.1.2



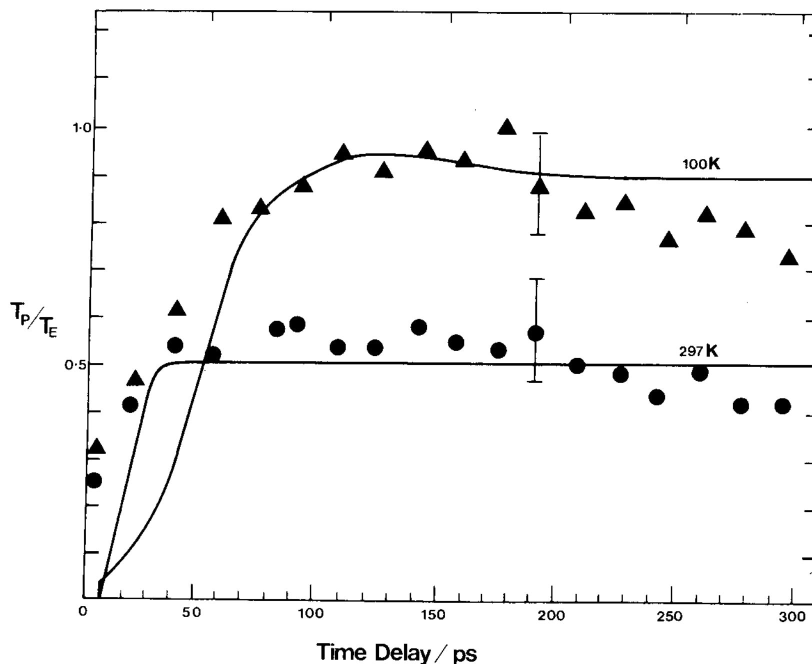
Plot of experimentally measured probe pulse transmission in Ge as a function of relative time delay between pump and probe beam. (After Shank and Auston 1975.)

hole plasma. In this process, some of the excite beam is scattered in the direction of the probe because of the phase grating formed by the interference of excite and probe which produces a spatial modulation in the electron–hole plasma and hence also in the refractive index. We will return to a discussion of this in § 4.

Saturation of the absorption certainly occurs over longer time scales. Figure 3.1.2 shows a more gradual rise and fall of the probe transmission after zero delay and has been interpreted as band-filling saturation. The rise here follows the integrated optical pulse energy. The decay could be a result of the reduction of the plasma density by diffusion of the carriers away from the thin region (approximately 1 μm thickness) near the surface where most of the absorption takes place or alternatively it could result from Auger recombination at these high carrier densities. Auston and Shank (1974) also measured the time evolution of the optically generated plasma by ellipsometry, making use of the excite-probe technique in reflection. The change in ellipticity of the probe gave a peak refractive index change of $-0{\cdot}05$. From these results, a calculation of the ambipolar diffusivity gave a value three times larger than the low density value ; this is consistent with diffusion in the degenerate regime although Jamison *et al.* (1976), in studies of the plasma reflectivity at 10 μm (see § 6.1) has concluded that the non-degenerate diffusion constant is still maintained up to carrier concentrations of 10^{19}–10^{20} cm^{-3}. Auston *et al.* (1975) used a 1·55 μm probe to measure free carrier absorption due to the electron hole plasma created by a 1·06 μm excite pulse. While diffusion redistributes the carriers, they are still available for free carrier absorption. The time decay of the probe transmission could be analysed in terms of band-to-band Auger recombination in this case.

The precise reason for the rise in transmission after the sharp spike in pulse-probe experiments had to be re-analysed in the light of the 105 K data of Smirl *et al.* (1976) shown in fig. 3.1.3. While the room temperature probe transmission follows the integrated optical pulse energy in less than 40 ps, the rise in transmission at 105 K takes place in a significantly longer time, (approximately 100 ps), than the pulse duration, 10 ps. Elci *et al.* (1977) subsequently presented a very detailed model which accounted for this slow rise in terms of the cooling of the hot electron–hole plasma generated by the excite pulse. The mechanism which was proposed to account for the time dependence of the enhanced probe transmission can be described briefly as follows. Direct optical absorption creates electron–hole pairs with the electrons initially occupying the central conduction valley near Γ, fig. 3.1.1. However, an electron in the Γ-valley can rapidly emit or absorb a phonon and make a transition to one of the valleys at X or L. The rate of scattering to the side valleys, approximately 10^{14} s^{-1}, is normally larger than the direct absorption rate. As the states in the valence band become depleted at high intensities, direct absorption becomes weaker. Carrier–carrier scattering events ensure that the electron and hole distributions will be Fermi-like. Both carrier types will have a common temperature through electron–hole scattering which may be in excess of 1500 K due to the initial excess energy of the carriers (approximately 0·5 eV). Free carrier absorption will excite generated electrons and holes still higher within their respective bands during the excite pulse thus raising the distribution temperature. The contribution of plasmon-assisted

Fig. 3.1.3



Probe pulse transmission in Ge versus delay between the excite pulse at 1·06 μm and the probe pulse at 1·06 μm for sample temperatures of 100 and 297 K. The data are plotted as the normalized ratio of probe pulse transmission to excite pulse transmission, T_P/T_F, in arbitrary units. The solid lines are theoretical curves from Elci *et al.* (1977) and the data are from Smirl *et al.* (1976).

recombination to carrier heating was also considerd to be significant in this model. The carrier temperature will be reduced by intraband optical phonon emission. Thus, the slow rise in probe transmission can be attributed primarily to this cooling of the hot electron–hole plasma. Immediately after the passage of the excite pulse, the probe transmission is low since the electrons (holes) are located high (low) in the conduction (valence) band because of the high distribution temperature. As the distribution cools and carriers fill states needed for absorption, the transmission increases. At longer time delays, the transmission falls again due to carrier recombination and/or diffusion. Plasmon-assisted interband electron–hole recombination was considered dominant by Elci *et al.* and so radiative recombination, Auger recombination and diffusion were neglected.

The solid lines in fig. 3.1.3 show the theoretical fits by Elci *et al.* (1977). Experiments by Van Driel *et al.* (1976) in which the energy band-gap was tuned by hydrostatic pressure have also been accounted for by this model but the predicted dependence of the saturated transmission on excitation pulse width has not been confirmed by experiment (Bessey *et al.* 1978).

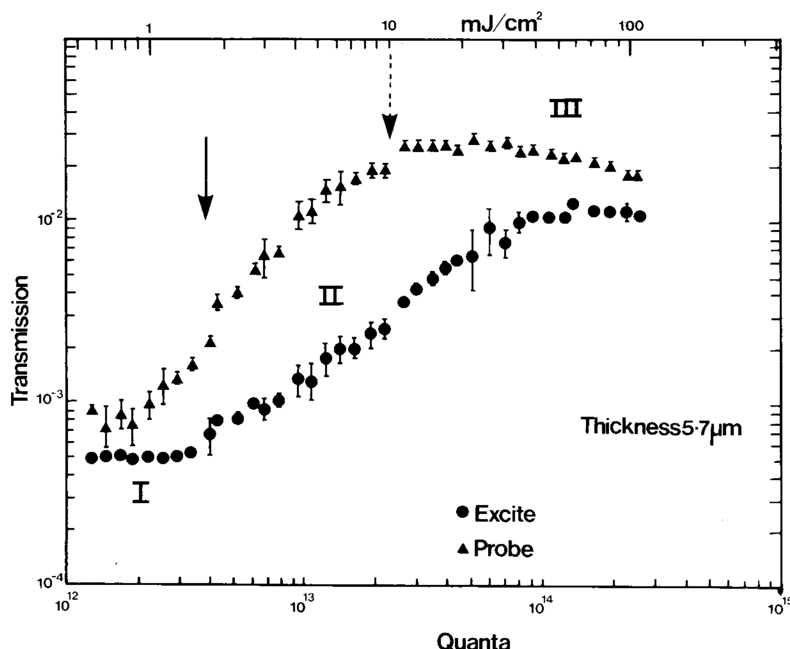
Further theoretical analysis by Latham *et al.* (1978) discussed the previous approximation of parabolic energy bands and the effects of the uncertainties

in both the electron–phonon coupling constant and the phonon assisted recombination in the model of Elci *et al.* (1977). In particular, it was pointed out that the values of electron–phonon coupling constant needed to explain the probe transmission rise are significantly lower than most measured values. In a further analysis, Elci *et al.* (1978) concluded that longitudinal diffusion can also cause significant changes in the probe transmission with time.

Some doubts thus exist for the details of the Elci *et al.* (1977) model and an alternative explanation of the excite and probe experiments was forwarded by Auston *et al.* (1978). These authors suggested enhanced intervalence band and Coulomb-assisted indirect absorption as being important mechanisms for consideration and stated that Auger recombination combined with these absorption processes at high carrier densities might account for the slow rise in the probe transmission of fig. 3.1.3. Further experiments by Smirl *et al.* (1978 a, b) clarified this. In these studies, a 1·06 μm excite pulse generated a carrier density in the non-linear regime while a 1·55 μm probe pulse (where the photon energy is too small to allow direct valence to conduction band transitions at the zone centre) was used to monitor the combined absorption due to free carriers, indirect transitions and intervalence band transitions at time delays after the excitation pulse. The result was a *decreased* transmission of the 1·55 μm radiation due to the 1·06 μm excite pulse showing that the change of transmission is opposite in sign and smaller in magnitude than changes caused by saturation of the direct absorption. Thus, the slow rise in transmission cannot be attributed solely to an additional absorption mechanism which increases at high carrier densities combined with Auger recombination.

An indication that absorption of 1·06 μm radiation due to intervalence band transitions becomes significant at high carrier densities has been given by Bosacchi *et al.* (1978 a, b) by the observation of an increase in reflectance at high pulse energies. However, this only occurred very close to damage and must therefore give rise to some doubt as to the cause (see § 6.1). More recently, Miller, Perryman and Smirl (1981 b) have observed a non-linear absorption in transmission measurements of germanium at 34 K which gives strong evidence for induced intervalence band absorption. Figure 3.1.4 shows the transmission of an excite pulse as a function of intensity and also the corresponding transmission of a probe pulse at a fixed delay of 62 ps after the excite pulse. By choosing this relatively long probe delay, one can assume that the generated carrier distribution will have cooled to near the lattice temperature but on the other hand very little recombination will have taken place. The transmission of the probe pulse can be seen to limit at a maximum value of approximately 3%, well below that expected for complete saturation of the direct valence-to-conduction band absorption by band filling. Using known band parameters for germanium, Miller *et al.* (1981 b) calculated that for low lattice and carrier temperatures, it requires 7×10^{18} and 4×10^{19} cm^{-3} holes to move the quasi-Fermi level to the initial states in the heavy and light hole bands respectively for valence-to-conduction band transitions. More than 10^{20} cm^{-3} holes lowers the Fermi level in the valence band to a position where the final states are available for direct transitions between the split-off band and the heavy and light hole bands. An estimate of the intensity required to attain this density near the front surface of the sample is shown by the solid arrow in fig. 3.1.4, in good agreement with the point of inflection in the data. An

Fig. 3.1.4



Transmission of excite and probe pulses as a function of intensity in Ge at 34 K. Probe pulse with fixed delay of 62 ps. (After Miller, Perryman and Smirl 1981 b.)

estimate of the excitation level required to produce this density throughout the crystal is shown by the dotted arrow.

The significance of this extra feature in the data is that it allows a more exact modelling of the carrier density and temperature as a function of time for given excitation conditions. Initial calculations which do not depend on the absolute magnitudes of the electron phonon coupling constants have shown that a carrier cooling model is still required to explain the temporal dependence of the absorption (Miller, Perryman and Smirl 1981 a, b, and Smirl, Miller *et al.* 1981). This leaves the interesting question as to why the cooling of the carrier distribution by the emission of optical phonons is so slow, approximately 100 ps (two orders of magnitude longer than expected). Van Driel (1979) has calculated the influence of hot phonons on the carrier energy relaxation times. He points out that the anomalously long relaxation time can be explained by a phonon bottleneck due to the relatively slow relaxation of optical phonons into acoustic phonons. Screening of the electron–phonon coupling constant may also be significant at these high carrier densities as pointed out by Yoffa (1980, 1981) for the cases of silicon and gallium arsenide.

Interband absorption non-linearities have also been observed in high transparency germanium using a pulsed CO_2 laser by Yuen *et al.* (1979). A decrease in the 9·6 μm transmission resulted at intensities in excess of 12 MW/cm^2. This was interpreted in terms of the absorption by non-equilibrium carriers excited by laser-induced energetic electrons via the carrier

Coulomb interaction (neither one- nor two-photon absorption are energetically possible at this wavelength). This mechanism has been used to describe pump laser depletion in small band-gap semiconductors (see § 3.5), although, in this case, a rate equation analysis failed to give quantitative agreement. The transmission of a second, weak 10·6 μm CO_2 laser probe which was delayed with respect to the 9·6 μm beam showed that the probe beam suffers the same drop in transmission and takes about 100 μs for the transmission drop to decay to $1/e$, in good agreement with the recombination time for optically generated electron–hole pairs in germanium.

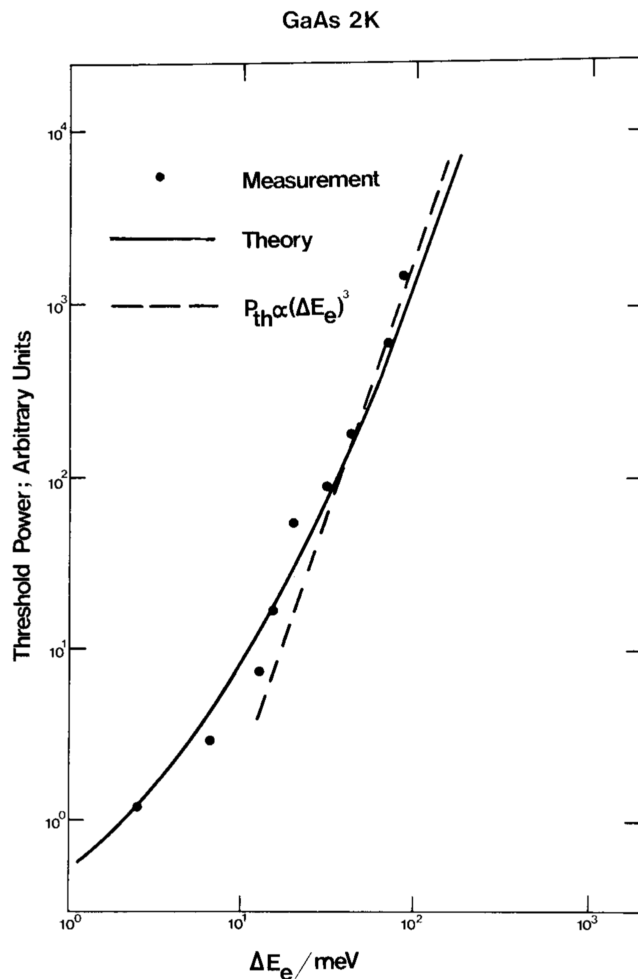
3.2. *GaAs*

Gallium arsenide is in many respects a simpler semiconductor to investigate than germanium because of its direct energy band-gap. Interband saturation studies have recently attracted considerable interest because of the strong excitonic properties near the band-edge. The gap is larger than germanium at 1·5 eV, making the 1·06 μm Nd lasers unsuitable for saturation measurements unless two-photon excitation is employed. However, this band-gap energy is in a regime accessible with dye lasers which, through recent advances, can be sychronously mode-locked with ion lasers and amplified to provide high-power sub-picosecond pulses at high repetition rates. In this section, we will review high-power saturation measurements in GaAs. Non-linear absorption and refraction measurements in this material at lower powers fall into a different category in the context of this article and will be discussed in § 5.2 although there is considerable overlap in the physics involved in both cases.

Saturated absorption was observed in GaAs as long ago as 1965. Using a GaAs injection laser at intensities in excess of 10^5 W/cm^2, Michel and Nathan (1965) observed increases in transmission by a factor of up to 14 in manganese-doped slices of GaAs. The photon energy of 1·47 eV was sufficiently large to excite electrons on the compensated acceptors into the conduction band. Similarly, Dapkus *et al.* (1970, a, b) demonstrated increased transmission at large intensities in germanium-doped GaAs at 77 K using a $GaAs_{1-x}P_x$ laser diode operating at 1·527 eV, just above the band-gap energy of the gallium arsenide sample.

The first detailed investigation of saturation in nominally pure GaAs was made by Shah *et al.* (1976) on a liquid phase epitaxially grown 3 μm thick layer sandwiched between two layers of $Al_xGa_{1-x}As$ and cooled to 2 K. A nitrogen laser pumped dye laser allowed the frequency dependence of the saturation to be followed. The observed dynamic Burstein–Moss shift due to interband optical transitions gave an order of magnitude change in transmission at photon energies just above the band-gap and at intensities exceeding 1 MW/cm^2. Figure 3.2.1 shows the threshold power at the onset of saturation as a function of the excess energy above the bottom of the conduction band. A theoretical calculation of the band filling also required inclusion of band-gap renormalization and carrier heating for a good fit to the data. Carrier densities exceeding 2×10^{16} cm^{-3} were estimated to result in a band-gap reduction due to carrier–carrier interactions. Thus, for photon energies just above the band-gap energy, more power is required for the onset of saturation than a simple density of states theory would suggest as illustrated by the comparison of the solid and dotted lines in fig. 3.2.1. The effect of carrier heating is to cause some population

Fig. 3.2.1



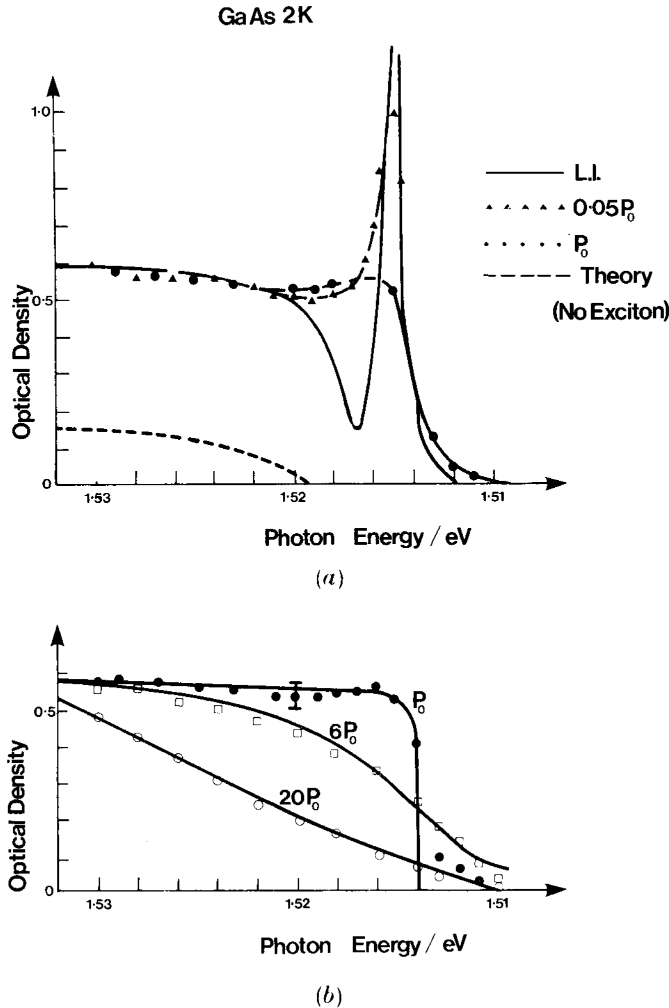
The variation of threshold power in GaAs measured at the onset of absorption saturation as a function of excess electron energy measured with respect to the unperturbed band-gap. (After Shah *et al.* 1976.)

of the states connected by the pump photons due to the tail of the Fermi distribution so that the onset of saturation is less abrupt.

In further experiments, Shah *et al.* (1977), Leheny and Shah (1978) used a 0·5 μm thick GaAs layer to study how Coulomb interactions influence the absorption spectrum under intense optical excitation close to the band-gap energy. The exciton peak in the absorption spectrum at 2 K, fig. 3.2.2 (*a*), was observed to broaden and then disappear at intensities above 500 W/cm^2 but the enhancement of optical absorption above the band-gap caused by Coulomb effects persisted to high intensities. While both exciton–exciton interactions and screening of excitons due to free carriers are possible mechanisms proposed as explanations the latter was used to interpret the data in this

case. It was suggested that free carrier screening is a dynamic response of the steady-state population of carriers to the creation of a new charge pair by photon absorption. It was argued that a time on the order of an inverse plasma frequency is required for the steady-state population to respond. Thus, if the kinetic energy of the created carriers is large, the carriers move so rapidly

Fig. 3.2.2



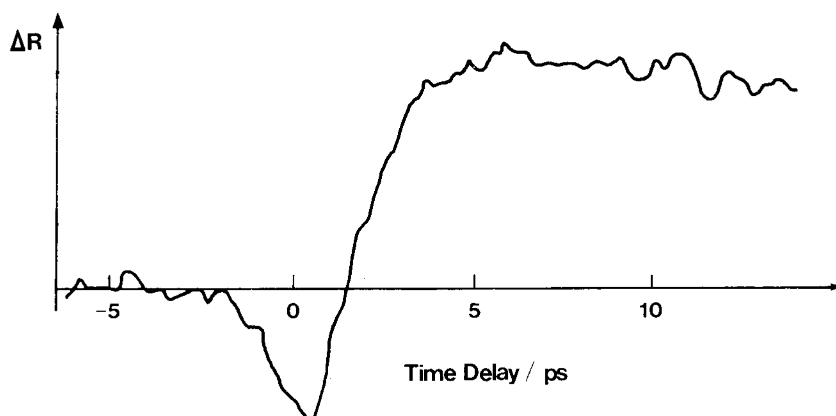
(*a*) Optical density of GaAs determined by transmission measurements. The low intensity (L.I.) spectrum was measured with a tungsten lamp, while the remaining spectra were measured with a dye laser. P_0 corresponds to incident intensity of $\approx 5\times10^2$ W/cm². The dashed curve illustrates the absorption coefficient due to parabolic bands determined from the L.I. measurement by subtracting the exciton contribution to the total absorption coefficient. (*b*) Data for high excitation intensity. The solid lines are calculated. Values of carrier density and temperature used in the calculation were, for P_0, $N=7\times10^{15}$ cm^{-3}, $T=20$ K ; for $6P_0$, $N=5\times10^{16}$ cm^{-3}, $T=40$ K : for $20P_0$, $N=1{\cdot}5\times10^{17}$ cm^{-3}, $T=60$ K. (After Shah *et al.* 1977.)

that there is insufficient time to screen their interactions and the enhancement of the absorption will persist. On the other hand, for excess carrier energies less than the plasma frequency the excitonic enhancement will be screened.

At higher intensities, fig. 3.2.2 (*b*), the absorption varies gradually with photon frequency and the absorption is seen to occur at energies lower than the exciton energy. The data could again be interpreted by a model which includes band-gap renormalization and carrier heating. Luminescence and Raman scattering measurements of the relaxation of the photo-excited carriers to the band extrema in GaAs by interaction with other carriers and by the emission of longitudinal optical phonons have been reviewed by Shah (1978) for high incident intensities and by Ulbrich (1978, 1979) for low incident intensities.

The advent of picosecond pulse dye lasers over the last few years has allowed time resolution studies of the non-equilibrium carrier and phonon dynamics in GaAs (Auston 1979). The first of these measurements was carried out in reflection (Auston *et al.* 1978, Shank *et al.* 1978) using a mode-locked c.w. dye laser with a 1 ps pulse duration. The second harmonic at 3075 A acted as a pump pulse to create a large excess density of carriers at the crystal surface while a time delayed probe beam at the laser fundamental frequency, 6150 A, monitored the change in reflectivity. At excitation energies of 10^{-4} J/cm^2 the reflectivity was observed to first decrease and then approximately 2 ps later increased and levelled off after 5 ps until recombination returned it to its initial value, fig. 3.2.3. This phenomenon was interpreted by carrying out a dispersion analysis of the perturbed optical dielectric function taking account of the non-equilibrium carriers. When the carrier distribution is very hot at short time delays, numerous high-energy condution band states will be occupied and corresponding valence band states will be vacant. Consequently the imaginary part of the dielectric constant, $\epsilon''(\omega)$, will be reduced at frequencies which are high compared to the probing frequency. From dispersion theory, this will reduce the real part of $\epsilon(\omega)$ at the probing frequency, and hence the

Fig. 3.2.3

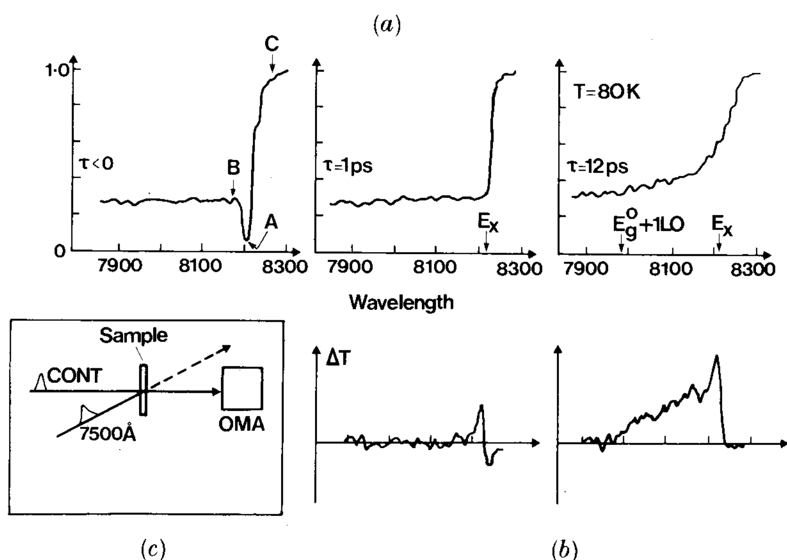


Time-resolved reflectivity from non-equilibrium carriers in GaAs following excitation at 4 eV. The probing photon energy was 2 eV. Optical pulse durations were approximately 1·0 ps. (After Shank *et al.* 1978.)

reflectivity will decrease. When the carrier distribution has relaxed to the band-edge, $\epsilon''(\omega)$ will be decreased at frequencies near the band-edge (below the probe frequency) and $\epsilon'(\omega)$ will increase at the probe frequency causing the reflectivity to increase.

Saturated transmission near the band-gap of GaAs on a sub-picosecond timescale has been used (Ippen and Shank 1978, Shank *et al.* 1979 a, b) to time revolve the effect of high densities of carriers on exciton and interband absorption. The technique used a passively mode-locked dye laser operating at 6150 A and cavity dumped at 10 Hz. The pulses were amplified with a dye amplifier pumped by the frequency doubled output of a Nd : YAG laser to give 0·5 ps pulses at powers of 200 MW. The amplifier output was split into two beams ; Raman shifting of one beam in ethanol supplied a pump at 7500 A, while a broadband continuum was generated as a probe by focusing the second beam in water. After transmission through a 1·5 μm thick layer of GaAs at 80 K, the probe was dispersed in a spectrometer and all wavelengths detected simultaneously using an optical multichannel analyser, fig 3.2.4 (*c*). The transmission results for three time delays, $\tau = 0$, 1, 12 ps, of the probe beam reproduced from Shank *et al.* (1979 a) in fig. 3.2.4 show the speed of the carrier effects on the transmission. The sharp exciton absorption at 8200 A has disappeared due to carrier screening within 1 ps while the absorption edge has shifted to lower energy. Band filling is observed by the increased transmission

Fig. 3.2.4

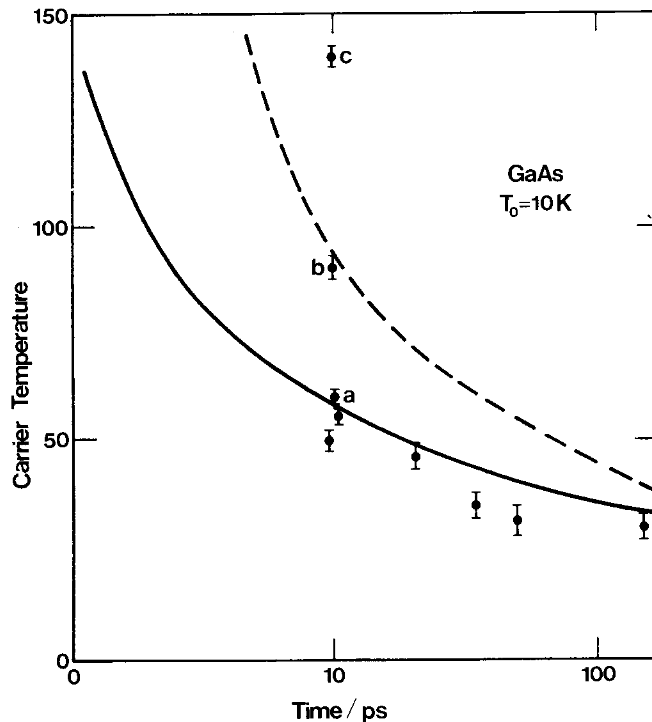


(*a*) Transmission spectra of GaAs corrected for wavelength variation of probe beam and OMA detector response. Spectra were recorded over a range of delay times between the pump and probe beams. (*b*) Difference spectra showing the difference between spectra obtained after a fixed delay and the spectra obtained for $\tau > 0$. These spectra were obtained by electronically subtracting the data stored in two memories of the multichannel analyser. (*c*) Schematic of experiment. The OMA consisted of a spectrometer and an intensified silicon vidicon with multichannel signal averager. (After Shank *et al.* 1979 a.)

at $\tau = 12$ ps. For the excitation densities generated (approximately 10^{17} cm^{-3}) and for the excess excitation energy of 120 meV, it was concluded that Coulomb interactions scatter the carriers out of their initial states in less than a picosecond since no preferential occupation of states at high energy was observed. Thus, a thermal distribution of carriers with some effective temperature greater than the lattice temperature evolved within 1 ps following excitation and then cooled to the lattice temperature in approximately 4 ps as a result of LO-phonon emission by the hot carriers.

Additional transmission measurements at 10 K by Leheny *et al.* (1979), using the same technique shows that at excitation densities above 10^{17} cm^{-3}, the carrier–lattice interaction cools the carriers less effectively and for excitation densities greater than 10^{18} cm^{-3} the carriers can no longer be described by a thermal distribution. Figure 3.2.5 shows the measured carrier temperature from saturated transmission data at fixed time delays after the excitation pulse for three carrier excitation densities. The results for $n \sim 10^{17}$ cm^{-3}, fig. 3.2.5, curve (a), agree with theory. Here the calculation assumes initial cooling is predominantly via LO-phonon scattering until about 55 K when the fraction

Fig. 3.2.5

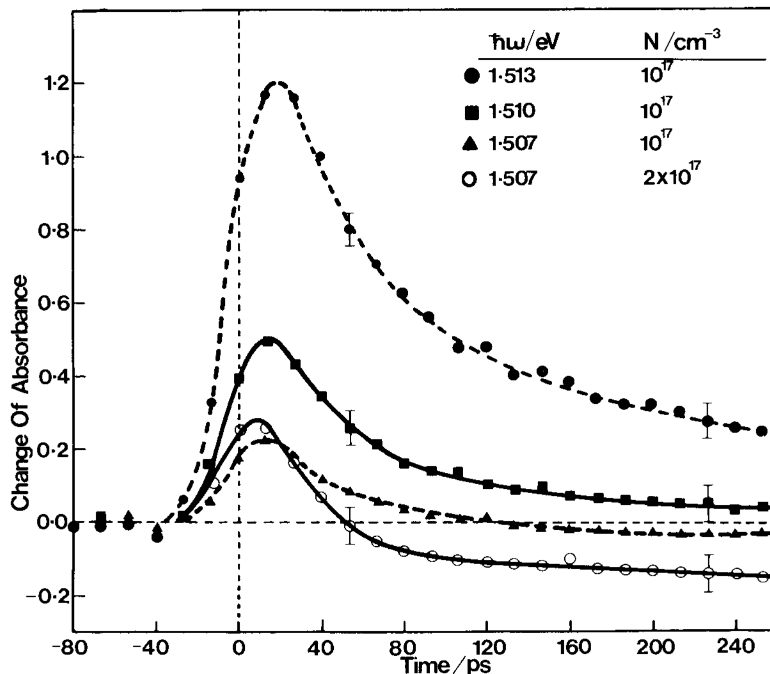


Measured hot-carrier temperatures in GaAs at various time delays following pulsed excitation. The solid curve is carrier temperature calculated assuming an initial excess energy of 120 meV. ($T_c(0) = 465$ K) and the dashed curve indicates the temperature variation expected if the carrier-LO-phonon scattering rate is decreased by a factor of 5. (After Leheny *et al.* 1979.)

of carriers with energy above the LO-phonon energy becomes small. At this point carrier–acoustic–phonon scattering dominates so that the rate of cooling drops. For higher densities, $n \sim 3 \times 10^{17}$ cm^{-3}, fig. 3.2.5, curve (b), a slowing of the carrier cooling was observed and by $n \sim 10^{18}$ cm^{-3} evidence of a non-equilibrium distribution was observed in the transmission. Leheny *et al.* pointed out that in this high excitation case, the plasmon energy is approximately equal to the LO-phonon energy with the result that strong coupling between carrier and phonon systems can screen the polar optical phonons although this mechanism alone does not explain the non-thermal distribution.

An alternative picosecond saturation technique has been employed by Von der Linde and Lambrich (1978, 1979 a, b). An excite-and-probe scheme was again used in which an intense pulse at 1·06 μm with 25 ps width from a mode-locked Nd : YAG laser created a hot electron–hole plasma via two-photon excitation in an 8 μm thick GaAs slab at 6 K. A second delayed frequency-tunable pulse from a synchronously mode-locked optical parametric oscillator measured the change in transmission. Figure 3.2.6 shows the change of absorption at three different probe wavelengths below the exciton ground

Fig. 3.2.6



Induced absorbance versus delay time in GaAs. Time zero refers to the maximum of the pump pulse ; $\hbar\omega$, photon energy of the probe pulses ; N, density of e–h pairs determined from two-photon transmission measurements. Dotted curves, guide to the eye only ; full curves, examples of the calculated temporal evolution of the absorbance. The calculated temperatures at the maxima of the curves are 160 and 120 K for 10^{17} cm^{-3} and 2×10^{17} cm^{-3} respectively. (After Von der Linde and Lambrich 1979 a.)

state as a function of delay time for generated carrier densities estimated in excess of 10^{17} cm $^{-3}$. An initial rapid increase in absorption is observed following the integrated pump pulse. After 20 ps, the induced absorption begins to decrease non-exponentially. At 1·507 eV, the induced absorption decays to zero and then becomes negative resulting in optical gain. The amount of gain is sensitive to the excited carrier density. This switch from absorption to gain may be understood in the following way. As observed by Shank *et al.* (1979 a), fig. 3.2.4, thermal equilibrium among the electrons and holes is reached very rapidly, while the plasma cools more slowly by LO phonon emission. Band-gap renormalization reduced the energy gap causing the intial increase in absorption below the exciton frequency. Later, as the distribution cools, the photoexcited carriers fill the states up to the Fermi energies, $E_f^{e}(n)$ and $E_f^{h}(n)$. Transitions between occupied states results in gain for photon energies above the reduced band-gap E'_g but below the chemical potential, μ, given by $\mu = E'_g(n) + E_f^{e}(n) + E_f^{h}(n)$. Like Leheny *et al.* (1979), Von der Linde and Lambrich determined the plasma temperature, T, as a function of time, t, from their measurements. Also, a calculation of the temporal dependence of T assuming relaxation by emission of longitudinal optical phonons led to the equation,

$$\tfrac{3}{2}k\,\frac{dT}{dt} + \left(\frac{\hbar\omega_{LO}}{\tau_0}\right)\exp\left(-\frac{\hbar\omega_{LO}}{kT}\right) = H(t), \qquad (3.2.1)$$

where $\hbar\omega_{LO}$ is the energy of the phonon, τ_0 is the time constant of the phonon relaxation and $H(t)$ is the rate of energy increase due to two-photon pumping. This equation describes the absorption results in fig. 3.2.6 very well (solid lines) when the finite pulse widths are taken into account in the calculation of the absorption change. The time constant, determined from the best fit to the data gave $\tau_0 = 0{\cdot}12$ ps, in good agreement with the theoretical value. The conclusions of this analysis disagree with the interpretation given by Leheny *et al.* (1979) for long delays after the excitation pulse. While the Leheny model accounts for the slow cooling by the small fraction of LO phonons above 55 K making the slower acoustic-phonon scattering dominant, in this case Von der Linde and Lambrich accounted for the slow cooling only by the exponential decrease in energy loss rate with temperature for the optical phonons.

The picosecond continuum spectroscopic technique developed by Ippen and Shank (1977) has recently been extended to characterize carrier transport in thin GaAs–$Al_{0\cdot2}Ga_{0\cdot8}As$ double heterostructures grown by liquid phase epitaxy (Fork *et al.* 1980, Shank *et al.* 1981). Carriers were optically injected near the band-edge and the Franz–Keldysh effect was used to monitor the evolution of the electric field as the holes and electrons drift to opposite ends of the region of an applied field. A small space-charge field induced by these carriers perturbed the optical absorption. The absorption spectrum was measured as a function of time delay after excitation and the spectra before excitation and after a given delay were subtracted and analysed to give carrier velocities. For GaAs, which has a low effective mass central valley, carriers introduced into this valley for moderately large electric fields exhibited a large initial velocity followed by rapid relaxation to the saturation velocity observed after long times. An overshoot electron velocity of $4{\cdot}4 \times 10^{7}$ cm/s

was observed on timescales of a few picoseconds compared to $1{\cdot}2 \times 10^7$ cm/s at longer delays.

Carrier dynamics in microstructure layers sufficiently thin that quantum size effects become important were also explored by Fork *et al.* (1980). With GaAs layers between 50 and 205 Å thick, induced bleaching of the absorption peaks at each of the two-dimensional ($n = 1, 2, 3$) band edges was attributed to screening of the excitons associated with those energies ; an overall bleaching at short times was attributed to band filling ; and a change in the spectrum with time was believed to be due to carrier recombination and cooling to the $n = 1$ level.

3.3. $InAs_xP_{1-x}$ *and* $Ga_{1-y}In_yAs$

Reintjes *et al.* (1975) has studied saturated absorption in two III–V alloys using a mode-locked Nd : glass laser. The alloy compositions of the $InAs_xP_{1-x}$ and $Ga_{1-y}In_yAs$ were adjusted to give an energy gap approximately 10 meV less than the laser photon energy of 1·17 eV. This resulted in the parameters $x = 0{\cdot}19$ and $y = 0{\cdot}2$. Both materials have direct band-gaps and had absorption coefficients of about 10^4 cm^{-1} at 1·17 eV. Thin epitaxial layers a few microns thick were used for single pulse and pulse-probe measurements. In single pulse experiments, the transmission increased from 1 to 15% in a 25 μm thick $InAs_xP_{1-x}$ layer and from 1 to 5% in a 2 μm $Ga_{1-y}In_yAs$ layer. The saturation of these alloys is due to a dynamic Burstein–Moss shift of the energy gap as discussed for germanium and gallium arsenide in §§ 3.1 and 3.2. The scattering of the generated free carriers from the optically coupled states will be fast compared to the laser pulse width of 20 ps. A pulse-probe measurement on $Ga_{1-y}In_yAs$ showed a rapid rise and fall of the probe transmission over a time scale of 20–40 ps, approximately the pulse duration.

3.4. *CdS and CdSe*

CdS has shown induced absorption at energies below the band-gap ($E_g =$ 2·41 eV) when excited by a high-power nitrogen laser at 337·1 nm (3·68 eV). Goto and Langer (1971) studied the change in transmission of thin samples during the 10 ns laser pulse excitation as a function of wavelength by using light from a Xe flash lamp as a probe. The decreased transmission exhibited two distinct time constants. The slower one had a decay time of the order of 10 μs and was thought to be due to lattice heating by the laser pulse. The fast component formed a continuous background, the magnitude of the absorption decreasing at longer wavelengths, and decayed within the time resolution of the equipment. The A_1 exciton absorption line was observed to broaden at high intensities. It was suggested that this non-linear absorption is consistent with the process of electron-assisted transitions where the relaxation of the hot electrons, which were produced by the absorption of the laser light, will couple to photons of below-band-gap energy with the effect of creating additional electron–hole pairs.

More recently, Bruckner *et al.* (1976) have investigated enhanced transmission of 0·526 μm picosecond radiation in $CdS_{0{\cdot}75}Se_{0{\cdot}25}$ at 80 K. The composition of the mixed crystals and the temperature were chosen to achieve photon resonance with the excitons. In a 2 mm thick crystal, 5 and 20 ps duration

pulses showed a transmission higher by a factor of 10^3 compared to continuous radiation and was accompanied by a time delay of the output pulse. These observations are consistent with the phenomenon of self-induced transparency which may occur if the pulse width, τ, is less than the exciton momentum relaxation time, T_2. In this case, an intense laser pulse can propagate with anomalously small loss and small velocity as a result of induced re-radiation of the medium. The character of the observed pulse deformation was quite different for longer pulses, $\tau \sim 1$ ns.

3.5. *InSb, InAs, $Cd_xHg_{1-x}Te$ and PbTe*

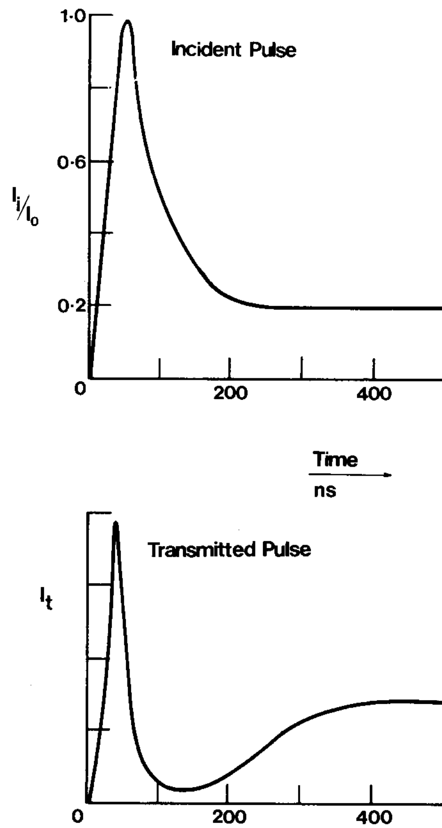
Small band-gap semiconductors can have rather unique optical properties because of their direct gaps and very low conduction band effective masses. As a result, they continue to provide unusual and often unexpected effects. Non-linear behaviour has found application in the spin-flip Raman laser (Smith *et al.* 1977) and, more recently, optical bistability and phase conjugation have been demonstrated (see §§ 6.4 and 6.5). We will review work on saturation and non-linear refraction obtained in InSb with c.w. lasers in § 5.1 but here we will concentrate on some rather novel non-linear transmission effects discovered with high-power pulsed lasers.

Nurmikko (1976 a) was the first to demonstrate a dynamic Burstein–Moss shift in InSb by using frequency doubled pulses from a mode-locked CO_2 laser at 5·3 μm in a 25 μm thick crystal slice. Experiments were carried out at 297 and 110 K. At the lower temperature, where the difference between the pump photon energy and the absorption edge is less than 15 meV, the absorption saturated at a peak intensity of 10 kW/cm^2 and gave three orders of magnitude self-induced increase in transparency.

$Cd_xHg_{1-x}Te$ is a particularly useful material for saturation studies because the band-gap can be chosen anywhere between 0 and 1·5 eV by variation of the alloy composition. Matter *et al.* (1976) carried out saturation measurements on material with an x value of 0·61 which corresponds to a room temperature band-gap of 0·75 eV. Picosecond 1·06 μm (1·17 eV) pulses of energy up to 0·1 mJ gave a change in transmission of more than a factor of 10 at 297 and 105 K in a 5 μm thick slice. A pulse-probe measurement showed that the enhanced transmission caused by the intense pump pulse decayed within 50 ps. Nurmikko (1976 b) used a mode-locked CO_2 laser to saturate the transmission of $Cd_xHg_{1-x}Te$ ($x \sim 0{\cdot}2$) at 77 K. In this case the photon energy was only 10 meV higher than the absorption edge and a 200 μm thick slice showed increased transparency of more than three orders of magnitude for intensities up to 1 MW/cm^2 because of band filling.

At low temperatures, 4 K, the temporal dependence of transmitted 10·6 μm pulses in InSb has shown strong pump depletion effects (Figueira *et al.* 1976) at intensities greater than 2 MW/cm^2. Figure 3.5.1 shows the incident and transmitted pulses for a high-power CO_2 laser. The transmitted radiation rises to a peak in 35–40 ns and then falls essentially to zero in another 35–40 ns. Nee *et al.* (1978) have proposed a model explaining this transmission cut-off in terms of the Kane process (Kane 1967). At 4 K, the band-gap is more than twice as large as the photon energy, ruling out carrier generation by single or two-photon absorption. However, the samples used have initial electron

Fig. 3.5.1



Incident pulse and transmitted pulse profiles in InSb at 10·6 μm. Experimental result of 2·2 MW/cm² incident pulse. (After Figueira *et al.* 1976.)

densities of 10^{16} cm^{-3}. These electrons will be excited to higher energy states due to absorption of photons from the incident laser beam. These electrons excited to the states with excess energy greater than the band-gap can decay back to the bottom of the conduction band by exciting an additional electron from the valence band to the conduction band. Nee *et al.* estimated that the pair-excitation rate for InSb exceeds both the conduction electron relaxation rate and the recombination rate for carriers of sufficient excess energy. A simplified rate equation analysis predicts transmitted pulse shapes of the type observed.

Jamison and Nurmikko (1978a) and Nurmikko and Jamison (1979) have subsequently observed a high intensity transmission limit in several small band-gap semiconductors using 10·6 μm pulses of nanosecond duration. *n*-type InAs, InSb and $Cd_xHg_{1-x}Te$ ($x = 0{\cdot}23$) were studied using pulses significantly shorter than the experiments of Figueira *et al.* (1976) and the 'induced opacity' was observed to occur at sub-nanosecond speeds. This was explained in terms of an infrared field induced avalanche carrier generation similar to the description given by Nee *et al.* (1978). In the Kane process the threshold excess

energy for an electron to inelastically scatter another electron to the conduction band is approximately equal to the band-gap energy. To reach this excess energy, intraband absorption of a single quantum, two successive and three successive quanta of 10·6 μm radiation are required for the $Cd_xHg_{1-x}Te$, InSb and InAs respectively. Thresholds for the limiting behaviour were 0·2, 2 and 20 MW/cm² respectively and the transmissions are shown in fig. 3.5.2. The limiting effect corresponds to an abrupt falling edge of the transmitted pulse which could not be resolved on a nanosecond timescale. Coincident with this intensity change, a spectral broadening of the transmitted pulse was observed and interpreted as being caused by self-phase modulation from a rapidly time-varying refractive index. Jamison and Nurmikko (1979) constructed a model employing a set of coupled rate equations in a conduction band which has been approximated as a discreet multilevel system (the level separation being equal to the LO-phonon energy) in order to estimate the hole–electron distribution after some multistage free-carrier absorption and up to the initial stages of avalanche formation. In contrast to Nee *et al.*, they find that near the threshold for the Kane process, the collision rate is smaller than for both phonon emission and electron–electron scattering, but this mechanism can nevertheless lead to the very rapid build up of an electron hole plasma. The transmission cut-off is then due to the strong intervalence band hole absorption.

Fig. 3.5.2

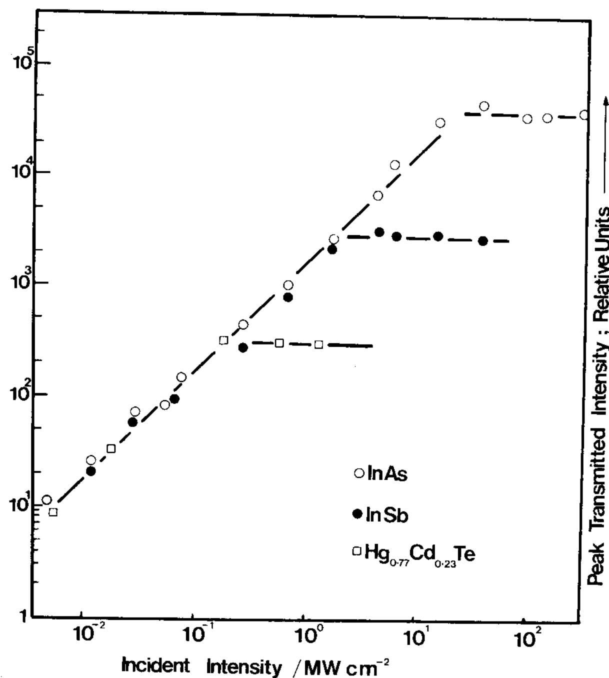
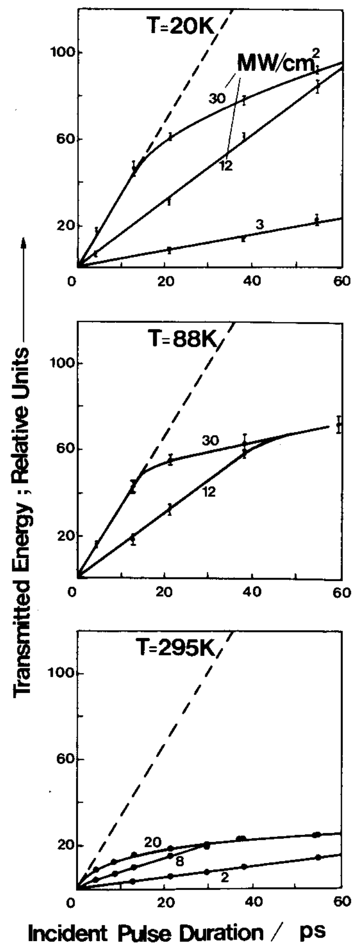


Illustration of the high-intensity transmission limit in InAs ($T=300$ K), InSb ($T=20$ K) and $Hg_{0.77}Cd_{0.23}Te$ ($T=20$ K). (After Jamison and Nurmikko 1979 a.)

Schwartz *et al.* (1980) have extended these investigations into the picosecond regime. 10·6 μm pulses of 5–60 ps duration were generated by utilizing two germanium optical switches opened and closed by pulses from a Nd : glass laser (see § 6.1). The transmission of these ultra-short 10·6 μm pulses through a 350 μm thick crystal of InSb was studied as a function of intensity, pulse width and temperature. The results (fig. 3.5.3) show that transmission limiting occurs at lower intensities for pulses of longer duration and for the sample at higher temperatures. At 20 and 77 K, reduced transmission occurs for pulses longer than 12 ps while pulses of 5 ps duration gave non-linear transmission at 295 K. The 20 K results were found to be qualitatively consistent with the avalanche process described previously (Jamison and Nurmikko 1979). At higher temperatures, additional carrier generation by two-photon absorption strengthens the effect.

Fig. 3.5.3



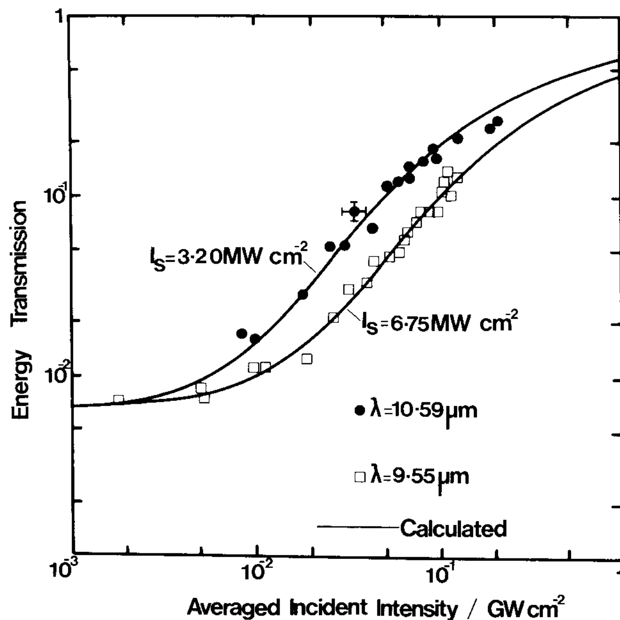
High-intensity 10·6 μm transmission limit in InSb for ultra-short pulses at different temperatures. The lines are drawn to aid the eye. The dashed lines depict transmission of 30 MW/cm² pulses in the absence of non-linearities. (After Schwartz *et al.* 1980.)

Schwartz and Nurmikko (1980) have also recently investigated 0·5 μm thick epitaxial PbTe using the picosecond infrared germanium switch. A dense electron–hole gas was created in PbTe using 1·06 μm radiation above the band-gap energy while a time delayed 10 ps pulse probed the transmission or reflection at either 10·6 or 5·3 μm. At 10·6 μm, the transmission decreased rapidly after the excite pulse was incident on the crystal and then recovered on a longer time scale (approximately 50 ps). The transmission appeared to be mainly affected by reflectivity changes at the first surface due to strong coupling of the infrared radiation with the photoexcited carriers. With 5·3 μm radiation, saturation of the absorption was observed through band-filling effects since this wavelength can couple valence and conduction band states in PbTe.

3.6. *Intervalence band saturation in p-type semiconductors*

Transmission saturation of 10·6 μm radiation in *p*-type semiconductors was first discovered by Gibson *et al.* (1972) using a mode-locked CO_2 laser. Germanium, silicon and gallium arsenide were found to exhibit saturation at intensities in excess of 10, 20 and 50 MW/cm² respectively due to the excitation of holes from the heavy to the light hole band. In silicon, the split-off band is

Fig. 3.6.1



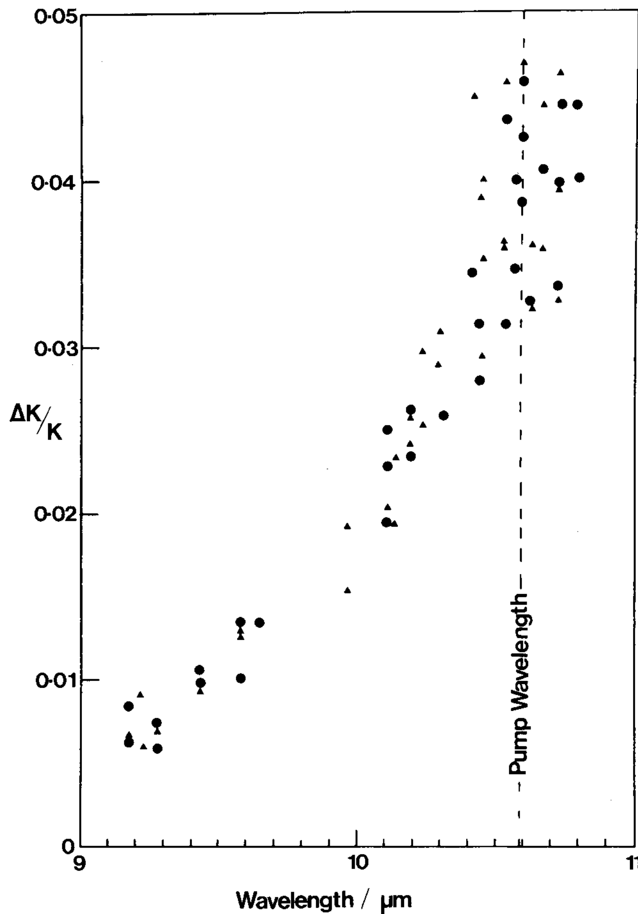
Typical space-time averaged transmission for Ge : Ga versus $W_{in}/(\pi\omega_0^2 T_p)$ at two wavelengths. Here ω_0 is the $1/e^2$ input intensity radius, T_p is the FWHM duration of the input pulse, and W_{in} is the input energy. Intensities refer to the plane perpendicular to the propagation vector within the sample. Small signal transmission was 0·0062 at 10·59 μm and 0·0070 at 9·55 μm. (After Phipps and Thomas 1977.)

also important. Further experimental work has concentrated on germanium and has led to its use as an ultra-fast mode-locking element for CO_2 lasers and also as a broadband saturable absorber for pulse shaping and isolator applications (see § 6.2).

In the 9–11 μm wavelength region, the hole absorption in germanium varies by only 6%. A careful study of the wavelength-dependence of the saturation effect by Phipps and Thomas (1977) shows that the non-linearity is also fairly uniform between 10·6 and 9·2 μm, fig. 3.6.1 ; the saturation intensity was found to vary from 3·2 MW/cm^2 for the longer wavelength to 6·8 MW/cm^2 for the shorter wavelengths.

For wavelength regions where short pulse lasers are not readily available, the exite-probe technique can be used in the frequency domain rather than the

Fig. 3.6.2



The fractional change ($\Delta K/K$) in the absorption coefficient of p-type germanium as a function of wavelength due to illumination at 10·59 μm ; ●, pump laser and probe laser polarizations parallel ; ▲, polarizations at right angles to each other. Pump intensity 1 MW cm^{-2}. (After Bishop *et al.* 1976.)

time domain to give insight into the nature of the saturation and the time scale of the relaxation processes. Bishop *et al.* (1976) used a 100 kW CO_2 laser operating at 10·59 μm and producing 200 ns pulses to excite a germanium crystal while a second CO_2 probe laser of 100 W power and generating 100 μs pulses monitored the change of transmission due to the higher power laser. The probe was tunable between 9·2 and 10·8 μm and the measured transmission change is shown in fig. 3.6.2. It can be seen that the pump causes an approximately lorentzian hole burning with a half-width at half-maximum of about 50 cm^{-1}. This is in agreement with a theoretical analysis by Keilman (1976) which predicted a 57 cm^{-1} spectral width at small saturation.

Keilman's model considers a distribution of inhomogeneously broadened two-level systems. This type of saturation model for semiconductors has been discussed in § 2.2. A broad smooth distribution of resonance frequencies with a width broad compared to the lorentzian width for a given transition is assumed. For inhomogeneous broadening, the dependence of absorption coefficient, α, on intensity, I, is given by

$$\alpha \sim \left(1+\frac{I}{I_s}\right)^{-1/2} \tag{3.6.1}$$

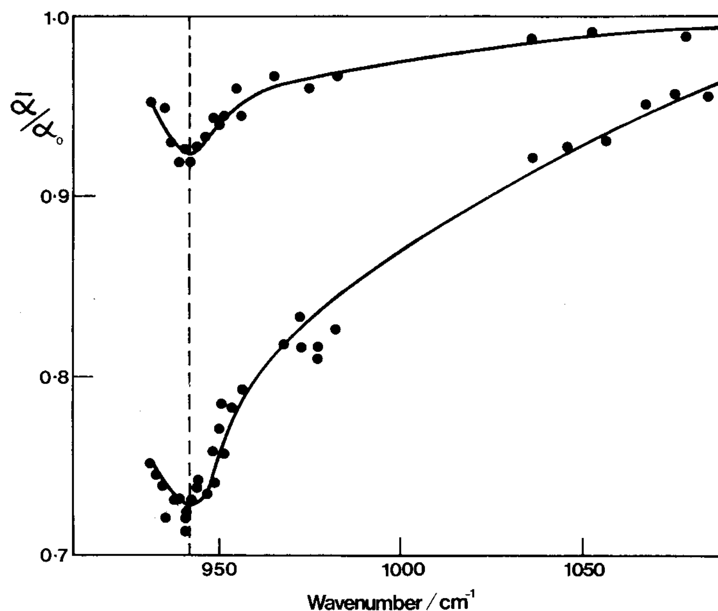
resulting in a saturation intensity, I_s, for germanium from Gibson *et al.* (1972) of 4 MW/cm^2 at 10·6 μm. The saturation intensity is proportional to $(T_1T_2)^{-1}$ and the net population density where T_1 is the relaxation time for the holes and T_2 is the dephasing time (§ 2.2). Thus, a single pulse saturated absorption measurement gives only the product of these times but a measurement of the lorentzian hole-burning width, $\Delta\nu$, leads to a determination of T_2 through

$$\Delta\nu = \left(1+\frac{I}{I_s}\right)^{1/2} \Big/ \pi T_2. \tag{3.6.2}$$

Keilman calculated the density of populated states in the light and heavy hole bands for a hole concentration of 10^{16} cm^{-3} at room temperature using Kane's model for the valence bands and deduced that $(T_1T_2)^{1/2} = 0{\cdot}55$ ps from the absorption measurements of Gibson *et al.* (1972). Relaxation of the holes most probably takes place through scattering by three optical phonons followed by several acoustic phonon interactions to take up the remaining energy. This leads to an expected value of T_1 on the order of 2·6 ps. The phase relaxation is dominated by lattice scattering of both lower and upper levels giving $T_2 = 0{\cdot}9$ ps is good agreement with the spectral hole-burning width observed by Bishop *et al.* (1976).

In further work, Keilman (1977) has observed a double dip resonance in excite-probe experiments, fig. 3.6.3. This shows a broad absorption dip similar to that observed by Bishop *et al.* (1976) due to hole burning and in this case gives a lorentzian (FWHM) width of 80–110 cm^{-1} corresponding to a value of $T_2 \sim 0{\cdot}12$ ps. The narrow absorption band superimposed on this structure was interpreted as being due to scattering of the pump into the probe beam as a result of the grating produced by interference of the two laser beams in the germanium (see § 4). The grating will have a lifetime and characteristic frequency width which depends on the population relaxation time, T_1. The width, given by $\Delta\nu = (1+I/I_s)^{1/2}/\pi T_1$, was measured to be between 15 and 25 cm^{-1} (FWHM) resulting in $T_1 \simeq 0{\cdot}6$ ps.

Fig. 3.6.3

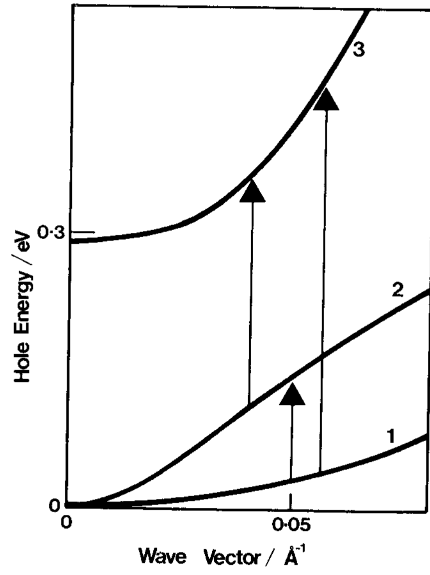


Experimental saturated absorption spectra in p-type germanium showing the grating dip on a broad hole-burning background. The saturating beam at 942 cm^{-1} has an intensity of 11 and 95 MW/cm^2 at its centre just inside the sample, for the upper and lower spectrum respectively. (After Keilman 1977.)

Keilman and Kuhl (1978) have shown also that 3 μm radiation can be modulated by 10 μm radiation in p-type germanium. Figure 3.6.4 shows the valence band structure of germanium. As previously described, 10 μm light can saturate hole transitions from the heavy to the light valence band ($1 \rightarrow 2$) at intensities in excess of 1 MW/cm^2. This causes bleaching of the states in the heavy hole band so that an increase of the transmission in the 3 μm wavelength region may be expected where these states are coupled to the split-off band, 3. A $LiNbO_3$ parametric oscillator pumped by the frequency doubled output of a Q-switched Nd : YAG laser with peak powers up to 500 W, was employed as the frequency tunable probe at 3 μm. With 9·6 μm pumping, a broad saturation dip was observed centred at 3360 cm^{-1} with a width (FWHM) of 1060 cm^{-1} as shown in fig. 3.6.5. An estimate of the dephasing time from these results gave $T_2 \sim 0{\cdot}016$ ps compared to $T_2 \sim 0{\cdot}057$ ps calculated for lattice interactions of this energy in the split-off band.

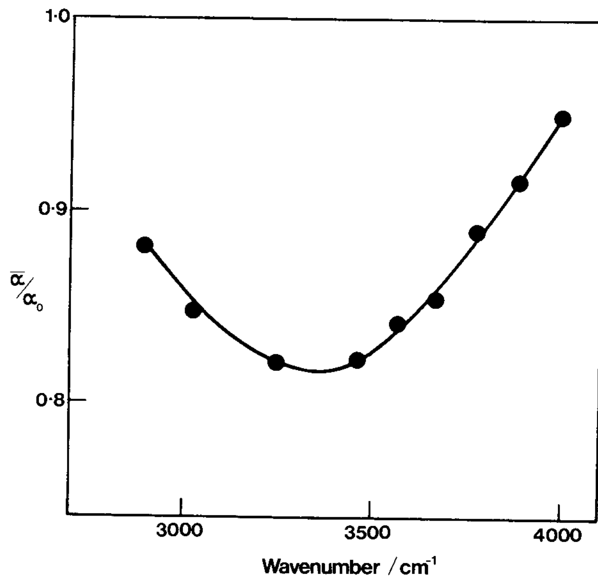
James and Smith (1979) have recently recalculated the wavelength dependence of the saturation of the intervalence band transitions in germanium at 10 μm. Previous calculations (Komolov *et al.* 1977, Sargent 1977) failed to predict the correct wavelength dependence. These authors used $\mathbf{k} \cdot \mathbf{p}$ theory taking account of the anisotropy and non-parabolicity of the bands. The results show good agreement with the previously observed increase in saturation intensity with increased photon energy. The temperature dependence was also calculated showing an increase in I_s with temperature due to the

Fig. 3.6.4



Spherical approximation of valence band structure of germanium. (After Keilman and Kuhl 1978.)

Fig. 3.6.5



Saturated absorption lineshape obtained near 3 μm wavelength in the presence of strong laser radiation at 1045 cm^{-1}. (After Keilman and Kuhl 1978.)

increased rate of phonon scattering. These calculations have been extended to other p-type semiconductors (James and Smith 1980 a, b) and the systematic dependence of the saturation intensity analysed in terms of the material parameters.

§ 4. Transient gratings

Previous sections of this article have described experimental studies of transient phenomena in semiconductors by making use of various non-linear optical interactions with laser radiation. In particular, double beam excite-probe methods have been described in § 3. A special situation, neglected so far, occurs when the two beams are coincident in time as well as space at the sample surface. Interference of the two coherent radiation fields can cause a spatial modulation of the optical properties of the material with a period, Λ dependent on the angle, 2θ, between the two beams :

$$\Lambda = \lambda/2 \sin \theta. \tag{4.1}$$

A moving interference pattern is produced by superposition of two light waves with different wavelengths but most studies have been carried out with two beams from the same laser. The resulting transient diffraction grating gives a powerful tool for the study of various relaxation and diffusion processes on short timescales by observing either the self-diffraction of the creating beams or the diffraction of a third, delayed pulse at the same or some other wavelength. Laser-induced gratings have practical applications in holography for real-time storage and processing of optical data as well as in distributed feedback lasers (Eichler 1977). The closely related application of phase conjugation by four-wave mixing will be discussed in § 6.4. Eichler (1977, 1978) and Vinetskii *et al.* (1979) have reviewed techniques and research into laser-induced gratings in solids, plasmas, laser materials and dye solutions. In this section we will concentrate only on recent grating studies in semiconductors.

Gratings arise because the absorption and the refractive index depend on the intensity of the incident radiation. The interaction may be described by third and higher-order coefficients of the non-linear susceptibility in the usual formalism of non-linear optics just as for four-photon mixing. The time constants of the material are accounted for in the frequency dependence of the susceptibility. Two extreme cases are usually considered for the production of diffracted radiation. One is the ‘ thin ’ grating or Raman–Nath region in which all of the diffracted waves have an optical path difference which is negligible compared to the wavelength. This applies to the most commonly met situation in semiconductors when a free carrier grating is generated at the surface by single photon interband transitions. When this condition is not satisfied, the grating is termed ‘ thick ’ and the Bragg condition (phase-matching) must be fulfilled for efficient generation of a diffracted beam.

In semiconductors, the grating is normally created by the generation of an excess carrier concentration due to the interfering light. These carriers may cause a modulation of both the absorptive and refractive properties of the material (i.e. amplitude and phase gratings respectively) so that a complete understanding of the diffraction process requires the determination of which

of these will dominate in any instance. Woerdman (1971) has distinguished between amplitude and phase gratings by comparing the self-image of a silicon grating in a probe beam with the self-image produced by a photographic amplitude grating. In most cases investigated so far, the phase grating is thought to dominate the diffraction. The change of refractive index, Δn, can be calculated to a first approximation from the Drude model for free carriers, § 2.4,

$$\Delta n = -\frac{4\pi Ne^2}{2nm^*\omega^2} \tag{4.2}$$

where N is the density of the excited electron–hole pairs, e is the electron charge, n is the refractive index, m^* is the reduced effective mass of an electron–hole pair, and ω is the frequency of the diffracted light.

Monitoring the change in diffraction efficiency of a probe beam at delays after the grating formation by the two excite beams can determine the diffusion or relaxation constants of the material. Recombination and diffusion affect the grating decay independently. For the case of a diffusion dominated grating, the decay time τ_D will be a function of spatial period, Λ so that the ambipolar diffusion coefficient, D_a, may be extracted from,

$$\tau_D = \frac{\Lambda^2}{4\pi^2 D_a}. \tag{4.3}$$

In self-diffraction studies, the diffusion coefficients have been calculated from the variation in peak diffracted intensity with the inverse square of the grating period. At high excitation levels, the diffusion coefficient may become carrier density dependent and the recombination rate increase due to the Auger and other density dependent processes.

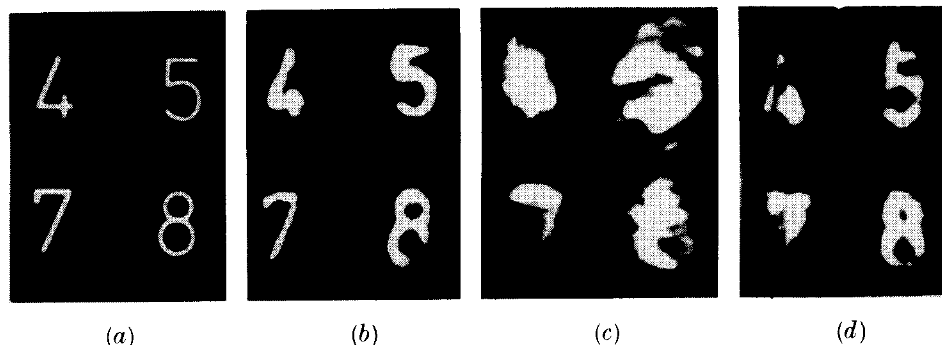
4.1. *Si*

The first semiconductor grating experiments were carried out in silicon by Woerdman and Bolger (1969), Woerdman (1970 a, b, 1971). Q-switched Nd : YAG laser pulses were split into two beams and then superimposed with some small angular separation on a silicon slice. A series of spots were observable on a screen behind the sample due to self-diffraction of the radiation (Woerdman and Bolger 1969). Intensities of approximately 1 MW/cm^2 gave estimated carrier densities of 10^{17}–10^{18} cm^{-3} and a refractive index modulation of about 10^{-4}, resulting in a phase grating with a diffraction efficiency of a few per cent (Woerdman 1970 a). This was used to create transient holograms, fig. 4.1.1, with lifetimes of 20–30 ns, consistent with a grating decay mechanism due to carrier diffusion.

Jarasiunas and Vaitkus (1977) found that the threshold for self-diffraction in silicon using a Q-switched Nd : glass laser was 0·1 MW/cm^2. The dependence of self-diffraction intensity on grating period was used to measure diffusion coefficients in two samples of 12 and 18·5 cm^2/s in good agreement with those deduced from Hall measurements.

Shifted holograms have been produced in silicon by placing the crystal in crossed electric and magnetic fields. The Lorentz force shifts the distribution of electrons and holes relative to the generating light field and resulted in a 30% increase in the intensity of the self-diffracted light (Vinetskii *et al.* 1977).

Fig. 4.1.1



Reconstructions from free carrier hologram in silicon. (*a*) Original transparency. (*b*) Reconstruction with $2L/c = 3$ ns. (*c*) Same as (*b*), but camera focused behind the reconstruction. (*d*) Reconstruction with $2L/c = 22$ ns. (After (Woerdman 1970 a.)

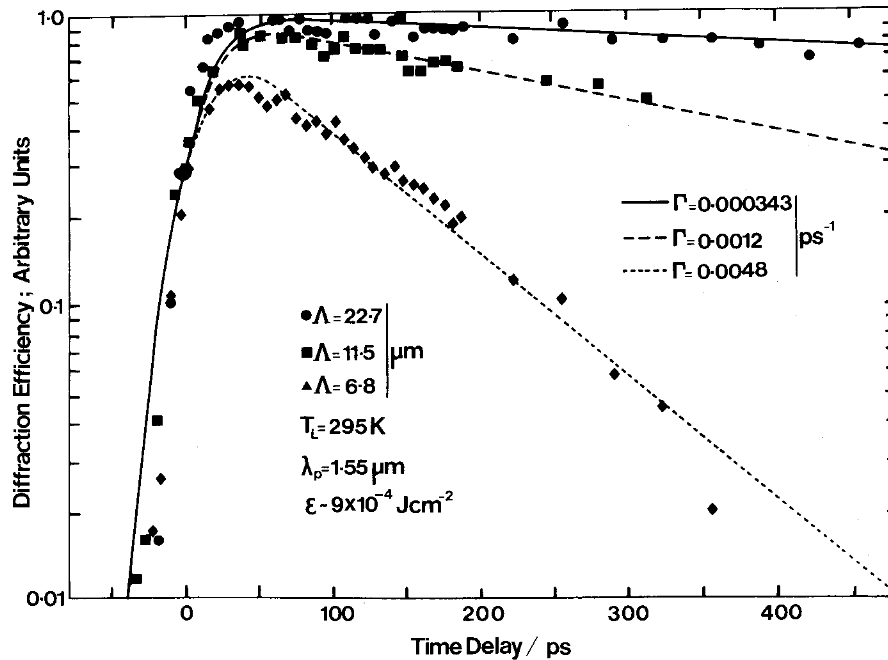
4.2. *Ge*

Temporary and permanent gratings have been formed on the surface of germanium by Wiggins and co-workers (1974, 1978, 1979) using two beams from a ruby laser producing 30 ns pulses at 6943 Å and the diffraction observed in reflection. From these studies, it was concluded that the change in the optical properties was due to heating of the surface. This was also the deduction of similar work by Vaitkus *et al.* (1978) using both ruby and *Q*-switched Nd : glass lasers with nanosecond pulse durations. Using mode-locked neodymium lasers with picosecond pulse widths, heating is avoided and the grating formed by the change in the refractive index is due to the spatial modulation in the large density of carriers generated by the laser excitation, § 3.1 (Kennedy *et al.* 1974, Shank and Auston 1975, Lindle *et al.* 1979).

Recently, Smirl, Moss and Lindle (1981 a, b), Moss *et al.* (1981) have used the transient grating technique to separate contributions from the density dependent diffusion coefficient and non-linear recombination coefficients of germanium using 35 ps pulses from a Nd : YAG laser. The interference between two spatial and temporally coincident excite pulses produced a spatially modulated free carrier grating close to the surface and the decay of this grating was monitored by measuring the first-order diffracted light in transmission from a third (probe) pulse at 1·55 μm as a function of time delay. Since the photon energy of the probe at 1·55 μm is below the direct band-gap of germanium (see § 3.1), the probe ‘ sees ’ an index grating. Figure 4.2.1 shows the room temperature results for three different angular separations, 2θ, of the excite pulses, and hence different grating spacings, Λ. Here the peak surface carrier density was 4×10^{19} cm^{-3}. Notice that the grating lifetime becomes smaller with decreasing grating spacing indicating increasing importance of diffusion for these smaller spacings. The calculated fits to the data using a Drude model for the free carrier plasma yields a diffusion coefficient, D, of 53 cm^2/s and a recombination lifetime larger than a nanosecond. At 100 K, similar measurements gave $D = 142$ cm^2/s. Experiments carried out at higher carrier densities, $N > 10^{20}$ cm^{-3}, (Smirl, Moss and Lindle 1981 b), showed that the Auger

recombination process becomes significant. Smirl *et al.* deduced from a numerical analysis of these high density transient gratings that the assumption of a density-independent diffusivity agreed best with the data contrary to that expected from simple models which do predict a density dependence.

Fig. 4.2.1



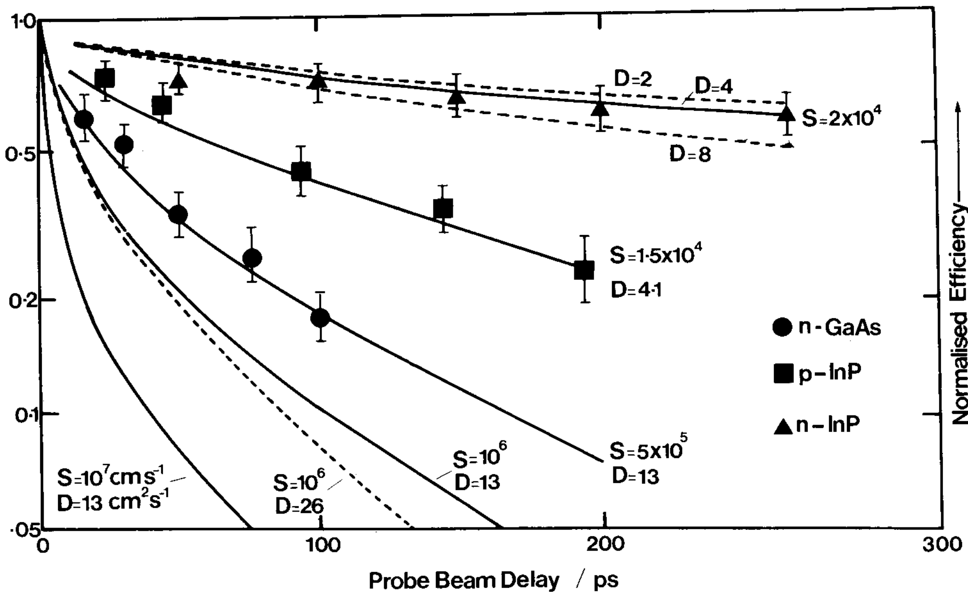
Probe diffraction efficiency in Ge, in arbitrary units, as a function of time delay between pump and probe pulses. (After Smirl *et al.* 1981.)

4.3. *GaAs, InP and InAs*

Hoffman *et al.* (1978), Jarasiumas, Hoffman *et al.* (1978) have used transient gratings to investigate surface recombination in GaAs and InP. The gratings were generated by the interference between two beams of 0·53 μm radiation from a frequency doubled Nd : glass laser with 6 ps pulse duration. A typical carrier concentration of $\sim 10^{20}$ cm^{-3} was estimated just after excitation due to single-photon interband absorption. A delayed probe pulse at 1·06 μm from the same laser (photon energy less than the band-gap) monitored the decay of the induced phase grating in time. Figure 4.3.1 shows the results for n-GaAs, p-InP and n-InP. Theoretical fits to the data for various values of diffusion coefficient, D, and surface recombination velocity, S, show that surface recombination is the dominate mechanism for the grating decay. (Bulk recombination rates such as Auger recombination would deplete the plasma on a time scale of hundreds of picoseconds.) This method thus gives a sensitive, direct method for measurements of recombination at semiconductor surfaces.

A ruby laser with 20 ns pulse duration was used by Wiggins and Qualey (1979 b) to assess InAs as a candidate for optical amplification by self-diffraction. Like the previous measurements in germanium using a nanosecond pulsed laser, the observed enhanced reflectivity due to a transient phase grating on the surface arises due to change in refractive index with radiation heating.

Fig. 4.3.1



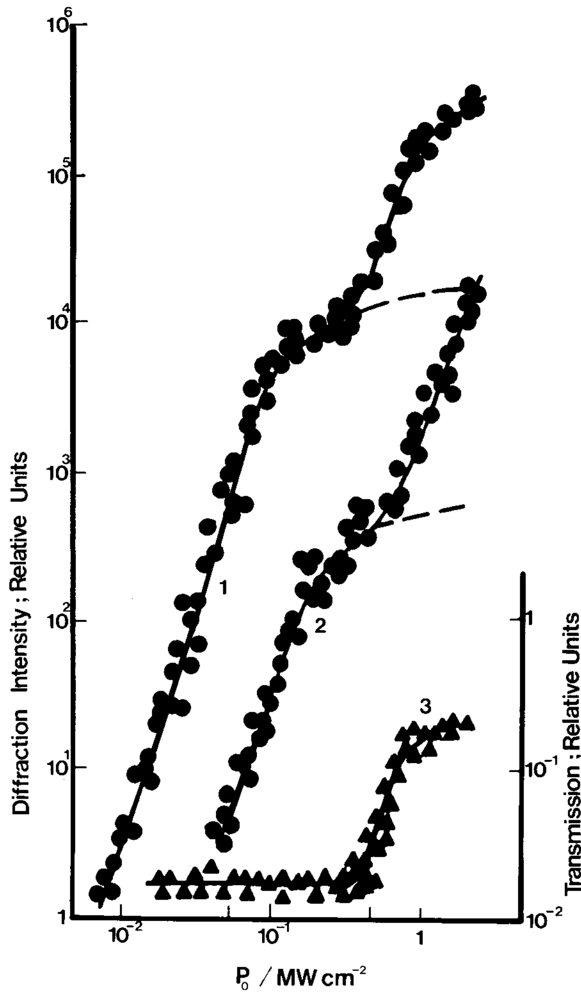
Experimental (points) and calculated (curves) diffraction efficiency versus probe delay for n-InP, p-InP and n-GaAs. (After Hoffman *et al.* 1978.)

4.4. *CdS, CdSe, CdTe and ZnSe*

Two-photon excitation has been used to create free carrier gratings in CdS (Borshch *et al.* 1973, Jarasiunas and Gerritsen 1978), CdSe (Jarasiunas and Vaitkus 1974, 1977, Jarasiunas *et al.* 1978), ZnSe (Jarasiunas 1977) and CdTe (Kremenitskii *et al.* 1979) with Q-switched ruby and Nd : glass lasers. The relatively high two-photon absorption coefficients of CdS ($E_g = 2{\cdot}41$ eV) and CdSe ($E_g = 1{\cdot}8$ eV) at the ruby laser photon energy allowed Jarasiunas and Gerritsen (1978) to observe several self-diffracted orders in sub-millimetre thick crystals showing that the thin grating criterion was obeyed. ZnSe ($E_g = 2{\cdot}7$ eV) has a smaller two-photon coefficient and a thick grating was created in this case. An analysis of the dependence of the self-diffracted peak intensity as a function of grating period gave ambipolar diffusion coefficients of 3 cm^2/s (300 K) and 20 cm^2/s (62 K) for CdS, 65 cm^2/s (13 K) for CdSe and 2·5 cm^2/s (300 K) for ZnSe. At large grating periods ($\sim 100\ \mu$m), recombination dominates diffusion for the grating decay in CdS. A measured lifetime of 9 ns at 62 K is consistent with recombination via exciton states.

Transient gratings have also been generated in CdS by single photon absorption of frequency doubled Nd : YAG at 0·53 μm (2·34 eV) by Eichler *et al.* (1978). The grating was studied by detecting the diffraction of pulses from a flashlamp pumped dye laser after creation by the YAG laser. The long pulse length of the dye laser allowed the investigation of the grating decay as a function of time. The photon energy was not sufficiently high to generate the grating by valence to conduction band transitions. At low pulse energies, the intensity dependence of the observed Bragg diffraction indicated that the absorption was probably due to transitions involving shallow impurity levels rather than by two photon absorption. Use of a 1 cm thick crystal of CdS

Fig. 4.4.1



Transmission (3) and lux-diffraction characteristics (1, 2) of CdSe single crystals in the exciton–phonon part of the absorption spectrum : (1, 2) grating periods 30 and 6·2 μm. The dashed curves are the lux-diffraction characteristics after subtraction of the absorption non-linearity. (After Baltrameyunas, Vaitkus and Yarashyunas 1976.)

allowed a diffraction efficiency of 25% with dye laser depletion of up to 50%. In this case, the decay time of the grating was found to be independent of grating spacing. Two characteristic decay time constants were also obtained for the decay of red luminescence from CdS.

A first order diffraction efficiency of over 30% has been measured in CdSe by Baltrameyunas, Vaitkus *et al.* (1976), Baltrameyunas, Jarasiunas *et al.* (1976), and Vaitkus *et al.* (1980) using a ruby laser which operates at an energy just below the fundamental absorption edge of this II–VI compound. The temperature of the crystal was adjusted to give a resonance between the photon frequency and the energy of the free $A(n=1)$ exciton minus the energy of one optical phonon. Efficient diffraction was observed at only 100 kW/cm^2 at 86 K, as shown in fig. 4.4.1, but the amount of scattered radiation decreased rapidly when the temperature was altered. A detuning of 4·5 meV decreased the absorption corrected diffraction efficiency by a factor of two. Decreasing the grating period reduced the number of diffracted orders because of the change from Raman–Nath to Bragg diffraction condition.

§ 5. Weak excitation studies

Many of the experiments involving lasers and semiconductors make use of the very large intensities available from pulsed lasers with $I \gg 1$ kW/cm^2 being usual; such work was reviewed in § 3. However, the linear optical properties of semiconductors have historically been determined using conventional spectroscopic techniques in which intensities are likely to have been on the order of 1 mW/cm^2 or less. This leaves a range of more than six orders of magnitude to be investigated, with the intensities required readily available from c.w. lasers.

Of course the distinction between weak and strong excitation studies (the latter discussed in the previous section) is to some extent arbitrary, but the experimental work in this section generally proceeds gradually from the linear regime to the beginnings of the non-linear regime, and is performed primarily with c.w. lasers.

The work we will discuss here is mainly concerned with phenomena near the band-gap energy in the direct band-gap III–V semiconductors InSb and GaAs. In both cases, large effects on both absorption coefficient and refractive index are seen which would be masked entirely at larger intensities. Although the two materials are very closely related to each other, the physics of the optical spectrum near the band-gap energy can be quite different, particularly at low temperatures. The difference results from the difference in the conduction band effective mass, m_c, which in turn can be regarded as being due to the difference in band-gap energy E_g, with (approximately) $m_c \propto E_g$: at 0 K in GaAs $E_g \sim 1{\cdot}52$ eV, $m_c \sim 0{\cdot}07$ m and in InSb $E_g \sim 0{\cdot}24$ eV, $m_c \sim 0{\cdot}014$ m. This has two well-known consequences: the calculated exciton orbit radius in InSb is ~ 700 Å compared to ~ 140 Å for GaAs so that the electron–hole correlation (which can be regarded as strongly screened if there is one free carrier per exciton volume) is screened at very low free carrier concentration in InSb ($\sim 10^{14}$ cm^{-3}); donor impurities (whose orbit radius is almost the same as the exciton radius) 'overlap' at the same low concentration in InSb leading to 'banding' of the donor electrons and preventing freeze out of free carriers in

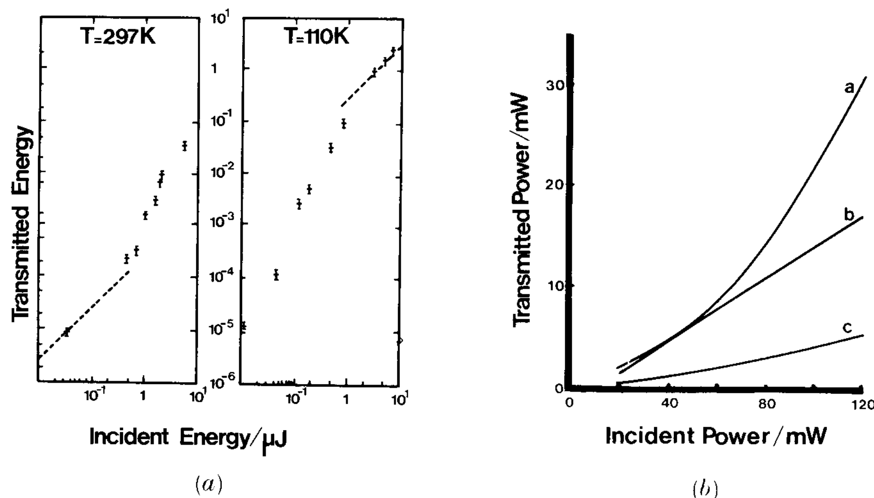
n-InSb. Consequently unless magnetic fields are applied, excitonic structure is absent in all but the purest samples of InSb at very low temperatures (the exciton binding energy in InSb is $\sim 0{\cdot}4$ meV $\equiv kT$ at $T \sim 5$ K) and even the remanent Coulomb correlation between electron and hole appears to play little part in determining the absorption spectrum just above the band-gap energy. In contrast, GaAs particularly below ~ 50 K, shows both strong excitonic structure below the band-gap and modification of the absorption spectrum above the band-gap due to Coulomb correlation. This constrast between the linear spectra of the materials is reflected in the non-linear properties as will be discussed below, suggesting at the very least that a variety of dynamic non-linear optical effects will exist in other materials and at other temperatures. Weak excitation studies remain to be investigated in most other direct gap materials although this should prove a fruitful area of research. Hildebrand *et al.* (1978) have already commented that some-non-linear absorption can be observed under c.w. excitation in GaSb, and as early as 1969 Ashkin *et al.* observed non-linear absorptive and refractive effects in CdS.

It should be emphasized that in both InSb and GaAs the non-linear effects are very large (e.g. effective $\chi^{(3)}$ $(\omega : \omega, -\omega, \omega) \sim 1$ e.s.u. has been measured in InSb) and consequently they may have immediate applications in other fields; for example, both InSb and GaAs have been used to demonstrate optical bistability at low c.w. laser powers (see § 6.5).

5.1. *InSb*

The saturation of optical absorption in InSb at 2 K just above the band-gap energy was investigated by Nurmikko (1976 a) with sub-nanosecond $5{\cdot}3$ μm pulses (fig. 5.1.1 (a)) and by Lavallard *et al.* (1977) (fig. 5.1.1 (b)) using a c.w.

Fig. 5.1.1



(a) (b)

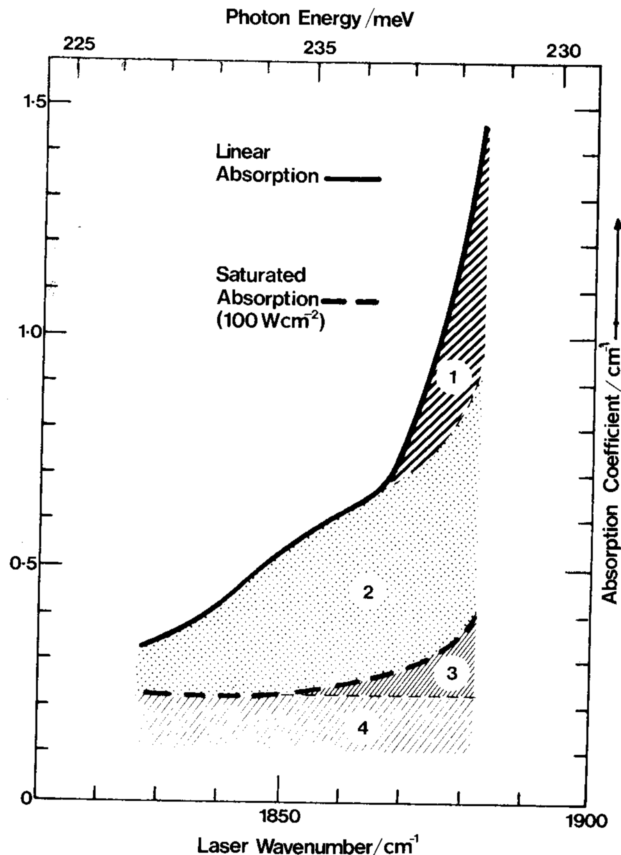
(a) The transmitted optical energy versus incident optical energy in a 25 μm thick InSb sample at $5{\cdot}3$ μm. (After Nurmikko 1976 a.) (b) Transmitted power versus incident power for CO line $\hbar\nu = 238{\cdot}2$ meV. The a, b, c curves are obtained for different positions of the laser spot on the sample. (After Lavallard *et al.* 1977.)

CO laser. The CO laser is particularly suited to studying InSb as it produces ~ 60 lines, approximately evenly spaced in the spectral region $5{\cdot}2$–$6{\cdot}0\ \mu$m which covers the band-edge region in InSb at both liquid helium and liquid nitrogen temperatures (although the band-gap is too narrow at room temperature). Lavallard *et al.* were able to saturate the absorption above the band-gap with intensities ~ 100 W/cm^2 in samples $\sim 60\ \mu$m thick. Furthermore they demonstrated that the CO beam transmission could be enhanced by shining a Nd : YAG laser beam ($1{\cdot}06\ \mu$m wavelength) coincidently on the sample and deduced that consequently the saturation in InSb was not explicable as self-induced transparency or ' state-filling ' (hole-burning) saturation. They were able to explain their results by assuming that the excited carriers relaxed to thermal Fermi distributions (through carrier–carrier scattering). The carrier temperature was assumed to result from the balance between the excess energy input for each created e–h pair (i.e. the photon energy, $\hbar\omega$, less the band-gap energy, E_g) and the energy loss due to recombination ; energy loss to the lattice was estimated to be negligible. The band-gap energy itself was measured from the luminescence spectrum so that band-gap renormalization (~ 3 meV in these experiments with carrier densities $\sim 4 \times 10^{15}$ cm^{-3}) was included. The luminescence spectrum was also consistent with the Fermi distribution.

Miller, Mozolowski *et al.* (1978) investigated the intensity-dependence of absorption coefficient and refractive index in low concentration *n*-InSb at 4 and 77 K, also with a CO laser, in the region just below the band-gap energy, where there is a linear absorption ' tail '. They observed a strong saturation of absorption in the tail at 4 K in the intensity range 1 mW/cm^2 to 30 W/cm^2 (fig. 5.1.2.) At 77 K, where the absorption tail is much longer (i.e. extends to lower energies), the absorption is linear until above ~ 30 W/cm^2 when it starts to rise ; Miller *et al.* ascribed this to induced free carrier absorption. At both temperatures, however, they observed that increasing the power in the Gaussian laser beam focused onto the sample led to a progressive distortion of the transmitted laser beam shape (seen in the diffraction far field behind the sample) at intensities $\gtrsim 30$ W/cm^2, the beam gradually broadening and acquiring a multiple ring structure. They explained this as a consequence of intensity-dependence of refractive index resulting in either self-focusing or self-defocusing of the beam. (The two are not distinguishable from their far-field behaviour). Thermal causes were shown to be unlikely, and the qualitative explanation suggested was non-linear refraction due to interband saturation by analogy with the well-known intensity dependence of refractive index seen in atomic vapours when pumping just below resonance ; in this case a self-defocusing non-linearity (i.e. refractive index decreasing with increasing intensity) is predicted. Holah *et al.* (1979) showed that this intensity-dependence resonated strongly as the band-gap energy was approached at 4 K, as expected from the saturation model.

The absorption at liquid helium temperatures was observed through photoconductivity by Seiler and Hanes (1979) in the band tail region (again using a CO laser). They observed a photoconductive signal even when pumping below the band-gap and proposed that this signal was due to transitions between acceptor states, just above the valence band, and the conduction band, although it seems likely that the experiments were performed at intensities

Fig. 5.1.2



InSb absorption tail at 5 K, showing linear absorption and saturated absorption results (heavy continuous and heavy dashed lines respectively) together with speculative divisions (light dashed lines) between portions of absorption due to differing causes. See test for discussion of regions 1–4. (After Miller 1981 b.)

where in the experiments of Miller, Mozolowski *et al.* (1978) this absorption had reached saturation.

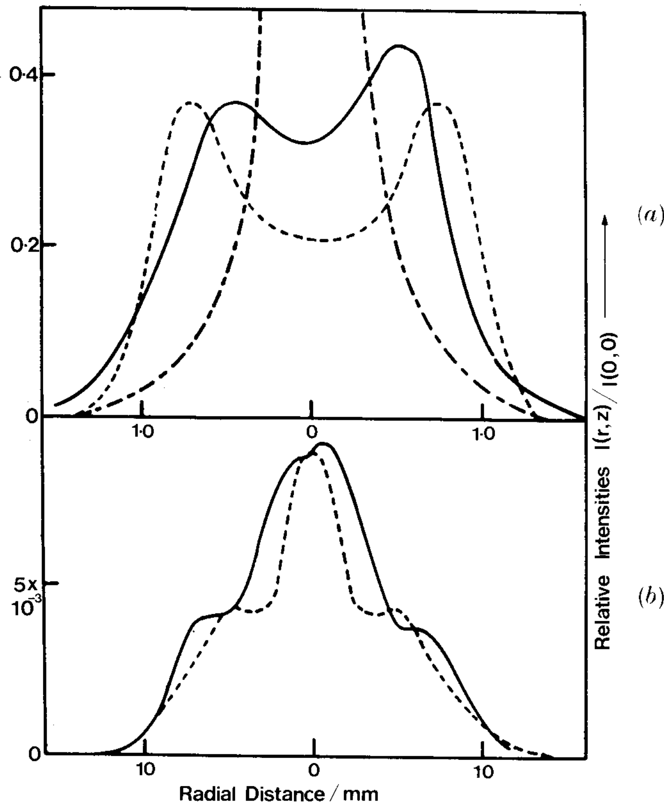
The absorption saturation at liquid helium temperatures was examined in greater detail by Miller (1981 b). He found that in the spectral region where the linear absorption had been satisfactorily explained by acceptor-to-conduction band transitions, the saturating absorption could be explained with good qualitative and quantitative agreement, by a simple model which predicts a homogeneous saturation for this type of process, with saturation intensities ~ 1 W/cm^2. However, other parts of the absorption tail were resolved in these experiments by their different saturation behaviour (see fig. 5.1.2). Region 4 is the background absorption measured at longer wavelengths and assumed to extend uniformly in this spectral region. Region 2 is the absorption supposed to result from acceptor-to-conduction band transitions. The absorption in

region 1 is distinguished by having a very low saturation intensity $\sim 0{\cdot}1$ W/cm^2. (The precise division between regions 1 and 2 is conjectural as it could not be accurately determined from experiments). By contrast, the absorption in region 3 which remains after saturating out regions 1 and 2, is substantially linear and appears to rise steeply as the band-gap energy (nominally $\sim 235{\cdot}5$ meV) is approached.

The absorption in regions 1 and 3 remains to be explained. Region 1 would be consistent with a relatively small number of states with large absorption cross-sections. Region 3 could be explained by, for example, scattering-assisted transitions coupling relatively large numbers of final states, and the absence of saturation in the band tail absorption at 77 K (Miller, Mozolowski *et al.* 1978) is consistent with a similar type of process. Hanes and Seiler (1980) demonstrated that absorption in this region below the band-gap at liquid helium temperatures caused a cooling of the electron distribution in the conduction band. They ascribed this effect to acceptor-to-conduction band transitions creating free carriers in the conduction band with excess energies below the mean energy per carrier in the existing electron distribution thereby reducing the mean energy per carrier. However, to explain their results within ~ 3 meV of the band-gap, they invoked a further acceptor impurity level 3 meV above the valence band for which there is apparently no independent evidence at present. Miller (1981 b) suggested that the cooling could also be explained by invoking carrier–carrier scattering to assist the absorption just below the edge in which case the energy to complete the transition would be extracted from the existing electron distribution resulting in an overall reduction of the electron distribution thermal energy. It is also likely that at the intensities used by Hanes and Seiler (1980) the acceptor-to-conduction band absorption was again saturated.

Weaire *et al.* (1979) demonstrated that the non-linear refraction at 4 K was indeed self-focusing (i.e. $dn/dI \equiv n_2 < 0$), as expected from the saturation model. They compared measured beam shapes in the near and far field behind the sample with theoretical profiles, calculated on the assumption that the incident gaussian intensity profile was maintained in propagating through the crystal but a gaussian phase ‘ lead ’ (or ‘ lag ’) was acquired. This model confirmed that the beginnings of self-focusing and self-defocusing are indeed totally indistinguishable in the far field but easily separable in the near field, fig. 5.1.3, and explained the characteristic ring structure ; the calculated profiles agreed well with experiment. They also demonstrated that beam profile distortion is a simple and comparitively accurate technique for measuring non-linear refraction in bulk materials with $n_2 \simeq -6 \times 10^{-5}$ cm^2/W being measured for InSb at ~ 5 K at a laser wavenumber of 1886 cm^{-1} ; as a general rule significant broadening and distortion of the beam will be seen in the far field for $\lambda/2$ to λ non-linear length change in the beam centre regardless of beam size. These observations also confirmed that the non-linear refraction is not thermal in origin since increasing temperature is expected to close up the band-gap leading to an increased refractive index (i.e. self-focusing) rather than the observed reduction. Miller, Smith and Johnston (1979) provided further evidence for this non-linear refraction through the observation of non-linear Fabry–Perot action and optical bistability (see § 6.5) ; this was extended to demonstrate ‘ optical transistor ’ action by Miller and Smith (1979) who also

Fig. 5.1.3



Experimental (solid lines) and theoretical (broken lines) intensity profiles (*a*) in the near fields at 7 cm from the sample and (*b*) in the far field at 189 cm. Data for 130 mW beam of 1·67 mm spot diameter on laser line at 1886 cm^{-1}; InSb sample and temperature 5 K. Theoretical profiles are shown for the self-defocusing condition ($x=3{\cdot}5$, dashed lines) in the near and far field and for the self-focusing condition ($x=-3{\cdot}5$, dot-dashed line) in the near field to emphasize that the experimental results originate from a defocusing. In the far field, the focusing and defocusing results are practically identical. The theoretical and experimental plots are normalized to give the same power levels. (After Weaire *et al.* 1979.)

presented the first attempt at a quantitative model for non-linear refraction by saturation in InSb by scaling from the atomic case, although the relative influences of interband and intraband relaxation were not clear. Also, the form of the resonance of the non-linear refraction at 4 K was quantitatively a poor fit (although displaying the correct qualitative features) and the absolute magnitude required an essentially unknown parameter (T_1/T_2).

One major difficulty in any *a priori* model for the non-linear refraction around and expecially below the band-gap is that the basic mechanism for the absorption is not understood in sufficient detail as has already been discussed above in this section, and yet the absorption is fundamental to obtaining any saturation effect. The basic qualitative model for the non-linear refraction

is simple: the finite absorption of radiation below or about the nominal band-gap creates free carriers inside the bands by some mechanism (e.g. scattering-assisted absorption): these carriers alter the absorption for frequencies above the band-gap and, by causality, the refractive index below the band-gap is altered as a result. The difficulties with any complete model lie in knowledge of (i) the precise absorption mechanism, (ii) the distribution of the excited carriers among the states which in turn requires (iii) a knowledge of the scattering mechanisms and even to some extent (iv) the precise nature of the states they occupy as these may be influenced by many body effects.

Miller, Smith and Wherrett (1980) (MSW) considered two limiting cases of the saturation model—' direct saturation ' and Burstein–Moss shifting. In the ' direct saturation ' case, the scattering between k-states is neglected entirely, and absorption is presumed to result from broadening of the direct interband transitions (T_2 broadening): thus the semiconductor bands (of which only heavy hole and conduction bands are considered dominant) can effectively be divided into separate two-level systems, one for each k. This is a generalization of the scaling model of Miller and Smith (1979). In the Burstein–Moss shifting model, the carriers are excited by an unspecified mechanism through a phenomenological absorption coefficient α and now intraband scattering is assumed strong giving a thermal distribution (in this case a 0 K Fermi distribution) and a resultant total blocking of a part of the absorption spectrum. It is shown that when the models are chosen with the same (measured) absorption coefficient, the results are identical within a small factor, and both can explain the observed magnitude of the steady-state non-linear refraction at 4 K assuming interband recombination times (approximately hundreds of nanoseconds) are the dominant relaxation mechanism, τ_R. This simple Burstein–Moss blocking model gives for n_2 ($\equiv dn/dI$)

$$n_2 = -\frac{2\pi}{3n}\left(\frac{eP}{\hbar\omega}\right)^2 \frac{\alpha\tau_R}{\hbar(\omega_g-\omega)\hbar\omega}, \tag{5.1.1}$$

where P is the usual momentum matrix element, n is the linear refractive index, $\hbar\omega$ is the photon energy and $\hbar\omega_g$ is the band-gap energy.

The reason for the poor resonance agreement of the ' direct saturation ' mechanism lies in its poor prediction of the form of the absorption below the gap for T_2 independent of $\mathbf{k}$. However, the Burstein–Moss shifting model makes no prediction at all of absorption coefficient. Because the pumping is below the band-gap, the connection with the work of Lavallard *et al.* (1977) is not clear.

Wherrett and Higgins (1981) have investigated the direct saturation and Burstein–Moss shifting models in greater detail, including the effects of finite temperature and initial carrier concentration, and the light-hole band. They also demonstrate theoretically that it is indeed comparatively difficult to saturate the absorption below the band-gap when it is due to scattering-assisted transitions (in this particular case T_2 broadening) and also that ironically the non-linear refraction, which is due to saturation, can occur in the presence of substantially linear absorption at the pumping frequency, as is observed experimentally. In the experimental measurements it is usually the first-order coefficient, n_2, of refractive index change with intensity that is measured. In the saturation models to lowest order the induced populations are created by

linear absorption, giving a refractive index change proportional to intensity ; hence any intensity dependence of absorption only enters to give higher order changes in refractive index. MSW also interpreted the data of Miller, Smith and Johnston (1979) on non-linear Fabry–Perot action as implying an eventual saturation of the change in refractive index itself, consistent with a saturation model.

The basic theory of the Burstein–Moss shift effect on refractive index has also subsequently been discussed by Moss (1980) who points out that there is empirical evidence in, for example, Ge, that free carriers can be created (and give rise to photoconductivity) for photon energies as low as 0·1 eV below the band-gap. His analysis differs from that of MSW and Wherrett and Higgins (1981) by assuming that the absorption edge is approximately ' square ' and moves upwards with increasing carrier concentration at a rate determined by the Burstein–Moss shift in a parabolic band whereas the former authors treat the density of states and the form of the saturable absorption edge self-consistently using a parabolic density of states throughout.

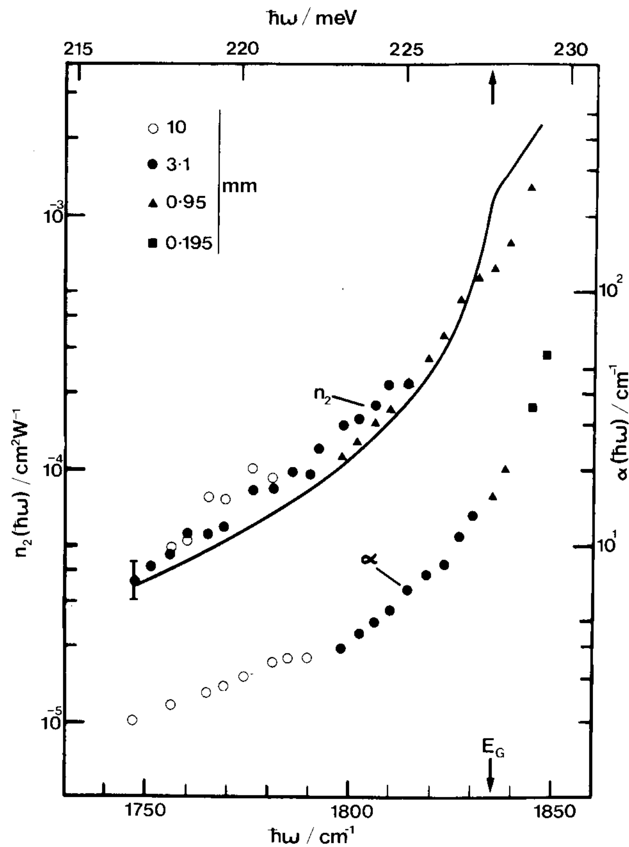
MSW also showed that non-resonant induced free carrier plasma refraction was unlikely to explain the observed non-linear refraction because it is too small. This was confirmed by transferred beam distortion experiments (Miller, Smith and Seaton 1981 b) which showed that, whereas self-defocusing caused by a strong pump beam caused a similar effect in a weak probe beam at the same wavelength, as the photon energy of the probe was moved away from the band-gap the induced distortion reduced and could only be restored by increasing the pump power ; if the effect had been due to the non-resonant plasma contribution the transferred distortion would have increased for lower probe photon energies.

The discussion so far has concentrated on the results at liquid helium temperatures but Miller, Mozolowski *et al.* (1978) also observed non-linear refraction at 77 K. Miller, Smith and Seaton (1981 a, b) measured $n_2 \simeq -3\times10^{-3}$ cm^2/W at 1856 cm^{-1} in n-InSb at 77 K, equivalent to an effective degenerate four-wave $\chi^{(3)}(\omega:\omega,-\omega,\omega)\sim 1$ e.s.u. (typical non-resonant values of $\chi^{(3)}$ are $\sim10^{-10}$ e.s.u.) and ascribed this to the same basic mechanism of non-linear refraction by band-gap resonant saturation. Miller, Seaton *et al.* (1981) investigated the non-linear refraction at 77 K in greater detail and measured n_2 between 1750 and 1850 cm^{-1} (the nominal band-gap at 77 K is $\sim$1835 cm^{-1}). They proposed a saturation model, based in this case on Boltzmann rather than 0 K Fermi statistics as appropriate to these conditions, but otherwise similar to the Burstein–Moss shift model of MSW.

They were able to obtain a good fit to their experimental results using no fitted parameters and measured absorption coefficients (see fig. 5.1.4). Using measured absorption coefficients, Smith and Miller (1980) were also able to obtain good agreement on the Burstein–Moss shift model for the n_2 measurements at 5 K (see fig. 5.1.5) although in this case the good agreement must be regarded as partly fortuitous due to the difficulty in measuring the relevant absorption coefficients.

Koch *et al.* (1981 a) have recently calculated the plasma density dependence of refractive index in InSb near the band-gap, including band-gap renormalization, for specific cases at both 5 and 77 K. They find for the conditions they choose (chosen to simulate the experimental conditions used by Weaire *et al.*

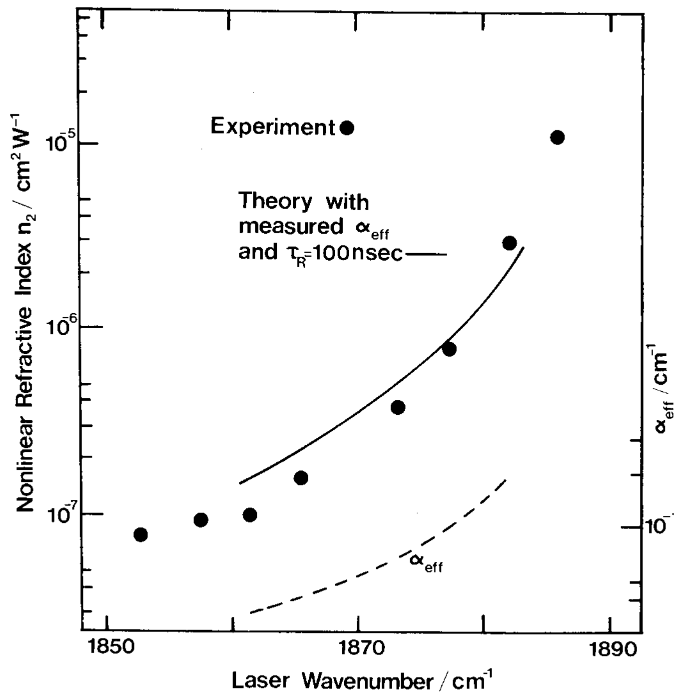
Fig. 5.1.4



Measured n_2 and absorption coefficient α as a function of photon energy $\hbar\omega$ compared with theory (solid line) for n_2 (Miller *et al.* 1981 a) for various lengths of InSb samples at 77 K.

(1979) and Miller, Smith and Seaton (1981 a, b), that the refractive index just below the band-gap decreases with increasing carrier concentration, as observed experimentally, with magnitudes in approximate agreement with experiment. In this calculation the band-gap renormalization is assumed to shift the entire conduction and valence bands uniformly relative to one another and the collision-broadening of the states is neglected. Non-parabolicity is included in the band structure, but it is not clear if this makes any significant difference for this calculation compared to a parabolic approach. The change in refractive index is deduced through the Kramers–Krong relations from the calculated change in the absorption spectrum (which includes both band-gap renormalization and saturation effects). Band-gap renormalization always leads to a reduction in effective one-electron band-gap energy with increasing carrier concentration. Therefore for frequencies below the band-gap, it should certainly result in an *increase* in refractive index for higher concentrations (as may

Fig. 5.1.5

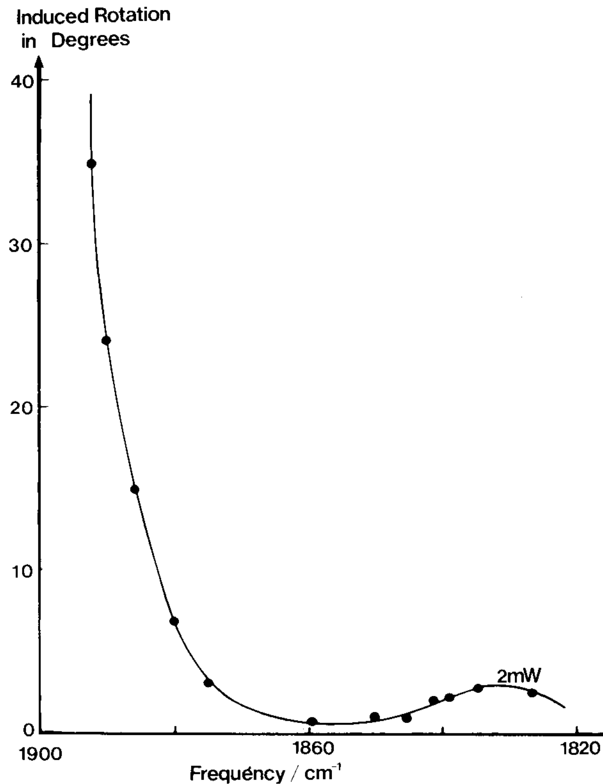


Measured n_2 and absorption coefficient α as a function of photon energy compared with theory (solid line) for n_2 (Smith and Miller 1980) for InSb at 5 K, fitted using a recombination time of 100 ns.

readily be verified from the Kramers–Krong relations). The observed *decrease* in refractive index in their calculations therefore implies that, in this region, the saturation contribution is dominant. In the limit of negligible renormalization contribution, the calculation of Koch *et al.* (1981 a) should reduce to the Burstein–Moss shift model of MSW at 0 K, the Boltzmann model of Miller, Seaton *et al.* (1981) at high temperatures and the general formalism presented by Wherrett and Higgins (1981) at arbitrary temperatures (with a small correction for non-parabolicity).

The discussion so far has centred on non-linear effects without applied magnetic field. MacKenzie *et al.* (1980) observed intensity dependence of Faraday rotation in cooled InSb in a magnetic field using a c.w. CO laser, showing resonances as the laser radiation approached the band-gap energy and also at the energy associated with acceptor-to-conduction band transitions (see fig. 5.1.6) ; this phenomenon appears to be the analogue of the non-linear refraction at zero magnetic field and can be qualitatively explained as follows. With the applied field, the conduction band in particular shows strong splitting into Landau levels ; alternative levels in the scheme correspond to opposite electronic spins with levels of one spin having a preferentially strong optical interaction with one sense of circular polarization. The direct saturation of

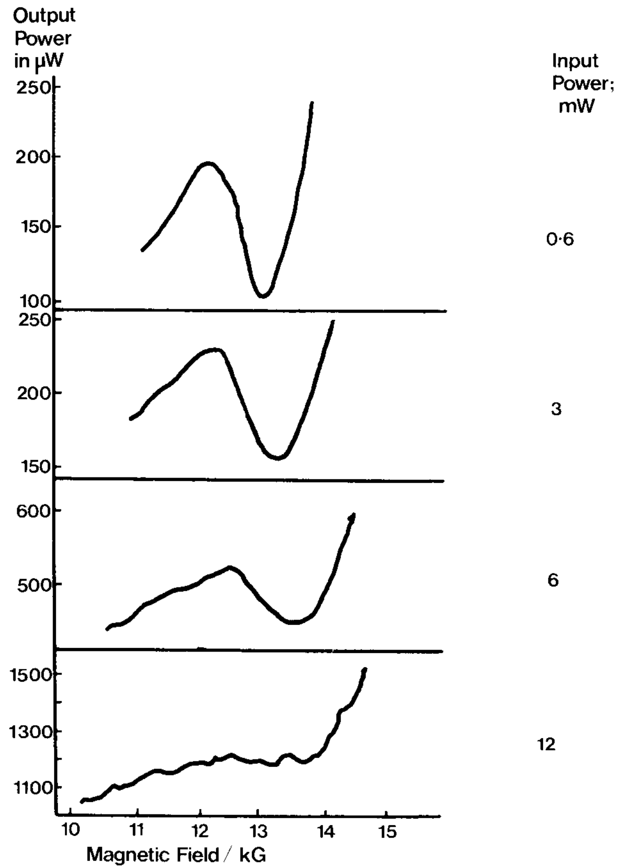
Fig. 5.1.6



Intensity induced Faraday rotation at a fixed magnetic field of ~ 6 kG as a function of radiation frequency in InSb at ~ 15 K with discrete points joined by line for clarity. (After MacKenzie *et al.* 1980.)

such an interband magneto-optical transition has been observed and is shown in fig. 5.1.7. When linearly polarized light is partly absorbed near to the gap, the resultant saturation will be predominantly that of the lowest available spin level, resulting in a larger change in the refractive contribution from one spin level compared to the other. This imbalance in changes in the refractive contributions results in a net change in the relative velocities of the two circular polarizations resulting in a change in the (Faraday) rotation of the linearly polarized light passing through the crystal. As this rotation (and hence the relative phase change) can readily be measured to better than 1°, this represents a very sensitive method of measuring these bulk non-linearities, effectively much more sensitive than the beam profile distortion methods used at zero field which are sensitive only to $\sim 180°$ phase changes. Induced rotations as large as 10^3 degrees per W/cm^2 per cm length were observed, and could also be made to saturate at higher powers, qualitatively consistent with a saturation origin for the effect. Theoretical calculations based on a ' direct saturation ' model (Dennis *et al.* 1980, 1981) were able to predict the overall form of the observed non-linear Faraday rotation.

Fig. 5.1.7

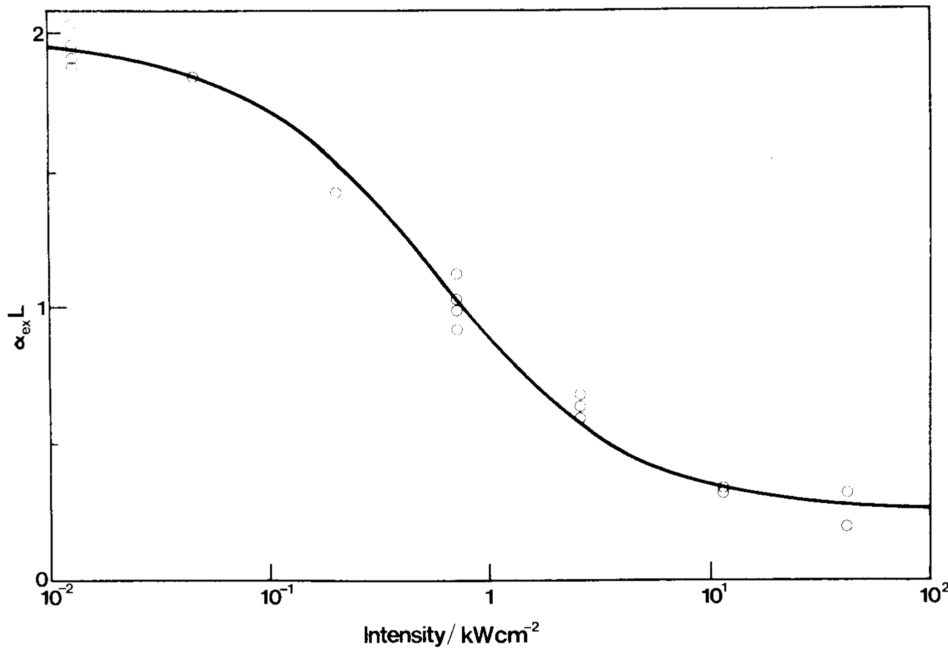


Transmission of InSb (4×10^{14} cm^{-3}, thickness 125 μm) for incident radiation of 1920 cm^{-1} showing the lowest interband transition into the 0↑ Landau level being progressively saturated up to incident powers of 12 mW (~ 100 W/cm^2). (After MacKenzie *et al.* 1981.)

5.2. *GaAs*

Gibbs *et al.* (1979 a) investigated the saturation of the excitonic resonance in GaAs at 10 K with intensities between 10 W/cm^2 and 100 kW/cm^2 using a c.w. dye laser. Their investigation was motivated partly in an effort to find saturable absorbers with low saturation intensities, extending the earlier work of Shah *et al.* (1977) who has observed excitonic saturation (see § 3.2 and fig. 3.2.2), and led to the empirical observation of a Bloch-like saturation of the excitonic absorption with a saturation intensity of ~ 150 W/cm^2 (when corrected for gaussian beam form and finite optical thickness) (see fig. 5.2.1). They demonstrated that this intensity is approximately that required to generate one carrier pair or exciton for each excitonic ' volume ' in the crystal (i.e. the volume of the lowest excitonic orbit) which implies a density $\sim 10^{17}$ cm^3. Thus on order of magnitude grounds for the saturation intensity they were

Fig. 5.2.1



Saturation of the excitonic absorption in a 0·42 μm thick GaAs sample at 10 K, plotted as $\alpha_{ex}L(\equiv \ln(1/\text{transmission}))$, where α_{ex} is the effective absorption coefficient and L the sample length, against intensity. The solid curve is a fitted theoretical curve on a Bloch-like (homogeneous) saturation model. (After Gibbs *et al.* 1979 a.)

unable to distinguish between three possible mechanisms for the excitonic saturation ; (1) exciton–exciton collisions, (2) screening of the electron–hole Coulomb interaction by free carriers and (3) reduction in the number of potential exciton-forming electrons available in the valence band ; each of these three mechanisms would be characterized by the same induced carrier and/or exciton density. They also observed that there was a significant component of this below band-gap absorption which did not saturate, and hence limited the practical usefulness as a saturable absorber, particularly for absorptive bistability (see § 6.5). Gibbs *et al.* (1979 c) investigated the change in refractive index in cooled GaAs (seen on a weak probe beam below the exciton resonance), induced by a mode-locked argon ion laser pumping above the band-gap. The refractive index changes were seen through the induced tuning of the Fabry–Perot cavity formed from the reflection-coated faces of the crystal itself. They were able to separate the effects of thermal changes and those due to excitonic effects by the difference in their temporal behaviour and the sign of dn/dI (positive for thermal, negative for excitonic). They saw the excitonic phase shift falling off approximately as $1/\Delta\lambda$ from 100 to 300 Å below the gap, as expected for a Block-like effect, and noted that the excitonic index change saturated at intensities consistent with the observed excitonic absorption saturation ; they also found the observed phase shifts in agreement with

approximate estimates for excitonic effects, and concluded they were observing the consequences of carriers, created by the argon laser pulse, progressively screening out the exciton resonance with the consequent change in refractive index. The excitonic effect disappeared gradually with increasing temperature above ~ 100 K as expected because kT is then much greater than the binding energy. Gibbs *et al.* (1979 d) subsequently observed optical bistability in GaAs between 5 and 120 K (see § 6.5) thus confirming that the excitonic saturation effect was indeed associated with a non-linear refraction (created by the same beam). They were also able to switch this device in less than 1nS using a 0·6 nJ 5900 Å pulse, thus corroborating their previous measurements of such two-beam refractive index changes (Gibbs *et al.* 1979 c).

Koch *et al.* (1981 b) have attempted to calculate the intensity-dependent changes of refractive index in GaAs in conditions appropriate to those of Gibbs *et al.* (1979 d) including many body effects. Since the actual observation of bistability was at comparatively high intensities (10^5W/cm^2) Koch *et al.* suggest that the change in refractive index is due to plasma effects, these having previously been observed by Hildebrand *et al.* (1978) in GaAs at 10 K. In their calculation, they include the effects of saturation, band-gap renormalization, collision broadening and excitonic enhancement. They calculate the change of absorption coefficient due to the change in plasma concentration and deduce the change in refractive index through the Kramers–Kronig relations. They estimate that the change in refractive index due to excitons (which can exist only below the critical Mott density) is much smaller than that due to the plasma effect. Gibbs *et al.* (1981) however, estimate, on the basis of the observed Bloch-like resonance, that relatively large intensity-dependence of refractive index should also exist at low intensities up to the onset of saturation, with $n_2 \sim 4 \times 10^{-4}$ cm^2/W. At these intensities (i.e. < 100 W/cm^2) it is unlikely that the strong plasma effects would exist. An *a priori* theory for the Bloch-like saturation remains to be developed for the GaAs case, although Goll and Haken (1980) do predict Bloch-like saturation under a variety of approximations.

§ 6. Applications

In this section we describe some of the ways in which the physical mechanisms responsible for the dynamic optical non-linearities discussed above are being employed. This research area may itself be described as dynamic with many inventive experiments being carried out recently and may well lead to a variety of fast optical device applications in the future.

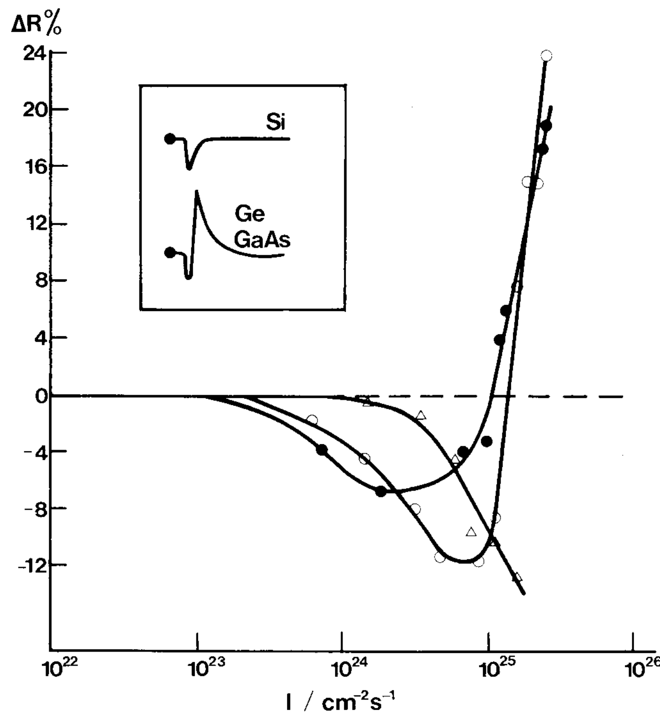
6.1. *Optical gating*

The creation of large densities of non-equilibrium carriers at the surface of a semiconductor by the absorption of short pulse laser radiation with photon energies above the band-gap energy can alter the reflection and transmission properties of the material on a very short time scale. Soon after the development of glass lasers, it was found that the rapid reflectivity change in germanium at large intensities allowed this material to be used as a passive Q-switch for a ruby laser when incorporated as a cavity mirror (Carmichael and Simpson 1964). Other semiconductors, Si, InP, InSb, and $GaAs_xP_{1-x}$ have also

successfully Q-switched a Nd : glass laser (Sooy *et al.* 1964). However, surface damage has limited their usefulness for this purpose. Indeed, it is quite probable that the change in reflectivity in the visible region of the spectrum at high intensities is due to a molten surface layer. (Birnbaum 1965, Birnbaum and Stocker 1966, 1968, Blinov *et al.* 1967 a, b, Bonch-Bruevich *et al.* 1968, Galkin *et al.* 1968.)

More recently, dynamic semiconductor plasma reflectivity has been successfully employed to gate sub-nanosecond pulses of 10 μm radiation. This spectral region lacks a convenient wide gain bandwidth laser for the generation of picosecond pulses. In the gating technique, a short pulse of laser radiation incident on the semiconductor creates enough carriers at the surface for the plasma energy to exceed 0·117 eV (10·6 μm). Figure 6.1.1 shows the changes in reflectivity for Si, Ge and GaAs at 10·6 μm as a function of intensity of Q-switched ruby laser pulses. These results, by Galkin *et al.* (1968), show the same characteristic shape as the reflectivity of doped semiconductors as a function of carriers concentration. Germanium, for instance, exhibits a minimum in reflectance for 10·6 μm radiation at a carrier concentration of 2×10^{19} cm^{-3} as the plasma frequency approaches the radiation frequency. For slightly higher concentrations, the refractive index tends to zero resulting in a large reflectivity. An analysis of the change of reflectivity such as that shown in fig. 6.1.1. required the correct interpretation of the contributions of

Fig. 6.1.1



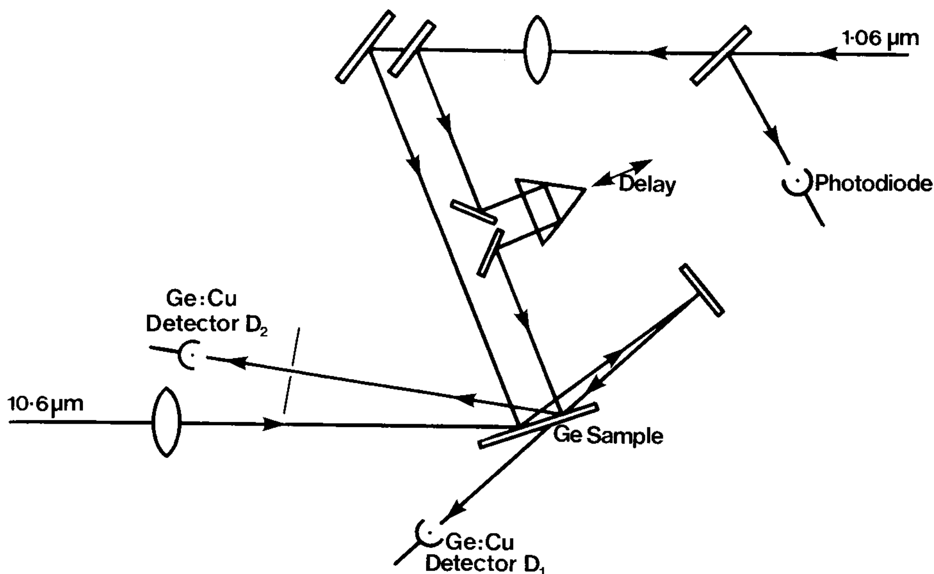
Plot of ΔR at a probing radiation wavelength 10·6 μ versus the ruby-laser exciting-light intensity. ▲▲▲, Si, ●●●, Ge, ○○○, GaAs. The insert shows the form of the ΔR oscillograms. (After Galkin *et al.* 1968.)

diffusion and recombination to the carrier number just as in the case of dynamic laser induced free carrier gratings as discussed in § 4.

High speed switching of CO_2 laser radiation using the reflection and transmission properties of an optically generated plasma in germanium has been reported by Alcock *et al.* (1975). In these experiments, 10·6 μm pulses were incident on a slab of polycrystalline germanium at Brewster's angle to minimize reflection to $\sim 3\%$. These 200 ns duration infrared pulses were synchronized with 2 ns visible pulses from a ruby laser so that when both beams overlapped simultaneously on the semiconductor, essentially complete reflection of the 10 μm radiation was observed for ruby intensities greater than 15 MW/cm^2. The rise and fall of the reflected infrared pulses took place in a time approximately equal to that of the ruby laser pulse. The transmission was also substantially decreased on the time scale of the ruby laser pulse rise time but had a longer recovery time which was dependent on the visible laser intensity. This may be understood by considering the rapid diffusion of carriers into the crystal bulk. The surface density is reduced in a time dependent on the diffusion constant of the carriers so that the reflectivity will drop when the plasma frequency falls below the photon energy, however, these carriers continue to contribute to the absorption for a time dependent on the recombination rate. For high CO_2 laser intensities, a second reflected pulse was observed several nanoseconds after the first which may be due to heating of the carriers by the 10 μm radiation when the plasma becomes underdense.

Jamison and co-workers (1976, 1978) extended this technique into the picosecond regime by using 6 ps mode-locked Nd : glass laser pulses to create the high density plasma in germanium. The Brewster angle reflected radiation

Fig. 6.1.2



Experimental arrangement for measuring reflected energies and the temporal autocorrelation function of the reflectivity. (After Jamison *et al.* 1976.)

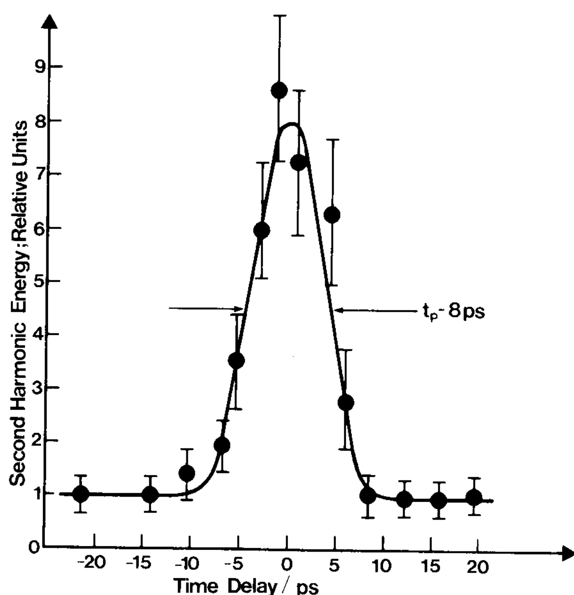
from a 1 W c.w. CO_2 laser increased to 85% for incident 1·06 μm energies of $\sim$0·19 mJ/cm^2 [5]. An optical variable delay gate was used to measure the temporal duration of the transient reflectivity as shown in fig. 6.1.2. A second delayed 1·06 μm beam was arranged to illuminate another part of the germanium slab close by. The reflected CO_2 laser pulse caused by the first 1·06 μm beam was returned to the germanium for a second time and could be reflected again if it was synchronized with this second 1·06 μm beam. By varying the time of arrival of this second laser, a measure of the autocorrelation function of the transient reflectivity was obtained, resulting in a measurement of $\sim$100 ps for the reflected pulse length at moderate 1·06 μm power densities. Larger pulse lengths were observed at the highest intensities. The fall in 10·6 μm reflectivity after this interval of time could be interpreted in terms of the accepted diffusion constant for germanium (65 cm^2/s) even for carrier densities up to 10^{19}–10^{20} cm^{-3}.

A very useful variable length picosecond gate is possible using the same experimental set-up as is shown in fig. 6.1.2 except that the *transmitted* light is employed instead of the reflected light after the 10·6 μm radiation has been incident on the germanium for a second time. Jamison and Nurmikko (1978b) used a TEA CO_2 laser in conjunction with a mode-locked Nd : glass laser to create megawatt picosecond pulses at 10·6 μm by this method. The reflection of 10·6 μm radiation at the first location generates a pulse with a rapidly rising leading edge and a decay tail of $\sim$100 ps in duration. The 10·6 μm beam transmitted through the germanium at the second location was abruptly cut off due to reflection losses and intervalence band absorption of excess holes after the 1·06 μm excitation. The time interval between the opening and closing of the switch was determined by the delay in arrival between the picosecond 1·06 μm pulses. The ratio of the intensity of the switched pulses to the background was 250 and was further increased by passing the pulses through a hot CO_2 cell acting as a narrow band absorber. Pulse widths of less than 5 ps were measured using a Michelson interferometer and second harmonic generation. Figure 6.1.3 shows a result of the autocorrelation measurement. Variable width ultra-short pulses at 10 μm have considerable potential for the study of ultra-fast processes in the infrared. For instance, the avalanche absorption process in InSb and other small gap semiconductors has been investigated using this technique as mentioned in § 3.5.

Corkum *et al.* (1978) have demonstrated an alternative method for obtaining high power ultra-short pulses at 10·6 μm using a germanium optical switch. In this case, a pulsed dye laser operating at 0·6 μm controlled the reflection of a low power c.w. CO_2 laser to give sub-nanosecond pulses. These pulses were injected into a 7 atm TE CO_2 laser to induce mode-locking. Measurements using an optical Kerr effect detection system showed that these pulses had durations of less than 100 ps (Corkum and Alcock 1978).

Another laser switching technique employing a semiconductor involves the dynamic Burstein–Moss shift. As discussed in § 4, a large density of photo-generated carriers can shift the absorption edge to higher energy enabling the increased transmission of a probe beam at energies just above the band-gap. Nurmikko and Pratt (1975) used a 200 ns train of $\sim$1 ns mode-locked pulses from a CO_2 laser operating at 9·6 μm to cause a modulation of the absorption edge in a 5 μm thick film of $Pb_{1-x}Sn_xTe$. This excitation photon energy was

Fig. 6.1.3



A result of the 10·6 μm SHG autocorrelation measurements on Ge. (After Jamison and Nurmikko 1978 b.)

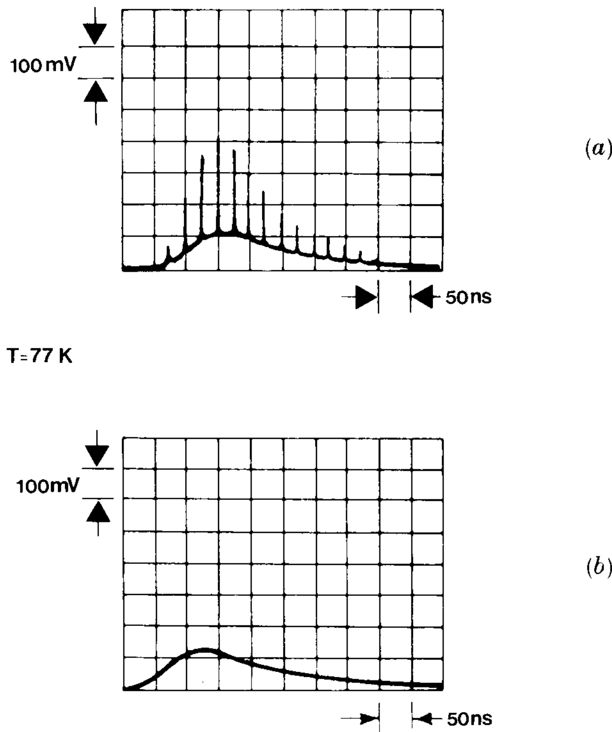
26 meV above the band-gap energy of the alloy at 77 K. A probe beam consisting of a 100 ns pulse with a smooth temporal envelope at a wavelength of 10·6 μm (10 meV above the band-gap energy) was simultaneously incident on the crystal and showed a corresponding rapid modulation in its transmission as shown in fig. 6.1.4. The ultimate speed limiting process of this type of shutter is the interband recombination rate, which in this case for $Pb_{1-x}Sn_xTe$ was of the order of 100 ps.

6.2. *Mode-locking*

One of the applications of saturable absorbers with fast recovery times is in the passive mode-locking of lasers. In § 3.6, we reviewed studies of the intervalence band saturation of *p*-type semiconductors. The very short relaxation times and the broadband saturation make them good candidates for mode-locking infrared lasers. Investigations on germanium around 10 μm showed that the hole-burning saturation has a T_2 relaxation time of less than a picosecond and is fairly uniform between 9·2 and 10·6 μm. Gibson *et al.* (1974) demonstrated mode-locking of an atmospheric pressure CO_2 laser with germanium, resulting in 1·5 ns pulses. A coated, *p*-type germanium crystal was utilized as both a mode-locking element and an output window for the laser and it was suggested that the same crystal could also be used simultaneously as a power meter via the photon drag effect. Using a double Rogowski-type of cavity with two amplifier, mode-locked pulses of 500 ps could be generated.

Feldman and Figueira (1974) used a 1 mm thick, AR coated, germanium slice to mode-lock a sub-atmospheric TE CO_2 laser. The shortest pulses were observed by operating the laser near threshold and placing the crystal close to the output window, fig. 6.2.1. Pulse compression was observed in the ~20 pulses of the train. Widths varied from 3 ns at the start of the train to 0·04 ns at the end. A saturable absorber will transmit the peak of a pulse preferentially with respect to leading and trailing edges leading to pulse narrowing on successive traverses of the cavity.

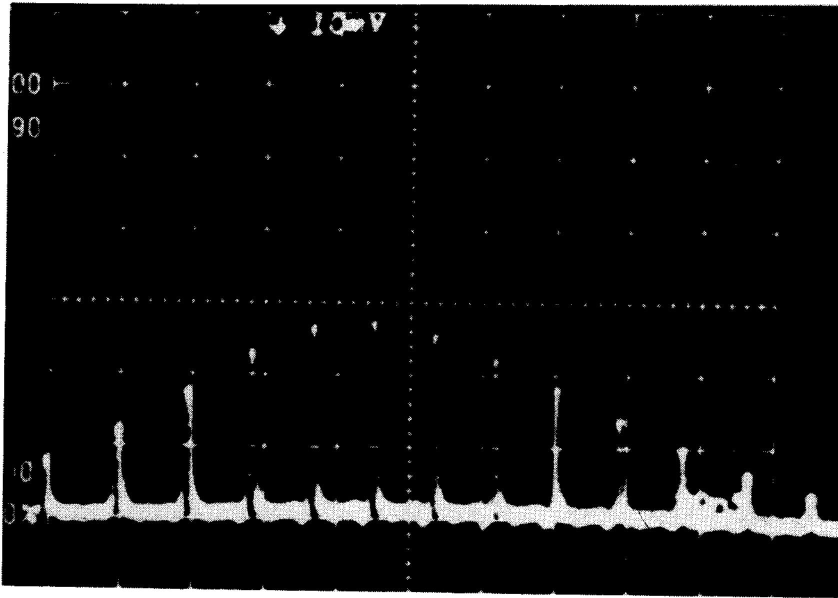
Fig. 6.1.4



Transmission of the 10·6 μm wavelength radiation in $Pb_{1-x}Sn_xTe$ (*a*) in the presence and (*b*) in the absence of pump beam. (After Nurmikko and Pratt 1975.)

Shorter pulse durations have been reported by Alcock and Walker (1974), by *p*-type germanium mode-locking a TE CO_2 laser operating at a gas pressure of 12 atm. (Increasing the pressure of the CO_2 laser increases the amplifying bandwidth by allowing broadening and overlapping of the rotational lasing lines.) Pulses of 75 ps duration were measured by upconversion of the CO_2 output by mixing with Nd : YAG laser pulses in a proustite crystal. The sum frequency output at 0·96 μm could be observed using a streak camera with S-1 response.

Fig. 6.2.1



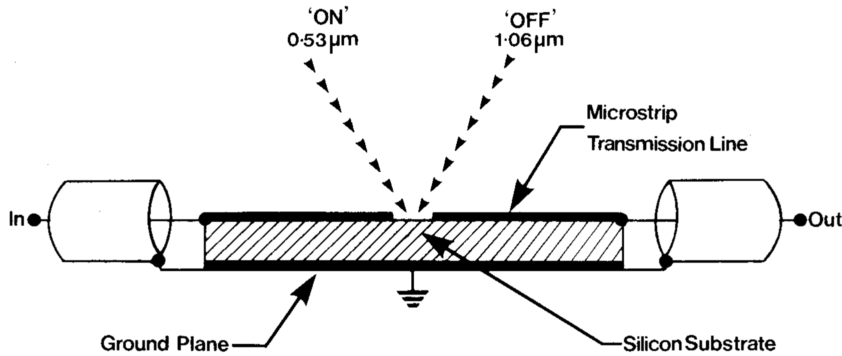
Passively mode-locked 10·6 μm pulse train at 20 ns/div using p-type Ge. (After Feldman and Figueira 1974.)

6.3. *Optoelectronic gating*

One of the most important developments to come from large excitation electron dynamics in semiconductors using short pulse lasers is electrical switching and gating. Conventional all electronic devices are normally limited to response times of $\sim 10^{-10}$ s and are more typically 10^{-9} s, although for particular applications some sampling gates can be as fast as 25 ps. Making use of the ultra-fast interband photoconductive response of semiconductors to short optical pulses, detectors, pulse generators and sampling gates can now be constructed with measured response times less than 10 ps (Auston, Lavallard, *et al.* 1980). These laser operated devices are also capable of switching voltages in the kilovolt range (Le Fur and Auston 1976, Morou *et al.* 1980) ; can have a maximum jitter of as little as ± 2 ps, (Mourou *et al.* 1980) ; and have found application in gating of micro- and millimeter waves (Johnson and Auston 1975), streak camera triggering (Mourou *et al.* 1980, Mourou and Knox 1980), high-speed sampling (Lawton 1976, Antonetti, Malley *et al.* 1977, Leonberger and Moulton 1979, Auston *et al.* 1980), active shaping of optical pulses (Le Fur and Auston 1976, Antonetti, Migus *et al.* 1977, Agostinelli *et al.* 1979, Stavola *et al.* 1979), and measurement of carrier lifetimes, (Lee *et al.* 1977, Johnson *et al.* 1980, Moyer *et al.* 1980).

The laser activated photoconductive switch was introduced by Auston (1975, 1976) making use of a high resistivity silicon transmission line with aluminium conductors evaporated on the top and bottom surfaces as shown in fig. 6.3.1. A sub-millimetre gap in the upper conductor prevented transmission of a d.c.

Fig. 6.3.1



An optoelectronic gate based on surface photoconductivity and volume photoconductivity due to the 0·53 and 1·06 μm optical pulses on the transmission properties of the silicon microstrip structure. (After Auston 1975.)

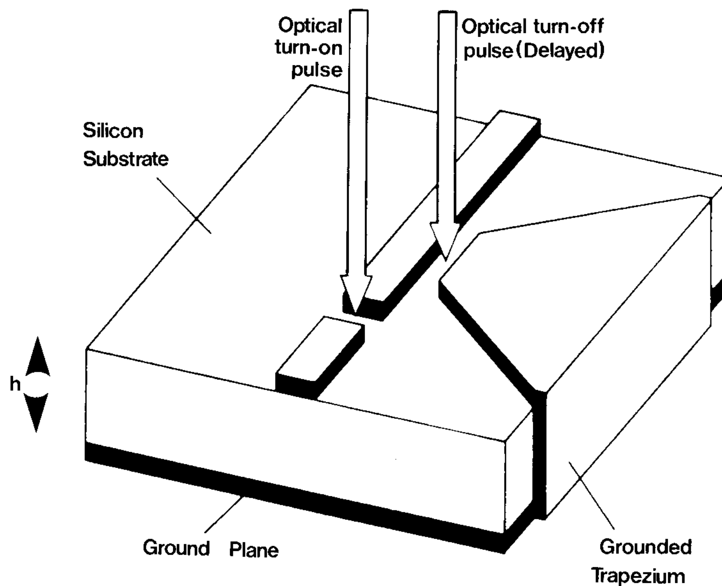
voltage across the device. Absorption of a 5 ps, 0·53 μm optical pulse from the frequency doubled output of a mode-locked Nd : glass laser produced a thin layer of high conductivity in the gap, turning the switch on and allowing transmission of the electrical signal. To turn the switch off, a delayed 8 ps pulse at 1·06 μm from the same laser created carriers throughout the depth of the silicon because of the smaller absorption coefficient at this wavelength, thus shorting the transmission line and terminating the electrical signal. To determine the response time of the device, two transmission gates were operated in tandem. The first gate generated an electrical impulse of 35 V and a duration of 15 ps, while the second gate (also with a 15 ps aperture) was used to measure the electrical pulse transmission for different delays between the two gates. Rise times of 10 ps were recorded.

This first device was capable of switching d.c. voltages of up to 100 V. By using pulsed electrical voltages to avoid breakdown of the bias signal, Le Fur and Auston (1976) demonstrated switching of 1·5 kV. Improvement in hold-off voltage has also been demonstrated by cooling the silicon to increase the resistance and avoid thermal runaway (Mourou *et al.* 1980).

An alternative technique for rapidly terminating the electrical pulse is to use the second delayed laser pulse to short circuit the transmission line as shown in fig. 6.3.2. Because only a surface conducting layer is required to activate the shunt, this method allows the same wavelength to be used for both optical pulses. Operation of a silicon optoelectronic gate with a shunt to terminate the electrical signal has been demonstrated using GaAs diode lasers (Platte 1976) and dye lasers (Auston *et al.* 1976, Castagne *et al.* 1976).

Although the electrical switching time for silicon is fast, the recombination time for the carriers is rather slow, thus limiting the repetition rate of this device to less than a megahertz. Several other semiconductors with shorter carrier life times have been investigated. Lawton and Scavannec (1975), Lee (1977) and Mourou and Knox (1979) have constructed and tested the performance of semi-insulating GaAs with a carrier lifetime of less than 100 ps which

Fig. 6.3.2



Electro-optic switch showing use of delayed optical pulse at the same wavelength to rapidly terminate the electrical pulse. (After Platte 1976.)

should allow repetition rates in the GHz range. This also means that no 'switch off' optical pulse is necessary in order to achieve electrical pulses of approximately 100 ps in duration. GaAs has the added advantage of offering higher room temperature mobility and dark resistivity than silicon. Lee used a high-strength dielectric insulator between the line and ground plate to allow d.c. voltages up to 5 kV to be applied without breakdown. However, the maximum output voltage obtained was 0·6 kV, thought to be due to the reduction of mobility via intervalley scattering at high electric fields (the mechanism which gives rise to the Gunn effect). This would appear to limit the usefulness of GaAs for high voltage operation but more recently Mourou and Knox (1979) have demonstrated switching of voltages up to 8 kV with 90% efficiency using GaAs in a configuration in which the semiconductor replaces a section of the central conductor of an HN connector.

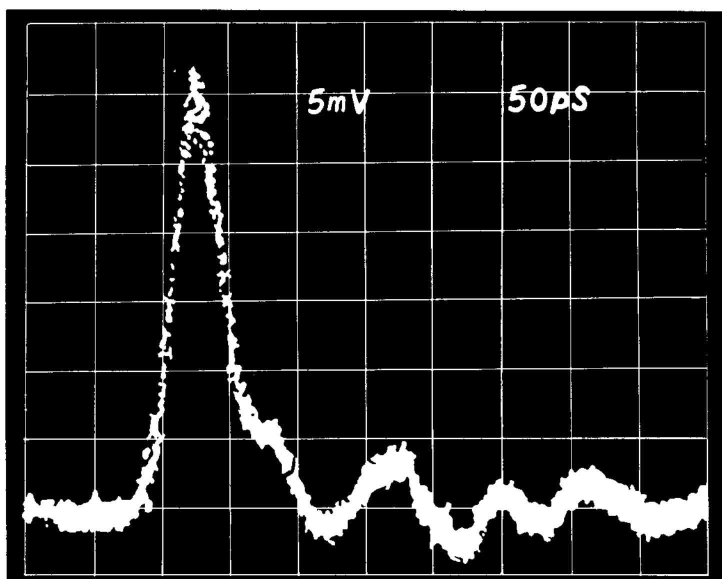
A successful optoelectronic switch has been fabricated from Fe-doped InP by Leonberger and Moulton (1979). The design incorporated a gap of only 3 μm width and the response was analysed with various c.w. mode-locked lasers (frequency doubled Nd : YAG, GaAs/GaAlAs and doubled $Nd_{0\cdot5}La_{0\cdot75}P_5O_{14}$). A switch risetime of 50 ps was achieved with 0·53 μm radiation and the results showed that the switch can be used to generate a pulse train having a 900 MHz repetition rate and a bandwidth of 6 GHz.

The II–VI alloy, $CdS_{0\cdot5}Se_{0\cdot5}$ has been used by Mak *et al.* (1980), because of its larger band-gap (2·0 eV) allowing a high dark resistance at room temperature. The 20 ns carrier lifetime gave relatively long electrical pulses after switch on by a 0·53 μm optical pulse but these pulses could be shortened to a nanosecond by using a coaxial pulse forming network.

Recently, Auston, Johnson and Smith (1980) have reported a reduction of the photoconductive response time of thin films of crystalline silicon from 1 ns to 8 ps by ion-bombardment with molecular oxygen atoms for the fabrication of fast photodetectors. Auston, Lavallard *et al.* (1980) have also used amorphous films of silicon prepared by thermal decomposition of silane to give a rapid photoconductive response. An autocorrelation technique using two identical strip line detectors has resulted in measured electrical pulse widths of 10 ps for amorphous films prepared in ultra-high vacuum, fig. 6.3.3.

Care must be taken in the construction of these devices to attain the minimum capacitance of the gap and the maximum transmission bandwidth. Several authors have discussed the optimum design requirements for the applications previously mentioned, Johnson and Auston (1975). See also, Castagne *et al.* (1978), De Fonzo (1978), De Fonzo *et al.* (1979), Lee *et al.* (1980), Meriau *et al.* (1977) and Platte and Appelhaus (1976).

Fig. 6.3.3



Response of am-Si detector to 8 ps pulses from a synchronously pumped Rh6G due laser. A sampling scope with a 25 ps rise time was used. Horizontal and vertical scales are 50 ps/div and 5 mv/div respectively. (After Auston, Lavallard, Sol and Kaplan 1980.)

6.4. *Phase conjugation*

The subject of phase conjugation has excited considerable interest recently (see, for example, Yariv 1978, 1979 for a review). It represents a concept in applications of non-linear optics entirely different from the classic frequency mixing and harmonic generation because it deals with the ability of certain non-linear effects to process optical images. The particular process of ‘ phase-conjugation ’ is where perturbations on an optical phase front are reversed in

sense, i.e. phase leads are converted to phase lags and vice versa. The classic demonstration of phase conjugation is to aberrate a plane by passing it through, for example, a piece of roughened glass, and then to generate a 'phase conjugate' beam from this aberrated beam and pass it, exactly counterpropagating to the original beam, back through the aberrating glass; the phase conjugate beam will have exactly the opposite phase aberrations to those originally introduced by the glass and the phase conjugate beam after passing through the glass will be totally unaberrated as can be verified by viewing it directly. The term 'phase conjugate' arises from the fact that the reversal of phase aberrations can be formally expressed by taking the complex conjugate of the spatial variation of the phase, e.g. if a field can be described by the form

$$\mathbf{E}_0 \exp [i(\omega t - kz + \phi(x, y)],$$

where $\phi(x, y)$ describes the lateral variation of phase then its phase conjugate could be expressed as

$$\mathbf{E}_0 \exp [i(\omega t + kz - \phi(x, y))]$$

which describes a counter-propagating phase conjugate. (Reflection from an ordinary mirror would give a field of form

$$\mathbf{E}_0 \exp [i(\omega t + kz + \phi(x, y))].$$

Phase conjugation is also sometimes called the 'time reversal' although this terminology can be misleading as there is no reversal of causality involved. If the field is expressed in the form

$$\mathbf{E}_0 \cos [\omega t - kz + \phi(x, y)]$$

then its phase conjugate if counter-propagating to the original field, can be written as

$$\mathbf{E}_0 \cos [\omega t + kz - \phi(x, y)]$$

or equivalently

$$\mathbf{E}_0 \cos [\omega(-t) - kz + \phi(x, y)],$$

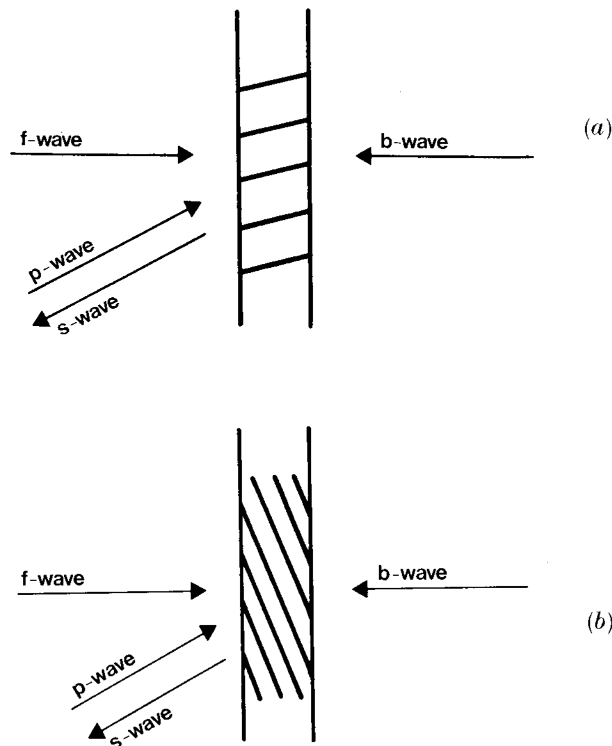
hence the term 'time reversal'.

There are various ways of generating phase conjugates, and the term is generalized to cover not only those processes in which the conjugate is generated counterpropagating to the original field. The important point is that the sense of the phase 'aberrations' should be reversed; this cannot be accomplished by a linear process but can be achieved through, for example, stimulated Brillouin scattering, three-wave mixing or (degenerate) four-wave mixing (see, for example, Yariv 1978, 1979) and it is this latter process where dynamic non-linear optical processes in semiconductors become particularly relevant because degenerate four-wave mixing (DFWM) requires only waves of *one* frequency and hence active processes (e.g. saturating absorption and/or refraction) can give rise to it.

Phase conjugation through DFWM can be given a simple physical interpretation. When two beams of the same frequency interfere they produce a 'grating' of regions of high and low net intensity; if there is some intensity-dependence of absorption or refraction in the medium in which the interference

takes place then a real absorption or phase ' grating ' is created in the medium. If a third beam of the same frequency scatters off this grating, then the resultant (fourth) scattered beam is the DFWM signal (it will also be of the same frequency). The classic configuration for phase conjugation by DFWM involves counterpropagating forward and backward ' pump ' beams and a ' probe ' beam incident on the medium and in this case two separate ' gratings ' are formed, one between front pump and probe off which the back pump is scattered to produce a ' signal ' beam, and the other between back pump and probe off which the front pump scatters to produce a ' signal ' beam (see fig. 6.4.1). In the special case that the pump waves are uniform plane waves and the non-linear effect created in the medium is entirely local, i.e. not subject to diffusion of any kind, then each ' grating ' is actually a volume hologram of the ' object ' which gives rise to the probe beam and true phase conjugation will occur with the signal beam as the phase conjugate. The phase conjugate can then be thought of to arise from the fact that the reconstruction beam (e.g. the backward pump in the case of the grating which arises from front pump/probe interference) is actually the (spatial) complex conjugate of the construction

Fig. 6.4.1



Schematic representation of degenerate four-wave mixing showing the ' gratings '. (*a*) The signal, *s*, resulting from the scattering of backward pump wave, *b*. (*b*) The signal, *s*, resulting from the scattering of forward pump wave, *f*. *p* is the probe beam. (After Steel *et al.* 1979.)

beam (in this example the front pump). In these ideal conditions the two contributions to the signal beam add equally in the final result.

Some authors like to reserve the term DFWM for the specific case where the non-linearity is entirely 'local' and the non-linear response is exactly proportional to intensity. In this case the non-linearity can be exactly described by the degenerate four-wave third-order susceptibility $\chi^{(3)}(\omega:\omega,-\omega,\omega)$. In the case of active non-linear processes in semiconductors this is only ever approximately true, particularly due to the influence of diffusion, but the term DFWM is often retained for convenience, even when the intensity-dependence of the signal wave does not follow the power laws expected for $\chi^{(3)}$ effects (see, for example, Watkins *et al.* 1981).

Another configuration of DFWM is the so-called forward DFWM in which there is no back pump beam ; the signal beam is now generated by scattering of the front pump beam off the grating created by the front pump/probe beam interference (see, for example Hopf *et al.* 1981). This interaction is not in principle exactly phase-matched, but in practice with short semiconductor samples and narrow angles between pump and probe this can be effectively phase matched and strong forward DFWM seen. This beam is also a phase conjugate if the pump beam is uniform and diffusion is negligible, only now it will travel in a forward direction ; such a beam can be made into a backward propagating conjugate by the use of ordinary mirrors and beamsplitters. The grating formed in the medium can in principle be either an absorption or phase grating but in the case of nearly all of the experimental work to be discussed here (with the exception only of Smirl *et al.* 1980) the grating is predominantly a phase grating (i.e. due to refractive effects). It will be clear that the DFWM experiments discussed here are very closely related to the laser-induced gratings discussed in § 4. Phase conjugation by DFWM is a special case of these grating phenomena, being restricted to the Bragg regime (as opposed to Raman–Nath scattering) with all beams at the same frequency.

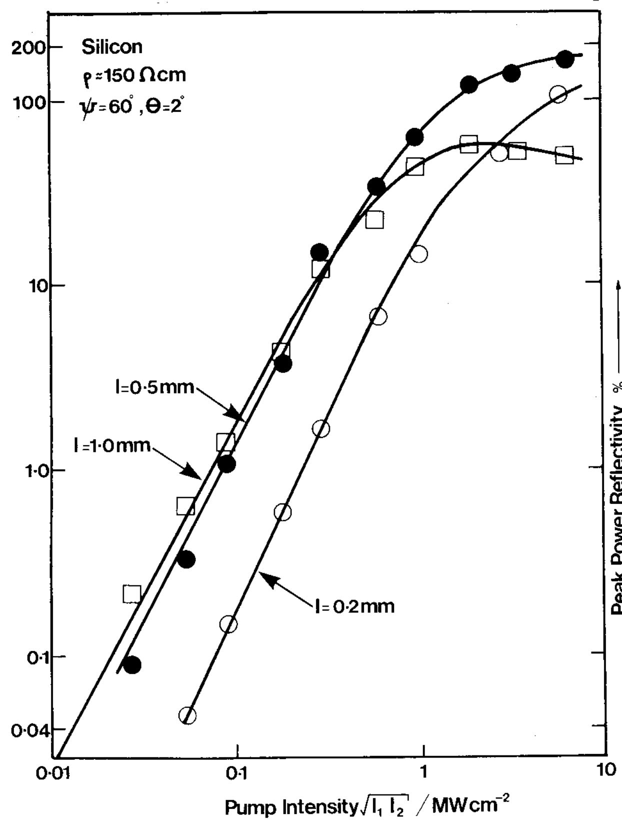
The first semiconductor in which real-time holography was investigated was Si by Woerdmann (1970 b) whose work has also been discussed in § 4.1. The configuration he used for his experiments (which employed a Q-switched Nd : YAG laser) was exactly that of backward DFWM (except that the back pump was delayed) although the phase conjugate properties of the backward signal were not investigated. The grating created in these experiments was a phase grating due to the dispersion of the free carriers created by one-photon absorption. He was able to demonstrate image reconstruction and its progressive degradation with increasing angle between the 'probe' and 'front pump' beams and also the decay of the hologram in time by delaying the back pump beam. The degradation in angle was attributed to 'washing out' of the grating in the material by diffusion ; this becomes stronger as the angle between pump and probe increases because the grating spatial period becomes shorter. The time decay (~ 25 ns) was also attributed to diffusion and found consistent with other measured diffusion coefficients. In his experiments there was probably no grating formed by the interference of back pump and probe because the back pump was delayed in time.

A similar process was subsequently investigated by Jain and Klein (1979) and Jain *et al.* (1979) who examined the magnitude of the backward DFWM signal both experimentally and on the basis of a simple model, using the known

form of free carrier dispersion and assuming ambipolar diffusion. The agreement between experiment and theory (Jain and Klein 1979) was good, with measured effective $\chi^{(3)} \sim 8 \times 10^{-8}$ e.s.u. for their conditions. Using narrow angles between front pump and probe ($\sim 2°$) and good temporal overlap between front and back pump pulses, they were able to obtain amplified reflectivity (150%) (Jain *et al.* 1979) in the signal beam for $\sim 6{\cdot}8$ MW/cm² pump intensity. Results from these experiments are shown in fig. 6.4.2. At low intensities these results show the rise is reflectivity proportional to the product of front and back pump intensities as would be expected for a $\chi^{(3)}$ non-linearity (see, for example, Yariv 1978). They ascribe the fall-off in reflectivity at high intensities to induced free carrier absorption.

In their calculations, Jain and Klein (1979) discount the grating which would be formed by interference of back pump and probe because diffusion would effectively destroy it. The spatial periods for the back pump/probe

Fig. 6.4.2



Log–log plots of DFWM signal power 'reflectivity' versus the mean pump intensity $\sqrt{(I_1 I_2)}$ (I_1 and I_2 are the front and back pump intensities respectively) for various thicknesses, l, of Si sample. The angle of incidence was 60° to reduce surface reflections and the angle between pump and probe was 2°. (After Jain *et al.* 1979.)

grating would be $\sim 0{\cdot}15\ \mu m$ compared to $61\ \mu m$ for the front pump/probe grating. Both Woerdmann (1971) and Jain and Klein (1979) pointed out that narrow-gap semiconductors would be particularly favourable for this kind of process at longer wavelengths becuase of the increased free carrier dispersion at longer wavelengths and the smaller effective masses found in direct narrow-gap semiconductors such as InSb and $Hg_{1-x}Cd_xTe$.

Si has subsequently been examined by Hopf *et al.* (1981) to assess the quality of phase conjugation in semiconductors; previous studies had concentrated only on the magnitude of the DFWM signal. Hopf *et al.* investigate primarily the forward DFWM signal although they also observed backward DFWM. They used a variety of interferometric techniques to assess the phase conjugate quality and deduced that at higher probe to signal conversion efficiencies (e.g. $\sim 50\%$) there can be $\sim \lambda/2$ distortion in the conjugate wave due to self-defocusing which may be unacceptable in some applications. They also observed a higher order scattered beam (produced by scattering the probe (rather than the pump) off the (front) pump/probe grating) which showed 'phase doubling' rather than phase conjugation. There is some discrepancy between the results of Hopf *et al.* (1981) and those of other workers in that Hopf *et al.* find their data inconsistent with the diffusion coefficients of other workers, requiring instead a value some twenty times larger.

Ge has been investigated at $1{\cdot}06\ \mu m$ by Smirl *et al.* (1980) using 30 ps pulses from a mode-locked Nd : YAG laser. In contrast to the Si work where the absorption at $1{\cdot}06\ \mu m$ is by the weak indirect process, $1{\cdot}06\ \mu m$ radiation in Ge leads to very strong direct absorption in thin ($\sim 6\ \mu m$) Ge wafers. With pump fluences of $\sim 10^{-2}\ J/cm^2$, Smirl *et al.* demonstrated forward DFWM which they claim is due to the absorption 'grating'. They demonstrated by scattering with a second pump pulse that the grating thus created persists with a time constant ~ 250 ps.

Ge has also been investigated using a CO_2 laser at $10{\cdot}6\ \mu m$. The first observation by Bergmann *et al.* (1978) utilized the intrinsic 'passive' $\chi^{(3)}$ and is beyond the general scope of this review. Subsequently Watkins *et al.* (1981), using intensities in the range $100\ kW/cm^2$ to $100\ MW/cm^2$ in both n and p-type samples, observed a very steeply rising DFWM signal associated with a multiple photon valence-conduction band absorption process. This latter phenomenon led to observed 'reflectivities' $> 100\%$, with a pump intensity-dependence of reflectivity $\sim I^{11}$ (as opposed to I^2 for the passive $\chi^{(3)}$ effect also observed by these authors). This intensity-dependence is consistent with the observed $I^{5{\cdot}5}$ dependence for plasma creation by the multiple photon excitation process; the DFWM is thought to arise from the free carrier plasma refraction grating created by this process. The multiple photon excitation has itself been studied by Yuen *et al.* (1980) (see § 3). Reflectivities as high as 800% were seen in p-type Ge. For completeness, we should mention also the work of Depatie and Haueisen (1980) who observed phase conjugation at $4\ \mu m$ in Ge, although again the 'passive' $\chi^{(3)}$ was used.

DFWM has also been studied in the narrow direct-gap semiconductors InSb and $Hg_{1-x}Cd_xT_e$. These systems have provided the lowest power DFWM in semiconductors. The only c.w. observation of DFWM in semiconductors so far reported is in InSb cooled to 5 K (Miller, Harrison *et al.* 1980), using a c.w. laser near the band-gap energy and the backward DFWM configuration.

The source of the non-linearity is thought to be non-linear refraction by band-gap resonant saturation (see § 5.1) giving a phase grating associated with the carrier density grating created by absorption; pumping with ~ 12 mW front pump and ~ 5 mW back pump, ' reflectivities ' $\sim 1\%$ were observed in a 7·5 mm sample with focused spot sizes ~ 200 μm (i.e. intensities < 100 W/cm^2). Using a TEA CO_2 laser at 10·6 μm with intensities in the range 10–500 kW/cm^2. Jain and Steel (1980) observed backward DFWM in $Hg_{1-x}Cd_xT_e$ at room temperature with reflectivities $\sim 10\%$ at 160 kW/cm^2. This DFWM was ascribed to the phase grating due to free-carrier plasma refraction from the carrier density grating due to absorption; these experiments were also performed near the band-gap energy. The fall-off in ' reflectivity ' at intensities above ~ 160 kW/cm^2 was attributed to strong intervalence band absorption by optically-created holes. DFWM has also been reported by Khan *et al.* (1980) in $Hg_{1-x}Cd_xTe$ at 12, 77 and 295 K also using a TEA CO_2 laser with ' reflectivities ' up to 9% observed. In this case, however, the photon energy was apparently significantly below the band-gap (different x-values were used from those of Jain and Steel (1980)) and the non-linearity used was apparently the ' passive ' $\chi^{(3)}$ associated with conduction band non-parabolicity which is beyond the scope of this review.

Phase conjugation, including a demonstration of aberration correction, has also been seen in ZnSe (Borshch *et al.* 1980). Although the configuration used was essentially similar to DFWM, the authors called their observation more precisely degenerate six-photon mixing as the absorption mechanism for the generation of the free carrier grating is two-photon absorption of the exciting ruby laser. The scattering off the phase grating created by this absorption gives the final signal in the same direction as for DFWM. Using pump intensities ~ 50 MW/cm^2 they were able to observe $\sim 200\%$ ' reflection '. They also demonstrated that to obtain the highest ' reflections ', it was necessary to deliberately misalign the front and back pump beams slightly. They were able to explain this as a consequence of ' self-action ' effects (i.e. self-focusing/defocusing) induced by the laser beams.

In summary, semiconductors offer a variety of opportunities for phase conjugation and/or DFWM. The systems described here are all far from any practical device for any of the applications of phase conjugation (see, for example, Yariv 1978, Yariv *et al.* 1979) in, for example, aberration correction or dispersion correction. The observations of DFWM gain (i.e. $> 100\%$ ' reflectivity ') offer some promise for ' optical transistor ' devices although the detail of the processes which limit the reflectivity are in general not well understood and c.w. DFWM ' gain ' has yet to be reported for semiconductor systems.

6.5. *Optical bistability and the optical transistor*

Optical bistability (OB) is, as it name implies, the existence of two stable states for one set of optical input conditions. The concept has been in existence for some 12 years since the suggestion of Szöke *et al.* (1969) but was first demonstrated by Gibbs *et al.* (1976) using Na vapour as the active medium inside a Fabry–Perot cavity. The first type of OB to be proposed was the so-called ' absorptive ' OB in which a saturable absorber is placed inside a Fabry–Perot cavity tuned on-resonance, the idea being that as the field inside the cavity increased, the absorber would ' bleach ', thus further increasing the

field inside the cavity by increasing the finesse (or 'Q-factor'). Ideally this process would be self-sustaining leading to a 'switched-on' action at one intensity with a 'switching-off' only occurring at lower intensities, thus creating a bistable region between the switch-on and switch-off intensities. However, this turns out to be difficult to achieve because the requirement on the degree of bleaching is rather severe (i.e. ratio of >8 between low and high intensity absorption limits (see, for example, Gibbs *et al.* 1979 a)). In practice, it is much easier to observe OB through intensity dependence of the refractive index as was first observed by Gibbs *et al.* (1976), usually called 'dispersive' (or 'refractive') OB; the most common method is to use the non-linear medium inside a Fabry–Perot, this time tuned somewhat off-resonance. Then as the incident intensity is increased the cavity is pulled towards resonance by the intensity-dependent refractive index changing the optical length. However, as the cavity is pulled towards resonance a larger fraction of the incident light gets into the cavity, thus enhancing the change in optical length. This process also can be self-sustaining, leading to a switching action with 'switch-on' and 'switch-off' intensities defining the boundaries of the bistable region. A model, with a simple graphical interpretation of the bistable region, was proposed by Felber and Marburger (1976) almost simultaneously with the observation of dispersive OB by Gibbs *et al.* (1976) (who also presented a simple model).

A large amount of theoretical activity was stimulated by the work of Bonifacio and Lugiato (1976) who developed an analytical technique for handling the coupled Maxwell–Bloch equations. This work is generally concerned with the rather specific conditions obtainable with two-level atomic systems and will not concern us further here. OB has been the subject of a number of short reviews (Gibbs *et al.* 1979 b, Collins and Wasmundt 1980, Farina *et al.* 1980, Gibbs *et al.* 1980, Smith 1980). For collections of recent papers on the subject see Bowden *et al.* (1981) and Smith (1981).

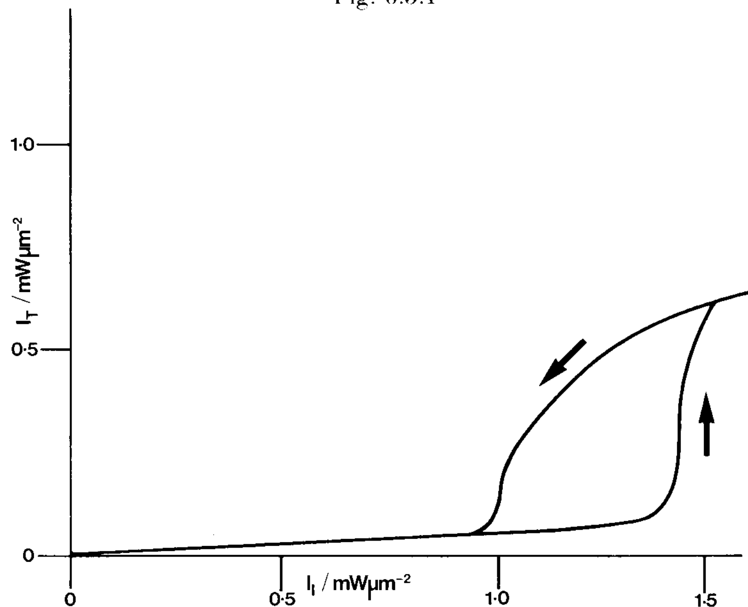
The subject of OB has grown very rapidly since 1976, stimulated both by its fundamental theoretical aspects and the opportunity it offers for a new class of all-optical devices of logical operations in a fashion loosely analogous to the electronic transistor. Despite this interest, there have been comparitively few observations of 'intrinsic' OB, i.e. truly all-optical devices with no electronic components. Of those which have been demonstrated, those based on near-band-gap non-linear refractive effects in semiconductors (Gibbs *et al.* 1979 d, 1981, Miller, Smith and Johnston 1979, Miller, Smith and Seaton 1981 a, b), arguably currently offer the best prospects for practical devices especially for low-power logic operations because they offer small size and potentially fast operation at low switching energies combined with some prospect that they might be usable in integrated optical circuits although presently in both cases the materials have to be cooled to liquid cryogen temperatures. Optical bistability has also been reported at room temperature in a system using two tellurium crystals (Staupendahl and Schindler 1980) and Jain and Steel (1980) mentioned the observation of some non-linear Fabry–Perot effects in $Hg_{1-x}Cd_xTe$ in their room temperature experiments on degenerate four-wave mixing (see § 6.4).

Gibbs *et al.* (1979 d) prepared a GaAlAs–GaAs–GaAlAs 'sandwich' structure, of thicknesses 0·21–4·1–0·21 μm, by molecular beam epitaxy on a GaAs

substrate and etched away a 1–2 mm diameter hole leaving the thin GaAs layer. This layer was coated for 90% reflectivity on each side, and the resulting cavity could be tuned by moving the focused spot across the crystal which was slightly non-parallel after the etching. The laser spot was $\sim 10\ \mu m$ diameter and OB was observed at intensities ~ 50–$100\ kW/cm^2$ (i.e. ~ 100 mW laser power) at 15–K (see fig. 6.5.1). The device could switch in ~ 40 ns, corresponding approximately to the excitonic lifetime but the turn-on time could be greatly reduced by using a short pulse to switch on. Working at a ' holding ' intensity within the bistable region the device turned on in < 1 ns using an additional 200 ps 5900 Å pulse with an effective absorbed switching energy of ~ 24 pJ in the active region. They predicted a limiting switch-on time of < 1 ps, based on the measurements of Shank *et al.* (1979) (discussed in § 3), and limiting switching energies of ~ 1 fJ (10^{-15} J) on the assumption that a $(0{\cdot}2\ \mu m)^3$ device could be made to switch (i.e. one cubic wavelength inside the material). The proposed mechanism for the non-linear refraction in this dispersive bistability is associated with the saturation of excitonic absorption (discussed in § 5). At the intensities used for bistability the excitonic resonance is almost totally saturated.

In confirmation of this interpretation, this relatively fast OB could be observed only up to ~ 120 K (where kT is approximately twice the exciton binding energy). Thereafter a slower thermal effect took over as the samples were not heat-sunk. The existence of bistability due to excitonic saturation has also been considered theoretically by Goll and Haken (1980).

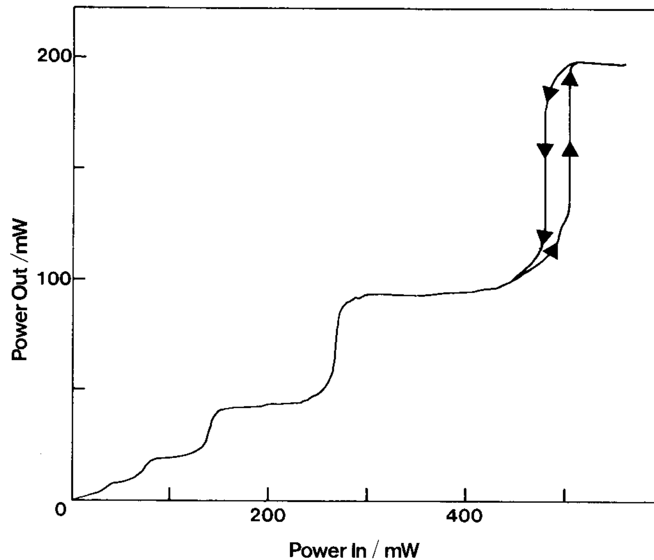
Fig. 6.5.1



Optical bistability in GaAs at 15 K at 819·9 nm laser wavelength (Gibbs *et al.* 1979 d) seen as a plot of transmitted intensity (I_T) against incident intensity (I_I).

Simultaneously with the GaAs work (the papers of Gibbs *et al.* (1979 d) and Miller, Smith and Johnston (1979) were received by the same journal on the same day !). Miller, Smith and Johnston (1979) demonstrated non-linear Fabry–Perot action in InSb at 5 K using a 500 μm thick InSb sample, polished plane-parallel and using only the natural reflectivity of the crystal faces (36%) to form the cavity mirrors. With this simple system, they obtained five successive non-linear orders of the Fabry–Perot cavity (i.e. 5 $\lambda/2$ change in effective optical length) with increasing intensity, with clear bistability in the fifth non-linear order, using spot sizes ~ 200 μm diameter and laser powers up to 500 mW (fig. 6.5.2). Subsequently they were able to demonstrate

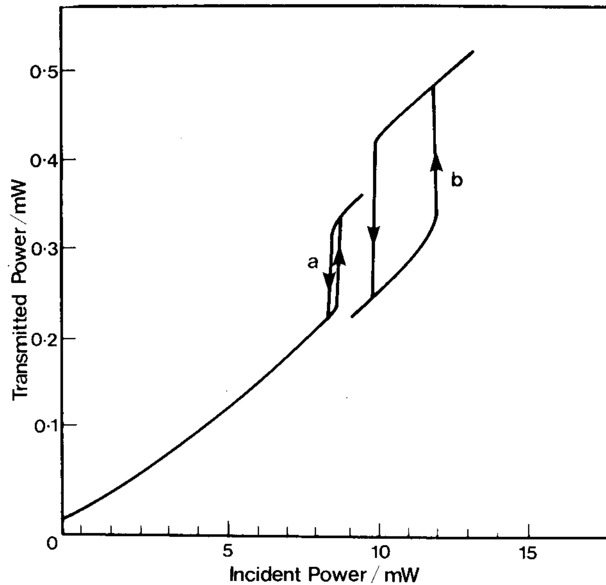
Fig. 6.5.2



Optical bistability in InSb at 5 K at 5·277 μm laser wavelength (Miller, Smith and Johnston 1979) seen as a plot of output power against input power. Five non-linear ' orders ' of Fabry–Perot action are seen with clear bistability in the fifth order. Spot size is ~ 180 μm diameter.

bistability in two successive non-linear orders (third and fourth) in both transmission and reflection at powers ~ 100–200 mW in similar samples (Miller, Smith and Seaton 1981) and in a 130 μm crystal at 77 K coated to 70% reflectivity, bistability was seen in the first non-linear order at ~ 8 mW (~ 80 W/cm^2) (fig. 6.5.3). The measured switching times were detector-system limited but at $\lesssim 500$ ns consistent with interband recombination as the longest time constant in these devices. The proposed mechanism for the non-linear refraction causing this dispersive OB is the band-gap resonant effect discussed in § 4. Miller and Smith (1979) extended the observations of non-linear Fabry–Perot action in InSb to demonstrate true two-beam ' optical transistor ' action. They focused two beams, a main beam and a weak beam, at slightly different

Fig. 6.5.3

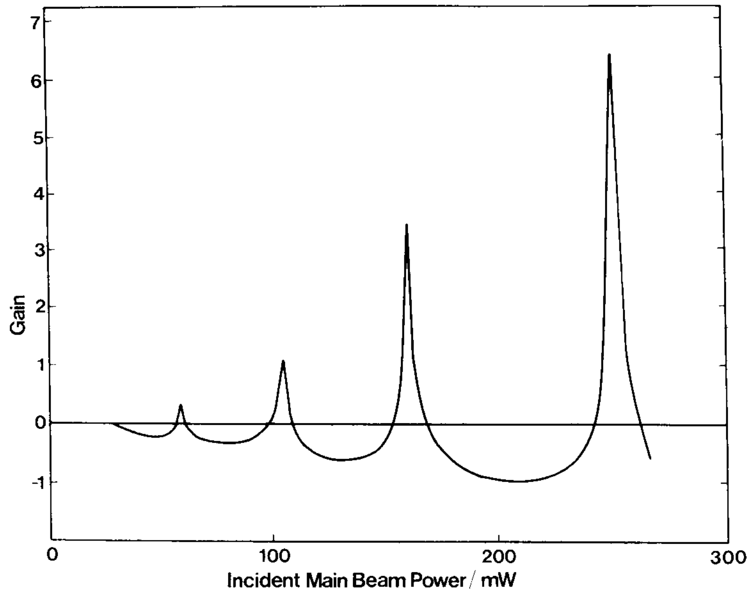


Transmitted power plotted against incident power for c.w. CO laser beam (wave-number 1827 cm^{-1}, spot size $\sim 150\ \mu m$) passing through a polished polycrystalline InSb slice (5×5 mm $\times$ 130 μm thick) coated to $\sim 70\%$ reflectivity on both faces, held at ~ 77 K. Onset of bistability is seen at ~ 8 mW (trace (a)) with clear bistability at slightly higher power with different cavity detuning (trace (b)). (After Miller, Smith and Seaton 1981 a.)

angles, coincidently on the sample. They made small changes in the incident weak beam power and observed the change in the transmitted main beam power, ratioing the latter to the former to define a ' gain '. The results of this experiment taken under otherwise identical conditions to the results in fig. 6.5.2 are shown in fig. 6.5.4. At the powers corresponding to the sharp rises in the results in fig. 6.5.2, the gain reaches a peak, ultimately $\gg 1$, showing real optical signal gain of up to 10. This measurement of two-beam gain is physically different from the one-beam differential gain which would result from the derivative of the results in fig. 6.5.2, as in the presence of two beams and a cavity degenerate four-wave mixing is possible (see, for example, Miller, Smith and Seaton 1981 b) and has indeed been observed in InSb at 5 K at similar powers (Miller, Harrison *et al.* 1981). Because the device operates mainly by ' transfer of phase thickness ' it was termed the ' transphasor ' by analogy with the transistor.

Gibbs *et al.* (1979 d) and Miller, Smith and Seaton (1981 b) have considered various ways in which device switching times and energies may be altered, including alteration of switch-off times by doping or diffusion. Miller (1981 a) has analysed the design of non-linear Fabry–Perot cavities in the presence of linear absorption (as is present in practice in the InSb devices) and deduced the effective material figure of merit for minimum switching intensity is $n_2/\lambda\alpha$ where α is the linear absorption coefficient. The ratio of n_2/α is due to the fact

Fig. 6.5.4



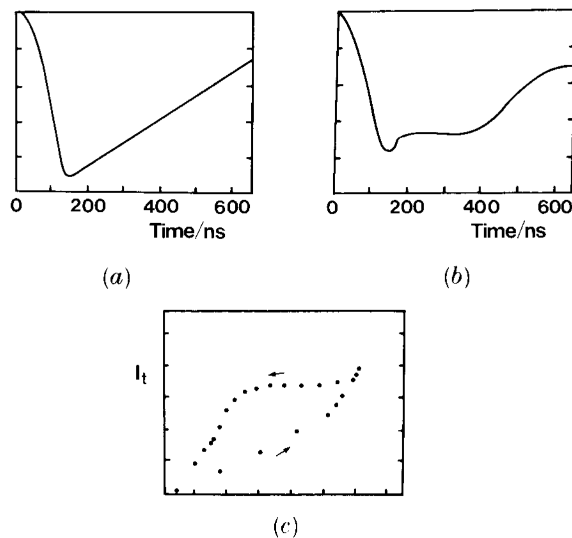
Differential amplification between two laser beams (Miller and Smith 1979) under similar conditions to the bistability results in InSb at 5 K (Miller, Smith and Johnston 1979). The peaks in gain correspond to the rising ' steps ' in the non-linear Fabry–Perot action.

that for a more highly absorbing material, the cavity has to be made shorter to preserve the finesse and consequently n_2 has to be larger. This parameter $n_2/\lambda\alpha$ is physically identical (except for fundamental constants) with σ, the refractive index change for one excited system per unit volume (see, for example, Miller, Smith and Seaton 1981 b). On the basis of this analysis these authors deduced limiting switching energies also of $\sim 10^{-15}$ J, with 1 pJ being more readily feasible. It is important to emphasize that the switching energy is not dependent (at least in these simple models) on the switch-on speed desired; faster switch-on requires higher power, but this may be achieved by using a shorter pulse of the same energy. The limits on switch-on time are liable to depend on the speed of intraband processes.

Whether it is possible to scale these devices to faster switch-off and/or higher temperature operation depends to some extent on the ' robustness ' of the non-linearity. Excitons favour pure, low-temperature materials, and it may not be possible to dope heavily or increase the temperature. The non-linearity in InSb has, however, survived to 77 K (already ~ 20 times the theoretical exciton binding energy) and on the model of Miller, Seaton *et al.* (1981) should exist at room temperature, weakening only in proportion to temperature. The observations of OB at 77 K by Miller, Smith and Seaton (1981 a, b) were also made on relatively impure polycrystalline material. However, it seems likely that the cavity field build-up times can be kept very short in semiconductors with $\lesssim 1{\cdot}2$ ps calculated for GaAs and $\lesssim 20$ ps for InSb in the devices already demonstrated, so cavity times should not be a problem.

Neither the GaAs nor InSb devices have been optimized and much work remains to be done to develop a practical and useful bistable device. Nevertheless, the observations in these two materials offer considerable promise, especially as they result from apparently different physical mechanisms thus offering a variety of ways of achieving the necessary non-linearity and suggesting the existence of related non-linearities in other materials. Miller, Smith and Seaton (1981 a, b) have considered the scaling of the interband saturation contribution to the non-linearity and conclude that on this contribution alone similar effects should be observable for similar energies in other materials, because although the interband saturation non-linearity decreases as $1/E_g^2$ (E_g is the band-gap energy) in going to wider gap materials, this is exactly compensated by the possibility of tighter focusing (by a factor $1/\lambda^2$ in intensity for a given power) due to weaker diffraction at shorter wavelengths thus making the effective material figure of merit $n_2/\lambda^3\alpha$ or σ/λ^2. The scaling of other effects (e.g. excitonic and band-gap renormalization) is not yet clear, although it is likely they will be of a different form and hence it is highly unlikely that the various effects will always conspire to cancel one another. Therefore despite the current limited understanding of the band-gap-resonant non-linearities, the future is very promising for device applications. Recently, the possibility of OB utilizing yet another non-linearity near the band-gap of wider band direct-gap semiconductors, namely that due to the creation of excitonic molecules in for example CdS or CuCl (Koch and Haug 1981), has also been suggested.

Fig. 6.5.5



Typical modulation in Te of self-controlled 10·6 μm optical–optical modulator with external 5·3 μm pump. (*a*) An example of well-defined pulse shaping (saw tooth pulse). (*b*) A modulated pulse offering the behaviour of optical bistability and limiter action. The bistable operating mode is illustrated by plotting the transmitted intensity (I_t) versus the incident intensity (I_0) (*c*). (After Staupendahl and Schindler 1980.)

A different approach to optical bistability has been demonstrated using the semiconductor tellurium by Staupendahl and Schindler (1980). They used two crystals of tellurium, both pumped by the same Q-switched CO_2 laser pulse ; one crystal was arranged to generate second harmonic radiation through the conventional passive $\chi^{(2)}$ of tellurium whereas the other crystal was used to absorb by two-or three-photon absorption various combinations of 10·6 μm photons (directly from the CO_2 laser) and 5·3 μm photons (produced by the second harmonic generation in the first crystal). The band-gap of tellurium ($\sim$0·35 eV) is such that three 10·6 μm photons or one 5·3 μm + one 10·6 μm photon or two 5·3 μm photons can bridge the gap. The second crystal was arranged as a Fabry–Perot cavity, by polishing the crystal plane-parallel and utilizing the natural reflectivity of the crystal faces. The creation of free carriers in this crystal by multiphoton absorption was presumed to alter the refractive index through the usual Drude-model free carrier refraction and hence tune the cavity. With both the 5·3 μm and 10·6 μm beams coincident on the crystal they were able to demonstrate a variety of modulation effects (fig. 6.5.5 (*a*), (*b*)) at intensities $\sim$1–10 MW/cm^2 and also showed clear hysteresis in the output which they ascribed to optical bistability (fig. 6.5.5 (*c*)). Although this system is a compound one it still represents an all-optical method of optical modulation, utilizing at least three different non-linear optical processes simultaneously in one material.

§ 7. Conclusions

The series of non-linear optical effects presented in this review covers a wide range of magnitudes both in timescales and size of non-linearity. At one extreme we have experiments which are giving new information on the energy loss mechanisms of excited carriers in semiconductors on timescales of picoseconds or less, and at the other we have continuous wave effects which can be excited with milliwatt powers. The latter effects in which a refractive index change of the order of unity can be readily achieved heralds the beginning of ‘ milliwatt power non-linear optics ’. From a device point of view switching, amplifying and logic elements are emerging with a similar range of operating parameters. If it proves possible to combine milliwatt power levels with near-picosecond operating time constants the possibilities for all-optical electronics will be considerable.

Acknowledgments

This article originated with an invitation to one of us (S.D.S.) during a period as Visiting Professor at the Optical Sciences Centre, University of Arizona. The assistance of the United States Air Force and the provision of a Senior Fellowship under the NATO Senior Scientists Programme, and the hospitality extended by Dr. Peter Franken and his colleagues is acknowledged. Likewise, Alan Miller benefitted greatly from many stimulating discussions with colleagues during his period as Visiting Assistant Professor at North Texas State University and acknowledges the support of the Office of Naval Research. We also acknowledge numerous conversations with workers in the field, in particular those who attended the Conference on Excited States and Multiresonant Non-linear Optical Processes in Solids, arranged by the European Physical Society Quantum Electronics Division, at Aussois in March 1981.

REFERENCES

AGOSTINELLI, J., MOUROU, G., and GABEL, C. W., 1979, *Appl. Phys. Lett.*, **35,** 731.

ALCOCK, A. J., and WALKER, A. C., 1974, *Appl. Phys. Lett.*, **25,** 299.

ALCOCK, A. J., CORKUM, P. B., and JAMES, D. J., 1975, *Appl. Phys. Lett.*, **27,** 680.

ANTONETTI, A., MALLEY, M. M., MOUROU, G., and ORSZAG, A., 1977, *Optics Commun.*, **23,** 435.

ANTONETTI, A., MIGUS, A., MALLEY, M. M., and MOUROU, G., 1977, *Optics Commun.*, **21,** 211.

ASHKIN, A., TELL, B., and DZIEDZIC, J. M., 1967, *I.E.E.E. Jl quant Electron.*, **3,** 400.

AUSTON, D. H., 1975, *Appl. Phys. Lett.*, **26,** 101 ; 1976, *High Speed Optical Techniques*, SPIE, Vol. 94 ; 1979, *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 73.

AUSTON, D. H., JOHNSON, A. M., LEFUR, P., SHANK, C. V., IPPEN, E. P., and TESCHKE, O., 1976, *Proc. IXth Int. Conf. Quant. Electron.*, Amsterdam, in *Optics Commun.*, **18,** 84.

AUSTON, D. H., JOHNSON, A. M., and SMITH, P. R., 1980, Picosecond phenomena II, *Springer Series in Chemical Physics*, Vol. 14, edited by R. Hockstrasser, W. Kaiser and C, V. Shank (New York : Springer-Verlag), p. 71.

AUSTON, D. H., LAVALLARD, P., SOL, N., and KAPLAN, D., 1980, *Appl. Phys. Lett.*, **36,** 66.

AUSTON, D. H., MCAFEE, S., SHANK, C. V., IPPEN, E. P., and TESCHKE, O., 1978, *Solid-St. Electron.*, **21,** 147.

AUSTON, D. H., and SHANK, C. V., 1974, *Phys. Rev. Lett.*, **32,** 1120.

AUSTON, D. H., SHANK, C. V., and LEFUR, P., 1975, *Phys. Rev. Lett.*, **35,** 1022.

BALTRAMEYUNAS, R., JARASIUNAS, K., VAITKUS, J., and VELECKAS, D., 1976, *Optics Commun.*, **18,** 47.

BALTRAMEYUNAS, R., VAITKUS, YU., and YARASHYUNAS, K., 1976, *Soviet Phys. Semicond.*, **10,** 572.

BASSANI, F., and PARRAVICINI, G. P., 1975, *Electronic States and Optical Transitions in Solids* (Oxford : Pergamon Press).

BAUER, G., 1974, *Springer Tracts in Modern Physics*, Vol. 74 (Berlin : Springer-Verlag).

BERGMANN, E. E., BIGIO, I. J., FELDMAN, B. J., and FISHER, R. A., 1978, *Optics Lett.*, **3,** 82.

BESSEY, J. S., BOSACCHI, B., VAN DRIEL, H. M., and SMIRL, A. L., 1978, *Phys. Rev.* B, **17,** 2782.

BIRNBAUM, M., 1965, *J. appl. Phys.*, **36,** 657.

BIRNBAUM, M., and STOCKER, T. L., 1966, *Br. J. appl. Phys.*, **17,** 461 ; 1968, *J. appl. Phys.*, **39,** 6032.

BISHOP, P. J., GIBSON, A. F., and KIMMITT, M. F., 1976, *J. Phys.* D, **9,** L101.

BLAKEMORE, J., 1962, *Semiconductor Statistics* (Oxford : Pergamon Press).

BLINOV, L. M., VAVILOV, V. S., and GALKIN, G. N., 1967 a, *Soviet Phys. solid St.* **9,** 666 ; 1967 b, *Soviet Phys. Semicond.*, **1,** 1124.

BLOEMBERGEN, N., 1965, *Non-Linear Optics* (New York : W. A. Benjamin Inc.).

BONCH-BRUEVICH, A. M., KOVALEV, V. P., ROMANOV, G. S., IMAS, YA. A., and LIBENSON, M. N., 1968, *Soviet Phys. tech. Phys.*, **13,** 507.

BONIFACIO, R., and LUGIATO, L. A., 1976, *Optics Commum.*, **19,** 172.

BORSHCH, A. A., BRODIN, M. S., OVCHAR, V. V., ODULOV, S. G., and SOSKIN, M. S., 1973, *Soviet Phys. JETP Lett.*, **18,** 397.

BORSHCH, A. A., BRODIN, M., VOLKOV, V., and KIKHTAREV, N., 1980, *Optics Commun.*, **35,** 297.

BOSACCHI, B., LEUNG, C. Y., and SCULLY, M. O., 1978 a, *Optics Commun.*, **27,** 475 ; 1978 b, Picosecond phenomena, *Springer Series in Chemical Physics*, Vol. 4, edited by C. V. Shank, E. P. Ippen and S. L. Shapiro (Berlin : Springer-Verlag), p. 244.

BOWDEN, C. M., CIFTAN, M., and ROBL, H. R. (editors), 1981, *Optical Bistability* (New York : Plenum Press).

BRINKMAN, W. F., and RICE, T. M., 1973, *Phys. Rev.* B, **7,** 1508.

BRUCKNER, V., DNEPROVSKII, V. S., FOKIN, V. S., KOSHCHUG, D. G., KRAEVSKII. M. V., OAK, S. M., SILINA, E. K., and ZHUKOV. E. A., 1976, Physics of semiconductors, *Proc. 13th Int. Conf.*, Rome, p. 845.

CARMICHAEL, C. H., and SIMPSON, G. N., 1964, *Nature, Lond.*, **202,** 787.

CASTAGNE, R., LAVAL, S., and LAVAL, R., 1976, *Electronics Lett.*, **12,** 438.

CASTAGNE, R., LAVAL, R., LAVAL, S., and MERRIAUX, A., 1978, Picosecond phenomena. *Springer Series in Chemical Physics*, Vol. 4, edited by C. V. Shank. E. P. Ippen and S. L. Shapiro (Berlin : Springer-Verlag). p. 182.

CHANG, T. Y., 1981, *Opt. Engng*, **20,** 2, 220.

CHEMLA, D. S., 1980, *Rep. Prog. Phys.*, **43,** no. 10, 1191.

COLLINS, S. A., and WASMUNDT, K. C., 1980, *Opt. Engng*, **19,** 478.

CONWELL, E. M., 1967, *Solid State Physics*, Suppl. 9, edited by F. Seitz, D. Turnbull and H. Ehrenreich (New York : Academic Press).

CORKUM, P. B., and ALCOCK, A. J., 1978, Picosecond phenomena, *Springer Series in Chemical Physics*, Vol. 4, edited by C. V. Shank, E. P. Ippen and S. L. Shapiro (Berlin : Springer-Verlag), p. 308.

CORKUM, P. B., ALCOCK, A. J., ROLLIN, D. F., and MORRISON, H. D., 1978, *Appl. Phys. Lett.*, **32,** 27.

DAPKUS, P. D., HOLONYAK, N., BURNHAM, R. D., and KEUNE, D. L., 1970 a, *Appl. Phys. Lett.*, **16,** 93 ; 1970 b, *J. appl. Phys.*, **41,** 4194.

DEFONZO, A. P., 1978, *NRL*, Washington, Report No. 8239.

DEFONZO, A. P., LEE, C. H., and MAK, P. S., 1979, *Appl. Phys. Lett.*, **35,** 575.

DENNIS, R. B., HIGGINS, N. A., MACKENZIE, H. A., SMITH, S. D., WANG, W. L., VOGE, D., and WHERRETT, B. S., 1981, *SPIE*, **236,** 428.

DENNIS, R. B., MACKENZIE, H. A., SMITH, S. D., WHERRETT, B. S., and VOGE, D., 1980, *J. phys. Soc. Japan*, Suppl. A, **49,** 605.

DEPATIE, D., and HAUEISEN, D., 1980, *Optics Lett.*, **5,** 252.

DIMMOCK, J. O., 1967, *Semiconductors and Semimetals*, Vol. 3, edited by R. K. Willardson and A. C. Beer (New York : Academic Press), p. 259.

EICHLER, H. J., 1977, *Optica Acta*, **24,** 631 ; 1978, *Festkorperprobleme*, **18,** 241.

EICHLER, H. J., HARTIG, CH., and KNOPF, J., 1978, *Phys. Stat. Sol. (a)*, **45,** 433.

ELCI, A., SCULLY, M. O., SMIRL, A. L., and MATTER, J. C., 1977, *Phys. Rev.* B, **16,** 191.

ELCI, A., SMIRL, A. L., LEUNG, C. Y., and SCULLY, M. O., 1978, *Solid-St. Electron.*, **21,** 151.

FAN, H. Y., 1956, *Rep. Prog. Phys.*, **19,** 107 ; 1967, *Semiconductors and Semimetals*, Vol. 3, edited by R. K. Willardson and A. C. Beer (New York : Academic Press), p. 406.

FARINA, J. D., NARDUCCI, L. M., YUAN, J. M., and LUGIATO, L. A., 1980, *Opt. Engng*, **19,** 469.

FELBER, F. S., and MARBURGER, J. H., 1976, *Appl. Phys. Lett.*, **28,** 731.

FELDMAN, B. J., and FIGUEIRA, J. F., 1974, *Appl. Phys. Lett.*, **25,** 301.

FIGUEIRA, J. F., CANTRELL, C. D., FORMAN, P. R., and RIAK, J. P., 1976, *Appl. Phys. Lett.*, **28,** 398.

FLYTZANIS, C., 1975, Theory of nonlinear optical susceptibilities, *Quantum Electronics*, Vol. 1A, edited by H. Rabin and C. L. Tang (New York : Academic), p. 9.

FORK, R. L., SHANK, C. V., GREENE, B. I., REINHART, F. K., and LOGAN, R. A., 1980, Picosecond phenomena II, *Springer Series in Chemical Physics*, Vol. 14, edited by R. Hochstrasser, W. Kaiser and C. V. Shank (Berlin : Springer-Verlag), p. 280.

FOSSUM, H. J., CHEN, W. S., and ANCKER-JOHNSON, B., 1973, *Phys. Rev.* B, **8,** 2857.

GALKIN, G. N., BLINOV, L. M., VAVILOV, V. S., and GOLOVASHKIN, A. G., 1968, *Soviet Phys. JETP Lett.*, **7,** 69.

GIBBS, H. M., MCCALL, S. L., and VENKATESAN, T. N. C., 1976, *Phys. Rev. Lett.*, **36,** 1135 ; 1979 b, *Optics News*, **5,** 6 ; 1980, *Opt. Engng*, **19,** 463 ; 1981, *Optical Bistability*, edited by C. M. Bowden, M. Ciftan and H. R. Robl (New York : Plenum Press), p. 109.

GIBBS, H. M., GOSSARD, A. C., MCCALL, S. L., PASSNER, A., WIEGMANN, W., and VENKATESAN, T.N.C., 1979 a, *Solid St. Commun.*, **30,** 271.
GIBBS, H. M., VENKATESAN, T. N. C., MCCALL, S. L., PASSNER, A., and WIEGMANN, W., 1979 c, *Appl. Phys. Lett.*, **34,** 511.
GIBBS, H. M., MCCALL, S. L., VENKATESAN, T. N. C., GOSSARD, A. C., PASSNER, A., and WIEGMANN, W., 1979 d, *Appl. Phys. Lett.*, **35,** 451.
GIBSON, A. F., KIMMITT, M. F., and NORRIS, B., 1974, *Appl. Phys. Lett.*, **24,** 306.
GIBSON, A. F., ROSITO, C. A., RAFFO, C. A., KIMMITT, M. F., 1972, *Appl. Phys. Lett.*, **21,** 356.
GOLL, J., and HAKEN, H., 1980, *Phys. Stat. Sol. (b)*, **101,** 489.
GOTO, T., and LANGER, D. W., 1971, *Phys. Rev. Lett.*, **27,** 1004.
GREENAWAY, D. L., and HARBEKE, G., 1968, *Optical Properties and Band Structure of Semiconductors* (Oxford : Pergamon Press).
HANES, L. K., and SEILER, D. G., 1980, *Optics Commun.*, **34,** 89.
HEDIN, L., and LUNDQVIST, S., 1969, *Solid State Physics*, Vol. 23, edited by F. Seitz, D. Turnbull and H. Ehrenreich (New York : Academic Press), p. 1.
HENSEL, J. C., PHILLIPS, T. G., and THOMAS, G. A., 1977, *Solid State Physics*, Vol. 32, edited by H. Ehrenreich, F. Seitz and D. Turnbull (New York : Academic Press), p. 87.
HILDEBRAND, O., GOEBEL, E. O., ROMANEK, K. M., WEBER, H., and MAHLER, G., 1978, *Phys. Rev.* B, **17,** 4775.
HOFFMAN, C. H., JARASIUNAS, K., GERRITSEN, H. J., and NURMIKKO, A. V., 1978, *Appl. Phys. Lett.*, **33,** 536.
HOLAH, G. D., DEMPSEY, J., MILLER, D. A. B., WHERRETT, B. S., and MILLER, A., 1979, *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 505.
HOPF, F. A., TOMITA, A., and LIEPMANN, T., 1981, *Optics Commun.*, **37,** 72.
IPPEN, E. P., and SHANK, C. V., 1977, *Topics in Applied Physics*, Vol. 18, edited by S. L. Shapiro (Berlin : Springer-Verlag), p. 83 ; 1978, Picosecond phenomena, *Springer Series in Chemical Physics*, Vol. 4, edited by C. V. Shank, E. P. Ippen and S. L. Shapiro (New York : Springer-Verlag), p. 103.
JAIN, R. K., and KLEIN, M. B., 1979, *Appl. Phys. Lett.*, **35,** 454.
JAIN, R. K., KLEIN, M. B., and LIND, R. C., 1979, *Optics Lett.*, **4,** 328.
JAIN, R. K., and STEEL, D. G., 1980, *Appl. Phys. Lett.*, **37,** 1.
JAMES, R. B., and SMITH, D. L., 1979, *Phys. Rev. Lett.*, **42,** 1495.
JAMES, R. B., and SMITH, D. L., 1980 a, *Phys. Rev.* B, **21,** 3502 ; 1980 b, *J. appl. Phys.*, **51,** 2836.
JAMISON, S. A., and NURMIKKO, A. V., 1978 a, *Appl. Phys. Lett.*, **33,** 182 ; 1978 b, *Ibid.*, **33,** 598 ; 1979, *Phys. Rev.* B, **19,** 5185.
JAMISON, S. A., NURMIKKO, A. V., and GERRITSEN, H. J., 1976, *Appl. Phys. Lett.*, **29,** 640.
JARASIUNAS, K., and GERRITSEN, H. J., 1978, *Appl. Phys. Lett.*, **33,** 190.
JARASIUNAS, K., HOFFMAN, C., GERRITSEN, H., and NURMIKKO, A. V., 1978, Picosecond phenomena, *Springer Series in Chemical Physics*, Vol. 4, edited by C. V. Shank, E. P. Ippen and S. L. Shapiro (Berlin : Springer-Verlag), p. 327.
JARASIUNAS, K., and VAITKUS, J., 1974, *Phys. Stat. Sol. (a)*, **23,** K19 ; 1977, *Ibid.*, **44,** 793
JAVAN, A., and KELLY, P. L., 1966, *I.E.E.E. Jl quant. Electron.*, **2,** 470.
JOHNSON, A. M., and AUSTON, D. H., 1975, *I.E.E.E. Jl quant. Electron.*, **11,** 283.
JOHNSON, A. M., AUSTON, D. H., SMITH, P. R., BEAN, J. C., HARBINSON, J. B., and KAPLAN, D., 1980, Picosecond phenomena II, *Springer Series in Chemical Physics*, Vol. 14, edited by R. Hochstrasser, W. Kaiser and C. V. Shank (Berlin : Springer-Verlag), p. 285.
JOHNSTON, A. M., PIDGEON, C. R., and DEMPSEY, J., 1980, *Phys. Rev.* B, **22,** 825.
KAISER, W., and LAUBEREAU, A., 1977, *Nonlinear Optics*, edited by P. G. Harper and B. S. Wherrett (London : Academic Press), p. 257.
KANE, E. O., 1957, *J. Phys. Chem. Solids*, **1,** 249 ; 1967, *Phys. Rev.*, **159,** 624.
KEILMAN, F., 1976, *I.E.E.E. Jl. quant. Electron.*, **12,** 592 ; 1977, *Appl. Phys.*, **14,** 29.

KEILMAN, F., and KUHL, J., 1978, *I.E.E.E. Jl quant. Electron.*, **14,** 203.
KENNEDY, C. J., MATTER, J. C., SMIRL, A. L., WEICHE, H., HOPF, F. A., PAPPU, S. V. and SCULLY, M. O., 1974, *Phys. Rev. Lett.*, **32,** 419.
KHAN, M. A., KRUSE, P. W., and READY, J. F., 1980, *Optics Lett.*, **5,** 261.
KNOX, R. S., 1963, *Theory of Excitons, Solid State Physics*, Suppl. 5, edited by F. Seitz and D. Turnbull (New York : Academic Press).
KOCH, S. W., and HAUG, H., 1981, *Phys. Rev. Lett.*, **46,** 450.
KOCH, S. W., SCHMITT-RINK, S., and HAUG, H., 1981 a, *Phys. Stat. Sol.* (*b*), **106,** 135 ; 1981 b, *Solid St. Commun.* (to be published).
KOMOLOV, V. L., YAROSHETSKII, I. D., and YASSIEVICH, I. N., 1977, *Soviet Phys. Semicond.*, **11,** 48.
KREMENITSKII, V., ODULOV, S., and SOSKIN, M., 1979, *Phys. Stat. Sol.* (*a*), **51,** K69.
LANDSBERG, P. T., and BEATTIE, A. R., 1960, *Proc. R. Soc.*, **258,** 486.
LATHAM, W. P., SMIRL, A. L., ELCI, A., and BESSEY, J. S., 1978, *Solid-St. Electron.*, **21,** 159.
LAVALLARD, P., BICHARD, R., and BENOIT A LA GUILLAUME, C., 1977, *Phys. Rev.* B, **16,** 2804.
LAWTON, R. A., 1976, *I.E.E.E. Trans. Instrum. Meas.*, **25,** 56.
LAWTON, R. A., and SCAVANNEC, A., 1975, *Electronics Lett.*, **11,** 74.
LEE, C. C., and FAN, H. Y., 1974, *Phys. Rev.* B, **9,** 3502.
LEE, C. H., 1977, *Appl. Phys. Lett.*, **30,** 84.
LEE, C. H., ANTONETTI, A., and MOUROU, G., 1977, *Optics Commun.*, **21,** 158.
LEE, C. H., MAK, P. S., DEFONZO, A. P., 1980, Picosecond phenomena II, *Springer Series in Chemical Physics*, Vol. 14, edited by R. Hochstrasser, W. Kaiser and C. V. Shank (Berlin : Springer-Verlag), p. 88.
LE FUR, P., and AUSTON, D. H., 1976, *Appl. Phys. Lett.*, **28,** 21.
LEHENY, R. F., and SHAH, J., 1978, *Solid-St. Electron.*, **21,** 167.
LEHENY, R. F., SHAH, J., FORK, R. L., SHANK, C. V., and MIGUS, A., 1979, *Solid St. Commun.*, **31,** 809.
LEONBERGER, F. J., and MOULTON, P. F., 1979, *Appl. Phys. Lett.*, **35,** 712.
LINDLE, J. R., MOSS, S. C., and SMIRL, A. L., 1979, *Phys. Rev.* B, **20,** 2401.
LIOUVILLE, J., 1838, *J. Math.*, **3,** 349.
MACKENZIE, H. A., DENNIS, R. B., SMITH, S. D., VOGE, D., and WANG, WEI-LI., 1981, *Proc. Int. Conf. Excited States and Multiresonant Non-linear Optical Processes in Solids*, CNRS, Aussois, p. 42.
MACKENZIE, H. A., DENNIS, R. B., VOGE. D., and SMITH, S. D., 1980, *Optics Commun.*, **34,** 205.
MAK, P. S., MATHUR, V. K., and LEE, C. H., 1980, *Optics Commun.*, **32,** 485.
MATTER, J. C., SMIRL, A. L., and SCULLY, M. O., 1976, *Appl. Phys. Lett.*, **28,** 507.
MERIAU, A., CASTAGNE, R., LAVAL, R., and LAVAL, S., 1977, *Electronics Lett.*, **13,** 245.
MICHEL, A. E., and NATHAN, M. I., 1965, *Appl. Phys. Lett.*, **6,** 101.
MILLER, A., 1979, *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 63.
MILLER, A., JOHNSTON, A. M., DEMPSEY, J., SMITH, J., PIDGEON, C. R., and HOLAH, G. D., 1979, *J. Phys.* C, **12,** 4839.
MILLER, A., PERRYMAN, G. P., and SMIRL, A. L., 1981 a, *Proc. Int. Conf. Excited States and Multiresonant Non-linear Optical Processes in Solids*, CNRS, Aussois, p. 2 ; 1981 b, *Optics Commun.*, **38,** 289.
MILLER, D. A. B., 1981 a, *I.E.E.E. Jl quant. Electron.*, **17,** 306 ; 1981 b, *Proc. R. Soc.* (to be published).
MILLER, D. A. B., HARRISON, R. G., JOHNSTON, A. M., SEATON, C. T., and SMITH, S. D., 1980, *Optics Commun.*, **32,** 478.
MILLER, D. A. B., MOZOLOWSKI, M. H., MILLER, A., and SMITH, S. D., 1978, *Optics Commun.*, **27,** 133.
MILLER, D. A. B., SEATON, C. T., PRISE, M. E., and SMITH, S. D., 1981, *Phys. Rev. Lett.*, **47,** 197.
MILLER, D. A. B., and SMITH, S. D., 1978, *Appl. Optics*, **17,** 3804 ; 1979, *Optics Commun.*, **31,** 101.
MILLER, D. A. B., SMITH, S. D., and JOHNSON, A., 1979, *Appl. Phys. Lett.*, **35,** 658.

MILLER, D. A. B., SMITH, S. D., and SEATON, C. T., 1981 a, *Optical Bistability*, edited by C. M. Bowden, M. Ciftan and H. R. Robb (New York : Plenum Press), p. 115 ; 1981 b, *I.E.E.E. Jl quant. Electron.*, **17,** 312.
MILLER, D. A. B., SMITH, S. D., and WHERRETT, B. S., 1980 *Optics Commun*, **35,** 221.
MOSS, S. C., LINDLE, J. R., MACKEY, H. J., and SMIRL, A. L., 1981, *Appl. Phys. Lett.* (in the press).
MOSS, T. S., 1980, *Phys. Stat. Sol.* (*b*), **101,** 555.
MOSS, T. S., BURRELL, G. J., and ELLIS, B., 1973, *Semiconductor Opto-Electronics* (London : Butterworths).
MOUROU, G., and KNOX, W., 1979, *Appl. Phys. Lett.*, **35,** 492 ; 1980, *Ibid.*, **36,** 623.
MOUROU, G., KNOX, W., and STAVOLA, M., 1980, Picosecond phenomena II, *Springer Series in Chemical Physics*, Vol. 14, edited by R. Hochstrasser, W. Kaiser and C. V. Shank (Berlin : Springer-Verlag), p. 75.
MOYER, R., AGMON, P., KOCH, T., and YARIV, A., 1980, Picosecond phenomena II, *Springer Series in Chemical Physics*, Vol. 14, edited by R. Hochstrasser, W. Kaiser and C. V. Shank (Berlin : Springer-Verlag), p. 84.
MULLER, G. O., and ZIMMERMANN, R., 1979, *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 165.
NEE, T. W., CANTRELL, C. D., SCOTT, J. F., and SCULLY, M. O., 1978, *Phys. Rev.* B, **17,** 3926.
NURMIKKO, A. V., 1976 a, *Optics Commun.*, **16,** 365 ; 1976 b, *Ibid.*, **18,** 522.
NURMIKKO, A. V., and JAMISON, S. A., 1979, *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 315.
NURMIKKO, N. V., and PRATT, G. W., 1975, *Appl. Phys. Lett.*, **27,** 83.
PANKOVE, J. I., 1975, *Optical Processes in Semiconductors* (New York : Dover).
PHIPPS, C. R., and THOMAS, S. J., 1977, *Optics Lett.*, **1,** 93.
PIDGEON, C. R., WHERRETT, B. S., JOHNSTON, A. M., DEMPSEY, J., and MILLER, A., 1979, *Phys. Rev. Lett.*, **42,** 1785.
PLATTE, W., 1976, *Electronics Lett.*, **12,** 437.
PLATTE, W., and APPELHAUS, G., 1976, *Electronics Lett.*, **12,** 271.
RABI, I. I., 1937, *Phys. Rev.*, **51,** 652.
REINTJES, J. F., MCGRODDY, J. C., and BLAKESLEE, A. E., 1975, *J. appl. Phys.*, **46,** 879.
REYNOLDS, D. C., and COLLINS, T. C., 1981, *Excitons, Their Properties and Uses* (New York : Academic Press).
RICE, T. M., 1977, *Solid State Physics*, Vol. 32, edited by H. Ehrenreich, F. Seitz and D. Turnbull (New York : Academic Press), p. 1.
ROGACHEV, A. A., 1980, *Prog. quant. Electron.*, **6,** 141.
SARGENT, M., III, 1977, *Optics Commun.*, **20,** 298.
SARGENT, M., SCULLY, M. O., and LAMB, W. E., 1974, *Laser Physics* (Reading, Mass. : Addison–Wesley).
SCHMITT-RINK, S., TRAN THOAI, D. B., and HAUG, H., 1980, *Z. Phys.* B, **39,** 25.
SCHWARTZ, B. D., FAUCHET, P. M., and NURMIKKO, A. V., 1980, *Optics Lett.*, **5,** 371.
SCHWARTZ, B. D., and NURMIKKO, A. V., 1980, Picosecond phenomena II, *Springer Series in Chemical Physics*, Vol. 14, edited by R. Hochstrasser, W. Kaiser and C. V. Shank (Berlin : Springer-Verlag), p. 303.
SEILER, D. G., and HANES, L. K., 1979, *Optics Commun.*, **28,** 326.
SHAH, J., 1978, *Solid St. Electron.*, **21,** 43.
SHAH, J., LEHENY, R. F., and LIN, C., 1976, *Solid St. Commun.*, **18,** 1035.
SHAH, J., LEHENY, R. F., and WIEGMANN, W., 1977, *Phys. Rev.* B, **16,** 1577.
SHANK, C. V., and AUSTON, D. H., 1975, *Phys. Rev. Lett.*, **34,** 479.
SHANK, C. V., AUSTON, D. H., IPPEN, E. P., and TESCHKE, O., 1978, *Solid St. Commun.*, **26,** 567.
SHANK, C. V., FORK, R. L., GREENE, B. I., REINHART, F. K., and LOGAN, R. A., 1981, *Appl. Phys. Lett.*, **38,** 104.

SHANK, C. V., FORK, R. L., LEHENY, R. F., and SHAH, J., 1979 a, *Phys. Rev. Lett.*, **42,** 112.

SHANK, C. V., FORK, R. L., LEHENY, R. F., and SHAH, J., 1979 b, *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 493.

SMIRL, A. L., 1980, *Physics of Non-linear Transport in Semiconductors*, edited by D. Ferry, J. R. Barker and C. Jacobini (New York : Plenum Press), pp. 367 and 517.

SMIRL, A. L., BOGGESS, T. F., and HOPF, F. A., 1980, *Optics Commun.*, **34,** 463.

SMIRL, A. L., LINDLE, J. R., and MOSS, S. C., 1978 a, *Phys. Rev.* B, **18,** 5489 ; 1978 b, Picosecond phenomena, *Springer Series in Chemical Physics*, Vol. 4, edited by C. V. Shank, E. P. Ippen and S. L. Shapiro (Berlin : Springer-Verlag), p. 174.

SMIRL, A. L., MATTER, J. C., ELCI, A., and SCULLY, M. O., 1976, *Optics Commun.*, **16,** 118.

SMIRL, A. L., MILLER, A., PERRYMAN, G. P., and BOGGESS, T. F., 1981, *Proc. 3rd Int. Conf. Hot Carriers in Semiconductors* (in the press).

SMIRL, A. L., MOSS, S. C., and LINDLE, J. R., 1981 a, *Proc. Int. Conf. Excited States and Multiresonant Non-linear Optical Processes in Solids*, CNRS, Aussois, p. 73 ; 1981 b, *J. appl. Phys.* (to be published).

SMITH, P. W., 1980, *Opt. Engng*, **19,** 456 ; 1981, Special issue on optical bistability, *I.E.E.E. Jl quant. Electron.*, **17,** 300.

SMITH, R. A., 1978, *Semiconductors*, 2nd ed. (Cambridge : Cambridge University Press).

SMITH, S. D., DENNIS, R. B., and HARRISON, R. G., 1977, *Prog. quant. Electron.*, **5,** 205.

SMITH, S. D., and MILLER, D. A. B., 1979, Laser advances and applications, *Proc. 4th Nat. Quant. Electron. Conf.*, edited by B. S. Wherrett, (Chichester : Wiley), p. 231 ; 1980, *J. phys. Soc. Japan*, **49,** Suppl. A, 597.

SOOY, W. R., GELLER, M., and BORTFELD, D. P., 1964, *Appl. Phys. Lett.*, **5,** 54.

STAUPENDAHL, G., and SCHINDLER, K., 1980, *Proc. 2nd Int. Symp. Ultrafast Phenomena in Spectroscopy*, Jena, p. 437.

STAVOLA, M., AGOSTINELLI, J. A., and SCEATS, M. G., 1979, *Appl. Optics*, **18,** 4101.

STEEL, D. G., LIND, R. C., LAM, J. F., and GIULIANO, C. R., 1979, *Appl. Phys. Lett.*, **35,** 376.

STERN, F., 1963, *Solid State Physics*, Vol. 15, edited by F. Seitz and D. Turnbull (New York : Academic Press), p. 300.

SZÖKE, A., DANEU, V., GOLDHAR, J., and KUMIT, N. A., 1969, *Appl. Phys. Lett.*, **15,** 376.

ULBRICH, R. G., 1978, *Solid St. Electron.*, **21,** 51 ; 1979, *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 11 ; 1980, *Physics of Non-linear Transport in Semiconductors*, edited by D. Ferry, J. R. Barker and C. Jacobini (New York : Plenum), p. 327.

VAITKUS, YU, GAUBAS, E., and YARASHYUNAS, K., 1978, *Soviet Phys. Solid St.*, **20,** 1824.

VAITKUS, YU., JARASHUMAS, K., and BALTRAMEYUNAS, R., 1980, *I.E.E.E. Jl quant. Electron.*, **16,** 616.

VAN DRIEL, H. M., 1979, *Phys. Rev.* B, **19,** 5928.

VAN DRIEL, H. M., ELCI, A., BESSEY, J. S., and SCULLY, M. O., 1976, *Solid St. Electron.*, **20,** 837.

VINETSKII, V. L., KUKHTAREV, N. V., ODULOV, S. G., and SOSKIN, M. S., 1979, *Soviet Phys. Usp.*, **22,** 742.

VINETSKII, V. L., ZAPOROZHETS, T. E., KUKHTAREV, N. V., MATVILCHUK, A. S., SOSKIN, M. S., and KHOLEDAR, G. A., 1977, *Ukr. fiz. Zh.*, **22,** 1141.

VON DER LINDE, D., and LAMBRICH, R., 1978, Picosecond phenomena, *Springer Series in Chemical Physics*, Vol. 4, edited by C. V. Shank, E. P. Ippen and S. L. Shapiro (Berlin : Springer-Verlag), p. 232 ; 1979 a, *Phys. Rev. Lett.*, **42,** 1090 ; *Inst. Phys. Conf. Ser.*, Vol. 43, Physics of semiconductors, 1978, edited by B. L. H. Wilson (Bristol : Institute of Physics), p. 517.

Walker, A. C., and Alcock, A. J., 1974, *Optics Commun.*, **12,** 430.
Watkins, D. E., Phipps, C. R., and Thomas, S. J., 1981, *Optics Lett.*, **6,** 76.
Weaire, D., Wherrett, B. S., Miller, D. A. B., and Smith, S. D., 1979, *Optics Lett.*, **4,** 331.
Wherrett, B. S., 1977, *Nonlinear Optics*, edited by P. G. Harper and B. S. Wherrett (London : Academic Press).
Wherrett, B. S., and Higgins, N. A., 1981, *Proc. R. Soc.* (to be published).
Wiggins, T. A., Bellay, J. A., and Carrieri, A. H., 1978, *Appl. Optics.*, **17,** 526.
Wiggins, T. A., and Carrieri, A. H., 1979, *Appl Optics.*, **18,** 1921.
Wiggins, T. A., and Qualey, J. R., 1979, *Appl. Optics.*, **18,** 960.
Wiggins, T. A., and Salik, A., 1974, *Appl. Phys. Lett.*, **25,** 438.
Woerdman, J. P., 1970 a, *Phys. Lett.* A, **32,** 305 ; 1970 b, *Optics Commun.*, **2,** 212 ; 1971, *Philips Res. Rep.*, Suppl. **7.**
Woerdman, J. P., and Bolger, B., 1969, *Phys. Lett.* A, **30,** 164.
Yariv, A., 1975, *Quantum Electronics*, 2nd ed. (New York : Wiley) ; 1978, *I.E.E.E. Jl quant. Electron.*, **14,** 650 ; 1979, *Ibid.*, **15,** 524.
Yariv, A., Fekete, D., and Pepper, D. M., 1979, *Optics Lett.*, **4,** 52.
Yoffa, E. J., 1980, *Phys. Rev.* B, **21,** 2415 ; 1981, *Ibid.*, **23,** 1909.
Yuen, S. Y., Aggarwal, R. L., and Lax, B., 1980, *J. appl. Phys.*, **51,** 1146.
Yuen, S. Y., Aggarwal, R. L., Lee, N., and Lax, B., 1979, *Optics Commun.*, **28,** 237.