Proton-Implanted Optical Waveguides and Integrated Optical Detectors in Gallium Arsenide

Thesis by

Harold M. Stoll

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1974
(Submitted May 6, 1974)

To Keren

ACKNOWLEDGMENTS

First, I would like to thank my thesis advisor, Dr. Amnon Yariv, for providing the direction, encouragement, and support I enjoyed while pursuing this study. The solutions to many difficult problems encountered during the course of writing this thesis owe their existence to his keen physical insight. I would also like to offer a special thanks to Dr. Elsa Garmire whose original idea inspired the present investigation. Her enthusiasm in and encouragement of my work will always be deeply appreciated. The professional assistance of Dr. Robert Hunsperger of Hughes Research Laboratories is also gratefully acknowledged; numerous important contributions to this thesis resulted from our many stimulating discussions.

It is also a pleasure to acknowledge the skillful technical assistance of Desmond Armstrong and the secretarial proficiencies of Dian Rapchak and Ruth Stratton, who did such a marvelous job of typing the manuscript. Additionally, I would like to acknowledge the fine technical support provided by the Hughes Research Laboratories and the generous financial support received from the National Science Foundation, the Northrop Corporation, General Telephone and Electronics Laboratories, the Baker Foundation, the ARCS Foundation and the California Institute of Technology.

Finally, for buoying my spirits when the going got tough and for cheering me on when I had succeeded, I would like to thank my wife, Keren. Her love and understanding combined with the encouragement and

support received from our parents contributed immeasurably to the success of my graduate education.

ABSTRACT

Defect-associated energy levels which appear within the forbidden energy gap of proton-irradiated gallium arsenide give rise both to free carrier compensation and to additional near band-edge optical attenuation. These damage-induced changes in the electrical and optical properties of gallium arsenide have been exploited in the fabrication of passive optical waveguides and waveguide-compatible detectors suitable for use in integrated optical circuits. In order to understand the physical processes which underlie the operation of these structures and devices and, thereby, to suggest ways in which the irradiation and post-irradiation fabrication parameters may be optimized, a model has been constructed which explains the experimentally observed electrical and optical properties of proton-irradiated gallium arsenide. Using this model and data obtained by other investigators, a self consistent analysis of the optical confinement and attenuation properties of waveguides formed by irradiating n-type gallium arsenide with 300 keV protons has been made: Optical confinement is found to result from the so-called plasma depression effect; optical attenuation is found to be due to a combination of dissipative scattering and absorption by thermal spikes and dipole-assisted transitions between defect levels and the band continua. The model is also used to make a preliminary analysis of an integrated optical detector fabricated using the same proton-irradiation technique.

TABLE OF CONTENTS

INTRODUCTION 1

CHAPTER 1 - RADIATION DAMAGE 5

I. Introduction 5

II. Disorder Production 6

A. Absolute Defect Concentrations 6

B. Defect Distributions 14

C. Channeling Effects 17

CHAPTER 2 - OPTICAL WAVEGUIDING AND ATTENUATION 20

I. Introduction 20

II. One-Dimensional (Planar) Waveguiding 22

A. The Wave Equation 22

B. Four Media Waveguide 24

III. Loss Calculations-The Perturbation Technique 34

IV. The Kramers-Kronig Relations 39

V. Absorption Mechanisms 41

A. Free Carrier Intraband Transitions 41

B. Transitions Involving Discrete States and Bands 43

CHAPTER 2 - APPENDIX 50

I. Metallic Overlayer Losses of A Three Media Guide 50

A. Perturbation Technique Solution 50

B. Solution of the Complex Wave Equation 55

II. Substrate Losses of a Three Media Guide 58

A. Perturbation Technique Solution 58

B. Solution of the Complex Wave Equation 60

CHAPTER 3 - THE OPTICAL AND ELECTRICAL PROPERTIES OF PROTON-IMPLANTED GALLIUM ARSENIDE 63

I. Introduction 63

II. Ion-Implanted Gallium Arsenide 63

III. Proton-Implanted Gallium Arsenide 64

A. Optical Properties of Proton-Implanted Gallium Arsenide 68
1. Absorption and Scattering by Thermal Spikes 68
2. Absorption by Impurity-Defect Complexes 69
3. Absorption Involving Bandtails 82
B. Absorption-Induced Refractive Index Changes 95
C. Electrical Properties of Proton-Implanted Gallium Arsenide 98

CHAPTER 4 - OPTICAL WAVEGUIDING AND ATTENUATION IN PROTON-IMPLANTED GALLIUM ARSENIDE 108
I. Introduction 108
II. The Depth Distribution of Damage 108
III. Optical Waveguiding 117
IV. Optical Attenuation 120
V. Data Analysis 126
VI. Discussion 158
VII. Conclusion 160

CHAPTER 5 - A PROTON-IMPLANTED INTEGRATED OPTICAL DETECTOR 165
I. Introduction 165
II. Device Construction and Geometry 166
III. Device Operation 168
IV. Device Performance 168

CHAPTER 6 - EXPERIMENTAL TECHNIQUES 176
I. Proton Implantation 176
II. Post Implantation Annealing 176
III. Mode Profile and Attenuation Measurements 178
IV. Free Carrier Concentration Measurements 181
V. Step Etching 181

REFERENCES 184

INTRODUCTION

The advantages offered by communications systems operating in the visible and near-infrared regions of the electromagnetic spectrum are essentially two-fold: (1) a considerable reduction in system size and weight and (2) a tremendous increase in bandwidth and, therefore, in system capacity. Currently available optical communications systems (i.e., those operating over relatively short distances and using incoherent energy sources) appear adequate to satisfy present needs wherein only reductions in system size and weight are desired. As human and machine populations increase in size and sophistication, however, the bandwidth advantages of optical communication systems will eventually have to be exploited in order to accommodate attendant increases in the flow of information. The required large-bandwidth optical communications systems are envisioned as being composed of compact optical processing stations interconnected by fibre-optical transmission links.

In the interests of compactness and reliability the optical processing stations referred to above will be composed of integrated optical circuits whose elements (coherent energy sources, modulators, detectors, etc.) will have dimensions on the order of a wavelength ($\sim 1\mu$) and will be fabricated on a common substrate. With regard to the integration of individual device components two approaches are presently being considered. The first, or heterogeneous approach, would combine the best available components from different materials systems while the second, or homogeneous approach, would require all components to be fabricated within the same material system. Although the second approach might, under certain circumstances, be disadvantageous from the standpoint of device performance, the

potential advantages it offers in the way of production simplicity and overall system reliability would appear to make it the more attractive alternative.

Because its properties are well understood and highly controllable as a result of efforts in the electronics industry and because it has been used to construct many of the necessary device elements, the gallium arsenide material system appears particularly well-suited for the construction of homogeneous integrated optical circuits of the type described above. Having at one's disposal a suitable material system, however, only solves one-half of the problem; in order to realize fully the inherent advantages of the homogeneous approach, compatible processes must be developed for fabricating both the individual device elements and the waveguides which interconnect them. It is toward the goal of fabricating homogeneous gallium arsenide integrated optical circuitry that the present effort has been directed.

The investigations to be described here were motivated, as mentioned above, by the need to develop a process by which planar integrated optical devices could be simply and accurately fabricated in gallium arsenide. In particular, the goal was undertaken to fabricate waveguiding structures in n-type gallium arsenide by making use of the free carrier compensation effects caused by proton irradiation. The technique of ion implantation or elementary particle irradiation lends itself naturally to the fabrication of mass produced integrated optical circuits. One can, for example, imagine a sequential implantation process wherein different projectile ions, ion masks, and annealing procedures are used to fabricate entire optical circuits in gallium arsenide. In this regard, the present study has demonstrated that both active (integrated optical detectors) as well as passive (wave-

guides) devices may be constructed by selectively varying the dose and heat treatment history of proton-irradiated gallium arsenide.

The planar (one dimensional) waveguides and integrated optical detectors referred to above were fabricated by irradiating n-type gallium arsenide with 300 keV protons. Defects generated by the impinging protons gave rise to defect levels within the forbidden energy gap which served both to compensate free carriers and to generate loss at the wavelengths investigated (1.06μ and 1.15μ). Following suitable post-irradiation heat treatment in order to reduce the defect-associated losses to an acceptable level ($\sim$4 db cm^{-1}) the former effect was used to produce waveguiding structures via the free carrier plasma depression effect. Conversely, by not annealing the irradiated material, the latter effect, involving dipole transitions between defect levels and the band continua, was exploited in order to fabricate integrated optical detectors.

The six chapters of this thesis may be conceptually divided into two parts. The first part, consisting of Chapters 1 and 2, deals with certain general principles and techniques which will be found useful in analyzing the optical confinement and attenuation properties of radiation-generated, multilayer waveguides. Chapter 1 is begun with derivations of the proportionality constants which relate absolute defect concentrations to proton fluence for both high and low energy interactions; it is concluded with a brief discussion of defect concentration distributions and channeling effects. Chapter 2 is concerned with the optical confinement and attenuation properties of the three and four media planar waveguiding structures which are found to exist as a result of irradiating gallium arsenide with high energy protons.

The second part of the thesis, consisting of Chapters 3, 4, and 5, is concerned with both an analysis of the physical changes which occur in gallium arsenide after irradiation by high energy protons and an interpretation, within this analytical framework, of the devices which were fabricated. The model of proton-irradiated gallium arsenide is developed in Chapter 3. Using this model as a foundation, the optical mode profiles and attenuation coefficients of gallium arsenide waveguides formed by proton-irradiation are analyzed in Chapter 4. A preliminary investigation of the properties of integrated optical detectors formed by irradiating epitaxial gallium arsenide waveguides with protons is presented in Chapter 5.

Chapter 6 concludes the thesis with a description of the experimental techniques used.

CHAPTER 1

RADIATION DAMAGE

I. Introduction

Atomic projectiles, in coming to rest within a target material, create disorder which, in turn, modifies to some extent the electrical, thermal, and optical properties of the resulting projectile-target amalgam. Depending on the size and charge state of the projectile, the physical changes induced in the target arise either from the equilibrium disorder created, from the chemical properties of the projectile-dopant or from a combination of both. If the size of the implanted species is such as to allow substitutional occupation of lattice sites, the resulting physical properties of the implanted material will largely be determined by the charge state of the implanted ion and any residual disorder which remains after heat treatment. On the other hand, if the projectile is very small, as is the case with protons, neutrons and electrons, no substitutional occupation will occur and any physical changes which obtain will be a function of how much disorder the projectile generated in the course of spending its energy. Since proton projectiles were used in the present study, only disorder-associated modifications of the target, gallium arsenide, will be considered.

We begin our treatment by deriving proportionality constants which relate defect concentration to incident particle flux. Two such proportionality constants, one representing the effects of high energy interactions and the other the effects of low energy interactions, will

be obtained. Their regimes of validity will be investigated by considering, from a qualitative point of view, the dynamic behavior of projectiles coming to rest within a target. Some recently developed techniques used to obtain a quantitative representation of radiation-generated disorder distributions will then be discussed. We conclude the chapter with a brief description of channeling effects.

II. Disorder Production

An implanted ion (proton in our case) may lose energy through two types of interactions [1]: those involving inelastic collisions with target electrons and those involving inelastic collisions with the target nuclei; only the second mechanism results in atomic disorder. The dynamic evolution of a proton of given kinetic energy may therefore be characterized by two parameters both of which are energy dependent: the cross section for electronic scattering and the cross section for nuclear scattering. Knowledge of a third parameter, the number of secondary displacements per primary nuclear event, allows, in principle, a complete characterization of disorder production to be made.

A. Absolute defect concentrations

The number of primary displacements (i.e., displacements caused by projectile-target collisions) per unit volume and unit time generated by incident radiation of energy E is given by

$$n_p = \phi N_o \sigma_n(E) \qquad (1\text{-}1)$$

where N_o is the number of target atoms per unit volume, $\sigma_n(E)$ is

the nuclear scattering cross section and ϕ is the number of incident particles crossing a unit area per unit time. If the bombardment is terminated after a time t , the number of primary displacements is given (for $N_p << N_o$) by:

$$N_p = \phi t\, N_o \sigma_n(E) = \Phi N_o \sigma_n(E) \tag{1-2}$$

where $\Phi (\equiv \phi t)$ is termed the total integrated flux or dose.

Primary displacements (termed primary knock-ons), if sufficiently energetic, produce additional secondary displacements, the whole process resulting in a disorder cascade. The average number of displaced atoms in a cascade, $\bar{\nu}(E)$, is a function of the energy spectrum of the primary knock-ons which, in turn, is dependent on the original energy, E , of the incident radiation. The total number of displaced atoms including primary and secondary displacements therefore becomes:

$$N_d = N_p \bar{\nu}(E) = \Phi N_o \sigma_n(E) \bar{\nu}(E) \tag{1-3}$$

Kinchin and Pease [2] calculate $\nu(T)$, the number of displacements generated by a primary knock-on of energy, T , to be:

$$\nu(T) = \begin{cases} 0 & T < E_d \\ 1 & E_d \leq T \leq 2E_d \\ T/2E_d & 2E_d \leq T \leq E_i \\ E_i/2E_d & T \geq E_i \end{cases} \tag{1-4}$$

where E_d is the threshold for nuclear displacements and E_i is the electronic ionization energy of the target atoms. Equation (1-4) was

derived by assuming hard sphere collisions and neglecting replacement processes which involve the re-introduction of displaced atoms into vacated lattice sites. That the rather gross assumption of simple hard sphere scattering has little effect on the value of $\nu(T)$ is demonstrated by the results of Snyder and Neufeld [3]. Assuming Rutherford type scattering and counting replacements as displacements, they obtained for $\nu(T)$:

$$\nu(T) = \begin{cases} 1 & 0 \leq T < 2E_d \\ 0.56\,(T/E_d) & T \geq 2E_d \end{cases} \tag{1-5}$$

Finally, Harrison and Seitz [4], assuming a sharp threshold for displacements but otherwise approaching the problem in the manner of Snyder and Neufeld, obtained for $\nu(T)$:

$$\nu(T) = \begin{cases} 1 & 0 \leq T < E_d \\ 0.56(T+E_d)/E_d & T \geq E_d \end{cases} \tag{1-6}$$

The average number of displacements in a cascade $\bar{\nu}(E)$ may now be found by evaluating the following integral [5]:

$$\bar{\nu}(E) = \int_0^{T_m} \nu(T)P(T)\,dT \tag{1-7}$$

where $P(T)$, the probability that a primary knock-on received energy T, is given by [5]:

$$P(T) = \frac{\sigma_n(T)p(T)}{\int_0^{T_m} \sigma_n(T)p(T)dT} \tag{1-8}$$

and T_m , the maximum energy that may be imparted to a primary knock-on in an elastic collision, is given by:

$$T_m = \frac{4m_p M}{(m_p+M)^2} E \quad . \tag{1-9}$$

$\sigma_n(T)dT = d\sigma_n(T)$ is the differential nuclear scattering cross section, E is the energy of the incident protons, m_p is the projectile (proton in our case) mass, M is the averaged mass of the target atoms (taken to be 72.3 a.m.u. for GaAs) and $p(T)$, the probability that an atom becomes displaced after receiving energy T is

$$p(T) = \begin{cases} 1 & T > E_d \\ 0 & T < E_d \end{cases} \quad . \tag{1-10}$$

As mentioned above, the value of $\nu(T)$ is relatively insensitive to the form of the scattering law assumed. To determine $\bar{\nu}(E)$, however, one must decide whether to use a hard sphere scattering cross section or a Rutherford scattering cross section in Eq. (1-8). One may obtain a rule-of-thumb solution to this dilemma by first assuming that the interaction potential between moving and stationary particles is of the Thomas-Fermi, screened coulomb type [6]

$$V(r) = (Z_p Z e^2/r)e^{-r/a} \tag{1-11}$$

for which the impact parameter, or point of closest approach, is given by [6]

$$b = Z_p Z e^2 (m_p + M)/(E \cdot M) \quad . \tag{1-12}$$

Here, Z_p is the proton charge (=1), Z is the averaged target charge (=32 for $\overline{\text{GaAs}}$), r is the instantaneous separation of target and projectile, E is the incident projectile energy, and a is the screening distance given by [6]

$$a = 0.8853 a_o (Z_p^{2/3} + Z^{2/3})^{-1/2} \tag{1-13}$$

where a_o = the hydrogen Bohr radius (=5.3 x 10^{-9}cm). The rule for deciding which type of interaction to choose is the following: For impact parameters smaller than a, the projectiles (protons) will be strongly affected by the interaction potential and Rutherford scattering will be assumed. For impact parameters larger than a, the projectiles will not be able to penetrate the screened target and hard sphere scattering will be assumed. The incident proton energy, E_A, below which hard sphere scattering predominates and above which Rutherford scattering prevails may be found by first setting $E = E_A$ in Eq. (1-12) and then setting a equal to b. The result is [5]

$$E_A = \frac{Z_p Z (Z_p^{2/3} + Z^{2/3})^{1/2} \cdot e^2 \cdot (m_p + M)}{0.8853\, a_o\, M} \tag{1-14}$$

For protons incident on GaAs, $E_A = 3.3 \times 10^3$eV . Unfortunately, therefore, neither the hard sphere scattering approximation nor Rutherford scattering is entirely appropriate for dealing with 300 keV protons. Lindhard, Scharff, and Schiott (LSS) have, however, obtained an exact expression for the nuclear scattering cross section, based on the Thomas-Fermi interaction potential, which is applicable for all energies [7]:

$$d\sigma_n(T) = \frac{\pi a T_m^{1/2} Z_p Z e^2 (m_p + M)}{2E \cdot M T^{3/2}} f(T) dT \qquad (1\text{-}15)$$

$f(T)$ in Eq. (1-15) is a numerically computed function (See reference [7].). Therefore, in order to use Eq. (1-15) in Eq. (1-8) one must perform a numerical integration. Because of this complication Eq. (1-15) will not be used to derive $\bar{\nu}(E)$. Instead $\bar{\nu}(E)$ will be computed using both hard sphere and Rutherford cross sections. The regimes of validity of both of these approaches will then be investigated as a prelude to the discussion of defect distributions which follows in Section II-B.

Proceeding with the analysis, the necessary cross sections are given by [5]

Rutherford: $$d\sigma_r(T) = \frac{\pi Z_p^2 Z^2 e^4 m_p}{M \cdot E} \cdot \frac{dT}{T^2} \qquad (1\text{-}16a)$$

Hard Sphere: $$d\sigma_{h.s.}(T) = \frac{\pi a^2}{T_m} dT \quad . \qquad (1\text{-}16b)$$

Using Eq. (1-16a) and Eq. (1-16b) to evaluate Eq. (1-8) yields:

$$P_r(T) = \frac{T_m E_d}{T_m - E_d} \cdot \frac{1}{T^2} \qquad (1\text{-}17a)$$

$$P_{h.s.}(T) = \frac{1}{T_m - E_d} \; . \qquad (1\text{-}17b)$$

Taking, for simplicity, the value of $\nu(T)$ obtained by Kinchin and Pease, Eqs. (1-17a) and (1-17b) together with Eq. (1-7) yield for the average number of displacements for both Rutherford and hard sphere scattering:

$$\bar{\nu}_r(E) = \frac{1}{2}\left(\frac{T_m}{T_m - E_d}\right)[1 + \ln(T_m/2E_d)] \qquad (1\text{-}18a)$$

$$\bar{\nu}_{h.s.}(E) = \frac{T_m^{\,2}/(4E_d)}{T_m - E_d} \qquad (1\text{-}18b)$$

Finally, Eqs. (1-18a), (1-18b) and (1-3) give for the total number of displacements:

$$(N_d)_r = \frac{1}{2}\,\frac{\pi Z_p^{\,2} Z^2 e^4 m_p}{M \cdot E \cdot E_d} \cdot [1 + \ln(T_m/2E_d)] \cdot N_o \Phi \qquad (1\text{-}19a)$$

$$(N_d)_{h.s.} = \frac{\pi a^2 T_m}{4E_d} \cdot N_o \Phi \qquad (1\text{-}19b)$$

Taking, for the case of gallium arsenide, $N_o = 4.42 \times 10^{22}/cm^3$ and $E_d = 17.5$ eV [8], yields, for $E = 300$ keV, $(N_d)_r = 277 \times 10^4 \Phi$ and $(N_d)_{h.s.} = 6.37 \times 10^7 \Phi$. Fig. 1-1, which shows some theoretical damage distributions [12], provides a general means by which these values may be interpreted. As the projectiles come to rest they pass through

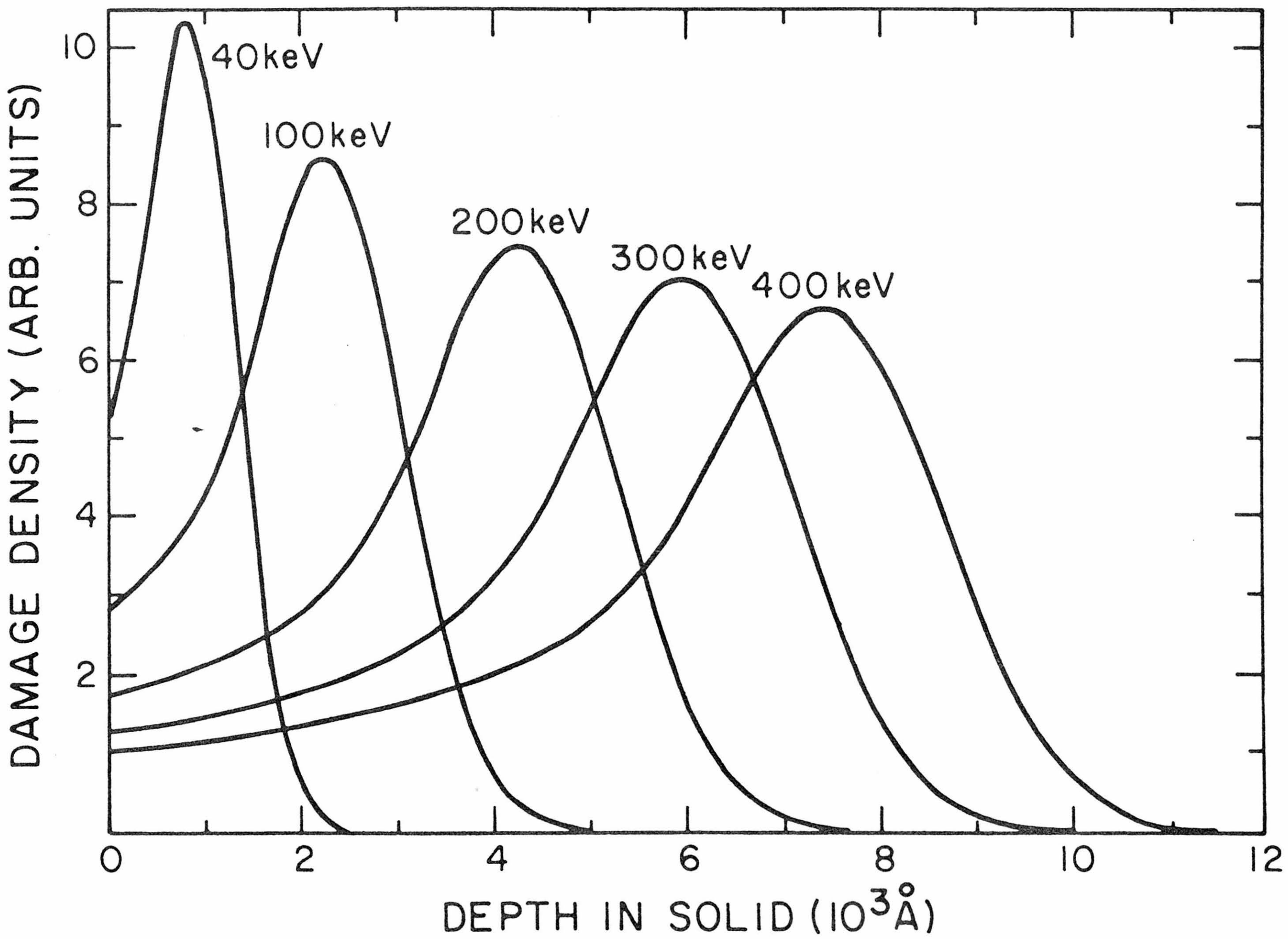


Figure 1-1. Depth distribution of energy deposited into atomic processes as calculated by the two step method for B^{11} incident on an amorphous silicon target (After Brice [12]).

two energy regimes: Near the beginning of their travel they are very energetic and most encounters are of the Rutherford type. Once their energy has descended below a certain threshold, 3.3×10^3eV for the case of protons incident on gallium arsenide, hard sphere scattering predominates and a damage peak results. Finally, it will be noted that the values of $(N_d)_r$ and $(N_d)_{h.s.}$ which we have obtained almost certainly overestimate the actual amount of damage [9]. Reasons cited for this discrepancy include: (1) channeling effects not included because of the assumption of target atom randomness (2) partial self annealing and (3) electronic collisions which were not considered in the analysis. The following section outlines several, more sophisticated techniques used to characterize implantation damage wherein electronic collisions are explicitly considered.

B. Defect distributions

Three techniques have been developed to calculate ion and disorder depth distributions. The first [10], which utilizes a Monte Carlo approach, traces the history of a specified number of incident ions and the secondaries they generate in coming to rest. Histograms, with statistical scatter corresponding to the number of events considered, are obtained and then smoothed using a computer program to yield final ion and disorder distributions. The second method [11] begins with integral equations for the distributions desired and approximates the final solution by calculating certain disorder distribution moments. The third technique [12], similar in most

respects to the second, seeks to find the depth distribution of energy deposited into atomic processes; the resulting distribution is then correlated with the amount of disorder as a function of depth. The first procedure requires a great amount of numerical calculation in order to reduce the amount of statistical scatter in the resulting histograms,and a solution using the second technique is overly complicated owing to the inherent complexity of the integral equations and their attendant boundary conditions. Reliance is therefore placed on the third,or so-called two-step method,developed by Brice.

Brice's technique involves first determining the spatial distribution of the incident ions after they have come to rest. Then, knowing the nuclear scattering cross section and the relative partition of energy deposition between atomic and electronic processes, the final depth distribution of energy expended in producing atomic disorder is calculated. The procedure, in some detail, is as follows: letting $Q(\vec{r})$ be the amount of energy deposited in atomic processes as a function of position, $P_3(E,E',\vec{r})$ be the probability that an ion incident with energy E finds itself at position $\vec{r}$ with energy $E'(<E)$, f be the fraction of E' available for nuclear interactions, and $d\sigma$ be the differential nuclear scattering cross section, the spatial differential $d^3Q(\vec{r})$ is given by

$$d^3Q(r) = \int_E^0 \left\{ P_3(E,E',\vec{r})dr\ N(\overrightarrow{dR/dE'}) \left| \int_\sigma fd\sigma' \right| \right\} dE' \qquad (1\text{-}20)$$

where $d\vec{R}$ is the distance traveled by an ion in losing the amount of energy dE' and N is the number of (randomly-situated) target atoms per unit volume.

The first step of Brice's two step method requires the determination of $P_3(E,E',\vec{r})$, the spatial distribution of spent projectiles. This distribution is provided by the solution of certain integro-differential equations similar to those originally derived by Lindhard, Scharff, and Schiott [7]. The distribution which emerges is:

$$P_3(E,E',\vec{r}) = P_1(E,E',x)(\pi\overline{\Delta R_\perp}^2)^{-1}\exp[-(y^2+z^2)/\overline{\Delta R_\perp}^2] \qquad (1\text{-}21)$$

where

$$P_1(E,E',x) = (2\pi\overline{\Delta R_p}^2)^{-1/2}\exp[-(x-\overline{R}_p)^2/2\overline{\Delta R_p}^2]. \qquad (1\text{-}22)$$

The parameters $\overline{R}_p$, $\overline{\Delta R_p}^2$, and $\overline{\Delta R_\perp}^2$ represent the average projectile range, the projectile range scatter, and the latteral projectile scatter, all of which are derived from the solution of the aforementioned integro-differential equations.

The second step involves the solution of Eq. (1-20) which may be obtained once f, $d\sigma$, and $d\overline{R}/dE$ are known. Using a Thomas-Fermi screened coulomb interaction potential to calculate $d\sigma$ and tabulated values of f and $d\overline{R}/dE$, Brice has calculated the disorder profiles for a number of projectile-target combinations.The theoretical disorder profiles generated by B^{11} implanted into Si are shown in Fig. 1-1, this projectile-target combination being the one which most closely parallels the proton-gallium arsenide combination under consideration.

While the present study does not require a rigorous know-

ledge of the disorder profile (an empirically derived distribution will be found to be adequate), further investigations involving sequential and/or masked implantations will necessitate a more exact knowledge of the disorder distribution and a method such as Brice's will have to be implemented. Furukawa et al. [13], for example, have shown that the lateral spread of protons implanted through a mask is a considerable fraction of their total range. Such uncertainty in edge definition would have an appreciable effect on the behavior of coupled, channel optical waveguides [14] fabricated in a similar manner.

C. Channeling effects

All the treatments of disorder production described thus far assume that the atoms of the target are randomly distributed. When the effects of target crystallinity are taken into consideration, however, the important phenomenon of channeling emerges [1].

When an ion is incident along a direction parallel to or nearly parallel to a set of crystal planes, it becomes "channeled" or trapped between rows of the target lattice. Atomic collisions become less likely under these circumstances and most of the ion's original energy is lost through glancing electronic encounters (See Fig. 1-2). The consequences of such behavior are twofold: residual atomic disorder is found at penetration depths exceeding those predicted by the theories LSS and Brice, and the absolute disorder concentrations found in Section II-A are reduced because fewer primary knock-ons are generated. Oen and Robinson [15] modified the displacement cascade

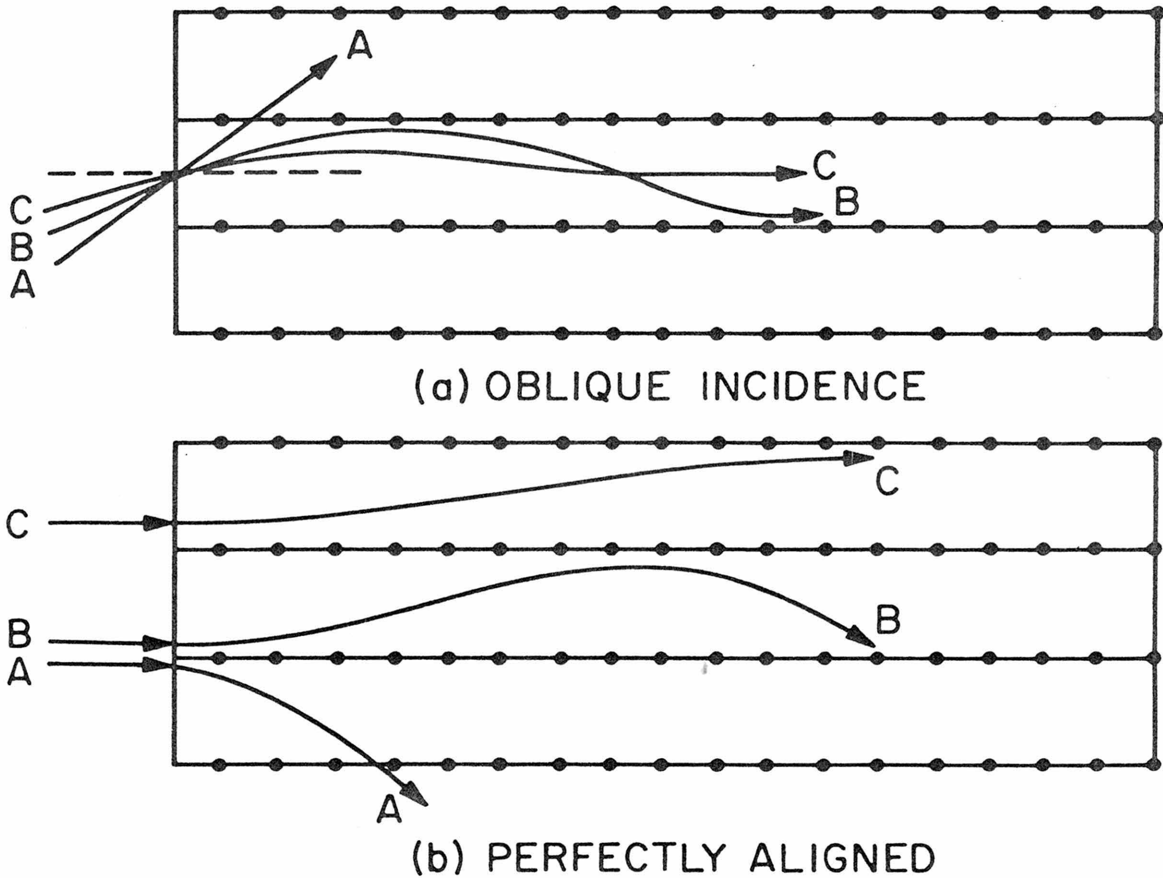


Figure 1-2. Schematic representation of ion trajectories for axial channeling. (From Mayer and Marsh [16]) The crystal lattice is depicted as a set of atomic chains:
(a) Trajectories for various angles of incidence relative to the lattice row: B and C represent trajectories for angles less than the critical angle, and A for values greater than the critical angle.
(b) Trajectories for parallel incidence as a function of impact position. (After Mayer, et al [1])

model used by Kinchin and Pease to include the effects of channeling. In place of Eq. (1-4), they obtained

$$\nu(t) = \frac{1}{1 - 2P}\left[(1-P)\left(\frac{T}{2E_d}\right)^{(1 - 2P)} - P\right] \qquad T > 2E_d \qquad (1\text{-}23)$$

where P is the probability that an ion becomes channeled. Instead of varying with the first power of T , as in Eq. (1-4), $\nu(T)$ is now seen to vary as $T^{(1 - 2P)}$. At least part of the discrepancy between measured values of disorder production and those theoretically predicted may now presumably be eliminated if one knows P . Unfortunately, the degree of channeling is a very sensitive function of ion beam alignment and target surface preparation and, hence, P remains an elusive parameter. However, even though P may not be reliably determined theoretically, one supposes that it will be large for the case of small, light projectiles such as protons implanted into relatively open lattices such as possessed by gallium arsenide. As will be seen in Chapter 4, channeling appears to play a significant role in the creation of proton-implanted GaAs waveguides.

CHAPTER 2

OPTICAL WAVEGUIDING AND ATTENUATION

I. Introduction

Electromagnetic radiation is said to be waveguided when its natural tendency to diffract is overcome by the presence of refracting media arranged in a specific geometry. Accordingly, a waveguiding structure consists of a guiding medium surrounded, in either or both dimensions transverse to the direction of wave propagation, by cladding/substrate media whose refractive indices are such as to effect total internal reflection at the cladding/substrate guide interfaces. The interfacial index discontinuities and waveguide dimensions are chosen according to the mode of propagation desired and the wavelength to be confined.

This study has been concerned with the waveguiding of near-infrared radiation $(\lambda_0 \sim 1\mu)$ in planar (one-dimensional) structures formed by proton-implanting gallium arsenide. Typical guide widths were on the order of $3 - 5\mu$ as determined by the maximum penetration depth of 300 keV protons. The guide-substrate index discontinuity required for mode confinement resulted from material changes in the proton-damaged gallium arsenide, the origins of which will be examined at length in the following chapter. Either air or metal (See Chapter 5) formed the overlaying medium. Since the refractive index discontinuity of the overlayer-guide interface was considerably larger than that of the guide-substrate interface, waveguiding characteristics were predominantly determined by the latter. Figure 2-1 shows a typical guide configuration.

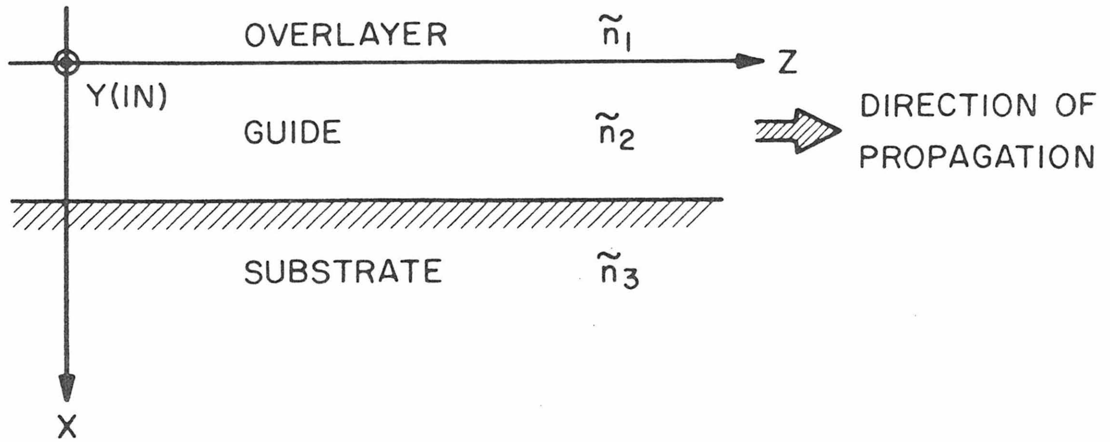


Figure 2-1. Typical dielectric waveguide configuration.

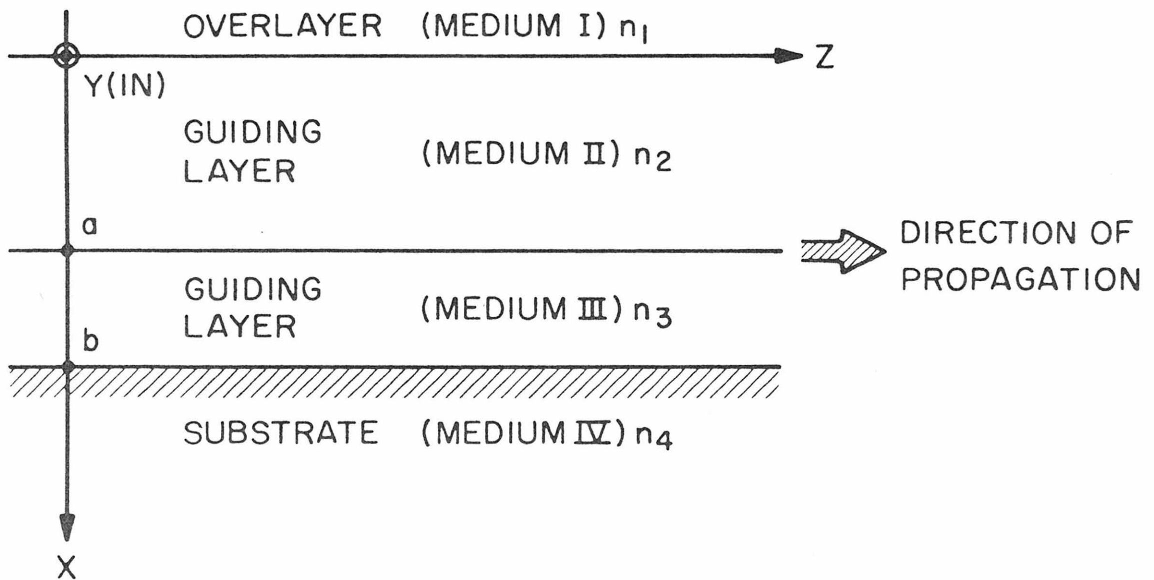


Figure 2-2. Four media waveguiding structure.

A review of the equations which govern waveguiding and attenuation in general followed by a mode analysis of the four media waveguiding structure to be encountered in Chapter 4 constitute Section II of this chapter. In Section III, a perturbative technique is developed for calculating the attenuation coefficients of specified waveguide geometries. The chapter is concluded with a discussion of the interrelationship which exists, via the Kramers-Kronig equations, between the real and imaginary parts of the complex dielectric constant; free carrier and dipole-associated absorption mechanisms, which may be treated within this formalism, are then dealt with extensively.

II. One-Dimensional (Planar) Waveguiding

A. The Wave Equation

Maxwell's equations for radiation propagating in a homogeneous, charge-free, non-magnetic, isotropic medium of dielectric constant ε and conductivity σ are given by:

$$\nabla \times \vec{E} = -\mu_0 \partial\vec{H}/\partial t \tag{2-1a}$$

$$\nabla \times \vec{H} = \sigma\vec{E} + \varepsilon\, \partial\vec{E}/\partial t \tag{2-1b}$$

$$\nabla \cdot \vec{H} = 0 \tag{2-1c}$$

$$\nabla \cdot \vec{E} = 0 \quad . \tag{2-1d}$$

By combining Eqs. (2-1a) and (2-1b) and using Eq. (2-1d), the familiar wave equation emerges:

$$\nabla^2\vec{E} = \mu_o\sigma\partial\vec{E}/\partial t + \mu_o\varepsilon\ \partial^2\vec{E}/\partial t^2 \quad , \tag{2-2}$$

A similar equation is obtained for the magnetization vector, $\vec{H}$.

Guided radiation is assumed to be confined in the x-direction only and to be propagated in the z direction; all quantities involving the differential $\frac{\partial^2}{\partial y^2}$, therefore, equal zero. Furthermore, modes of propagation are distinguished as being either T.E. (transverse electric --no $\vec{H}$ component in the y-direction), or T.M. (transverse magnetic --no $\vec{E}$ component in the y-direction). When solving for T.E. modes, the $\vec{E}$ field component in the y-direction will be sought first. Similarly, when solving for T.M. modes, the y-component of the H-field will be solved for first. The other field components are then readily obtained using Eqs. (2-1). Following this convention, a trial solution of Eq. (2-2) is chosen as:

$$\vec{E}_y(x,z;t) = \hat{y}\ E_y(x)e^{i(\omega t-\beta z)} \tag{2-3}$$

where $\hat{y}$ is a unit vector in the y-direction.

Substitution of Eq. (2-3) into Eq. (2-4) yields:

$$\frac{d^2\ E_y}{dx^2} + (\beta^2 - \tilde{n}_i^2 k_o^2)E_y = 0\ ;\ i = 1,2,3 \tag{2-4}$$

with an identical equation resulting for H_y . The complex refractive index, $\tilde{n}$ ($\equiv$ n-ik), of Eq. (2-4) is related to the complex dielectric constant, $K(\equiv K_r - iK_i)$, in the following way:

$$\tilde{n}^2 \equiv (n-ik)^2 \equiv K \equiv K_r - iK_i \equiv \frac{1}{\varepsilon_0} (\varepsilon - i \frac{\sigma}{\omega}) . \qquad (2\text{-}5)$$

When $\sigma \neq 0$, the propagation constant, β, becomes imaginary, giving rise to optical attenuation. For a medium of infinite extension, $d^2E_y/dx^2 = 0$ and the resulting T.E.M. (transverse electromagnetic) wave attenuates as:

$$e^{2 \cdot Im[\beta]z} = e^{-\alpha z} \qquad (2\text{-}6)$$

where:

$$\alpha \equiv 2Im[\beta] = 2\, k \cdot k_o = K_i k_o / n \quad . \qquad (2\text{-}7)$$

$k_o (=2\pi/\lambda_o)$ in Eq. (2-7) is the free-space propagation constant. α, the attenuation coefficient, will be found to be an important parameter when calculating guided mode attenuation using the perturbative technique to be developed in Section III.

B. Four Media Waveguide

It will be found in Chapter 4 that a combination of channeling (See Section 1-IIC) and defect migration effects leads to an additional waveguiding layer located adjacent to the main damage layer. In addition, it will be found that the number of free carriers originally present in these layers is reduced according to the amount of proton-induced damage generated (or remaining after heat treatment). Consequently, because of free carrier polarizability (to be discussed in Section IV), the partially compensated, damaged layers have indices of refraction which are, by differing amounts, greater than that of

the undamaged substrate material. Figure 2-2 depicts the four-media waveguide which results from this combination of circumstances.

Optical modes of the structure shown in Fig. 2-2 will now be solved for, assuming all media are lossless. A perturbation technique, which may be used to find mode attenuation when some or all of the guide media have complex refractive indices, will be developed in Section III.

T.E. Modes:

The wave equation, (2-2), becomes for $\sigma = 0$:

$$\nabla^2 \vec{E} = \mu_o \varepsilon \, \partial^2 \vec{E} / \partial t^2 \, . \tag{2-8}$$

Solutions of Eq. (2-8) are assumed to be of the form:

$$\vec{E}(x,z;t) = \hat{y} \, E_y^{(m)}(x) e^{i(\omega t - \beta_m z)} \tag{2-9}$$

where m is the mode order. Referring to Fig. 2-2, the electric field variation in the x-direction, $E_y^{(m)}(x)$, is taken to be:

$$E_y^{(m)}(x) = \begin{cases} Ae^{qx} & x \leq 0 \\ B \sin(hx) + C \cos(hx) & 0 \leq x \leq a \\ D \sin(\ell x) + E \cos(\ell x) & a \leq x \leq b \\ F e^{-px} & b \leq x \end{cases} \tag{2-10}$$

Substitution of Eqs. (2-9) and (2-10) into Eq. (2-8) yields the following relations:

$$n_1^2 k_o^2 = \beta_m^2 - q^2 \tag{2-11a}$$

$$n_2^2 k_o^2 = \beta_m^2 + h^2 \tag{2-11b}$$

$$n_3^2 k_o^2 = \beta_m^2 + \ell^2 \tag{2-11c}$$

$$n_4^2 k_o^2 = \beta_m^2 - p^2 \tag{2-11d}$$

The boundary conditions are that the tangential field components, $E_y^{(m)}(x)$ and $H_z^{(m)}(x)$, (or, from Eqs. (2-1a) and (2-1b), $E_y^{(m)}(x)$ and $\partial E_y^{(m)}(x)/\partial x$) be continuous. Applying these boundary conditions to Eq. (2-10) yields the following homogenous set of equations for A, B, C, D, and E:

$$A = C \tag{2-12a}$$

$$qA = hB \tag{2-12b}$$

$$B\cdot\sin(ha)+C\cdot\cos(ha)=D\cdot\sin(\ell a)+E\cdot\cos(\ell a) \tag{2-12c}$$

$$hB\cdot\cos(ha)-hC\cdot\sin(ha)=\ell D\cdot\cos(\ell a)-\ell E\cdot\sin(\ell a) \tag{2-12d}$$

$$D\cdot\sin(\ell b)+ E\cdot\cos(\ell b)=F\cdot e^{-pb} \tag{2-12e}$$

$$\ell D\cdot\cos(\ell b)-\ell E\cdot\sin(\ell b)=-pFe^{-pb} \tag{2-12f}$$

Nontrivial values for A, B, C, D, E, and F obtain when the determinant of their coefficients vanishes:

$$\begin{vmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ q & -h & 0 & 0 & 0 & 0 \\ 0 & \sin(ha) & \cos(ha) & -\sin(\ell a) & -\cos(\ell a) & 0 \\ 0 & h\cos(ha) & -h\cos(ha) & -\ell\cos(\ell a) & \ell\sin(\ell a) & 0 \\ 0 & 0 & 0 & \sin(\ell b) & \cos(\ell b) & -e^{-pb} \\ 0 & 0 & 0 & \ell\cos(\ell b) & -\ell\sin(\ell b) & pe^{-pb} \end{vmatrix} = 0 \tag{2-13}$$

This determinant may be reduced to the following characteristic equation:

$$h\left[\frac{q - h\tan(ha)}{h + q\tan(ha)}\right] = -\ell\left[\frac{p - \ell\tan(b-a)\ell}{\ell + p\tan(b-a)\ell}\right] \tag{2-14}$$

Eqs. (2-11) and (2-14) form a set of five equations, the solution of which uniquely determines the five unknowns q, h, ℓ, p, and β_m.

Before continuing, it should be noted that in deriving Eq. (2-14) it was assumed that the E-field variation in the x-direction of both medium II and medium III was sinusoidal, i.e., from Eq. (2-11) that

$$\beta_m^2 < n_2^2k_o^2 \ , \quad n_3^2k_o^2 \quad . \tag{2-15}$$

When either of the conditions stated by Eq. (2-15) is violated, the appropriate waveguide parameter (either h or ℓ from Eqs. (2-11b) and (2-11c)) must be multiplied by $i(=\sqrt{-1})$ in Eq. (2-14) thereby causing the corresponding E-field variation to become exponential.

T.M. Modes:

The solution for the T.M. modes is formally identical to that for the T.E. modes. Boundary conditions require that $H_y^{(m)}(x)$ and $E_z^{(m)}(x)$, or $\frac{1}{\varepsilon}\,\partial H_y^{(m)}(x)/\partial x$, be continuous. Accordingly, the characteristic equation for the T.M. modes is:

$$\frac{h}{n_2^2}\left[\frac{(q/n_1^2) - (h/n_2^2)\tan(ha)}{(h/n_2^2) - (q/n_1^2)\tan(ha)}\right] = -\frac{\ell}{n_3^2}\left[\frac{(p/n_4^2) - (\ell/n_3^2)\tan[(b-a)\,\ell]}{(\ell/n_3^2) + (p/n_4^2)\tan[(b-a)\,\ell]}\right] \tag{2-16}$$

The simultaneous solution of Eqs. (2-11) and (2-16) uniquely determines the T.M. field distribution; those comments made above regarding (2-15) hold for the T.M. case also.

Equations (2-14) and (2-16) may be checked for accuracy by considering the limiting situations described in Table 2-1.

Table 2-1

Limit	T.M.	T.E.
$\varepsilon_3 \to \varepsilon_2$	$\tan(bh) = \dfrac{(qn_2^2/hn_1^2) + (pn_2^2/hn_4^2)}{1 - (qn_2^2/hn_1^2)(pn_2^2/hn_4^2)}$	$\tan(bh) = \dfrac{q/h + p/h}{1 - (q/h)(p/h)}$
$b \to a$	$\tan(ah) = \dfrac{(qn_2^2/hn_1^2) + (pn_2^2/hn_4^2)}{1 - (qn_2^2/hn_1^2)(pn_2^2/hn_4^2)}$	$\tan(ah) = \dfrac{q/h + p/h}{1 - (q/h)(p/h)}$
$a \to 0$	$\tan(b\ell) = \dfrac{(qn_3^2/\ell n_1^2) + (pn_3^2/\ell n_4^2)}{1 - (qn_3^2/\ell n_1^2)(pn_3^2/\ell n_4^2)}$	$\tan(b\ell) = \dfrac{q/\ell + p/\ell}{1 - (q/\ell)(p/\ell)}$

In all instances, the characteristic equations which result are those of the familiar three-media structure (See Appendix 2A.). Unfortunately, unlike the three-media equations given in Table 2-1, Eqs. (2-14) and (2-16) do not yield to a simple graphical analysis [17]; the field distributions of the four media structure must be solved numerically by computer. An approximate analysis may, however, be made of the four media structure near second order mode cutoff, a propagation regime wherein Eqs. (2-11) and (2-14) (only T.E. modes will be considered) may be used in conjunction with experimentally observed mode profiles to demonstrate the internal consistency of the waveguide model which has been chosen.

Figure 2-3 shows the development of the T.E. second order mode profile of sample 2B (See Table 4-1) as a function of isochronal annealing. As the mode approaches cutoff, two features of its behavior are to be noted: (1) the height of the interior lobe begins to exceed that of the surface lobe and (2) the interior lobe becomes broader than the surface lobe. As will be shown, this behavior may be reconciled with that of a four media waveguide whose interior waveguiding layer (medium III of Fig. 2-2) is of lower refractive index than that of its surface waveguiding layer (medium II of Fig. 2-2). As mentioned in Chapter 1 a structure such as this may be expected to result, at least in part, from the effects of channeled protons.

We begin an analysis of the shape and extrema of the second order mode profile by using Eqs. (2-12) to cast Eq. (2-10) into a different form:

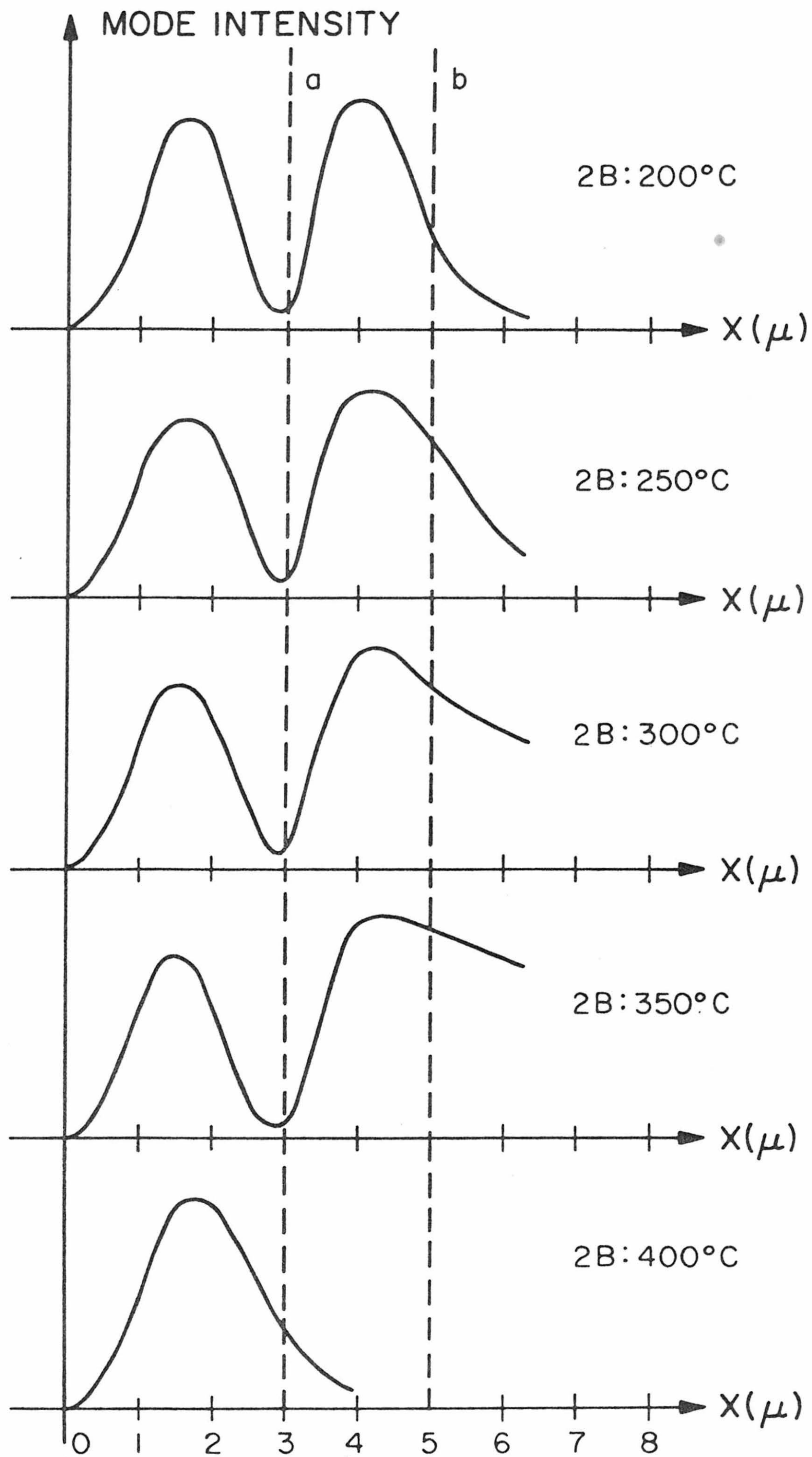


Figure 2-3. Mode profiles of sample 2B as a function of isochronal annealing (0.5 hour anneal at each temperature). The values which have been assigned the parameters a and b as well as symbol definitions will be explained in Chapter 4.

$$E_y^{(m)}(x) = \begin{cases} B_m(\frac{h}{q} e^{qx}) & x \leq 0 \\ B_m[\sin(hx)+(h/q)\cos(hx)] & 0 \leq x \leq a \\ B_m\{[\sin(ha)+(h/q)\cos(ha)]\cos(x-a)\ell \\ \quad + (h/\ell)[\cos(ha)-(h/q)\sin(ha)]\sin(x-a)\ell\} & a \leq x \leq b \\ B_m\{[\sin(ha)+(h/q)\cos(ha)]\cos(b-a)\ell \\ \quad + (h/\ell)[\cos(ha)-(h/q)\sin(ha)]\sin(b-a)\ell\}e^{-p(x-b)} & b \leq x \end{cases} \tag{2-17}$$

The points x_{II} and x_{III} at which extrema of $E_y^{(2)}(x)$ occur may be found by differentiating the appropriate field distributions given in Eq. (2-17) and setting the results equal to zero. Accordingly, the extremal points are solutions of:

$$\tan hx_{II} = \frac{q}{h} \tag{2-18a}$$

$$\tan \ell(x_{III}-a) = \frac{h}{\ell} \frac{\cos(ha) - (h/q)\sin(ha)}{\sin(ha) + (h/q)\cos(ha)} \tag{2-18b}$$

Simplified representations of the quantities of interest, $E_y^{(2)}(x_{II})$ and $E_y^{(2)}(x_{III})$, may be obtained by first deriving two inequalities involving the ratios h/q and ℓ/q . We begin with the derivation involving h/q . A manipulation of Eqs. (2-11) yields:

$$h^2 + q^2 = (n_2^2 - n_1^2)k_o^2 \qquad (2\text{-}19a)$$

$$h^2 + p^2 = (n_2^2 - n_4^2)k_o^2 \qquad (2\text{-}19b)$$

Near cutoff, $p \approx 0$, corresponding to infinite or near-infinite mode penetration into medium IV, the substrate. Taking $p = 0$, Eq. (2-19b) gives for the maximum possible value of h^2 , $(n_2^2 - n_4^2)k_o^2$. Simultaneously, Eq. (2-19a) gives for the minimum possible value of q^2 , $(n_4^2 - n_1^2)k_o^2$. As will be seen in Section IV, the plasma depression effect causes $(n_2^2 - n_4^2)$ to be on the order of 10^{-2} (for $\lambda_o \sim 1\mu$) Assuming, for the present, that the plasma effect is the operative waveguiding mechanism and taking air to be the waveguide overlayer (medium I), the following inequality therefore emerges:

$$(n_4^2 - n_1^2)k_o^2 = 10.5\ k_o^2 \gg (n_2^2 - n_4^2)k_o^2 \approx 10^{-2}k_o^2 \qquad (2\text{-}20)$$

where n_2 , the refractive index of gallium arsenide, has been taken equal to 3.4 for wavelengths of $\sim 1\mu$. Therefore, from Eqs. (2-19) and (2-20) it may be concluded that:

$$h^2/q^2 \ll 1 \quad . \qquad (2\text{-}21a)$$

And by identical reasoning that:

$$\ell^2/q^2 \ll 1 \quad . \qquad (2\text{-}21b)$$

Referring to Fig. 2-3, it is to be observed that the second order mode approaches zero near $x = a$. This fact suggests that

$$ha \approx \pi \quad . \tag{2-22}$$

Using Eq. (2-22), Eq. (2-18b) may be approximated by:

$$\tan \ell(x_{III} - a) \approx q/\ell \quad . \tag{2-23}$$

Equations (2-21a) and (2-21b) may now be used in conjunction with Eqs. (2-18a) and (2-23) to deduce that

$$hx_{II} \approx \pi/2 \tag{2-24a}$$

and that

$$\ell(x_{III} - a) \approx \pi/2 \quad . \tag{2-24b}$$

Finally, Eqs. (2-22), (2-24a), and (2-24b) may be used together with Eq. (2-17) to obtain the mode extrema:

$$\begin{aligned} E_y^{(2)}(x_{II}) &= B_2[\sin(hx_{II}) + (h/q)\cos(hx_{II})] \\ &\approx B_2 \end{aligned} \tag{2-25a}$$

$$\begin{aligned} E_y^{(2)}(x_{III}) &= B_2\Big\{[\sin(ha) + (h/q)\cos(ha)]\cos\ell(x_{III}-a) \\ &\quad + (h/\ell)[\cos(ha) - (h/q)\sin(ha)]\sin\ell(x_{III}-a)\Big\} \\ &\approx B_2\cdot(h/\ell) \end{aligned} \tag{2-25b}$$

As indicated by Fig. 2-3, the interior maximum, $E_y^{(2)}(x_{III})$, of the second order mode exceeds the surface maximum, $E_y^{(2)}(x_{II})$, by an amount which increases as the mode approaches cutoff. From Eqs. (2-25a) and (2-25b), this observation leads to the conclusion that $h > \ell$. Moreover, subtraction of Eq. (2-11c) from Eq. (2-11b)

yields

$$\ell^2 = h^2 - (n_2^2 - n_3^2)k_o^2 \quad . \tag{2-26}$$

From Eq. (2-26) and the fact that $h > \ell$ near cutoff, the deduction is made that $n_2 > n_3$ near cutoff. This result will be seen in Chapter 4 to be consistent with the general features of the ancillary layer (region III) as they relate to those of the primary damage layer (region II) of proton-implanted gallium arsenide. Finally, the internal consistency of the semi-empirical approach which has been used is demonstrated by the agreement between predicted and observed interior lobe broadening: based on the fact that $h > \ell$ near cutoff, Eq. (2-17) predicts that the field distribution in region III will be broader than the field distribution in region II. (Further comparisons of experimental and theoretical mode profiles will be made in Chapter 4.)

III. Loss Calculations - The Perturbation Technique

The wave equation, for the case of structural variation in the x-direction only, is given by:

$$\nabla^2\vec{F}_j(x,z;t) = +\mu_o[\varepsilon_r(x) - i\varepsilon_i(x)]\partial^2\vec{F}_j(x,z;t)/\partial t^2 \tag{2-27}$$

where $\vec{F}_j$ is either the T.E. (transverse electric) field amplitude, $\vec{E}_y$, or the T.M. (transverse magnetic) field amplitude, $\vec{H}_y$, and the dielectric constant, $\varepsilon(x)$, has been taken to be imaginary. Treating $-i\mu_o\varepsilon_i(x)$ as a perturbation, the solution of the unperturbed

wave equation, with $\varepsilon_i(x) = 0$, is given by the eigenmode set:

$$\vec{F}_j^{(m)}(x,z;t) = \hat{j}\ F_j^{(m)}(x)e^{i(\omega t-\beta_m z)} \qquad (2\text{-}28)$$

where $\vec{F}_j^{(m)}(x,z;t)$ represents the mth mode of polarization $\hat{j}$. The eigenmode set (2-28) includes both radiation modes (unconfined) and guided modes (confined) [18] which together form a complete orthogonal set. Perturbations to be considered here will be of the "dc" type, viz. without variation in the z-direction. As will be seen shortly, under these circumstances the confined and unconfined eigenmodes of a waveguide remain uncoupled. Since only guided-mode losses will be of interest, field expansions involving the set (2-28) will therefore contain guided mode terms only.

Members of the set (2-28) may be normalized according to the power they carry in the positive z-direction. Taking the power to be 1 watt per unit width in the y-direction, the normalization conditions become:

$$\delta_{m,n} = \frac{1}{2}\,\mathrm{Re}\left[\int_{-\infty}^{\infty} (-E_y^{(m)}\cdot\ (H_x^{(n)})^*)dx\right] = \frac{\beta_m}{2\omega\mu_o}\int_{-\infty}^{\infty} |E_y^{(m)}(x)|^2 dx \qquad (2\text{-}29a)$$

for the T.E. modes and

$$\delta_{m,n} = \frac{1}{2}\,\mathrm{Re}\left[\int_{-\infty}^{\infty}(H_y^{(m)}\cdot\ (E_x^{(n)})^*)dx\right] = \frac{\beta_m}{2\omega}\int_{-\infty}^{\infty}\frac{|H_y^{(m)}(x)|^2}{|\varepsilon_r(x)|}\ dx \qquad (2\text{-}29b)$$

for the T.M. modes. $\delta_{m,n}$ is the Kronecker delta,

$$\delta_{m,n} = \begin{cases} 1 & m = n \\ 0 & m \neq 0 , \end{cases} \tag{2-30}$$

and * denotes the complex conjugate.

Under the assumption that the perturbative term $-i\mu_0\varepsilon(x)$ has a small effect on the mode structure, the general solution of Eq. (2-27) may be written as a sum over the set (2-28):

$$\vec{F}_j(x,z;t) = \hat{j} \sum_m C_m(z) F_j^{(m)}(x) e^{i(\omega t - \beta_m z)} . \tag{2-31}$$

Here, $|C_m(z)|^2$ represents the power carried by the mth mode in the positive z-direction per unit length in the y-direction.

Substituting Eq. (2-31) into Eq. (2-27) and making use of the normalization conditions, Eqs. (2-29), yields the following system of coupled equations for the expansion coefficients, $C_m(z)$:

$$\frac{\partial^2 C_m}{\partial z^2} - 2i\beta_m \frac{\partial C_m}{\partial z} = K_m(z) \tag{2-32}$$

where, depending on whether T.E. or T.M. modes are being considered,

$$K_m(z) = \begin{cases} + \dfrac{i\omega\beta_m}{2} \left[\sum\limits_n C_n(z) e^{i(\beta_m - \beta_n)z} \displaystyle\int_{-\infty}^{\infty} \varepsilon_i(x) E_y^{(m)} \cdot (E_y^{(n)})^* dx \right] & \text{T.E.} \\ + \dfrac{i\omega\mu_o\beta_m}{2} \left[\sum\limits_n C_n(z) e^{i(\beta_m - \beta_n)z} \displaystyle\int_{-\infty}^{\infty} \frac{\varepsilon_i(x)}{\varepsilon_r(x)} H_y^{(m)} \cdot (H_y^{(n)})^* dx \right] & \text{T.M.} \end{cases} \tag{2-33}$$

The driving term, $K_m(z)$, contains many non-synchronous members which vary as $\exp[i(\beta_m-\beta_n)z]$, $(m \neq n)$. Keeping the single synchronous T.E. and T.M. driving terms $(m = n)$ and assuming that the expansion coefficients vary slowly with z (i.e., that $|\partial^2 C_m/\partial z^2| << |2i\beta_m \partial C_m/\partial z|$) , Eqs. (2-32) and (2-33) together yield:

$$C_m(z) = \begin{cases} \exp\left[-\frac{\omega}{4} \int_{-\infty}^{\infty} \varepsilon_i(x) |E_y^{(m)}(x)|^2 dx \right] z & \text{T.E.} \\ \exp\left[-\frac{\omega\mu_o}{4} \int_{-\infty}^{\infty} \frac{\varepsilon_i(x)}{\varepsilon_r(x)} |H_y^{(m)}(x)|^2 dx \right] z & \text{T.M.} \end{cases} \qquad (2\text{-}34)$$

Substituting the expressions for $C_m(z)$ into Eq. (2-31), the (power) attenuation coefficients are recognized as being:

$$\alpha_{T.E.}^{(m)} = -\frac{\omega}{2} \int_{-\infty}^{\infty} \varepsilon_i(x) |E_y^{(m)}(x)|^2 dx \qquad (2\text{-}35a)$$

$$\alpha_{T.M.}^{(m)} = -\frac{\omega\mu_o}{2} \int_{-\infty}^{\infty} \frac{\varepsilon_i(x)}{\varepsilon_r(x)} |H_y^{(m)}(x)|^2 dx \quad . \qquad (2\text{-}35b)$$

Eqs. (2-35) may be checked by considering the attenuation coefficient of an infinite plane wave propagating in a homogeneous medium. In this case ε_i = constant and can be moved outside the integrals of Eqs. (2-35). Using the same normalization conditions,

that were derived for the case of confined modes, the result for both T.E. and T.M. polarization states is

$$\alpha_{T.E.M.} = \frac{K_i k_o}{n} \tag{2-36}$$

where $\varepsilon = \varepsilon_r - i\varepsilon_i = \varepsilon_o(K_r - iK_i) = \varepsilon_o(n - ik)^2$. Equation (2-36) is identical to Eq. (2-7) which was derived for the case of a T.E.M. wave propagating in a homogeneous medium of infinite extent.

As was mentioned earlier the perturbation technique is valid only when the introduction of loss does not substantially modify the eigenmodes calculated in the absence of loss. The exponential coefficients in Eqs. (2-34) represent imaginary, first-order corrections to the unperturbed eigenvalues, β_m, of the unperturbed wave equation. As long as these corrections are much less than β_m, the original criterion for validity is satisfied. Therefore, the integrals of Eq. (2-35) should remain small, requiring either that $\varepsilon_i(x)$ be small when mode penetration into the lossy region is large or that mode penetration be small into the region where $\varepsilon_i(x)$ is large.

Another approach which may be used to calculate waveguide losses involves solving the imaginary wave equation, (2-27), directly. The introduction of a single lossy medium forces all guide parameters to become imaginary in this case and the exact solution for the modes becomes extremely complicated, even for simple geometries. In Appendix 2A, results obtained using this direct method are compared with those obtained using the perturbation technique for two geo-

metries which satisfy the conditions enumerated in the preceding paragraph: a metallic cladding for which ε_i is large but mode penetration is small and a lossy substrate for which ε_i is small but mode penetration is large.

IV. The Kramers-Kronig Relations

Losses encountered in thin film optical systems can arise from several mechanisms depending on the geometry and materials of the configuration under consideration. Transitions by free carriers within a conduction band can occur either in the guiding region or in the cladding regions which surround it (e.g., metallic overlayer or semiconducting substrate). Band tailing and defect-associated energy levels generated, for example, by ion implantation may cause material which was previously transparent at a given wavelength to become absorptive. Finally, scattering, either intentional (coherent coupling from periodic structures [19]) or unintentional (incoherent volume or surface scattering) may contribute to the optical attenuation.

The presence of loss (or gain) in a material causes its dielectric constant to become complex. If the loss is generated by absorption (as opposed to scattering), the real and imaginary parts of the dielectric constant may be related to each other through the Kramers-Kronig relationship. In Chapter 3, two types of absorp-

tion, free-electron and defect-associated, will be encountered which may be dealt with using this relationship. In anticipation of their need we therefore proceed with a derivation of the Kramers-Kronig relations.

As Eq. (2-5) indicates, the dielectric constant is a complex quantity. Noting, in addition, that ε_r and σ are dispersive as well, the dielectric constant may be written as

$$K(\omega) = K_r(\omega) - iK_i(\omega) \quad . \tag{2-37}$$

Cauchy's integral formula states that the values of an analytic function, (in our case $K(\omega)$ [20]) along a closed contour in a region in which the function remains regular uniquely determine the values of the function at any point within the contour [21]. In particular:

$$K(\omega) = \frac{1}{2\pi i} \int_C \frac{K(\omega')}{\omega' - \omega} \, d\omega' \tag{2-38}$$

Substituting Eq. (2-37) into Eq. (2-38), assuming that ω is real, and performing the required integration yields:

$$K_r(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega' K_i(\omega')}{(\omega')^2 - \omega^2} \, d\omega' \tag{2-39a}$$

$$K_i(\omega) = - \frac{2\omega}{\pi} P \int_0^\infty \frac{K_r(\omega')}{(\omega')^2 - \omega^2} \, d\omega' \tag{2-39b}$$

where it has been recognized that the free space value of $K_r = 1$ and the Cauchy principal value, P , has been taken prior to integrating. Equations (2-39) are known as the Kramers-Kronig relations.

Using Eqs. (2-5) and (2-36), a relation equivalent to that given by (2-39a) may be obtained which relates the real part of the refractive index to the absorption coefficient [22]:

$$n(E) - 1 = \frac{ch}{2\pi^2} P \int_0^\infty \frac{\alpha(E')}{(E')^2 - E^2} dE \qquad (2\text{-}40)$$

where h = Plank's constant, c = the speed of light in vacuo, and $E = h\nu$, the photon energy. Using Eq. (2-40), one can, in principle, determine the refractive index at a given photon energy if the absorption is known for all photon energies.

V. Absorption Mechanisms

A. Free Carrier Intraband Transitions

Quantum mechanics predicts that free carriers moving in a perfectly periodic crystal lattice cannot absorb energy. Thermal effects, defects, and impurities, however, disrupt the lattice periodicity enough to allow the scattering of electrons from one energy state in the conduction band to another [23]. Accordingly the motion of an electron in the presence of an applied field, $E_o e^{i\omega t}$, may be described as follows:

$$m^* \frac{d^2x}{dt^2} + m^* g \frac{dx}{dt} = -e\, E_o e^{i\omega t} \tag{2-41}$$

where g is the damping coefficient, m^* is the effective mass and e is the electronic charge.

The steady-state solution of Eq. (2-41) is

$$x = \frac{e\, E_o/m^*}{\omega^2 - i\omega g} e^{i\omega t} \tag{2-42}$$

Taking the resulting ac polarization to be $P = -Nex$, the complex dielectric constant becomes:

$$K = K_r - i\, K_i = n^2 - \frac{Ne^2/(m^* \varepsilon_o)}{\omega^2 - i\omega g} \tag{2-43}$$

or:

$$K_r = n^2 - \frac{Ne^2/(m^* \varepsilon_o)}{\omega^2 - g^2} \tag{2-44a}$$

$$K_i = \frac{Ne^2 g/(m^* \omega \varepsilon_o)}{\omega^2 - g^2} \tag{2-44b}$$

where N is the number of free carriers cm^{-3} and n is the refractive index in the absence of free carriers. Equations (2-44) can be shown to obey the Kramers-Kronig relationship [24].

In the presence of a steady electric field, $m^* g(dx/dt) = eE$. Also, by definition of the mobility, μ, $dx/dt = \mu E$. Therefore, one has for the damping factor,

$$g = e/\mu m^*. \tag{2-45}$$

Using Eq. (2-45), Eqs. (2-44a) and (2-44b) therefore become:

$$K_r \approx n^2 - \frac{Ne^2}{m^* \varepsilon_o \omega^2} \tag{2-46a}$$

$$K_i \approx \frac{Ne^3}{(m^*)^2 \varepsilon_o \omega^3 \mu} \tag{2-46b}$$

where it has been assumed that, for those cases of interest, $\omega >> g$. The free carrier-associated attenuation coefficient is, therefore:

$$\alpha_{f.c.} = \frac{k_o K_i}{n} = \frac{Ne^3}{(m^*)^2 n \varepsilon_o \omega^2 \mu} \quad . \tag{2-46c}$$

It is to be noted from Eq. (2-46a) that the presence of free carriers lowers the dielectric constant (the so-called plasma depression effect). This mechanism will be seen in Chapter 4 to be responsible for optical waveguiding in proton-implanted GaAs.

B. Transitions Involving Discrete States and Bands

The quantum electrodynamical Hamiltonian for an electron in the presence of an electromagnetic field is given by

$$H = H_e + H_r + H_I \tag{2-47}$$

where H_e, H_r, and H_I, the electronic, radiation, and interaction energy terms, respectively, are

$$H_e = \hbar^2 \nabla^2 / 2m^* + V(\vec{r}) \tag{2-48a}$$

$$H_r = \sum_\lambda \hbar\omega_\lambda (a_\lambda^+ a_\lambda^- + \tfrac{1}{2}) \tag{2-48b}$$

$$H_I = e\vec{E}\cdot\vec{r} \quad . \tag{2-48c}$$

Here $\vec{E}$ is the optical electric field vector and $\vec{r}$ is the position

vector of an electron; a_λ^+ and a_λ^- are creation and annihilation operators, respectively, which possess the following commutation and operational properties

$$[a_\ell^-, a_m^-] = [a_\ell^+, a_m^+] = 0; \quad [a_\ell^-, a_m^+] = \delta_{\ell,m} \qquad (2\text{-}49)$$

$$a_\lambda^+ |n_\lambda\rangle = \sqrt{n_\lambda + 1} \; |n_\lambda + 1\rangle \qquad (2\text{-}50a)$$

$$a_\lambda^- \; |n_\lambda\rangle = \sqrt{n_\lambda} \; |n_\lambda + 1\rangle \qquad (2\text{-}50b)$$

$$a_\lambda^+ a_\lambda^- |n_\lambda\rangle = n_\lambda |n_\lambda\rangle \quad . \qquad (2\text{-}50c)$$

$|n_\lambda\rangle$ in Eqs. (2-50) is a radiation state vector whose occupation number is n_λ.

The electric field, $\vec{E}$, may be expressed as a sum over radiation modes, λ, and polarization directions, $\hat{j}$, in the following way [20]:

$$\vec{E} = -2i \sqrt{\frac{\pi\hbar\omega}{V\varepsilon_r}} \sum_{j,\lambda} \sin\left(\frac{\omega_\lambda \sqrt{\varepsilon_r}}{c} z\right) (a_\lambda^+ - a_\lambda^-)\hat{j} \qquad (2\text{-}51)$$

where the direction of propagation has been chosen to be parallel to the z=axis; V is the volume occupied by the modes and ε_r is the real part of the (in general) complex dielectric constant. Substituting Eq. (2-51) into Eq. (2-48c), the interaction Hamiltonian becomes:

$$H_I = -2ie \sqrt{\frac{\pi\hbar\omega}{V\varepsilon_r}} \sum_{j,\lambda} \sin\left(\frac{\omega_\lambda \sqrt{\varepsilon_r}}{c} z\right) (a_\lambda^+ - a_\lambda^-)\hat{j}\cdot\vec{r} \qquad (2\text{-}52)$$

The total absorption and emission rates per unit volume may now be obtained by applying Fermi's golden rule to the transitions induced by H_I. Using the relations given by Eqs. (2-50a) through (2-50c) in order to evaluate the appropriate matrix elements, they are:

$$W_{ab} = \frac{8e^2\pi^2}{V\varepsilon_r} \sum_{j,\lambda} \sum_{u,\ell} \omega_\lambda n_\lambda |\langle u|\sin\left(\frac{\omega_\lambda\sqrt{\varepsilon_r}}{c} z\right) \hat{j}\cdot\vec{r}|\ell\rangle|^2 \delta(E_u - E_\ell - \hbar\omega_\lambda) \quad (2\text{-}53a)$$

$$W_{em} = \frac{8e^2\pi^2}{V\varepsilon_r} \sum_{j,\lambda} \sum_{u,\ell} \omega_\lambda (n_\lambda+1) |\langle u|\sin\left(\frac{\omega_\lambda\sqrt{\varepsilon_r}}{c} z\right) \hat{j}\cdot\vec{r}|\ell\rangle|^2 \delta(E_\ell - E_u + \hbar\omega_\lambda) \quad (2\text{-}53b)$$

where summations now extend over all upper, $|u\rangle$, and lower, $|\ell\rangle$, electronic states differing in energy by $\hbar\omega_\lambda$. $\delta(x-x_o)$ in Eqs. (2-53) is the dirac delta function

$$\int_a^b \delta(x-x_o)dx = \begin{cases} 1 & a \leq x_o \leq b \\ 0 & \text{otherwise} \end{cases} \quad . \quad (2\text{-}54)$$

Equations (2-53) may be made more compact by performing the following manipulation [25]: taking D_u to be the number of states per unit volume at energy E_u and D_ℓ to be the number of states per unit volume at energy E_ℓ, the summations over u and ℓ in Eqs. (2-53) are converted into sums over energies by making the following definition:

$$\sum_{u,\ell} |\langle u|\sin\left(\frac{\omega_\lambda\sqrt{\varepsilon_r}}{c} z\right) \hat{j}\cdot\vec{r}|\ell\rangle|^2 = \sum_{E_u,E_\ell} D_u D_\ell \langle |H_{u,\ell}|^2\rangle_{av} \, . \quad (2\text{-}55)$$

where $<|H_{u,\ell}|^2>_{av}$ represents the strength of the dipole transition averaged over transitions between degenerate upper and lower energy levels. Letting $P_i (i = u,\ell)$ be the probability that a state is occupied and P_i' be the probability that it is vacant, Eq. (2-55) may now be used to recast Eqs. (2-53) into the form:

$$W_{ab} = \frac{8e^2\pi^2}{V\varepsilon_r} \sum_{j,\lambda} \sum_{E_u,E_\ell} \omega_\lambda n_\lambda n_\ell n_u' <|H_{u,\ell}|^2>_{av} \delta(E_u - E_\ell - \hbar\omega_\ell) \qquad (2\text{-}56a)$$

$$W_{em} = \frac{8e^2\pi^2}{V\varepsilon_r} \sum_{j,\lambda} \sum_{E_u,E_\ell} \omega_\lambda (n_\lambda+1) n_\ell' n_u <|H_{u,\ell}|^2>_{av} \delta(E_\ell - E_u + \hbar\omega_\lambda) \qquad (2\text{-}56b)$$

where:

$$n_\ell = D_\ell P_\ell \qquad (2\text{-}57a)$$

$$n_\ell' = D_\ell P_\ell' \qquad (2\text{-}57b)$$

$$n_u = D_u P_u \qquad (2\text{-}57c)$$

$$n_u' = D_u P_u' \quad . \qquad (2\text{-}57d)$$

Furthermore, specializing to the case of singly-polarized, monochromatic radiation of angular frequency ω, Eqs. (2-56a) and (2-56b) become:

$$W_{ab}(\hbar\omega) = \frac{8e^2\pi^2}{V\varepsilon_r} \omega n_\omega \sum_{E_u,E_\ell} n_\ell n_u' <|H_{u,\ell}|^2>_{av} \delta(E_u - E_\ell - \hbar\omega) \qquad (2\text{-}58a)$$

$$W_{em}(\hbar\omega) = \frac{8e^2\pi^2}{V\varepsilon_r} \omega(n_\omega+1) \sum_{E_u,E_\ell} n_\ell' n_u <|H_{u,\ell}|^2>_{av} \delta(E_\ell - E_u + \hbar\omega) \quad . \qquad (2\text{-}58b)$$

$<|H_{u,\ell}|^2>_{av}$ in Eqs. (2.58) when evaluated using Bloch wavefunctions for the initial and final states, imposes the familiar conservation of momentum conditions:

$$k_u = k_\ell + \frac{\omega}{c} \simeq k_\ell \quad \text{(for absorption)} \qquad (2\text{-}59a)$$

$$k_\ell = k_u - \frac{\omega}{c} \simeq k_u \quad \text{(for emission)} \quad . \qquad (2\text{-}59b)$$

Equations (2-58a) and (2-58b) are then evaluated using a so-called reduced density of states function [25]. When the periodicity of the crystal lattice has been significantly perturbed, however, either by degenerate impurity or degenerate defect populations, the resulting electronic states become admixtures of those of the previously perfect structure and the k-selection rules of Eqs. (2-59) no longer apply. Under these circumstances, $<|H_{u,\ell}|^2>_{av}$ becomes independent of the photon energy, $\hbar\omega$, and may be moved outside the summations of Eqs. (2-58). Two types of transitions allow the use of this procedure: (1) interband transitions wherein band-tailing caused by lattice perturbations becomes important [26] and (2) transitions between discrete levels wherein one or both of the levels lies too far from a band edge to be described adequately by a single Bloch function.

Two examples of non-momentum conserving transitions, one involving transitions between the ground and excited states of vacancy complexes and the other involving transitions between disorder-induced bandtails and the band continua, will be encountered in Chapter 3. The products of density of states functions needed to calculate these transitions are conveniently enumerated within the present context. Accordingly, we

have for transitions between two discrete levels which are assumed to have arisen from monovalent charge centers (e.g., transitions between singly-ionized donor and acceptor levels):

$$n_\ell n'_u = \left(\frac{g_\ell N_\ell \delta(E_\ell - E_A)}{1 + (1/g_\ell)e^{(E_\ell - F_p)/kT}}\right)\left(\frac{g_u N_u \delta(E_u - E_D)}{1 + g_u e^{(F_n - E_u)/kT}}\right) \tag{2-60a}$$

$$n'_\ell n_u = \left(\frac{N_\ell \delta(E_\ell - E_A)}{1 + g_\ell e^{(F_p - E_\ell)/kT}}\right)\left(\frac{N_u \delta(E_u - E_D)}{1 + (1/g_u)e^{(E_u - F_n)/kT}}\right) . \tag{2-60b}$$

And for transitions between continuous energy level distributions:

$$n_\ell n'_u = \left(\frac{\rho_\ell(E_\ell)}{1 + e^{(E_\ell - F_p)/kT}}\right)\left(\frac{\rho_u(E_u)}{1 + e^{(F_n - E_u)/kT}}\right) \tag{2-61a}$$

$$n'_\ell n_u = \left(\frac{\rho_\ell(E_\ell)}{1 + e^{(F_p - E_\ell)/kT}}\right)\left(\frac{\rho_u(E_u)}{1 + e^{(E_u - F_n)/kT}}\right) . \tag{2-61b}$$

Here, P_i has been replaced by the Fermi-Dirac distribution function. F_p is the quasi-Fermi level for holes; F_n is the quasi-Fermi level for electrons; E_A and E_D are acceptor and donor ionization energies, respectively; g_ℓ and g_u are degeneracy factors; and N_ℓ and N_u are the number of charge centers per unit volume. In addition, it has been recognized, in deriving Eqs. (2-60), that when ionized, the number of participating states is given by $g_\ell N_\ell$ (or $g_u N_u$) and, when neutral, the number is given by N_ℓ (or N_u) .

The total emission and absorption transition rates to be associated with the products of density of states functions given by Eqs.

(2-60a) through (2.-61b) may be found by converting sums to integrals in Eqs. (2-56a) and (2-56b):

$$W_{ab}(\hbar\omega) = \frac{8e^2\pi^2}{V\varepsilon_r}\omega n_\omega \langle|H_{u,\ell}|^2\rangle_{av} \int_{E_u}\int_{E_\ell} n_\ell(E_\ell)n_u'(E_u)\delta(E_u-E_\ell-\hbar\omega)dE_u dE_\ell \quad (2\text{-}62a)$$

$$W_{em}(\hbar\omega) = \frac{8e^2\pi^2}{V\varepsilon_r}\omega(n_\omega+1)\langle|H_{u,\ell}|^2\rangle_{av} \int_{E_u}\int_{E_\ell} n_\ell'(E_\ell)n_u(E_u)\delta(E_\ell-E_u+\hbar\omega)dE_u dE_\ell \quad (2\text{-}62b)$$

The absorption coefficient, $\alpha(\hbar\omega)$, is given by the power absorbed per unit volume divided by the power crossing a unit area. The power per unit area (energy flux) is

$$I_\omega = \frac{c\hbar\omega}{\sqrt{\varepsilon_r}}\left(\frac{n_\omega}{V}\right) . \quad (2\text{-}63)$$

Therefore, using Eqs. (2-62a) and (2-62b):

$$\alpha(\hbar\omega) = \frac{\hbar\omega[W_{ab}(\hbar\omega) - W_{st}(\hbar\omega)]}{I_\omega}$$

$$= \frac{8e^2\pi^2\omega}{\sqrt{\varepsilon_r}C}\langle|H_{u,\ell}|^2\rangle_{av}\left\{\int_E [n_\ell(E-\hbar\omega)n_u'(E)-n_\ell'(E-\hbar\omega)n_u(E)]dE\right\} \quad (2\text{-}64)$$

where:

$$W_{em}(\hbar\omega) = W_{st}(\hbar\omega) + W_{sp}(\hbar\omega) = (n_\omega+1)W_{sp}(\hbar\omega) \quad (2\text{-}65)$$

$W_{st}(\hbar\omega)$ is the stimulated emission rate and $W_{sp}(\hbar\omega)$ is the spontaneous emission rate.

Appendix 2A

I. Metallic Overlayer Losses of a Three-Media Guide

A. Perturbation Technique Solution

The unperturbed eigenmodes of the three media structure shown in Fig. 2A-1a may be obtained by substituting Eq. (2-28) into Eq. (2-27) with $\varepsilon_i(x) = 0$. Using the normalization conditions (2-29) and the appropriate boundary conditions, they are given by

T.E.:

$$E_y^{(m)}(x) = \begin{cases} A_m \cdot \frac{h}{q} \exp(-qx) & 0 \geq x \geq -\infty \\ A_m \cdot [\frac{h}{q} \cos(hx) - \sin(hx)] & t \geq x \geq 0 \\ A_m \cdot [\frac{h}{q} \cos(ht) + \sin(ht)] \exp[p(x+t)] & \infty \geq x \geq t \end{cases} \quad (2A-1)$$

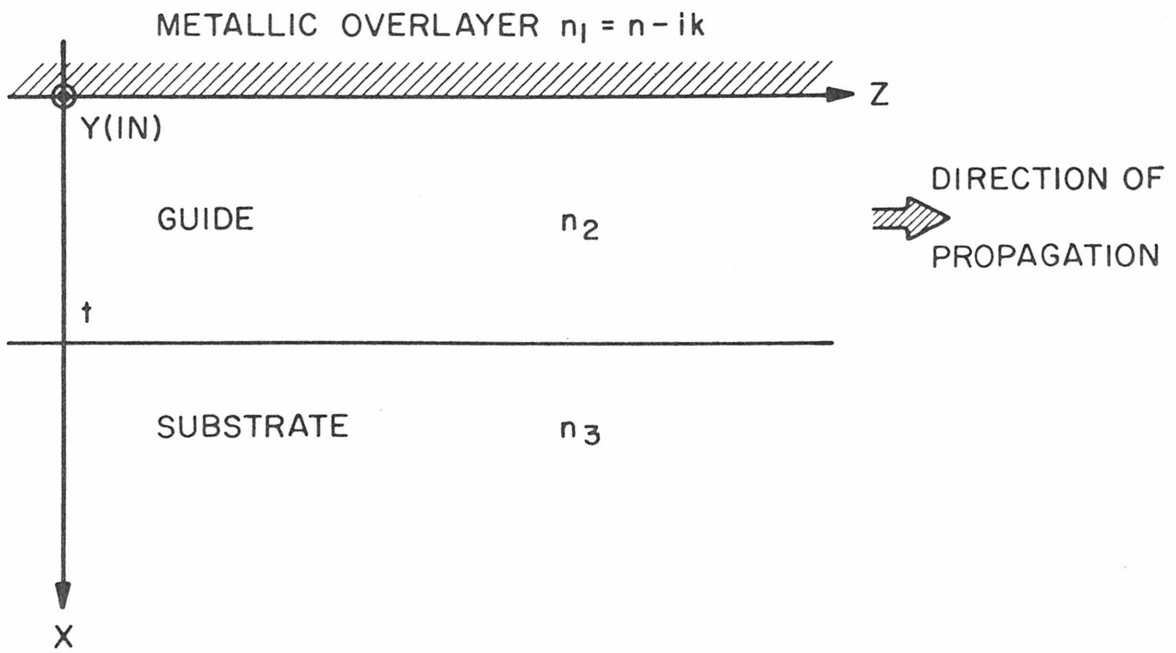
where:

$$A_m = 2 \left\{ \frac{\omega\mu_o}{\beta_m (t + \frac{1}{q} + \frac{1}{p})(1 + \frac{h^2}{q^2})} \right\}^{1/2} \quad (2A-2)$$

T.M.:

$$H_y^{(m)}(x) = \begin{cases} B_m \cdot \frac{h}{\bar{q}} \exp(-qx) & 0 \geq x \geq -\infty \\ B_m \cdot [\frac{h}{\bar{q}} \cos(hx) - \sin(hx)] & t \geq x \geq 0 \\ B_m \cdot [\frac{h}{\bar{q}} \cos(ht) + \sin(ht)] \exp[p(x+t)] & \infty \geq x \geq t \end{cases} \quad (2A-3)$$

where [27]:



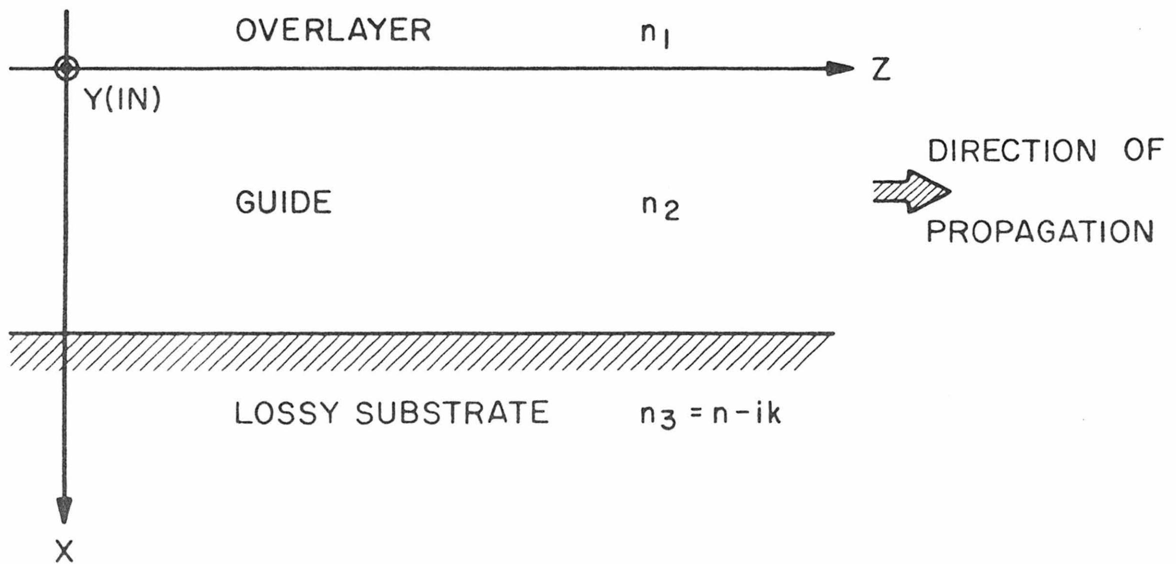


Figure 2A-1. (a) Three media guide with lossy metallic overlayer;
(b) Three media guide with lossy substrate.

$$B_m = 2\sqrt{\frac{\omega\varepsilon_o}{\beta_m t_{eff}}} \tag{2A-4}$$

t_{eff}, $\bar{p}$, and $\bar{q}$ in Eqs. (2A-3) and (2A-4) are defined as follows:

$$t_{eff} \equiv \frac{\bar{q}^2 + h^2}{\bar{q}^2}\left[\frac{t}{n_2^2} + \frac{q^2 + h^2}{\bar{q}^2 + h^2}\,\frac{1}{n_1^2 q} + \frac{p^2 + h^2}{\bar{p}^2 + h^2}\,\frac{1}{n_3^2 p}\right] \tag{2A-5a}$$

$$\bar{p} \equiv (n_2^2/n_3^2)p \tag{2A-5b}$$

$$\bar{q} \equiv [n_2^2/(n^2 - k^2)]q \quad . \tag{2A-5c}$$

Using only the real part of $n_1^2 (\equiv n-ik)^2$ to solve for the unperturbed waveguide modes, Eqs. (2A-1) and (2A-3), when substituted into the wave equation, yield the following relations for both the T.E. and the T.M. modes:

$$(n^2-k^2)k_o^2 = \beta_m^2 - q^2 \tag{2A-6a}$$

$$n_2^2\,k_o^2 = \beta_m^2 + h^2 \tag{2A-6b}$$

$$n_3^2\,k_o^2 = \beta_m^2 - p^2 \quad . \tag{2A-6c}$$

Equations (2A-6) together with the following eigenvalue equations (see Table 2-1) uniquely determine the T.E. and T.M. field distributions:

<u>T.E.</u>:

$$\tan(ht) = \frac{h(p+q)}{h^2 - pq} \tag{2A-7a}$$

<u>T.M.</u>:

$$\tan(ht) = \frac{h(\bar{p}+\bar{q})}{h^2 - \bar{p}\cdot\bar{q}} \tag{2A-7b}$$

The attenuation coefficients of the three-media structure may now be evaluated by using Eqs. (2A-1) through (2A-5) in Eqs. (2-35a) and (2-35b):

$$\alpha_{TE}^{(m)} = -\frac{\omega}{2}\int_0^\infty \varepsilon_i(x)|E_y^{(m)}(x)|^2 dx = -\frac{nkk_o^2}{\beta_m(t+\frac{1}{q}+\frac{1}{p})(1+\frac{h^2}{q^2})}\frac{h^2}{q^3} \qquad (2A\text{-}8a)$$

$$\alpha_{TM}^{(m)} = -\frac{\omega\mu_o}{2}\int_0^\infty \frac{\varepsilon_i(x)}{\varepsilon_r(x)}|H_y^{(m)}(x)|^2 dx = -\frac{nk\,k_o^2}{\beta_m t_{eff}(n^2-k^2)}\frac{h^2}{q\cdot\bar{q}^2} \qquad (2A\text{-}8b)$$

where $\varepsilon_1(x) \equiv \varepsilon_r(x) - i\varepsilon_i(x) \equiv \varepsilon_o n_1^2 \equiv \varepsilon_o(n^2-k^2) - 2i\varepsilon_o nk$.

For well-confined modes, the following limiting conditions apply:

$$\frac{1}{p} \rightarrow 0 \qquad (2A\text{-}9a)$$

$$\frac{1}{q} \rightarrow 0 \qquad (2A\text{-}9b)$$

$$h \rightarrow (m+1)\frac{\pi}{t} \qquad (2A\text{-}9c)$$

$$\beta_m \rightarrow n_2 k_o \quad . \qquad (2A\text{-}9d)$$

As discussed in Section 2-II, for $\lambda \simeq 1\mu$, $q >> h$. Moreover, when $|n_1^2| >> n_2^2$ and $k^2 >> n^2$ (see Table 2A-1), $q \simeq k\cdot k_o$. Under these circumstances, the attenuation coefficient for well-confined T.E. modes becomes:

$$\alpha_{TE}^{(m)} = -\frac{n}{2n_2k^2}\left(\frac{m+1}{2}\right)^2\frac{\lambda^2}{t^3} \quad . \qquad (2A\text{-}10a)$$

If, in addition to those conditions just mentioned, $|\bar{q}| >> |h^2|$ (a condition which is not always met for wavelengths approaching 10μ and certain metallic overlayers (see Table 2A-1)), t_{eff} may be approximated by t/n_2^2 , yielding for the attenuation coefficient of well-confined T.M. modes:

TABLE 2A-1

λ(microns)	n	k
0.5	0.5	6.0
1.0	1.8	10.0
10.0	25.0	67.0

The real and imaginary parts of the complex dielectric constant ($n = n - ik$) for gold [28].

$$\alpha_{TM}^{(m)} = -\frac{n}{2n_2^3}\left(\frac{m+1}{2}\right)^2 \frac{\lambda^2}{t^3} \quad . \tag{2A-10b}$$

B. Solution of the Complex Wave Equation [29]

Equations (2A-7) may be rewritten in the form

$$ht = \tan^{-1}[(p/h)\eta_{23}] + \tan^{-1}[(q/h)\eta_{21}] + m\pi$$

$$(m = 0,1,2,\cdots) \tag{2A-11}$$

where $\eta_{23} = \eta_{21} = 1$ for T.E. modes and $\eta_{ij} = n_i^2/n_j^2$ for T.M. modes. Furthermore, retaining both the real and the imaginary parts of n_1^2, equations (2A-6) may be recast as:

$$n_1^2 k_o^2 = \beta_m^2 - q^2 \tag{2A-12a}$$

$$n_2^2 k_o^2 = \beta_m^2 + h^2 \tag{2A-12b}$$

$$n_3^2 k_o^2 = \beta_m^2 - p^2 \tag{2A-12c}$$

where, since n_1^2 is a complex quantity, all guide parameters become complex and, as a matter of notation, will be written in the form $h = h_r + ih_i$.

In order to make Eq. (2A-11) tractable, it is necessary to assume that:

$$|(q/h)\eta_{21}| \gg 1 \tag{2A-13a}$$

$$n_2^2/n_3^2 \simeq 1 \quad . \tag{2A-13b}$$

Using Eqs. (2A-13), Eq. (2A-11) may be simplified and separated into real and imaginary parts, yielding:

$$C_{23} = \frac{H_r}{\sin H_r}(-1)^m \tag{2A-14a}$$

$$H_i = \frac{H_r \mathrm{Im}(\frac{1}{C_{21}\eta_{21}}) C_{23}\cos H_r}{(-1)^m - C_{23}\cos H_r} \tag{2A-14b}$$

where $H \equiv ht$, $C_{23} \equiv \sqrt{(n_2^2 - n_3^2)}\, k_o t$ and $C_{21} \equiv \sqrt{(n_2^2 - n_1^2)}\, k_o t$.

Since it is known that, for the mth order mode, $(m+\frac{1}{2})\pi \lesssim H_r \lesssim (m+1)\pi$, Eqs. (2A-14) may be solved by using the following algorithm: an arbitrary value lying between the limits mentioned is assigned to H_r ; C_{23} is solved for using Eq. (2A-14a); knowing H_r and C_{23} , H_i is solved for using Eq. (2A-14b). The imaginary part of β_m is then obtained from Eq. (2A-12b):

$$\begin{aligned} \mathrm{Im}[\beta_m^{TE,TM}] &= - \frac{h_r h_i}{\mathrm{Re}[\beta_m^{TE,TM}]} \\ &= - \frac{\mathrm{Im}\left[\frac{1}{n_2 \eta_{21}\sqrt{n_2^2 - n_1^2}}\right]}{t^3 k_o^2} \frac{H_r^3}{k_o^2[H_r - \tan(H_r)]} \end{aligned} \tag{2A-15}$$

where $\eta_{21} = 1$ (TE); $\eta_{21} = n_2^2/n_1^2$ (TM). Finally, from Eq. (2-6), the attenuation coefficients are given by

$$\alpha^{(m)}_{TE,TM} = 2\ \mathrm{Im}[\beta_m^{TE,TM}]. \tag{2A-16}$$

If it is assumed once again that $|n_1|^2 >> n_2^2$ and that $k^2 >> n^2$, we have for the case of well confined modes (Recalling conditions (2A-9)):

$$\alpha_{TE}^{(m)} = - \frac{n}{2n_2k^2} \left(\frac{m+1}{2}\right)^2 \frac{\lambda^2}{t^3} \qquad (2A\text{-}17a)$$

$$\alpha_{TM}^{(m)} = - \frac{n}{n_2^3} \left(\frac{m+1}{2}\right)^2 \frac{\lambda^3}{t^3} \qquad (2A\text{-}17b)$$

We note that Eqs. (2A-17) are identical to Eqs. (2A-10) which were derived using an entirely different approach. More revealing, however, is a comparison of Eqs. (2A-8) and Eq. (2A-15). In order to solve Eq. (2A-11) and hence to obtain Eq. (2A-15), it was necessary to assume that $n_2^2/n_3^2 \simeq 1$, an approximation not needed to derive Eqs. (2A-8). That the perturbation technique did not require that $n_2^2/n_3^2 \simeq 1$ is not surprising, however, since the ratio n_2^2/n_3^2, by itself, does not determine the degree of mode confinement or penetration into the lossy overlayer--the thickness t being an equally important, independent parameter. From this observation we conclude, therefore that the perturbation technique solution is of a more general nature in situations where the other approximations common to both approaches are valid.

When conditions (2A-13) are not met, an approximate solution for $Im[\beta_m^{TE,TM}]$ may still be obtained if $C_{23}^2 - H_r^2 \gg H_i^2$ [29]. Assuming that this condition has been met, one obtains, after separating the real and imaginary parts of Eq. (2A-11), the following coupled, transcendental equations:

$$H_r = \tan^{-1} \eta_{23} \sqrt{\frac{C_{23}^2}{H_r^2} - 1} + Re[\tan^{-1}(C_{21}\eta_{21}/H_r)] + m\pi$$

$$(m = 0,1,2,\cdots) \qquad (2A\text{-}18a)$$

$$H_i = \frac{\mathrm{Im}[\tan^{-1}(C_{21}\eta_{21}/H_r)]}{1 + f^{-1}(C_{23}^2 - H_r^2)^{-1/2}} \qquad \text{(2A-18b)}$$

where $f = \eta_{23} + (1/\eta_{23} - \eta_{23})H_r^2/C_{23}^2$. After solving Eqs. (2A-18) for H_i, $\mathrm{Im}[\beta_m^{TE,TM}]$ and $\alpha_{TE,TM}^{(m)}$ are obtained from Eqs. (2A-12b) and (2A-16).

We conclude by noting that, since condition (2A-13a) was not required in order to derive Eqs. (2A-18), Eqs. (2A-18) may be used to calculate the optical loss when penetration into the metallic overlayer is large--a situation wherein the perturbation technique is not applicable.

II. Substrate Losses of a Three Media Guide

A. Perturbation Technique Solution

Figure 2A-1b depicts the geometry. The unperturbed field distributions and mode normalizations for $\varepsilon_i(x) = 0$ are identical to those of Section 2A-I. Consequently, setting $\varepsilon_i(x) = 0$ for $x \geq +t$, Eqs. (2-35) immediately yield for the attenuation coefficients:

$$\alpha_{TE}^{(m)} = -\frac{\omega}{2}\int_{-\infty}^{-t} \varepsilon_i(x)|E_y^{(m)}(x)|^2 dx = -\frac{nkk_o^2[\frac{h}{q}\cos(ht) + \sin(ht)]^2}{\beta_m(t + \frac{1}{q} + \frac{1}{p})(1 + \frac{h^2}{q^2})}\frac{1}{p} \qquad \text{(2A-19a)}$$

$$\alpha_{TM}^{(m)} = -\frac{\omega\mu_o}{2}\int_{-\infty}^{-t} \frac{\varepsilon_i(x)}{\varepsilon_r(x)}|H_y^{(m)}(x)|^2 dx = -\frac{nkk_o^2[\frac{\bar{h}}{\bar{q}}\cos(ht) + \sin(ht)]^2}{\beta_m t_{eff}(n^2 - k^2)}\frac{1}{p} \qquad \text{(2A-19b)}$$

where it has been assumed that $n^2 \gg k^2$. This restriction is necessary since mode penetration into the substrate may be large for those guide-substrate index discontinuities of interest.

As the substrate-guide index discontinuity is decreased (or held constant while the guide is made thinner), the optical mode becomes progressively less well confined, i.e., closer to cutoff. In the limit of optical cutoff, substantially all of the energy propagates in the lossy substrate and optical attenuation must approach the bulk, substrate value. Making the assumptions that $|q|$ and $|\bar{q}| >> h$, and that, near cutoff,

$$p \to 0 \tag{2A-20a}$$

$$\beta_m \to nk_o \tag{2A-20b}$$

$$ht \to (\ell + \tfrac{1}{2})\pi \qquad (\ell=0,1,2,\ldots) \tag{2A-20c}$$

we have for the limiting T.E. and T.M. attenuation coefficients:

$$\alpha_{TE}^{(m)} \to \alpha_{TM}^{(m)} \to \frac{k_o K_i}{n} \quad . \tag{2A-21}$$

The attenuation coefficients (2A-21) are, as required, seen to approach the bulk attenuation coefficient given by Eq. (2-7).

In the limit of very good confinement, substrate penetration becomes negligible and the optical attenuation should go to zero. Now:

$$p \to \infty \tag{2A-22a}$$

$$\beta_m \to n_2 k_o \tag{2A-22b}$$

$$ht \to (\ell+1)\pi \qquad (\ell=0,1,2,\cdots) \quad . \tag{2A-22c}$$

Trivially, from Eqs. (2A-19a) and (2A-19b) we find that:

$$\alpha_{TE}^{(m)} \to \alpha_{TM}^{(m)} \to 0 \quad . \tag{2A-23}$$

B. Solution of the Complex Wave Equation

The notation used will be that of Section 2A-IB. In addition, the following definitions are made:

$$n_1^2 = K_1 \tag{2A-24a}$$

$$n_2^2 = K_2 \tag{2A-24b}$$

$$n_3^2 = K_r - iK_i \tag{2A-24c}$$

$$\Delta K = K_2 - K_r \quad . \tag{2A-24d}$$

Once again, in order to make progress, the assumption must be made that $|(q/h)\eta_{21}| >> 1$. (It will be noted that this condition is violated as guided, free-space wavelengths approach 10μ.) Doing so yields:

$$\tan^{-1}[(q/h)\eta_{21}] \simeq \frac{\pi}{2} \ . \tag{2A-25}$$

Using Eq. (2A-25), Eq. (2A-11) becomes:

$$-\cot(ht) \cong \eta_{23}(p/h) \quad . \tag{2A-26}$$

The loss constant, $\mathrm{Im}[\beta_m]$, is understood (in the case of free carrier absorption, for example) to be much smaller than $\mathrm{Re}[\beta_m]$. Hence, from Eq. (2A-12b), it is safe to assume that $h_r >> h_i$. Allowing these approximations, the right hand side of Eq. (2A-26) becomes

$$\eta_{23}(p/h) = \eta_{23}\sqrt{\frac{(\Delta K + iK_i)k_o^2 - h^2}{h^2}}$$

$$\simeq \sqrt{\frac{\Delta K k_o^2 - h_r^2}{h_r^2}}\left[1 - i\,\frac{[-K_i k_o^2 h_r^2 + 2h_r h_i \Delta K k_o^2]}{2(\Delta K k_o^2 - h_r^2)h_r^2}\right] \tag{2A-27a}$$

for $\eta_{23} = 1$ (T.E.) and:

$$\eta_{23}(p/h) \approx \sqrt{\frac{\Delta K k_o^2 - h_r^2}{h_r^2}} \left[1 - i\,\frac{[-K_i h_r + 2h_i \Delta K]k_o^2}{2p_r^2 h_r}\right] \cdot \frac{K_2}{K_r^2}(K_r - iK_i) \quad \text{(2A-27b)}$$

for $\eta_{23} = K_2/(K_r - iK_i)$ (T.M.) . Assuming, as we did above, that $h_i << h_r$, the left hand side of Eq. (2A-26) may be approximated by:

$$- \cot(h_r t + ih_i t) = \frac{\sin(2h_r t) - i \sinh(2h_i t)}{\cosh(2h_i t) - \cos(2h_r t)}$$

$$\approx - \cot(h_r t) + i\,\frac{h_i t}{\sin^2(h_r t)} \quad \text{(2A-28)}$$

Equating the real and imaginary parts of Eqs. (2A-27) and (2A-28) yields:

$$- \cot(h_r t) = \eta_{23}(p_r/h_r) \quad \text{(2A-29)}$$

and:

$$I_m[\beta_m^{T.E.}] = - \frac{K_i k_o}{2p_r n_2}\left[\frac{t}{\sin^2(h_r t)} + \frac{\Delta K k_o^2}{P_r h_r^2}\right]^{-1} \quad \text{(2A-30a)}$$

$$I_m[\beta_m^{T.M.}] = - \frac{K_i k_2}{n_2 K_r k_o}\left[\frac{P_r}{K_r} + \frac{k_o^2}{2p_r}\right] \cdot \left[\frac{t}{\sin^2(h_r t)} + \frac{\Delta K k_o^2}{P_r h_r^2}\right]^{-1} \quad \text{(2A-30b)}$$

where use has been made of the following relations:

$$p_r^2 \approx \Delta K k_o^2 - h_r^2 \quad \text{(assuming that } p_i << p_r) \quad \text{(2A-31a)}$$

and

$$\beta_m^2 = K_2 k_o^2 - h^2 \quad . \tag{2A-31b}$$

Finally, we examine the usual limiting conditions. Near cutoff:

$$\alpha^{(m)}_{T.E.M.} = 2 \cdot Im[\beta_m^{TE,TM}] \to \frac{k_o K_i}{n} \tag{2A-32}$$

where $K = K_r - iK_i = (n - ik)^2$. For the case of well-confined modes, conditions (2A-22) together with Eqs. (2A-29) and (2A-30) yield:

$$I_m[\beta_m^{TE,TM}] \to 0 \quad . \tag{2A-33}$$

These results are seen to be in agreement with those obtained earlier.

We conclude by noting that, in order to derive Eqs. (2A-30), the tacit assumption was made that $\Delta K >> K_i$. As Eq. (2-46b) indicates, this assumption may not be true for the case of free carrier absorption at long wavelengths. It is, of course, also violated when ΔK is small (for example, thin waveguides near cutoff). For these reasons and the obvious one of simplicity, the perturbation technique will henceforth be used to calculate free-carrier-associated waveguide attenuation.

CHAPTER 3

THE OPTICAL AND ELECTRICAL PROPERTIES OF PROTON-IMPLANTED GALLIUM ARSENIDE

I. Introduction

The physical mechanisms responsible for the nature and performance of proton-implanted optical waveguides in gallium arsenide are governed by the damage-induced defect levels which appear within the forbidden energy gap of the irradiated material. In order to develop a model capable of predicting the behavior of such structures it is therefore necessary to gain a thorough understanding of the electronic defect level spectrum of proton-irradiated gallium arsenide. The most convenient way of obtaining the defect level spectrum or defect density-of-states function, and the one which will be used here, is to infer it from various absorption data; as will be shown in a later section, the validity of the spectrum so obtained may be checked by comparing theoretical and experimental curves of free carrier concentration versus defect concentration. Since only those defects which are electrically active (viz., those in which free carriers are trapped), as opposed to those which are strictly dissipative in nature, yield absorption data useful within this context, our analysis will be begun by considering the different types of absorption to be expected under varying implantation conditions.

II. Ion-Implanted GaAs

Radiation-generated disorder in semiconductors may modify the energy level spectrum of the implanted material in either or both of the

following ways, depending on the energy, fluence, and species of projectile ion used: the fundamental absorption edge may suffer an apparent shift to lower energies and/or discrete levels may be introduced into the forbidden gap. When implanted into GaAs, heavy ions such as Xe and Se [30], as well as reactor neutrons of sufficient fluency [31] ($\geq 10^{18} cm^{-2}$), cause band edge tailing which may be approximated over a certain energy range by an E^2 dependence. In addition, implants of this type are known to give rise to at least two discrete energy levels with ionization energies of about 0.2 eV and 0.5 eV [32]. Energetic electrons ($\sim$ 1 MeV), on the other hand, have a less drastic effect on GaAs, producing a smaller edge shift with a nearly exponential energy variation and a single, barely discernible level located 1.3 eV from one of the band edges [32]. Intermediate between these two extremes, the damage caused by 150-400 keV protons, in doses between 10^{14} and $10^{16} cm^{-2}$, has been shown to introduce a continuous bandtailing which begins deep within the band gap [33] and ends with a slight shift of the fundamental edge toward lower energies. The existence of two discrete defect levels in proton-implanted gallium arsenide has been inferred by electrical measurements [34] but has not been substantiated by optical excitation techniques.

III. Proton-Implanted GaAs

Wohlleben and Beck, from measurements made on the carrier removal rates of both n- and p-type GaAs as a function of proton fluence, concluded that quasi-continua of levels exist both just below the conduction band edge and just above the valence band edge [35]. Their

hypothesis is in general agreement with the results of Stein who measured the optical absorption coefficient of proton-implanted GaAs from 0.2 eV to the band edge [33]; he found the absorption spectrum to resemble a structureless, exponential function of energy. Contrary to these findings and to those to be presented below, Pruniaux, North, and Miller inferred the presence of two discrete levels located 0.4 eV and 0.8 eV above the valence band edge based on resistivity versus temperature measurements of n-type GaAs [34]. (It should be mentioned that, with the exception of the data of Wohlleben and Beck who used 3 MeV protons, all of the results quoted are based on the effects of 150-400 keV proton implants.)

The photoresponse of a proton-implanted GaAs detector (to be described in detail in Chapter 5) is shown in Fig. 3-1a together with the response of an unimplanted control sample; Stein's data are shown in Fig. 3-1b for comparison. Different theories, to be discussed below, predict that the bandtailing of ion-damaged material will have either an E^2, $A \exp(BE^2)$ or $C \exp(DE)$ energy dependence. As can be seen in Fig. 3-1, however, no single functional form adequately describes the experimental data throughout the range of energy variation. It is supposed, therefore, that some superposition of the three energy dependencies mentioned, the individual contributions of which being dependent on the dynamics of the implanted ion, will best describe the experimental data.

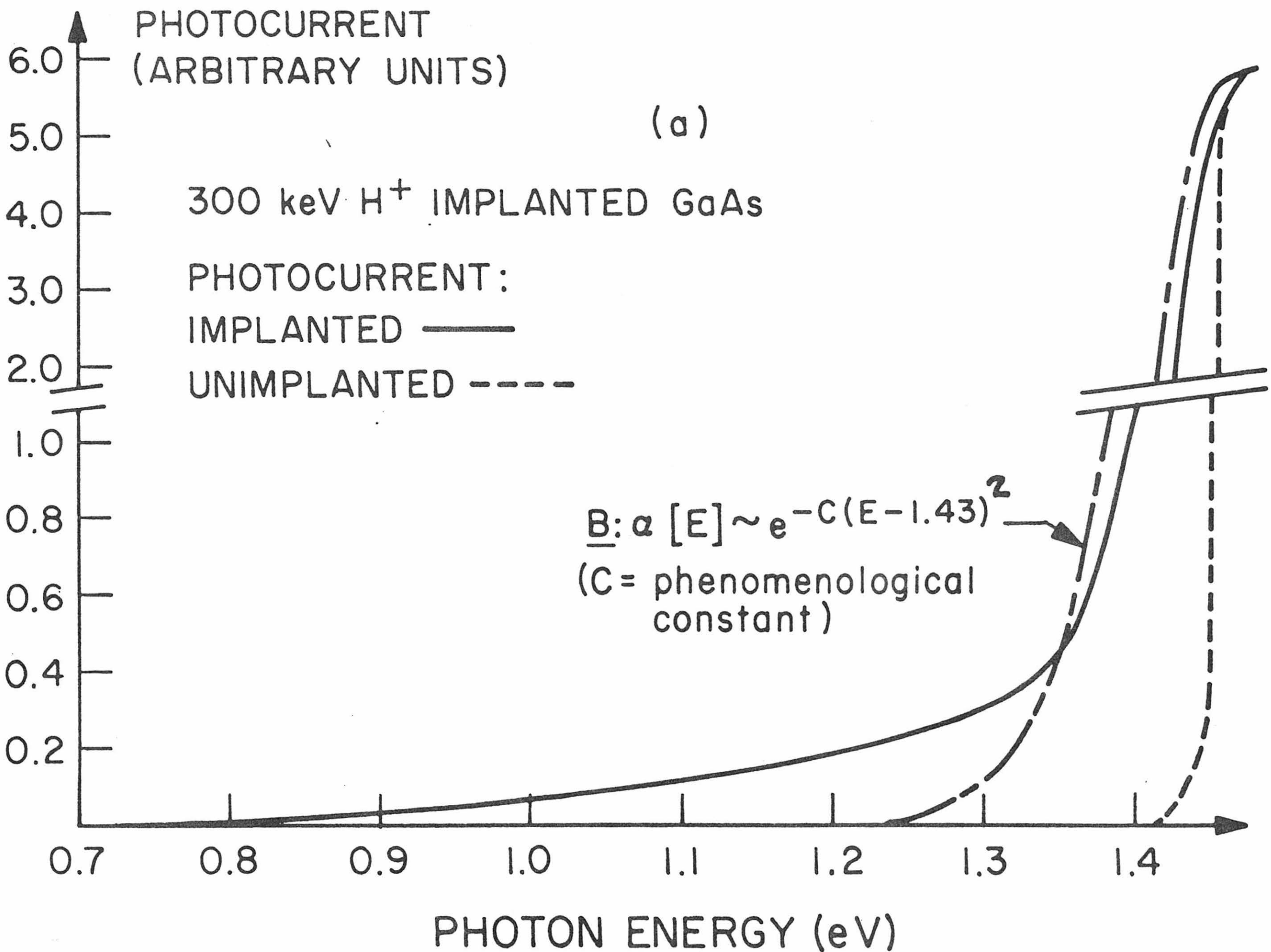


Figure 3-1. (a) Photoresponse of implanted and unimplanted GaAs; curve B is the theoretical excitation spectrum of gallium vacancy-donor complexes.

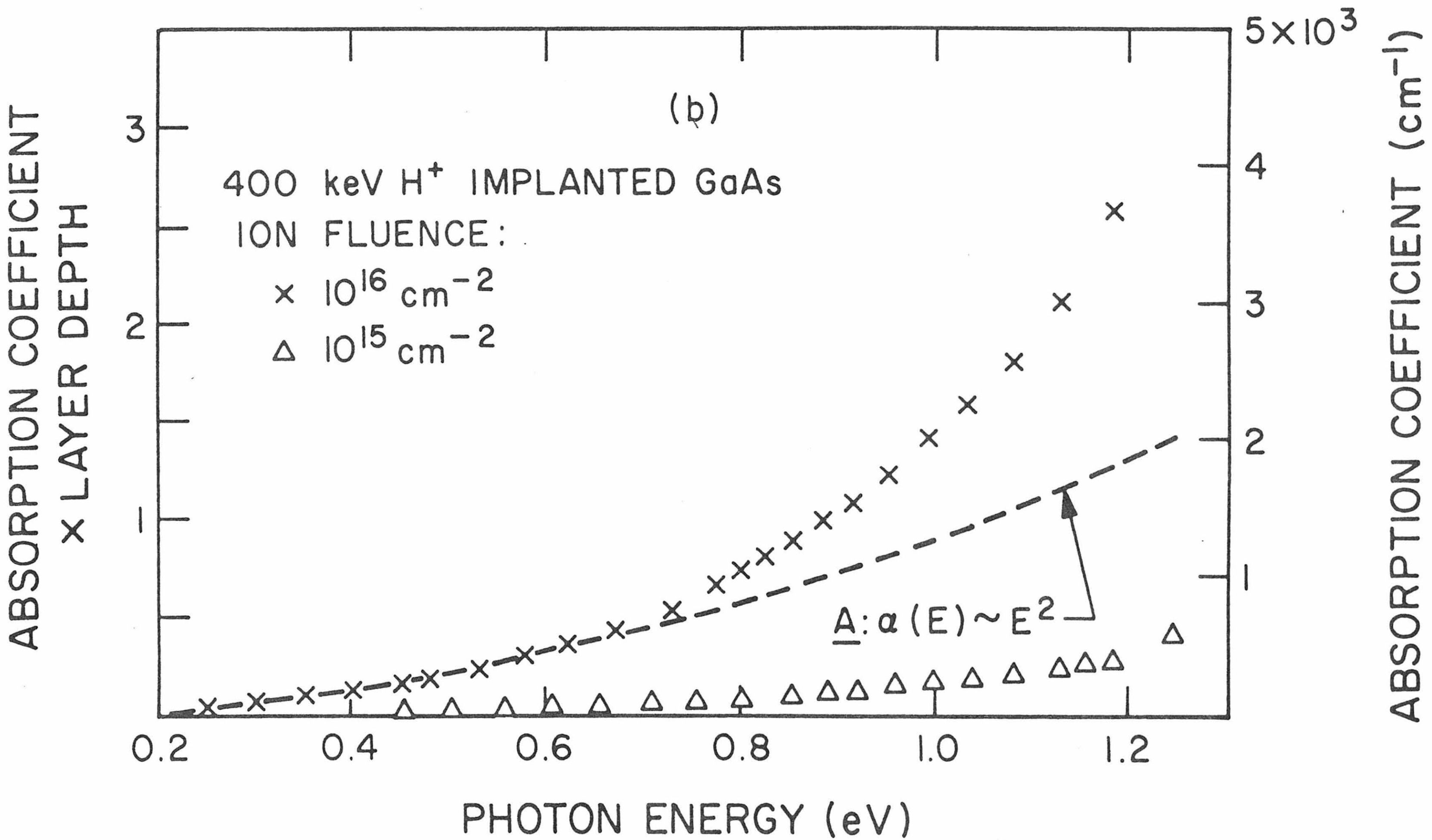


Figure 3-1. (b) Total absorption spectra of implanted GaAs for two different fluences (after Stein [33]); curve A represents the absorption due to thermal spikes.

A. Optical Properties of Proton-Irradiated GaAs

1. Absorption and Scattering by Thermal Spikes

As was mentioned above, large projectile ions such as Xe or Se or energetic ones such as reactor neutrons are known to induce a band edge tail which varies over a certain energy range as the energy squared. Aukerman, et al [36] suggested that fast neutrons incident on GaAs create so-called thermal spikes, or superheated regions in which metallic precipitates are formed. Subsequently, McNichols and Ginell [31] developed a theory based on these observations in which the optical attenuation was attributed to scattering and absorption by small, metallic inhomogeneities. Their expression for the attenuation coefficient (measured in cm^{-1}) for the case of neutron-irradiated GaAs is :

$$\alpha(E) = 2.2 \times 10^{3}(f)E^{2} \qquad (3\text{-}1)$$

where E is measured in electron volts and f is the volume fraction occupied by metallic zones. As Se and Xe produce bandtailing in GaAs which has the form given by (3-1), it is presumed that implants of these species (as well as of other heavy ions) also give rise to thermal spiking. One may logically conclude, therefore, that in order to generate an absorption spectrum which bears an E^2 energy dependence, the dynamics of the implant projectile must be such that localized regions of the target material become highly disordered or converted to a different phase. This situation apparently occurs when the kinetic energy of the incident ion is above a certain threshold. As shown by curve A in Fig. 3-1b, that portion of the absorption spectrum lying

between 0.2 eV and 0.65 eV may be fitted by the functional form of Eq. (3-1), indicating that a fraction of the 400 keV protons used by Stein fulfilled this requirement.

It will be further noted that the photoresponse curve of Fig. 3-1a cuts off sharply near the center of the (unirradiated) band gap at approximately 0.7 eV while the absorption curve obtained by Stein tails to energies less than 0.2 eV. According to the bandtailing model to be presented in a later section, disorder-induced, electrically active trapping centers give rise to absorption only for photon energies greater than about 0.6 eV. Bearing this fact in mind, it may be concluded that the absorption curve of Fig. 3-1b represents a superposition of both strictly dissipative (as discussed here) and electrically active absorption mechanisms over the energy interval between 0.2 eV and 1.3 eV while only electrically active centers leading to photoconductivity are manifested in Fig. 3-1a. The dissipative contribution at 1.08 eV (1.15μ), the energy at which optical attenuation measurements were made here, represents approximately 60% of the total loss and, as will be seen in Chapter 5, imposes a limit on the maximum quantum efficiency to be expected from a detector based on the induced long-wavelength photoconductivity of proton-implanted gallium arsenide.

2. Absorption by Impurity-Defect Complexes

Much experimental evidence exists in support of the contention that n-type gallium arsenide, either grown under conditions which favor the creation of vacancies or subjected to treatment, such as

irradiation, which forcibly introduces vacancies, contains so-called vacancy-donor complexes. As will be shown below, the shoulder, or shift toward lower energies, of the fundamental edge of proton-irradiated gallium arsenide closely matches the excitation spectrum of gallium vacancy-donor complexes, thereby lending additional credence to this supposition. Moreover, in the presence of an externally applied electric field [37], the photoluminescence of such centers may be quenched, leading to induced photoconductivity. As indicated in Fig. 3-1a, this effect appears to be manifested by a shift toward lower energies of the photoresponse curve of irradiated GaAs.

Photoluminescence measurements of structurally imperfect GaAs reveal emission spectra which can, in principle, be attributed to defect levels within the bandgap. Such measurements made on n-type GaAs show a broad emission centered at about 1.2 eV whenever the likelihood of gallium vacancies has been enhanced during sample preparation [38]. Conversely, this band is substantially reduced in intensity or completely absent if measures are taken to discourage the formation of gallium vacancies. It has been observed that the 1.2 eV emission peak is almost always present if the n-type dopant is one of the group VI elements [39]; heavy doping with a group VI element, which is substitutional on arsenic sites, should reduce the arsenic vacancy population and, hence, increase the gallium vacancy concentration since the product of the two vacancy concentrations must remain constant in thermal equilibrium. Group IV dopants, on the other hand, which primarily occupy gallium sites (for n-type GaAs), are known to lead to a

reduction in the intensity of the 1.2 eV emission band [39], all other conditions being the same (for example, the respective impurity concentrations). Finally, the 1.2 eV band is not present whenever the GaAs has been doped heavily by either cadmium or zinc [40] (both substitutional on gallium sites) or whenever it has been grown from a gallium-rich melt [41].

Owing to the wide range of projectile parameters and target histories (largely unspecified) chosen by independent investigators, attempts to form a causal relationship between photoluminescence spectra and defects generated by elementary particle irradiation have been less successful than those aimed at linking growth conditions to the 1.2 eV emission line. Mitchell and Norris [42], using reactor neutrons, and Jeong et al [43], using electrons, found that both the 1.2 eV (Mitchell and Norris measured it to be 1.28 eV) and the band edge emission of n-type GaAs were partially quenched by the irradiations. Tkachev et al [44], using 50 eV electrons, and Harris and Eisen [45], using 450 keV protons, observed an increase in an emission centered at 1.35 eV and an accompanying decrease in the band edge luminescence of n-type GaAs after bombardment; both of these investigations attributed the emission to arsenic vacancies or arsenic vacancy-impurity complexes. Harris and Eisen based their conclusion on the fact that this emission (which could also be induced by annealing an unirradiated GaAs specimen) could be inhibited by depositing a layer of SiO_2 on the sample prior to heat treatment -the SiO_2 overlayer presumably preventing an out-diffusion of arsenic atoms. In opposition to these findings, Tuck noted [46] that by <u>increasing</u> the arsenic overpressure during heat treatment and,

therefore decreasing the number of arsenic vacancies (or increasing the number of gallium vacancies in thermal equilibrium), the intensity of the 1.2 eV line of n-type GaAs was made two to three times greater (temperature = 1000°C, As pressure = 0.1 atm).

If gallium vacancies, which are presumed to act as acceptors in the complexing and compensation processes, are responsible for the 1.2 eV emission of n-type material, then one would expect arsenic vacancies to play a similar role for the case of p-type GaAs. This has been observed to be the case: the group II impurities, Zn and Cd, which are substitutional on gallium sites have been shown to produce a complex line near 1.37 eV [47]. Based on these observations, which appear to confirm the supposition that gallium vacancies are responsible for the defect spectra of n-type GaAs, and those of Williams [39] who measured spectra of n-type material doped by six different impurity species (S, Si, Ge, Se, Sn, and Te) (see Fig. 3-2) it may be concluded that gallium vacancies are positively charged and give rise to an emission centered near 1.2 eV.

Self-activated luminescent centers in ZnS [48], formed by the association of a zinc vacancy with a negatively charged impurity ion, are exactly analogous to the gallium vacancy-impurity comples presently under consideration. The latter center is depicted in Fig. 3-3. Since the configurational-coordinate (CC) model was so successful in predicting the behavior of ZnS luminescent centers, it will also be applied here. The one electron CC diagram is shown in relation to the bandgap of GaAs in Fig. 3-4.

The gallium vacancy-impurity complex may be regarded as a molecular-like entity bound together by the coulombic attraction between a

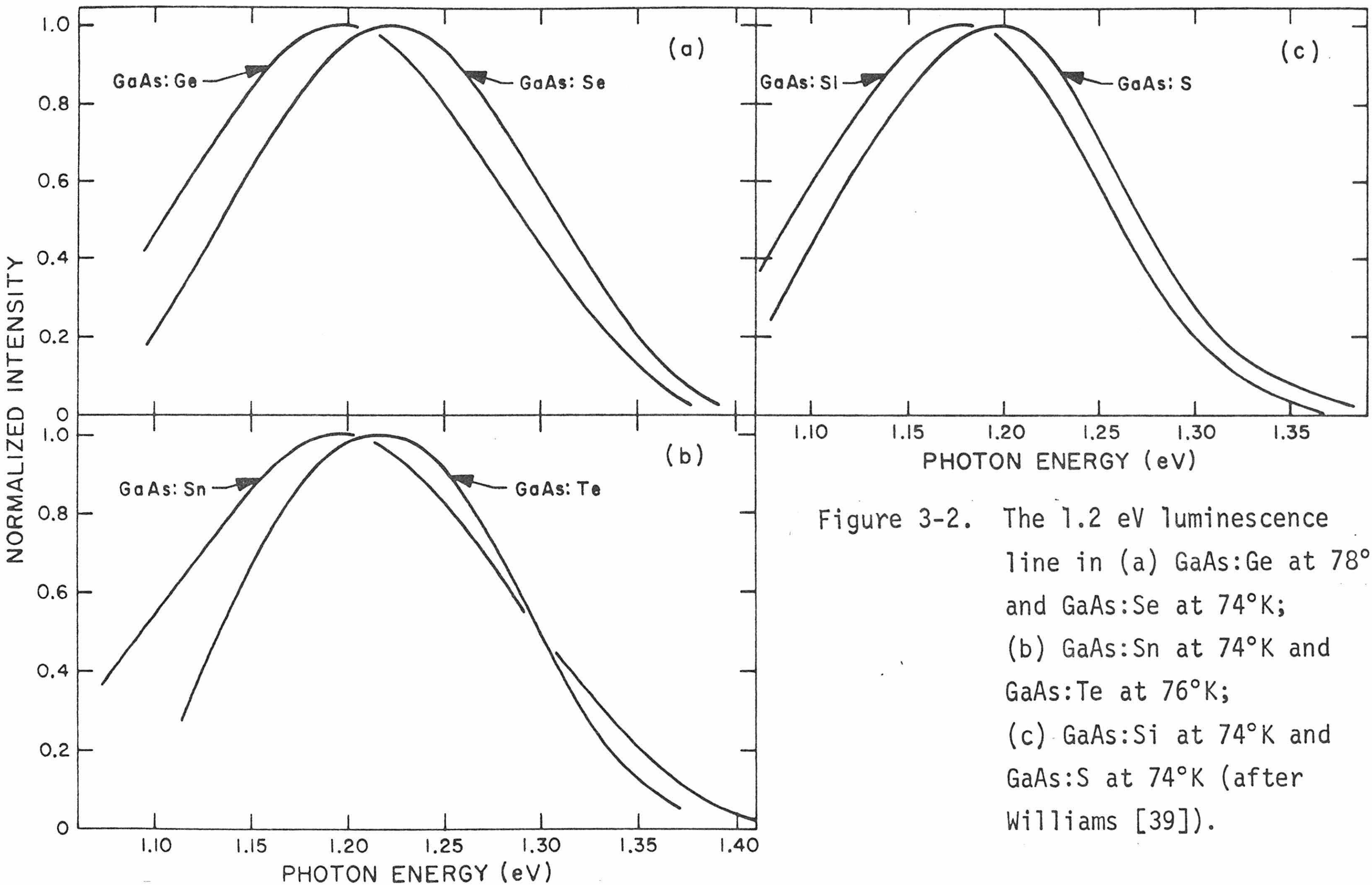


Figure 3-2. The 1.2 eV luminescence line in (a) GaAs:Ge at 78°K and GaAs:Se at 74°K; (b) GaAs:Sn at 74°K and GaAs:Te at 76°K; (c) GaAs:Si at 74°K and GaAs:S at 74°K (after Williams [39]).

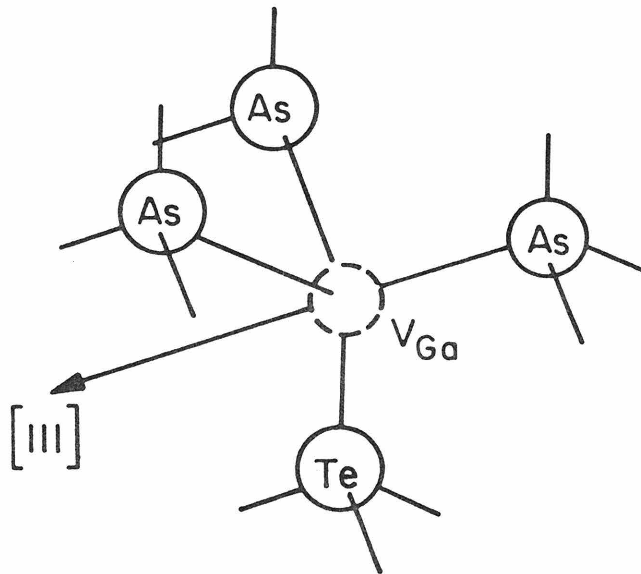


Figure 3-3. The gallium vacancy-donor (in this case tellurium) complex. (After Williams [39])

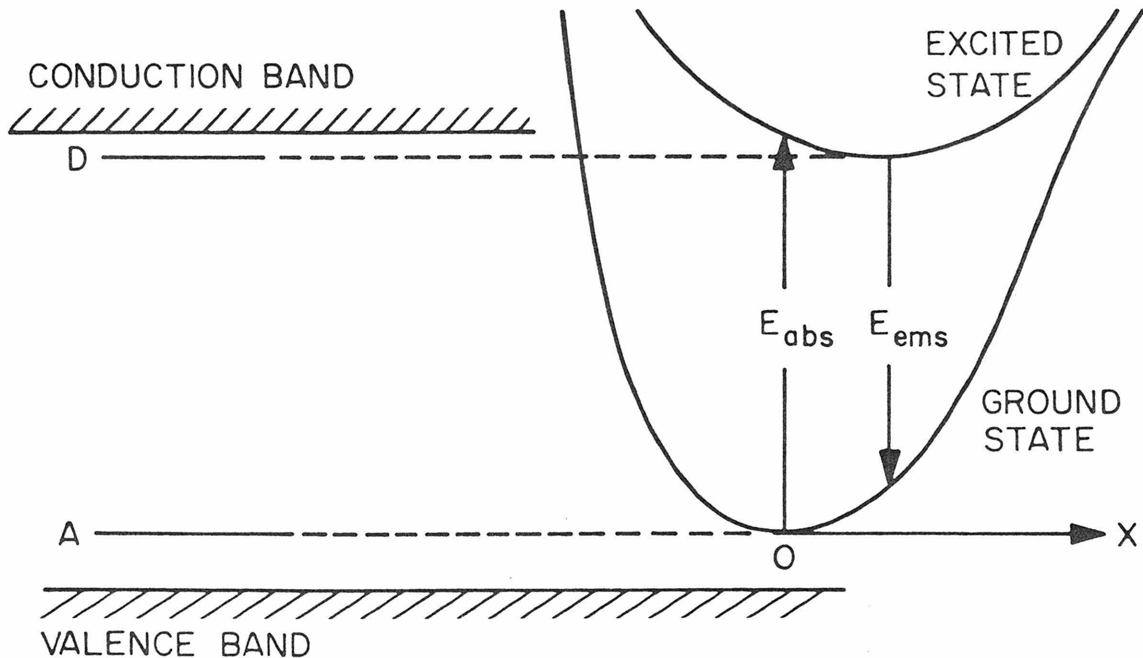


Figure 3-4. The one electron configurational coordinate model for vacancy complexes in GaAs. Also shown is the band gap and donor-like (D) and acceptor-like (A) levels of the zero-point energies of the ground and excited states, respectively, which lie within the band gap. E_{abs}, the absorption energy is greater than E_{ems}, the emission energy, and the difference is the Stokes shift. The electron transitions are vertical because of the Franck-Condon principle. (After Williams [39]).

positively charged vacancy and a negatively-charged donor atom. Accordingly, it may be treated as an harmonic oscillator whose ground and excited electronic state, vibrational potential energies are given by :

$$E_g = \frac{1}{2} K_g x^2 \qquad (3\text{-}2a)$$

$$E_e = \frac{1}{2} K_e (x-x_o)^2 + E_o \qquad (3\text{-}2b)$$

where $K_{g,e}$ are the effective spring constants and, x, the configurational coordinate, represents the instantaneous separation of charge centers. The minima of the ground and excited states are taken to be the gallium vacancy level and the donor level, respectively. It will be noted, however, that because of the coulombic attraction which exists between donor and "acceptor", these levels will lie somewhat further from the band edges than would be the case for isolated charge centers. Moreover, when the center is in its excited state, it is expected that the degree of molecular binding will be modified slightly; as reflected in Eqs. (3-2a) and (3-2b), the modifications thereby induced are a small change, x_o, in the equilibrium separation and a different spring constant, K_e. In addition, because the equilibrium separation between charge centers is different for the excited and ground states, the absorption energy, E_{abs}, is greater than the emission energy, E_{ems}, by an amount termed the Stokes shift [49]. This situation is most easily visualized at 0°K in the classical approximation as shown in Fig. 3-4. Since the bonding between atoms in GaAs is covalent, however, the displacement between minima, x_o, and, therefore, the Stokes shift, should be small. Finally, since electronic transitions between the ground end excited states are

assumed to occur much faster than the complex can assume a different configuration (Franck-Condon principle [49]), they are taken to be vertical as depicted in Fig. 3-4.

Ignoring the effects of stimulated emission, a valid assumption considering the weak excitation fields used here, a modified form of Eq. (2-64) may be used to describe absorption by complex centers:

$$\alpha_c(\hbar\omega) = \frac{8e^2\pi\omega}{\sqrt{\varepsilon_r}c} \int_x P(x) \quad \langle|H_{u,\ell}|^2\rangle_{av} \quad n_\ell(E_\ell)n_u'(E_u)\delta(E_u - E_\ell - \hbar\omega)dx \qquad (3\text{-}3)$$

where the range of integration is now taken over all values of the configuration coordinate. $P(x)$ in Eq. (3-3) is the probability that a particular configuration is assumed and is given classically by [49]:

$$P(x) = (K_g/2\pi kT)^{1/2} \exp(-K_g x^2/ 2kT) \qquad (3\text{-}4)$$

Equation (3-4) is correct in the quantum-mechanical case if T is replaced by an effective temperature [50]:

$$T_{eff} = \theta \coth(\theta/T) \qquad (3\text{-}5)$$

where:

$$\theta = h\nu_g/2k \quad . \qquad (3\text{-}6)$$

Here h is Planck's constant and ν_g is the natural frequency of the ground state using the simple harmonic oscillator model. The matrix element $\langle|H_{u,\ell}|^2\rangle_{av}$ in Eq. (3-3) must be evaluated for initial and final states which are products of molecular and electronic state functions; it will be assumed to vary slowly with x compared with $P(x)$ (which is

strongly peaked even at room temperature) and therefore will be moved outside the integral of Eq. (3-3).

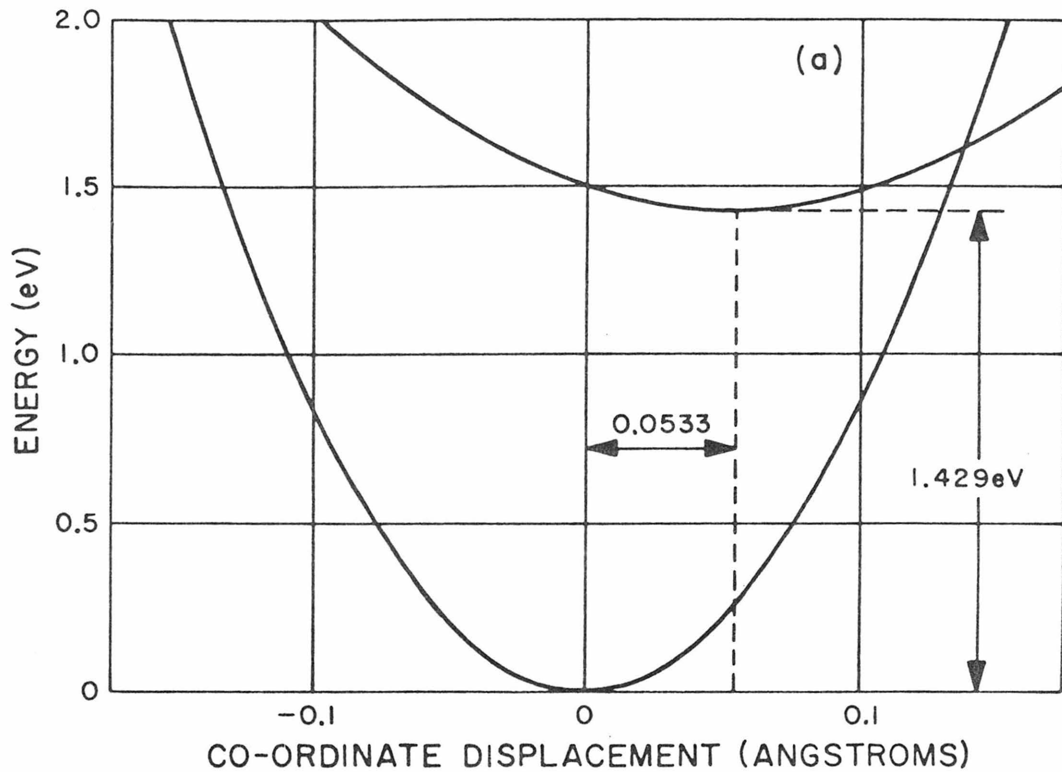
Following the method of Klick and Schulman [49], Williams and White [51] have determined the CC parameters at $80^{o}K$ for the case of Ge-doped GaAs to be:

$$
\begin{aligned}
h\nu_e &= 0.022 \text{ eV} \\
h\nu_g &= 0.040 \text{ eV} \\
E_o &= 1.429 \text{ eV} \\
x_o &= 5.33 \times 10^{-10} \text{cm} \\
K_e &= 8 \times 10^{5} \text{ dyne/cm} \\
K_g &= 27 \times 10^{5} \text{dyne/cm} \quad .
\end{aligned}
\qquad (3\text{-}7)
$$

The configuration coordinate diagram for the germanium-gallium vacancy center plotted using these parameters is shown in Fig. 3-5a. Judging from the data presented in Fig. 3-2 the CC curves for the case of tellurium complexes (most samples used here were tellurium-doped) should not differ substantially from those shown in Fig. 3-5a and therefore Te-complexes will be treated using the values given in (3-7). Because the 1.2 eV emission peak for Te-doped GaAs shifts by only 0.028 eV (toward shorter wavelengths) (See Fig. 3-5b) in going from 0°K to 300°K, the additional assumption will be made that the CC parameters remain approximately constant over this temperature range.

For transitions between the ground and excited states, Eq. (2-60a) becomes:

$$n_\ell n'_u \cong 4N_c^2 \qquad (3\text{-}8)$$



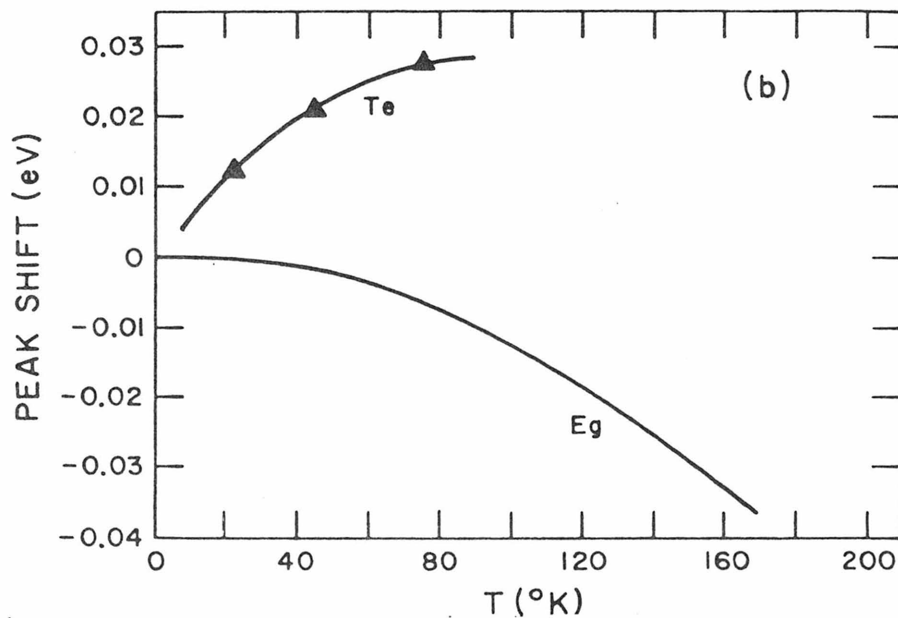


Fig. 3-5. (a) The configurational coordinate diagram for the gallium vacancy-germanium complex plotted using (3-7) (After Williams and White [51]); (b) peak shift from 0°K as a function of temperature for the 1.2 eV line in GaAs:Te. The curve E_g shows the change in band gap. (After Williams [39]).

where it has been assumed that the quasi-Fermi levels for holes and electrons are far from either band edge (valid when the GaAs has been compensated as a result of implantation). The degeneracies of the upper and lower states have been set equal to two and $N_\ell = N_u = N_c$, the number of complexes per cubic centimeter.

That the absorption should appear to vary as the square of the number of complexes is at first surprising until recollection is made of the original meaning of $\langle |H_{u,\ell}|^2 \rangle_{av}$. Equation (2-55), in a more general form, states that:

$$\sum_{u,\ell} |H_{u,\ell}|^2 = \sum_{E_u, E_\ell} (D_u D_\ell \langle |H_{u,\ell}|^2 \rangle_{av}) \tag{3-9}$$

or that:

$$\sum_{d_u, d_\ell}^{D_u, D_\ell} |H_{u,\ell}|^2 = D_u D_\ell \langle |H_{u,\ell}|^2 \rangle_{av} \tag{3-10}$$

when it is remembered that the summation over all upper and lower states may be broken up into individual summations over energy states and their respective degeneracies. Now, it will be recognized that $|H_{u,\ell}|^2$, contains selection rules which reduce the total number of transitions from $D_u \cdot D_\ell$ to D_u since only transitions between ground and excited states of the same complex are allowed. Taking account of this, Eq. (3-10) becomes:

$$N_c \sum_{d_u, d_\ell = 1}^{g_u, g_\ell} |H_{u,\ell}|^2 = g_u g_\ell N_c^2 \langle |H_{u,\ell}|^2 \rangle_{av} \tag{3-11}$$

where $D_u = N_c g_u$ and $D_\ell = N_c g_\ell$. Finally, taking $g_u = g_\ell = 2$:

$$\langle |H_{u,\ell}|^2 \rangle_{av} = \frac{1}{N_c} |H_{u,\ell}|^2 \quad (3\text{-}12)$$

Using Eqs. (3-2), (3-4), (3-5), (3-8) and (3-12), the absorption coefficient may now be written as:

$$\alpha_c(\hbar\omega) = D\omega \int_x \exp(-K_g x^2/2kT_{eff}) \delta(E_o - h\omega) dx \quad (3\text{-}13)$$

where:

$$D = \frac{32\pi e^2 N_c}{\sqrt{\varepsilon_r} c} \left(\frac{K_g}{2\pi k\ T_{eff}}\right)^{1/2} \langle |H_{u,\ell}|^2 \rangle_{av} \quad (3\text{-}14)$$

and:

$$E_o = E_u - E_\ell = \frac{(K_e - K_g)}{2} x^2 - K_e x_o x + \frac{K_e x_o^2}{2} + E_o \quad (3\text{-}15)$$

Equation (3-13) may be evaluated by making use of the following relationship:

$$\int_a^b g(x)\delta[f(x)]dx = \sum_{x_o} g(x_o)/|df(x)/dx|_{x_o} \quad (3\text{-}16)$$

where the sum over x_o includes all of the roots of $f(x)$ between a and b. $\alpha(h\omega)$ was computed using this procedure and the constants given in (3-7). The result is shown as curve B in Fig. 3-1a; the constant D in Eq. (3-13) was adjusted to match the maximum absorption observed near the band edge . Curve B (calculated at 300°C) may be compared with Fig. 3-6 which shows the absorption and emission spectra attributed to the gallium vacancy complex as measured by Williams and White at 80°K [51]. Finally, we note that, because all other terms vary slowly in comparison, $P(x)$, which has a Gaussian shape, deter-

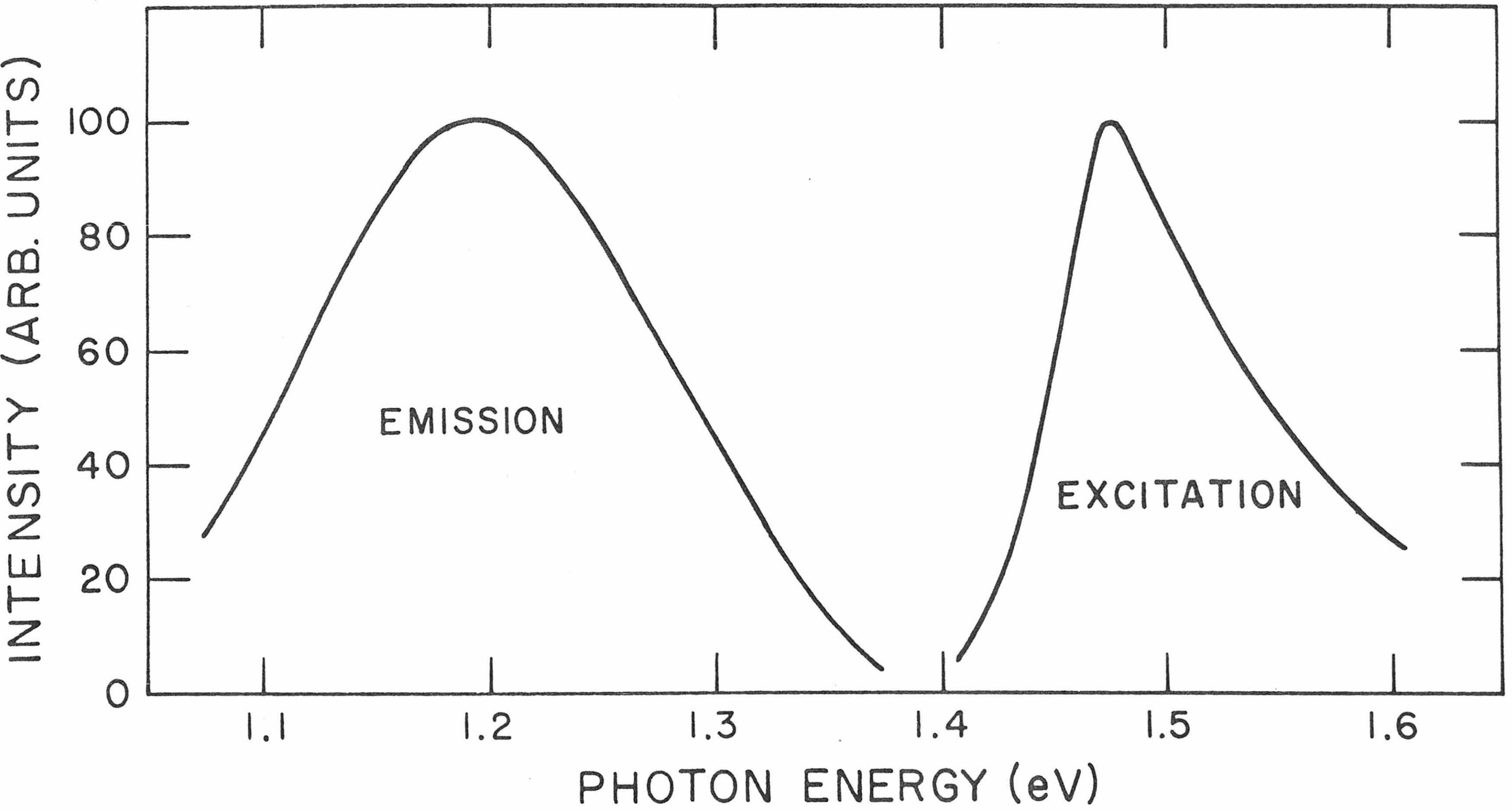


Fig. 3-6. Excitation and emission spectra due to the localized Ga-vacancy-donor complex in Ge-doped GaAs. (After Williams and White [51])

mines the principal spectral variation of $\alpha(\hbar\omega)$.

3. Absorption Involving Bandtails

Vacancies and interstitials created by proton implantation give rise to regions of local compression and dilation as illustrated in Figs. 3-7a and 3-7b. Interstitial atoms cause the local lattice constant to become smaller and, hence, in GaAs, the direct energy gap to become larger; vacancies have the opposite effect. For a small change, Δa, in the lattice constant, the energy gap, E_g, therefore changes by:

$$\Delta E_g = -(E_{1_c} + E_{1_v})\Delta a \qquad (3\text{-}17)$$

where E_{1_c} and E_{1_v} are the pressure coefficients for the edges of the conduction and valence bands, respectively. Assuming E_{1_c} and E_{2_c} to be approximately equal [52], the resulting deformation potentials generated by compression and dilation are shown in Figs. 3-7c and 3-7d. The assumption of a symmetrical or nearly symmetrical deformation of the band edges with respect to the center of the band gap will be justified in a later section which examines the similarity of carrier removal rates for both p- and n-type GaAs as a function of integrated proton flux.

Figure 3-7e shows how the band edge deformations introduced by a vacancy defect give rise to an acceptor level of ionization energy, E_a, and a donor level of ionization energy, $E_d(\cong E_a)$. Here we may note that absorption measurements made on electron-irradiated GaAs indicate the existence of a single defect-associated energy level located approximately 1 eV from one of the band edges [36]. This

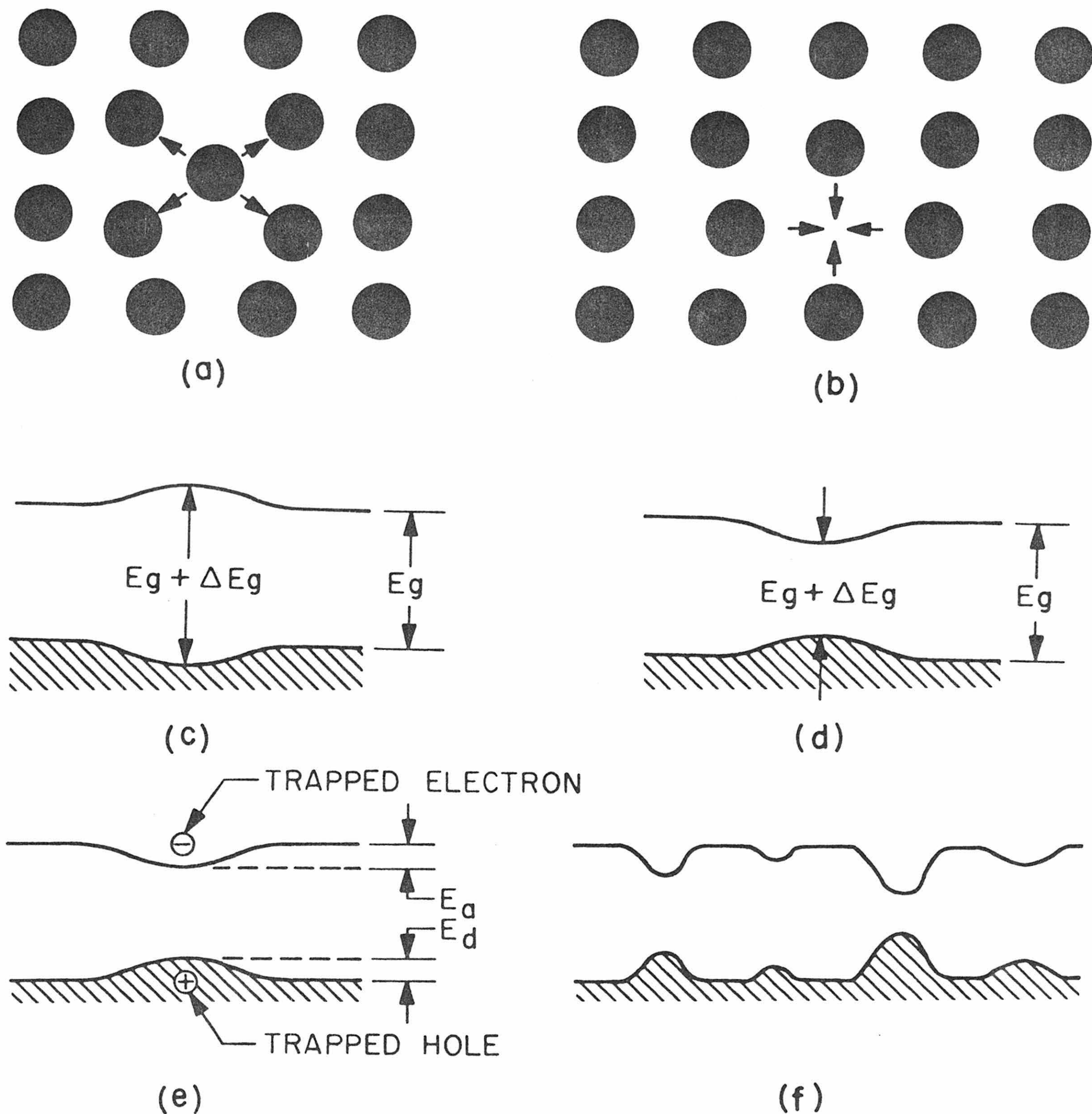


Figure 3-7. (a) and (b) illustrate the effects of interstitials and vacancies, respectively, on the local lattice structure; (c) and (d) show the effects of interstitials and vacancies, respectively, on the local band gap of GaAs; (e) donor and acceptor levels created within the band gap of GaAs by a vacancy; (f) band edge deformations caused by vacancy clusters of different size.

observation is consistent with the hypothesis that energetic electrons, possessing a relatively small amount of kinetic energy, are able to produce only isolated Frenkel (interstitial-vacancy) defect pairs [53]. The defect energy levels which result, therefore, might correspond to those shown in Fig. 3-7e. Protons in the 300-400 keV range, on the other hand, are likely to introduce much more extensive damage in the form of vacancy or interstitial clusters in addition to single Frenkel defects. In this case, a spectrum of defect levels is created, with each pair of acceptor-donor levels corresponding to a different configuration and local concentration of defect species. This situation is depicted in Fig. 3-7f. Since both the entropy and the electrostatic energy (each type of isolated defect or defect cluster gives rise to a local space charge) of a defect cluster increase with its size, large defect clusters thermodynamically are less favored than smaller clusters. In this regard, we note that the absorption and photoresponse spectra shown in Fig. 3.1 seem to vary in an exponential-like manner over an appreciable energy interval, suggesting that the number of defect clusters decreases exponentially with cluster size. Finally, based on these observations and the fact that the lowest acceptor-like defect level (or the highest donor-like defect level) appears to reside 0.58 eV from the valence band edge [35], the defect density of states functions are taken to be those shown schematically in Fig. 3-8.

Setting the energy of the pre-irradiated conduction band edge equal to zero and taking the room temperature energy gap to be 1.43 eV, the density of states functions referred to above together with the

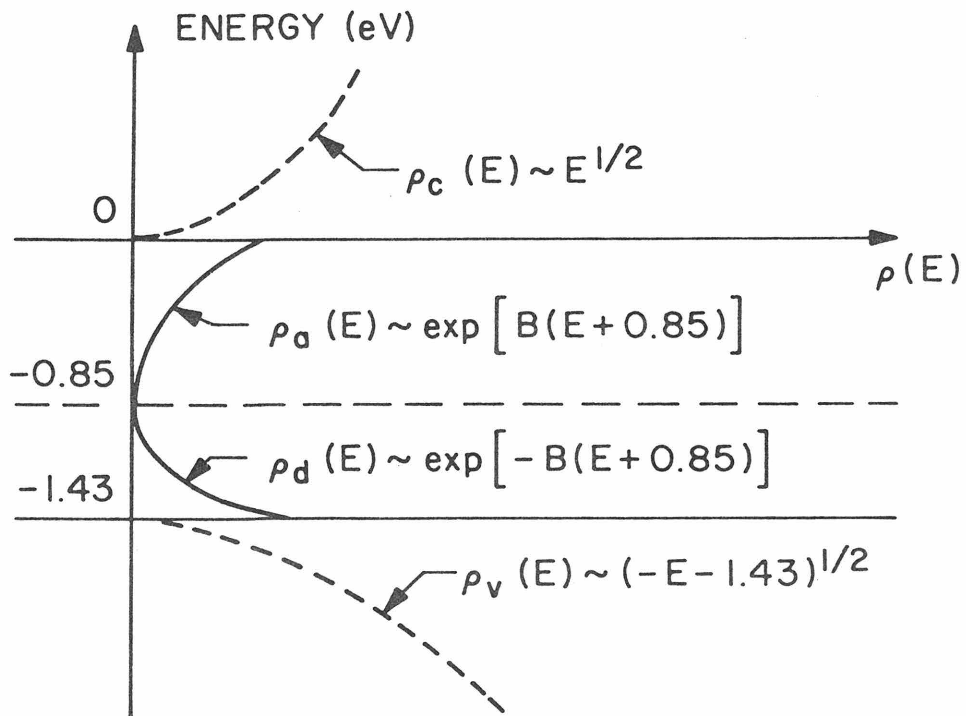


Figure 3-8. Conduction band, valence band and defect density of states functions of proton-implanted gallium arsenide as inferred from photoresponse, absorption [33], and free carrier removal [35] data.

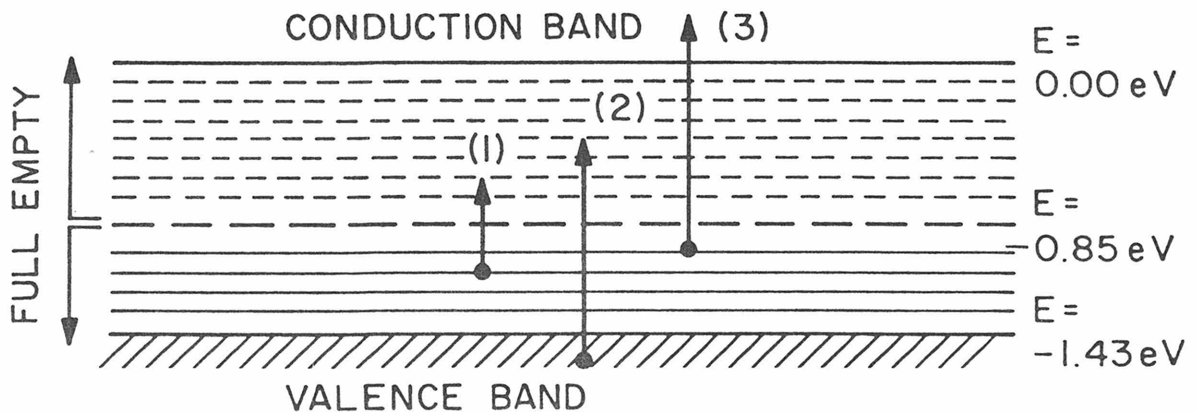


Figure 3-9. Schematic representations of the three types of dipole transitions made possible of proton-induced defect levels in gallium arsenide.

valence and conduction band density of states functions are given by:

$$\rho_c(E) = \frac{1}{2\pi^2}\left(\frac{2m_c}{\hbar^2}\right)^{3/2}(E)^{1/2} \quad : \text{ conduction band} \qquad (3\text{-}18a)$$

$$\rho_a(E) = A_a N \exp[B(E + 0.85)] \quad : \text{ acceptor defects} \qquad (3\text{-}18b)$$

$$\rho_d(E) = A_d N \exp[-B(E + 0.85)] \quad : \text{ donor defects} \qquad (3\text{-}18c)$$

$$\rho_v(E) = \frac{1}{2\pi^2}\left(\frac{2m_v}{\hbar^2}\right)^{3/2}(-E - 1.43)^{1/2} : \text{ valence band} \qquad (3.18d)$$

where

m_c = the electron effective mass

m_v = the hole effective mass

N = the number of defects per unit volume

and the normalization constants,

A_a = B/[exp(0.85B - 1)]

A_d = B/[exp(0.58B - 1)],

have been chosen such that $\int_{-0.85}^{0} \rho_a(E)dE = \int_{-1.43}^{-0.85} \rho_d(E)dE = N$.

Before proceeding with calculations of the various defect associated contributions to the absorption coefficient, we draw on several experimental observations to infer the probable location of the Fermi level. First we note that the photoresponse curve of Fig. 3-1a was obtained from n-type ($n = 10^{16}cm^{-3}$) material which had been

given a total dose of 2×10^{15} protons cm^{-2} and then annealed for 0.5 hours at 500°C. The results of Chapter 4 indicate that, under these circumstances, $N \cong 1.2 \times 10^{19}$ defects cm^{-3}.

Given this defect concentration and the original free carrier concentration, the Fermi level is found, using Eqs. (3-18) and (3-38) (See Section III-C) to lie very near its asymptotic value of $E_v + 0.58$ eV. In addition,with regard to the absorption spectrum obtained by Stein, we note that although no mention was made of the original free carrier concentrations of the irradiated specimens, because of the doses used (between 10^{15} and 10^{16} protons cm^{-2}) and because of the lack of post-implantation annealing, the Fermi level is once again expected to lie near its limiting value of $E_v + 0.58$ eV.

As a further prelude to the absorption calculations it may be recognized that, although both Stein's absorption data and the photo-response measurements were taken at room temperature, very little error will be introduced by using 0°K occupation statistics. While this approximation becomes invalid, for transitions which begin or terminate within a few kT of the Fermi level (it being seriously violated in the case of low energy transitions between defect levels close to and on either side of the Fermi level), it will be recalled that the lowest energy investigated (by Stein), was 0.2 eV ($\gg kT = 0.025$ eV at room temperature).

We now calculate the defect-associated absorption spectrum. To do so three types of transitions (shown schematically in Fig. 3-9) involving energies less than $E_g = 1.43$ eV, must be considered. Referring to Fig. 3-9, it is seen that with the exception of those

levels which lie within a few meV of the valence and conduction bands, the defect levels may be considered "deep". Accordingly, their associated electronic wave functions become (as discussed in Chapter 2) admixtures of free-electron wave functions and, as a result, the k-selection rule for transitions involving them is relaxed [26]. Recalling the above assumption regarding occupation statistics, Eq. (2-64) for the absorption coefficient therefore becomes:

$$\alpha(\hbar\omega) = \xi(\omega) \int_E n_\ell(E - \hbar\omega) n_u'(E) dE \qquad (3\text{-}19)$$

where:

$$\xi(\omega) = \frac{8e^2\pi^2\omega}{\sqrt{\epsilon_r}\, c} \langle |H_{u,\ell}|^2 \rangle_{av} \qquad (3\text{-}20)$$

Transitions between the valence band and defect levels above the Fermi level and transitions between filled defect levels and the conduction band may be treated in a straightforward manner. The resulting absorption coefficients, using Eqs. (3-18), are given by:

$$\alpha^{(2)}(\hbar\omega) = \xi(\omega) \frac{AN}{2\pi^2} \left(\frac{2m_v}{\hbar^2}\right)^{3/2} \int_{-0.85}^{-1.43+\hbar\omega} [-(E - \hbar\omega) - 1.43]^{1/2}$$

$$\times \exp[B(E + 0.85)] dE$$

$$= \frac{AN}{2\pi^2} \left(\frac{2m_v}{\hbar^2}\right)^{3/2} \exp[-0.58B] \left\{\xi(\omega) \int_0^{\hbar\omega - 0.58} y^{1/2} \exp(-By) dy\right\} \exp(B\hbar\omega) \qquad (3\text{-}21)$$

and:

$$\alpha^{(3)}(\hbar\omega) = \frac{AN}{2\pi^2}\left(\frac{2m_c}{\hbar^2}\right)^{3/2} \exp[-0.85B] \left\{ \xi(\omega) \int_0^{\hbar\omega - 0.85} y^{1/2} \times \exp(-By)dy \right\} \exp(B\hbar\omega) \ . \qquad (3\text{-}22)$$

Equations (3-21) and (3-22) are seen to account for a portion of the absorption spectrum lying between 0.58 and 1.43 eV.

Finally, we calculate the absorption due to transitions between filled and empty defect levels. These transitions involve initial and final electronic states which are both (for levels more than a few meV from either band edge) highly localized. In order to characterize these deep levels we use the quantum defect method [54]. Accordingly, the radial part of the ground (S) state solution $(F_\nu(\vec{r}) = P_\nu(r)Y_0^0(\theta,\phi))$ of the semi-phenomenological Schrödinger's equation,

$$[-\hbar^2\nabla^2/2m^* - e^2/\kappa r - \varepsilon(OBS)]F_\nu(\vec{r}) = 0 \ , \qquad (3\text{-}23)$$

is given by:

$$P_\nu(\nu) = N_\nu(2r/\nu a^*)^{\nu-1} e^{-r/\nu a^*} \qquad (3\text{-}24)$$

where:

$$N_\nu = (\nu a^*)^{1/2}/[2\nu\Gamma(\nu)] \ . \qquad (3\text{-}25)$$

Γ in Eq. (3.25) is the gamma function and κ is the relative dielectric constant. The quantity ν (referred to as the effective principal quantum number) is related to the observed binding energy through the following relation [25]:

$$\varepsilon(OBS) = -R^*/\nu^2 \qquad (3\text{-}26)$$

where R^* is the hydrogenic Rydberg $e^2/2\kappa a^*$.

Making the ad hoc assumption, to be justified shortly, that transitions between pairs of defect levels located symmetrically with respect to the Fermi level (see Fig. 3-10) are the most likely to occur and noting from Fig. 3-1a that the exponential-like spectral behavior terminates (or becomes masked) at approximately 1.3 eV, the smallest binding energy with which we must concern ourselves is 0.065 eV (=(1.43 - 1.3)/2). Given that $a^* = \kappa(m/m^*)a_0$ (a_0 = the Bohr radius = 0.53Å) and that $R^* = (m^*/m\kappa^2)R$ (R = the ground state ionization energy of hydrogen in vaccuo = 13.65 eV), we have,therefore, for the level nearest the conduction band:

$$a_c^* = (12/0.072)(.53\text{Å}) = 88.5\text{Å} \qquad (3\text{-}27a)$$

$$R_c^* = (0.072/144)(13.65\text{ eV}) = 6.82\text{ meV} \quad . \qquad (3\text{-}27b)$$

and for the level nearest the valence band:

$$a_v^* = (12/.5)(.53\text{Å}) = 12.7\text{Å} \qquad (3\text{-}28a)$$

$$R_v^* = (0.5/144)(13.65\text{ eV}) = 47.4\text{ meV} \quad . \qquad (3\text{-}28b)$$

κ in Eqs. (3-27a) - (3-28b) has been taken equal to 12 for GaAs at $\lambda \sim 1\mu$. Using Eqs. (3-27b) and (3-28b) and taking $\varepsilon(OBS) = 0.065$ eV, Eq. (3-26) yields for the two shallowest defect levels of interest:

$$\nu_c = 0.324 \qquad (3\text{-}29a)$$

$$\nu_v = 0.854 \quad . \qquad (3\text{-}29b)$$

The approximate extension (e^{-1} points) of the corresponding radial wave functions may now be calculated using Eq. (3-24):

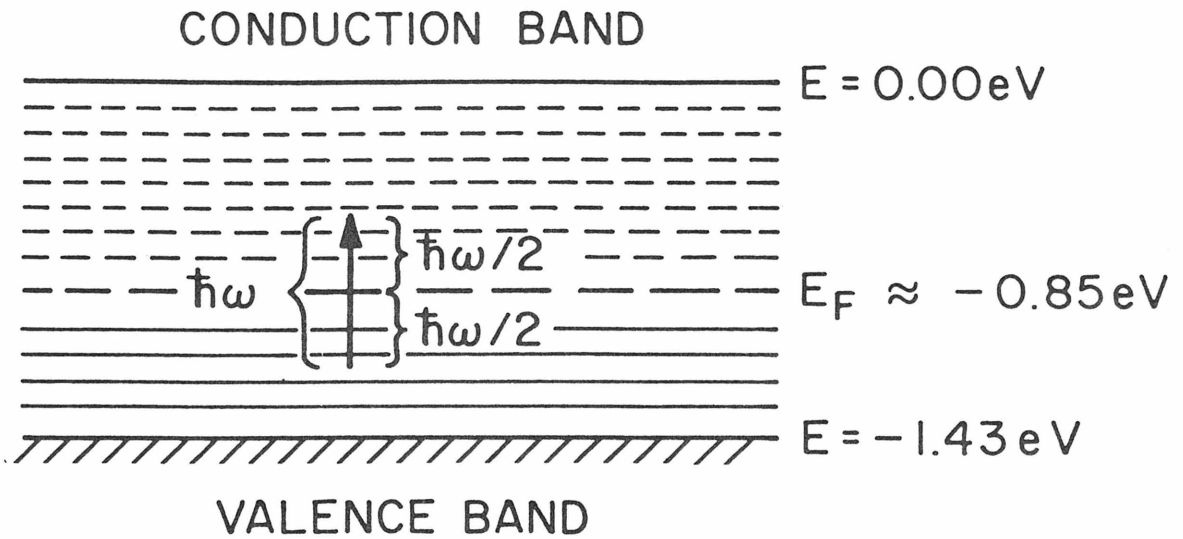


Figure 3-10. Allowed transition between defect levels located symmetrically with respect to the Fermi level of proton-implanted gallium arsenide.

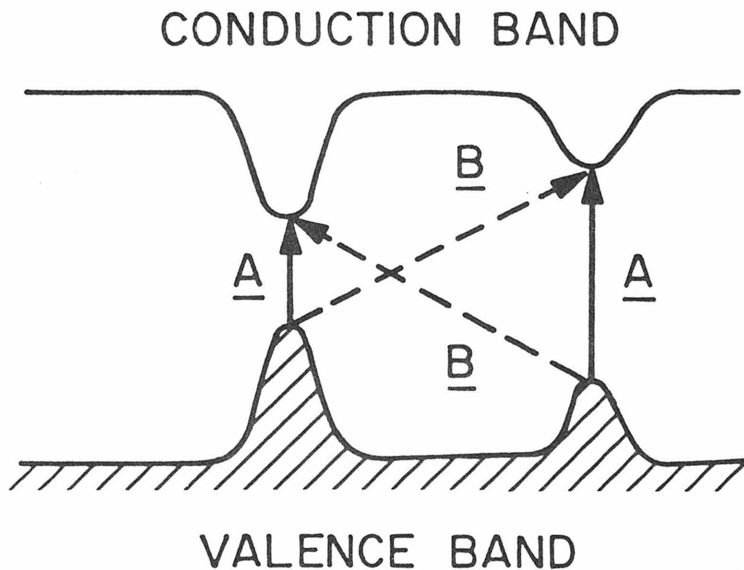


Figure 3-11. Transitions between defect states of proton-implanted gallium arsenide. Transitions between donor- and acceptor-like levels originating from the same defect (labeled A) are favored over those between states originating from different defects (labeled B).

$$r_c = \nu_c a_c^* = 28.7\text{Å} \quad (3\text{-}30a)$$

$$r_v = \nu_v a_v^* = 40.5\text{Å} \quad . \quad (3\text{-}30b)$$

Recalling that the number of defects remaining in the sample used to obtain the photoresponse curve was taken to be $\sim 10^{19}\text{cm}^{-3}$ and assuming for them a random, homogeneous distribution, the average distance between defects is found to be:

$$r_d = (3/4\pi N)^{1/3} = 34.7\text{Å} . \quad (3\text{-}31)$$

Bearing in mind the values given by Eqs. (3-30a) and (3-30b), Eq. (3-31) indicates that for the smallest ionization energies of interest, defect wave functions just begin to overlap. Moreover, as one moves farther into the band gap, ν, decreases rapidly and defect wave functions become progressively more localized. As a result of this localization, the matrix element which connects states associated with spatially separated defects will be much smaller than the matrix element which connects states associated with the same defect. The situation is illustrated in Fig. 3-11: transitions labeled A are favored over those labeled B. According to this "selection" rule, $\langle|H_{u,\ell}|^2\rangle_{av}$, as in the case of vacancy-complex excitation, will be equal to $\langle|H_{u,\ell}|^2\rangle$ divided by the total number of defects per unit volume N.

Following the above discussion, the absorption coefficient for transitions between pairs of defect levels is given by:

$$\alpha^{(1)}(\hbar\omega) = \frac{8e^2\pi^2\omega|H_{u,\ell}|^2}{\sqrt{\epsilon_r}c} A_a A_d N \exp[B\hbar\omega] . \quad (3\text{-}32)$$

Taking $N = 10^{20}cm^{-3}$ as the greatest number of defects to be expected under the implantation conditions to be considered here (See Chapter 4), $\alpha^{(1)}$ is seen to be orders of magnitude smaller than either $\alpha^{(2)}$ or $\alpha^{(3)}$ which contain the valence and conduction band effective density-of-states, respectively. Moreover, since $m_v >> m_c$, $\alpha^{(2)} >> \alpha^{(3)}$ giving, to good approximation:

$$\alpha(\hbar\omega) = \alpha^{(1)}(\hbar\omega) + \alpha^{(2)}(\hbar\omega) + \alpha^{(3)}(\hbar\omega)$$

$$\cong \alpha^{(2)}(\hbar\omega) \quad . \quad (3\text{-}33)$$

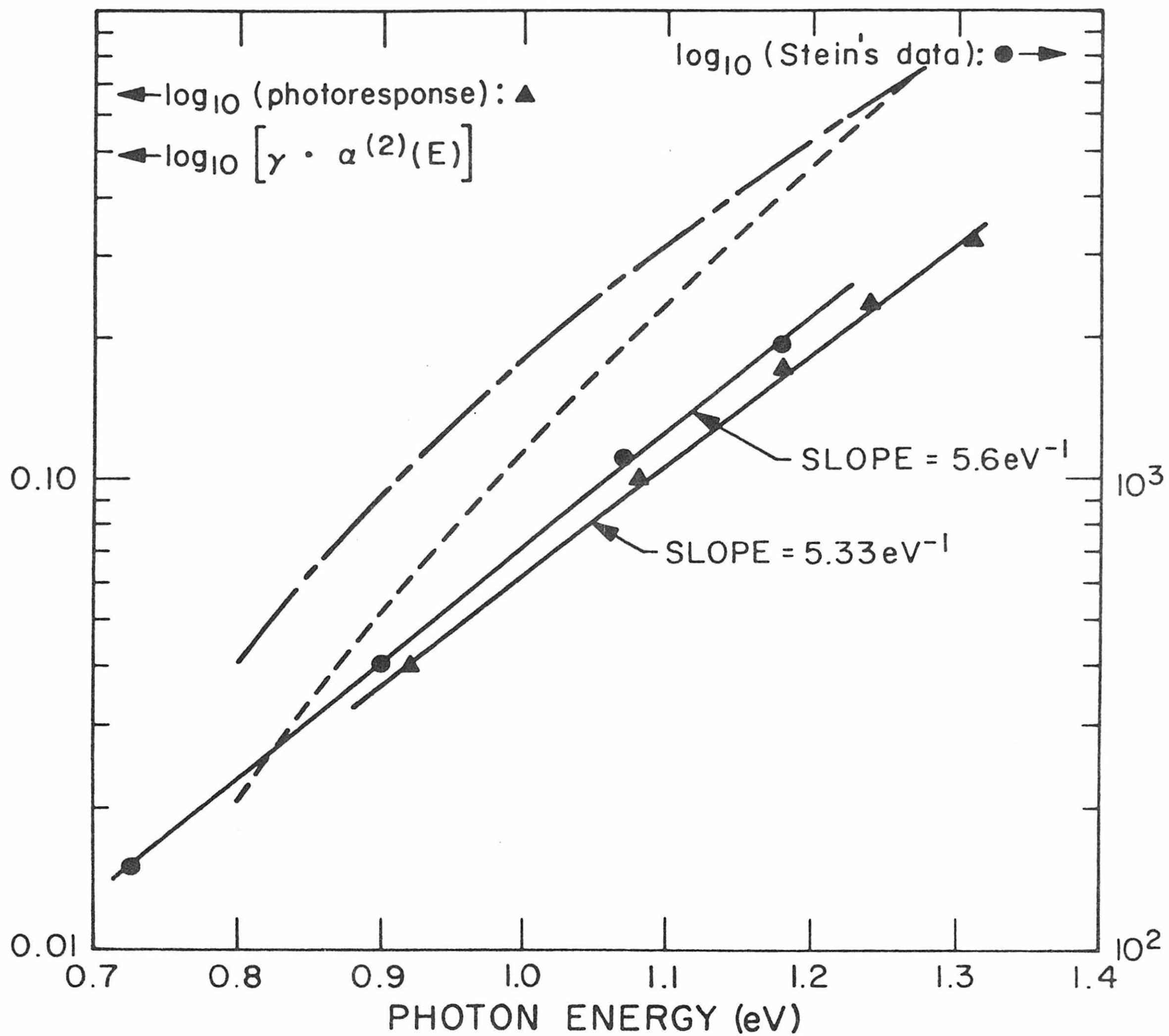
The expression for $\alpha^{(2)}(\hbar\omega)$,

$$\alpha^{(2)}(\hbar\omega) = C\cdot\hbar\omega\cdot\exp[B(\hbar\omega - 0.58)] \times \left\{\int_0^{\hbar\omega-0.58} y^{1/2}\exp(-By)dy\right\} \quad (3\text{-}34)$$

where:

$$C = \frac{ANe^2}{4\sqrt{\epsilon_r}c\hbar} \left(\frac{2m_v}{\hbar^2}\right)^{3/2} \langle|H_{u,\ell}|^2\rangle_{av} \quad (3\text{-}35)$$

was fitted to the absorption and photoresponse data of Fig. 3-1 in the following way: Semilogarithmic plots of the photoresponse curve (between 0.8 eV and 1.35 eV) and of the electrically active contribution to the absorption spectrum (between 0.7 eV and 1.2 eV) were used to obtain an initial estimate of the exponential coefficient B of Eq. (3-34). These plots are shown in Fig. 3-12; the electrically



$\log_{10}$[photoresponse data]: ▲

$\log_{10}$[(Total abs.)-(curve A)]: ● (from Stein)

$\log_{10}[\gamma \cdot \alpha^{(2)}(E)]$: - - - - - - ($\gamma = 10^{-2}$, B = 5.4 eV^{-1})

$\log_{10}[\gamma \cdot \alpha^{(2)}(E)]$: —— – —— ($\gamma = 10^{-1}$, B = 3 eV^{-1})

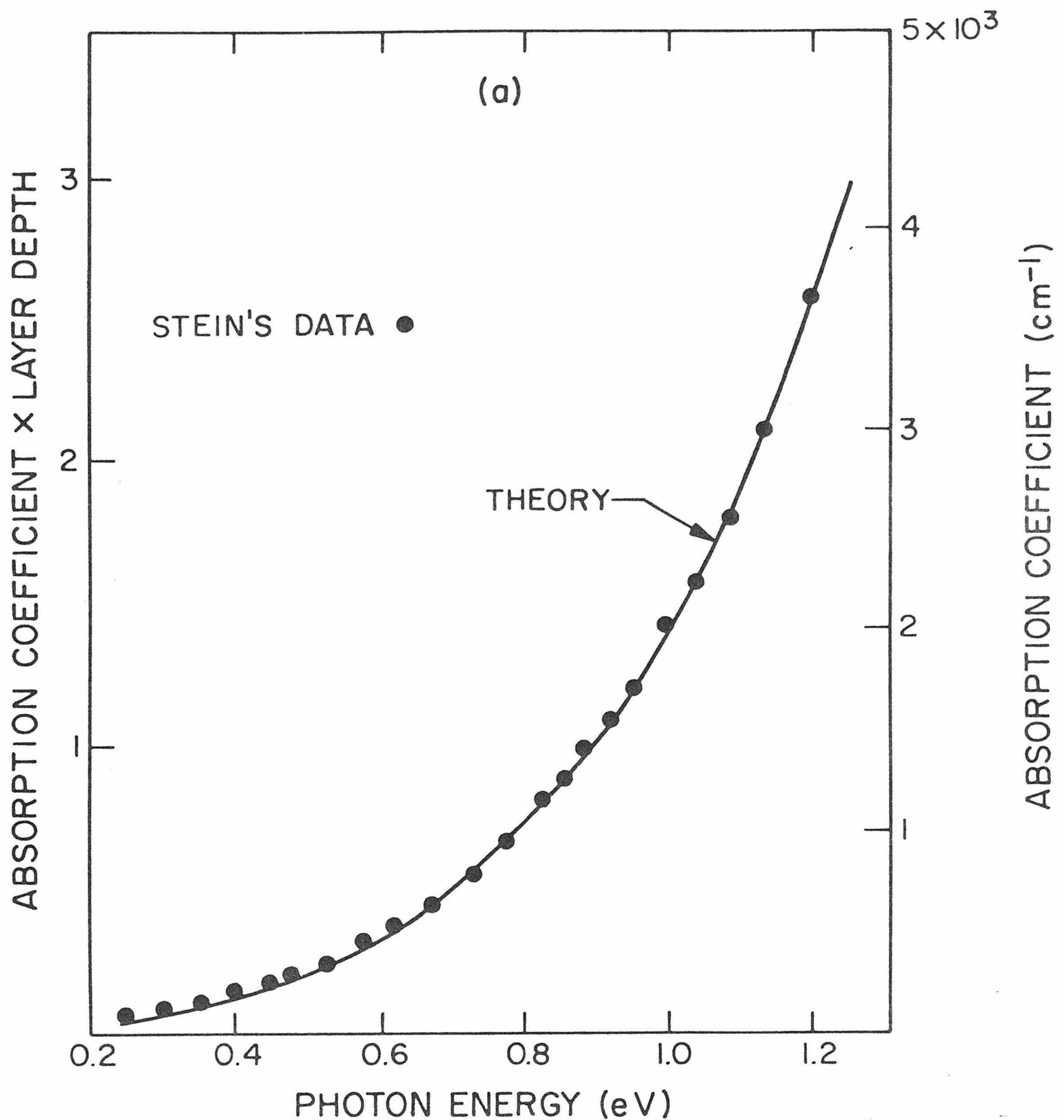
Figure 3-12. Semi-logarithmic plots used to find the exponential constant B. C in Eq. (3-34) has been set equal to one.

active portion of the absorption spectrum was obtained by subtracting curve A (the strictly dissipative contribution due to absorption and scattering by thermal spikes) of Fig. 3-1b from the total absorption spectrum of Fig. 3-1b. The slopes of the photoresponse and absorption plots yield for B values of 5.33 eV^{-1} and 5.6 eV^{-1}, respectively. Letting $B = 5.4\ eV^{-1}$ in Eq. (3-34) (and taking the constant $C \equiv 1$ for convenience), however, led to an absorption spectrum which increased too rapidly with energy (See Fig. 3-12); a best fit was eventually found, by trial and error, to result when $B = 3\ eV^{-1}$.

Finally, as a means of summarizing the results of this section, Fig. (3-13) compares the theoretical and experimental photoresponse and absorption spectra of proton-implanted GaAs. The theoretical photoresponse spectrum was obtained by summing Eqs. (3-13) and (3-34); the theoretical absorption spectrum was obtained by summing Eqs. (3-1), (3-13), and (3-34). In both cases, multiplicative constants were chosen to obtain the best agreement between theory and experiment.

B. Absorption-Induced Refractive Index Changes

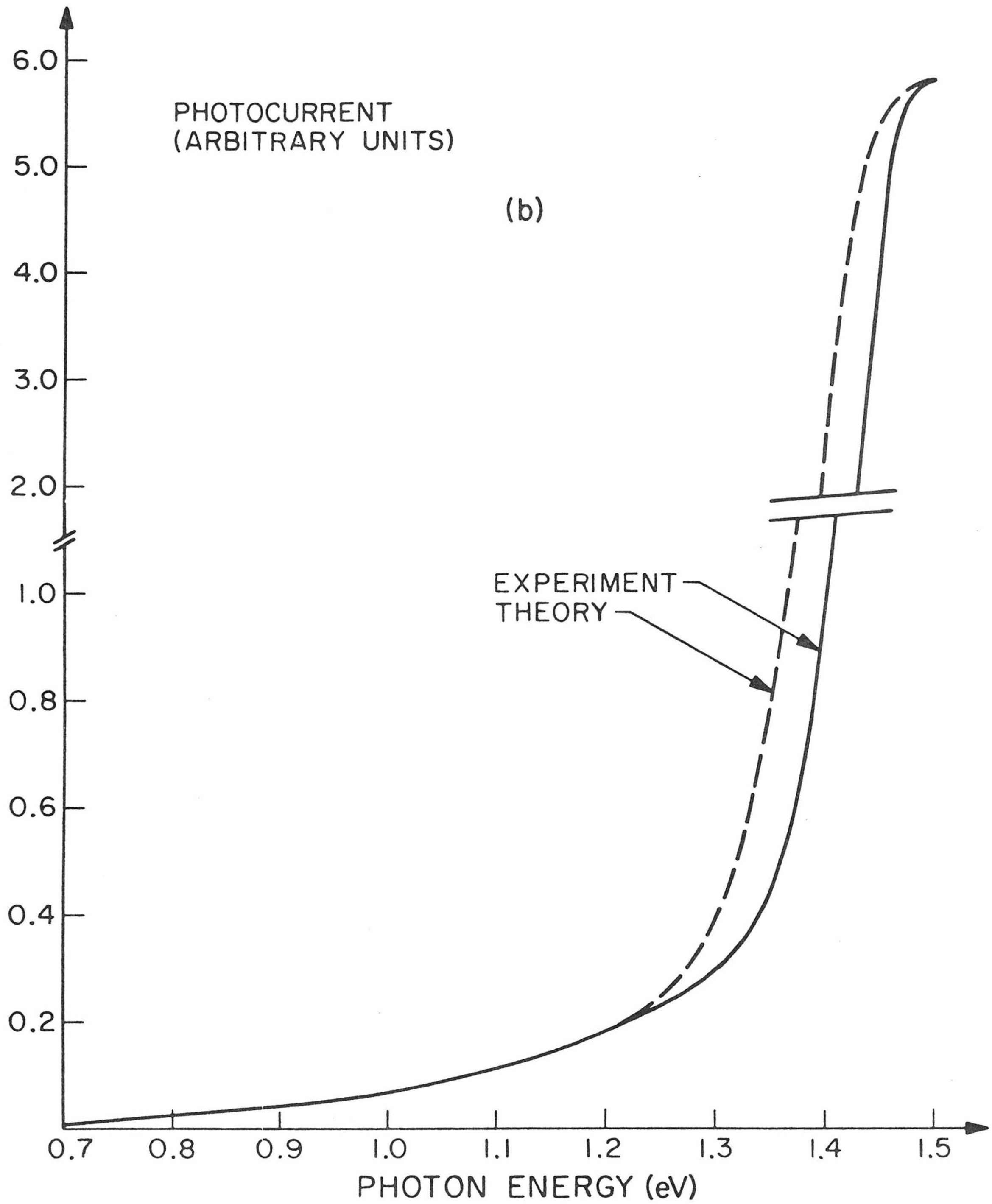
According to the Kramers-Kronig relationship given in Chapter 2, the additional absorption introduced by proton implantation will give rise to a refractive index change. An upper bound for the largest contribution, that due to the shift toward lower energies of the band edge, may be easily calculated by assuming that the edge shift is constant between 1.35 and 1.45 eV. Taking the absorption coefficient in this region to be 600 cm^{-1}, the (positive) change in the refractive index at 1.15μ (1.08 eV) is given by:



Theoretical Absorption Spectrum:

$$\alpha(E) = 1.35 \times 10^3 E^2 + 3.7 \times 10^2 \times E \times \exp[3(E-0.58)] \times \left\{ \int_0^{E-0.58} y^{1/2} \exp(-3y)\,dy \right\}$$

Figure 3-13. (a) Theoretical absorption spectrum fitted to Stein's data; (b) theoretical and experimental photoresponse curves. (See Eqs. (3-1), (3-13), and (3-34)).



Theoretical Photoresponse Curve:

$$\alpha(E) = 2.31 \times 10^3 \int_x \exp(-Kg\, x^2/2k\, T_{eff})\delta(E_o - E)dx$$

$$+ 0.0345 \times E \times \exp[3(E-0.58)] \times \{\int_0^{E-0.58} y^{1/2}\exp(-3y)dy\}$$

$$\Delta n_d(1.08 \text{ eV}) = \frac{ch}{2\pi^2} \int_{1.35}^{1.45} \frac{600}{(E')^2 - (1.08)^2} dE'$$

$$= 1.43 \times 10^{-5} \quad . \tag{3-36}$$

As is obvious from an inspection of Fig. 3-1a, inclusion of the exponential and quadratic portions of the absorption spectrum will change the above result very little, particularly since, at 1.08 eV, integration over these quantities (as in (3-36)) produces both positive and negative contributions which tend to cancel.

Δn_d may be compared with the refractive index change, Δn_{fc}, caused by free carrier absorption. Equation (2-46a) yields, for a representative free carrier concentration of $10^{18} cm^{-3}$ and at a wavelength of 1.15μ:

$$\Delta n_{fc}(1.08 \text{ eV}) = \frac{Ne^2}{2n\omega^2 m^* \varepsilon_o}$$

$$= 2.34 \times 10^{-3} \tag{3-37}$$

where the relation, $K_r(N=0) - K_r(N=10^{18}) \equiv n_r^2(N=0) - n_r^2(N=10^{18}) \sim 2n\Delta n$, has been used. Since Δn_d is nearly two orders of magnitude smaller than Δn_{fc}, refraction index changes will henceforth be assumed to arise only from free carrier effects.

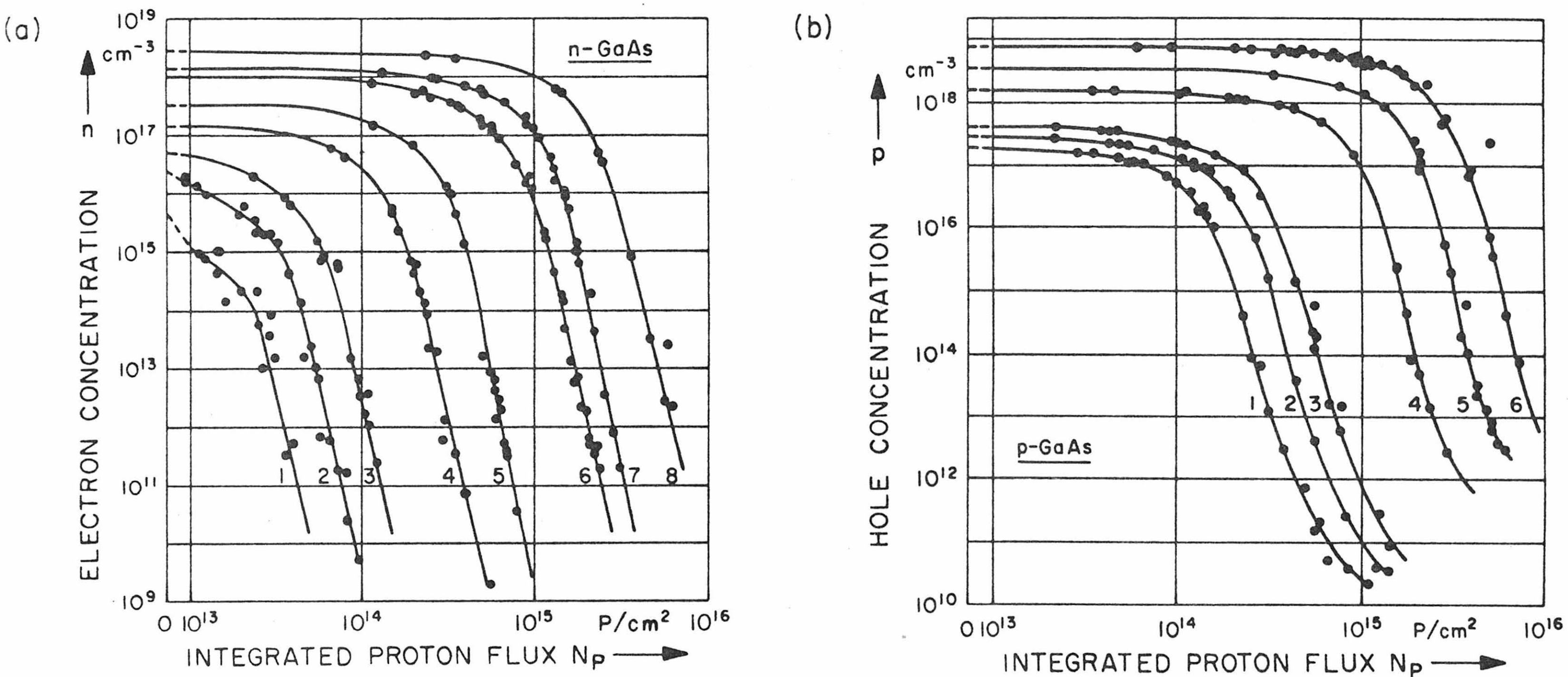
C. Electrical Properties of Proton-Irradiated GaAs

All implantations into GaAs involving sub-atomic projectiles are known to cause some modification of the targets' electrical properties,

the extent by which these properties are modified being dependent on the projectile species, energy, and fluence. The conductivity and Hall mobility of electron [55] and neutron-irradiated GaAs [36] have been extensively investigated; in comparison very little information regarding changes in these parameters as a result of proton-implantation may be found in the literature. On the other hand, measurements of free carrier concentration as a function of incident particle flux, data highly relevant to the present study have, to the best of the author's knowledge, been reported only for the case of proton-implantation [35].

Several conclusions regarding the nature of proton-induced damage in GaAs may be reached by examining the carrier removal curves referred to above [35]. These data, obtained for both n- and p-type GaAs by Wohlleben and Beck (W-B) using 3 MeV protons, are shown in Fig. 3-14. By calculating the integral carrier removal rates, $(n_o-n)/N_p$ and $(p_o-p)/N_p$, (n_o and p_o are initial electron and hole concentrations; n and p are final electron and hole concentrations after irradiation by N_p protons) and extrapolating to the case of infinite N_p (see Fig. 3-15), W-B obtained as limiting electron and hole concentrations, $3.6 \times 10^4 cm^{-3}$ and $10^9 cm^{-3}$, respectively, regardless of whether the material was originally n- or p-type. This means, for example, that n-GaAs is converted to p-GaAs under continued proton irradiation while the Fermi level adjusts to a final value of $E_F^* = E_V + 0.58$ eV. In addition to the integral carrier removal rates, W-B computed the differential carrier removal rates for both

Figure 3-14. Electron (a) and hole (b) concentration of p- and n-type GaAs with different initial doping as a function of integrated proton flux for 3 MeV protons. (After Wohlleben and Beck [35]).



n-type GaAs:

curve	$n_0 (cm^{-3})$
1	5.0×10^{15}
2	2.5×10^{16}
3	5.0×10^{16}
4	1.5×10^{17}
5	3.5×10^{17}
6	1.0×10^{18}
7	1.5×10^{18}
8	3.0×10^{18}

p-type GaAs:

curve	$p_0 (cm^{-3})$
1	2.0×10^{17}
2	3.0×10^{17}
3	4.3×10^{17}
4	1.7×10^{18}
5	3.8×10^{18}
6	8.0×10^{18}

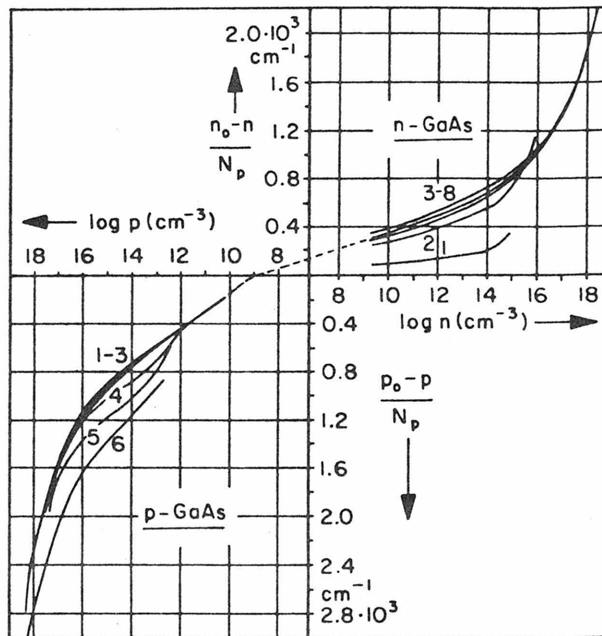


Fig. 3-15. Integral carrier removal rates for n- and p-type GaAs of different initial doping. (After Wohlleben and Beck [35])

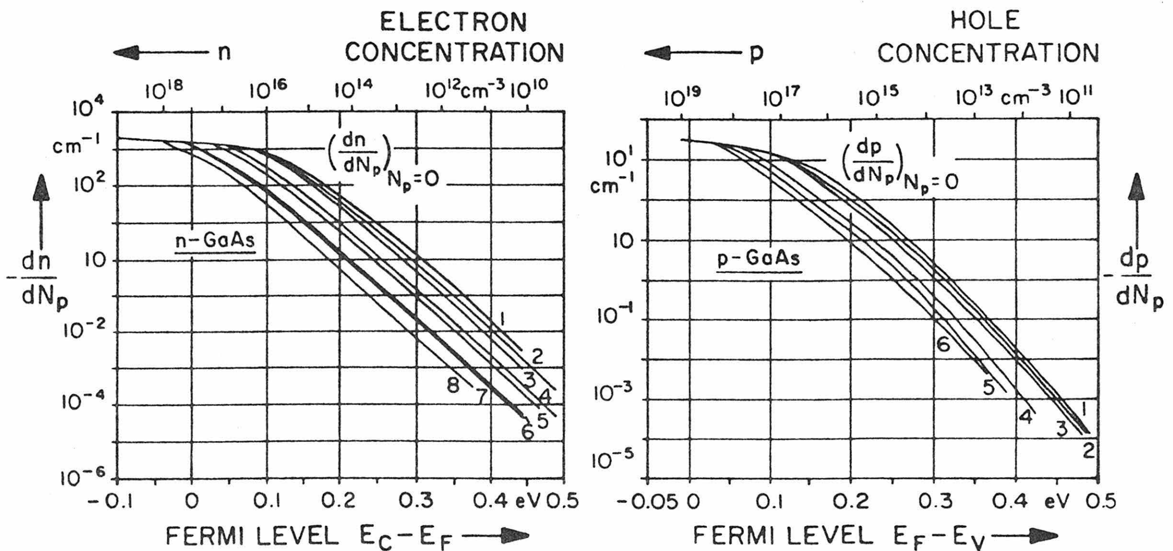


Figure 3-16. Differential carrier removal rates for n- and p-type GaAs of different initial doping. (After Wohlleben and Beck [35])

n- and p-type specimens by differentiating the curves of Fig. 3-14. Their results are shown in Fig. 3-16. The similarity between the differential carrier removal rates for both electrons and holes suggests the existence of nearly identical defect-generated acceptor- and donor-like energy level distributions.

The results of W-B described above may now be used to justify further the defect density-of-states functions which were proposed to explain the absorption and photoresponse data. The functional forms chosen to characterize both the donor- and acceptor-like defect density of states functions were taken to have identical exponential coefficients in order to reflect W-B's findings that the carrier removal rates for both holes and electrons were nearly identical. (The relationship between the exponential coefficient, B, of Eqs. (3-18) and the carrier removal rates will be made explicit shortly.) Moreover, the constant energy, -0.85 eV, which appeared in Eqs. (3-18b) and (3-18c) may now be understood as representing the limiting Fermi level as deduced by W-K, viz. the lowest-lying acceptor-like level or the highest-lying donor-like defect level.

As mentioned above, the exponential coefficient B of the defect density of states functions is intimately related to the carrier removal rate curves of Fig. 3-14. In order to demonstrate this fact, the exponential coefficient B, obtained previously by fitting Eq. (3-34) to the absorption and photoresponse data, will now be determined by analyzing the carrier removal curves of Fig. 3-14. Using Eqs. (3-18a)-(3-18d), the charge-balance equation for proton-implanted n-type GaAs is given by:

$$\int_0^\infty \frac{\rho_c(E)}{1 + e^{(E-\mu)/kT}}\,dE + \int_{-0.85}^{0} \frac{\rho_a(E)}{1 + e^{(E-\mu)/kT}}\,dE$$

$$= N_D + \int_{-\infty}^{-1.43} \frac{\rho_v(E)}{1 + e^{(\mu-E)/kT}}\,dE + \int_{-1.43}^{-0.85} \frac{\rho_d(E)}{1 + e^{(\mu-E)/kT}}\,dE \qquad (3\text{-}38)$$

where μ is the Fermi level and N_D is the number of donor atoms. (Owing to the effects of impurity banding encountered at the concentrations considered ($\sim 10^{18}$ cm^{-3}) combined with small donor ionization energies, all donors are assumed to be ionized.) The total negative charge (left-hand side of Eq. (3-38)) is made up of free electrons and negatively ionized acceptor-like centers while ionized donors, free holes, and ionized donor-like centers (right-hand side of Eq. (3-38)) constitute the total positive charge. At this point it will be noted that we have tacitly assumed the defect concentration to be proportional to the proton fluence. Using the results of Chapter 1, this assumption will soon be justified.

Continuing with our analysis, the carrier removal curves of Fig. 3-14, below the region of degeneracy, are seen to become linear over a wide range of free carrier concentration. Within this regime, Boltzmann statistics may be used to yield for the number of free electrons:

$$n_f = n_c\, e^{\mu/kT} \qquad (3\text{-}39)$$

where the effective density of conduction band states, $n_c = 2(2\pi m_c kT/h^2)^{3/2}$. In addition, within the linear regime, the

number of free electrons, free holes, and ionized donor-like defects may be neglected in comparison with the number of ionized donors. Under these conditions, Eq. (3-38) becomes:

$$\int_{-0.85}^{0} \frac{\rho_a(E)}{1 + e^{(E-\mu)/kT}}\, dE = N_D \ . \qquad (3\text{-}40)$$

The integral which appears in Eq. (3-40) may be evaluated approximately by using the absolute zero, step function form of the Fermi-Dirac distribution. Doing so and using Eq. (3-18b) yields

$$\frac{0.26N}{B}\,[e^{B(\mu+ 0.85)} - 1] = N_D \ . \qquad (3\text{-}41)$$

Finally, using Eq. (3-39) and restricting our attention to values of N for which

$$N_D \gtrsim 0.26N/B \qquad (3\text{-}42)$$

(3-41) becomes:

$$\left\{\frac{0.26}{B}\, e^{0.85B}(n_c)^{-kT}\right\} N\cdot n_f^{BkT} = N_D$$

or:

$$\log_{10}n_f = -(1/BkT)\log_{10}N + C \qquad (3\text{-}43)$$

where $C = \log_{10}(BN_De^{-0.85B}(n_c)^{kT}/0.26)$.

Taking kT = 38.7 meV (T = 300°K) and measuring the slopes of the carrier removal curves for both n- and p-type material, Eq. (3-43) yields, in both cases, $B = 2.97\ eV^{-1}$. The agreement between this value and the one obtained as a result of fitting Eq. (3-34) to the

absorption and photoresponse data (3 eV^{-1}) is remarkable and lends substantial credence to the proposed defect density of states functions.

Returning now to the question of how the number of defects is related to proton fluence, we again draw on the results of Wohllenben and Beck [35]. (The carrier removal curves given here were obtained by Wohllenben and Beck who, it will be recalled, used 3 MeV protons.) Using Eq. (1-19a) and accounting for the fact that the 3 MeV protons emerged from their specimens at 1.1 MeV, W-B found the (averaged) proportionality constant relating defect concentration to proton fluence to be $(N_d)_r = 5.8 \times 10^3 \Phi$.

With regard to the number of residual defects it is reasonable to assume, and has been suggested experimentally that, because of the high degree of spatial correlation between vacancies and interstitials, many of the isolated Frenkel defects recombine spontaneously [53]. According to the model of Kinchin and Pease [2] (Eq. (1-4)) single defect pairs appear for primary displacement energies between E_d and $2E_d$. The number of such defects may be estimated by determining what fraction of the primary knock-ons, for a given incident proton energy, possess energies within this range. Using Eq. (1-8), the probability that a primary displacement has energy between E_d and $2E_d$ is given by

$$P(E_d \leq T \leq 2E_d) = \frac{\int_{E_d}^{2E_d} \sigma_n(T)p(T)dT}{\int_0^{T_m} \sigma_n(T)p(T)dT} . \tag{3-44}$$

Using Eq. (1-17a), we find that $P(E_d \leq T \leq 2E_d) = 0.5$, i.e., for every primary collision, the probability is 0.5 that a single interstitial-vacancy pair will result. Assuming that <u>all</u> single Frenkel defects recombine, we have, finally, for the total number of displacements or defects (belonging to clusters of size greater than one), $2.9 \times 10^3 \Phi$ $(=5.8 \times 10^3 \Phi/2)$.

Taking 2.9×10^3 as the proportionality constant relating defect density to proton dose, both experimentally and theoretically determined curves of free carrier concentration as a function of defect concentration (for $10^{18} cm^{-3}$ n-type gallium arsenide) are shown in Fig. 3-17. The theoretical curve was obtained by solving Eq. (3-38) exactly on a computer; the experimental curve was obtained by using the proportionality constant just derived in conjunction with curve 6 of Fig. 3-14. Generally speaking, the agreement between theory and experiment is excellent; at least part of the discrepancy is probably due to not taking into consideration the additional compensating effects of the vacancy-donor complexes which were hypothesized earlier in order to explain the small, rigid shift of the fundamental absorption edge toward lower energies.

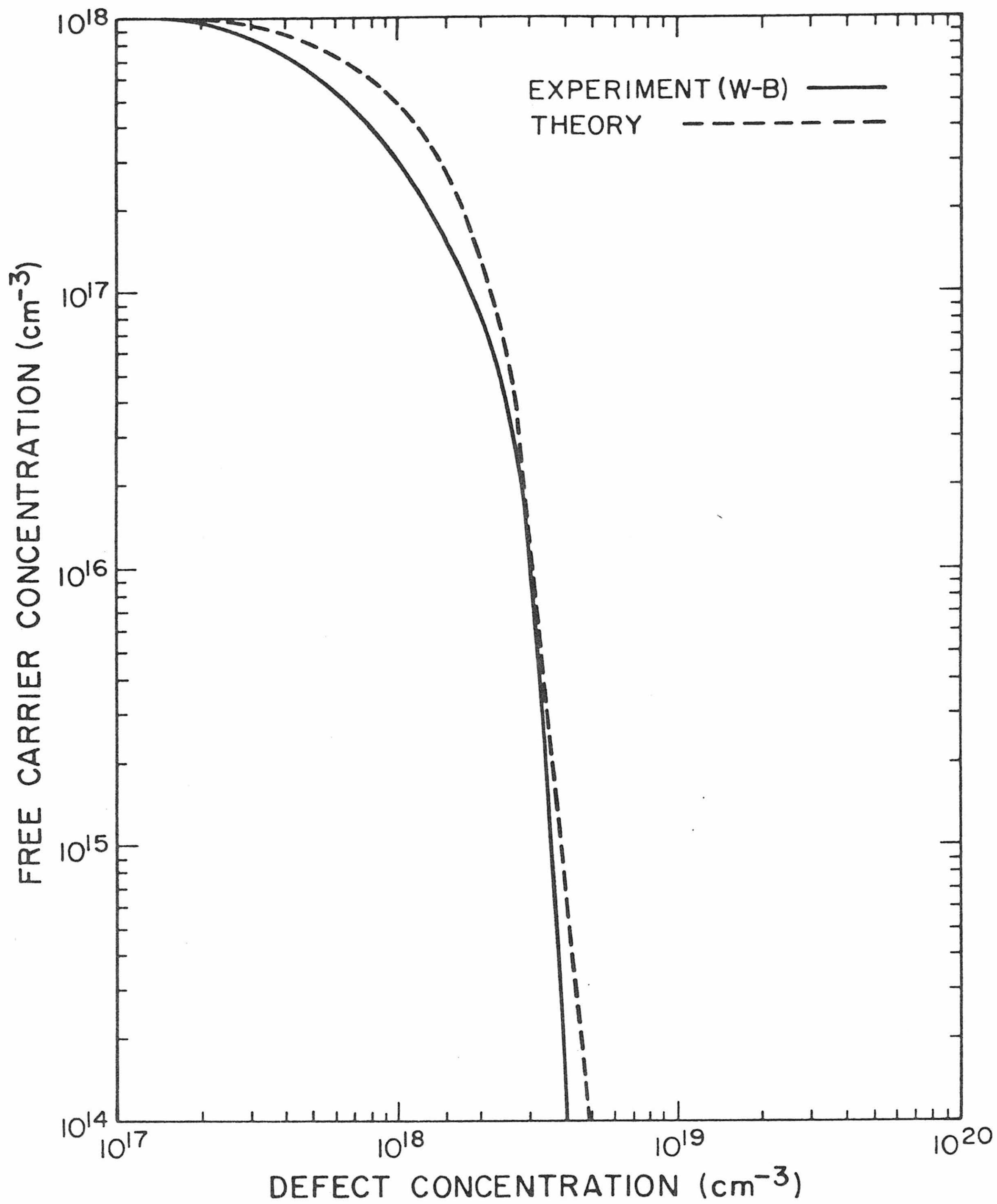


Figure 3-17. Theoretical and experimental curves of free carrier concentration as a function of defect concentration.

CHAPTER 4

OPTICAL WAVEGUIDING AND ATTENUATION IN PROTON-IMPLANTED GALLIUM ARSENIDE

I. Introduction

This chapter will examine the optical confinement and attenuation properties of planar waveguides formed by proton-implanting gallium arsenide. The physical mechanisms responsible for these properties were analyzed extensively in Chapter 3; the experimental results to be given below have been interpreted within this analytical framework. As a necessary point of departure, we begin this chapter with a description of the proton-induced damage distribution and the possible effects its nonuniformity imposes on the structure and attenuation of guided modes; a comparison between the theoretical and experimental results then follows.

II. The Depth Distribution of Damage

The mode profile and attenuation analysis to be presented in Section V presupposes a knowledge of the relationship which exists between the electrical compensation center and absorption center depth distributions generated by 300 keV protons incident on gallium arsenide. Described below is the procedure which was used to infer the electrical compensation center distribution; the relationship between this distribution and the absorption center distribution will be explored in Section IV. For the time being, we anticipate the

conclusions to be reached in Section IV and assume that one distribution may be related to the other through a simple proportionality constant. Consequently, the distribution obtained below will be understood to represent the damage distribution in general; the terms "electrical compensation center distribution," "defect center distribution," and "damage distribution" will, accordingly, be used interchangeably.

A curve such as the one shown in Fig. 3-17 together with a plot of the residual free carrier concentration versus depth of proton-implanted gallium arsenide may be used to determine the depth distribution of proton-generated, electrical compensation centers. Using reverse-biased Schottky barriers, Pruneaux, et al obtained plots of the free carrier concentration versus depth of proton-implanted gallium arsenide [34]. Their specimens were mounted in the sample chamber of the accelerator and a series of profiles made for each specimen, succeeding profiles corresponding to successively greater proton doses. Fig. 4-1 is an example of some of their data: The incident proton energy used was 150 keV and doses ranged from 10^{10} protons cm^{-2} to 1.2×10^{11} protons cm^{-2} in increments of $10^{10} cm^{-2}$; the initial gallium arsenide free carrier concentration was $1.3 \times 10^{16} cm^{-3}$.

A partially successful attempt was made to obtain similar profiles for some of the substrate-dose-anneal combinations used in the present study. Figure 4-2 shows free carrier concentration versus depth curves for an 8.5×10^{17} cm^{-3} n-type sample which was given a

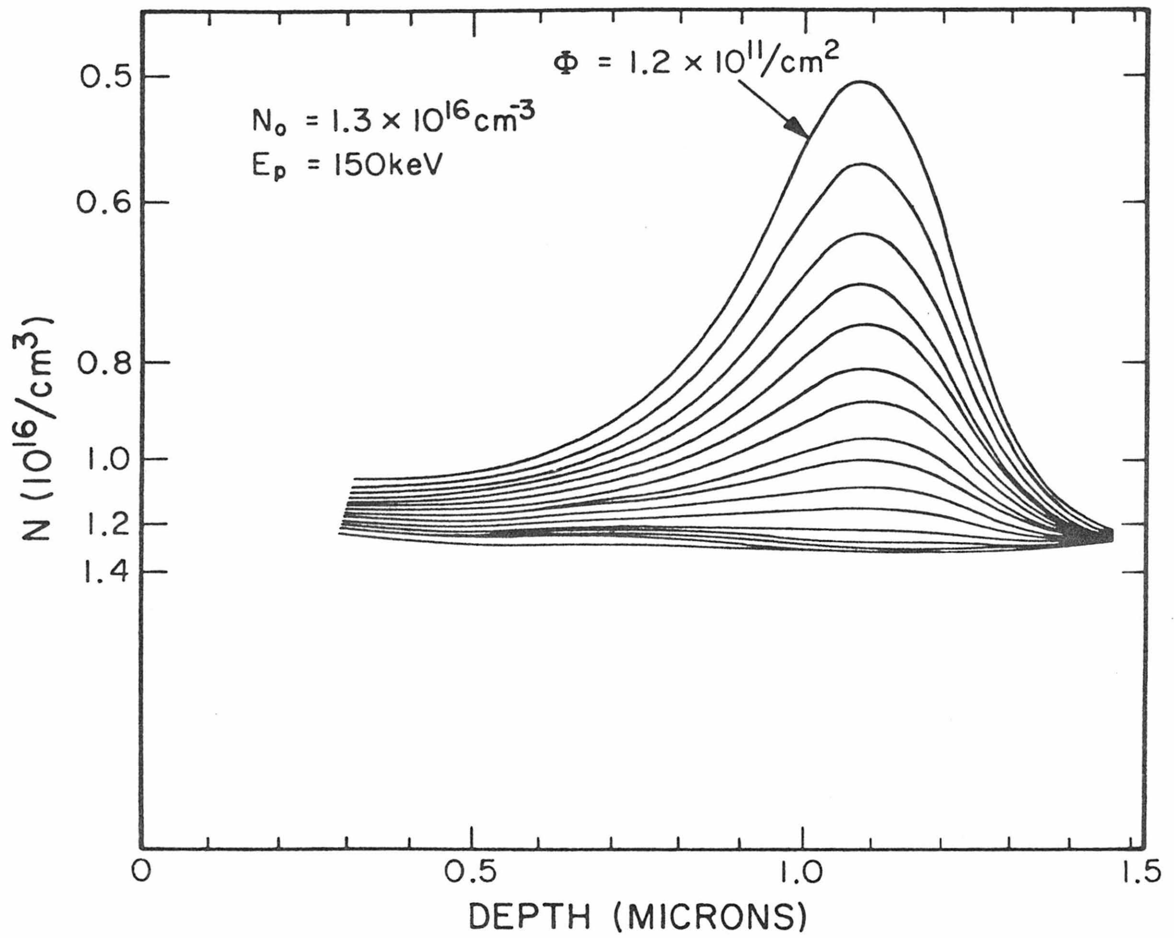


Figure 4-1. Reciprocal carrier concentration versus depth of proton-implanted GaAs. Incident proton energy was 150 keV. The n-type epitaxial GaAs had an initial carrier concentration of 1.3 x $10^{16}cm^{-3}$. Each incremental proton dose was $\sim 10^{10}cm^{-2}$. (After Pruniaux, North, and Miller [34])

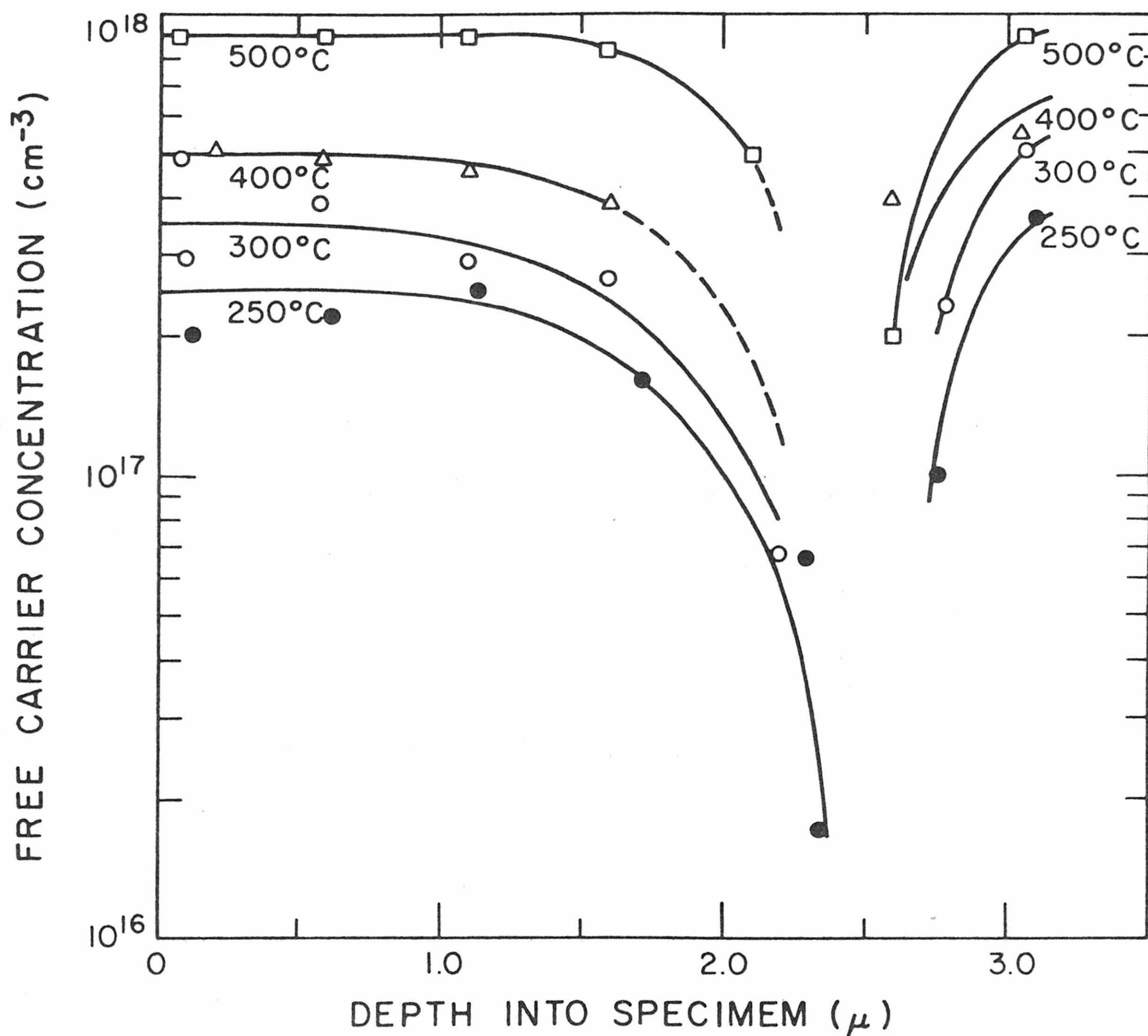


Figure 4-2. Free carrier concentration vs. depth curves of proton-implanted and annealed GaAs. Initial free carrier concentration was $8.5 \times 10^{17} cm^{-3}$. The proton dose was $1.5 \times 10^{14} cm^{-2}$. The annealing time at each temperature was 0.5 hour.

dose of 1.5 x 10^{14} protons cm^{-2}; the parameter is the annealing temperature. Two difficulties were encountered in making these measurements: after low temperature annealing ($\leq$ 350°C for 0.5 hours) excessive leakage current caused by damage-induced electrical generation centers prevented reliable capacitance versus voltage data to be taken over a meaningful reverse bias range. On the other hand, after high enough annealing to remove the generation centers, while acceptable Schottky barriers resulted, the free carrier concentration, by this time, had risen to such an extent that the depletion layers generated were too shallow to allow the carrier concentration to be determined over more than a few hundred angstroms. No solution was found for the first difficulty. However, by step-etching the specimen (see Chapter 6) following a high enough anneal, it was possible to determine the free-carrier concentration at discrete points corresponding roughly to the step depths; the data of Fig. 4-2 were obtained in this manner.

While, for the most part, these profile measurements did not yield data reliable enough for waveguide mode structure analysis (see Section V), they did reveal two important facts. The first was that damage annealed _in situ_; macroscopic diffusion effects were not apparent from the profiles. This fact greatly simplified the optical mode analysis: Since no geometrical changes occurred in the waveguides, only the pertinent refractive indices needed to be considered when calculating the optical mode profiles after annealing. The second fact to emerge was that the width of the damage peak generated by 300 keV protons appeared to be approximately equal to the width

of the damage peak produced by 150 keV protons (as used by Pruniaux, et al). This observation is in agreement with the theoretical calculations of Brice [12]: Fig. 1-1, for example, illustrates how the damage peak, according to Brice's theory, maintains approximately the same width for increasingly energetic implantations of B^{11} into Si .

The fact that the width of the damage peak remains almost the same for both 150 and 300 keV proton implants in accordance with the theoretical predictions of Brice (his example of boron-implanted silicon most closely parallels our case) suggests that certain other features of the proton-induced damage profile may be inferred from the results of his analysis. In particular, the damage distribution generated by 300 keV protons over the last 1.5μ of their travel will be assumed to have the same shape as the damage distribution caused by 150 keV protons over their entire penetration range; the damage caused by 300 keV protons between 0 and 1.5μ will then be found by extrapolation. The reasoning upon which this approach is based is the following: since protons are known to penetrate gallium arsenide approximately 1μ for every 100 keV of incident energy [56], it is reasonable to assume that after 1.5μ of travel, the residual proton energy is 150 keV, viz. that the protons lose energy linearly with depth. If it is further assumed that nearly all of the protons originally incident survive and maintain their direction of travel to this depth, then the damage distribution of 300 keV protons between 1.5μ and 3μ becomes the same as that

generated by 150 keV protons. To find the disorder distribution of 300 keV protons between 0 and 1.5μ , the distribution between 1.5μ and 3.0μ is smoothly extrapolated back to the target surface; the ratio of disorder at the surface to disorder at 1.5μ is taken to be one-half since Rutherford (nuclear) scattering varies inversely with projectile energy.

Following the above discussion, the damage profile for 300 keV protons is shown as curve A of Fig. 4-4; the damage or defect concentration is given in arbitrary units. The profile was obtained using the following procedure: The scaled (see Section V) free carrier concentration versus proton dose curve shown in Fig. 4-3 (derived from curves 1, 2, and 3 of Fig. 3-14) was used in conjunction with one of Pruniaux's free carrier concentration versus depth curves to generate an "effective" proton dose (α defect concentration) depth distribution; the resulting curve was then extrapolated, in the manner described in the preceding paragraph, 1.5μ back to the sample surface. As mentioned, arbitrary damage units were retained for convenience; had the abscissa of the scaled free carrier concentration versus proton dose curve first been multiplied by the proportionality constant obtained in Section 3-II C , absolute damage units would have resulted.

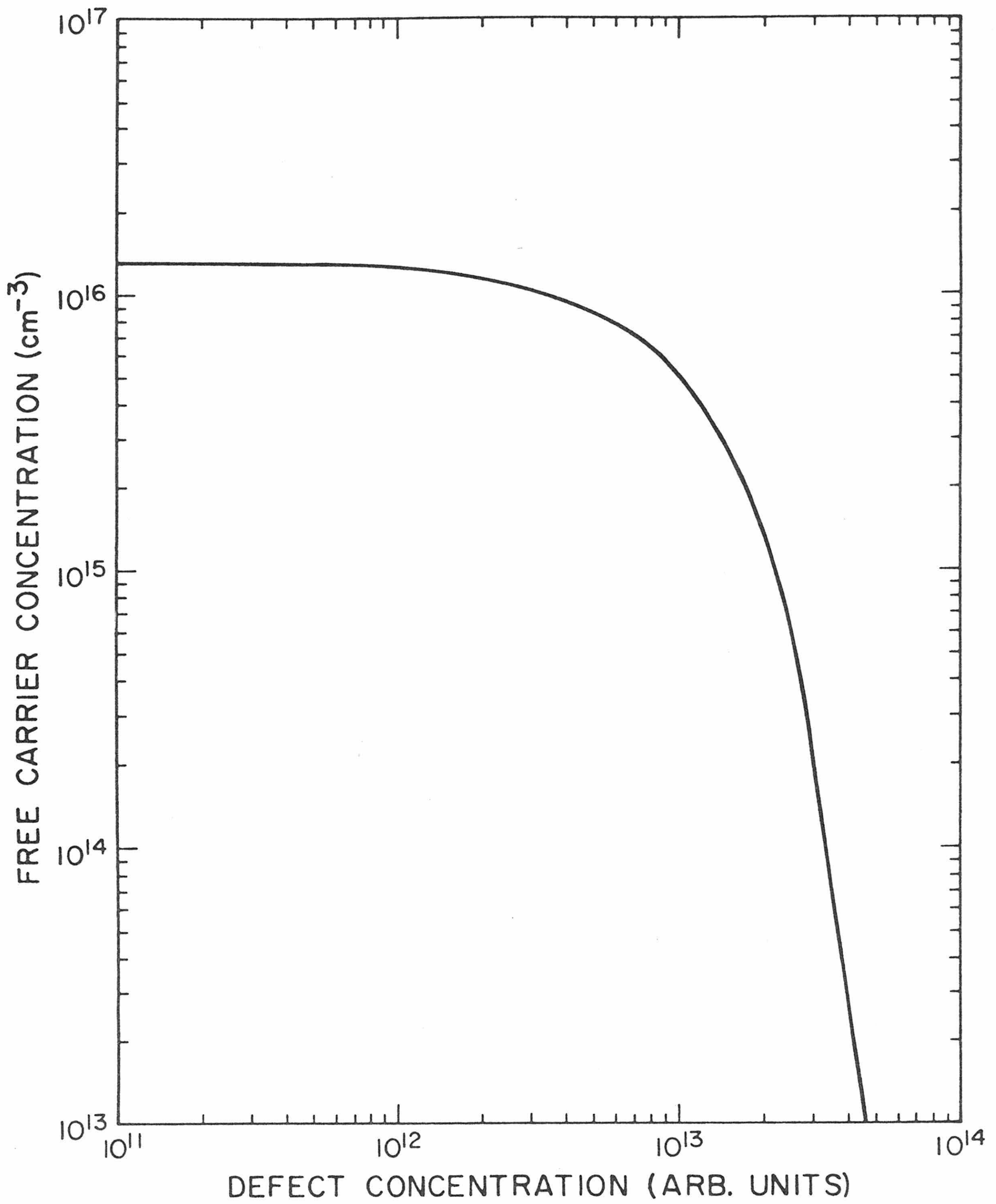


Figure 4-3. Scaled curve of free carrier concentration vs. defect concentration for initial free carrier concentration of 1.3×10^{16} (cm^{-3}).

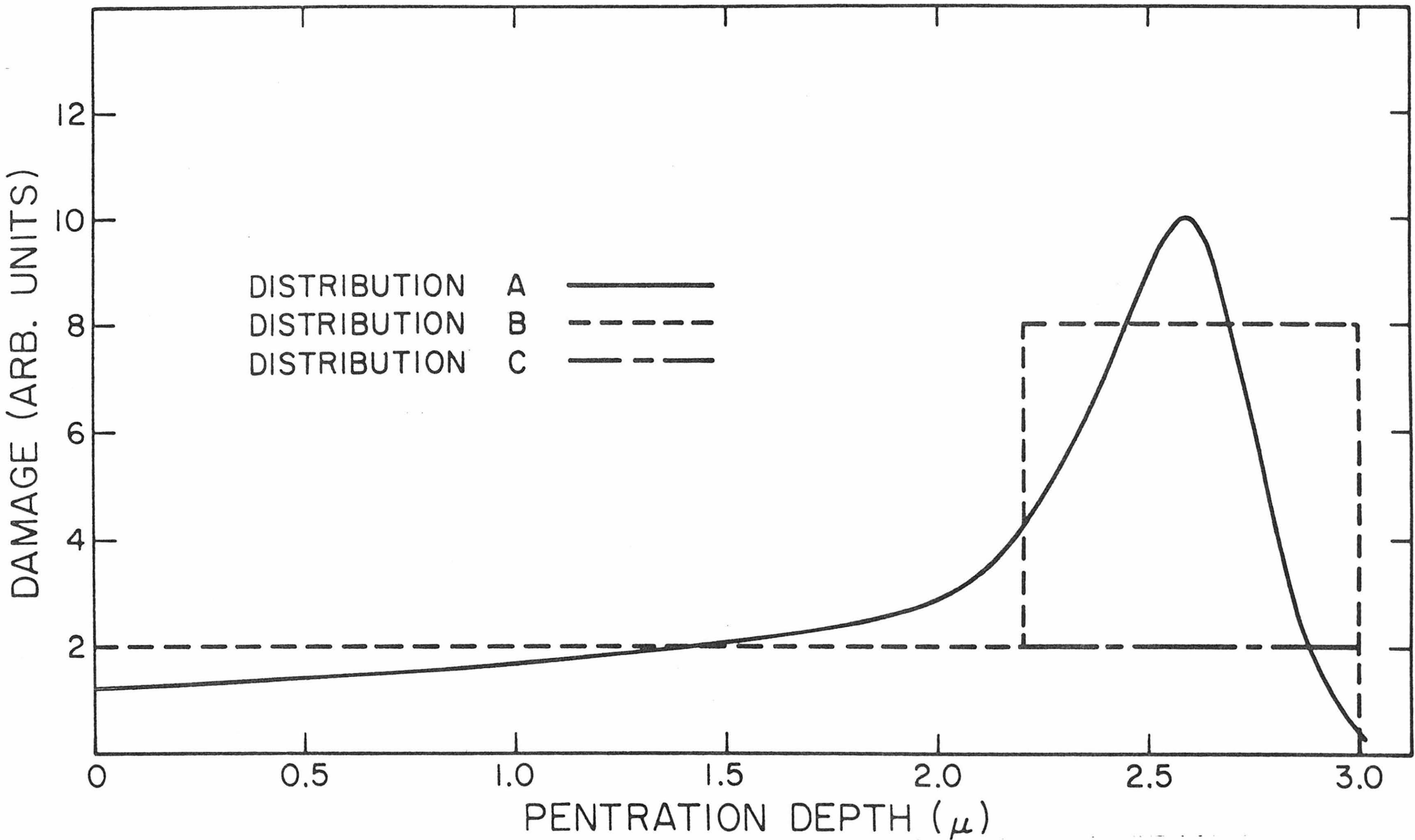


Figure 4-4. Distribution A: Electrical compensation center distribution as derived from the data of Wohlleben and Beck and Pruniaux et al. Distributions B and C: Approximate distributions used for mode behavior analysis.

III. Optical Waveguiding

In order to assess the effect that the non-uniform damage profile has an optical waveguiding and in particular the cutoff condition, the damage profile will be approximated by curve B of Fig. 4-4.

The equations which were derived in Chapter 2 for the case of a four media waveguide will be used to evaluate the waveguiding properties of the approximated damage profile. For convenience, they are repeated here. Referring to Fig. 4-4, the T.E. electric field distribution is given by:

$$E_y^{(m)}(x) = \begin{cases} Ae^{qx} & x \leq 0\mu \\ B\sin(hx) + C\cos(hx) & 0\mu \leq x \leq 2.2\mu \\ D\sin(\ell x) + E\cos(\ell x) & 2.2\mu \leq x \leq 3\mu \\ F\,e^{-px} & 3\mu \leq x \end{cases} \qquad (2\text{-}10)$$

where m is the mode order. Substitution of Eq. (2-10) into Maxwell's equations and use of the appropriate boundary conditions yields:

$$n_1^2 k_o^2 = \beta_m^2 - q^2 \qquad (2\text{-}11a)$$

$$n_2^2 k_o^2 = \beta_m^2 + h^2 \qquad (2\text{-}11b)$$

$$n_3^2 k_o^2 = \beta_m^2 + \ell^2 \qquad (2\text{-}11c)$$

$$n_4^2 k_o^2 = \beta_m^2 - p^2 \qquad (2\text{-}11d)$$

and:

$$h\left[\frac{q - h\tan(ha)}{h + q\tan(ha)}\right] = -\ell\left[\frac{p - \ell\tan(b-a)\ell}{\ell + p\tan(b-a)\ell}\right] \qquad (2\text{-}14)$$

where β_m is the phase vector of the mth mode.

Eliminating β_m from Eqs. (2-11) gives:

$$q = \sqrt{(n_2^2 - n_1^2)k_o^2 - h^2} \qquad (4\text{-}1a)$$

$$\ell = \sqrt{(n_3^2 - n_2^2)k_o^2 + h^2} \qquad (4\text{-}1b)$$

$$p = \sqrt{(n_2^2 - n_4^2)k_o^2 - h^2} \qquad (4\text{-}1c)$$

$$= \sqrt{(n_3^2 - n_4^2)k_o^2 - \ell} \quad . \qquad (4\text{-}1d)$$

Equation (2-46a) may now be used to relate those relative dielectric constant discontinuities, $K_i - K_j \equiv n_i^2 - n_j^2$, which arise solely from the plasma depression effect to the appropriate free carrier concentrations. (The present assertion that the plasma depression effect gives rise to optical confinement is substantiated in Section V.) Doing so, Eqs. (4-1c) and (4-1d) become:

$$p = \sqrt{1.24 \times 10^{-20}(N_4 - N_2)k_o^2 - h^2} \qquad (4\text{-}2a)$$

$$= \sqrt{1.24 \times 10^{-20}(N_4 - N_3)k_o^2 - \ell^2} \qquad (4\text{-}2b)$$

where N_i is the free carrier concentration of medium i and an optical wavelength of 1.15μ has been assumed.

It is apparent from Eqs. (4-2) that for $N_2, N_3 \lesssim N_4/10$,

$$p \approx 1.24 \times 10^{-20} N_4 - h^2 \qquad (4\text{-}3a)$$

$$\approx 1.24 \times 10^{-20} N_4 - \ell^2 \qquad (4\text{-}3b)$$

and that, under these conditions, $h \approx \ell$. Therefore, when N_2, $N_3 \lesssim N_4/10$, the optical mode profile is governed almost exclusively by N_4 (realizing that the refractive index discontinuity between medium 1--usually air-- and medium 2 is orders of magnitude larger than those index discontinuities caused by the plasma depression effect) and, hence, is virtually independent of the detailed shape of the damage distribution. However, as the damage profile is reduced in magnitude by, for example, annealing, and the waveguide approaches cutoff, N_2 will become comparable to N_4 while N_3 may remain much smaller than N_4. As a result, the cutoff condition may be modified slightly.

The cutoff conditions of the four media guide under consideration and those of a three media guide, such as would result from the uniform damage distribution shown as curve C of Fig. 4-4, are now compared assuming that $N_4 = 10^{18}$ free electrons cm^{-3}-- a representative substrate free carrier concentration.

Cutoff occurs when total internal reflection is no longer possible at one of the guide boundaries, in our case at the guide-substrate interface. The conditions for cutoff of the four media guide may be found by solving Eqs. (2-11) and (2-14) when $p = 0$; they may be for the three media guide by first setting $N_2 = N_3$

and $b = a$ and then taking $p = 0$. The resulting cutoff value of N_2 for the three media waveguide is found to be $4.4 \times 10^{17} cm^{-3}$. In order to determine the cutoff values of N_2 and N_3 for the four media structure, curve B of Fig. 4-4 and Fig. (3-17) must be used in conjunction with Eqs. (2-11) and (2-14). The procedure is as follows: surface and peak defect densities are chosen in the ratio of one to four in accordance with curve B of Fig. 4-4; Fig. 3-17 is then used to find corresponding surface and peak free carrier concentrations; using a computer program, Eqs. (2-11) and (2-14) are solved for p; the routine is repeated until $p \cong 0$. Doing so yields $N_2 = 6.5 \times 10^{17} cm^{-5}$ and $N_3 = 10^{17} cm^{-3}$. Hence, at cutoff, N_2 for a four media guide is nearly equal to the cutoff value of N_2, $4.4 \times 10^{17} cm^{-3}$, for a three media guide. Experimentally, it is found that differences in free carrier concentrations this small cannot be inferred from mode profile measurements; a guided mode near cutoff becomes masked by so-called substrate radiation modes [18]. From a waveguiding standpoint, therefore, the damage distribution may be taken to be uniform (ref: curve C of Fig. 4-4) and equal in magnitude to an appropriate average over the leading (nearly level) portion of the nonuniform damage distribution.

IV. Optical Attenuation

In Section V the attenuation coefficient of a proton-implanted waveguide, measured under isochronal annealing conditions, will be used to determine the absorption center concentration in the leading, nearly uniformly damaged portion of the waveguiding layer. From this knowledge, the free carrier concentration in this region will

be obtained and, following the results of Section III, a theoretical optical mode profile calculated. Theoretical and experimental mode profiles will then be compared. In order for this procedure to be valid, the attenuation coefficient must be related through a proportionality constant to the magnitude of the electrically active defect center distribution. For this to be the case two requirements must be met: First, the absorption center distribution and the optical mode intensity profile (within the damaged layer) must maintain their shapes subsequent to heat treatment. Second, the electrically active defect distribution must be proportional to the absorption center distribution. The extent to which these requirements are satisfied will now be considered.

Only the attenuation coefficients of first order (single lobe) modes were measured. An examination of Fig. 4-11 reveals that within the guiding (damaged) layer the shape of the first order modes varied only slightly throughout the entire annealing range (300°C-500°C) with mode intensity peaks generally appearing 2.5-3μ from the waveguide surface. On the other hand, because of those experimental difficulties discussed in Section II, a similarly detailed knowledge of the compensation center profile as a function of annealing parameters is not available. Nevertheless, based on the data shown in Fig. 4-2, which indicate that the damage peak remains approximately stationary after annealing, the assumption will be made that the damage distribution also retains its shape subsequent to heat treatment, simply being scaled down further and further following succeeding annealing cycles. Adequate justification of this point of view is offered by

Figs. 4-11a to 4-11f which compare experimentally observed mode profiles with those obtained using the procedure (which is dependent on the present assumption for its validity) outlined at the beginning of this section. For example, it is to be noted that second order mode cutoff, a sensitive indicator of the number of residual free carriers in the waveguiding region, occurs at exactly those temperatures predicted by the theoretical approach based on the present assumption.

The question of whether the compensation center distribution is proportional to the overall absorption center distribution remains moot. It will be recalled that in Chapter 3 two contributions to the absorption at 1.15μ (the wavelength used here to measure guide parameters) were proposed; dipole-induced transitions between electrically active (compensating) defect levels and the band continua and strictly dissipative absorption and scattering by regions of local precipitation. The cross section for producing these two types of defects are no doubt different functions of projectile energy: precipitate formation seems to increase with the kinetic energy of the projectile (see Section 3-111A1) whereas the cross section for Rutherford scattering varies inversely with projectile energy (ref: Eq. 1-16a). Consequently, it seems likely that a larger fraction of the total absorption near the surface of the damaged layer is due to precipitate-induced attenuation than is the case near the damage peak where the kinetic energy of the incident protons has probably fallen below the threshold required for precipitate formation. A convenient way of characterizing the resulting mix of absorption mechanisms is to define an average contribution for each over the

entire damage layer. Since gallium arsenide specimens approximately twice as thick ($\sim$10μ) as the maximum proton penetration depth ($\sim$4μ) were used to obtain the absorption spectrum shown in Fig. 3-1b, the absorption data presented therein represent just such an average. Fig. 3-1b indicates that at 1.15μ (1.08 eV) about 40% of the absorption may be attributed to electrically active centers. We therefore assume that the electrical and absorption center distributions are proportional to one another with the proportionality constant taken to be 0.4.

Having discussed the two requirements mentioned at the beginning of this section, we now examine the effect which the non-uniform damage profile has on the optical attenuation coefficient. As was seen in Section III its effect on the optical waveguiding properties was minimal. Figure 4-5 shows a representative first order mode profile (curve A) together with the actual damage profile (curve B) and the approximate damage profile (curve C) which may be used, in accordance with the conclusions reached in Section III, to calculate optical confinement. Multiplying the ordinate of curve A by, in turn, the ordinates of curves B and C yields the product curves shown in Fig. 4-6, the areas of which are approximately in the ratio of one to one and one-half. It is seen, therefore, that while both the uniform (curve C) and the non-uniform (curve B) damage profiles give use to nearly the same degree of optical confinement, the peak of the non-uniform distribution increases the attenuation coefficient by roughly a factor of 1.5 over that value which would obtain in its absence. The significance of this fact with regard to the possible use of proton-implanted waveguides as active devices will be discussed at the end of this chapter.

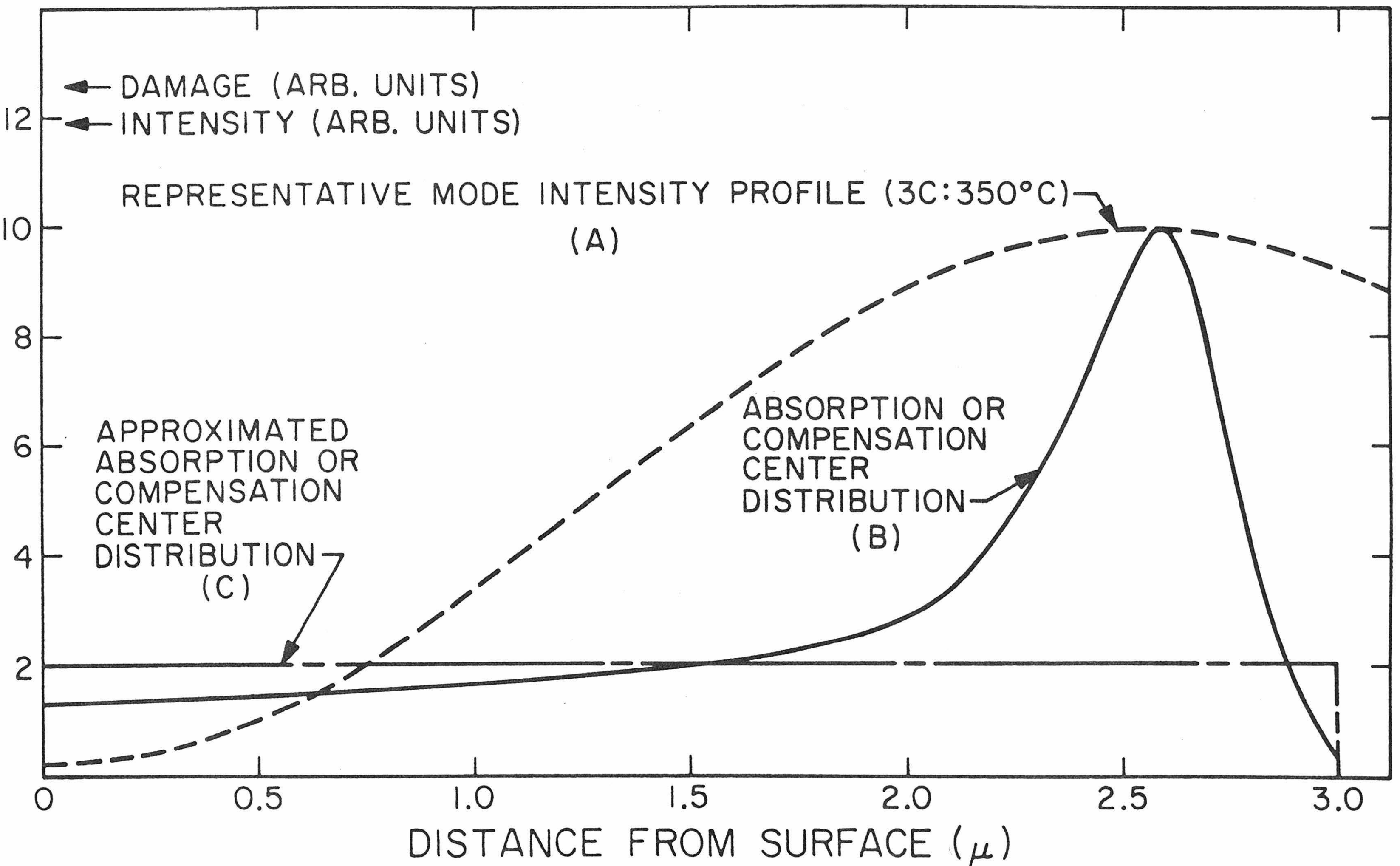


Figure 4-5. Mode intensity profile and defect center distributions used to calculate relative absorption coefficients.

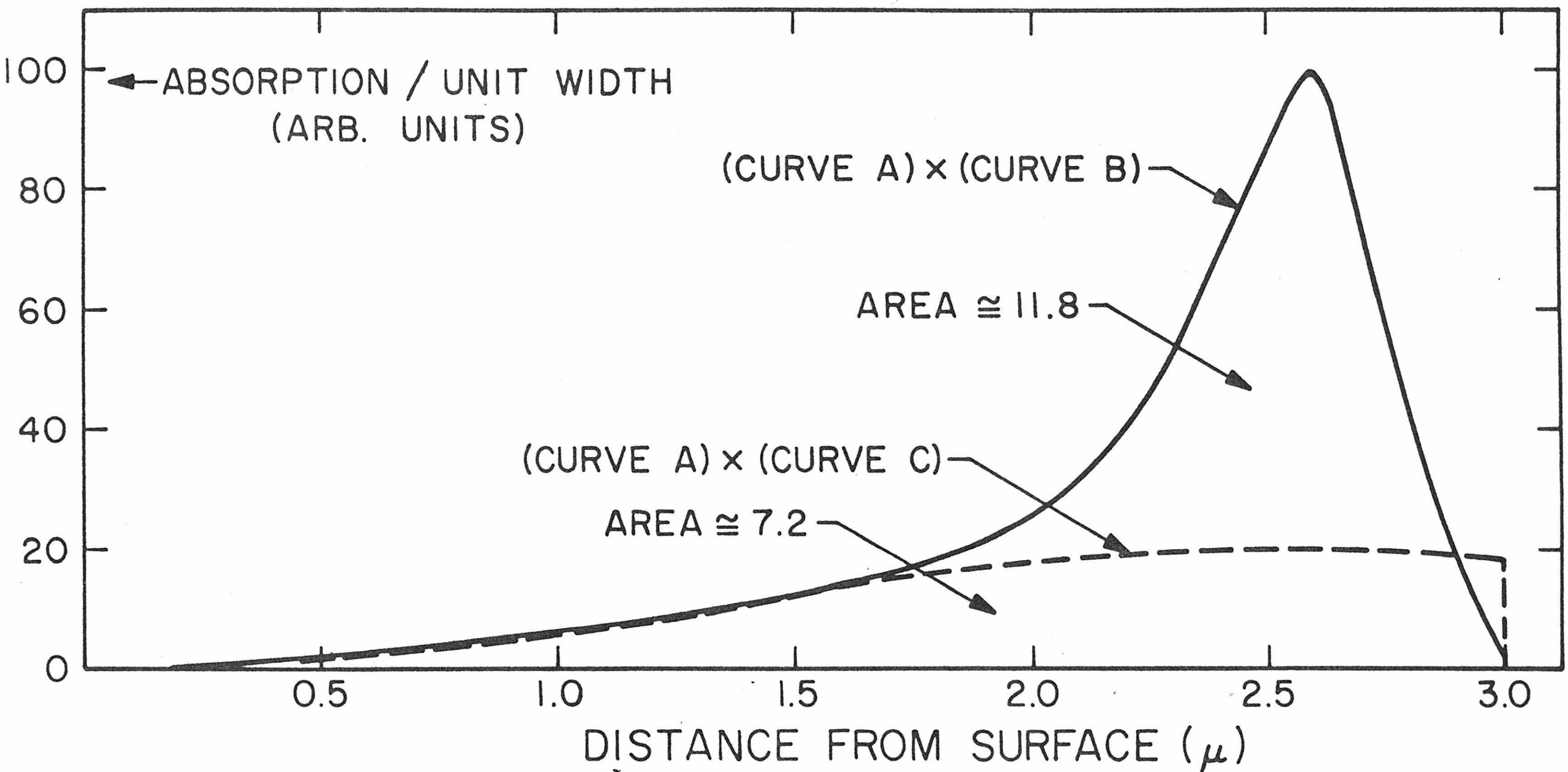


Figure 4-6. Product plots obtained by multiplying curve A, a representative mode intensity profile, by the inferred and approximated defect center distributions. Areas under the curves are proportional to absorption.

V. Data Analysis

One of the main objectives of this study has been the identification of the mechanism responsible for the refractive index discontinuity which gives rise to optical waveguiding in proton-implanted gallium arsenide. It has already been established theoretically (see Section 3-IIIB) that the guide substrate refractive index discontinuity arising from the plasma depression effect is one to two orders of magnitude greater than the contribution arising, through the Kramers-Kronig relationship, from damage-induced modifications of the absorption spectrum. Discounting any changes in the refractive index caused by disorder-associated density fluctuations [57], it would appear, therefore, that the plasma depression-effect is the operative mechanism. The most direct way of substantiating this hypothesis would be first to measure the free carrier concentration of an implanted layer and then to calculate a theoretical mode profile; a comparison between theoretically and experimentally obtained mode profiles could then be made. Unfortunately, because of the experimental difficulties mentioned in Section II, this method could not be used. An alternative approach, however, would be to infer the free carrier concentration in the guiding layer from the attenuation coefficient in the manner described at the beginning of Section IV; this is the procedure which will be adopted.

We preface the data analysis by describing how a universal free carrier concentration versus defect concentration curve was obtained from the data of Wohlleben and Beck (Fig. 3-14). Curves 4 through 8 of Fig. 3-14, representing n-type material which spanned the

substrate free carrier concentrations encountered here, were scaled to obtain a universal curve in the following way: Curve 6 ($1 \times 10^{18} cm^{-3}$ material) was chosen as a representative reference. Curves 4, 5, 7, and 8 were then overlayed over curve 6 and the factor by which their abscissas had to be multiplied to achieve coincidence noted. The logarithms of these factors were then plotted as functions of the logarithms of the original free carrier concentrations as shown in Fig. 4-7. A best fit straight line having the form:

$$\log_{10}(C/10^{18}) = M_a \log_{10} F_a \qquad (4\text{-}4)$$

was then drawn through the resulting points. Here C represented a given substrate free carrier concentration, F_a was the desired abscissa scaling factor, and M_a was the corresponding slope. Measuring the slope of this straight line yielded:

$$F_a = (C/10^{18})^{-0.9} . \qquad (4\text{-}5)$$

The universal carrier removal curve is shown in Fig. 4-8. It was obtained from curve 6 of Fig. 3-14 by using the proportionality constant, derived in Section 3-IIIC which relates proton dose to defect concentration. An an example of how this curve is used, in order to obtain the carrier removal curve for 3×10^{18} material, its abscissa must be multiplied by 2.66(= $(C/10^{18})^{0.9}$). As a check on the universal curve, when it was used to re-obtain curves 4, 5, 7, and 8 of Fig. 3-14, curve 4 ($1.5 \times 10^{17} cm^{-3}$) deviated most from the original, the deviation being about 20% near the knee

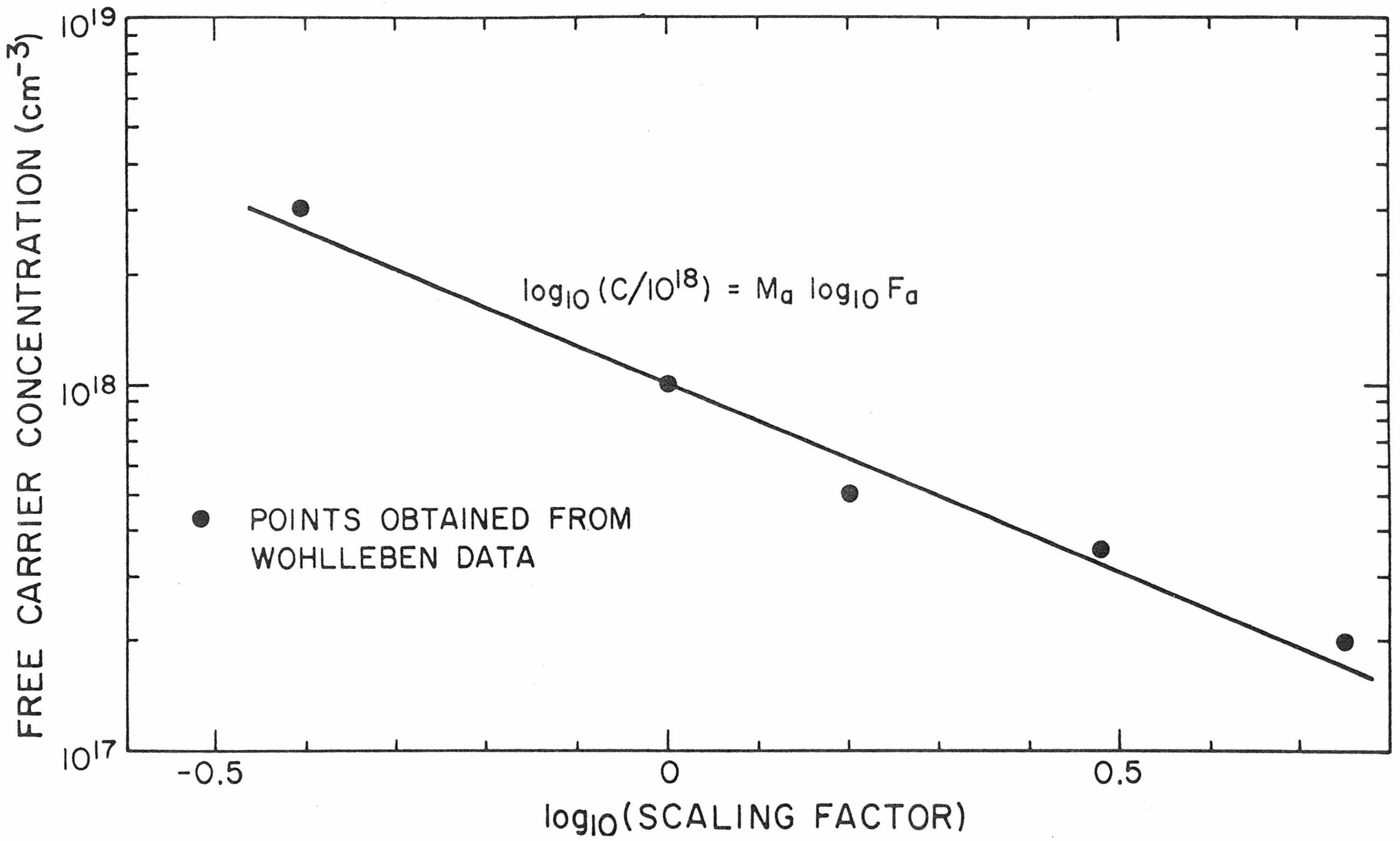


Figure 4-7. Data used to find abscissa scaling factor for universal carrier removal curve.

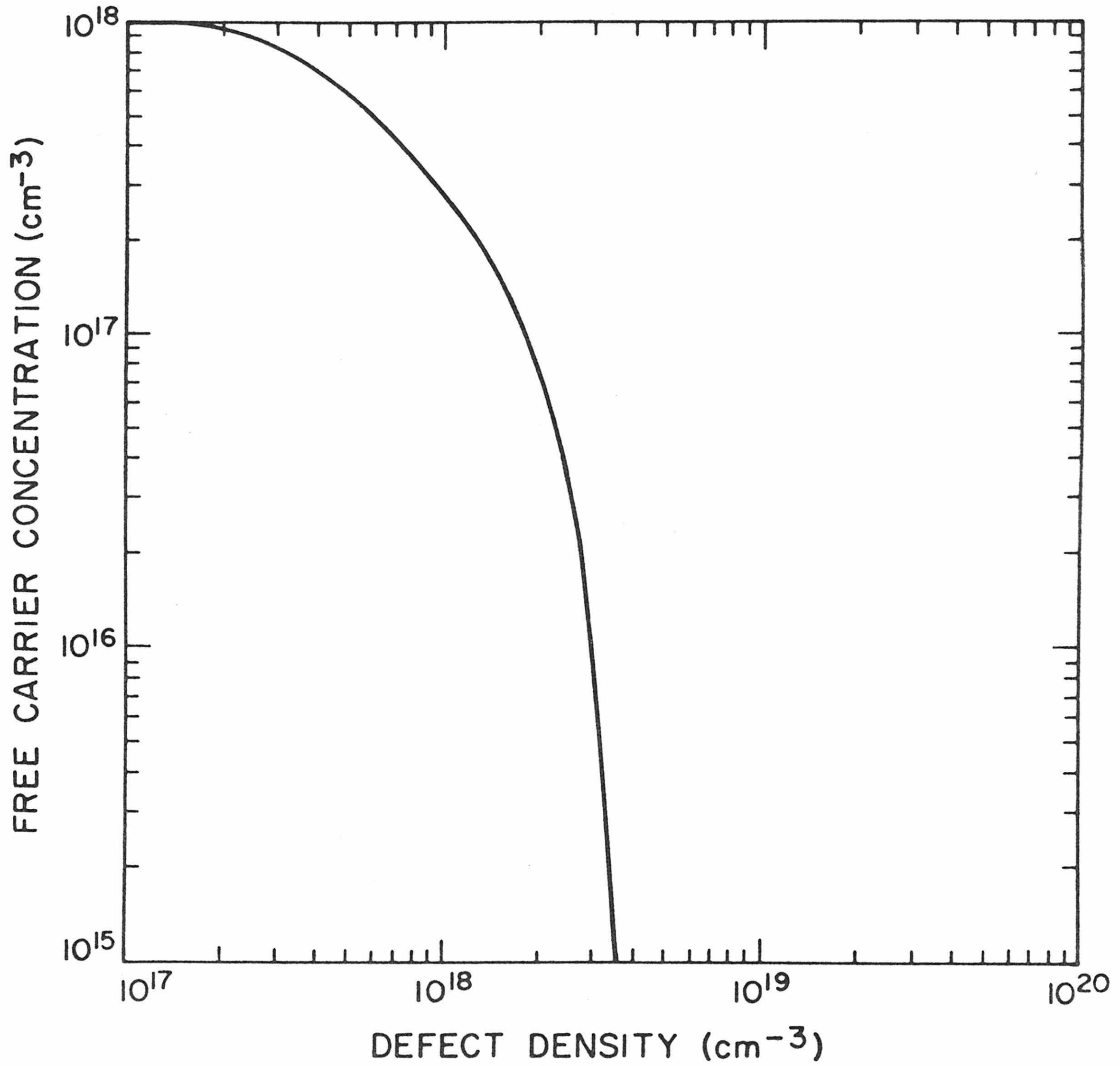


Figure 4-8. Universal free carrier removal curve (for n-type GaAs). To use multiply ordinate by $(C/10^{18})$ and abscissa by $(C/10^{18})^{0.9}$ where C is the new (unimplanted) free carrier concentration.

of the scaled curve. Accuracy to within experimental error is assured for our purposes, however, since the lowest free carrier concentration under consideration (see Table 4-1) was $6.3 \times 10^{17} cm^{-3}$.

In addition to the universal carrier removal curve derived above, the analysis which follows depends on the validity of three important assertions. They are the following: (1) that the disorder-induced energy level spectrum anneals uniformly, viz. that the defect density of states function maintains the same functional form subsequent to heat treatment; (2) that the defect density of states is proportional to the incident proton flux; and (3) that the rate at which the defect levels anneal is independent of donor concentration. The first assertion has been partially verified by the experimental results of Stein. Figure 4-9, after Stein, shows the isochronal annealing behavior of both the integrated absorption spectrum (from 0.2 eV to 1.3 eV) and the absorption at 1.06μ of proton-irradiated gallium arsenide. The fact that these curves are indistinguishable from one another indicates that the annealing characteristics are not dependent upon the particular photon energy selected for observation, at least between 0.2 eV and 1.3 eV. Hence, recalling the functional intimacy which exists between the absorption spectrum in this energy range and the exponential defect density of states function, it is concluded that the density of states function also maintains its shape following heat treatment. We may infer the validity of the second assertion from investigations performed on electron [58] and neutron [59] irradiated GaAs. Both of these studies (See Fig. 4-10) found the absorption

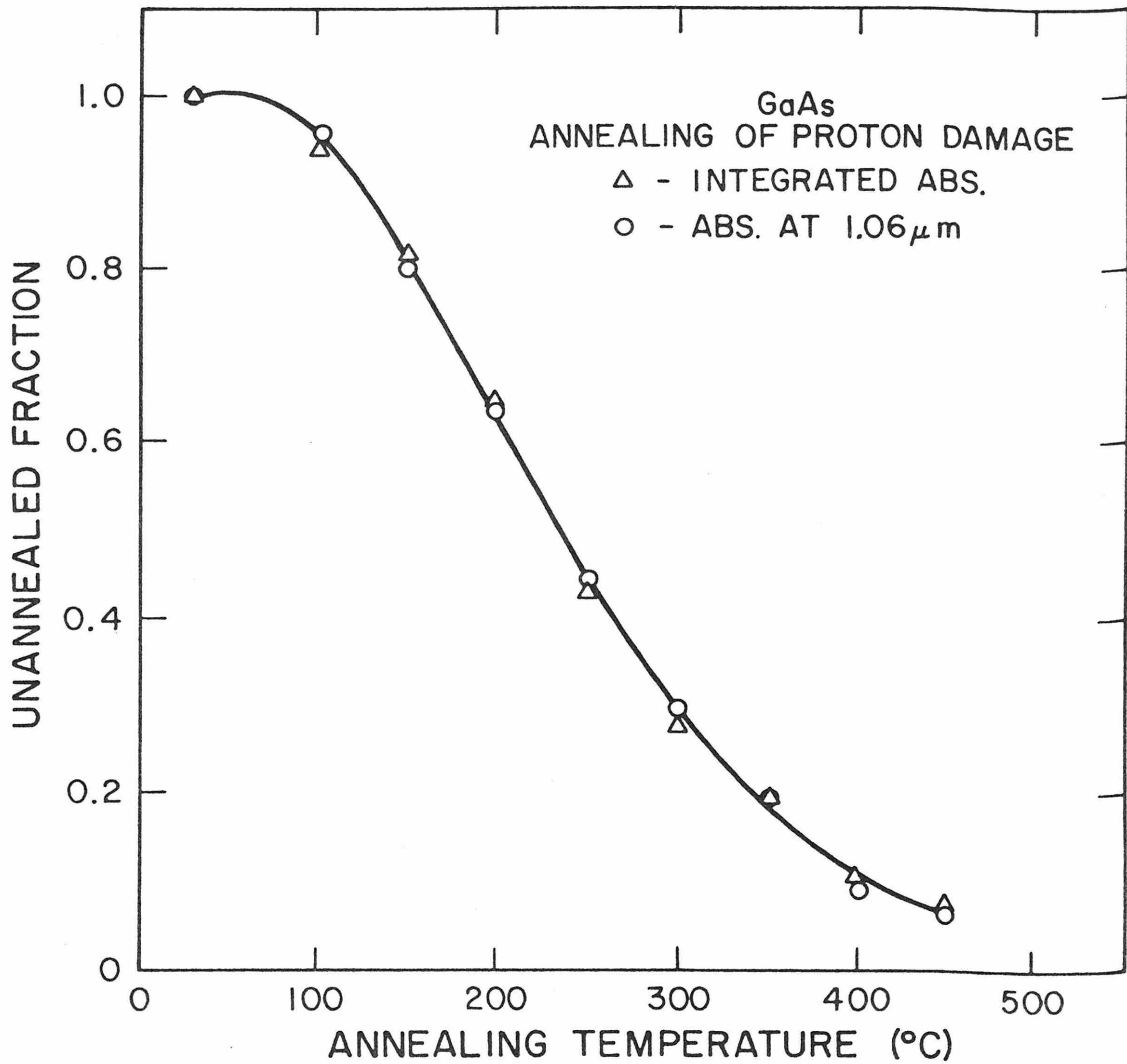
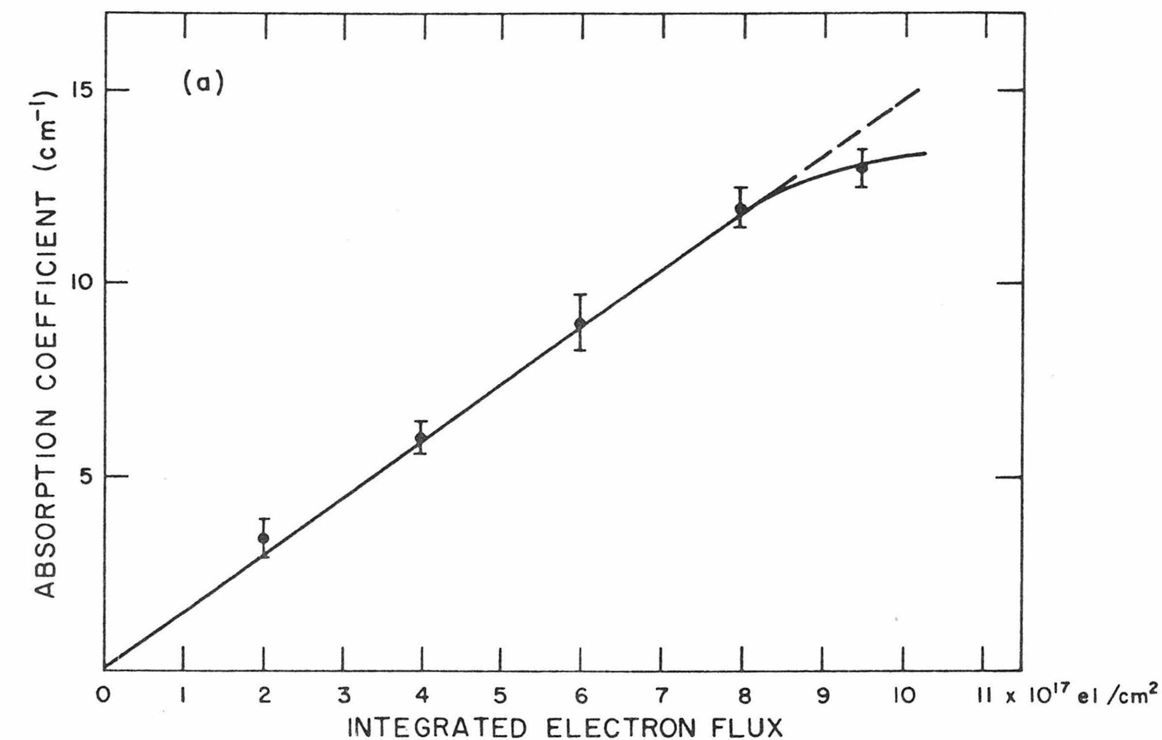


Figure 4-9. Isochronal annealing of proton-induced absorption at 1.06μ and integrated between 0.25 and 1.25 eV. (After Stein [33])



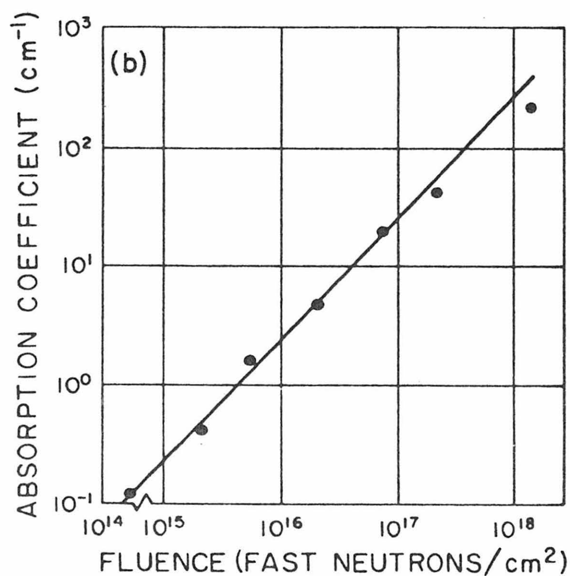


Figure 4-10. a: The absorption coefficient of electron-irradiated GaAs as a function of electron fluence. (After Vaidyanathan and Watt [58]); b: The absorption coefficient of neutron-irradiated GaAs as a function of neutron fluence. (After Burkig et al [59])

coefficient of irradiated GaAs to be proportional to incident particle flux. Since the absorption caused by protons has been found to be the sum of both exponential (electron-like) and squared (neutron-like) energy dependencies, the conclusion may be drawn that proton-induced damage is also proportional to proton fluence. Unfortunately, the validity of the third assertion may not be ascertained from Stein's data; the behavior under annealing conditions of the fundamental edge shift toward lower energies, which was attributed to donor-vacancy complexes in Chapter 3, was not investigated. (In this regard it will be noted that no mention was made of the initial free carrier (donor) concentrations used to obtain Figs. 4-9 and 3-1b--much less whether they were varied during the experiments.) For the time being, however, the assumption will be made that assertion (3) is true; this assumption will be re-examined following the data analysis wherein certain discrepancies between experiment and theory seem attributable to its partial falsehood.

We begin the data analysis by choosing waveguiding specimens whose guiding characteristics may be used to find the proportionality constant which relates the defect density to the waveguide attenuation coefficient; the fact that such a proportionality constant may be defined was established in Section IV. Table 4-1 shows the substrate-dose combinations which were investigated. All implanted waveguides were analyzed under isochronal annealing conditions: one-half hour temperature pulsing in 50°C increments, beginning at 50°C and ending at 500°C--the temperature at which gallium arsenide begins to decompose. Figures 4-11 and 4-12 show the results of the mode profile

TABLE 4-1

D \ N_s	3.25×10^{18} (1)	2×10^{18} (2)	9.5×10^{17} (3)	8.5×10^{17} (4)	6.3×10^{17} (5)
2.0×10^{15} (A)	X	X	X	X	X
8.0×10^{14} (B)	X	X	X	X	X
3.0×10^{14} (C)	X	X	X	X	X
1.5×10^{14} (D)				X	

NOTE: N_s $\equiv$ substrate free carrier concentration (cm^{-3})

D $\equiv$ proton dose (cm^{-2})

Numbers and letters in parentheses indicate substrate concentrations and proton doses, respectively. All combinations shown as "X's" were examined under isochronal annealing (one-half hour temperature pulsing in 50°C increments beginning at 50°C and ending at 500°C).

and mode attenuation measurements which were made following each anneal cycle; details of the experimental procedures used will be described in Chapter 6. Specimens 1C and 5C (see Table 4-1 for symbol definitions), annealed to 450°C, were chosen for the purpose of finding the proportionality constant. Under the initial assumption that the plasma depression effect gave use to the entire dielectric constant discontinuity between guide and substrate, the mode profiles of these specimens were used to find the residual defect concentration of a "C" dose after a 450°C anneal in the following way: Fig. 4-8, the universal free carrier concentration versus defect concentration curve was first scaled to obtain curves appropriate to substrates 1 and 5. A defect concentration was then chosen, by trial and error, which gave, according to the scaled curves, residual free carrier concentrations in the guiding regions of specimens 1C and 5C adequate to produce, as closely as possible, the observed mode profiles. The trial and error theoretical mode profiles for samples 1C:450°C and 5C:450°C were obtained by numerically solving the following equations (see Appendix 2A for a complete description of the parameters) appropriate to the three media waveguiding structure shown in Fig. 4-13:

T.E. modes (only T.E. modes were evaluated):

$$E_y^{(m)}(x) = \begin{cases} A_m \cdot \frac{h}{q} \exp(-qx) & -\infty \leq x \leq 0 \\ A_m \cdot [\frac{h}{q} \cos(hx) - \sin(hx)] & t \geq x \geq 0 \\ A_m \cdot [\frac{h}{q} \cos(ht) + \sin(ht)] \exp[p(x + t)] & \infty \geq x \geq t \end{cases} \quad (2A\text{-}1)$$

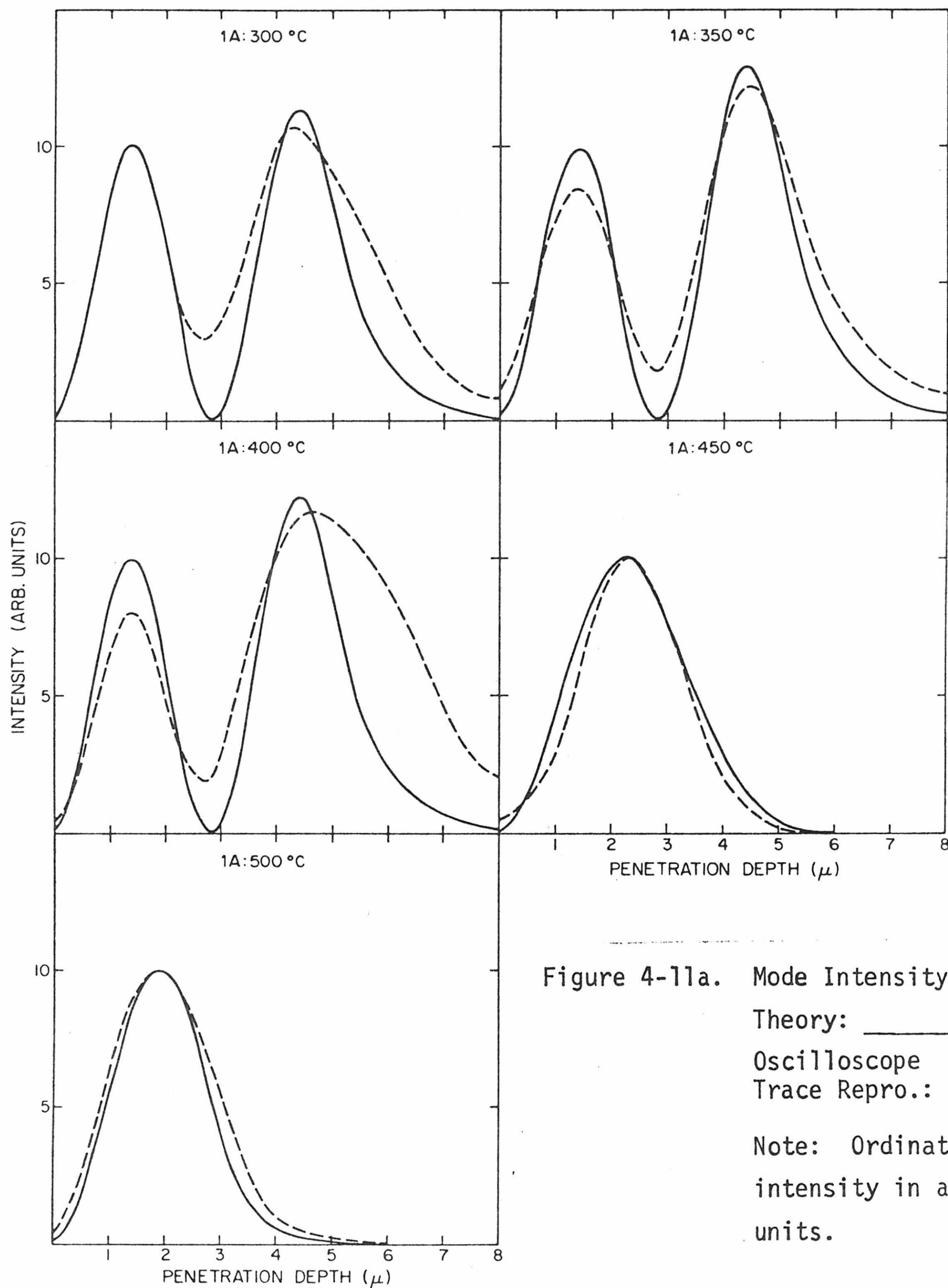


Figure 4-11a. Mode Intensity Profiles

Theory: ____________

Oscilloscope Trace Repro.: ---------

Note: Ordinates are intensity in arbitrary units.

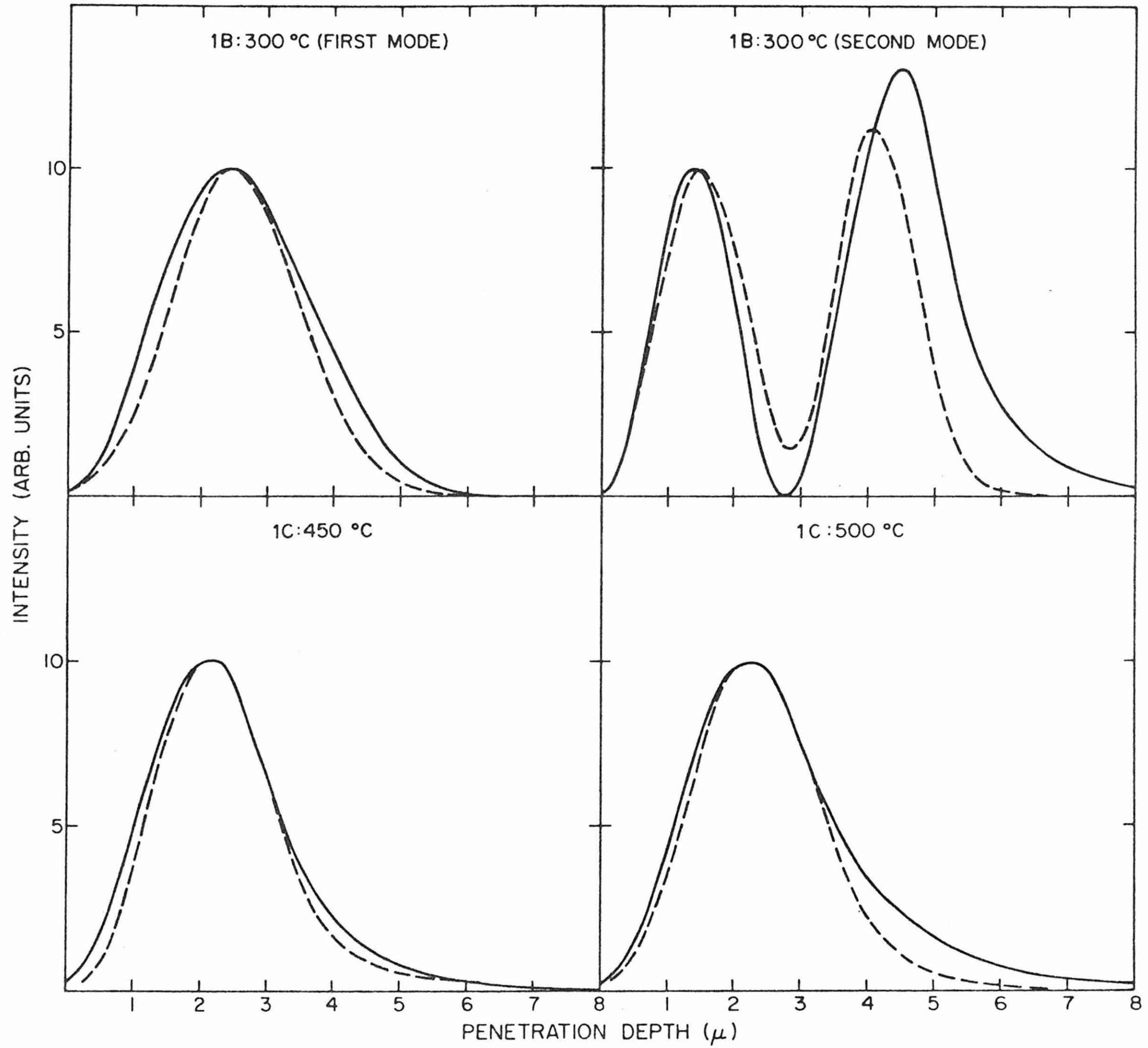


Figure 4-11b. Mode Intensity Profiles

Theory: ____________

Oscilloscope

Trace Repro.: -------

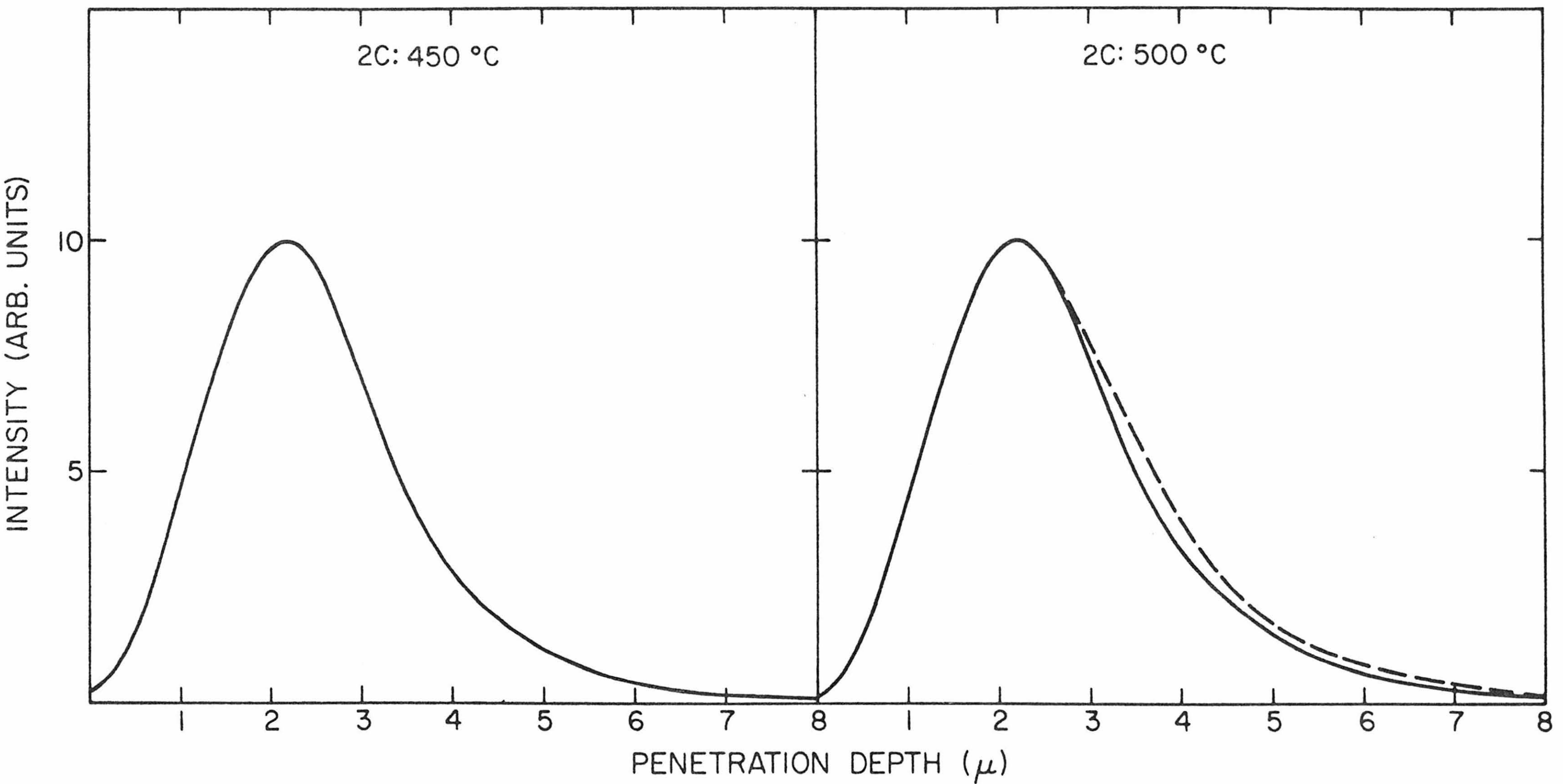


Figure 4-11c. Mode Intensity Profiles

Theory: ————

Oscilloscope

Trace Repro.: --------

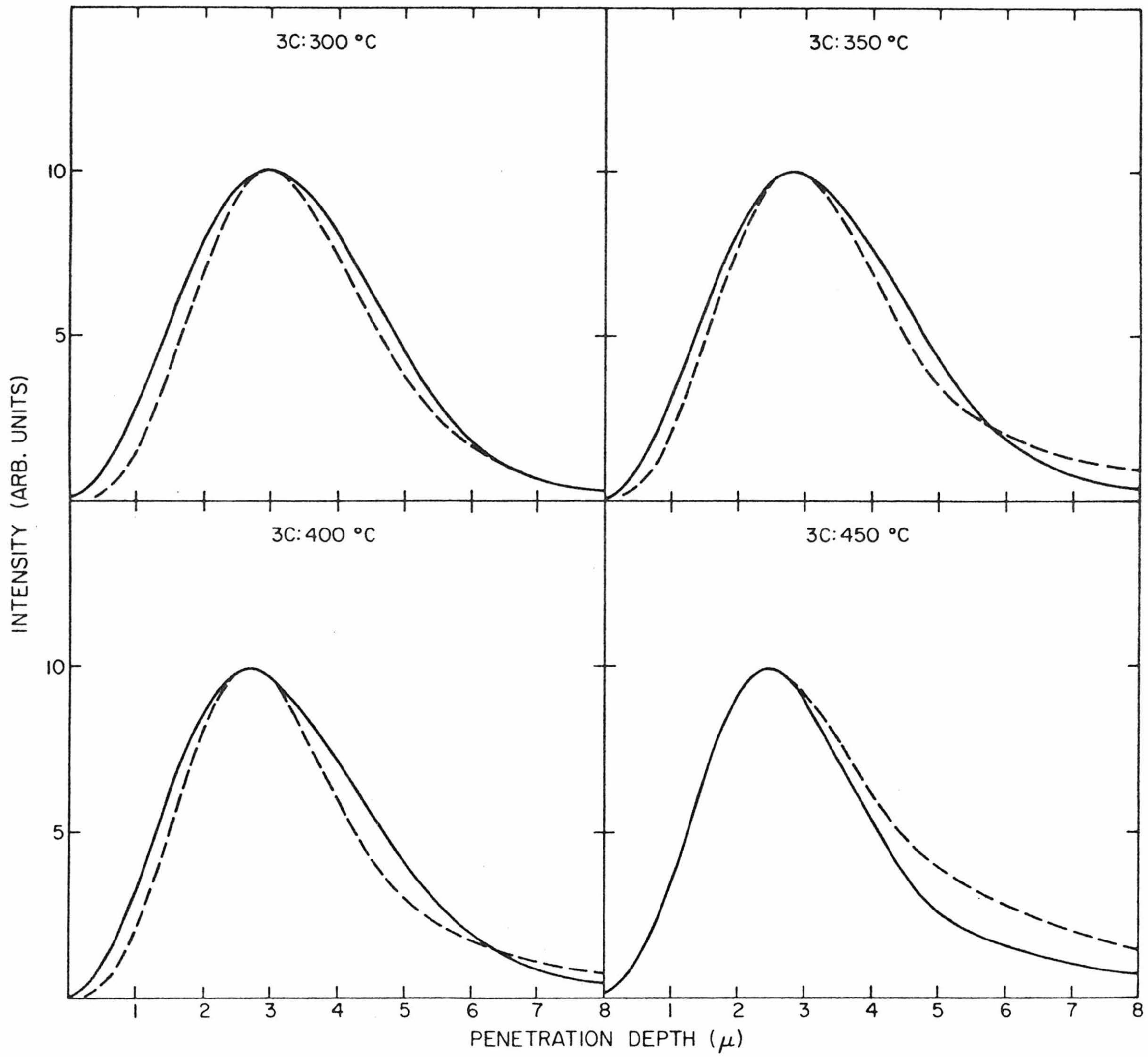


Figure 4-11d. Mode Intensity Profiles

Theory: ______________

Oscilloscope

Trace Repro.: -------

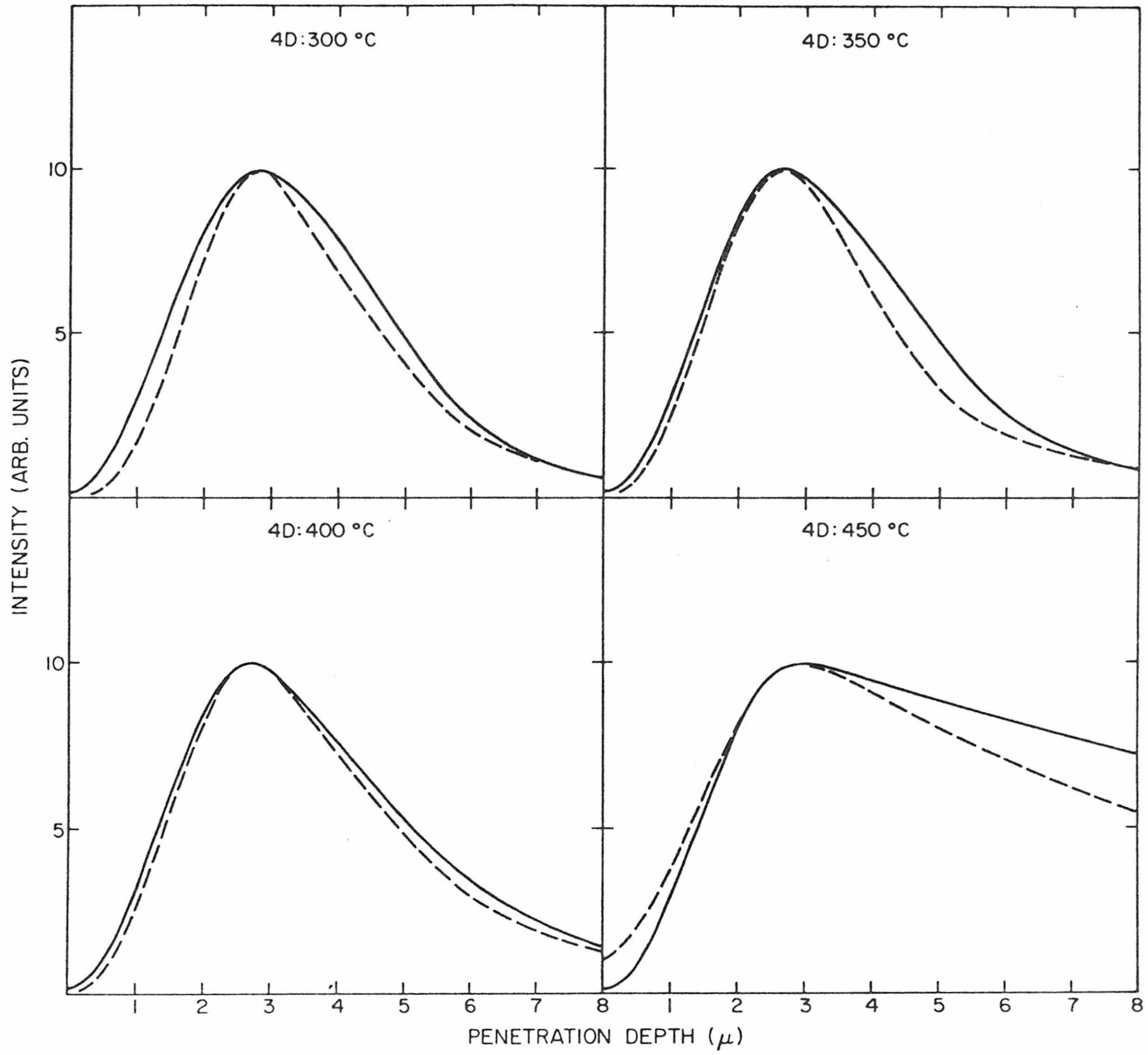


Figure 4-11e. Mode Intensity Profiles

Theory: ______________

Oscilloscope

Trace Repro.: --------

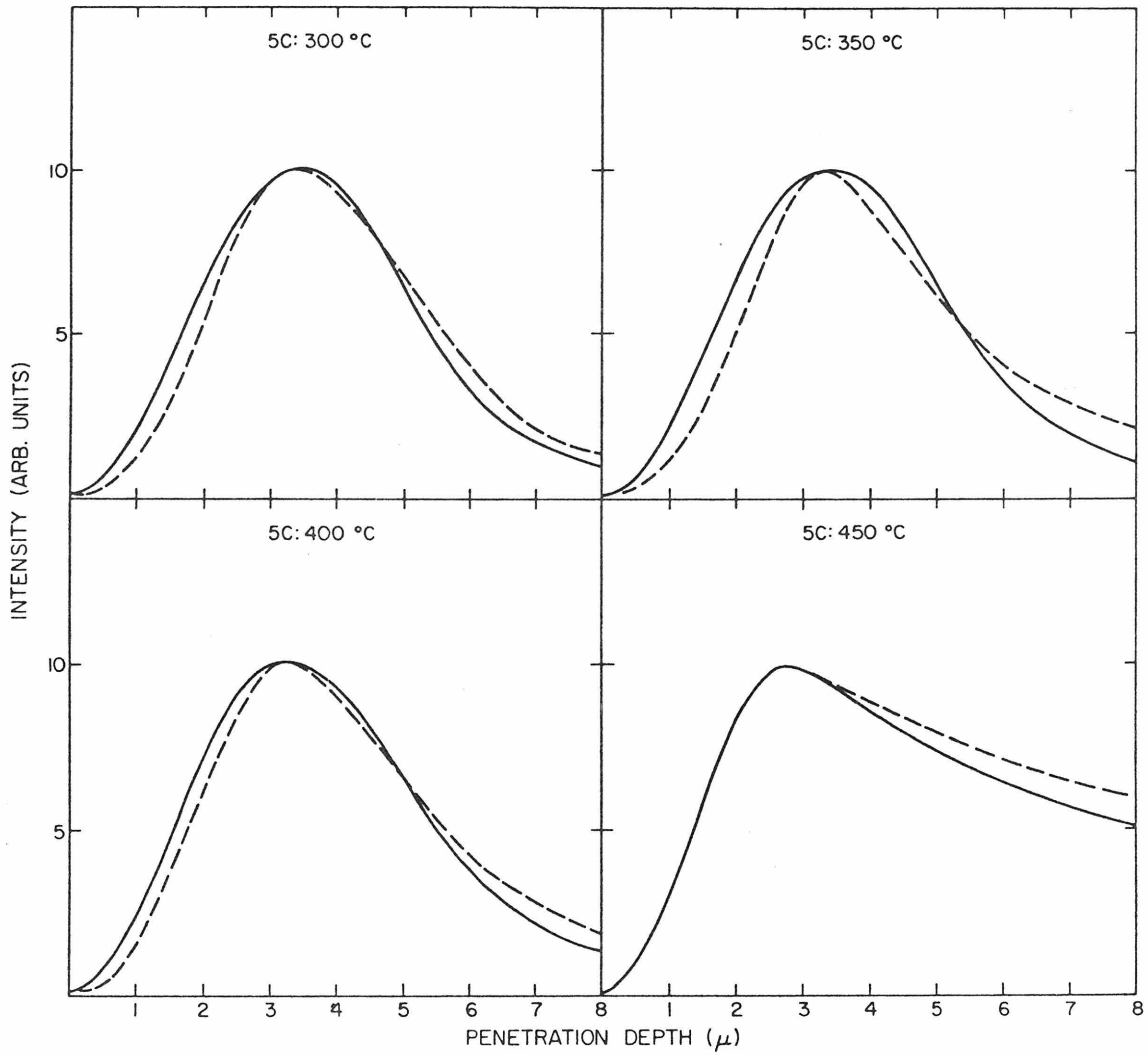


Figure 4-11f. Mode Intensity Profiles

Theory: ________________

Oscilloscope

Trace Repro.: ---------

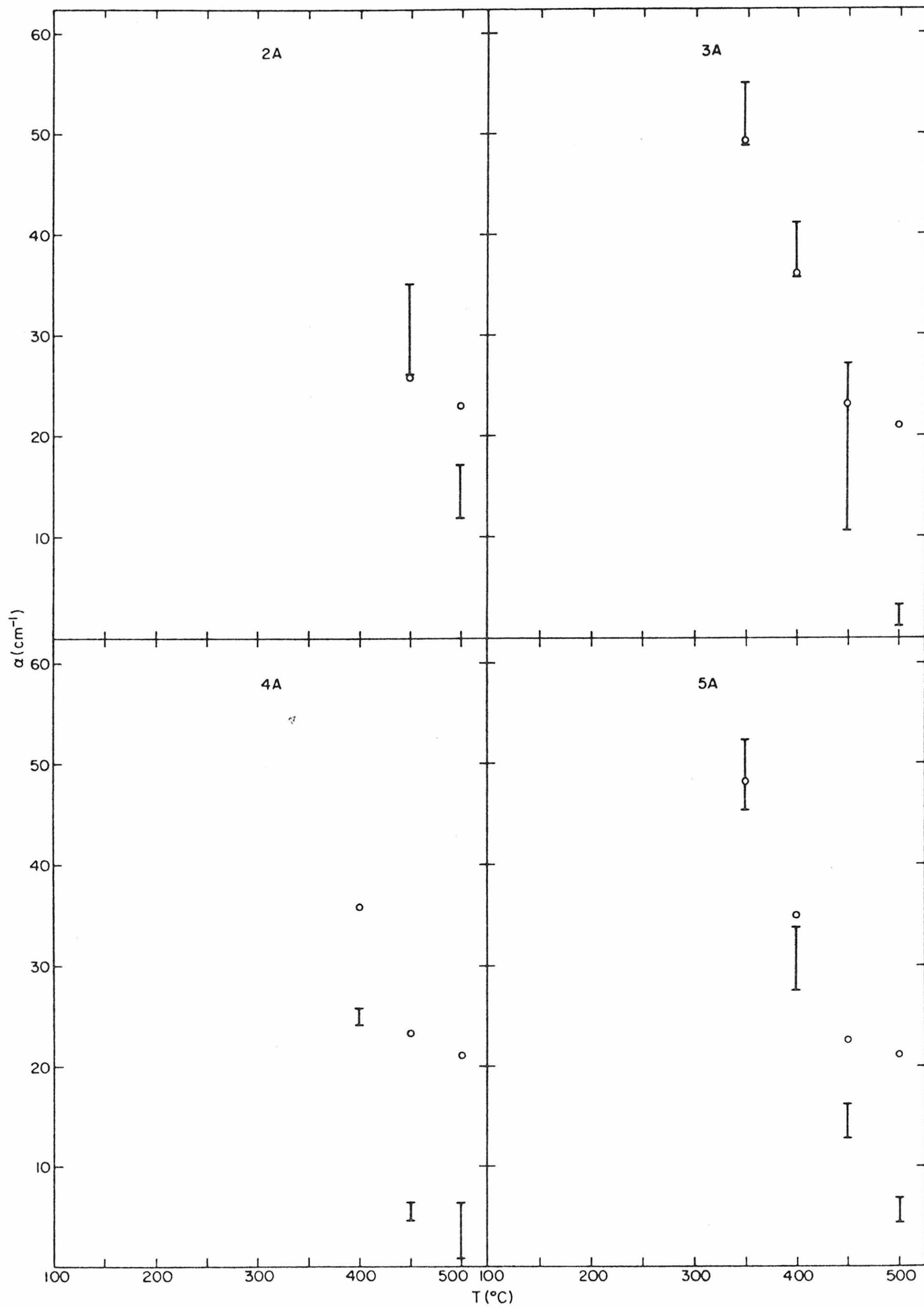


Figure 4-12a. First order mode attenuation coefficient as a function of isochronal annealing (0.5 hour each anneal).

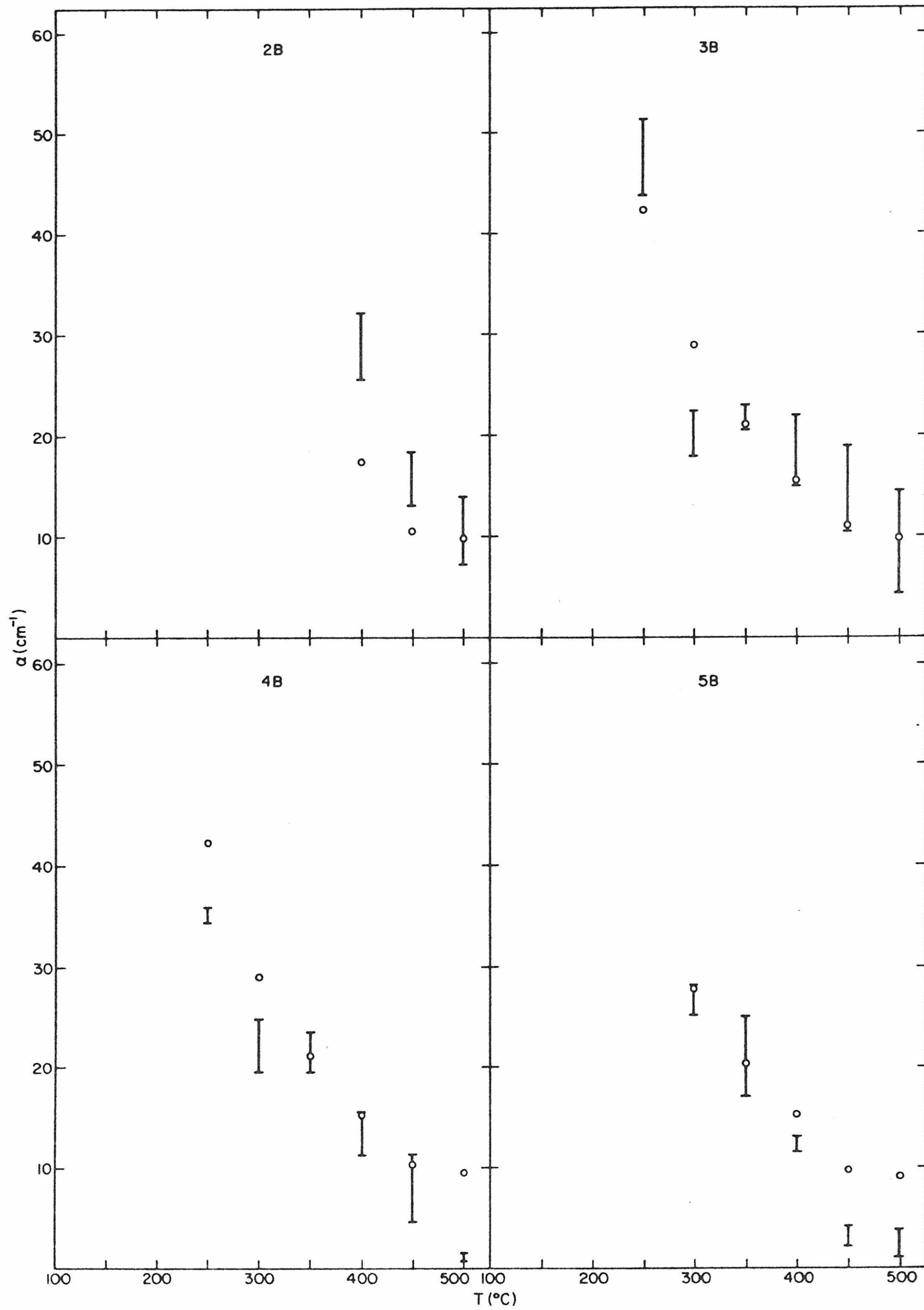


Figure 4-12b. First order mode attenuation coefficient as a function of isochronal annealing (0.5 hour each anneal).

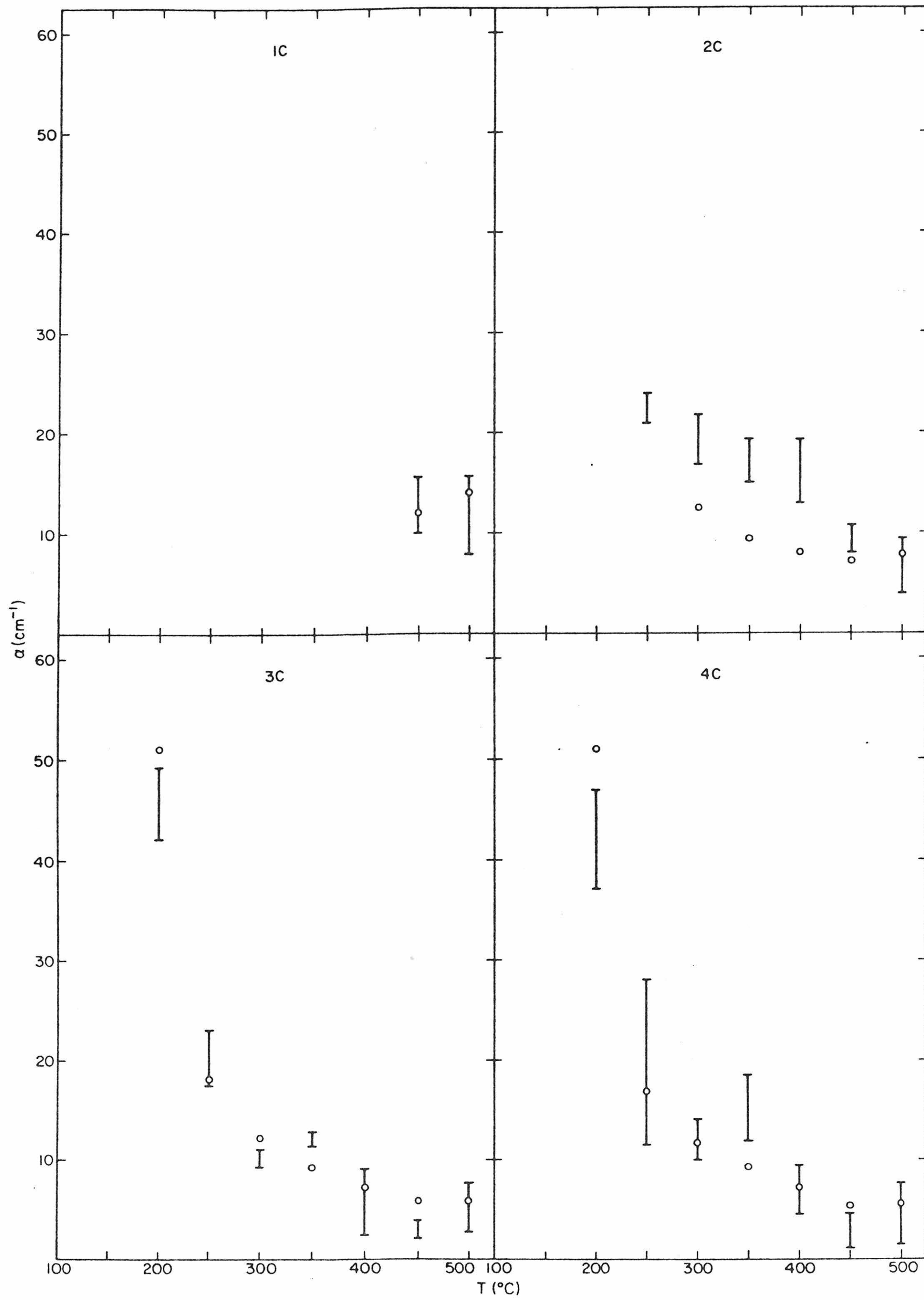


Figure 4-12c. First order mode attenuation coefficient as a function of isochronal annealing (0.5 hour each anneal).

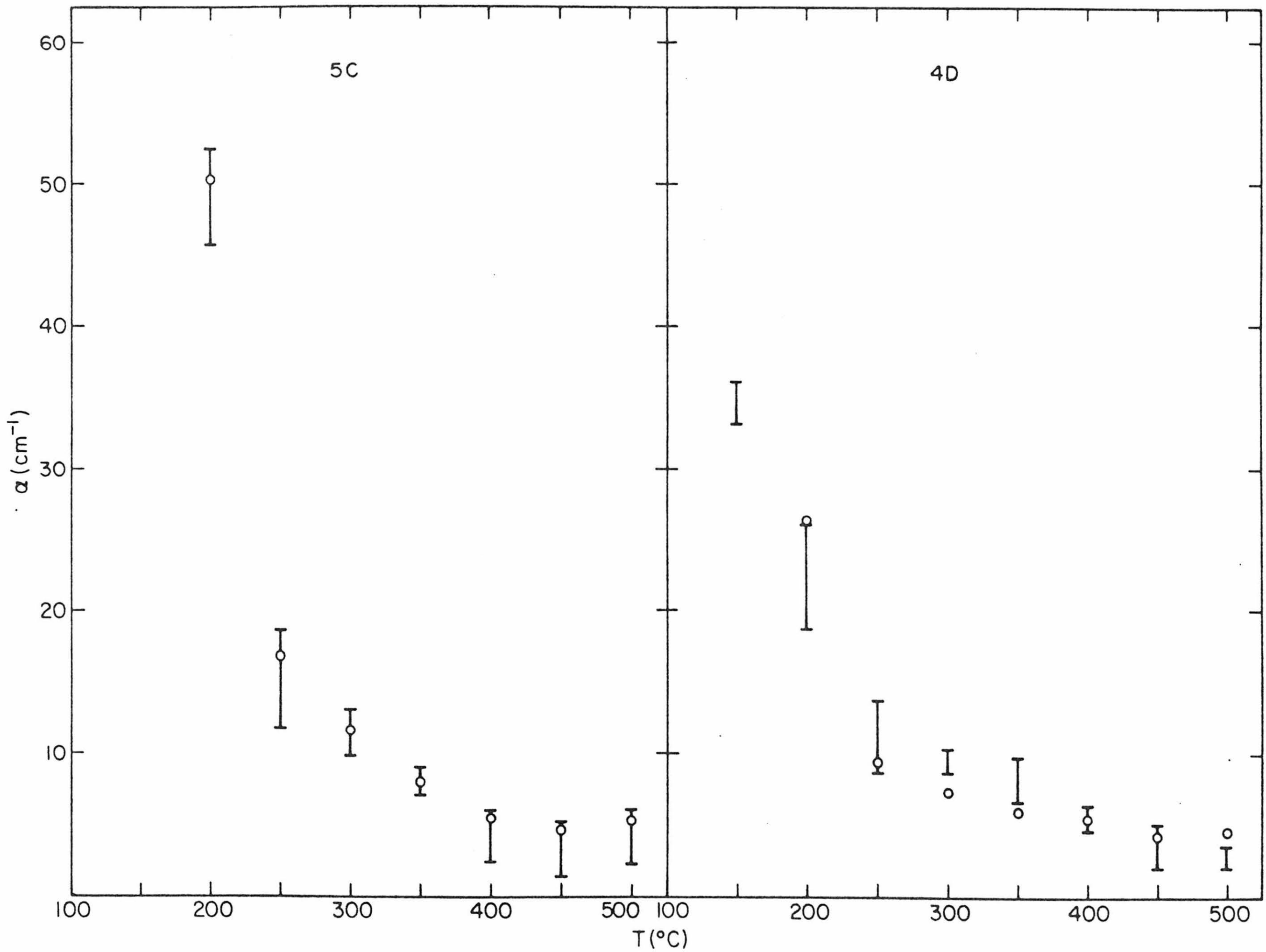


Figure 4-12d. First order mode attenuation coefficient as a function of isochronal annealing (0.5 hours each anneal).

$$n_1^2 k_o^2 = \beta_m^2 - q^2$$

$$n_2^2 k_o^2 = \beta_m^2 + h^2 \qquad (2A\text{-}6)$$

$$n_3^2 k_o^2 = \beta_m^2 - p^2$$

$$\tan(t) = \frac{h(p + q)}{h^2 - p \cdot q} \; . \qquad (2A\text{-}7)$$

The defect concentration which produced the best agreement between experiment and theory was found to be 2×10^{18} defects cm^{-3}. (Recall that, according to the discussion of Section II, this value represents an averaged defect concentration over the nearly uniform region which precedes the damage peak.)

The next step was to find the attenuation coefficient to be associated with a defect population of $2 \times 10^{18} cm^{-3}$. Two mechanisms contribute to the attenuation: free carrier absorption and defect-generated absorption. The bulk free carrier absorption coefficient is given by:

$$\alpha_{f.c.} = \left(\frac{e^3 \lambda_o^2}{4\pi^2 n (m^*)^2 \mu \varepsilon_o C^3} \right) N \qquad (2\text{-}46c)$$

or

$$\alpha_{f.c.}(cm^{-1}) = 1.07 \times 10^{-18} N(cm^{-3}) \qquad (4\text{-}6)$$

where the following values were used: $\lambda_o = 1.15\mu$, $n = 3.4$, $m^* = 0.08m$, and $\mu = 2.3 \times 10^3 cm^2 \cdot volt^{-1} \cdot sec^{-1}$. The values of the effective mass and the mobility correspond to $10^{18} cm^{-3}$ Te-doped material. Figure

4-14 shows experimentally obtained attenuation coefficients of the substrates used in the present study. Taking the slope of a straight line fitted to the data yields:

$$\alpha_{f.c.}(cm^{-1}) = 2.2 \times 10^{-18} N(cm^{-3}) \qquad (4\text{-}7)$$

The discrepancy between (4-6) and (4-7) can probably be attributed to uncertainties in the values of the effective mass and mobility used to obtain (4-6). Because of these uncertainties, (4-7) was considered to be the most reliable relationship and was, therefore, the one used.

Since, as indicated by Eq. (4-7), the magnitude of the wave vector, $|\beta|$ (corresponding to a free space wavelength of $\sim 1\mu$), of a wave propagating in gallium arsenide is modified only slightly by the presence of dispersive free carriers, it was possible to use the perturbative technique developed in Section 2-III to find the free carrier absorption coefficient for waveguided modes. Accordingly, having already obtained the mode profiles of waveguides 1C:450°C and 5C:450°C, the free carrier contributions to their attenuation coefficients were calculated by using the following formula:

$$\alpha_{f.c.}(cm^{-1}) = \frac{2.2 \times 10^{-18}}{\int_{-\infty}^{\infty} |E_y^{(1)}(x)|^2 dx} \left[N_g \int_{-0}^{3} |E_y^{(1)}(x)|^2 dx + N_s \int_{-3}^{-\infty} |E_y^{(1)}(x)|^2 dx \right] \qquad (4.8)$$

where N_g = guide free carrier concentration, N_s = substrate free carrier concentration, and the geometry of Fig. 4-13 was assumed. Using those values of N_g obtained previously for samples 1C:450°C and 5C:450°C, $1.6 \times 10^{18} cm^{-3}$ and $2.2 \times 10^{15} cm^{-2}$, respectively, the

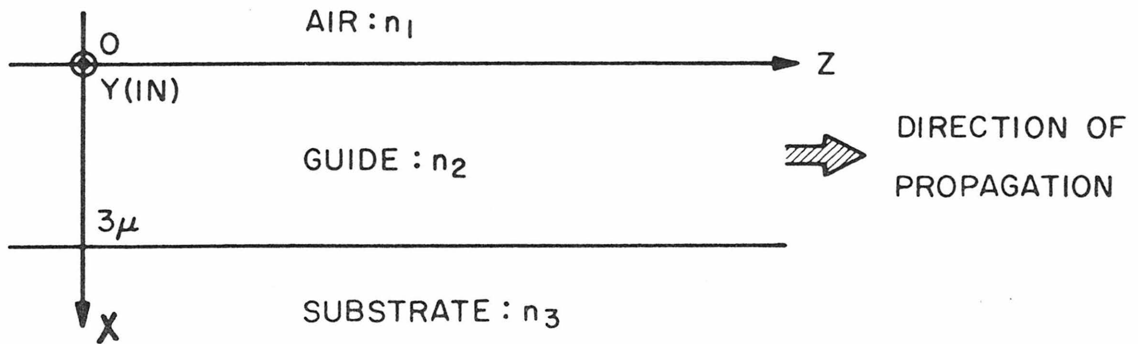


Figure 4-13. Geometry of proton-implanted gallium arsenide waveguide (see Section II for a justification of guide depth)

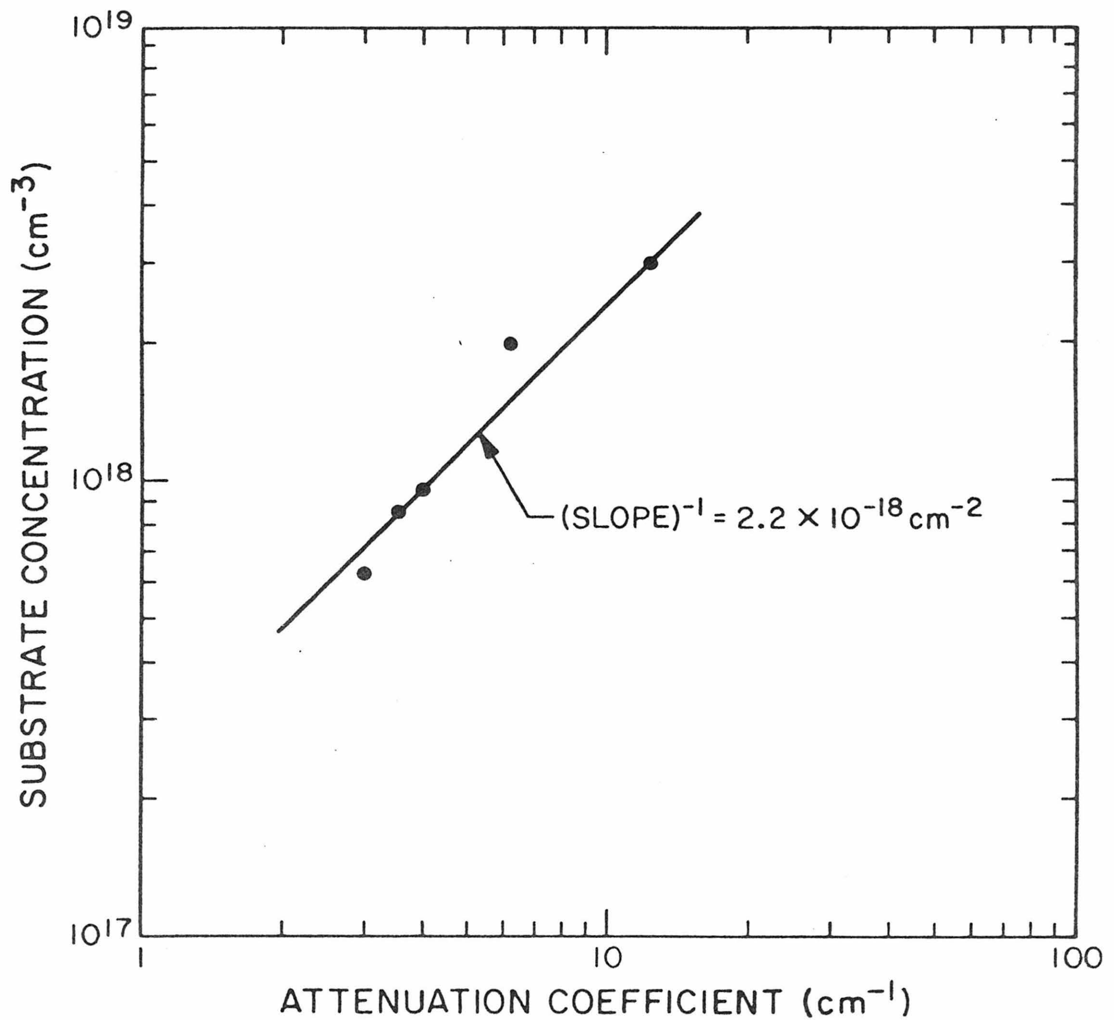


Figure 4-14. Measured attenuation coefficients of the substrates used in the present study.

free carrier absorption contributions for these waveguides were calculated to be:

$$\alpha_{f.c.} = 8.5 \text{ cm}^{-1} \quad (1C{:}450°C) \tag{4-9a}$$

$$\alpha_{f.c.} = 1.7 \text{ cm}^{-1} \quad (5C{:}450°C) \tag{4-9b}$$

The average, total attenuation coefficients of samples 1C:450°C and 5C:450°C, from Figs. 4-12c and 4-12d, were taken to be:

$$\alpha_{tot} = 11.7 \text{ cm}^{-1} \quad (1C{:}450°C) \tag{4-10a}$$

$$\alpha_{tot} = 4.9 \text{ cm}^{-1} \quad (5C{:}450°C)\,. \tag{4-10b}$$

The defect-associated contribution to the attenuation coefficient of a "C" dose waveguide annealed to 450°C was therefore deduced to be ~ 3.2 cm^{-1}. Finally, the relationship between defect concentration and defect-associated attenuation was found to be:

$$\alpha_d(\text{cm}^{-1}) = 1.6 \times 10^{-18} N_d(\text{cm}^{-3})\,. \tag{4-11}$$

Assuming the validity of the three assertions discussed earlier, Eq. (4-11) together with the attenuation coefficients of sample 5C, (see Fig. 4-12d) were used to infer the defect concentrations remaining after each anneal cycle. Sample 5C was chosen for this purpose for the following reasons: (1) Free carrier contributions to the total attenuation coefficients were the smallest of those samples investigated; any uncertainty in knowing the former lead, therefore, to the smallest possible uncertainty in ascertaining the latter.

(2) The free carrier contribution for this sample could be assumed to remain relatively constant over the temperature range of interest (200°C - 500°C): In the complete absence of free carriers in the guiding layer (assumed to be the case after the lowest temperature anneal), for example, Eq. (4-8) yields for the free carrier attenuation coefficient, 1.6 cm^{-1}. This value is to be compared with an estimated $\alpha_{f.c.}$ of 1.8 cm^{-1} at 500°C (based on the fact that sample 5C approached cutoff after the 500°C anneal). (3) Attenuation coefficients were measured for the 5C sample over the entire temperature range investigated (200°C - 500°C).

Following the reasoning given above, Eq. (4-11) together with Fig. 4-12d were used to obtain the defect concentration versus anneal temperature plot (of a C dose waveguide) shown in Fig. 4-15. (In deriving this plot, the free carrier contribution to the attenuation coefficient was estimated to be 1.6 cm^{-1} between the 200°C and the 400°C anneals, inclusive, and 1.7 cm^{-1} and 1.8 cm^{-1} after the 450°C and 500°C anneals, respectively.) In addition, taking the defect density to be proportional to the proton flux, in accordance with the second assertion mentioned earlier, Fig. 4-15 was scaled to obtain defect concentration versus anneal temperature plots for doses A, B, and D. These plots together with scaled free carrier concentration versus defect concentration curves, derived from Fig. 4-8, were then used to find the residual free carrier concentrations in the wave-guiding regions of all substrate-dose combinations after every anneal cycle. When Eqs. (2A-1) were used to obtain optical mode profiles

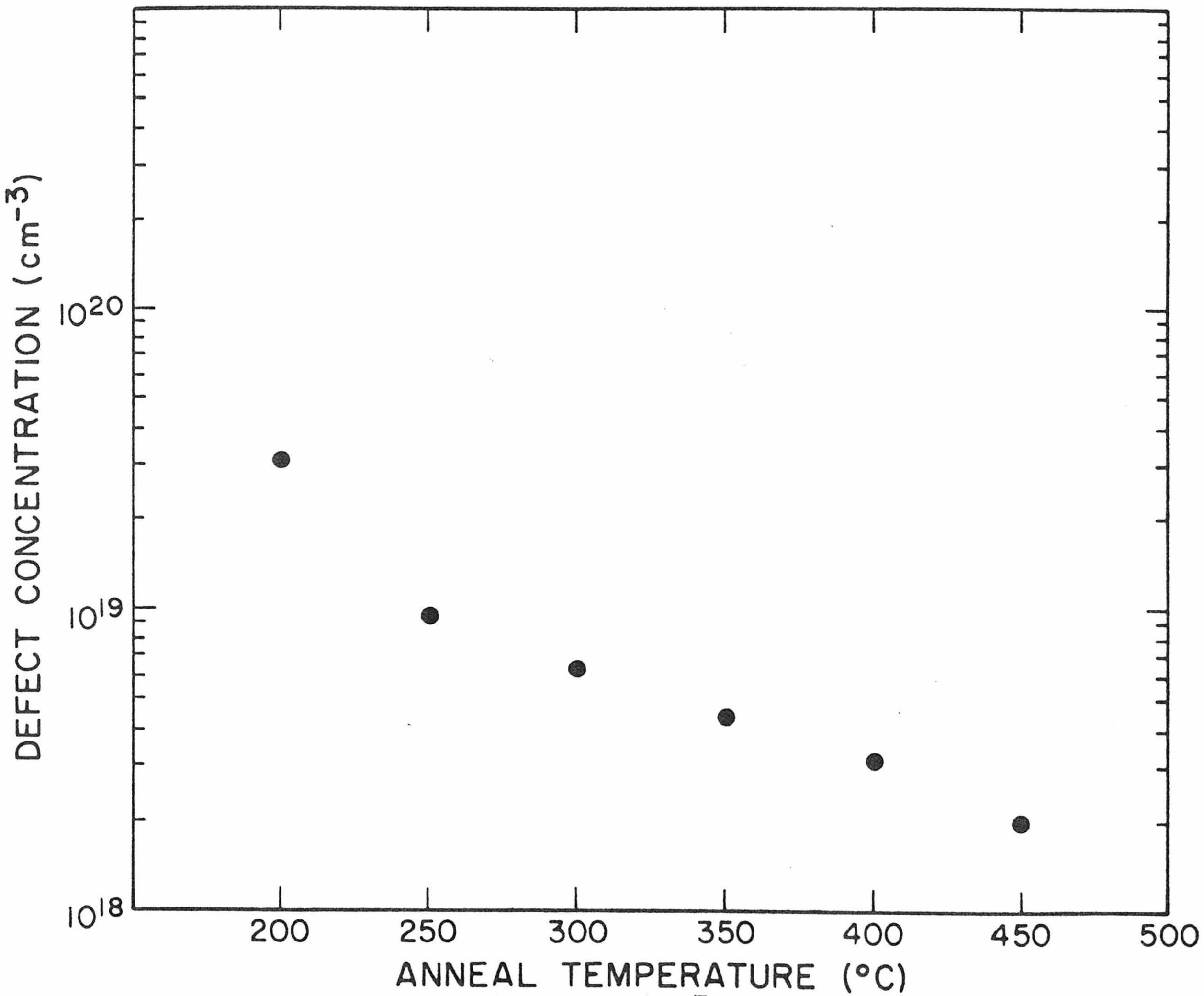


Figure 4-15. Defect concentration versus anneal temperature for a "C" dose implant. (Represents average defect concentration in leading, nearly uniform portion of defect distribution curve).

from this information, however, a major discrepancy arose: the model did not predict the second order mode confinement which was demonstrated by samples 1A, 1B, 1C, 2A, and 2B.

An examination of the experimental data revealed that the second peak of the second order modes fell outside the 3μ depth which had been assumed to represent the maximum proton penetration. In order to explain this behavior, an ancillary guiding layer lying adjacent to the region of primary damage was proposed. (The origin of this layer is thought to be a combination of channeling (See Section 1-IIC) and defect migration effects [60].) Characterization of this new waveguiding model required the determination of two additional parameters: the defect concentration in and the spatial extent of the ancillary layer. It was found that parameters such as these were most easily inferred by analyzing the mode profile of a specimen undergoing second order mode cutoff. Sample 2B:300°C met this requirement and was, therefore, chosen for analysis.

The four media waveguiding structure proposed above has already been discussed in Section 2-II. Consequently, Eq. (2-46b) which relates the dielectric constant to the free carrier concentration together with EQs. (2-11) through (2-14) which describe T.E. mode propagation in the four media waveguide (See Fig. 4-16) were solved by trial and error using a computer program until a mode profile was obtained which matched, as closely as possible, the second order mode profile of sample 2B:300°C. An initial reduction in the number of unknown parameters (the defect concentration in the primary damage layer, the defect concentration in the ancillary layer, and the

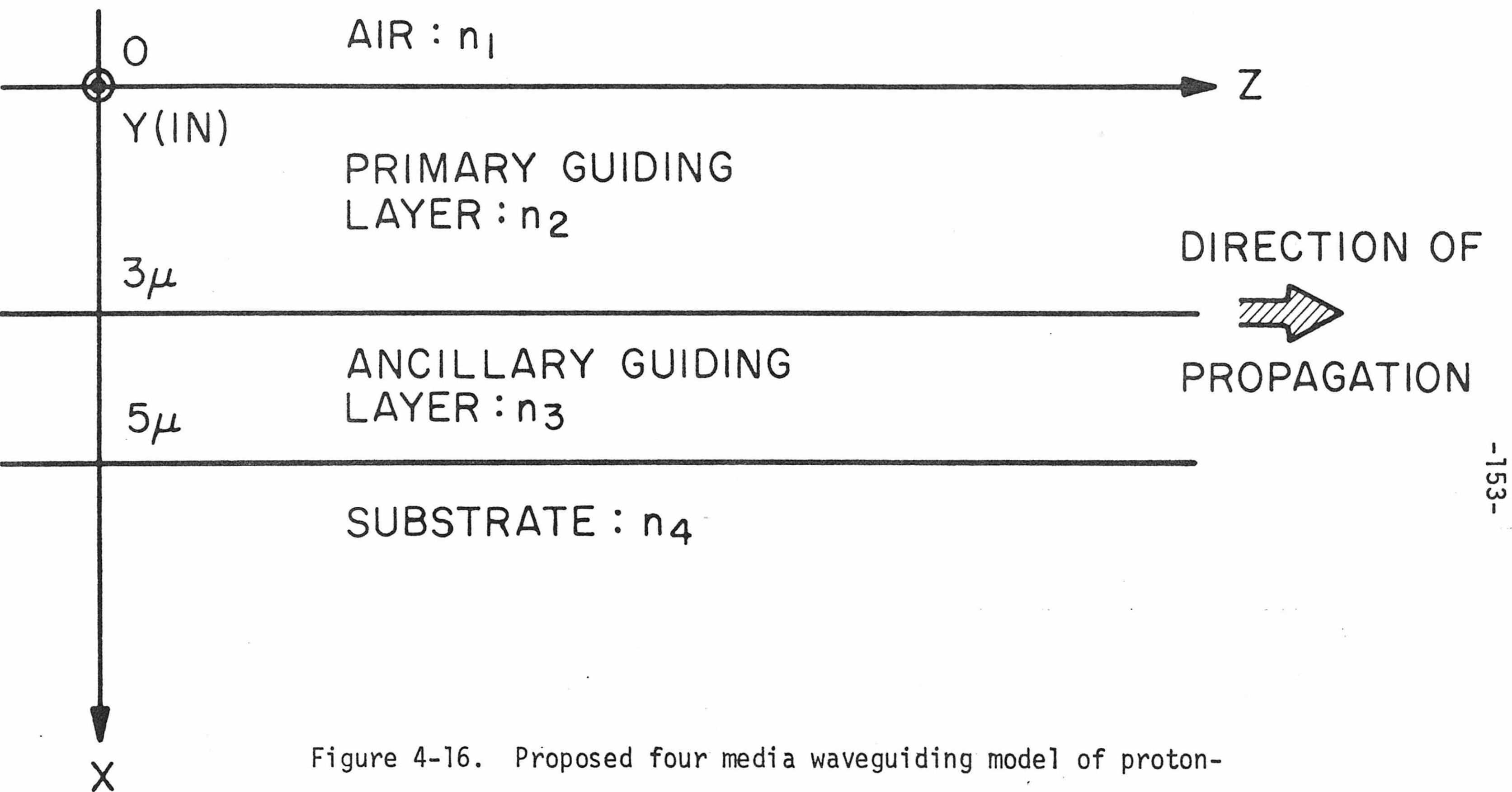


Figure 4-16. Proposed four media waveguiding model of proton-implanted gallium arsenide.

width of the ancillary layer) was accomplished by assuming that the amount of damage generated in the ancillary layer was small in comparison with the damage which resided in the primary (3μ wide) guiding region. The contribution by the ancillary region to the overall attenuation coefficient was accordingly assumed to be inconsequential and, moreover, was assumed to have disappeared completely after 450°C or 500°C annealing. Using these assumptions, it was therefore possible to retain, temporarily, the defect concentration versus anneal temperature plots obtained earlier for the case of three media waveguiding. Accordingly, Figs. 4-8 and 4-15 (properly scaled) yielded for the defect concentration in the primary guiding layer of sample 2B:300°C a value of 7×10^{18} defects cm^{-3}. Finally, the best-fit, theoretical second order mode profile for sample 2B:300°C (shown together with the experimentally observed profile in Fig. 4-17) was obtained by taking for the ancillary layer a defect concentration of 2.8×10^{18} defects cm^{-3} and a width of 2μ. The best fit between experiment and theory was determined by the satisfaction of the following criteria: positional coincidence of mode extrema, equality of the ratios of lobe maxima, and similarity of substrate penetration tails.

The apparent falsehood of the assumption regarding the relative inconsequentiality of the ancillary layer with respect to the primary damage layer (The ratio of defect concentration in the ancillary layer to defect concentration in the primary damage layer was found, above, to be $0.41 = 2.8 \times 10^{18}/7 \times 10^{18}$ for sample 2b:300°C.) will be discussed below. We temporarily circumvent the

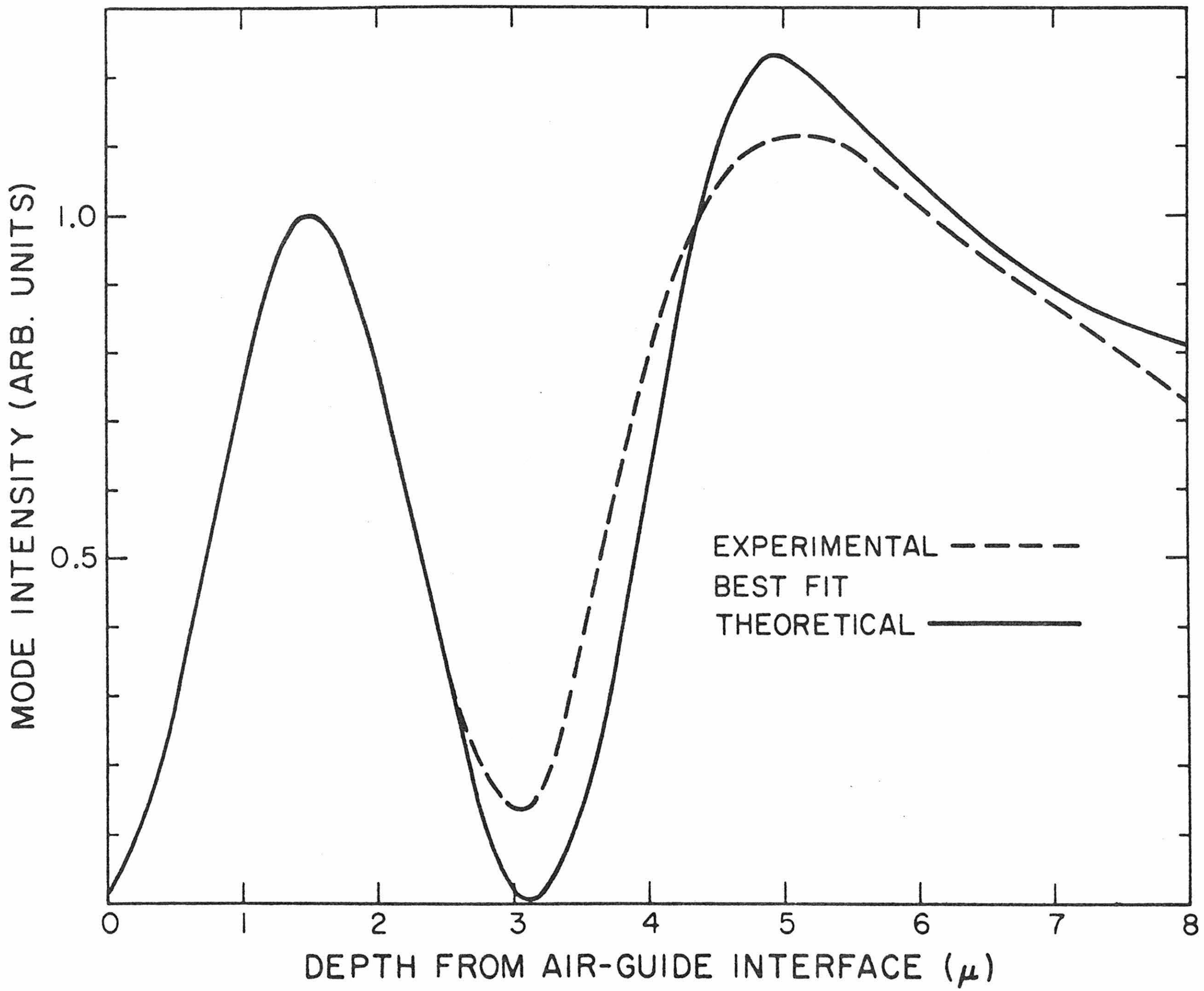


Figure 4-17. Experimental and best-fit theoretical mode intensity profiles for specimen 2B:300°C.

issue and proceed with a description of the final theoretical mode profiles. To obtain them, the assumption was made that the ratio of defect concentration in the ancillary layer to defect concentration in the primary damage layer remained 0.4 throughout the annealing temperature range. Figure 4-8, scaled as before for each of the five substrates, was then used to find the free carrier concentrations in the ancillary waveguiding regions of each waveguiding specimen after each anneal cycle. These free carrier concentrations together with the free carrier concentrations determined earlier for the case of three media waveguiding were used in Eqs. (2-11) - (2-14) to obtain the theoretical mode profiles shown in Fig. 4-11. A discussion of these mode profiles will follow in Section VI.

With the above results in hand, a re-examination of our earlier premise regarding the magnitude of the ancillary layer is in order. The assumption that the defect concentration in this layer was relatively small was made so that the defect and free carrier concentrations obtained for the case of three media waveguiding could be retained. In particular, it was hoped that the first order mode profile shapes and cutoff conditions would not be perturbed sufficiently by the added layer to affect the determination of the proportionality constant which related the attenuation coefficient to the defect concentration. An inspection of Fig. 4-11 reveals that the first order mode profiles did approximately retain their shapes within the 5μ waveguiding region over the temperature range investigated. Furthermore, after the 450°C anneal cycle, the free

carrier concentrations in the ancillary layers of all the substrates examined had very nearly returned to their original, pre-irradiation values. Hence, the proportionality constant relating the attenuation coefficient to the defect concentration, which was obtained by assuming only a 3μ wide guide for samples 1C and 5C, remained meaningful. It was to be concluded, therefore, that the procedure used to obtain the theoretical mode profiles shown in Fig. 4-11 (which involved using the defect concentration versus anneal temperature curve of Fig. 4-15) was still valid.

To complete the analysis, semi-theoretically derived attenuation coefficients were compared with those observed experimentally. Using Fig. 4-15 together with Eq. (4-11), the defect-associated attenuation coefficient was determined for a "C" dose guide between 200^{0}C and 500^{0}C. Recalling that the defect concentration was proportional to the proton dose, the defect-associated attenuation coefficients for doses A, B, and D throughout this temperature range were computed by multiplying the C dose coefficients by the appropriate scaling factors. The free carrier absorption contribution for each guide was then obtained by using the four-media analog of Eq. (4-8):

$$\alpha_{f.c.}(cm^{-1}) = \frac{2.2 \times 10^{-18}}{\int_{-\infty}^{\infty} |E_y^{(1)}(x)^2| dx} \left[N_{g1} \int_{-0}^{3} |E_y^{(1)}(x)|^2 dx + N_{g2} \int_{-3}^{-5} |E_y^{(1)}(x)|^2 dx + N_s \int_{-5}^{-\infty} |E_y^{(1)}(x)|^2 dx \right] \quad (4\text{-}12)$$

where $E_y^{(1)}(x)$ is given by Eqs. (2-10) with $m = 1$ (first order) and N_{g1} and N_{g2} are the free carrier concentrations of the primary and ancillary layers, respectively. Adding together the defect and free carrier-associated contributions yielded the attenuation coefficients shown as open circles in Fig. 4-12. Experimentally observed attenuation coefficients have been represented in Fig. 4-12 by heavy dots with bars joining the lowest and highest coefficients measured for each temperature.

VI. Discussion

The agreement between theoretically and experimentally obtained mode profiles, as demonstrated by Fig. 4-11, is good and lends substantial credence to the supposition that the plasma depression effect gives rise to optical waveguiding in proton-implanted gallium arsenide. Several important observations regarding this agreement may be made: Of particular significance is the fact that cutoff of the second order mode of specimen 1B, noted above to be a sensitive indicator of guide parameters, occurred at exactly that temperature, 400°C, predicted by the model; it will be recalled that cutoff of the second order mode of specimen 2B, the only other specimen which demonstrated second order mode cutoff, was used to fix the ancillary layer parameters. In addition to this cutoff behavior, the mode profiles of samples 1C and 5C offer persuasive evidence that the plasma depression effect is the sole mechanism responsible for refractive index change since, for these samples, the only parameter which was varied was the substrate free carier concentration. Of further note is the good coincidence of mode extrema

obtained for the second order modes of samples 1A and 1B. Here it should be realized that, owing to the method by which radiation was coupled to the waveguides (see Chapter 6), both first and second order modes were, in general, excited simultaneously, the degree of admixture being a function of the exact position and size of the "edge-coupled" input beam. (The hybrid character of the "second order" modes excited in this fashion is betrayed by the lack of nodal behavior.) The fact that the extrema of the theoretical and experimental second order mode profiles agreed as well as they did was due, in part, to the fact that the first order modes were greatly attenuated by the damage peak which occurred near their maxima. These points are illustrated by Fig. 4-11b which shows the first and second order mode profiles of specimen 1B:300°C.

The agreement between the semi-phenomenologically obtained attenuation coefficients and those measured experimentally was good in general but broke down for several important cases. The disagreement arose for specimens given large proton doses and annealed past 400°C. As Fig. 4-12 indicates, samples 1A, 2A, 3A, 4A, and 5A, when annealed past 400°C, showed the most disagreement, with some disagreement being shown by samples 4B and 5B above 450°C and 500°C, respectively. The importance of understanding these discrepancies is emphasized by the fact that two of the samples mentioned, 3A:500°C and 4A:500°C, appeared to suffer only minimal, free carrier-associated losses and, hence, offered promise as good passive waveguiding structures.

A possible explanation for the disagreement mentioned above has already been alluded to in Section V. It will be recalled that in that section the validity of one of the fundamental assertions upon which the entire analysis has been based was left unresolved. This assertion, assumed until now to be true, stated that the absorption spectrum annealed uniformly. It is the present contention that for certain dose-anneal combinations, those defects responsible for the shift of the fundamental absorption edge toward lower energies (identified in Section 3-IIIA2 as gallium vacancy-donor complexes) remain after high temperature annealing while those defects responsible for the long wavelength absorption tail disappear.

The anomaly of well-confined, low-loss modes (such as demonstrated by samples 3A:500°C and 4A:500) may therefore be explained by the presence of relatively stable gallium vacancy-donor complexes which, because of their large ionization energy, serve to compensate free carriers without contributing appreciably to the attenuation at the wavelength used (1.15μ).

VII. Conclusion

A model has been developed in the last two chapters which describes the mechanisms that give rise to and affect the waveguiding of near infrared radiation in proton-implanted gallium arsenide. Optical attenuation and mode structure measurements have been made on both one-dimensional (planar) and two-dimensional (channel) waveguides [14] fabricated in this manner. Optical waveguides such as these, envisioned as eventually forming the basis of monolithic integrated circuits, may be categorized according to their eventual use as follows: passive

guides to be used for the simple transmission of information between integrated optical devices and active guides to be used as information processing elements. To be classed as an acceptable passive waveguide only one criterion, that of low loss, must be satisfied. The present effort has demonstrated that such waveguides may be readily and reproducibly made using the technique of proton implantation: losses measured, as low as 4.3 db cm^{-1}, are comparable to those encountered in the best available GaAs epitaxial waveguides of similar configuration [61]. A GaAs waveguiding structure which is to double as an active device, on the other hand, must be capable of supporting an electric field over as large a fraction of its optical mode as possible in order to take efficient advantage of the operating speeds which result from electrooptic interactions; optical attenuation may or may not be a factor, depending on the type of active element under consideration. While no proton-implanted waveguides suitable for active applications have so far been obtained, as will be pointed out below, the experimental data and theoretical conclusions which have resulted from the present effort strongly suggest their possibility.

Having demonstrated the feasibility of producing acceptable passive waveguides, future efforts should therefore be directed toward the fabrication of proton-implanted gallium arsenide waveguides suitable for active applications. As mentioned above, in order to optimize their efficiency, electrooptic devices such as phase modulators [17] and mode selectors [62] require the maximum possible spatial overlap of the applied electric field and the guided mode intensity profile [17]. A typical electrooptic device configuration is shown in Fig. 4-18 wherein the electric field is generated by a

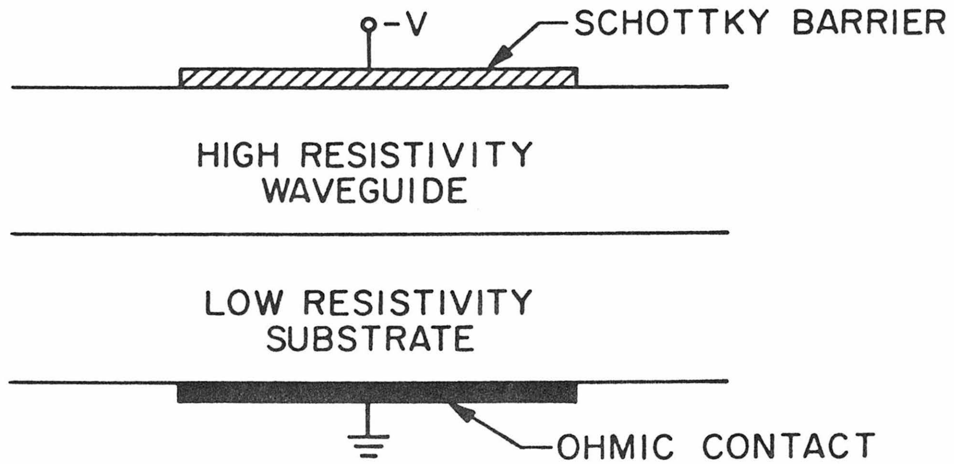


Figure 4-18. Typical electrooptic device configuration.

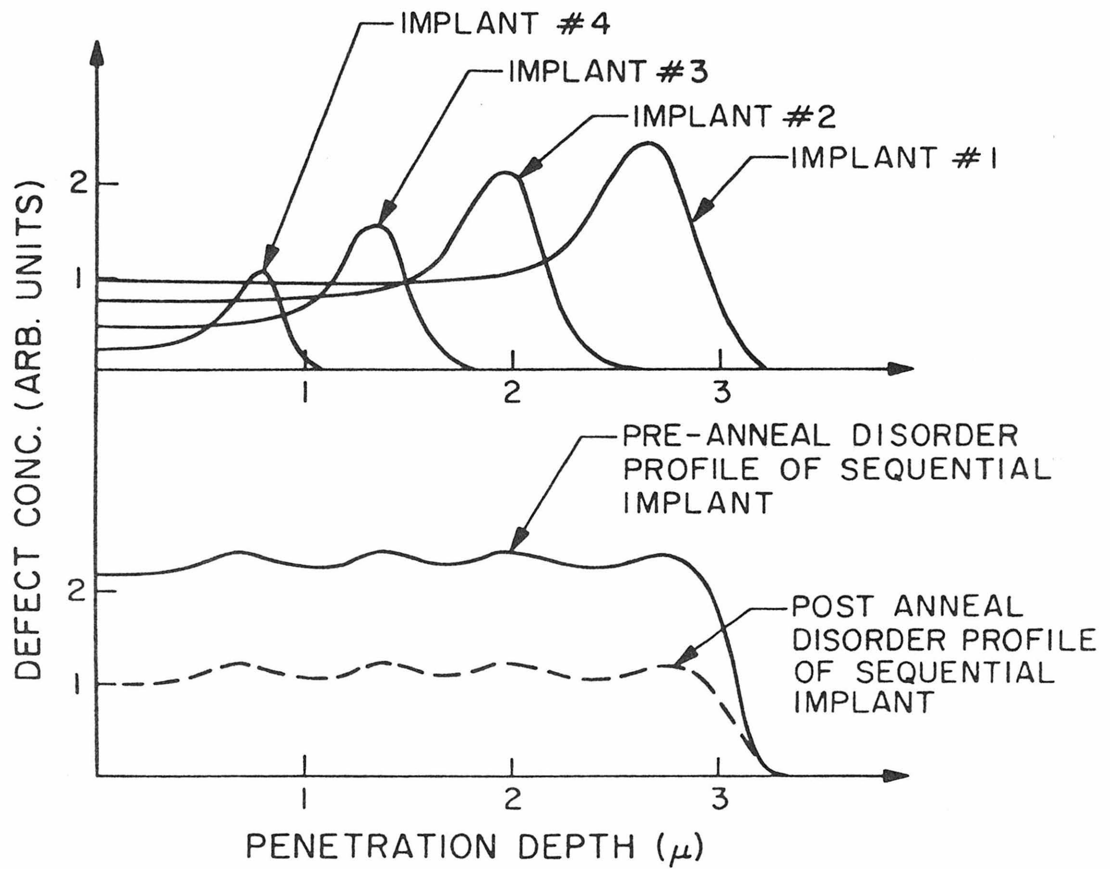


Figure 4-19. Schematic of proposed sequential implant. The desired end result is shown dotted.

reverse-biased Schottky barrier. The maximum penetration depth of the field of the reverse-biased Schottky barrier is given, in terms of the material's breakdown field and free carrier concentration, by

$$d_{max} = \frac{2\kappa\varepsilon_o E_{B.D.}}{Ne} \qquad (4\text{-}13)$$

where κ = relative dielectric constant, ε_o = permittivity of free space, e = electron charge, N = free carrier concentration and $E_{B.D.}$ = the material's breakdown field. Taking $\kappa = 11.5$ and $E_{B.D.} = 4 \times 10^5 V\cdot cm^{-1}$ for gallium arsenide, Eq. (4-13) becomes:

$$d_{max}(\mu) = \frac{8.15 \times 10^{16}}{N(cm^{-3})} \quad . \qquad (4\text{-}14)$$

In order to cause an electric field to appear across the entire width of a 3μ wide implanted gallium arsenide waveguide, therefore, requires the residual free carrier concentration in the waveguiding layer to be less than or equal to $2.7 \times 10^{16} cm^{-3}$. By comparison, the residual surface free carrier concentration of sample 5A:500°C, which was the least lossy waveguide measured ($\sim 1\ cm^{-1}$) and, hence, one of the most promising for active applications, was determined by capacitance-voltage measurements to be $2.04 \times 10^{17} cm^{-3}$. From Eq. (4-13), the maximum attainable field penetration in this case was only 0.4μ - far from the optimal value of 3μ .

In order to produce an implanted waveguide which has both low loss and a low residual free carrier concentration in the waveguiding layer, a sequential implantation, outlined schematically in Fig. 4-19,

is proposed. Recalling Fig. 4-6, the damage peak of the proton-generated damage profile was seen to increase the attenuation coefficient by approximately a factor of 1.5 over that value which would have obtained in its absence. By using the technique of sequential implantation and selective annealing, however, in order to produce a uniform damage profile equal in magnitude to one and one-half times that of the residual surface layer of a single implant profile, one should obtain, in principle, a waveguide with the same low attenuation coefficient as that of the annealed, single implant guide but with roughly 1.5 times the residual compensating defect center concentration. The resulting higher defect concentration would, in turn, lead to a lower residual free carrier concentration in the waveguiding layer and deeper penetration by externally applied electric fields.

The behavior, attributed above to stable, gallium vacancy-donor complexes, might also, if properly exploited, lead to low loss, well-compensated waveguides. The question of whether the shift of the fundamental edge toward lower energies behaves as suggested could be answered by monitoring the behavior of the entire damage-induced absorption spectrum under isochronal annealing conditions.

CHAPTER 5

A PROTON-IMPLANTED INTEGRATED OPTICAL DETECTOR

I. Introduction

Integrated optical circuits may be thought of as the optical analogs of integrated electronic circuits. As such, the components of an integrated optical circuit are envisioned as being fabricated on a common substrate, the whole constituting a completely self-contained information processing unit. With regard to the integration of individual device elements, two approaches are presently under consideration: The first is the hybrid approach which would combined the best available components fabricated from different material systems. The second approach advocates the use of a single material system from which all device elements are fabricated. The latter, or monolithic, approach offers the obvious advantages of greater compactness and ease of fabrication, as all elements may be constructed using compatible processes. As many of the necessary optical components (e.g., injection lasers and modulators) have been fabricated using gallium arsenide, this mateiral currently represents the most promising candidate for monolithic integration. Accordingly, described below is an integrated optical detector which was made by proton-implanting an epitaxial gallium arsenide waveguide.

II. Device Construction and Geometry

Figure 5-1 depicts the device geometry. The optical waveguiding structure consisted of a 3.5μ thick n-type (S-doped, $n = 2.8 \times 10^{16} cm^{-3}$) epitaxial film grown on a degenerate n-type substrate ($n = 1.25 \times 10^{18} cm^{-3}$). Good optical confinement for the guide thickness used resulted from the guide-substrate refractive index discontinuity generated by the plasma depression effect. Prior to implantation, optical attenuation at 1.15μ was measured to be (see Section 6-III) $1.3 \ cm^{-1}$ and could be accounted for by consideration of free carrier substrate penetration losses.

The waveguide was implanted with 300 keV protons, the total integrated flux of which was $2 \times 10^{15} cm^{-2}$. The sample was then annealed at 500°C for 30 minutes in order to allow some optical transmission through the damaged waveguide; residual, defect-associated losses were measured to be $\sim 15 \ cm^{-1}$ based on a comparison of the optical attenuation before and after implantation and annealing. The anomolously high defect-associated loss remains unexplained; losses measured in proton-implanted waveguides whose substrate free carrier concentrations and heat treatment histories were similar (ref: samples 3A:500°C and 4A:500°C) were only $1.3 - 1.8 \ cm^{-1}$.

Subsequent to the heat treatment, an ohmic contact was formed on the substrate side of the sample by vacuum depositing a thin layer of Au-Ge and then alloying in a flowing hydrogen atmosphere for five minutes at 250°C (see Section 6-II). 11 mil square Al Schottky barriers were then evaporated in a waffle pattern over the implanted waveguide as shown in Fig. 5-2.

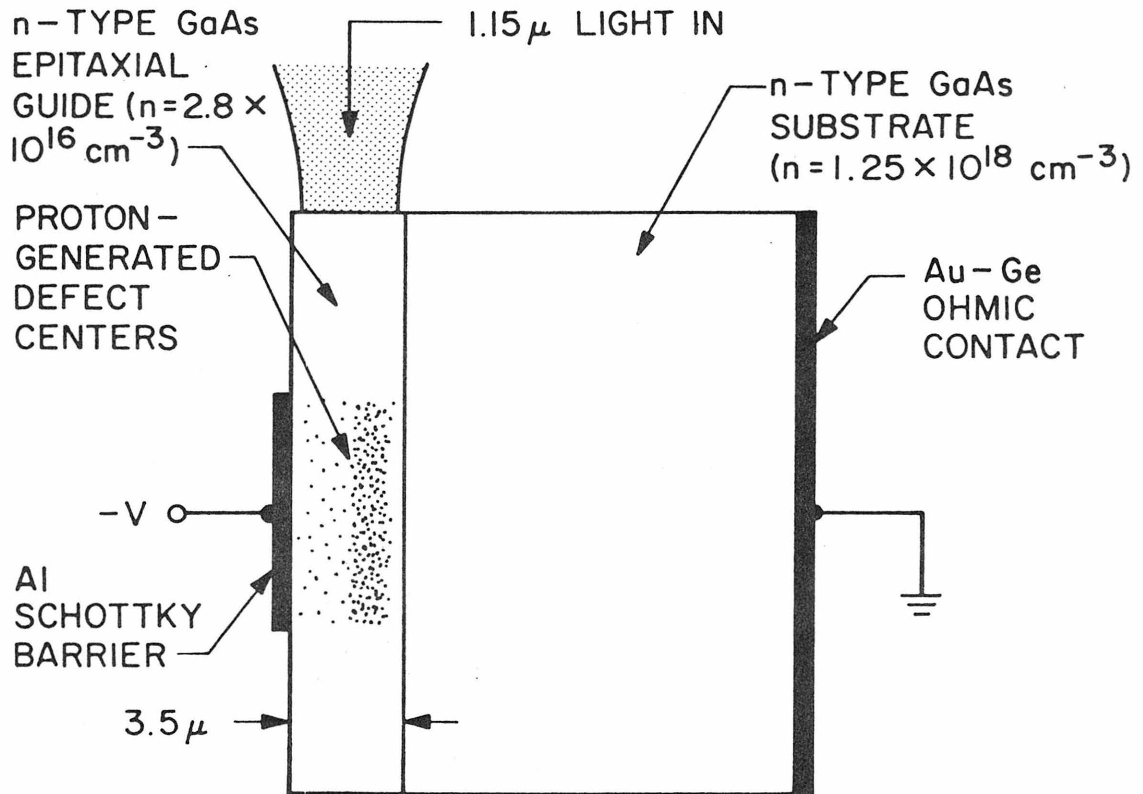


Figure 5-1. Proton-Implanted integrated optical detector-device geometry.

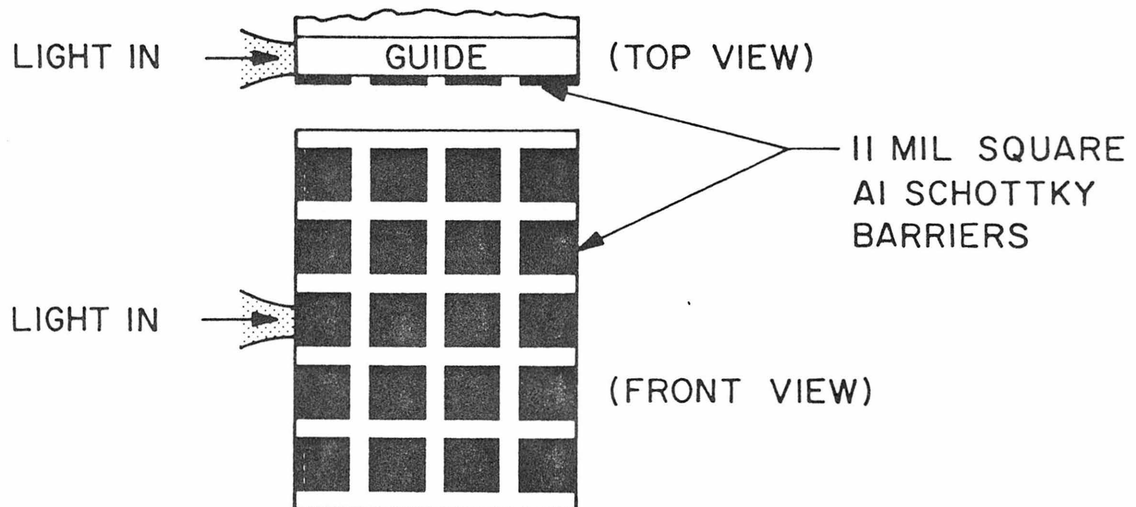


Figure 5-2. Aluminum Schottky barrier pattern.

III. Device Operation

The principle of operation of the detector is similar to that of conventional p-n or p-i-n junction photodetectors Upon the application of a reverse bias to the Schottky barrier, a depletion layer is produced which, given sufficient reverse bias, extends across the high resistivity waveguiding layer to the lower resistivity substrate. Any dipole transitions made possible by radiation-produced defect levels then generate free carriers which are swept out of the depletion layer, thereby causing current to flow through an external circuit. The situation is depicted schematically in Fig. 5-3. By insuring that the radiation-induced damage extends over most of the waveguiding layer and by choosing epitaxial material of high purity (in order to allow the widest possible depletion region), maximum detector efficiency can be obtained over a given interaction length.

IV. Device Performance

The experimental arrangement used to measure device performance is shown in Fig. 5-4. Radiation was end-coupled to one end of a cleaved specimen and examined for waveguiding at the opposing, cleaved output face. The technique used to monitor the presence of waveguiding is described in Section 6-III. After waveguiding had been established, a 3 mil diameter gold whisker was mechanically contacted to one of the 11 mil square Al Schottky barriers. A transistor curve tracer was then used to determine the detector's I-V characteristics under varying illumination conditions. Figure 5-5 shows a series of the reverse bias characteristics obtained; a

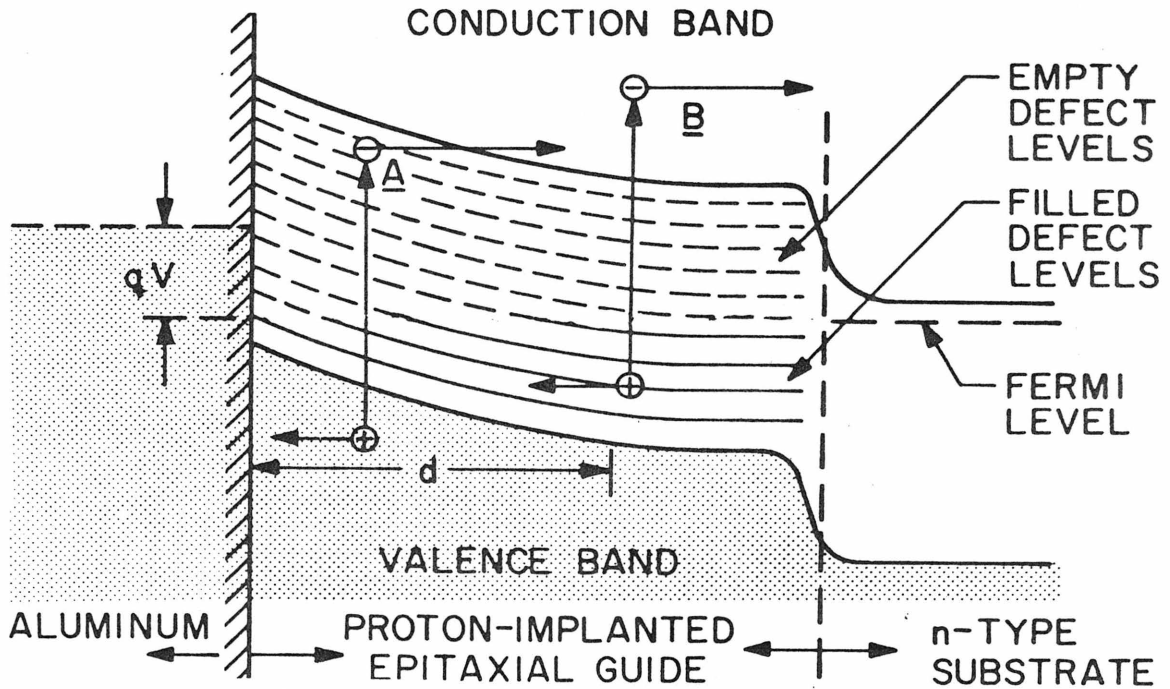


Figure 5-3. Principle of operation of integrated optical detector: upon the application of a reverse bias (V) electrons liberated by transitions labeled "A" and "B" are swept out of depletion layer d resulting in the flow of current.

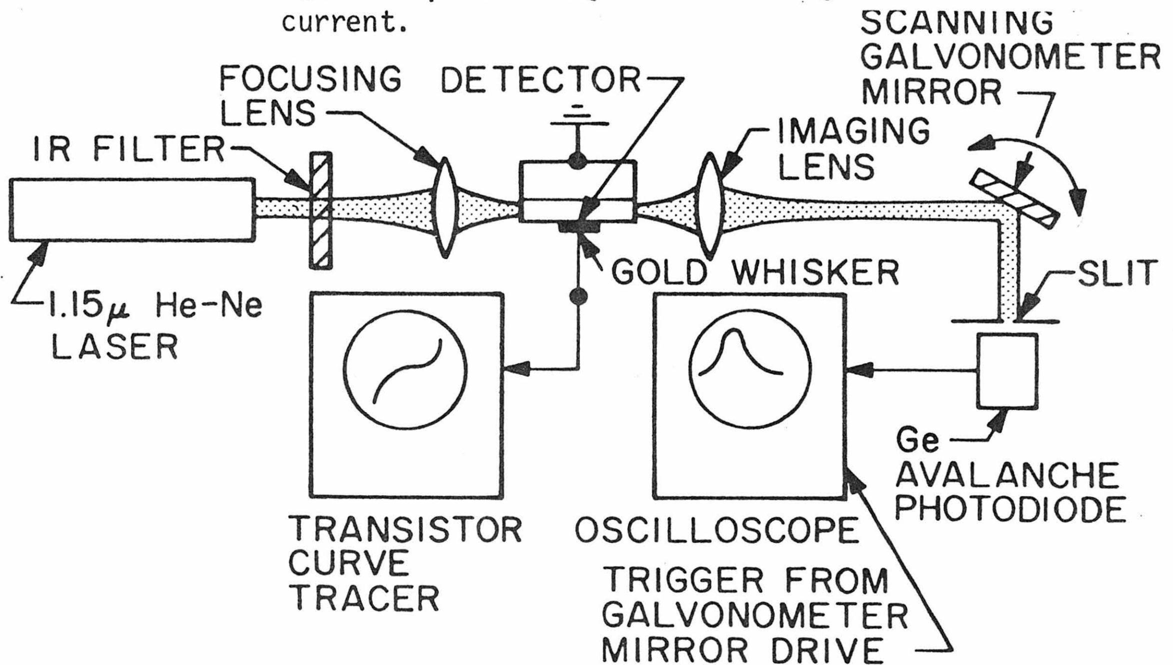


Figure 5-4. Experimental setup used to measure device performance. Schottky barrier I-V characteristics were monitored under both dark and illumination conditions using a transistor curve tracer. Waveguiding in the epitaxial layer was simultaneously monitored using an oscilloscope. (See Fig. 6-3).

Spectra-Physics model 120 He-Ne laser emitting at 1.15μ served as the illumination source.

The particular Schottky barrier used to obtain the traces given in Fig. 5-5 is shown with respect to the entire sample in Fig. 5-6. Taking the output power of the laser to be 0.75 mw and the overall waveguide insertion loss to be 50% (including both reflection at the input face, 30%, and waveguide mode coupling loss, 20%) and assuming that the entire guided beam passed under the Schottky barrier in question, the number of photons per second that were lost while traversing the detector is given by:

$$P = \frac{(0.75 \times 10^{-3})(0.5)\exp(-15 \times 0.0025 \times 8)[1 - \exp(-15 \times 0.0025 \times 11)]}{(1.08)(1.602 \times 10^{-19})}$$

$$= 5.42 \times 10^{14} \text{ photons/sec} \qquad (5\text{-}1)$$

where it has been recalled that the defect-associated attenuation coefficient was 15 cm^{-1} and noted that the Schottky barrier examined was 8 mils from the waveguide input face. The detector current generated under full illumination (Fig. 5-5a) was approximately 1.5×10^{-5} amps just before reverse breakdown. Hence, the number of electrons swept out of the depletion layer per second was:

$$e = \frac{1.5 \times 10^{-5}}{1.602 \times 10^{-19}} = 0.935 \times 10^{14} \text{ electrons/sec} . \qquad (5\text{-}2)$$

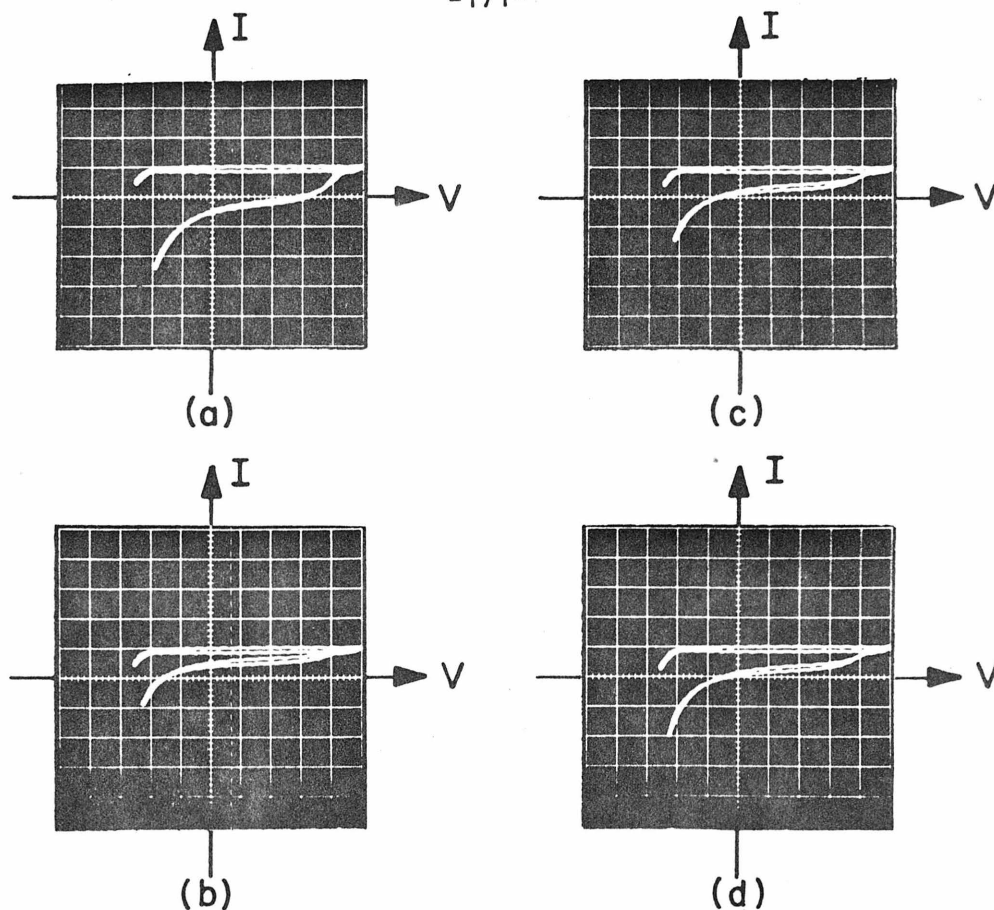


Figure 5-5. Reverse bias characteristics of detector under dark and illumination conditions: (a) (b) (c).

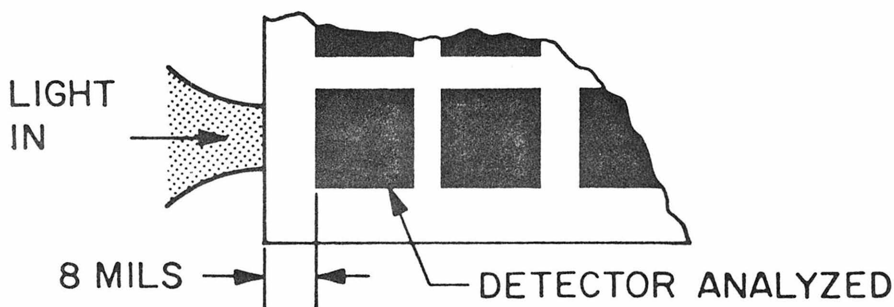


Figure 5-6. Position with respect to input beam of detector used to obtain data shown in Figs. 5-5 and 5-8.

The internal quantum efficiency of the detector was, therefore:

$$(e/p) \times 100 = 17\% \ . \qquad (5\text{-}3)$$

Two factors probably contributed to the relatively low quantum efficiency. First, Fig. (3-1a) indicates that only about 40% of the total defect-associated attenuation at 1.15μ (1.08 eV) results in the promotion of trapped electrons to the conduction band- the remaining 60% of the attenuation being attributed to dissipative scattering and absorption by thermal spikes. Second, the residual free carrier concentration in the implanted waveguided, as determined by the capacitance-voltage technique (see Section 6-IV), was found to be $\approx 2.7 \times 10^{16} cm^{-3}$, or roughly the original pre-implantation value. As a result, the width of the depletion layer (for a reverse bias of 30V) was, according to Eq. (4-14), only 2.9μ. Figure 4-6 indicates that, under these circumstances, approximately 90% of the electrons liberated from trapping levels are exposed to the influence of the electric field generated within the depletion region. Combining the two effects just discussed, we find that the limiting quantum efficiency at 1.15μ is given by

$$(0.4 \times 0.9) \times 100 = 36\% \ . \qquad (5\text{-}4)$$

Finally, an attempt was made to measure the response time of the implanted detector by using a Q-switched Nd:YAG laser; the 1.06μ laser pulses were end-coupled to the detector assembly in the manner described above. A schematic of the electronic circuitry used

is shown in Fig. 5-7; a series of the pulses observed using different load resistors is given in Fig. 5-8. As may be seen in Fig. 5-8, pulse rise times reached a minimum for load resistors less than or equal to 800Ω, indicating that the external electronics were no longer limiting the detector response. In order to determine whether the minimum 200ns rise time represented a limit of the implant detector or was, in fact, the rise time of the Nd:YAG pulses, the pulses were measured using a commercial Ge p-n detector. Figure 5-9 shows a representative pulse response obtained using the Ge detector. Since the Ge detector response was in the GHz range, it was concluded that the rise times measured for both the implant and the Ge detectors, 200 ns, were a function of the Q-switched laser and did not represent limits of the implant detector.

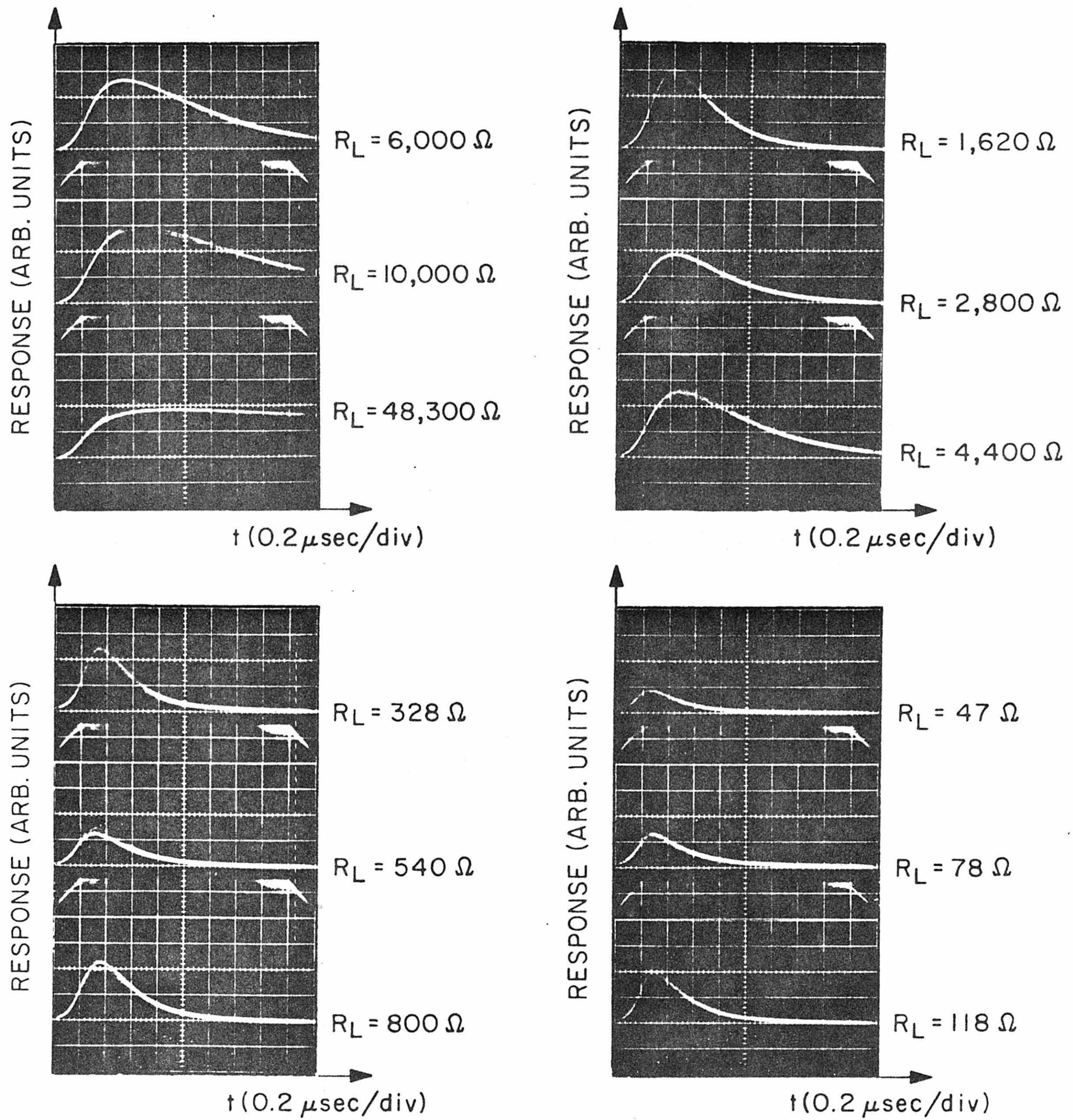


Figure 5-8. Pulse response of integrated optical detector as a function of load resistance.

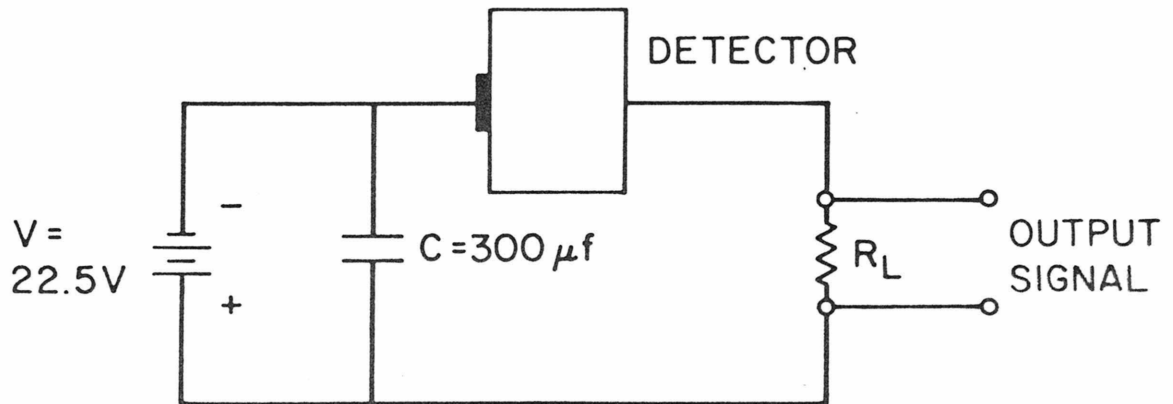


Figure 5-7. Circuit used to measure pulse response of integrated optical detector.

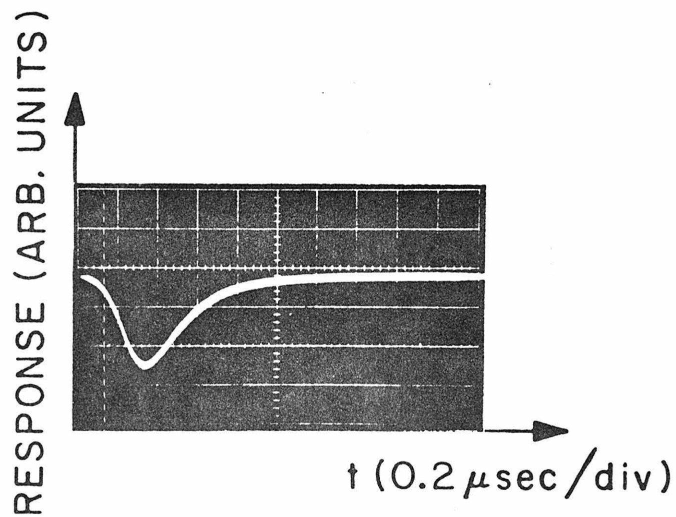


Figure 5-9. Response of commercial Ge avalanche diode to Nd:YAG laser pulse.

CHAPTER 6

EXPERIMENTAL TECHNIQUES

I. Proton Implantation

All implantations were performed using a modified Accelerators, Inc. implantation machine [63]. The proton source used was H_2S. Implantations were done at room temperature by scanning the focused H^+ beam in a raster pattern over the gallium arsenide targets.

II. Post Implantation Annealing

The apparatus shown in Fig. 6-1 was used to do all annealing. The samples were placed on a molybdenum strip and annealed in a flowing hydrogen atmosphere. Temperatures were monitored by a thermocouple which was brazed to the underside of the molybdenum strip. Since several specimens were usually annealed simultaneously, it was necessary to determine the temperature uniformity over the 1 x 2 cm area of the strip which was used. Three pieces of sample 2B, placed as shown in Fig. 6-2a, were used for this purpose. Temperature uniformity was demonstrated by annealing the three pieces for one-half hour at 300°C and then comparing their waveguide mode profiles. As may be seen in Fig. 6-2b, following this heat treatment all three specimens approached second order mode cutoff. Subsequent annealing at 350°C caused them all to undergo second order mode cutoff, thereby demonstrating the hoped for temperature uniformity.

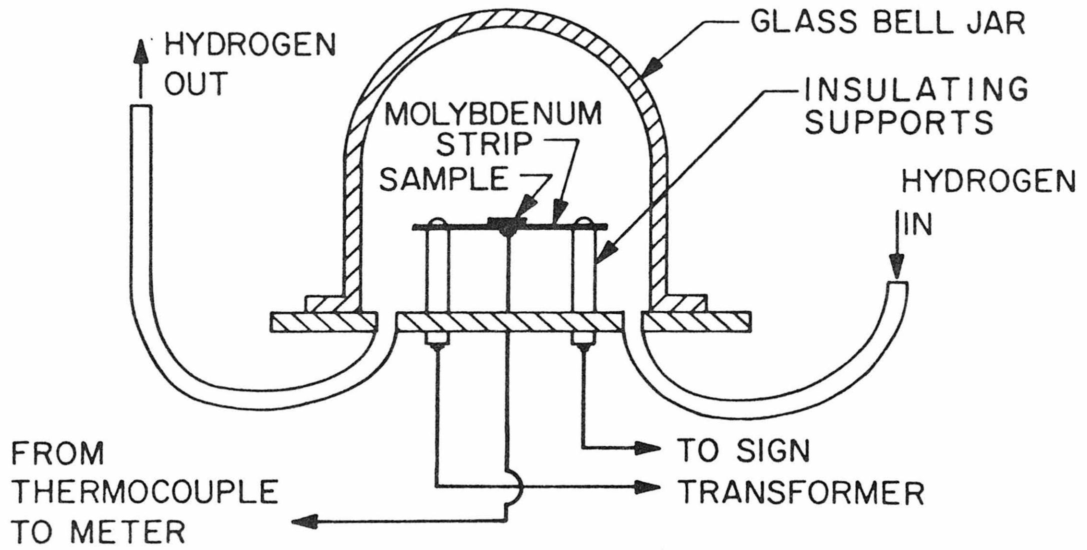
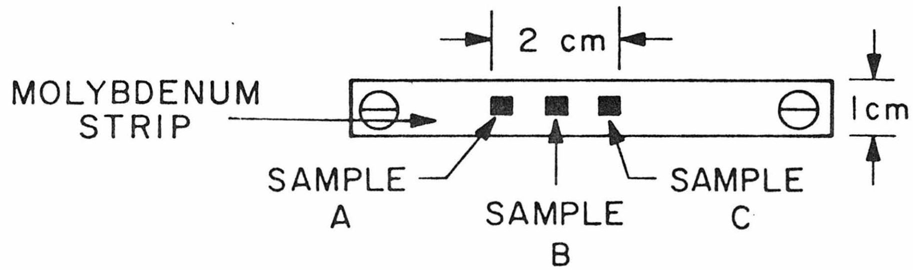


Figure 6-1. Annealing apparatus.



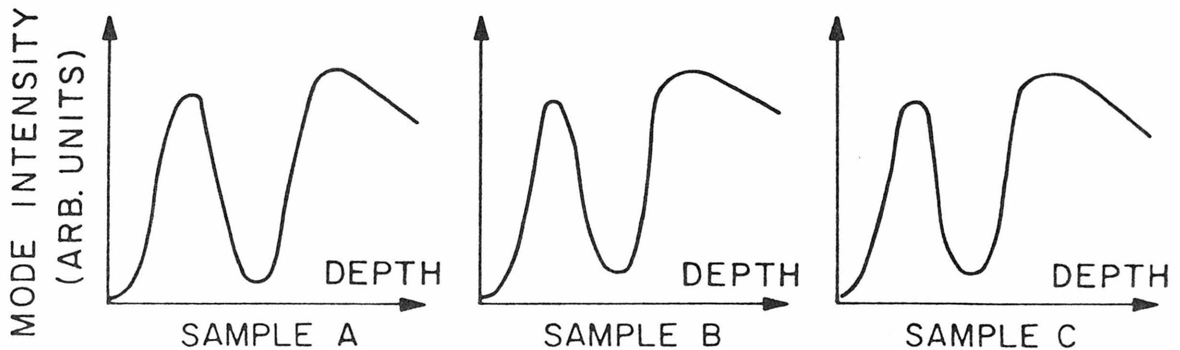


Figure 6-2. (a) Placement of samples A, B, and C of specimen 2B for temperature uniformity test. (b) Mode profiles of samples A, B, and C measured after 0.5 hour anneal at 300°C.

III. Mode Profile and Attenuation Measurements

The experimental setup shown in Fig. 6-3 was used to perform the mode profile and attenuation measurements. The lasing source, a Spectra Physics model 120 He-Ne laser emitting at 1.15μ, was end-coupled to the cleaved facet of a waveguide by means of a 43:1 microscope objective (L1); extraneous emission from the laser was blocked by a filter (F1) placed close to its output mirror. The effective focal length and lens diameter of the microscope objective were 5 mm and 4.7 mm, respectively. The near-field radiation pattern at the output of the cleaved waveguide was imaged, by means of an identical 43:1 microscope objective (L2), on a 30μ wide slit placed over a BNC-mounted, germanium avalanche photodiode (P1) located 0.38 meters from the imaging objective; the magnification factor was, accordingly, 76x. The resolution of the microscope objectives, as determined by the Rayleigh criterion [64], was:

$$d \cong \frac{1.22\ f\lambda}{D} = \frac{1.22(5 \times 10^3)1.15}{4.7 \times 10^3}$$

$$= 1.5\mu \qquad (6.1)$$

Therefore, since the diameter of the Rayleigh disc (Eq. 6-1), imaged onto the detector, was 114μ (= 76 x 1.5μ), the system resolution was determined by Eq. (6-1) as opposed to the width of the detector slit which was 30μ.

Real-time optical mode profiles were obtained by scanning, via galvanometer mirror (M1), the imaged, near field radiation patterns of the waveguided modes over the detector slit. The

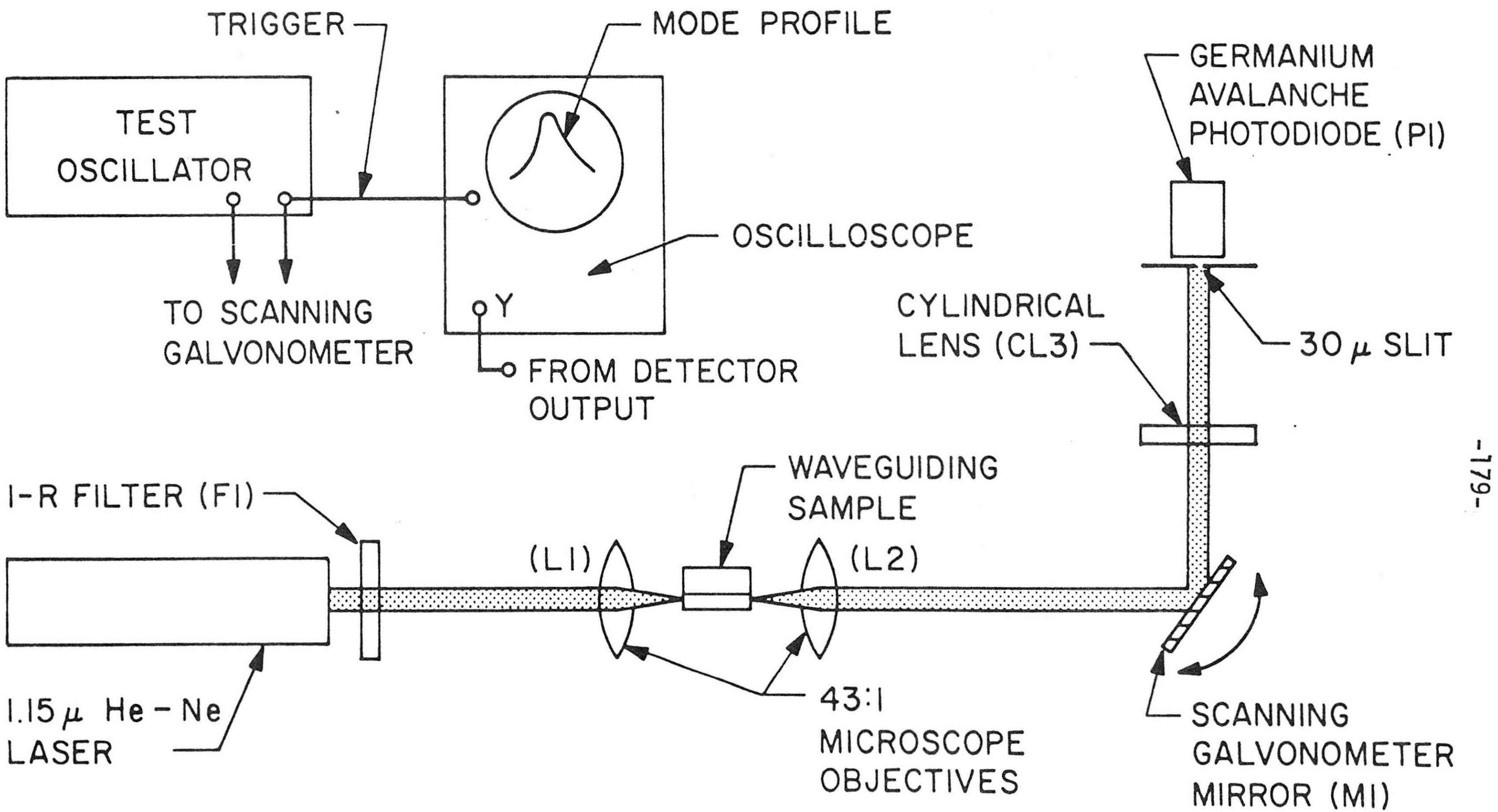


Figure 6-3. Apparatus used to measure waveguide mode profiles and waveguide attenuation coefficients.

galvanometer mirror was driven by a 38 Hz sine wave generated by a conventional test oscillator; the total scanning angle was 6° peak to peak. CRT displays of the waveguide mode profiles were obtained by coupling the detector output to the vertical drive of a Tektronix model 540 oscilloscope; triggering was accomplished by use of the same sine wave which drove the scanning galvanometer mirror. The linearity of the horizonal display scale was checked by comparing mode profiles obtained in the manner described here with those obtained using the sine wave galvanometer mirror signal as the horizontal oscilloscope drive. No differences between profile shapes were noted for mode patterns less than approximately 10μ in width. For reasons of convenience, the technique herein described was therefore retained. Finally, in order to determine the scaling factor which related distance to horizontal sweep speeds, a 1.5μ periodicity wire grid grating was focused via the two 43:1 microscope objectives onto the detector.

The attenuation coefficients of first order guided modes were measured by monitoring the outputs of waveguides of different lengths. With the exception of a cylindrical lens (CL3), the experimental arrangement used was identical to that described above. The cylindrical lens, whose focal length was 30mm, served to focus the near field radiation pattern down to a horizontal line approximately 10 mils wide, thereby allowing the detector, which was 40 mils in diameter, to collect all of the power emergent from a given waveguide. As a result, the total power coupled out of a waveguide could be determined by integrating the CRT-generated mode profiles. The power incident on the input face of each waveguide was measured

before and after the waveguide measurements by using the same procedure.

IV. Free-Carrier Concentration Measurements

The conventional capacitance-voltage technique was used to measure free carrier concentrations. Ohmic contacts were formed by evaporating a germanium-gold film over one side of a specimen and then alloying for five minutes at 250°C in the apparatus shown in Fig. 6-1. Schottky barriers were then made by evaporating gold over the opposing, polished specimen surface. Finally, a gallium arsenide sample prepared in this manner was cleaved into smaller pieces to insure the electrical isolation of the Schottky barriers.

V. Step Etching

The step etching technique mentioned in Chapter 4 is outlined in Fig. 6-4 [65]. A sample to be step-etched was first coated with a thin layer of black Apiezon wax. A narrow strip of the wax was then removed by treatment with trichloroethylene. Subsequent immersion of the entire smaple in a 1% bromine methanol etch resulted in the generation of a shallow channel where the specimen had been unprotected by the wax. By repeating this procedure, six such channels, or steps, were etched into each sample to be examined, the last step being one-sixth as deep as the first (assuming that successive emmersion times were equal). The step depths chosen for the present study were, as determined by interferometric measurements: 0.5,

1.0, 1.5, 2.0, 2.5, and 3.0μ ±(0.15μ).

Following the step-etching, specimens were prepared for Schottky barrier capacitance-voltage measurements as described in Section IV. Instead of a uniform Schottky barrier, however, gold dots 16 mils in diameter were deposited onto the step-etched specimen in order to isolate Schottky barriers which resided on different steps.

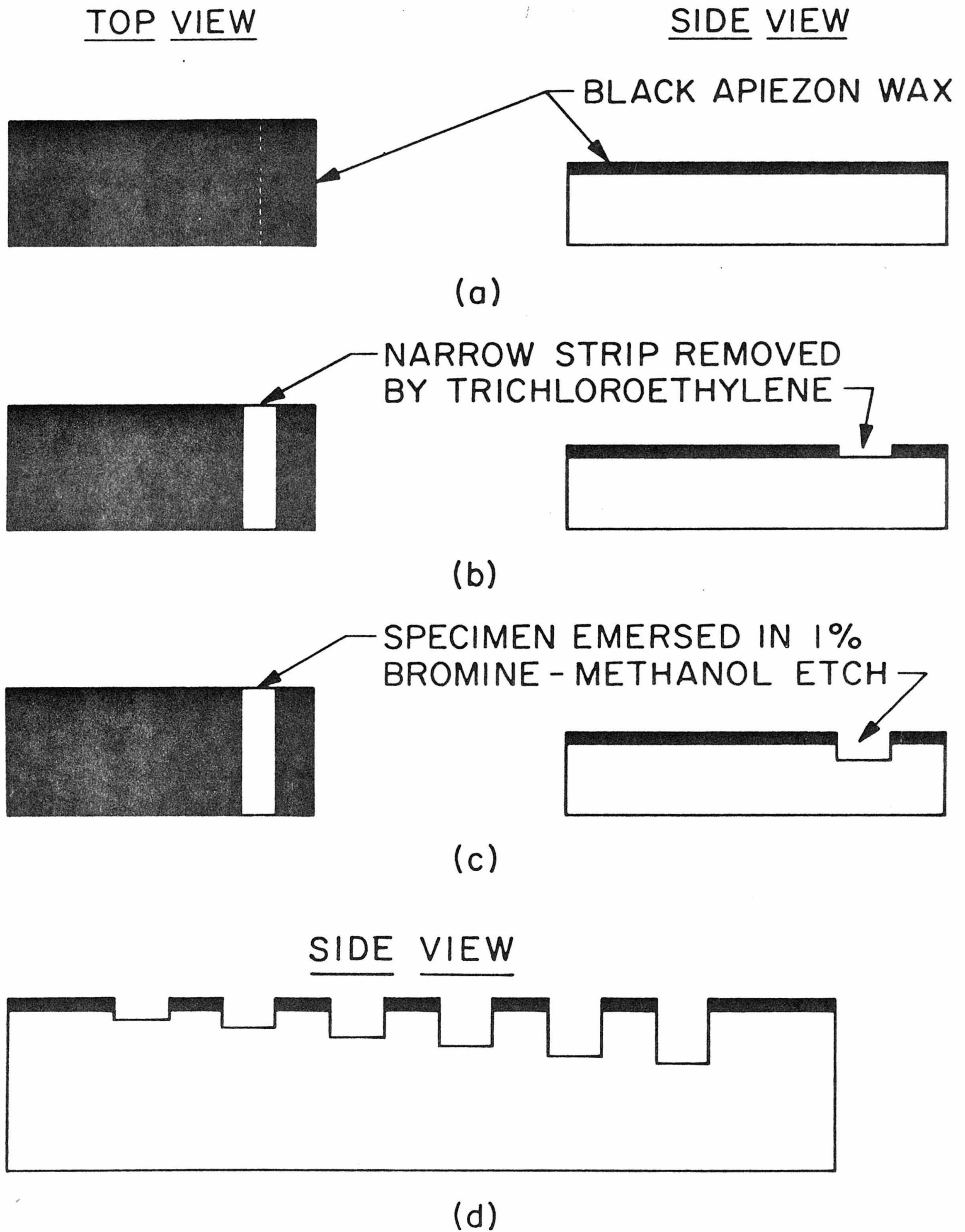


Figure 6-4. Step etching technique. (a) Specimen to be step-etched is first coated with black apiezon wax; (b) Narrow strip of apiezon wax is removed by trichloroethylene; (c) Specimen is immersed in 1% bromine-methanol etch; (d) Result of six consecutive step-etches.

REFERENCES

[1] J. W. Mayer, L. Eriksson, J. A. Davies, Ion Implantation in Semiconductors, Academic Press (1970).

[2] G. H. Kinchin and R. S. Pease, Rept. Prog. Phys., 18, 1 (1955).

[3] W. S. Snyder and J. Neufeld, Phys. Rev., 97, 1636 (1955); 103, 862 (1956).

[4] W. A. Harrison and F. Seitz, Phys. Rev., 98, 1530 (1955).

[5] G. J. Dienes and G. H. Vienyard, Radiation Effects in Solids, Vol. II, Interscience Publishers, Inc. (1957).

[6] N. Bohr, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd., 18, 8 (1948).

[7] J. Lindhard, M. Scharff, H. E. Schiott, Kgl. Danske Videnskab Selskab, Mat.-fys. Medd., 33, 14 (1963).

[8] J. A. Grimshaw and P. C.Banbury, Proc. Phys. Soc. of London, 84, 151 (1964).

[9] W. Köhler and W. Schilling, Nukleonik, 7, 389 (1965).

[10] P. V. Pavlov, D. I. Tetel'baum, E. I. Zorin, V. I. Alekseev, Soviet Physics-Solid State, 8, 2141 (1967).

[11] See P. Sigmund and J. B. Sanders, in Proceedings of the International Conference on the Applications of Ion Beams to Semiconductor Technology, Grenoble, p. 215 (1967).

[12] D. K. Brice, Radiation Effects, 6, 77 (1970).

[13] S. Furukawa, H. Matsumura, H. Ishiwara, Japanese Journal of Applied Physics, 11, 134 (1972).

[14] S. Somekh, E. Garmire, A. Yariv, H. L. Garvin, R. G. Hunsperger, Appl. Phys. Lett. 22, 46 (1973).

[15] O. S. Oen and M. T. Robinson, Appl. Phys. Lett. 2, 83 (1963).

[16] J. W.Mayer and O. J. Marsh, Appl. Solid State Sci. 1, 239 (1969).

[17] D. B. Hall, Thesis: "Waveguiding and Electrooptic Modulation of Light with GaAs Epitaxial Thin Films," California Institute of Technology (1971).

[18] D. Marcuse, Bell System Tech. J. 48, 3187 (1969).

[19] M. L. Dakss, L. Kuhn, P. F. Heidrich, B. A. Scott, Appl. Phys. Lett. 16, 523 (1970).

H. Kogelnik and T. Sosnowski, Bell System Tech. J. 49, 1602 (1970).

[20] A. Yariv, Quantum Electronics, Wiley, New York (1967).

[21] F. B. Hildebrand, Advanced Calculus for Engineers, Prentice-Hall, New York (1949).

[22] F. Stein, Solid State Physics, 15, 299 (1973).

[23] J. I. Pankove, Optical Processes in Semiconductors, Prentice-Hall, Englewood Cliffs, N. J. (1971).

[24] T. S. Moss, Optical Properties of Semiconductors, Academic Press New York (1959).

[25] H. B. Bebb and E. W. Williams, Semiconductors and Semimetals, Vol. 8, Academic Press (1972).

[26] G. Lasker and F. Stein, Phys. Rev. 133, A553 (1964).

[27] A. Gover, private communication.

[28] Applied Optics and Optical Engineering, R. Kingslake, Ed., Academic Press (1965).

[29] E. M. Garmire and H. Stoll, IEEE J. of Quantum Electronics, Vol. QE-8, No. 10 (1972).

[30] J. A.Borders, Appl. Phys. Lett. 18, 16 (1971).

[31] J. L. McNichols and W. S. Ginell, J. of Appl. Phys. 38, 656 (1967).

[32] L. W. Aukerman, P. W. Davis, R. D. Graft, and T. S. Shilliday, J. of Appl. Phys. 34, 3590 (1963).

L. K. Vodop'yanov and N. I. Kurdiani, Soviet Phys.-Solid State, 8, 204 (1966).

[33] H. J. Stein, Proc of the International Conf. on Ion Implantation in Semiconductors and Other Metals, Plenum Press (1973).

[34] B. R. Pruniaux, J. C. North and G. L. Miller, Ion Implantation in Semiconductors, I. Ruge and J. Granl, Eds., Springer-Verlag (1971).

[35] K. Wohlleben and W. Beck, Z. Naturforschg 21a, 1057 (1966).

[36] L. W. Aukerman, P. W. Davies, R. D. Graft, and T. S. Shilliday, J. of Appl. Phys. 34, 3590 (1963).

[37] G. A. Dussel and K. W. Böer, Phys. Stat. Sol. 39, 391 (1970).

[38] E. W. Williams and H. Barry Bebb, Semiconductors and Semimetals, Vol. 8, Academic Press (1972).

[39] E. W. Williams, Phys. Rev. 168, 922 (1968).

[40] E. W. Williams and D. M.Blacknall, Trans. Met. Soc. AIME 239, 387 (1967).

[41] M. B. Pannish, H. J. Queisser, L. Derick, and S. Sumski, Solid State Electronics 9, 311 (1966).

[42] E. W. J. Mitchell and C. Norris, Proc. of the International Conf. on the Physics of Semiconductors, Kyoto (1966), J. of the Phys. Soc. of Japan, 21, supplement (o966).

[43] M. U. Jeong, J. Shirafuji, and Y. Inuishi, Proc. Intern. Conf. Appl. Ion Beams Semiconductor Tech. (P. Glotin, ed.) p. 153, Grenoble (1967).

[44] V. D. Tkachev, D. S. Domanevskii, V. M. Lomako, and V. F. Stel'makh, Soviet Physics-Semiconductors, 1, 825 (1968).

[45] J. S. Harris and F. H. Eisen, Radiation Effects 7, 123 (1971).

[46] B. Tuck, Phys. Status Solidi 29, 793 (1968).

[47] C. J. Hwang, Phys. Rev. 180, 827 (1969).

[48] J. S. Prener and F. E. Williams, J. Chem. Phys. 25, 261 (1956).

[49] C. C. Klick and J. H. Schulman, Solid State Physics 5, 97 (1957).

[50] F. E. Williams and H. E. Heff, Phys. Rev. 84, 1181 (1951).

[51] E. W. Williams and A. M. White, Solid State Communications 9, 279 (1971).

[52] While experimental evidence exists in support of the contention that the valence band is perturbed much less than the conduction band by external compression and dilation of the crystal lattice (see, for example, R. W. Keyes in Semiconductors and Semi-Metals 4 (Academic Press, 1968)), the hypothesis adopted here (that the conduction and valence bands undergo comparable deformation as a result of catastrophic, internal crystalline disruption) appears to explain in a satisfactory manner the experimental results of Wohlleben and Beck [35] who found, for example, that both n- and p-type GaAs arsenide are similarly compensated by the effects of

high energy protons.

[53] L. W. Aukerman, Semiconductors and Semimetals, Vol. 4, Academic Press (1968).

[54] H. B. Bebb and E. W. Williams, Bull. Am. Phys. Soc. 13, 26 (1968).

[55] L. W. Aukerman and R. D. Graft, Phys. Rev. 127, 1576 (1962).

[56] A. G. Foyt, W. T. Lindley, C. M. Wolfe, and J. P. Donnelly, Solid State Electronics, Vol. 12, Pergamon Press (1969).

[57] D. Kleitman and H. J. Yearian, Phys. Rev. 108, 901 (1958).

[58] K. V. Vaidyanathan and L. A. K. Watt, Proc. Intern. Conf. Appl. Ion Beams Semiconductor Techn. (P. Glotin, ed.), p. 293, Grenoble (1967).

[59] V. C. Burkig, J. L. McNichols, and W. S. Ginell, J. of Appl. Phys. 40, 3268 (1969).

[60] R. G. Hunsperger, private communication.

[61] The attenuation coefficient of a (residual) n-type ($2 \times 10^{16} cm^{-3}$) epitaxial waveguide on an n-type ($1.25 \times 10^{18} cm^{-3}$) substrate was measured (according to the technique described in Chapter 6) to be 1.5 cm^{-1}.

[62] A. Yariv, IEEE J. of Quantum Mechanics, QE-9, 919 (1973).

[63] Located at Hughes Research Laboratories; Malibu, California.

[64] F. A. Jenkins and H. E. White, Fundamentals of Optics, 3rd Ed., McGraw-Hill, New York (1957).

[65] The author wishes to thank N. Hirsch of Hughes Research Laboratories for performing the step-etching.