Electro-Optic and Waveguide Properties of Reverse-Biased Gallium Phosphide *p-n* Junctions

F. K. REINHART, D. F. NELSON, AND J. MCKENNA
Bell Telephone Laboratories, Murray Hill, New Jersey 07974
(Received 8 August 1968)

Observations of light transmitted along the plane of reverse-biased GaP *p-n* junctions show the depletion layer to be birefringent and to produce mode confinement of the light. Extensive measurements were made on diodes having a variety of dopings ($N_D \approx 10^{17}$ to 10^{18} cm^{-3}) and several crystalline orientations ($\mathbf{E}_J \| [111]$ or [110] or [100]) with several laser wavelengths (632.8 nm and 1153 nm from a He-Ne laser and 896.5 nm from a GaAs laser). The measurements were mainly of three types: (1) measurement of the intensity on the exit face of the diode versus position, voltage, and polarization state; (2) measurement of the change of phase with junction voltage of each state of polarization by an optical heterodyne technique; (3) measurement of the phase difference between these two polarization states versus voltage and wavelength by optical compensation methods. We interpret these results in terms of a dielectric waveguide mode in which a higher optical dielectric constant exists in a layer whose width is fixed and approximately equal to the zero-bias junction width. The birefringence arises from the linear electro-optic (Pockels) effect within the junction width, which changes with bias. The optical dielectric constants on the *n* and *p* sides are assumed slightly different. Agreement with this model is generally satisfactory; a few discrepancies, including an anom alous modulation for $\mathbf{E}_{\text{light}} \| \mathbf{E}_J \| [100]$, appear to be best explained by the presence of a small quadratic electro-optic (Kerr) effect. The origin of the higher dielectric constant that gives mode confinement is not known.

I. INTRODUCTION

INITIAL observations[1] of the electro-optic effect in *p-n* junctions were made with visible light in GaP junctions having their electric fields oriented in the [111] direction. Observations[1,2] on the manner of transmission of the light along the plane of the junction led to the conclusion that a dielectric waveguide was present, that is, the index of refraction was higher in the junction region than outside.

A simple planar dielectric waveguide model of an anisotropic crystalline layer (representing the linear electro-optic or Pockels effect on the GaP) contained in an optically isotropic medium of a lower refractive index (representing normal GaP) was worked out. It showed that for typical parameter values only the lowest-order mode of each polarization could be transmitted. Their phase difference could be expressed by a power series in the electric field E. The term in E^2 (and higher-order terms) was, however, too small to explain the observed superlinearity with applied voltage, if the value of the electro-optic coefficient then available was used. However, if the width of the waveguide, which appeared squared in the coefficient of the term linear in E were allowed to depend upon voltage, just as the junction width did, then agreement with the voltage dependence of the phase difference was obtained. In this manner it was concluded that the boundaries of the dielectric waveguide coincided with the boundaries of the *p-n* junction *at all biases*.

Several subsequent observations, to be presented in this paper, do not support this last conclusion. (1) The change of phase with voltage of each state of polarization cannot be accounted for with such a model. This has been explored for diodes of several different crystalline orientations in order to distinguish contributions from the electro-optic effect and waveguide width changes. (2) The intensity distribution of the transmitted light is much less sensitive to voltage than expected from the above model. (3) Phase difference measurements on most diodes have shown discrepancies with the model for biases $\gtrsim -2$ V. Also, new measurements[3] of the linear electro-optic coefficient in bulk crystals have shown it to be of a size sufficient for the E^2 term in the phase difference expression to account for the superlinear voltage dependence.

These new observations indicate that the main dielectric discontinuities which cause mode confinement do not move as the applied voltage is changed. The junction boundaries, which delineate the electro-optic region, do, of course. Thus, a waveguide model which contains four boundaries is needed. Such a double-walled waveguide model was recently explored for the case of symmetric guides.[4] Observations on the intensity distribution, to be presented in this paper, indicate that the index of refraction generally differs on the two sides of a *p-n* junction. Asymmetric dielectric waveguides containing only two boundaries have been studied by several workers.[4–8] In Sec. II of this paper, we explore the combination model, an anisotropic, asymmetric, double-walled waveguide model, which we feel best

[1] D. F. Nelson and F. K. Reinhart, Appl. Phys. Letters **5**, 148 (1964).
[2] A. Ashkin and M. Gershenzon, J. Appl. Phys. **34**, 2116 (1963).
[3] D. F. Nelson and E. H. Turner, J. Appl. Phys. **39**, 3337 (1968).
[4] D. F. Nelson and J. McKenna, J. Appl. Phys. **38**, 4057 (1967).
[5] A. L. McWhorter, Solid-State Electron. **6**, 417 (1963).
[6] R. F. Kazarinov, O. V. Konstantinov, V. I. Perel, and A. L. Efros, Fiz. Tverd. Tela **7**, 1506 (1965) [English transl.: Soviet Phys.—Solid State **7**, 1210 (1965)].
[7] W. W. Anderson, IEEE J. Quantum Electron. **QE1**, 228 (1965).
[8] W. G. Oldham and Ali Bahraman, IEEE J. Quantum Electron. **QE3**, 278 (1967).

accounts for the experimental situation. Other idealizations contained in the original model—the use of abrupt dielectric changes, the use of the spatial average of the electric field, and the neglect of absorption—have been shown elsewhere to be adequate approximations for the lowest-order modes.[4] The problem of exciting the modes of the waveguide has been treated in a separate paper.[9] It shows that for typical diode lengths (~0.07 cm) a substantial intensity from continuum modes, necessitated by mode matching at the input surface, will exist near the junction on the output surface. Elimination of continuum modes from the measurement region by the fabrication of long diodes (~1 cm) has so far not been possible for GaP. A discussion of this problem is given in Sec. IV.

In this paper we present experimental evidence in support of the model just outlined. Section III gives a brief description of the experimental technique. In Sec. IV, we consider the question of whether we are dealing with discrete modes of a waveguide. The possibility of no discrete modes existing if the waveguide is too asymmetric[4–8] makes this an important consideration. Next we present measurements on the intensity of the transmitted light versus position, polarization, and voltage. In Sec. V, we present measurements on the voltage dependence of the phase of the individual polarizations obtained by an optical heterodyne technique. These are compared to phase difference measurements made by compensator methods. Both types of phase measurements have been made on diodes with the junction field in the [111], [110], and [100] directions. Discussion of an anomalous modulation found in the [100] diodes and of the possible presence of the Kerr effect are also presented there. In Sec. VI, we summarize the conclusions.

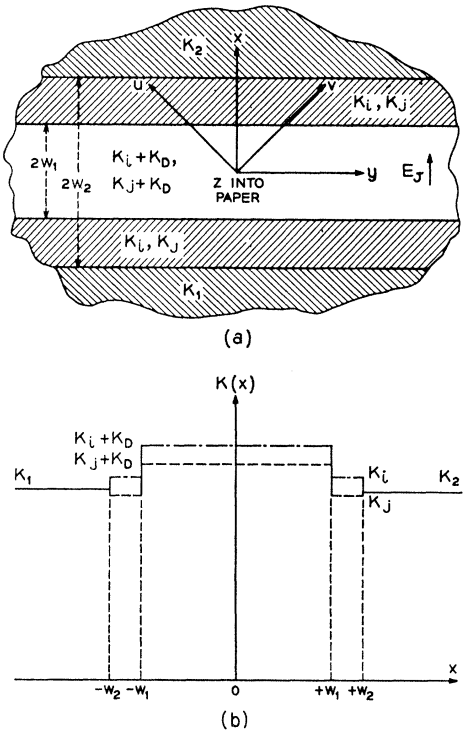

FIG. 1. (a) End-on view of the dielectric waveguide showing the geometry and coordinate system used. The principal axes for $\mathbf{E}_J \| [111]$ or $\mathbf{E}_J \| [100]$ are x, y, z, while for $\mathbf{E}_J \| [110]$ they are u, v, z. (b) Spatial variation of two dielectric constant components for the double-walled waveguide model.

II. ASYMMETRIC DOUBLE-WALLED WAVEGUIDE

In Fig. 1, we illustrate the dielectric constant profile of the model. The positions of the boundary planes, $x=\pm w_1$, are independent of voltage. It is reasonable to assume that these boundaries, which give mode confinement, coincide with the zero-bias junction boundaries. We will show later that only the product $w_1 D$, where D is the fractional increase in the dielectric constant in the region $-w_1 \leq x \leq w_1$, is readily measurable and hence the precise determination of w_1 is not presently possible. The planes $x=\pm w_2$ are the junction boundaries which will depend on voltage. They are also the electro-optic boundaries.

To complete the description of the model it is necessary to prescribe the direction of $\mathbf{E}_J$, the junction electric field, with respect to the crystallographic axes. Three orientations of $\mathbf{E}_J$ are considered. The reasons for these choices have been discussed elsewhere.[4] The first choice is to have $\mathbf{E}_J$ in the [111] direction, and to take the x, y, z axes in the [111], [$\bar{1}$10], [$\bar{1}\bar{1}$2] directions, respectively. A uniform crystal of point group $\bar{4}3m$, such as GaP, when subjected to a uniform field $\mathbf{E}_J$ in the [111] direction has a dielectric matrix diagonal in the above coordinate system with diagonal matrix elements[4,10]

$$K_x^{(0)} = n^2(1+2\delta), \tag{2.1}$$

$$K_y^{(0)} = n^2(1-\delta), \tag{2.2}$$

$$K_z^{(0)} = n^2(1-\delta). \tag{2.3}$$

The quantity δ is defined by

$$\delta = -\bar{E}_J r_{41} n^2/\sqrt{3}, \tag{2.4}$$

where r_{41} (<0 for GaP) is the only nonvanishing electro-optic coefficient for point group $\bar{4}3m$ crystals, n is the index of refraction, and $\bar{E}_J \equiv \langle E_J \rangle_{\text{av}}$ is the spatial average of the junction electric field.

The dielectric matrix describing the asymmetric double-walled waveguide with $\mathbf{E}_J$ in the [111] direction is diagonal in the above coordinate system, and the diagonal matrix elements $K_m (m=x, y, z)$ are defined by

[9] J. McKenna, Bell System Tech. J. 46, 1491 (1967).

[10] S. Namba, J. Opt. Soc. Am. 51, 76 (1961).

the equations

$$K_m = K_m^{(0)} + K_D, \quad |x| < w_1 \tag{2.5}$$

$$= K_m^{(0)}, \quad w_1 < |x| < w_2 \tag{2.6}$$

$$= K_1, \quad x < -w_2 \tag{2.7}$$

$$= K_2, \quad x > w_2. \tag{2.8}$$

In these equations, the quantities $K_m^{(0)}$ are defined in Eqs. (2.1)–(2.4). The quantity K_D represents a fixed addition to the dielectric constant arising from some as yet unknown mechanism during the formation of the junction. We write

$$K_D = n^2 D, \tag{2.9}$$

where we assume that $0 < D \ll 1$. On the p and n sides of the junction the crystal is isotropic and the optical dielectric constant differs from its intrinsic value only by the free carrier density contribution. This statement is supported by recent measurements[3] of the refractive index of GaP on crystals of various dopings up to 10^{18} cm^{-3} which indicated no dependence on doping other than could be expected from the free carrier effect. We write

$$K_j = n^2(1 - d_j), \quad j = 1, 2, \tag{2.10}$$

$$d_1 = d - \eta/2, \quad d_2 = d + \eta/2. \tag{2.11}$$

It should be noted that the definition of the dielectric tensor here differs from that of the double-walled model of Ref. 4.

For an absorptionless, nonmagnetic medium with a dielectric matrix $\mathbf{K}$, Maxwell's curl equations are

$$\nabla \times \mathbf{E} = -\mu_0 \dot{\mathbf{H}} \tag{2.12}$$

and

$$\nabla \times \mathbf{H} = \epsilon_0 \mathbf{K} \cdot \dot{\mathbf{E}}, \tag{2.13}$$

where ϵ_0 and μ_0 are, respectively, the permittivity and permeability of free space. We will always resolve the $\mathbf{E}$ and $\mathbf{H}$ vectors along the principal axes and we look for solutions of Eqs. (2.12) and (2.13) of the form

$$\mathbf{E} = \mathbf{e}(x) \exp i(\omega t - \beta z), \tag{2.14}$$

$$\mathbf{H} = \mathbf{h}(x) \exp i(\omega t - \beta z), \tag{2.15}$$

corresponding to waves traveling in the $+z$ direction with a propagation constant β. Here ω is the angular frequency of the light and t the time. Figure 1 illustrates the geometry assumed.

A straightforward solution of Eqs. (2.12) and (2.13) can be accomplished by the techniques of Ref. 4. However, because of the larger number of boundaries, the algebra involved soon becomes prohibitively complicated. To get around this problem we use a well-known technique from waveguide theory.[11] We resolve all vectors into their components parallel and perpendicular to the direction of propagation (the z direction). Thus,

$$\mathbf{E} = \mathbf{E}_T + E_z \mathbf{U}_z, \quad \mathbf{H} = \mathbf{H}_T + H_z \mathbf{U}_z,$$

$$\nabla = \nabla_T + \mathbf{U}_z(\partial/\partial z), \tag{2.16}$$

where $\mathbf{U}_z$ is a unit vector in the z direction and the subscript T denotes the transverse component. We can define a 2×2, transverse dielectric matrix $\mathbf{K}_T$, which for $\mathbf{E}_J$ in the [111] direction is diagonal in the x,y coordinate system with diagonal matrix elements K_x and K_y. For our assumed solutions (2.14) and (2.15) $(\partial/\partial z)$ can be replaced by $-i\beta$, and for any transverse vector $\mathbf{V}_T$,

$$\mathbf{V}_T \times \mathbf{U}_z = \boldsymbol{\sigma} \cdot \mathbf{V}_T \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{V}_T. \tag{2.17}$$

We can now resolve the Maxwell curl equations into their transverse and longitudinal components, solve for $\mathbf{E}_T$ and $\mathbf{H}_T$ in terms of E_z and H_z, and solve for E_z and H_z in terms of $\mathbf{E}_T$ and $\mathbf{H}_T$:

$$\mathbf{E}_T = (\beta^2 \mathbf{I} - k^2 \mathbf{K}_T)^{-1} \cdot [i\beta(\nabla_T E_z) + i\omega\mu_0 \boldsymbol{\sigma} \cdot (\nabla_T H_z)], \tag{2.18}$$

$$\mathbf{H}_T = (\beta^2 \mathbf{I} + k^2 \boldsymbol{\sigma} \cdot \mathbf{K}_T \cdot \boldsymbol{\sigma})^{-1} \cdot [i\beta(\nabla_T H_z) - i\omega\epsilon_0 \boldsymbol{\sigma} \cdot \mathbf{K}_T \cdot (\nabla_T E_z)], \tag{2.19}$$

$$E_z = (-i/\omega\epsilon_0 K_z) \mathbf{U}_z \cdot (\nabla_T \times \mathbf{H}_T), \tag{2.20}$$

$$H_z = (i/\omega\mu_0) \mathbf{U}_z \cdot (\nabla_T \times \mathbf{E}_T). \tag{2.21}$$

In these equations, $\mathbf{I}$ is the 2×2 unit matrix, $k = \omega(\epsilon_0\mu_0)^{1/2} = 2\pi/\lambda$ is the free-space wave number, and λ is the free-space wavelength. Equations (2.18) and (2.19) can be substituted into Eqs. (2.20) and (2.21) to get two partial differential equations from which E_z and H_z can be determined. The remaining components are then obtained from (2.18) and (2.19).

When $\mathbf{E}_J \| [111]$ we can have both TE $(\mathbf{E}_T \perp \mathbf{E}_J)$ and TM $(\mathbf{E}_T \| \mathbf{E}_J)$ modes. For the TE modes we write

$$H_z = A \cos bx + B \sin bx, \quad |x| \le w_1 \tag{2.22}$$

$$= C_1 \cos q(x + w_1) + D_1 \sin q(x + w_1), \quad -w_2 \le x \le -w_1 \tag{2.23}$$

$$= C_2 \cos q(x - w_1) + D_2 \sin q(x - w_1), \quad w_1 \le x \le w_2 \tag{2.24}$$

$$= \{C_1 \cos q(w_2 - w_1) - D_1 \sin q(w_2 - w_1)\} \times \exp\{p_1(x + w_2)\}, \quad x \le -w_2 \tag{2.25}$$

$$= \{C_2 \cos q(w_2 - w_1) + D_2 \sin q(w_2 - w_1)\} \times \exp\{-p_2(x - w_2)\}, \quad x \ge w_2. \tag{2.26}$$

The TM modes have E_z defined by Eqs. (2.22)–(2.26). Here A, B, C_1, D_1, C_2, and D_2 are constant amplitude factors, and the boundary conditions at $x = \pm w_1$ allow us to solve for C_1, D_1, C_2, and D_2 in terms of A and B. The real parameters β, b, p_1, and p_2 and the (real or imaginary) parameter q satisfy five eigenvalue equa-

[11] W. K. H. Panofsky and M. Phillips, *Classical Electricity and Magnetism* (Addison-Wesley Publishing Co., Inc., Cambridge, Mass., 1955), pp. 192–194.

tions. For the TE modes

$$\beta^2=k^2K_1+p_1^2, \tag{2.27}$$

$$\beta^2=k^2K_2+p_2^2, \tag{2.28}$$

$$\beta^2=k^2(K_y+K_D)-b^2, \tag{2.29}$$

$$\beta^2=k^2K_y-q^2, \tag{2.30}$$

$$2w_1b=F_{\mathrm{TE}}(b,p_1,q)+F_{\mathrm{TE}}(b,p_2,q), \tag{2.31}$$

where

$$F_{\mathrm{TE}}(b,p_j,q)=\tan^{-1}\left[\frac{q\{p_j-q\tan q(w_2-w_1)\}}{b\{q+p_j\tan q(w_2-w_1)\}}\right]. \tag{2.32}$$

Equations (2.27)–(2.30) are the conditions that H_z satisfy its partial differential equation. The quantities A and B satisfy a pair of linear, homogeneous equations arising from the boundary conditions at $x=\pm w_2$ (which we will not write down), and Eq. (2.31) is the condition that these equations have a nontrivial solution.

Certain restrictions must be imposed on the parameters of the waveguide, if solutions to these equations are to exist. First, in order that Eqs. (2.27)–(2.30) have solutions real in β, b, p_1, and p_2, it is necessary that

$$K_y+K_D>K_1(>K_2). \tag{2.33}$$

A single equation for b can be obtained by eliminating β, p_1, p_2, and q between Eqs. (2.27)–(2.32), and this equation can be solved graphically (see Ref. 4, Sec. II B). This equation need not have a solution if the difference between K_1 and K_2 is too great. However, it can be shown that Eqs. (2.27)–(2.32) will have a solution with β, b, p_1, and p_2 real and positive and q either real or imaginary if any one of the following inequalities is satisfied:

$$k(w_2-w_1)(K_y-K_1)^{1/2}>\tfrac{1}{2}\pi, \tag{2.34}$$

$$(K_1-K_2)^{1/2}<(K_y-K_1)^{1/2}\times\tan\{k(w_2-w_1)(K_y-K_1)^{1/2}\}, \tag{2.35}$$

$$2kw_1(K_y+K_D-K_1)^{1/2}>F_{\mathrm{TE}}[k(K_y+K_D-K_1)^{1/2},0,k(K_y-K_1)^{1/2}]+F_{\mathrm{TE}}[k(K_y+K_D-K_1)^{1/2},k(K_1-K_2)^{1/2},\times k(K_y-K_1)^{1/2}]. \tag{2.36}$$

For the TM modes we have Eqs. (2.27) and (2.28) and

$$\beta^2=k^2(K_z+K_D)-(K_z+K_D)b^2/(K_z+K_D), \tag{2.37}$$

$$\beta^2=k^2K_x-K_xq^2/K_z, \tag{2.38}$$

$$2w_1b=F_{\mathrm{TM}}(b,p_1/K_1,q)+F_{\mathrm{TM}}(b,p_2/K_2,q), \tag{2.39}$$

where

$$F_{\mathrm{TM}}(b,p_j/K_j,q)=\tan^{-1}\left[\frac{(K_z+K_D)q\{p_j/K_j-(q/K_z)\tan q(w_2-w_1)\}}{K_zb\{q/K_z+(p_j/K_j)\tan q(w_2-w_1)\}}\right]. \tag{2.40}$$

It can be shown that this set of equations will have an appropriate solution if

$$K_x+K_D>K_1(>K_2),$$

and if any one of the following inequalities is satisfied

$$k(w_2-w_1)[K_z(K_x-K_1)/K_x]^{1/2}>\tfrac{1}{2}\pi, \tag{2.41}$$

$$(K_1-K_2)^{1/2}<[K_2^2(K_x-K_1)/(K_xK_z)]^{1/2}\times\tan\{k(w_2-w_1)[K_z(K_x-K_1)/K_x]^{1/2}\}, \tag{2.42}$$

$$2w_1\tilde{b}>F_{\mathrm{TM}}(\tilde{b},0,\tilde{q})+F_{\mathrm{TM}}[\tilde{b},k(K_1-K_2)^{1/2}/K_2,\tilde{q}], \tag{2.43}$$

where

$$\tilde{b}=k[(K_z+K_D)(K_x+K_D-K_1)/(K_x+K_D)]^{1/2},\quad \tilde{q}=k[K_z(K_x-K_1)/K_x]^{1/2}. \tag{2.44}$$

For a GaP electro-optic diode modulator, typical values of the parameters at full reverse bias are $w_1=140$ nm, $w_2=400$ nm, $\lambda=632.8$ nm, $n=3.31$, $r_{41}=-0.97\times10^{-12}$ m/V, $\bar{E}_J=370$ kV/cm, $D=1.5\times10^{-3}$, $d=4.3\times10^{-5}$, $\eta=1.8\times10^{-5}$. It can be shown that just the lowest-order TE and TM modes exist for these values. Furthermore, for these values of the parameters, the small angle approximation can be made for the tangent wherever it appears in the eigenvalue equations. Treating δ, D, d, and η as small quantities, we can solve the eigenvalue equations for b, q, p_1, and p_2 to lowest order, and β to second order in these small quantities.

For the zero-order TE mode with $\mathbf{E}_J$ in the [111] direction, we find that the parameters are

$$b=kn(D+d-\delta)^{1/2}, \tag{2.45}$$

$$q=kn(d-\delta)^{1/2}, \tag{2.46}$$

$$p_1=k^2n^2\mathfrak{D}-\eta/(4\mathfrak{D}), \tag{2.47}$$

$$p_2=k^2n^2\mathfrak{D}+\eta/(4\mathfrak{D}), \tag{2.48}$$

$$\beta=kn[1-d/2-d^2/8+k^2n^2\mathfrak{D}^2/2+\eta^2/(32k^2n^2\mathfrak{D}^2)], \tag{2.49}$$

where

$$\mathfrak{D}=w_1D+w_2(d-\delta). \tag{2.50}$$

The inequalities (2.34)–(2.36) are to lowest order

$$nk(w_2-w_1)(d_1-\delta)^{1/2}>\tfrac{1}{2}\pi, \tag{2.51}$$

$$\sqrt{\eta}<nk(w_2-w_1)(d_1-\delta), \tag{2.52}$$

$$\sqrt{\eta}<2nk[w_1D+w_2(d_1-\delta)]. \tag{2.53}$$

For the typical parameter values listed earlier, inequalities (2.51) and (2.52) are not satisfied, but inequality (2.53) is satisfied.

For the zero-order TM mode, the expressions for b, q, p_1, p_2, and β, as well as inequalities (2.41)–(2.44) are obtained from Eqs. (2.45)–(2.53) by replacing δ by -2δ.

The phase difference $\Delta\varphi$ can now be written down.[4]

$$\Delta\varphi=l(\beta_{\mathrm{TE}}-\beta_{\mathrm{TM}})=\frac{(2\pi)^3\sqrt{3}n^5lr_{41}\bar{E}_Jw_2}{\lambda^3}\left\{w_1D+w_2\left(d-\frac{n^2r_{41}\bar{E}_J}{2\sqrt{3}}\right)\right\}$$
$$\times\left[1-\left(\frac{\lambda}{4\pi n}\right)^4\frac{\eta^2}{\{w_1D+w_2(d+n^2r_{41}\bar{E}_J/\sqrt{3})\}^2\{w_1D+w_2(d-2n^2r_{41}\bar{E}_J/\sqrt{3})\}^2}\right]. \quad (2.54)$$

In this expression l is the length of the diode.

Other simple orientation choices are $\mathbf{E}_J\|[100]$ with the x, y, z axes in the [100], [01$\bar{1}$], [011] directions, and $\mathbf{E}_J\|[110]$ with the x, y, z axes in the [110], [00$\bar{1}$], [$\bar{1}$10] directions. For the sake of brevity we will not work out these two cases in this paper, but we will give expressions for β and for the phase difference in both cases. We refer the reader to Ref. 4 for the contribution of the Pockels effect to the dielectric tensor, the analogs of Eqs. (2.1)–(2.3), and for the general form of the solutions.

For $\mathbf{E}_J\|[100]$, the expression for β for the zero-order TE mode is obtained from Eqs. (2.49) and (2.50) by replacing δ by $\sqrt{3}\delta$, while the expression for β for the zero-order TM mode is obtained from Eqs. (2.49) and (2.50) by setting $\delta=0$. The expression for $\Delta\varphi$ is

$$\Delta\varphi=\frac{(2\pi)^3n^5lr_{41}\bar{E}_Jw_2}{\lambda^3}\left\{w_1D+w_2\left(d+\frac{n^2r_{41}\bar{E}_J}{2}\right)\right\}\left[1-\left(\frac{\lambda}{4\pi n}\right)^4\frac{\eta^2}{\{w_1D+w_2(d+n^2r_{41}\bar{E}_J)\}^2\{w_1D+w_2d\}^2}\right]. \quad (2.55)$$

For the case of $\mathbf{E}_J\|[100]$ and propagation along [0$\bar{1}$1], the expression for β and $\Delta\varphi$ must be modified by replacing $r_{41}\bar{E}_J$ by $-r_{41}\bar{E}_J$.

For $\mathbf{E}_J\|[110]$ there are no longer TE and TM modes, but there are two modes designated the (+) mode and the (−) mode.[4] For the (+) mode β is obtained from Eqs. (2.49) and (2.50) by replacing δ by $-\sqrt{3}\delta$, and for the (−) mode β is obtained from Eqs. (2.49) and (2.50) by replacing δ by $\sqrt{3}\delta$. In this case the expression for $\Delta\varphi$ is

$$\Delta\varphi=(\beta_+-\beta_-)l=-\frac{2(2\pi)^3n^5lw_2\bar{E}_Jr_{41}(w_1D+w_2d)}{\lambda^3}$$
$$\times\left[1-\left(\frac{\lambda}{4\pi n}\right)^4\frac{\eta^2}{\{w_1D+w_2(d-r_{41}n^2\bar{E}_J)\}^2\{w_1D+w_2(d+r_{41}n^2\bar{E}_J)\}^2}\right]. \quad (2.56)$$

In all three expressions for the phase difference, Eqs. (2.54), (2.55), and (2.56), the last term in square brackets differs little from unity for the typical parameter values quoted earlier.

Throughout this section, we have assumed that the junction field $\mathbf{E}_J$ produces the anisotropy in the junction region by means of the linear electro-optic effect alone, and we have neglected the quadratic electro-optic effects.[12,13] This neglect is not justifiable in all the situations considered in this paper, and so we now consider how the formulas of this section must be modified in order to take into account the Kerr effect.

When referred to the crystallographic axes, the inverse of the dielectric matrix of a crystal of point group $\bar{4}3m$ has the components[13]

$$B_{11}=1/n^2+R_{11}E_1^2+R_{12}(E_2^2+E_3^2), \quad (2.57)$$
$$B_{22}=1/n^2+R_{11}E_2^2+R_{12}(E_3^2+E_1^2), \quad (2.58)$$
$$B_{33}=1/n^2+R_{11}E_3^2+R_{12}(E_1^2+E_2^2), \quad (2.59)$$
$$B_{23}=B_{32}=r_{41}E_1+R_{44}E_2E_3, \quad (2.60)$$
$$B_{31}=B_{13}=r_{41}E_2+R_{44}E_3E_1, \quad (2.61)$$
$$B_{12}=B_{21}=r_{41}E_3+R_{44}E_1E_2, \quad (2.62)$$

where r_{41} is the previously mentioned electro-optic coefficient, and R_{11}, R_{12}, and R_{44} are the three nonvanishing Kerr constants.

For $\mathbf{E}_J\|[100]$,

$$E_1^2=E_J^2, \quad E_2=E_3=0. \quad (2.63)$$

If the (B_{ij}) matrix with these values of E_1, E_2, and E_3 is diagonalized, the principal axes are in the [100], [01$\bar{1}$], and [011] directions, the same as when the Kerr effect is neglected. The corresponding dielectric constants are now

$$K_x^{(0)}=n^2(1-n^2R_{11}\langle E_J^2\rangle_{\mathrm{av}}), \quad (2.64)$$
$$K_y^{(0)}=n^2(1-\sqrt{3}\delta-n^2R_{12}\langle E_J^2\rangle_{\mathrm{av}}), \quad (2.65)$$
$$K_z^{(0)}=n^2(1+\sqrt{3}\delta-n^2R_{12}\langle E_J^2\rangle_{\mathrm{av}}). \quad (2.66)$$

If we compare Eqs. (2.64) and (2.65) with (2.1)–(2.3), it is easy to see that in order to take into account the Kerr effect for the zero-order TE mode with $\mathbf{E}_J\|[100]$,

[12] J. F. Nye, *Physical Properties of Crystals* (Clarendon Press, Oxford, 1957), p. 242.

[13] W. P. Mason, *Crystal Physics of Interaction Process* (Academic Press Inc., New York, 1966), p. 165.

we must modify Eqs. (2.49) and (2.50) for β by replacing δ by $\sqrt{3}\delta+n^2R_{12}\langle E_J^2\rangle_{\text{av}}$, while for the TM mode, β is obtained from Eqs. (2.49) and (2.50) on replacing δ by $n^2R_{11}\langle E_J^2\rangle_{\text{av}}$. Equation (2.55) for the phase difference is no longer valid and must be recalculated. We do not give the result here.

A similar calculation can be carried out when $\mathbf{E}_J$ is in the [111] and [110] directions. The principal axes in the [111] case remain unchanged when the Kerr effect is taken into account. In the [110] case a slight field-dependent rotation of the axes around the [110] direction occurs. We do not reproduce the results since they will not be needed here.

III. EXPERIMENTAL TECHNIQUE

Most of the diodes studied were prepared by Zn diffusion into Te-doped ($N_D\approx10^{17}$–10^{18} cm^{-3}) GaP crystals grown from a gallium solution. A few crystals grown by vapor deposition were also processed. They were carefully selected for optical and electrical homogeneity and oriented.[14] A few junctions were grown by the solution epitaxy method. The faces through which the light entered or left were also oriented crystallographically before being optically polished. A few of the diodes with the junction field in the [111] direction were cleaved along {110} planes. The crystals were mounted so as to prevent the light beam from striking the mounting. Some diodes were provided with SiO antireflection coatings.

The experimental arrangement for measuring the intensity distribution of transmitted light consisted of a 632.8-nm He-Ne laser focused onto the entrance face of the diode by the condenser of a Zeiss polarizing microscope. At the position of the real image (×200) formed of the exit face of the diode by the microscope objective a motor-driven slit was located. A photomultiplier tube detected the light passing through the slit and its signal was recorded on an x-y recorder whose horizontal coordinate represented the slit position.

The phase difference measurements were made in nearly the same experimental arrangement as just described. The polarizers in the microscope were crossed and oriented so as to introduce equal intensities into the two principal states of polarization of the junction waveguide. An Ehringhaus compensator was used to measure the phase difference. For some measurements a Sénarmont compensator was used with appropriate reorientation of the analyzer. The null in the light intensity as the compensator was adjusted was detected sometimes visually, sometimes with the slits and photomultiplier, but most often with the photomultiplier and no slits. When using the photomultiplier, either its signal could be minimized, which is an essentially static procedure, or a small alternating voltage (usually 1-kHz frequency) could be superimposed on the reverse bias of the diode and the ac component of the photomultiplier signal minimized. Most of these measurements were made with a He-Ne 632.8-nm laser beam though a few were also made with a He-Ne 1153-nm laser beam and a GaAs 896.5-nm laser beam.

The measurements of the change of phase with voltage were made by an optical heterodyne technique at 632.8 nm.[15] A small alternating voltage (frequency typically 7.23 MHz) was superposed on the reverse bias of the diode. The light beam emerging from the diode was combined with the beam from a second laser whose frequency was periodically swept (22-Hz repetition frequency) through the frequency range of the first laser beam. No slits were used in this arrangement. A photomultiplier detected the beating between the two beams. The strength of the sidebands created by the modulating voltage on the first beam can be interpreted in terms of the change of phase with voltage. Details of this method are discussed elsewhere.[14]

In order to obtain reliable measurements the whole system had to be aligned very carefully first without and then with the diode inserted. Since we have only indirect evidence of the quality of the alignment of a diode, a standard approach had to be adopted which was used for all types of optical measurements: (a) The diode was rotated about an axis that was approximately perpendicular to the junction plane until the entire diode exit face appeared sharply in focus at a large magnification; (b) the diode was rotated around the axis of light propagation until the junction plane was in the desired orientation with respect to the crosshair in the microscope ocular that was fixed in space; (c) the beam was carefully focused on the entrance face; (d) the diode was moved laterally until the beam hit the junction region by observing the exit face of the diode; (e) the junction plane was adjusted with respect to the beam axis until a "clean" looking mode pattern appeared when focusing on the exit face. Because of the diode movements involved, steps (c)–(e) had to be carefully repeated a number of times. When aligning the slits for phase difference measurements a reverse bias close to breakdown was applied to the diode. The slits were moved to coincide with the peak of the transmitted light when the polarizers were crossed. When the compensation was varied the deepest minimum occurred at this location also (see Sec. IV, Fig. 5). Alignment steps (c)–(e) were then repeated to obtain the maximum phase difference reading while retaining as deep a compensation minimum as possible.

To obtain a deep compensation minimum either with or without slits it was often helpful to change the laser beam spot size on the entrance face of the diode. An optimum spot size is one which couples the greatest amount of the light into the discrete modes and the least amount into the continuum modes. This is expected to lead to the deepest compensation minimum. Since the spot sizes were diffraction limited, alteration of the

[14] F. K. Reinhart, J. Appl. Phys. 39, 3426 (1968).

[15] I. P. Kaminow, Appl. Phys. Letters 7, 123 (1965).

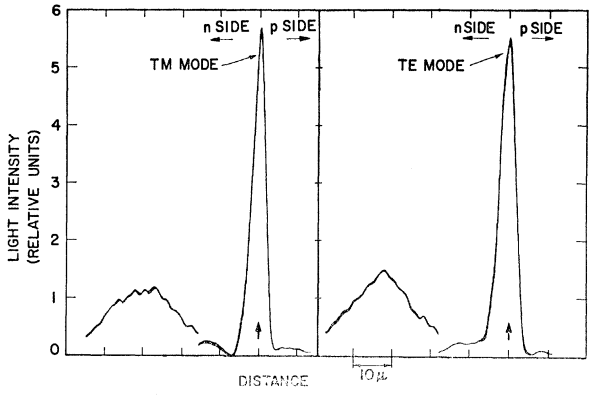

FIG. 2. Intensity distribution on the output surface of diode KC46CA ($E_J \| [111]$) of a 632.8-nm He-Ne laser beam focused on the input surface on the bulk n side of the crystal (low broad curves) and on the p-n junction (high narrow curves). The scales for the curves of each polarization are the same.

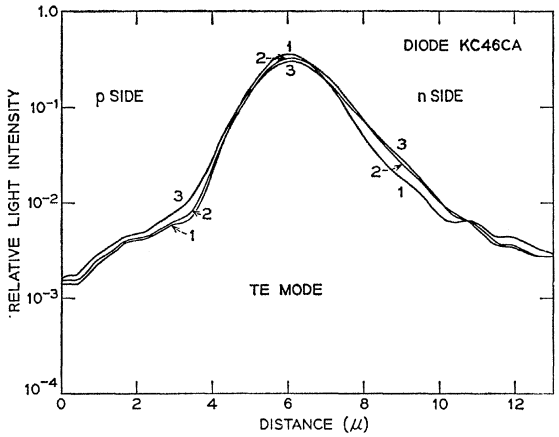

FIG. 3. Intensity distributions for the TE mode on the output surface of the diode for (1) 0 V, (2) −14 V, (3) −28 V.

convergence angle of the light at the diode with the aid of an auxiliary lens was used to change the spot size.

The coincidence in location of the intensity maximum and the deepest compensation minimum gives a strong empirical argument for the use of slits in phase difference measurements. The use of slits, however, makes the measurements sensitive to alignment procedure. Variations of the measured phase difference of 30% were often obtained if the entire alignment and measurement procedure were repeated. We attribute this to the fact that not all the light is coupled into the discrete waveguide modes but that some of the light is coupled into the continuum waveguide modes at the entrance face of the diode.[9] Interference of these two components of the transmitted light can cause a variation of the measured phase difference with position and with the alignment. A smaller variation of the measured phase difference with alignment was found when no slits were used. For this reason all the data presented in this paper on phase differences were obtained without the use of slits. It should be pointed out, however, that the data presented in an earlier letter[1] were obtained using slits aligned as described above or using visual judgment which weighs heavily the attainment of the deepest compensation minimum at a specific location just as the slits do.

A further problem with the phase difference measurements is the disagreement of phase differences measured by the static and the ac techniques. With slits being used this disagreement can easily be a factor of 2. Without slits the disagreement is much smaller but the ac technique always yields larger phase differences. The discrepancy can be 15% but is generally less than this. This disagreement is believed to be the direct result of the presence of the continuum mode light. Ideally when the compensator is varied through a phase change of 2π the transmitted light should vary sinusoidally, reaching zero at a particular compensation value. In practice the minimum reached is not zero and its value increases as the reverse bias is increased. Recent calculations[9] account for this effect qualitatively and show it to be due to an increasing proportion of the continuum modes. On this basis it can be shown that the ac measurements will be too high and the static measurements too low. For this reason all the phase difference data presented in this paper are an unweighted average of the results of the two techniques.

IV. INTENSITY MEASUREMENTS

Measurements of the dependence of the light intensity versus position on the exit surface of a diode modulator as a function of voltage and polarization are potentially capable of distinguishing between various waveguide models. With this goal in mind many such measurements have been made and will be discussed in this section. First we wish to show by the simplest of observations and theoretical arguments that discrete modes of the waveguide are transmitted. Some evidence has already been presented in favor of this conclusion.[1,2]

In Fig. 2 is shown the intensity of light on the exit surface of a diode under two conditions of input focusing and two polarizations. To obtain these data, a 632.8-nm He-Ne laser beam was focused on the input surface of the diode, and the diode, whose junction field was in the [111] direction, was in full reverse bias. The image of the output surface was scanned photoelectrically using a narrow slit. The broad low curves in each polarization were obtained when the input beam was focused on the surface of the homogeneous crystal on the n side of the junction. When the position of the focus was moved laterally over to the p-n junction, the peak intensity of the light at the exit surface was increased by a factor of 5 and the width of the distribution was narrowed by a similar factor. This factor is almost the same for each of the principal polarizations, being only slightly larger for the TM mode. This indicates that the higher dielectric constant inside the p-n junction that causes mode confinement is nearly the same for each

TABLE I. α_j values as defined by Eq. (5.3) for the orientations studied and measured values of the fractional optical dielectric constant increase (D) and the linear electro-optic coefficient (r_{41}).

Diode	Orientation E_J	Orientation z	α_j TE (−)	α_j TM (+)	$D\times10^3$ TE (−)	$D\times10^3$ TM (+)	$\Delta\varphi$	$r_{41}\times10^{12}$ [m/V] TE (−)	$r_{41}\times10^{12}$ [m/V] TM (+)	$\Delta\varphi$
KC71BB	[111]	⊥[111]	1	−2	1.1	1.1	1.1	−0.69	−0.78	−0.79
KC46CA	[111]	⊥[111]	1	−2	1.6	1.4	1.4	−0.68	−0.63	−0.97
KC50BA	[110]	‖[1̄10]	$\sqrt{3}$	$-\sqrt{3}$	2.0	2.2	1.8	−0.71	−1.18	···
KC50CA	[100]	‖[011]	$\sqrt{3}$	0	1.8	···	2.2	−0.52	···	+0.29
KC50CB	[100]	‖[01̄1]	$-\sqrt{3}$	0	1.6	···	1.2	−0.97	···	−0.68

polarization. This is true in spite of the fact that for the TM mode the linear electro-optic effect increases the dielectric constant, as seen from Eqs. (2.1), (2.4), and (2.5), while for the TE mode it decreases the dielectric constant, as seen from Eqs. (2.2), (2.4), and (2.5). From these facts we conclude that the "built-in" higher dielectric constant in the junction is large compared to the electro-optic contribution, that is $D\gg2\delta$. We calculate δ to be $\lesssim2\times10^{-4}$ from the junction electric field determined from capacitance measurements. Therefore $D\gg4\times10^{-4}$ which is consistent with the measurements of D to be presented in the next section. For $D=1.5\times10^{-3}$ (see Table I) and a waveguide width comparable to typical junction widths, one discrete mode of each polarization will exist unless $\eta/D\geq0.16$,[4] that is, unless the asymmetry is too large. Using the Drude formula to calculate the free carrier depression of the dielectric constant on the homogeneous n side of this diode we find $d_1=3.4\times10^{-5}$. From the diffusion profile determined from capacitance measurements we calculate $d_2=5.2\times10^{-5}$ at a distance 1.5 μm beyond the junction on the p side. These values lead to $\eta/D=1.2\times10^{-2}$, well within the tolerable limit for the existence of discrete modes which must thus be present in the diode of Fig. 2.

Though the above argument as well as previous results[1,2] are strong evidence for the presence of discrete modes, a quantitative demonstration that the transmitted intensity versus position, voltage, and polarization agrees with that expected for discrete modes would be very useful. Figures 3 and 4 show the transmitted light distribution on the output surface for the diode just discussed. To obtain these data the input focusing was adjusted to minimize the light intensity on the wings of the distributions. The decrease of intensity on each side of the distributions is roughly exponential as expected from the model of Sec. II. No direct comparison of experiment and theory is possible, however, since Fig. 5 demonstrates that the light distribution is not that of a single mode. The distributions of Fig. 5 were measured with the diode between crossed polarizers and the principal electro-optic axes at 45° to the transmission axes of the polarizers. The different curves correspond to the insertion of various amounts of optical phase difference by an Ehringhaus compensator also placed between the crossed polarizers. If the light distribution was due to a combination of a single mode of each polarization, the curves for each different amount of optical compensation should be identical in shape and only translated vertically with respect to each other. Small deviations from this are to be expected since the two modes have different shapes.[4] Clearly the deviations seen in Fig. 5 are too large to be explained in this manner.

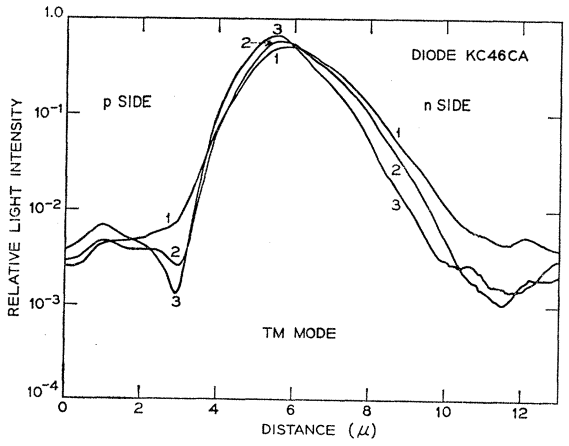

FIG. 4. Intensity distributions for the TM mode on the output surface of the diode for (1) 0 V, (2) −14 V, (3) −28 V.

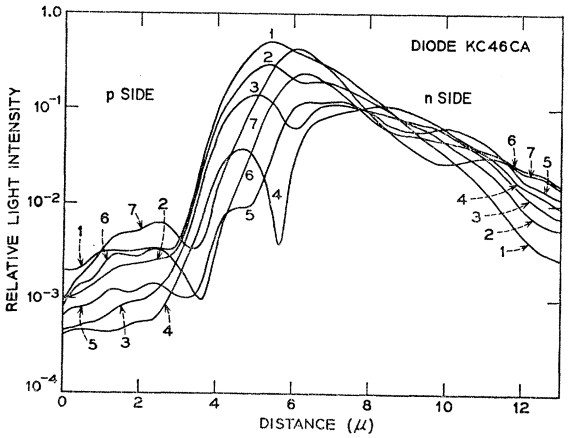

FIG. 5. Intensity distributions on the output surface of the diode which is at −28-V bias and placed between crossed polarizers for (1) 0°, (2) 30°, (3) 60°, (4) 90°, (5) 120°, (6) 150°, (7) 180° of phase difference introduced by an Ehringhaus compensator.

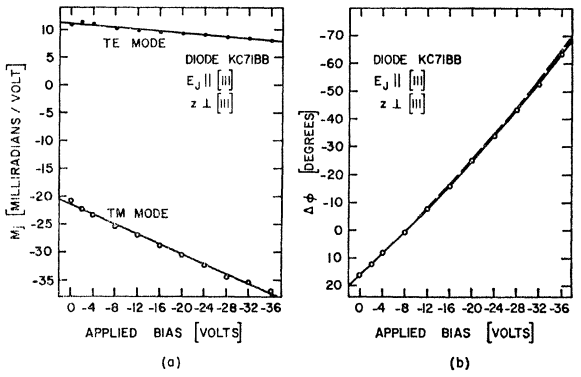

FIG. 6. Phase measurements versus applied bias at $\lambda = 632.8$ nm. (a) Modulation index per volt measurements with least-square-fitted straight lines. (b) Phase difference measurements with least-square-fitted parabola (solid) and integrated difference of modulation index per volt with a constant strain birefringence (dashed) added to agree with the $\Delta\varphi$ data.

The explanation, we believe, is that there is a significant admixture of continuum modes[9] to the two discrete modes. The continuum modes result from satisfying the electromagnetic boundary conditions at the input surface and from the output surface not being in the far field of the radiation pattern of the input surface for 0.07-cm-long diodes.[9] This problem could be largely eliminated, if diodes of sufficient length ($\gtrsim 1$ cm) could be fabricated. Unfortunately materials problems have prevented this so far.

In spite of our inability to verify numerically the shape of the intensity distributions predicted by the model of Sec. II, we can glean considerable information from the distributions of Figs. 3 and 4 in support of the model. First, the peak intensity of the TM mode is a little larger than the TE mode. Second, the peak intensity of the TM mode grows a little with increased reverse bias while the width of its distribution narrows a bit. Third, the peak intensity of the TE mode diminishes a little with increased reverse bias while the width of its distribution broadens a bit. These three facts are all consistent with an electro-optic contribution to the dielectric constant which is small compared to the dielectric constant contribution which causes mode confinement and which adds to the latter for the TM mode and subtracts from it for the TE mode. This is in agreement with the model. Fourth, the effect of voltage on the distribution is larger for the TM mode than the TE mode. This also agrees with expectations as seen from Eqs. (2.1), (2.2), and (2.51)–(2.54). Fifth, the small changes in the width of the distributions are consistent simply with electro-optic contributions and hence with the thickness of the higher dielectric constant region which produces mode confinement being independent of voltage as assumed in the model. If this thickness were to vary with voltage like the junction width does, as originally believed,[1] the intensity distributions would have to change by roughly a factor of 2 in the voltage range depicted in Figs. 3 and 4. Sixth, it is quite apparent that the intensity distributions are asymmetric as was accounted for in the model. The faster decrease on the p side indicates that the dielectric constant is depressed by the presence of free carriers there more than on the n side. This is expected since the junction was formed by diffusion of an acceptor into a uniform n-type crystal. Furthermore, the greater amount of continuum light evidenced on the n side in Fig. 5 is consistent with the predictions of the mode matching calculations.[9] These calculations predict that the continuum light is a superposition of two basic types of continuum light waves. The first type consists of plane waves propagating at an angle to the waveguide axis, starting on either the n or p sides. These waves are partially reflected and partially transmitted at the waveguide boundaries. The second type consists of plane waves starting on the n side of the junction which are totally reflected at the junction boundary on the p side. The energy in the continuum modes depends on how the junction is excited. However, all the energy in the totally reflected waves is on the n side, tending to increase the relative energy on the n side compared to the p side.

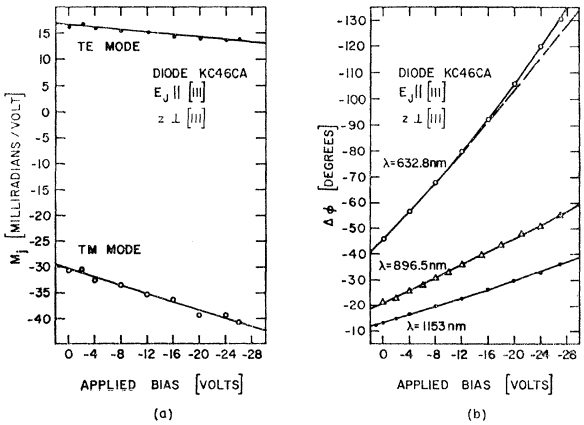

FIG. 7. Same caption as Fig. 6.

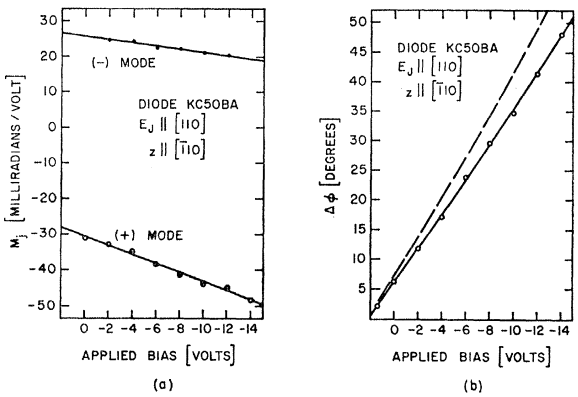

FIG. 8. Same caption as Fig. 6.

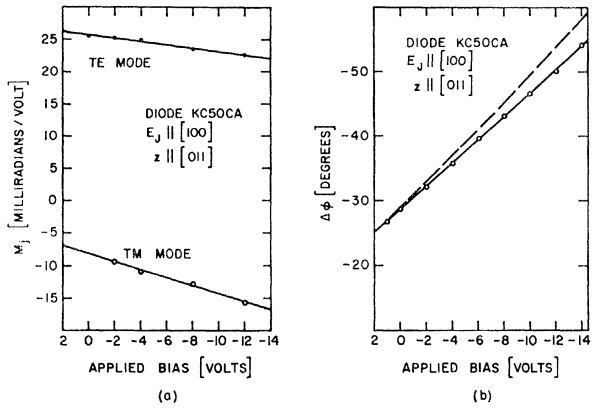

FIG. 9. Same caption as Fig. 6.

V. PHASE MEASUREMENTS

In this section we present measurements of the phase of light waves transmitted along the junction plane of GaP p-n junctions. We will show that these measurements are best interpreted in terms of the waveguide model of Sec. II, that is, by the linear electro-optic effect acting upon the discrete modes of a dielectric waveguide whose major mode-confining dielectric discontinuities are fixed in the crystal while the electro-optic effect boundaries move with the junction boundaries.

The phase measurements are of two types: (1) measurements of the change of phase with applied voltage for each of the two modes of differing polarization and (2) measurements of the phase difference between these two modes. The former are made by an optical heterodyne technique and the latter by an optical compensator method. Both methods were described in Sec. III. It is interesting to note that the application of SiO antireflection coatings did not influence the results of the phase difference measurements systematically. In Figs. 6(a)–10(a) are presented the results of the measurements of the modulation index per volt M_j defined by

$$M_j = d(\beta_j l)/dV \tag{5.1}$$

plotted versus the bias voltage on the diode. The index j denotes the mode. Within experimental error the results are describable by straight lines for both polarizations and for all of the diodes studied. The characteristics of the diodes for which data are presented in Figs. 6–10 are given in Table II.

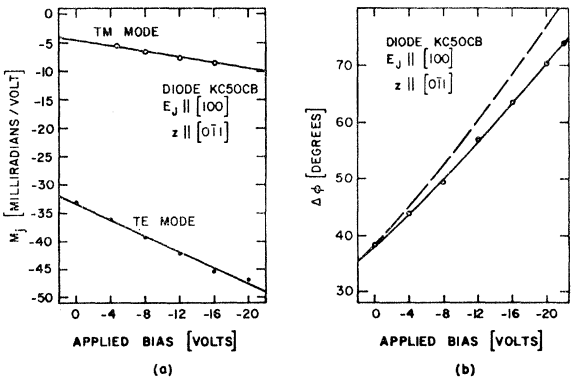

FIG. 10. Same caption as Fig. 6.

Because of the method of measurement of M_j only its absolute value is found for each polarization. The relative algebraic sign of the two used in plotting the results was determined by comparison with the phase difference measurements on the same diode. These results are shown in Figs. 6(b)–10(b). The comparison was made by integrating first the modulation index per volt data with respect to voltage for each polarization and then seeing whether a sum or a difference of the phases agreed with the phase difference results. The choices of signs determined in this way were consistent with the predicted signs of the model in each case. For the diode with $\mathbf{E}_J \| [110]$ where the phase changes are expected to be equal and opposite for the two modes a further check of the sign was possible. When an equal amount of both polarizations was sent through the diode, the two first-order side bands seen in the heterodyne technique were of vastly different strengths. This showed that the signs of the phases were opposite.

From Eqs. (2.49) and (5.1) the waveguide model of Sec. II yields

$$M_j = n^3 k^3 l \mathfrak{D}_j \left\{ 1 - \frac{\eta^2}{16(kn\mathfrak{D}_j)^4} \right\} \frac{d\mathfrak{D}_j}{dV}, \tag{5.2}$$

where

$$\mathfrak{D}_j = w_1 D + w_2(d - \alpha_j \delta), \tag{5.3}$$

and α_j is given in Table I for the various crystalline orientations and modes. In writing Eq. (5.2) we have assumed that l does not vary with voltage through the piezoelectric or electrostrictive effects. We will return to this point later. The term in η^2 arises from the asymmetry of the waveguide. For the typical values of η given in Sec. IV this term is negligible and may be dropped from Eq. (5.2).

From the insensitivity of the intensity distributions to voltage changes, shown in Sec. IV, we conclude that the increased dielectric constant in the junction, represented by D, and the width over which it extends, w_1, are voltage-independent. Thus the first term in Eq. (5.3) is independent of voltage. Since by definition $2w_2\bar{E}_J = V_0 - V$, where $V_0 \cong 2$ V for GaP is the built-in potential, the third term in Eq. (5.3) can be expressed as

$$\alpha_j w_2 \delta = -(\alpha_j n^2 r_{41}/2\sqrt{3})(V_0 - V), \tag{5.4}$$

which gives a linear contribution to the voltage dependence of M_j.

The second term $w_2 d$ in Eq. (5.3) requires closer inspection. First, its presence in Eq. (5.3) depends upon the assumption of the model that the electro-optic effect in the two regions, $w_1 \leq |x| \leq w_2$, changes the dielectric constant from its *intrinsic* value. If in these two regions the electro-optic effect were to change the dielectric

TABLE II. Diode characteristics: length (l); zero-bias junction half-width [$w_2(V=0)$]; average junction electric field at full reverse bias ($\bar{E}_{J\,\max}$); capacitance parameters (V_0',γ) determined from $C=C_0(1-V/V_0)^{-\gamma}$; doping gradient (a) determined from capacitance data; bulk electron concentration from conductivity measurements (n_{bulk}).

Diode	l [10^{-2} cm]	$w_2(V=0)$ [nm]	$\bar{E}_{J\,\max}$ [kV/cm]	V_0' [V]	γ	a [10^{22} cm^{-4}]	n_{bulk} [10^{17} cm^{-3}]
KC71BB	5.24	168	340	1.9	0.407	0.27	2.6
KC46CA	7.10	139	370	1.8	0.380	0.50	3.8
KC50BA	8.51	80.5	510	1.8	0.326	3.4	3.5
KC50CA	9.22	82.2	490	1.6	0.308	3.1	
KC50CB	7.15	148	325	1.7	0.355	0.42	

constant from the *depressed* value characteristic of the contiguous bulk region, the w_2d term would disappear from Eq. (5.3). Since, as we will see, we cannot positively verify the presence of this small term from experiments, either assumption is possible. For two reasons, however, we will use the assumption which retains the term. (1) Its presence improves the fit, as we will see, where $\mathbf{E}_T\|\mathbf{E}_J\|[100]$. (2) It uses only known mechanisms of altering the optical dielectric constant, that is, the plasma resonance from free carriers and the linear electro-optic effect. The alternate model would require a depression of the optical dielectric constant from the intrinsic value by some means other than the presence of mobile charge. Second, the maximum value of w_2d is only about 4% of w_1D and so can be neglected in the $\mathfrak{D}_j$ factor of Eq. (5.2). Third, the maximum of $|d(w_2d)/dV|$ is less than 10% of $|d(\alpha_j w_2\delta)/dV|$ for $V\leq 0$ and so may be neglected in the $d\mathfrak{D}_j/dV$ factor of Eq. (5.2) whenever $\alpha_j\neq 0$. The term $d(w_2d)/dV$ would be significant if $\alpha_j\neq 0$ only for a forward-biased abrupt junction whose n and p doping levels both exceed 5×10^{17} cm^{-3}.

With the approximations just outlined Eq. (5.2) can be simplified to be

$$M_j=-(\alpha_j n^5 k^3 l r_{41}/2\sqrt{3}) \times\{w_1D+\alpha_j n^2 r_{41}(V_0-V)/2\sqrt{3}\} \quad (5.5)$$

for all cases studied here except for $\mathbf{E}_J\|[100]$, where $\alpha_j=0$, when we expect

$$M_{\text{TM}}=n^3k^3lw_1D[d(w_2d)/dV]. \quad (5.6)$$

The only unknown in Eq. (5.5) is the product w_1D. It can be found for a particular diode from the intercept of the M_j versus V curve at $V=V_0$ for each polarization mode. The consistency of the values can then be checked. The slope of such a curve can be used to obtain a value of r_{41} from the second term of Eq. (5.5) for comparison with the value found on bulk GaP crystals.[3] In Eq. (5.6) w_1D is also unknown while d is known only approximately. The voltage dependence of w_2 is known from capacitance versus voltage measurements for each diode.

Within the validity of Eq. (5.5), Eqs. (2.54)–(2.56) for $\Delta\varphi$ can be simplified by putting $d=\eta=0$ and $2w_2\bar{E}_J=V_0-V$. In the cases of $\mathbf{E}_J\|[111]$ and $\mathbf{E}_J\|[100]$, $\Delta\varphi$ then contains just terms linear and quadratic in (V_0-V), and in the case of $\mathbf{E}_J\|[110]$, $\Delta\varphi$ contains only a term linear in (V_0-V). The presence of strain birefringence will add a voltage-independent term to $\Delta\varphi$. From the measured coefficient of the (V_0-V) term a value for w_1D can be found for the diode and from the measured coefficient of the $(V_0-V)^2$ term a value of r_{41} can be found for comparison with bulk crystal measurements of it.[3]

The value of D, the fractional dielectric constant increase in the junction, can be determined from the value of w_1D found from experiment, if the width w_1 within which D exists is known. We do not have any practical way of determining w_1 from experiment at present. However, it is reasonable that w_1 is approximately the zero-bias junction width since the dielectric constant increase was presumably created during the junction formation when no bias was on the junction. It was, of course, at an elevated temperature at that time. In Table I we present the values of D found by using the zero-bias junction half-width $w_2(V=0)$ for w_1. The clamped value[3] of $r_{41}=-0.97\times 10^{-12}$ m/V was used in the calculation, because of the bulk crystal clamping of the electro-optic layer. However, if the unclamped value of r_{41} were used, D would increase only by 13%. The agreement among the values of the fractional dielectric step height D found from the various measurements on a given diode is within 25%. In view of the measurement problems created by the continuum mode light this is regarded as satisfactory agreement.

Table I also gives the values of r_{41} found from M_j and $\Delta\varphi$ measurements for those orientations where this is possible. Except for one case (a positive value of r_{41}!) the agreement with the bulk value[3] is regarded as satisfactory. We also checked the voltage dependence of $\Delta\varphi$ for $\mathbf{E}_J\|[\bar{1}\bar{1}\bar{1}]$. The sublinear voltage dependence predicted by Eq. (2.54) after replacing $\bar{E}_J$ with $-\bar{E}_J$ was experimentally verified. All this agreement gives strong support to the model.

Though the majority of the data presented is in agreement with the model as presented so far, there are three observations that are deemed to fall outside the range of uncertainty that arises from continuum mode effects. (1) The modulation index per volt measured for the TM mode in diodes with $\mathbf{E}_J\|[100]$ is much larger than expected from Eq. (5.6) and is of a different

functional form. This we will refer to as the anomalous modulation. (2) The $\Delta\varphi$-versus-V curve for diode KC50CA $\{E_J\|[100]\}$ is slightly superlinear rather than slightly sublinear as expected. (3) The $\Delta\varphi$-versus-V curve for diode KC50BA $\{E_J\|[110]\}$ is slightly superlinear rather than linear as expected. This results from the unexpected dissimilarity of the slopes of M_j versus V for the (+) and (−) modes. A number of effects must be considered as possible explanations of these discrepancies.

(1) Waveguide width changes. This effect, represented by the w_2d term discussed previously, is too small in magnitude and has the wrong functional form to explain the anomalous modulation.

(2) Mode mixing. A sufficiently large strain through the elasto-optic effect can cause a coupling of energy between the two polarization modes.[4,16] Calculations indicate that strains of this size would also cause the optic axes in the electro-optic region to be misoriented by $\sim 10°$. Misorientations of this size have not been observed.

(3) Piezoelectric and electrostrictive length changes. A change of the junction electric field through these effects can cause a phase change of $n\Delta l$ in addition to the electro-optic phase change of $l\Delta n$. If the bulk crystal surrounding the junction did not constrain expansion or contraction of the junction layer, the piezoelectric effect would be much larger than the observed discrepancies on the basis of the recently determined piezoelectric coefficient of GaP.[3] However, an approximate calculation of the clamping effect of the bulk crystal (due to Tiersten) shows that the contribution of the piezoelectric effect to M_j is only 1% of the contribution of the electro-optic effect. It should also be remembered that piezoelectric length changes could only affect the M_j measurements and not the $\Delta\varphi$ measurements.

(4) Symmetry-violating indirect electro-optic effect.[3] If the application of the reverse bias to the junction excites acoustic resonances of the diode, the time- and space-dependent strains created will not possess the symmetry of those in an infinite crystal. They will be affected by the size and shape of the crystal. The indirect electro-optic effect arising from the combination of the piezoelectric and the elasto-optic effects will not then possess the symmetry of the direct or true electro-optic effect. The constraining or clamping effect of the bulk crystal on the junction layer prevents this mechanism from being significant as discussed above.

(5) Quadratic electro-optic effects. These include the Kerr effect and morphic effects. The latter may arise when the electric field is strong enough to lower the symmetry of the crystal. We will ignore them here. The Kerr effect will produce new terms in the phase expressions of the form given in Sec. II. The magnitude of the Kerr coefficients R_{11}, R_{12}, R_{44} of GaP are not presently known. An estimate of the Kerr effect can be made by an extension of the well-established Miller rule[17] for the second-order nonlinear susceptibilities (a third-rank tensor) to the third-order nonlinear susceptibilities (a fourth-rank tensor). Such an extension seems to be justified in the light of the data available at present. (1) The Kerr effect in some paraelectric pervoskites was measured by Geusic *et al.*[18] They found that the Kerr effect in these materials appears to be proportional to the square of the polarization with an apparently common proportionality constant. Since the refractive indices of these investigated materials are very similar, the above experimental finding gives support to an extension of Miller's rule concerning the third-order nonlinear susceptibilities (one of which is represented by the Kerr effect).[19] (2) Wynne and Boyd[20] have recently predicted third-order mixing constants for Ge based on a measured third-order nonlinear coefficient of calcite by making use of Miller's phenomenological rule. Their recent experiments have confirmed this prediction. They also have studied Si for which theory and experiment only differ by a factor of 3. (3) Patel *et al.*[21] also measured a third-order mixing coefficient in GaAs and also found good agreement with the phenomenological theory. These third-order mixing coefficients are closely related to the Kerr coefficients. Based on these facts we can justify a prediction of the magnitude of the Kerr effect for GaP by assuming that the coefficient of the third-order anharmonic force term is similar to the coefficient for Ge, Si, GaAs and the perovskites. The result of such a prediction is $|R| \approx 10^{-21}$ m^2/V^2. In the noncentrosymmetric GaP the second-order anharmonic forces that are thought to be responsible for the linear electro-optic effect also contribute to the Kerr effect. This contribution was estimated to be only 5×10^{-24} m^2/V^2 and can therefore be neglected.

The experimental discrepancies can be qualitatively explained by assuming Kerr constants of the magnitude discussed above. The experimental accuracy, however, is not sufficient to allow for any conclusions concerning the importance of the Kerr-effect contribution with possibly one exception, the anomalous modulation. It could be quantitatively accounted for if $R_{11}\approx -4\times10^{-21}$ m^2/V^2. In Figs. 11 and 12 are shown the improved fits to the anomalous modulation when the Kerr effect combined with the width change contribution to Eq. (5.6) for diodes KC50CA and KC50CB. A definitive proof of the existence of a Kerr effect would be the observation in a diode with $\mathbf{E}_J\|[110]$ of a rotation of the

[16] J. McKenna and J. A. Morrison, Bell System Tech. J. 47, 1933 (1968).

[17] R. C. Miller, Appl. Phys. Letters 5, 17 (1964).

[18] J. E. Geusic, S. K. Kurtz, L. G. Van Uitert, and S. H. Wemple, Appl. Phys. Letters 4, 141 (1964).

[19] S. K. Kurtz and F. N. H. Robinson, Appl. Phys. Letters 10, 62 (1967).

[20] J. J. Wynne and G. D. Boyd, Appl. Phys. Letters 12, 191 (1968).

[21] C. K. N. Patel, R. F. Slusher, and P. A. Fleury, Phys. Rev. Letters 17, 1011 (1966).

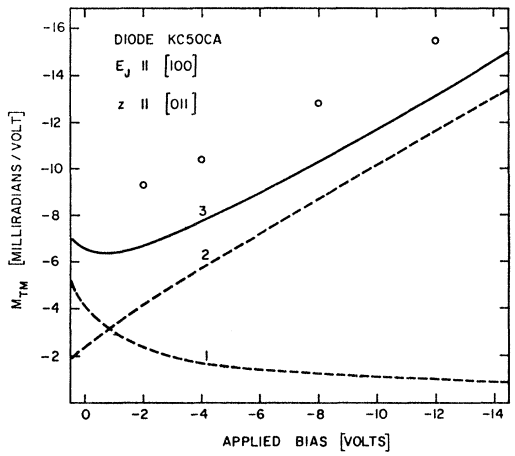

FIG. 11. Anomalous modulation versus applied bias (circles): (1) calculated waveguide width change contribution, (2) proposed Kerr-effect contribution, (3) combination of (1) and (2).

optic axes which are nominally at 45° to the junction plane by an angle which is proportional to the electric field in the junction. With the magnitude of R_{11} indicated above this expected rotation is less than 2° for full reverse bias. We have searched for this effect experimentally but have been unable to establish conclusively either its presence or absence because of problems presumably due to the presence of the continuum mode light. Thus, we feel that the Kerr effect is the most likely explanation of the discrepancies but, in view of the difficulties of measurement pointed out, we feel that the analysis in terms of a small additional Kerr effect falls short of a proof of its presence.

The origin of the larger dielectric constant in the junction, represented by D, is an interesting question. One possibility that suggests itself is the depression of the dielectric constant in the bulk crystal by the plasma resonance of the free carriers. In the depletion layer of the junction this depression would be absent. The magnitude of this effect is easily calculable and is found to be at least 50 times too small. It should be noted that, although w_1D is measured and w_1 is not known *a priori* to be the zero-bias junction width as assumed, any

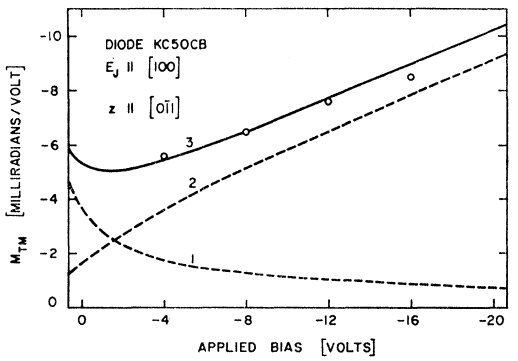

FIG. 12. Same caption as Fig. 11.

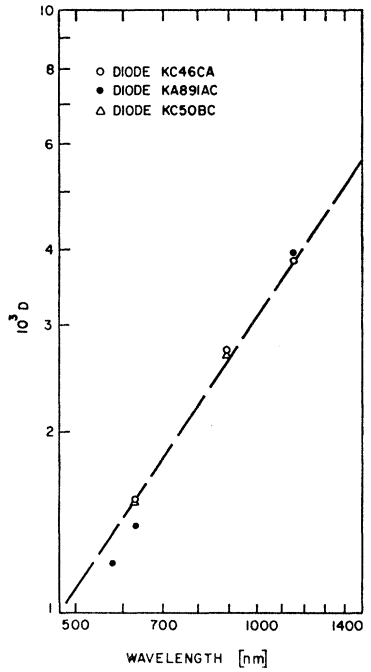

FIG. 13. Wavelength dependence of D. The line through the experimental points corresponds to $D \propto \lambda^{3/2}$.

drastic increase of w_1 from this value in order to decrease D would cause the expected light intensity distributions to be much broader than the measured ones presented in Sec. IV.

A number of characteristics of D can be derived from the experimental values of w_1D, if we assume that w_1 is the zero-bias junction width. Its boundaries do not move with voltage; its magnitude does not vary with voltage, crystal orientation, or junction-formation technique (junctions formed by liquid phase epitaxy have a comparable D to diffused junctions). D depends on the wavelength λ as roughly $\lambda^{3/2}$ according to the data presented in Fig. 13. This seems to rule out any simple band-gap shift as an explanation since it would cause D to decrease with increasing wavelength. Finally we have found w_1D to vary only from 2.3×10^{-8} to 1.8×10^{-8} cm when the dopant concentration, as deduced from the capacitance characteristics, varied from 4.5×10^{16} to 2.7×10^{17} cm^{-3} for a series of eight diodes. One other diode yielded $w_1D=3.6\times10^{-8}$ cm for 4.7×10^{17} cm^{-3}. We feel that these data are inconclusive in determining any dependence of w_1D on doping. At present we do not know of any mechanism to explain these facts and hence the origin of D.

VI. CONCLUSION

Our observations of light transmitted along the plane of reverse biased p-n junctions in GaP in conjunction with our studies of electromagnetic models of the junction have led us to the following conclusions: (1) An optical waveguide surrounding the junction occurs

naturally. (2) The waveguide arises from a higher refractive index in the junction plane. (3) The asymmetry in this dielectric waveguide is sufficiently small to allow propagation of discrete modes. (4) For parameter values characteristic of the GaP p-n junctions studied, only the lowest-order discrete mode of each polarization can propagate. (5) The product of the fractional optical dielectric constant increase and the width over which it extends is independent of bias. (6) The linear electro-optic (Pockels) effect creates a birefringent layer coincident with the junction depletion layer which expands with voltage. (7) Certain anomalies in the phase measurements are best explained as arising from a small quadratic electro-optic (Kerr) effect, though the agreement falls short of a proof of its presence.

Two aspects of this problem need further work. One is the fabrication of diodes at least 1 cm in length from optically and electrically homogeneous crystals. By so doing the intensity of continuum modes, which arise from the satisfaction of boundary conditions at the input surface, will be greatly reduced at the output surface. This should improve the accuracy of both intensity and phase measurements and allow more definitive studies of secondary characteristics such as the Kerr-effect contribution. The second aspect needing further attention is the understanding of the mechanism which creates the higher refractive index inside the p-n junction. The present situation regarding this was summarized at the end of Sec. V.

ACKNOWLEDGMENTS

We wish to thank F. A. Trumbore and C. J. Frosch for growing the crystals, E. Dickten and P. W. Foy for fabricating the diodes, H. F. Tiersten for calculating the effect of the clamping of the junction layer by the bulk crystal, K. F. Rodgers and J. J. Schott for the assistance in measurements, and J. K. Galt, R. C. Miller, and C. D. Thurmond for useful suggestions.

PHYSICAL REVIEW VOLUME 177, NUMBER 3 15 JANUARY 1969

Determination of the Gradient-Elastic Tensors for $A^{III}B^{V}$ Compounds Using Nuclear Acoustic Resonance*

R. K. SUNDFORS
Arthur Holly Compton Laboratory of Physics, Washington University, St. Louis, Missouri 63130
(Received 8 July 1968)

Acoustic waves in crystalline solids can couple energy to nuclear spin systems via the nuclear electric quadrupole interaction. For a known quadrupole moment, measurement of the acoustic attenuation due to this interaction allows the determination of the components of the fourth-order tensor connecting electric field gradients to elastic strains. By the use of nuclear acoustic resonance at 300°K, the magnitudes and relative signs of the gradient-elastic tensor components have been determined for type $A^{III}B^{V}$ semiconducting single crystals at the A and B nuclear positions in InAs, InSb, GaAs, and GaSb and at the B nuclear position in AlSb. The experimental error is $\pm 6\%$ in the measured product of quadrupole moment and gradient-elastic tensor component. Ionic contributions to the electric field gradients alone do not explain the relative signs and magnitudes of the components. A physical model is proposed to separate the ionic and covalent contributions to the measured gradient-elastic tensor components. This model predicts the ratio of the ionic contributions to the tensor components at the same nuclear positions in two different compounds and in the same compound. Such a model can be used to compare the several sets of effective ionic charges that have been proposed for these compounds. The ionic contributions found from this model have magnitudes that can be explained by antishielding coefficients smaller than 220. The covalent contributions at A and B nuclear positions have different magnitudes and different signs.

I. INTRODUCTION

ELASTIC strains have appreciable interaction with the nuclear spin systems in the $A^{III}B^{V}$ semiconducting compounds. In the presence of strains, the cubic symmetry at A and B nuclear positions of the unstrained perfect crystal is destroyed and electric field gradients are created. The electric field gradients interact via the large nuclear electric quadrupole moments with the nuclear spin systems. The over-all interaction between elastic strains and the nuclear spin systems can then be considered to consist of two parts: (i) the coupling between elastic strains and electric field gradients, and (ii) the nuclear electric quadrupole interaction. Our lack of knowledge about (i) is conveniently represented by a fourth-order tensor[1] $\mathbf{S}$ connecting the elastic strain tensor $\boldsymbol{\epsilon}$ to the electric field gradient tensor,

$$\nabla_i\nabla_j V = V_{i,j} = \sum_{k,l} S_{ijkl}\epsilon_{kl} \quad (i,j,k,l = x,y,z). \qquad (1)$$

*This work supported by the National Science Foundation under NSF Grant No. GP-6525.

[1] E. F. Taylor and N. Bloembergen, Phys. Rev. **113**, 431 (1959); E. F. Taylor, Ph.D. thesis, Harvard University, 1958 (unpublished).