

A theoretical model of the linear electro-optic effect

Chun-Ching Shih† and Amnon Yariv

Steele Laboratory, California Institute of Technology, Pasadena, CA 91125, USA

Received 18 May 1981

Abstract. A simple theoretical study of the linear electro-optic effect is presented. This semiclassical approach is based on the single-energy-gap model, the dielectric theory and the concepts of bond charge and effective ionic charge. A general expression is obtained for the electro-optic coefficient of a crystal and is applied to a wide variety of diatomic and ternary compounds including zincblende (GaAs, GaP, ZnSe, ZnS, ZnTe, CuCl), wurtzite (ZnS, CdS, CdSe), quartz (SiO₂), lithium niobate (LiNbO₃, LiTaO₃), KDP (KH₂PO₄, KD₂PO₄, NH₄H₂PO₄), chalcopyrite (AgGaS₂, CuGaS₂) and proustite (Ag₃As₃). The calculated results are generally in good agreement with experiment.

1. Introduction

In the weak-field approximation, the induced polarisation in matter is linearly related to the applied electric field as (Yariv 1975)

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} \quad (1)$$

where ϵ_0 is the electric permittivity in vacuum and χ is the susceptibility of the medium. \mathbf{P} and \mathbf{E} are vectors, so χ is a tensor of rank two. Due to the complex interaction between the electric field and matter, the susceptibility χ is, in general, field dependent. Such a nonlinearity becomes manifest at large field amplitudes. The polarisation can be expressed as an expansion in terms of the total field in the medium:

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} + \epsilon_0 \mathbf{r} \mathbf{E} \mathbf{E} + \epsilon_0 \mathbf{R} \mathbf{E} \mathbf{E} \mathbf{E} + \dots \quad (2)$$

The coefficients \mathbf{r} and \mathbf{R} relate \mathbf{P} and \mathbf{E} to second and third orders. \mathbf{r} is a tensor of rank three, and \mathbf{R} is a tensor of rank four. If the medium has inversion symmetry, it follows that the second-order tensor \mathbf{r} is identically zero. The lowest nonlinear effect in such a system is the third order. This happens in gas systems or in a centrosymmetric crystal such as rock salt or caesium chloride.

We can take the total field as

$$E_T = (E^\omega e^{i\omega t} + E^\Omega e^{i\Omega t} + \text{cc}) \quad (3)$$

where E^ω represents the amplitude of the optical electric field of frequency ω , E^Ω is the low-frequency electric field amplitude, $\omega \gg \Omega$. The term indicated by

$$P^{\omega+\Omega} = \epsilon_0 = 2\epsilon_0 \mathbf{r} E^\Omega E^\omega \quad (4)$$

† Present address: TRW, Plasma Physics Department, One Space Park, Redondo Beach, CA 90278, USA.

represents the electro-optic effect. Physically, the applied static field E^Ω can be viewed as causing a change of the optical susceptibility of the medium, and consequently, the index of refraction at a frequency Ω . This causes the optical field to acquire sidebands at $\omega \pm n\Omega$. The large interest in this physical property is due in part to its wide application in light modulators and numerous laser devices.

In many applications it is necessary to modulate the amplitude, phase, frequency or direction of a laser beam at high speed. The electro-optic effect is found to be an excellent method for performing these tasks due to its fast response and accurate control. The modulating low-frequency signal $E^\Omega e^{i\Omega t}$ is applied to the crystal through which the laser beam passes. With a choice of specific orientations of the crystal, we can modulate the amplitude, phase, frequency or direction of the incident beam. An excellent review article discussing these techniques in detail can be found in the paper by Kaminow and Turner (1966). Common to all the applications listed above is the need for crystals with high electro-optic constant r , so that the modulation power can be minimised. Many efforts have been made in this direction. These, however, were based on empirical approaches. It was concluded (Kaminow and Turner 1966) that 'perhaps the development of a theoretical understanding of the electro-optic effect will lead to the discovery of synthesis of the ideal substance for each application in a logical way . . .'. This is the motivation behind the investigation reported here.

Theoretical attempts to understand the electro-optic effect began in the late 1960s. Kaminow derived a simple relation between an electro-optic coefficient measured at radio frequencies and the corresponding Raman-scattering efficiencies (Kaminow 1967). Using the measured efficiencies, he calculated the electro-optic coefficients of LiNbO_3 and LiTaO_3 , which are in good agreement with experiment (Kaminow and Johnston 1967). With the macroscopic equations between polarisation, ion displacement and electric field, Kelly (1966) found an expression for the coefficient of zincblende crystals. The result was applied to CuCl and ZnS with satisfactory agreement with measurements. The electrostatic point-charge model and dielectric theory have been used to determine the electro-optic coefficients of III-V compounds (Flytzanis 1969) and II-VI crystals (Flytzanis 1971). The theoretical treatments are so far limited to diatomic crystals. The measurement of the electro-optic coefficient reveals that materials with large coefficients are actually complex crystals with more than two kinds of atoms. Therefore, a generalised theory to describe the electro-optic behaviour of simple crystals as well as complex systems is needed. A complete review of the formal theory of the nonlinear optical effect can be found in the paper by Flytzanis (1975). In this paper a simple model is derived for calculating the electro-optic coefficient of crystals with various structures and constituents (Shih and Yariv 1980).

Based on the quantum-mechanical approach of the one-gap model, the susceptibility of a diatomic crystal is found to depend on the energy gap. The energy gap is then phenomenologically interpreted as the combination of symmetric and asymmetric parts. From the microscopic point of view, these two are the result of the motion of bond-charge in the bond region. Its harmonic and anharmonic motions respond to the total field in the crystal, inducing the linear and nonlinear susceptibilities. In the low-frequency regions the ions are displaced from their normal sites due to the electric field applied. From the expression derived for the dependence of the optical susceptibility on bond-rotation and bond-stretch, we derive an expression for the electro-optic coefficient of crystals. The result is, hopefully, applicable to crystals with arbitrary complex structure. The theory is reviewed briefly in § 2.

In §§ 3-8 the theory is applied to the calculation of the coefficient of various crystals.

They include zincblende and wurtzite crystals, quartz, LiNbO_3 and LiTaO_3 , KDP family, chalcopyrite compounds, and Ag_3AsS_3 .

In § 9, the theory and calculation are summarised. A possible direction for seeking better electro-optic materials is pointed out. The limitation and the possible development of the theory are discussed.

2. Theory

The starting points for our semiclassical analysis are the dielectric description of diatomic crystals in the Phillips–Van Vechten (PV) theory (Phillips 1968a, b, Van Vechten 1969, Phillips and Van Vechten 1969) and the bond-charge calculation of the bond nonlinearity in the theory of Levine (1969, 1973a, b, c). With the concept of the effective ionic charge and the isotropic displacement, a general expression for the electro-optic coefficient can be obtained. Due to the uncertainty in the measurement of coefficients, a large enough margin is allowed for the accuracy of the theory in order to achieve its generality. However, the uncertainty in the theoretical estimate is usually comparable with or less than that in the experiment.

To meet the purpose of predicting the properties of new materials, the theory is required to employ as few physical parameters as possible. In order to serve as a guide to the crystal grower, it is important that the parameters entering the theory can be measured on small crystals or powder. This will obviate the need for expensive and lengthy growth of crystals often to find that the coefficients are disappointingly small.

The interaction between an external electric field and the electrons in a solid is responsible for its index of refraction. Using the nearly-free electron model in the semiconductor, it is found that the dielectric constant at long wavelengths can be calculated from (Penn 1962)

$$\epsilon(\infty) = 1 + (\hbar\omega_p/E_g)^2 [1 - (E_g/4E_F) + \frac{1}{3}(E_g/4E_F)^2] \quad (5)$$

where $\omega_p^2 \equiv 4\pi Ne^2/m$ is the plasma frequency. E_F is the Fermi energy level and E_g is the average energy gap of the semiconductor. $E_g/4E_F$ is usually about 0.1. So the susceptibility in this model has a very simple expression:

$$\chi(\infty) \simeq (\hbar\omega_p/E_g)^2. \quad (6)$$

The effective band gap E_g is given by

$$E_g^2 = E_h^2 + C^2. \quad (7)$$

It is clear that E_h is the symmetric part of the energy gap relating to the covalent bonding, while C is the asymmetric part relating to the ionic bonding. In diamond-like crystals, the asymmetric part vanishes and E_g is identically equal to E_h .

The expressions for E_h and C are given according to the PV theory by:

$$E_h = 39.74 d_0^{-s} \text{ (eV)} \quad d_0 \text{ is in \AA ngstroms, } s = 2.48 \quad (8)$$

$$C = be^2[(Z_\alpha/r_\alpha) - (Z_\beta/r_\beta)] \exp(-k_s R) \quad (9)$$

where $R = \frac{1}{2}d_0$ is the interatomic distance, b is a dimensionless constant depending on the row of atoms α and β , r_α is the atomic radius, Z_α is the core charge, and $\exp(-k_s R)$

is the Thomas–Fermi factor (Kittel 1976)

$$k_s^2 = \frac{4}{\pi a_0} (3\pi^2 N)^{1/3} \quad (10)$$

where a_0 is the Bohr radius and N is the electron density corresponding to eight electrons per diatomic volume.

In this theory the covalency and ionicity parameters of the bond are defined as

$$f_c = E_h^2/E_g^2 \quad (11)$$

$$f_i = C^2/E_g^2. \quad (12)$$

A complete list of values of C , E_h , f_c , and f_i for 68 diatomic crystals is given by Van Vechten (1969a, b).

The main reason for the appeal of the PV theory is that it depends on only two physical quantities: atomic radius and coordination number. We have avoided the complex calculation of the electron wavefunction by replacing the versatile band structure with a simple energy gap. The simplicity of this semiempirical result makes it a suitable starting point for the theory of the nonlinear optical effect. However, the division of the potential into symmetric and antisymmetric parts applies only to diatomic crystals. Any extension of the PV theory should be limited to diatomic crystals. In order to extend our theory to more complex materials we look into the unit cell from the microscopic point of view. It is assumed that the bulk susceptibility is due to the geometrical superposition of the susceptibilities of individual bonds: or equivalently, that the total dipole moment can be taken as the vector sum over a unit cell of individual dipoles, each associated with a single bond:

$$\chi_{ij} = \frac{1}{V} \sum_n \beta_n \alpha_{ni} \alpha_{nj} \quad (13)$$

where α_{ni} is the direction cosine of the bond n in the i th direction, β_n is the bond susceptibility along the bond direction, V is the volume of a unit cell, n indicates the individual bond, and the summation runs through the bonds in one unit cell. In the formulation of (13), we assume that the dipole is induced only along the bond direction, and we neglect the transverse contribution. Usually, the transverse susceptibility is much smaller than the longitudinal one, since it involves promotion of the electron into antibonding orbitals with high energies, so expression (13) is usually a good approximation. In diatomic crystals, there exists only one type of bond:

$$\chi_{ij} = \frac{\beta}{V} \sum_n \alpha_{ni} \alpha_{nj}. \quad (14)$$

It can be seen that the bond susceptibility is proportional to the bulk susceptibility in crystals of the same structure. We assume that relation (6) gives the bond susceptibility except for a proportionality constant depending on the crystal structure. The quantities E_h and C are fundamental properties of the bond. With this concept, we can calculate the linear and nonlinear susceptibility in an extended PV theory.

The physical interpretation of the E_h and C is based on the bond-charge model (Levine 1973b, c). It is thought that the overlap of the electron distribution in two adjacent atoms generates a bond charge in the bond region. It represents only a small amount of charge, but it possesses high mobility. Its harmonic and anharmonic motion responding to the applied total field is the source of linear and nonlinear dipole moment.

Although the calculation of E_h and C is insensitive to the individual atomic radius and depends on the bond length only, the anharmonic motion of the bond charge should be sensitive to its own position.

The expression for C given in equation (9) should be good for the nonlinear effect because it already contains the dependence on r_α and r_β and is thus sensitive to the motion of the bond charge. The homopolar energy gap E_h given in equation (8) is only a function of $r_\alpha + r_\beta$. This is approximately true for the harmonic motion of the bond charge, but it may not be correct for the anharmonic motion, especially for the highly unequal atomic radii. Considering the small contribution of core electrons to the bond susceptibility, it is proposed that a generalised homopolar part of the energy gap is (Levine 1973b, c)

$$E_h^{-2} = (E_h^{-2})_0 \frac{(r_\alpha - r_c)^{2s} + (r_\beta - r_c)^{2s}}{2(r_0 - r_c)^{2s}} \quad (15)$$

where $(E_h^{-2})_0$ is the homopolar gap when $r_\alpha = r_\beta = r_0$, r_c is the average core radius.

Using the assumption of the geometric superposition of the individual bond dipoles, the change of the macroscopic susceptibility is expressed in terms of the changes of the bond susceptibility and direction cosines:

$$\Delta\chi_{ij} = \frac{1}{V} \left[\sum (\Delta\beta_n) \alpha_{ni} \alpha_{nj} + \sum \beta_n (\Delta\alpha_{ni}) \alpha_{nj} + \sum \beta_n \alpha_{ni} (\Delta\alpha_{nj}) \right]. \quad (16)$$

The changes of direction cosines are due to the relative displacement of atoms. Assuming that an electric field applied in the k direction induces a displacement Δx_k , we have

$$\begin{aligned} (\Delta\alpha_i)_k &= \frac{\partial}{\partial x_k} (\alpha_i) \Delta x_k \\ &= \frac{\partial}{\partial x_k} (x_i/d_0) \Delta x_k \\ &= \frac{1}{d_0} [\delta_{ik} - \alpha_i \alpha_k] \Delta x_k \end{aligned} \quad (17)$$

where d_0 is the bond length, $d_0 = 2r_0$.

The change $\Delta\beta_n$ of the bond susceptibility is due to the displacement of the bond charge and the stretching of the bond. As $\beta \propto (\hbar\omega_p)^2/E_g^2$, where the small contribution from D and A in (11) has been neglected, we have:

$$\Delta\beta/\beta = [\Delta(\omega_p^2/\omega_p^2)] + f_c E_h^2 \Delta(E_h^{-2}) - 2f_i (\Delta C/C). \quad (18)$$

The variations of ω_p^2 , E_h^{-2} , and C depend on d_0 , r_α , and r_β , so we have to relate these quantities to the known physical parameters. When the bond length changes, it is assumed that the ratio of atomic radii remains constant. The two independent parameters, r_α and r_β , can be transformed into two parameters relating directly to the macroscopic properties of the crystal:

$$\begin{aligned} \Delta r_\alpha &= (r_\alpha/d_0) \Delta d_0 + \delta \\ \Delta r_\beta &= (r_\beta/d_0) \Delta d_0 - \delta \end{aligned}$$

where δ is the displacement of the bond charge independent of the ionic motion, and δ corresponds to the optical susceptibility, while Δd_0 corresponds to the low-frequency dielectric constant.

From equations (9), (15), (18) and (19) it is a straightforward procedure to obtain:

$$\Delta\beta/\beta = \{ [f_i(1 + \frac{1}{2}k_s r_0) + sf_c - \frac{3}{2}] \Delta d_0/r_0 + [4f(Z_\alpha + Z_\beta)/(Z_\alpha - Z_\beta) + s(2s - 1) f_c \rho d_0^2/(r_0 - r_c)^2] (\delta/d_0) \} \tag{20}$$

where $\rho = (r_\alpha - r_\beta)/(r_\alpha + r_\beta)$. In the first square bracket, the term $\frac{1}{2}k_s r_0$ is obtained because the screening wavenumber k_s is proportional to $r_0^{-1/2}$. The number $(-\frac{3}{2})$ is due to the fact that ω_p is proportional to $r_0^{-3/2}$. The expression in the second square bracket is the electronic contribution from the bond-charge response. It is identical to the result obtained by Levine (1973b, c) in his calculation of the nonlinear optical susceptibility and will not be considered here. Thus the expression in the first square bracket represents the ionic contribution of a single bond due to the bond stretching.

Substituting equations (17) and (20) into (16), the ionic contribution to $(\Delta\chi_{ij})_k$ is

$$\Delta\chi_{ijk}^{ion} = \left\{ \frac{1}{V} \sum_n (\beta_n/r_0) [f\alpha_{ni}\alpha_{nj}\alpha_{nk} + \frac{1}{2}(\alpha_{ni}\delta_{jk} + \alpha_{nj}\delta_{ik})] \right\} \Delta x_k \tag{21}$$

where

$$f = f_i(1 + \frac{1}{2}k_s r_0) + sf_c - 2.5 = (k_s r_0/2 - 1.48) f_i - 0.02 \tag{22}$$

f is the ionicity factor and δ_{jk} is the Kronecker delta function. The relative displacement of atoms Δx_k is related to the dielectric constant of the crystal as:

$$Ne_c^* \Delta x_k = \epsilon_0(\epsilon'_{dc} - \epsilon'_{\infty k}) E_k^i \tag{23}$$

where N is the number of pairs of atoms per unit cell, e_c^* is the Callen (1949) effective ionic charge, ϵ'_{dc} is the relative dielectric constant, ϵ'_{∞} is the relative optical permittivity, and E_k^i is the low-frequency electric field in the k direction. The effective charge e_c^* is related to the Szigeti effective charge e_s^* as (Callen 1949)

$$e_c^* = e_s^*(\epsilon'_{\infty} + 2)/3\epsilon'_{\infty} \tag{24}$$

Furthermore, the Szigeti effective charge has been found empirically to be equal to $(c/\hbar\omega_p)$ in diatomic crystals (Lawaetz 1971). If we adopt this relation for more complex crystals, it becomes possible to calculate the effective charge from the knowledge of the atomic radius and structure. The only physical quantity which we cannot calculate is the static dielectric constant ϵ'_{dc} . In this paper we use the experimental data of ϵ'_{dc} to calculate the electro-optic coefficient.

Next we will relate the change of the susceptibility to the electro-optic coefficient. It is conventional to define the coefficient in terms of the change of $1/n_{ij}^2$, i.e.

$$\Delta(1/n^2)_{ij} = r_{ijk} E_k \tag{25}$$

The matrix form for r must reflect the crystal symmetry. Highly symmetric crystals possess fewer independent nonzero elements. A complete list of the matrix form for various point groups can be found in the book by Yariv (1975).

The coefficient r_{ijk} is related to the change in optical susceptibility by (Yariv 1975)

$$r_{ijk} E_k^s = -\Delta\chi_{ijk}/\epsilon_i^s \epsilon_j^t \tag{26}$$

Combining equations (21), (23), and (26), the final expression of the electro-optic coefficient is

$$r_{ijk} = \frac{\epsilon_0(\epsilon'_{dc_k} - \epsilon'_{\infty_k})}{VNe_c^* \epsilon'_i \epsilon'_j} \sum_n \frac{\beta_n}{r_0} [f\alpha_{ni}\alpha_{nj}\alpha_{nk} + \frac{1}{2}(\alpha_{ni}\delta_{jk} + \alpha_{nj}\delta_{ik})]. \quad (27)$$

As the direction cosines appear only as α_i or $\alpha_i\alpha_j\alpha_k$, it is apparent that $r_{ijk} = 0$ if the crystal has inversion symmetry. Usually, $\alpha_i\alpha_j\alpha_k$ should be an order of magnitude smaller than α_i if $\sum\alpha_i \neq 0$. The ionicity factor has been calculated for many crystals. Its value is about 0.1–0.2 for diatomic crystals and never exceeds 0.3 for most complex crystals.

We have established a theoretical model and derived an expression for the electro-optic coefficients of crystals. In the following sections, the theory will be applied to crystals with two or three atoms per formula. First we consider zincblende and wurtzite crystals which are the basic lattice structures considered in the PV dielectric theory. The crystal quartz is chosen to extend the theory beyond the AB type.

In some of the ternary compounds, not all the bonds have to be taken into account in the calculation. Lithium niobate and lithium tantalate are used to demonstrate the negligible contribution of the highly ionic bond like Li–O. It is also shown that the distorted octahedron which results in two different bond lengths for Nb–O is responsible for the nonlinear effect. The KH_2PO_4 (KDP) family is interesting in its nonlinear properties. The calculation shows the dominant role played by P–O bonds. In order to investigate materials with a wider range of transparency in the infrared region, we calculate the coefficient of chalcopyrite crystals which have a unit cell structure evolved from the zincblende crystal. The last ternary crystal we consider is proustite, whose complex structure consists of 54 bonds per unit cell.

In the remainder of this paper we calculate the ionic part of the electro-optic coefficient. The electronic contribution is obtained from the measured second harmonic generator coefficients. The sum of these two parts is compared with the measured value. In general, our calculated values are in good agreement with those obtained experimentally.

3. Diatomic crystals—zincblende and wurtzite

The diatomic crystals can be divided into four groups, according to their structures: zincblende, wurtzite, rock salt, and CsCl type. The crystals with rock salt and CsCl structures possess inversion symmetry and, consequently, do not have a linear electro-optic effect. Zincblende crystals are cubic, while wurtzite crystals are hexagonal, but both are in tetragonal coordination. The crystal structure of the two types is identical if one is limited to the nearest neighbours. They can be distinguished only by comparing the position of the next-nearest neighbours.

If the slight distortion from the perfect tetragonal structure in wurtzite (Lawaetz 1972) is neglected, the coefficients of both types of crystals can be calculated in a similar way. In table 1, the parameters used in equation (27) are presented. For the calculation of $\sum\alpha_i\alpha_j\alpha_k$ we have chosen the coordinates such that one of the four bonds points in the direction of (111) for zincblende and in the +z direction for wurtzite. The sense of the positive polarisation of the bond is defined as the direction from the positive ion to the negative ion. Those definitions are in compliance with the conventional choices used in experimental measurements. With the quantities in table 1, the electro-optic coefficient

Table 1. Parameters for zincblende and wurtzite crystals.

	Zincblende	Wurtzite
Number of atoms in a unit cell	4	2
Volume of a unit cell	a_0^3	$\frac{1}{2}\sqrt{3}a_0^2c_0$
Number of atoms per unit volume	$4/a_0^3$	$4/\sqrt{3}a_0^2c_0$
$\Sigma \alpha_i$	0	0
$\Sigma \alpha_i^2$	$\frac{1}{3}$	$\frac{1}{3}$
$\Sigma \alpha_1\alpha_2\alpha_3$	$16/3\sqrt{3}$	0
$\Sigma \alpha_1^2\alpha_3 = \Sigma \alpha_2^2\alpha_3$	0	$-\frac{1}{3}$
$\Sigma \alpha_3^3$	0	$\frac{2}{3}$

of diatomic crystals is found to be:

$$\text{zincblende} \quad r_{14} = 0.3689 \frac{a_0^2 w f}{e_c^* / e} \tag{28}$$

$$\text{wurtzite} \quad r_{33} = -2r_{13} = 0.4260 \frac{a_{\text{eff}}^2 w f}{e_c^* / e} \tag{29}$$

$$w = (\epsilon' - 1) (\epsilon'_{\text{dc}} - \epsilon') / \epsilon'^2$$

and other coefficients are identically zero due to the symmetric property of the crystals. In expressions (28) and (29), r is in units of $10^{-12} \text{ m V}^{-1}$, while a_0 and a_{eff} are in units of \AA .

We have calculated the electro-optic coefficients of nine diatomic crystals using formulae (28) and (29). The results for the zincblende (GaAs, GaP, ZnSe, ZnS, ZnTe, CuCl) and wurtzite (ZnS, CdS, CdSe) are compared with experimental data in table 2.

Table 2. Parameters and results of equations (28) and (29). $a = a_0$ or a_{eff} . r 's represent r_{14} (zincblende) and r_{33} (wurtzite) and are in units of $10^{-12} \text{ m V}^{-1}$. $r_{\text{expt'l}}$ are measurements with clamped crystals. Their signs are not yet determined, unless so specified.

AB	Zincblende						Wurtzite		
	GaAs	GaP	ZnSe	ZnS	ZnTe	CuCl	ZnS	CdS	CdSe
a^a	5.65	5.45	5.67	5.41	6.09	5.41	5.39	5.85	6.08
ϵ'_{dc}	13.2 ^b	12.0 ^c	9.1 ^d	8.3 ^d	10.1 ^d	7.5 ^e	8.7 ^f	9.4 ^d	10.2 ^d
w	0.192	0.284	0.450	0.528	0.331	0.656	0.567	0.652	0.562
f_i^g	0.310	0.370	0.630	0.623	0.546	0.749	0.623	0.683	0.699
f	-0.091	-0.113	-0.163	-0.179	-0.119	-0.212	-0.181	-0.162	-0.147
e_c^* / e	0.20	0.23	0.33	0.35	0.26	0.27	0.35	0.41	0.36
r_{ionic}	+1.03	+1.53	+2.64	+2.93	+2.07	-5.56	+3.63	+3.75	+3.61
r_{elec}	-2.73 ^h	-3.20 ⁱ	-4.68 ^j	-4.77 ^k	-6.41 ^l	+2.66 ^m	-5.63 ^k	-6.71 ^k	-7.40 ^j
$r_{\text{sum}}^{\text{theo}}$	-1.7	-1.7	-2.0	-1.8	-4.3	-2.9	-2.0	-3.0	-3.8
$r_{\text{expt'l}}$	-1.6 ⁿ	-1.1 ⁿ	2.0 ^p	1.6 ^p	4.3 ^p	-2.4 ^p	1.8 ^p	3.0 ^p	4.3 ^p

^a Wyckoff (1963); ^b Jones and Mao (1968); ^c Kaminow and Turner (1966); ^d Berlincourt *et al* (1963); ^e Alomas *et al* (1968); ^f Kobayakov (1966); ^g Van Vechten (1969a, b); ^h Mooradian and McWhorter (1969); ⁱ Wynne and Bloembergen (1969); ^j Soref and Moos (1964); ^k Patel (1966); ^l Chang *et al* (1965); ^m Chemla *et al* (1971); ⁿ Nelson and Turner (1968); ^p Kaminow and Turner (1971).

Values of the parameters a_0 , a_{eff} , w , f_i , f , e_c^* are listed in the table and discussed in the following.

In table 2, a represents a_0 for zincblende, and $a_{\text{eff}} = (\sqrt{3}a_0^2c_0)^{1/3}$ for wurtzite. The values of dielectric constants, ϵ'_{dc} , are taken from the experimental data. We note that the crystals with a higher number of valence electrons (n_v) have larger ϵ'_{dc} . This behaviour is probably not related to the value of n_v directly, because it can be seen that the effective ionic charge is actually smaller for higher n_v crystals. It should be the result of the bond rigidity. The crystal with lower n_v is more ionic and has stronger bonds. The distortion of the crystal in response to an applied field is smaller, resulting in smaller ϵ'_{dc} . This argument is also applied to crystals with the same n_v . For example, $\text{ZnTe} > \text{ZnSe} > \text{ZnS}$ because the atomic radius and the covalency follow the sequence $\text{Te} > \text{Se} > \text{S}$. So the relative displacement is larger for ZnTe , resulting in a higher ϵ'_{dc} .

The ionicity factor is a very important parameter which is calculated from the screening wavenumber and the bond ionicity. It is involved (see equation (27)) in the magnitude of the contribution from the summation of the triple product of direction cosines. The value of f in diatomic crystals ranges from 0.1 to about 0.2. Because it is such a small value, the accuracy of the estimate of s becomes very important. However, f is in general a negative quantity.

The effective ionic charge is calculated from the Szigetzi effective charge. It is a very small value due to the local field correction factor. The ionic contribution of the electro-optic coefficient is obtained by using equations (28) and (29). The uncertainty of r_{ionic} is due mostly to two sources. One is the uncertainty in the measured values of the physical parameters. The other is in the assumptions and approximations of the theory. Overall, the uncertainty of r_{ionic} should be about 15–20%, which is usually also the standard deviation of the measurement.

The purely electronic contribution is obtained from the coefficient of the second harmonic generation (SHG) using the relation (Yariv 1975)

$$r_{ijk} = -4d_{ijk}/\epsilon_i\epsilon_j. \quad (30)$$

We have assumed that the coefficient d_{ijk} has no dispersion in the frequency. As r_{ijk} is obtained in the limit of long wavelength, d_{ijk} should be the value measured at long wavelengths where the nonlinear effect and linear effect are less dispersive. The value of r_{elec} seems to have no obvious relation to the number of valence electrons.

The predicted coefficient is the sum of the theoretical calculation r_{ionic} and the experimental data r_{elec} . The predicted value is compared with the measurements. r_{exptl} is obtained with clamped crystals where the strain-induced effect can be neglected.

The comparison between $r_{\text{sum}}^{\text{theo}}$ and r_{exptl} shows that the prediction is in good agreement with experiment including the sign. The worst discrepancy is in GaP, but it is interesting to note that the electronic contribution is about double the ionic contribution. This is in good agreement with the observation from Raman scattering (Faust and Henry 1966). Another interesting example is CuCl: it is found that r_{elec} of CuCl is positive and r_{exptl} is negative, which implies that r_{ionic} must be negative and large. If we assign a negative sign to our calculated result, the value of $r_{\text{sum}}^{\text{theo}}$ is close to what we expect it should be. The question arises as to why CuCl should possess a different sign for r_{ionic} compared with the other crystals. This anomalous behaviour has been explained by considering the d-electron contribution (Kaminow and Turner 1964). The d electrons in Cu atoms are bound loosely and can be considered as valence electrons. Since the Cu atoms have more valence electrons than the Cl atoms have, the Cu ions become more positive than the Cl ions. The situation is completely different for GaAs, where the Ga ions are more

negative than the As ions because the Ga atoms have fewer valence electrons. Therefore, when a low-frequency field is applied, it is very possible that the relative displacement in CuCl is in the opposite direction compared with what happens in GaAs. However, this argument can only be fully supported by a direct measurement of the relative displacement of ions in these crystals. The displacement is small ($\sim 0.01 \text{ \AA}$), but it might be detectable by the new technique of modulation x-ray diffraction (Fujimoto 1978).

It should not be surprising that agreement of the theory with experiment is satisfactory, since it is based on a theory describing the bond properties of diatomic crystals. However, the dc dielectric constant is not much larger than ϵ'_∞ . A small error in ϵ'_{dc} could lead to a large discrepancy in the value of w , which depends on $(\epsilon'_{dc} - \epsilon'_\infty)$. In other complex crystals with very large values of ϵ_{dc} such critical dependence disappears. The prediction becomes more reliable.

4. Diatomic crystal-quartz

The interest in quartz is due mostly to its availability in nature in large crystals of excellent optical properties. The transparency region extends from UV ($\sim 1800 \text{ \AA}$) to the infrared ($\sim 7 \text{ \mu m}$).

At room temperature, quartz (SiO_2) has the α structure. Its point group is 32 with three molecules in a unit cell. Although quartz has only two kinds of atoms, it is completely different from the zinblende and wurtzite structures. For the diatomic crystals of AB_n type, it has been suggested that the heteropolar part of the energy gap is replaced by the expression (Levine 1973a)

$$c = be^2(Z_\alpha/r_\alpha - nZ_\beta/r_\beta) \exp(-k_s R) \quad (31)$$

where n is the number of atoms β per formula unit. The reason for including n is that the valence electron spends about n times the period around atom β as around atom α . The effective screened Coulomb potential is thus multiplied by n for atom β .

From the crystallographic data, we find the bond length is 1.61 \AA and the volume of a unit cell is 113 \AA^3 . Every silicon atom bonds to its four nearest oxygen atoms, while every oxygen atom bonds to two silicon atoms. There are twelve bonds in a unit cell which are all identical except in direction.

Although quartz possesses birefringence, the difference in the index of refraction is only about 1%. For convenience, we use the average value $\Sigma \alpha^2 = 4.0$ and $\Sigma \alpha_1^3 = -\Sigma \alpha_1 \alpha_2^3 = 1$.

From the summations of the direction cosines, the coefficient r_{11} which has been measured experimentally is derived from (27):

$$r_{11} = \frac{\epsilon_0}{N e_c^*} \frac{\chi(\epsilon'_{dc} - \epsilon') f}{\epsilon'^2 r_0} \Sigma \alpha_1^3 / \Sigma \alpha^2 \quad (32)$$

For demonstration and clarity all physical quantities appearing in the calculation are presented in table 3. The prediction is compared with the measurement. The excellent agreement should not be over-emphasised. The uncertainty in r_{11}^{expt} is 10% and in r_{11}^{ion} is even higher. In table 3 two quantities having abnormal values should be discussed. The value f is much higher than the value in zinblende and wurtzite, and e_c^* is also much larger. The higher values of f and e_c^* are actually the result of a lower valence electron density. In a comparable size of the unit cell, quartz has twelve bonds, while zinblende crystals have sixteen bonds, which affects the value of the plasma frequency.

Table 3. Calculation of r_{11} for quartz.

Symbol	Value	Remarks
v_b	9.4 \AA^3	Bond volume
n_b	2	Number of valence electrons per bond
n	0.21 \AA^{-3}	$n = n_b/v_b$
$\hbar\omega_p$	17.1 eV	$\hbar\omega_p(\text{eV}) = 37.16(\sqrt{n(\text{\AA}^{-3})})$
k_F	1.85 \AA^{-1}	$k_F = (3\pi^2n)^{1/3}$
k_S	2.10 \AA^{-1}	$k_S = (4k_F/\pi a_B)^{1/2}$
E_h	12.2 eV	Homopolar energy gap ^a
C	14.1 eV	Heteropolar energy gap ^a
E_g	18.6 eV	$E_g^2 = E_h^2 + C^2$
f_i	0.57	$f_i = C^2/E_g^2$
f	-0.39	$f = (k_S r_0/2 - 1.48)f_i - 0.02$
e_s^*/e	0.82	$e_s^*/e = C/\hbar\omega_p$
e_c^*/e	0.50	$e_c^* = e_s^*(e' + 2)/3e'$
N	0.106 \AA^{-3}	Number of ionic charges per unit volume
ϵ'	2.4	Optical permittivity ^b
ϵ'_{dc}	4.5	Dielectric constant ^c
r_0	0.805 \AA	Half of bond length
r_{11}^{ion}	$0.61 \times 10^{-12} \text{ m V}^{-1}$	Use equation (27)
d_{11}	$0.4 \times 10^{-12} \text{ m V}^{-1}$	SHG coefficient at 1.06 \mu m^d
d'_{11}	$0.44 \times 10^{-12} \text{ m V}^{-1}$	Assume Miller index is constant at 0.633 \mu m
r_{11}^{elec}	$-0.32 \times 10^{-12} \text{ m V}^{-1}$	$r_{11}^{\text{elec}} = 4d_{11}/\epsilon^2$
r_{11}	$0.29 \times 10^{-12} \text{ m V}^{-1}$	$r_{11} = r_{11}^{\text{ion}} + r_{11}^{\text{elec}}$
r_{11}^{exptl}	$0.29 \times 10^{-12} \text{ m V}^{-1}$	r_{11}^{exptl} at 0.633 \mu m^e

^a Levine (1973a); ^b Kaminow and Turner (1971); ^c Zubov *et al* (1963); ^d Miller (1964); ^e Rosner *et al* (1967).

The only experimental value used in this calculation is the dielectric constant ϵ'_{dc} . This value has been measured over a wide range of temperatures. It is found that in pure crystals ϵ'_{dc} is constant up to about 600°C, where the phase change to β modification takes place (Zubov *et al* 1965).

5. Lithium niobate and tantalate

Up to this point we have calculated the electro-optic coefficients of crystals with only two kinds of atoms and only one type of bond. However, the principle of the geometrical superposition of susceptibilities does not limit the applicability of our method to diatomic crystals. In the following we will extend the theory to complex crystals.

The first crystals we consider are lithium niobate (LiNbO₃) and lithium tantalate (LiTaO₃). These are used extensively in integrated and electro-optics. Large crystals (>1 cm) with good optical and electrical qualities are available. Due to their high transition temperature both crystals are easy to handle, for example, cut, polish, press, without creating additional domains.

At room temperature the structures of LiNbO₃ (Abrahams *et al* 1966) and LiTaO₃ (Abrahams *et al* 1967) are rhombohedral with point group 3m. Oxygen atoms form octahedral coordinations with the three-fold axis along the z direction. There are two formulae per unit cell. Two lithium ions and two niobium ions occupy four out of six octahedral sites in a unit cell. The positions of the lithium and niobium ions are distorted

from the centre of the octahedron due to the occurrence of an empty site for every three octahedral structures. The distortion is actually responsible for the second harmonic generation and linear electro-optic effect. The electro-optic tensor has the nonvanishing components: r_{33} , $r_{13} = r_{23}$, $r_{22} = -r_{12} = -r_{16}$, and $r_{42} = r_{51}$.

Bond lengths in both crystals are found to be:

$$\begin{aligned} \text{Ta-O} & \quad 1.891 \text{ \AA} \text{ and } 2.071 \text{ \AA} \\ \text{Li-O} & \quad 2.076 \text{ \AA} \text{ and } 2.293 \text{ \AA}. \end{aligned}$$

We can see that for each kind of bond there are two different bond lengths. Naturally, different bond lengths result in different bond susceptibilities. Rigorously speaking,

Table 4. The properties of bonds, Nb-O and Ta-O in LiNbO₃ and LiTaO₃—LiMO₃.

M-O	(Nb-O) _S	(Nb-O) _L	(Ga-O) _S	(Ta-O) _L
d_0	1.889 Å	2.112 Å	1.891 Å	2.070 Å
E_h	8.20 eV	6.22 eV	8.18 eV	6.53 eV
C	17.54 eV	13.75 eV	19.24 eV	15.75 eV
E_g	19.36 eV	15.09 eV	20.91 eV	17.05 eV
f_i	0.821	0.830	0.847	0.853
f	-0.292	-0.241	-0.282	-0.238
$\Sigma\alpha_3$	2.84913	-4.01420	3.0438	-3.8760
$\Sigma\alpha_3\alpha_1^2$	1.10334	-1.10871	1.1302	-1.1293
$\Sigma\alpha_3$	0.396	-0.195	0.410	-0.261
$\Sigma\alpha_3^3$	0.64244	-1.79678	0.7833	-1.6175
$\chi^H/\chi_{\text{avg}}^H$	0.921	1.085	0.9350	1.070
$f\Sigma\alpha_3\alpha_1^2 + \frac{1}{2}\Sigma\alpha_3$	1.10236	-1.7405	1.2032	-1.6698
$f\Sigma\alpha_3^3 + \Sigma\alpha_3$	2.66154	-3.5821	2.8229	-3.4918

there should be four different bonds in a unit cell. However, it has been found that the contribution of the Li-O bond to the linear and nonlinear susceptibility is negligibly small (Jeggo and Boyd 1970). We can neglect the susceptibility due to these bonds but cannot neglect their existence. Li-O bonds still have to be taken into account when calculating the bond volume, plasma frequency and screening factor. So we need only be concerned with the structured information of the Nb-O bonds and the Ta-O bonds. The results are presented in table 4. In the table we use the subscript S to indicate quantities for shorter bonds and L for longer bonds. The structure information is applied only to obtain $(f\Sigma\alpha_3\alpha_1^2 + \frac{1}{2}\Sigma\alpha_3)$ and $(f\Sigma\alpha_3^3 + \Sigma\alpha_3)$ which correspond to the coefficients r_{51} and r_{33} , respectively, since these two coefficients are large and of most interest for a 3m crystal. The reason why r_{51} and r_{33} are large becomes obvious in table 4. In these two crystals $\Sigma\alpha_3 \neq 0$. Furthermore, $\Sigma\alpha_3$ has about four times the value of $\Sigma\alpha_3^3$ and nearly three times the value of $\Sigma\alpha_3\alpha_1^2$. As f is a small value ($f < 0.3$), the term $\Sigma\alpha_3$ actually dominates over the term containing f , which involves the information on the electronic structure.

The results of calculations are shown in table 5. The ionic charge of niobium is found from the average C and $\hbar\omega_p$. However, the ionic charge of lithium is set equal to one, due to the high ionicity of the Li-O bond. If it is assumed that under the applied electric field the displacements of the niobium and lithium ions have the same magnitude and are in the same direction, the effective ionic charge per formula is obtained to be $1.8e$

for LiNbO_3 and $2.0e$ for LiTaO_3 . The low-frequency dielectric constants which are different in z and x or y directions are listed to calculate the factor $(\epsilon'_{dc_k} - \epsilon'_k)/\epsilon'_i \epsilon'_j$. The value found is the ionic contribution to the electro-optic coefficient and is entered as r^{ionic} . The electronic contribution is obtained from the coefficient of SHG (Miller and Savage 1966). The theoretical prediction r^{sum} is taken as the sum of these two values and is compared with the experimental measurement (Turner 1966). It is found that the prediction is in good agreement with experiment. The difference is less than 10%, which is well within the uncertainty of the theory and the measurement.

Table 5. Calculations of r_{15} and r_{33} for LiNbO_3 and LiTaO_3 (in units of $10^{-12} \text{ m V}^{-1}$).

	LiNbO_3	LiTaO_3
C	15.5 eV	17.4 eV
$h\omega_p$	28.6 eV	29.3 eV
e_s^*/e	3.7	4.0
e_c^*/e	1.8	2.0
$\epsilon'_{dc,3}$	28	43
$\epsilon'_{dc,1,2}$	43	41
$(\epsilon'_{dc1} - \epsilon'_1)/\epsilon'_1 \epsilon'_3$	1.5419	1.6393
$(\epsilon'_{dc3} - \epsilon'_3)/\epsilon_3^2$	1.0245	1.7210
r_{15}^{ionic}	+19.7	+16.5
r_{15}^{elec}	+0.8	+0.2
r_{15}^{sum}	+20.5	+16.7
$r_{33}^{\text{exp}^{\text{r}1}}$	+23	+15
r_{33}^{ionic}	+19.9	+27.8
r_{33}^{elec}	+6.0	+3.7
r_{33}^{sum}	+25.9	+31.5
$r_{33}^{\text{exp}^{\text{r}1}}$	+28	+30

6. The KDP family

Potassium dihydrogen phosphate (KDP) and ammonium dihydrogen phosphate (ADP) are among the best known of the nonlinear materials. They can be grown easily from a water solution with dimensions as large as several centimetres. The crystals are usually of good optical quality and can be cut or polished without difficulty. At room temperature, KDP and ADP are piezo-electric, and belong to the point group $\bar{4}2m$. Although the SHG has been observed (Van der Ziel and Bloembergen 1964) below the Curie temperature, no electro-optic measurements have been made at low temperatures where ADP is antiferroelectric and KDP is ferroelectric. Above T_c the only nonvanishing electro-optic coefficients are $r_{41} = r_{52}$ and r_{63} . The transparent region for the crystal is from $0.2 \mu\text{m}$ to about $2 \mu\text{m}$. Both the electro-optic coefficients and the index of refraction are almost constant over this range.

Without changing the crystal structure, other members of the KDP family are obtained by replacing K, H and P with some atoms from the corresponding columns in the periodic table or with some equivalent clusters, e.g., K can be replaced by NH_4 . So far, only five members of the family have had their dielectric constants determined. They are KH_2PO_4 (KDP), KD_2PO_4 (KDDP), KH_2AsO_4 (KDA), RbH_2AsO_4 (RDA), and $\text{NH}_4\text{H}_2\text{PO}_4$ (ADP). Therefore, we will apply the theory to these crystals and compare the results with experiment.

In KDP, both K and P are in tetragonal coordination with oxygen atoms. The hydrogen atom is situated about 0.21 Å from the midpoint of the line joining the oxygens. There exist three different bonds: K–O, H–O, and P–O. Using an argument similar to that used when discussing the Li–O bond in LiNbO₃, the contribution of the K–O bond to the linear and nonlinear susceptibility can be neglected due to its high ionicity. The H–O bond is covalent and contributes to the linear susceptibility. However, the almost isotropic distribution of H–O bonds makes the contributions to the nonlinear susceptibility cancel each other, so the only contribution to the electro-optic coefficient comes from P–O bonds. Normally, we use the principle of the geometrical decomposition of susceptibility to obtain the bond polarisability. In KDP, the participation of H–O bonds in χ prevents us from calculating the polarisability of P–O. In order to avoid such difficulty, we can assume that the polarisability of a bond is almost the same in two crystals if the environments of the bond are similar. In the crystals without hydrogen atoms, AlPO₄ is the best candidate to find the polarisability of the P–O bond because the P atom in KDP and AlPO₄ has the same coordination structure. The bond properties of P–O in AlPO₄, including C , E_h , f_i , have been obtained by decomposition (Levine 1973a). Since all members in the KDP family have the same structure, the properties and direction cosines of P–O bonds should not deviate much from crystal to crystal. With this argument, the electro-optic coefficient for any crystals of the KDP type is represented by the formula

$$r = \frac{\epsilon_0}{Ne_c^*} \frac{\chi(\text{P-O})}{r_0} \left(\frac{\sum \alpha_1 \alpha_2 \alpha_3}{\sum \alpha_2} \right) \frac{\epsilon'_{dc} - \epsilon'}{\epsilon'^2} f. \quad (33)$$

The values of parameters appearing in the formula are:

$$C(\text{P-O}) = 12.7 \text{ eV}$$

$$E_h(\text{P-O}) = 13.2 \text{ eV}$$

$$f_i(\text{P-O}) = 0.481$$

$$\hbar\omega_p(\text{P-O}) = 25.1 \text{ eV}$$

$$e_s^*/e(\text{P}) = 2.53$$

$$e_s^*/e \text{ per formula} = 5.8$$

$$e_c^*/e \text{ per formula} = 3.2$$

$$N = 0.010 \text{ formula } \text{\AA}^{-3}$$

$$\chi(\text{P-O}) = 0.85$$

$$r_0(\text{P-O}) = 0.78 \text{ \AA}$$

$$\sum \alpha_1 \alpha_2 \alpha_3 (\sum \alpha^2)^{-1} = 0.5 \quad (\text{Shay and Wernick 1975})$$

$$k_s r_0 = 1.92$$

$$f = -0.271.$$

Using these values, the coefficient is found to be

$$r = (\epsilon'_{dc} - \epsilon') \epsilon'^{-2} 2.47 \times 10^{-12} \text{ m V}^{-1}. \quad (34)$$

The electronic contribution is calculated from the SHG coefficient and ranges from 0.2 to 0.4. It is negligible compared with the total value which is about 10. Thus the electro-optic effect in KDP is essentially ionic. Since the r_{41} coefficient in most crystals has not been measured at high frequencies, we present only the calculation of r_{63} in table 6. The predictions are in good agreement with experiment. In the calculation, ϵ' is taken to be 2.3 because it is almost a constant for various crystals. We also assume that all positive ions, K, H and P, are displaced by the same distance under the applied electric field. Without this assumption, the calculation of the displacement of P atoms from the knowledge of ϵ'_{dc3} is impossible.

Table 6. The electro-optic coefficient r_{63} of KDP, KDDP, and ADP. The values corresponding to constant stress are noted by (T) and constant strain (clamped) by (S).

	ϵ'_{dc3}	$\frac{\epsilon'_{dc3} - \epsilon'}{\epsilon'^2}$	r_{63}^{theor} (10^{-12} m V $^{-1}$)	$r_{63}^{expt'l}$ (10^{-12} m V $^{-1}$)
KDP	(T) 21 ^a	3.535	8.7	9.4 ^e
	(S) 21 ^b	3.535	8.7	8.8 ^f
KDDP	(T) 50 ^c	9.017	22.3	26.4 ^e
	(S) 48 ^d	8.639	21.3	24.0 ^g
ADP	(T) 15 ^a	2.401	6.5	8.5 ^h
	(S) 14 ^d	2.212	5.5	5.5 ^h

^a Berlincourt *et al* (1964); ^b Kaminow and Harding (1963); ^c Sliker and Burlage (1963); ^d Kaminow (1965); ^e Blokh (1962); ^f Rosner *et al* (1967); ^g Christmas and Wildey (1970); ^h Carpenter (1950, 1953).

7. Ternary chalcopyrite compounds

Ternary compounds are of special interest because of their large electro-optic coefficients. As we have shown previously for LiNbO₃, LiTaO₃, and the KDP family, the coefficient is as high as 30×10^{-12} m V $^{-1}$, which is an order of magnitude higher than the coefficient for diatomic crystals. However, the use of such materials is limited to relatively short wavelengths. The infrared absorption of KDP begins at 1.5 μ m due to the vibration of H⁺ ions. The upper limit of the transparency range of oxides is at about 5 μ m. In order to extend the applicability of electro-optic materials to longer wavelengths, we have to choose compounds with heavier atoms.

For example, oxygen can be replaced by other atoms in the same column, such as S, Se or Te. The simplest compound has the structure of chalcopyrite (CuFeS₂) with a point group $\bar{4}2m$. A unit cell of the chalcopyrite structure consists of two unit cells of the zincblende structure, such as GaAs. Sulphur atoms occupy the positions of As, while Cu and Fe share evenly the positions of Ga. With a little distortion from the perfect tetragonal coordination, the ratio of the lattice constants, c/a , is usually less than 2 (Shay and Wernick 1975).

In general, compounds with the chalcopyrite structure are written as ABC₂. Both A-C and B-C bonds are not extremely ionic and contribute comparably to both the linear and the nonlinear susceptibilities. These compounds are consequently used to test our theory of additive bond susceptibilities.

Both the dielectric constant and the linear electro-optic coefficients have been measured in the crystals AgGaS₂ (Cound *et al* 1970), CuGaS₂ (Turner *et al* 1974). Our

calculation will concentrate on these compounds. The results will be compared with experimental measurements.

For the chalcopyrite crystals, only the nonzero summation $\Sigma\alpha_1\alpha_2\alpha_3$ of both bonds contributes to the electro-optic coefficient. Since these crystals have birefringence, we have $r_{41} = r_{52} \neq r_{63}$ with different dielectric constants $\epsilon_1 = \epsilon_2 \neq \epsilon_3$ and, in general,

Table 7. The calculation of electro-optic coefficients r_{63} and r_{41} for the ternary chalcopyrite crystals. r 's are in units of 10^{-12} mV^{-1} .

Characteristics	AgGaS ₂	CuGaS ₂
r_0	1.28 Å	1.1945 Å
χ	3.68	4.43
v_b	12.45 Å ³	9.763 Å ³
n_v	0.1606 Å ⁻³	0.2049 Å ⁻³
k_s	2.012 Å ⁻¹	2.095 Å ⁻¹
A-C ^a E_h	3.87 eV	4.58 eV
C	9.87 eV	9.95 eV
f_i	0.867	0.825
f	-0.1868	-0.2086
$\hbar\omega_p$	21.71 eV	24.52 eV
e_s^*/e	0.91	0.81
r_0	1.14 Å	1.162 Å
χ	5.84	6.07
v_b	8.80 Å ³	8.987 Å ³
n_v	0.2557 Å ⁻³	0.2504 Å ⁻³
B-C ^a k_s	2.174 Å ⁻¹	2.166 Å ⁻¹
E_h	5.147 eV	4.91 eV
C	5.416 eV	5.60 eV
f_i	0.525	0.565
f	-0.1465	-0.1451
$\hbar\omega_p$	18.79 eV	18.59 eV
e_s^*/e	0.72	0.75
e_s^*/e	0.732	0.688
$(\chi_A/r_A)f_A + (\chi_B/r_B)f_B$	0.6438 Å ⁻¹	0.7658 Å ⁻¹
$\Sigma\alpha_1\alpha_2\alpha_3/\Sigma\alpha^2$	0.57735	0.57735
$\epsilon_0/Ne (10^{-12} \text{ m Å/V})$	47.0	41.5
ϵ_{dc3}	14 ^b	10 ^c
$(\epsilon_{dc} - \epsilon)/\epsilon^2$	0.192	0.096
r_{ion}	+4.58	+2.56
r_{63}^{elec}	-7.85 ^d	-1.55 ^e
r_{sum}	-3.27	+1.01
$r_{expt'l}$	3.0 ^b	+1.05 ^c
ϵ_{dc1}	10 ^b	9.3 ^c
$(\epsilon_{dc} - \epsilon)/\epsilon^2$	0.096	0.078
r_{ion}	+2.30	+2.08
r_{elec}	-7.58 ^d	-1.55 ^e
r_{41}^{sum}	-5.28	+0.53
$r_{expt'l}$	4.0 ^b	+1.1 ^c

^a Wyckoff (1964); ^b Cound *et al* (1970); ^c Turner *et al* (1974); ^d Kleinman (1962); ^e Boyd *et al* (1971).

$\Sigma\alpha_1^2 = \Sigma\alpha_2^2 \neq \Sigma\alpha_3^2$. However, we can assume the crystal is in perfect tetragonal coordination in the calculation of the summation of direction cosines without introducing significant error. Therefore, the value $\Sigma\alpha_1\alpha_2\alpha_3/\Sigma\alpha^2 = 0.57735$ will be used for both A-C and B-C bonds in all chalcopyrite crystals.

We next review the considerations involved in the calculation of the electro-optic coefficients. N , r_0 , $\Sigma\alpha'_s$, k_s , and k_F (to calculate $\hbar\omega_p$) are obtained from the crystallographic data. E_h is obtained from a knowledge of the bond length. We then determine f_i and e_c^*/e using the value of C . Finally, the bond susceptibility has to be determined. In the case of only one kind of bond, there is no problem in obtaining the bond susceptibility if we know the crystal susceptibility and its geometrical factor. However, if the crystal has two kinds of bonds, only their average value of the bond polarisation can be obtained from the knowledge of the crystal susceptibility and its bond directions. In order to obtain the bond polarisation separately, a new method, which considers the crystal ABC_2 as the superposition of two fictitious zincblende crystals AC and BC, has been suggested in the investigation of bond ionicities (Levine 1973a).

The values of parameters and the results of the calculation are shown in table 7 for the two crystals $AgGaS_2$ and $CuGaS_2$. The results are compared with experiment. The measurement of the coefficient is at $0.633 \mu\text{m}$ for $AgGaS_2$, while at $3.39 \mu\text{m}$ for $CuGaS_2$. Due to the dispersion of the optical permittivity, we use the value of $\epsilon'_3 = 6.50$ and $\epsilon'_1 = 6.25$ at $0.633 \mu\text{m}$ for $AgGaS_2$. The values of r_{63} are in good agreement with experiment, while the results for r_{41} are satisfactory. For the theory, the assumption of uniform displacement may not be suitable in the chalcopyrite compound. It is quite possible that the relative movements of Ag, Cu and Ga with respect to S have different magnitudes. However, we need more experimental data to check this assumption. At the same time, it is a challenge to find a method to account for the different displacements theoretically.

8. Other ternary compounds

The ternary compounds are of continuous interest due to their simple structure, large nonlinearity, and wide range of transparency. There are many different chemical compositions in the ternary system. However, only those of $A^{II}B_2^{III}C_4^{VI}$ type and $A_3B^{III}C_3^{VI}$ type have been investigated in some detail.

In the $A^{II}B_2^{III}C_4^{VI}$ type, A could be Zn, Cd or Hg, B could be Al, Ga or In and C could be S, Se or Te. Most of these compounds are in the point group $\bar{4}$ or $\bar{4}2m$, except $ZnIn_2S_4$, which is 3m (Berger and Prochakham 1969). The crystals of $\bar{4}$ or $\bar{4}2m$ (defect chalcopyrite) have similar structures in the zincblende or chalcopyrite compounds, except that there are 25% vacancies of atom positions. In the defect chalcopyrite, there are 24 bonds per unit cell while chalcopyrite has 32. From the principle of the susceptibility superposition, χ of the defect chalcopyrite should be in general smaller than χ of chalcopyrite. This fact is seen clearly by comparing the susceptibility of $CdGa_2S_4$ and $AgGaS_2$; $[CdGa_2S_4] = 4.25$, while $[AgGaS_2] = 4.76$. A detailed crystallographic study of the defect chalcopyrite structure has been done more than twenty years ago (Hahn *et al* 1955). Although the structures of $\bar{4}$ and $\bar{4}2m$ are different, the summations of direction cosines $\Sigma\alpha_i$, $\Sigma\alpha_i\alpha_j$ and $\Sigma\alpha_i\alpha_j\alpha_k$ are the same for both structures if considered in perfectly tetragonal coordination.

The compound $ZnIn_2S_4$ has a layer structure of the point group 3m with three formulae per unit cell (Lappe *et al* 1962). The lattice constants are $a = 3.85 \text{ \AA}$ and $c = 37.0 \text{ \AA}$. The two indium ions have different coordination numbers. One third of the In

atoms occupy slightly compressed octahedral voids, while two-thirds of the In and Zn atoms occupy enlarged tetrahedral spaces.

Although the structure information for $A^{II}B_2^{III}C_4^{VI}$ crystals is complete, no measurement of the dielectric constant has been carried out. The calculation of the ionic part of the electro-optic coefficient is still impossible.

The other type of interesting ternary compound is $A_3B_2^{III}C_4^{VI}$. Three crystals of this type, Ag_3AsS_3 , Ag_3SbS_3 and Tl_3AsSe_3 have been studied in some detail for their structures, optical properties and nonlinearities. However, only the dielectric constant and electro-optic coefficients of Ag_3AsS_3 have been measured (Warner 1968).

Table 8. Values of physical parameters used in the calculation of the electro-optical coefficients of proustite (Ag_3AsS_3).

Characteristics	As-S	Ag-S
$\Sigma \alpha_3$	7.844	2.137
$\Sigma \alpha_1^2 = \Sigma \alpha_2^2$	7.291	11.589
$\Sigma \alpha_3$	3.418	12.811
$\Sigma \alpha_3^3$	-1.490	-2.266
$\Sigma \alpha_2^3 = -\Sigma \alpha_2 \alpha_1^2$	3.190	2.005
$\Sigma \alpha_3 \alpha_1^2 = \Sigma \alpha_3 \alpha_2^2$	-3.177	0.089
χ_{avg}	6.9832	5.6383
r_0	2.293 Å	2.440 Å
v_b	14.305 Å ³	17.231 Å ³
v_v	0.2097 Å ⁻³	0.1741 Å ⁻³
k_s	2.4117 Å ⁻¹	2.4877 Å ⁻¹
E_h	5.0735 eV	4.3501 eV
C	3.9107 eV	5.2620 eV
f_i	0.3727	0.5940
f	-0.1222	-0.1603
$\hbar \omega_p$	17.0175 eV	15.5051 eV
e_s^*/e	0.5745	2.0362
e_s^*/e	1.1158	
ϵ_0/Ne	80.96 mÅ V ⁻¹	

Ag_3AsS_3 belongs to the point group 3m with six formula units per unit cell. The lattice constants of its hexagonal unit cell are $a = 10.80$ Å and $c = 8.69$ Å. The bond distance of an As atom to the nearest three S atoms is 2.293 Å. Every S atom is bonded to one As atom (2.293 Å) and two Ag atoms (2.44 Å). In all, there are 54 bonds in a unit cell, where As-S has 18 and Ag-S has 36. The summations of direction cosines for both bonds are listed in table 8. From the nonzero summations of $\Sigma \alpha_i \alpha_j \alpha_k$, we know the nonzero coefficients of Ag_3AsS_3 (proustite) are r_{33} , $r_{22} = -r_{12} = -r_{61}$, $r_{13} = r_{23}$, and $r_{51} = r_{42}$.

As in the chalcopyrite compounds, we can obtain the values of most physical parameters for both bonds from the structural information. However, we cannot find the bond susceptibility in the same way, because there is no equivalent structure in 3m which has only two atoms in the formula. So we try to decompose the susceptibility into the individual bonds by using the bond-additive principle:

$$\begin{aligned} \chi_{Ag-S}(\Sigma \alpha_{1,2})_{Ag-S} + \chi_{As-S}(\Sigma \alpha_{1,2})_{As-S} &= \chi_o \\ \chi_{Ag-S}(\Sigma \alpha_3)_{Ag-S} + \chi_{As-S}(\Sigma \alpha_3)_{As-S} &= \chi_e. \end{aligned} \quad (35)$$

From the measured ordinary and extraordinary susceptibilities χ_o and χ_e , we can solve for the contribution of the individual bonds χ_{Ag-S} and χ_{As-S} . The susceptibility in table 8 is taken as if all the bonds in the crystal are occupied by the same kind of bonds. The averaged susceptibility is calculated from $\chi_{avg} = \frac{1}{3}(2\chi_o + \chi_e)$. The values of all physical parameters are shown in table 8, including the effective ionic charge.

Table 9. Comparison of the theoretical prediction for the electro-optic coefficients of Ag_3AsS_3 with experiment. All coefficients are in units of $10^{-12} mV^{-1}$. For the coefficient r_{ijk} , the first two rows are calculated as

$$\sum \frac{\chi}{r_0} [f\alpha_i\alpha_j\alpha_k + \frac{1}{2}(\alpha_i\delta_{jk} + \alpha_j\delta_{ik})] \text{ and } (\epsilon'_{dc} - \epsilon'_i)/\epsilon'_i\epsilon'_j.$$

r'_s	r_{22}	r_{13}	r_{51}	r_{33}
$\Sigma \chi/r_0 [f\alpha^3 + \alpha]$	-0.2144	0.1277	1.7295	2.1375
$(\epsilon'_{dc} - \epsilon'_i)/\epsilon'_i\epsilon'_j$	0.1310	0.1504	0.1592	0.2220
r^{ionic}	(+)2.04	(-)1.39	(-)19.98	(-)34.43
$r^{\text{elec a}}$	(-)1.10	(-)0.97	(-)0.97	?
r^{sum}	(+)0.94	(-)2.36	(-)20.95	?
$r^{\text{expt}^{\text{1b}}}$	(+)1.05	(-)2.54	?	?

^a Hulme (1967), Boggett and Gibson (1968); ^b Warner (1968).

The calculated values of the electro-optic coefficients are shown in table 9. The sign of the coefficient depends on the definition of the crystal polarity. Since all measurements of SHG coefficients and electro-optic coefficients are in absolute value, we compare the values of r^{ion} , r^{elec} and r to deduce the possible sign for each quantity. The only thing we know about the sign is the relative sign between the ionic parts of the coefficients. The estimate of signs is entered in parentheses. The prediction is in good agreement with experiment. The comparison is on the smaller coefficients where both contributions are comparable. We also predict the value of r_{33} although d_{33} has not been measured experimentally, since the ionic effect dominates in the electro-optic response. It will be of interest to see if measurements confirm the large predicted values for r_{51} and r_{33} .

9. Discussion

Based on the PV dielectric theory and the bond-charge model, we have obtained a general expression for the ionic part of the electro-optic coefficient. The theory has been applied to the calculation of the electro-optic coefficient of many crystals with different structures. The crystals used in the comparison possessed tetragonal and octahedral coordinations, single-bond, one kind of bond with different bond lengths, two different kinds of bonds. In general, we find that the theoretical calculations are in good agreement with experiment. This proves the general applicability of the theory to different crystal classes.

Referring to the key result (equation (27)), we point out that the structural information is more important than the detailed electronic information in seeking better electro-optic materials. The ionicity factor, f , is a new factor in the theory and is very small for all the investigated bonds, so the term $\Sigma \alpha_i$ usually dominates over the term

$f \sum \alpha_i \alpha_j \alpha_k$ if $\sum \alpha_i$ is not equal to zero. As a consequence, the coefficient of a crystal with nonzero $\sum \alpha_i$ should have a higher value. This is fully demonstrated by the fact that r_{33} and r_{51} in the 3m crystals like LiNbO_3 , LiTaO_3 and Ag_3AsS_3 are an order of magnitude larger than in most other crystals. However, not all 3m crystals have higher coefficients. For example, ZnIn_2S_4 has a zero value of $\sum \alpha_i$ and is not expected to have a large electro-optic coefficient, but a distorted octahedral structure is surely a promising mechanism for a larger electro-optic response. Such octahedral structure usually occurs in the point group with a 3-fold or 6-fold axis. Therefore, the crystals belonging to those point groups are promising candidates for electro-optic applications.

Although the theory has proved successful in predicting the electro-optic constants of diatomic and ternary compounds, it remains on somewhat shaky ground in the case of quaternary crystals, where more than two kinds of bonds have to be taken into account. It is very important to point out some basic assumptions and soft spots inherent in the theory.

(i) The relation between the dielectric constant and the ion-displacement depends on the bond-strength and is not completely known. In the theory, we accept it as a parameter and use its measured value in the calculation. It will be very important to know the dependence of ϵ'_{dc} on the structure and atom information like the crystal susceptibility, χ , as well as ambient temperature;

(ii) The displacement of ions is assumed to be uniform. This is not true in practical situations. The understanding of the displacements for different ions is still beyond the scope of the theory, but becomes important in more complex crystals;

(iii) The crystal susceptibility is considered as the geometrical composition of only the bond susceptibilities. The transverse bond polarisation has been completely neglected. This is justifiable for most crystals, but not for highly anisotropic bonds;

(iv) In the ideal situation, it should be possible to calculate the bond susceptibility from knowledge of the crystal structure. However, we obtain the bond susceptibility for chalcopyrite and 3m crystals from the measured macroscopic crystal susceptibility. The reason is that we find the 'b' value used in the calculation of C cannot be determined theoretically with the same accuracy as in the case of diatomic crystals. In diatomic crystals, b is related to a high degree of accuracy to the coordination number. Qualitatively, b in the ternary compounds still follows the relations but deviates highly from the predicted value. So we rely on the principle of bond-additivity to find the bond susceptibility. A generalised method to obtain b becomes very important to the further improvement of the theory.

In conclusion, the theory has been applied successfully to the understanding of the electro-optic effect in the diatomic and ternary compounds. However, considerable further improvement is needed to meet the challenge of complex crystals and the task of seeking out new electro-optic crystals.

References

- Abrahams S C, Hamilton W C and Segueira A 1967 *J. Phys. Chem. Solids* **28** 1693
 Abrahams S C, Reddy J M and Bernstein J L 1966 *J. Phys. Chem. Solids* **27** 977
 Alomas P, Sherman G, Wittig C and Coleman P D 1968 *Appl. Opt.* **8** 2557
 Berger L I and Prochukhan F C 1969 *Ternary Diamond-Like Semiconductors* (New York: Consultant Bureau)
 Berlincourt D A, Curran D R and Jaffe H 1964 *Physical Acoustics* vol 1 part A ed W P Mason (New York: Academic Press)

- Berlincourt D A, Jaffe H and Shiozawa L R 1963 *Phys. Rev.* **129** 1009
- Blokh O G 1962 *Sov. Phys.—Crystallogr.* **7** 509
- Boggett D M and Gibson A F 1968 *Phys. Lett* **28** A 33
- Boyd G D, Kasper H and McFee J H 1971 *IEEE J. Quantum Electron.* **QE-7** 563
- Callen H B 1949 *Phys. Rev.* **76** 1394
- Carpenter R O'B 1950 *J. Opt. Soc. Am.* **40** 225
- 1953 *J. Opt. Soc. Am.* **25** 1145
- Chang R K, Ducuing J and Bloembergen N 1965 *Phys. Rev. Lett.* **15** 415
- Chemla D, Kupecek P, Schwartz c, Schwab C and Goltzene A 1971 *IEEEJ. Quantum Electron.* **QE-7** 126
- Christmas T M and Wildey C G 1970 *Electron. Lett.* **6** 152
- Cound V M, Davies P H, Hulme K F and Robertson D 1970 *J. Phys. C: Solid State Phys.* **3** L83
- Faust W L and Henry C H 1966 *Phys. Rev. Lett.* **17** 1265
- Flytzanis C 1969 *Phys. Rev. Lett.* **23** 1336
- 1971 *Phys. Lett.* **34** A 99
- 1975 *Quantum Electronics* ed H Rabin and C L Tang vol 1 part A (New York: Academic Press)
- Fujimoto I 1978 *Phys. Rev. Lett.* **40** 941
- Hahn H, Frank G, Klingler W, Störger A D and Störger G 1955 *Z. Anorg. (Allg.) Chem.* **279** 241
- Hulme K F, Jones O, Davis P H and Hobden M V 1967 *Appl. Phys. Lett.* **10** 133
- Jeggo C R and Boyd G D 1970 *J. Appl. Phys.* **41** 2741
- Jones S and Mao S 1968 *J. Appl. Phys.* **39** 4038
- Kaminow I P 1965 *Phys. Rev.* **138** A1539
- 1967 *Proc. Symp. Ferroelectricity* ed E F Weller (New York: Elsevier)
- Kaminow I P and Harding G O 1963 *Phys. Rev.* **129** 1562
- Kaminow I P and Johnston W D 1967 *Phys. Rev.* **160** 519
- Kaminow I P and Turner E H 1966 *Proc. IEEE* **54** 1374
- 1971 *Handbook of Lasers* ed R J Pressley (Cleveland, Ohio: Chemical Rubber Company) p 453
- 1972 *Phys. Rev. B* **5** 1564
- Kelly R L 1966 *Phys. Rev.* **151** 721
- Kittel C 1976 *Introduction to Solid State Physics* 5th ed (New York: Wiley)
- Kleinman D A 1962 *Phys. Rev.* **128** 1761
- Kobyakov I B 1966 *Sov. Phys.—Crystallogr.* **11** 369
- Lappe F, Niggli A, Nitsche R and White J G 1962 *Z. Krist.* **117** 146
- Lawaetz P 1971 *Phys. Rev. Lett.* **26** 697
- 1972 *Phys. Rev. B* **5** 4039
- Levine B F 1969 *Phys. Rev. Lett.* **22** 787
- 1973a *J. Chem. Phys.* **59** 1463
- 1973b *Phys. Rev. B* **7** 2591
- 1973c *Phys. Rev. B* **7** 2600
- Miller R C 1964 *Appl. Phys. Lett.* **5** 17
- Miller R C and Savage A 1966 *Appl. Phys. Lett.* **9** 169
- Mooradian A and McWhorter A L 1969 *Light Scattering Spectra of Solids* ed G B Wright (New York: Springer)
- Nelson D F and Turner E H 1968 *J. Appl. Phys.* **39** 3337
- Patel C K N 1966 *Phys. Rev. Lett.* **16** 613
- Penn D R 1962 *Phys. Rev.* **128** 2093
- Phillips J C 1968a *Phys. Rev. Lett.* **20** 550
- 1968b *Phys. Rev.* **166** 832
- 1968c *Phys. Rev.* **168** 905
- 1970 *Rev. Mod. Phys.* **42** 317
- Phillips J C and Van Vechten J A 1969 *Phys. Rev. Lett.* **22** 705
- Rosner R D, Turner E H and Kaminow I P 1967 *Appl. Opt.* **6** 779
- Shay J L and Wernick J H 1975 *Ternary Chalcopyrite Semiconductors* (Oxford: Pergamon)
- Shih C and Yariv A 1980 *Phys. Rev. Lett.* **44** 281
- Sliker T R and Burlage S R 1963 *J. Appl. Phys.* **34** 1837
- Soref R A and Moos H W 1964 *J. Appl. Phys.* **35** 2152
- Turner E H 1966 *Thursday A13, Optical Society of America Annual Meeting*
- Turner E H, Buehler E and Kasper H 1974 *Phys. Rev. B* **9** 558
- Van der Ziel J P and Bloembergen N 1964 *Phys. Rev.* **135** A1662
- Van Vechten J A 1969a *Phys. Rev.* **182** 891
- 1969b *Phys. Rev.* **187** 1007

- Warner J 1968 *Brit. J. Appl. Phys. (J. Phys. D) Ser. 2* **1** 1949
- Wyckoff R W G 1963 *Crystal Structures* 2nd ed vol 1 (New York: Wiley)
- 1964 *Crystal Structure* vol 2 (New York: Wiley)
- Wynne J J and Bloembergen N 1969 *Phys. Rev.* **188** 1211
- Yariv A 1975 *Quantum Electronics* 2nd ed (New York: Wiley)
- Zubov V G, Firsova M M and Glushkova T M 1965 *Sov. Phys.-Crystallogr.* **9** 729
- Zubov V G, Firsova M M and Molokova T M 1963 *Sov. Phys.-Crystallogr.* **8** 85