# Theory of Second Harmonic Generation of Light

D. A. KLEINMAN
*Bell Telephone Laboratories, Murray Hill, New Jersey*
(Received July 10, 1962)

A theory is given for the generation of second harmonic light in crystals analogous to the classical theory of crystal optics. Optical activity and absorption are neglected, but the anisotropy is treated rigorously. An approximate treatment is also given, valid for most crystals, which retains the essential effects of the anisotropy and neglects the minor effects. All coherence effects are treated exactly, and special attention is given to the case of index matching. An exact solution is given for the second harmonic waves produced in and outside the medium when the fundamental beam enters the medium at a plane surface. A brief discussion is given of the effects of the finite aperture of the beam. The theory is applied to recent experiments on KDP and quartz, and values are given for the second-order polarization coefficients of these materials. Finally, the theory is given for second harmonic generation by focused beams based upon the diffraction theory of an ideal focus.

## I. INTRODUCTION

ONE of the most interesting applications of the ruby optical maser[1] (or laser) is the investigation of the nonlinear interaction of light with matter. The second harmonic generation of light (SHG) has been observed in quartz,[2] triglycine sulfate,[3] potassium dihydrogen phosphate,[4,5] and many other piezoelectric crystals.[6] It may be inferred from data now accumulated that SHG occurs in all piezoelectric crystals except in cases where there is absorption of the radiations involved. Experiments[6] have also shown that the effect is either absent or very weak in nonpiezoelectric crystals. This supports the suggestion of Franken *et al.*,[2] that the primary nonlinear effect is a dielectric polarization of second order in the electric field

$$\mathbf{P}=d\cdot\mathbf{EE}, \tag{1}$$

where **EE** represents the 6 components $E_iE_j$ and $d$ is a tensor which we shall call the *second-order polarization tensor*. It follows that the selection rules imposed by crystal symmetry on $d$ are the same as for the piezoelectric tensor.[7] It has also been suggested[8] that $d$ may exhibit additional symmetry, which in many important cases would reduce it to a single independent constant. If the additional symmetry is found experimentally, it means that the origin of the second-order polarization is in electronic processes of high frequency compared to the laser. It has been suggested[2] that $d$ should be of the order of an inverse atomic electric field $\sim 10^{-8}$ esu. Experimental evidence[9] favors somewhat smaller values in the range $d\sim 10^{-11}$ to $10^{-9}$ esu. The problem of determining $d$ will be discussed later in this paper.

The problem considered here is to determine the nature of the electromagnetic field arising from a dielectric polarization of the kind that might be produced at the second harmonic frequency by a laser beam. Mathematically, the discussion is based upon Maxwell's equations for a lossless anisotropic medium free of charges and currents but containing a prescribed source in the dielectric polarization. All linear induced polarization is assumed to be included in the dielectric constant tensor. The theory involves the inhomogeneous Maxwell equations, and therefore is an extension of the theory[10,11] of the optics of anisotropic media which is based upon the homogeneous Maxwell equations. Previously the nonlinear interactions of electromagnetic waves have been considered extensively for magnetic materials[12] and plasmas in a magnetic field.[13] These considerations have been directed mainly toward defining the various nonlinear mechanisms and evaluating their effects as functions of parameters such as frequency and magnetic field which are under control. In both of these cases the nonlinearities arise from well-known terms in the equations of motion which lead to recurrence relations between coefficients in the Fourier expansion of the field. In making the Fourier expansion and evaluating the leading coefficients, it is assumed that the expansion converges rapidly. The leading term, therefore, represents the usual linearized theory in which all nonlinear terms are

[1] R. J. Collins, D. F. Nelson, A. L. Schawlow, W. Bond, C. G. B. Garrett, and W. Kaiser, Phys. Rev. Letters **5**, 303 (1960). T. H. Maiman, Phys. Rev. Letters **4**, 564 (1960); Nature **187**, 493 (1960); Brit. Commun. Electron. **7**, 674 (1960).

[2] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Phys. Rev. Letters **7**, 118 (1961). These authors define the harmonic having twice the frequency of the fundamental as the second harmonic, and this terminology will be used in this paper.

[3] M. Bass, P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Phys. Rev. Letters **8**, 18 (1962).

[4] J. A. Giordmaine, Phys. Rev. Letters **8**, 19 (1962).

[5] P. D. Maker, R. W. Terhune, M. Nisenoff, and C. M. Savage, Phys. Rev. Letters **8**, 21 (1962).

[6] D. F. Edwards, J. G. Mavroides, B. Lax, Bull. Am. Phys. Soc., **7**, 14 (1962). B. Lax, J. Mavroides, and D. Edwards, Phys. Rev. Letters **8**, 166 (1962). Also private communications by P. A. Franken, R. W. Terhune, J. A. Giordmaine, R. C. Miller, and A. Savage.

[7] "Standards on Piezoelectric Crystals," Proc. Inst. Radio Engrs. **46**, 764 (1958).

[8] D. A. Kleinman, Phys. Rev. **126**, 1977 (1962).

[9] I am indebted to J. A. Giordmaine and R. W. Terhune for discussions concerning the value of $d$.

[10] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, New York, 1959), Chap. XIV.

[11] G. N. Ramachandran and S. Ramaseshan, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1961), Vol. BXXV, p. 1.

[12] D. Douthett, I. Kaufman, and A. Risley, J. Appl. Phys., **32**, 1905 (1961), contains references to earlier work.

[13] R. F. Whitmer and E. B. Barrett, Phys. Rev., **121**, 661 (1961), contains references to earlier work.

neglected. The second harmonic then arises from the nonlinear response of the system to the fundamental. The same approach is used here by postulating that a dielectric polarization **P** exists, arising from a nonlinear material Eq. (1) and a strong fundamental field **E**. In general $d$ will be a function of frequency,[8] so that **P** and **E** in (1) represent Fourier amplitudes. In this paper **P** will represent the second harmonic component, and the zero-frequency[2] component predicted by (1) will be ignored. In this way a theory is obtained for SHG. In formal appearance the theory has little resemblance to the previous theories just mentioned, but strongly resembles the electromagnetic theory of optics[10,11] for anisotropic media.

The importance of anisotropy in SHG has recently been emphasized by Giordmaine[4] and by Maker *et al.*[5] From the first,[2] it was recognized that because of dispersion a "coherence volume" exists for SHG within which the second harmonic wave can remain in phase with the polarization producing it. By means of anisotropy it is possible to match the refractive indices of the second harmonic wave and the fundamental wave, and thereby increase the coherence volume to the entire illuminated volume of the crystal. This effect has been demonstrated experimentally[4,5] in potassium dihydrogen phosphate. In this way intense SHG can be obtained from unfocussed parallel laser beams. Without the use of index matching, the output is smaller by several orders of magnitude unless the laser intensity is increased by focusing. Both index matching and focusing are treated in detail in this paper. Particular attention is given to the case, very important in practice, in which indices are nearly matched. Also considered are the waves arising from the boundary conditions at the surface of the medium. In the interest of simplicity, optical activity and absorption are neglected. In anisotropic crystals optical activity is usually a very small effect compared to the birefringence and can be neglected except for propagation very close to an optic axis direction.[11] Even in this case the rotations are of the order of 10° per mm and could only be important if the coherence volume is large. On the other hand, the coherence volume is only large in the nearly matching case which never corresponds to propagation near an optic axis. The neglect of absorption is justified by the fact that crystals exist which exhibit SHG and have very low absorption. Obviously, absorption at the second harmonic frequency is undesirable and only serves to reduce SHG. Absorption at the fundamental frequency would probably lead to serious burning by the laser beam.

It has been pointed out that the nonlinear interaction can mix two light beams.[3–5] If the beams are plane waves with wave vectors $\mathbf{K}_1$ and $\mathbf{K}_2$, the polarization at the sum frequency has the wave vector[4,5] $\mathbf{K}=\mathbf{K}_1+\mathbf{K}_2$. The theory given here applies equally well to this type of process as to the second harmonic processes $\mathbf{K}=2\mathbf{K}_1$ or $\mathbf{K}=2\mathbf{K}_2$. In any case $\mathbf{K}$ can be written

$$\mathbf{K}=(\omega/c)n'\mathbf{s}_k, \tag{2}$$

where $\omega$ is either the sum or the second harmonic frequency, $\mathbf{s}_k$ is a unit vector in the direction of $\mathbf{K}$, and $n'$ is an effective refractive index. In general $n'$ will not equal the refractive index $n$ for a wave freely propagating in the medium in the direction $\mathbf{s}_k$ with frequency $\omega$. The effect of normal dispersion is to make $n'$ less than $n$, and in the case of mixing a nonvanishing angle between $\mathbf{K}_1$ and $\mathbf{K}_2$ further reduces $n'$. The coherence volume may be defined as the cross-sectional area of the beam multiplied by a coherence length

$$l_{\text{coh}}=\lambda[(n'/n)-1]^{-1}, \tag{3}$$

where $\lambda=(2\pi c/\omega n)$ is the wavelength in the medium of the second harmonic wave. A coherence length of $l_c\sim 14\mu$ has been observed in quartz by Maker *et al.*,[5] which is typical for nonmatching conditions ($n'\neq n$).

The theory of crystal optics[10] is based upon the vector wave equation for the electric field $\mathbf{E}(\mathbf{r})$:

$$\nabla\times\nabla\times\mathbf{E}-(\omega^2/c^2)\epsilon\cdot\mathbf{E}=0, \tag{4}$$

where $\epsilon$ is the dielectric constant dyadic. If $\mathbf{E}(\mathbf{r})$ is assumed to be a plane wave,

$$\mathbf{E}(\mathbf{r})=\mathbf{E}e^{i(\omega/c)n\mathbf{s}\cdot\mathbf{r}}, \tag{5}$$

where $n$ is a refractive index and $\mathbf{s}$ a unit vector, a set of algebraic equations is obtained which may be written

$$\alpha_s\cdot\mathbf{E}=0, \tag{6}$$

where $\alpha_s$ is the dyadic

$$\alpha_s=n^2(I-\mathbf{ss})-\epsilon, \tag{7}$$

with $I$ the unit dyadic or idemfactor. Since optical activity and absorption are being neglected, $\epsilon$ is real and symmetric. According to (6), $\alpha_s$ must have an eigenvalue zero, and the field $\mathbf{E}$ is the corresponding eigenvector. The condition that $\alpha_s$ possesses a zero eigenvalue is

$$\det|\alpha_s|=0, \tag{8}$$

which for a general $\mathbf{s}$ defines the two refractive indices $n$. For a given $\mathbf{s}$ two dyadics $\alpha_s$ are then defined and two fields $\mathbf{E}$, so that two waves $\mathbf{E}$, $\mathbf{s}$, $n$ are completely defined except for an arbitrary factor. An exceptional case is when $\mathbf{s}$ is along one of the two optic axes, in which case the two indices become equal and the field $\mathbf{E}$ is only required to lie in a certain plane which in general is not normal to $\mathbf{s}$. Crystals having a symmetry axis of threefold or higher symmetry are uniaxial, since the two optic axes are required to lie along the symmetry axis. Crystals having four threefold axes or three fourfold axes are optically isotropic, and any direction may be considered an optic axis. The possibility of index matching in SHG depends upon the existence of two different waves $\mathbf{E}$, $\mathbf{s}$, $n$, and therefore is limited to uniaxial or biaxial crystals.

Considerable use will be made in what follows of the surface, which will be called the *index surface*, obtained by constructing all vectors $\mathbf{s}$ from some origin and laying off the distance $n(\mathbf{s})$ along each $\mathbf{s}$. The resulting surface has two sheets, corresponding to the two indices, except for points on the optic axes. This surface is helpful in analyzing boundary conditions at the surface of the medium. Let the boundary surface be the plane

$$\mathbf{N}\cdot\mathbf{r}=0, \tag{9}$$

where $\mathbf{N}$ is a unit vector normal to the boundary and directed into the medium. Let $\mathbf{s}_{v1}$, $\mathbf{s}_{v2}$ be the directions of the waves in the vacuum outside the boundary, and let $\mathbf{s}_1, \mathbf{s}_2, \cdots$ be the directions of waves inside the medium. Associated with $\mathbf{s}_1, \mathbf{s}_2, \cdots$ are the refractive indices $n_1, n_2, \cdots$. In order to satisfy boundary conditions at each point $\mathbf{r}$ on the plane (9), it is necessary that the relative phase of all waves be independent of the tangential component of $\mathbf{r}$. It can be shown[14] that this requires

$$n_1\mathbf{s}_1\times\mathbf{N}=n_2\mathbf{s}_2\times\mathbf{N}=\cdots=\mathbf{s}_{v1}\times\mathbf{N}=\mathbf{s}_{v2}\times\mathbf{N}. \tag{10}$$

This is the generalization of Snell's law for anisotropic media. Now consider the index surface and construct $\mathbf{s}_{v1}$, which may represent an incident plane wave. According to (10) $\mathbf{s}_{v1}-n_1\mathbf{s}_1$ lies along $\mathbf{N}$ and similarly $\mathbf{s}_{v1}-n_2\mathbf{s}_2$. By proceeding from $\mathbf{s}_{v1}$ along the direction $\mathbf{N}$ to intersect the index surface, the refracted waves $\mathbf{s}_1$, $\mathbf{s}_2$ can be determined.

## II. GENERAL FORMULATION

In this section we obtain a formal solution to the problem of the radiation produced by a plane polarization wave

$$\mathbf{P}(\mathbf{r},t)=\mathbf{P}e^{i\mathbf{K}\cdot\mathbf{r}-i\omega t}. \tag{11}$$

The wave vector $\mathbf{K}$ may be written in terms of $n'$ and $\mathbf{s}_k$ as in (2), and $\omega$ is twice the angular frequency of the fundamental wave or the sum of the frequencies of two mixed waves. The direction and magnitude of $\mathbf{P}$ are determined by the tensor $d$ and the electric fields of the fundamental or the mixed waves. Given $\mathbf{P}$ and $\mathbf{K}$ we must determine the second harmonic electric field from the inhomogeneous vector wave equation

$$\nabla\times\nabla\times\mathbf{E}-(\omega^2/c^2)\epsilon\cdot\mathbf{E}=(4\pi\omega^2/c^2)\mathbf{P}e^{i\mathbf{K}\cdot\mathbf{r}}. \tag{12}$$

Since we assume that the permeability is the same as in free space, the magnetic field need not be considered explicitly. We seek initially a particular solution of (12) in the form

$$\mathbf{E}(\mathbf{r})=\mathbf{E}'e^{i(\omega/c)n'\mathbf{s}_k\cdot\mathbf{r}}. \tag{13}$$

Substituting (13) into (12) gives

$$\alpha_k\cdot\mathbf{E}'=4\pi\mathbf{P}, \tag{14}$$

where $\alpha_k$ is the dyadic

$$\alpha_k=n'^2(I-\mathbf{s}_k\mathbf{s}_k)-\epsilon. \tag{15}$$

[14] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), p. 491.

We see that $\alpha_k$ is analogous to $\alpha_s$ in (7) except that $\alpha_k$ is not required to possess a zero eigenvalue. If $n'=n_k$, where $n_k$ is one of the refractive indices at frequency $\omega$ for $\mathbf{s}_k$, $\alpha_k$ will possess a zero eigenvalue; otherwise it will not. If $n'\neq n_k$, as in the nonmatching or general case, the inverse dyadic $\alpha_k^{-1}$ exists and the solution of (14) is

$$\mathbf{E}'=\alpha_k^{-1}\cdot 4\pi\mathbf{P}. \tag{16}$$

The wave (13) with $\mathbf{E}'$ given by (16) is a particular solution of the inhomogeneous wave Eq. (12). It can exist only when the source $\mathbf{P}(\mathbf{r},t)$ is present, since it is proportional to $\mathbf{P}$. The waves (5), on the other hand, can exist in the absence of any sources in the medium. It is convenient to call the waves (5) *free waves* and (13) *forced waves*. Evidently, any free wave can be added to (13) to construct a solution of (12). The construction using the index surface shows that in general two free waves are required for satisfying boundary conditions on a plane boundary.

Formally the inverse $\alpha_k^{-1}$ may be constructed as follows: Define the eigenvalues[15] $\lambda_k^{(j)}$ and eigenvectors $\mathbf{U}_k^{(j)}$ of $\alpha_k$

$$\alpha_k\cdot\mathbf{U}_k^{(j)}=\lambda_k^{(j)}\mathbf{U}_k^{(j)}\quad(j=1,2,3). \tag{17}$$

Since $\alpha_k$ is real and symmetric, the $\mathbf{U}_k^{(j)}$ are mutually orthogonal and may be normalized to unit length. If none of the $\lambda_k^{(j)}$ vanishes, the inverse is then

$$\alpha_k^{-1}=\textstyle\sum_j(1/\lambda_k^{(j)})\mathbf{U}_k^{(j)}\mathbf{U}_k^{(j)}. \tag{18}$$

For later use we shall also define the eigenvalues and eigenvectors of $\alpha_s$:

$$\alpha_s\cdot\mathbf{U}_s^{(j)}=\lambda_s^{(j)}\mathbf{U}_s^{(j)}, \tag{19}$$

with

$$\lambda_s^{(1)}=0. \tag{20}$$

Thus, the electric field for the free wave $n$, $\mathbf{s}$ is in the direction $\mathbf{U}_s^{(1)}$.

Now consider the *matching case* $n'=n_k$ in which $\lambda_k^{(1)}=0$. From (18) we see that this case presents no difficulty if $\mathbf{P}\cdot\mathbf{U}_k^{(1)}=0$. If $\mathbf{P}\cdot\mathbf{U}_k^{(1)}\neq 0$, there is no forced wave in the form of a plane wave of constant amplitude. Therefore, we seek a solution in the form of a linearly growing wave. Let the field in the matching case be written

$$\mathbf{E}_M(\mathbf{r})=\mathbf{E}_1(\mathbf{N}\cdot\mathbf{r})e^{i(\omega/c)n_M\mathbf{s}_M\cdot\mathbf{r}}+\mathbf{E}_{2,3}e^{i(\omega/c)n_M\mathbf{s}_M\cdot\mathbf{r}}, \tag{21}$$

where

$$\begin{aligned}\mathbf{E}_1&=E_1\mathbf{U}_M^{(1)},\\ \mathbf{E}_{2,3}&=E_2\mathbf{U}_M^{(2)}+E_3\mathbf{U}_M^{(3)}.\end{aligned} \tag{22}$$

Here the subscript $k$ has been replaced by $M$ to indicate the matching case, and $n_M$ is the refractive index for $\mathbf{s}_M=\mathbf{s}_k$. In (21), $\mathbf{N}$ is an arbitrary unit vector indicating the direction in which $\mathbf{E}_M(\mathbf{r})$ grows. Later, it will be

[15] See, for example, H. Margenau and G. Murphy, The Mathematics of Physics and Chemistry (D. van Nostrand Company, Inc., Princeton, New Jersey, 1943), Chap. 10.

shown that $\mathbf{N}$ represents the normal to the surface of the medium. Substituting (21) into (12) gives

$$i(\omega/c)n_M[\mathbf{s}_M\times(\mathbf{N}\times\mathbf{E}_1)+\mathbf{N}\times(\mathbf{s}_M\times\mathbf{E}_1)] + (\omega^2/c^2)\alpha_M\cdot\mathbf{E}_{2,3}=(4\pi\omega^2/c^2)\mathbf{P}. \quad (23)$$

The solution of (23) is readily found to be

$$E_1=-2\pi i(\omega/c)n_M^{-1} \times[(\mathbf{N}\cdot\mathbf{U}_M^{(1)})(\mathbf{s}_M\cdot\mathbf{U}_M^{(1)})-\mathbf{N}\cdot\mathbf{s}_M]^{-1}\mathbf{P}\cdot\mathbf{U}_M^{(1)}, \quad (24)$$

$$E_2=-(4\pi/\lambda_M^{(2)})\mathbf{P}\cdot\mathbf{U}_M^{(2)} -i(c/\omega)n_M[(\mathbf{N}\cdot\mathbf{U}_M^{(2)})(\mathbf{s}_M\cdot\mathbf{U}_M^{(1)}) +(\mathbf{N}\cdot\mathbf{U}_M^{(1)})(\mathbf{s}_M\cdot\mathbf{U}_M^{(2)})](1/\lambda_M^{(2)})E_1. \quad (25)$$

The equation for $E_3$ is similar to (25) with $\mathbf{U}_M^{(3)}$, $\lambda_M^{(3)}$ replacing $\mathbf{U}_M^{(2)}$, $\lambda_M^{(2)}$. Note that $E_1$ has the dimensions of an electric field divided by a length. We see that in the matching case the free wave which is matched becomes a linearly growing forced wave. In addition there is a nongrowing forced wave with its electric field perpendicular to that of the free wave. The nongrowing component is a slowly varying function of $\mathbf{s}_k$, but the growing wave is rapidly varying and the form found here is only valid for exact matching $\mathbf{s}_k=\mathbf{s}_M$.

## III. THE NEARLY MATCHING CASE

The formulation of the preceding section is deficient in the respect that it provides no convenient means for treating cases which are arbitrarily close to matching. In practice the laser beam has a finite angular spread, so that all of the beam can never satisfy the matching condition exactly. A proper formulation of the nearly matching case requires that a forced wave be found which satisfies (12) and reduces to (21) in the limit as $n'$, $\mathbf{s}_k$ approaches a matching condition. We imagine that $n'$ is a function of $\mathbf{s}_k$, so that an *effective-index surface* can be constructed for the polarization waves. The simplest case to consider is SHG in which $\mathbf{P}(\mathbf{r},t)$ arises from a single plane wave from the laser in the direction $\mathbf{s}_k$. Then $n'$ is a refractive index for $\mathbf{s}_k$ at the fundamental frequency $\omega/2$, and the effective-index surface is just one sheet of the index surface of the medium at frequency $\omega/2$. Actually only a small part of the effective index surface is needed corresponding to the range of $\mathbf{s}_k$ present in the beam. The matching directions which we shall denote by $\mathbf{s}_M$ are determined by the intersection of the index surface and the effective-index surface

$$n'(\mathbf{s}_M)=n(\mathbf{s}_M)\equiv n_M. \quad (26)$$

In general the intersection will be a continuous curve on the index surface if matching is possible. For further comments on the effective-index surface see the end of this section.

We may expect that a proper forced wave for the nearly matching case can be constructed from the forced wave (13) and a suitable free wave. If the surface of the medium is plane only certain free waves can be combined with (13). If now $\mathbf{N}$ is the inward surface normal, any free wave $n$, $\mathbf{s}$ which can be combined with (13) must satisfy

$$n\mathbf{s}\times\mathbf{N}=n'\mathbf{s}_k\times\mathbf{N}, \quad (27)$$

because of the considerations which lead to (10). The two allowed waves can be found using the index surface by a construction similar to that which gives the refracted waves from a given incident wave. Construct the index surface and from the same origin draw $n'\mathbf{s}_k$; then proceed along $\pm\mathbf{N}$ to intersect the two sheets of the surface. We may write (27) in the form

$$n'\mathbf{s}_k-n\mathbf{s}=\varphi\mathbf{N}, \quad (28)$$

where $\varphi$ is a scalar function of $\mathbf{s}_k$, measuring how far $\mathbf{s}_k$ is from matching conditions. Of the two waves allowed by (27), we shall be interested in this section only in the one which lies on the sheet containing $\mathbf{s}_M$, and this is the wave meant by $n$,$\mathbf{s}$ in (28). As the matching condition is approached, $\mathbf{s}_k\rightarrow\mathbf{s}_M$ and also $\mathbf{s}\rightarrow\mathbf{s}_M$.

It is now clear that $n$,$\mathbf{s}$ must be added to (13) in such a way that the singularity in $\mathbf{E}'$ as $\lambda_k^{(1)}\rightarrow 0$ is removed. From (14) and (18) we find that the desired form is

$$\mathbf{E}_k(\mathbf{r})=\alpha_k^{-1}\cdot 4\pi\mathbf{P}e^{i(\omega/c)n'\mathbf{s}_k\cdot\mathbf{r}} -(1/\lambda_k^{(1)})\mathbf{U}_s^{(1)}\mathbf{U}_k^{(1)}\cdot 4\pi\mathbf{P}e^{i(\omega/c)n\mathbf{s}\cdot\mathbf{r}}. \quad (29)$$

From (18) and (28), this can be written

$$\mathbf{E}_k(\mathbf{r})=\{\textstyle\sum_{j=2,3}(1/\lambda_k^{(j)})\mathbf{U}_k^{(j)}\mathbf{U}_k^{(j)}\cdot 4\pi\mathbf{P} +(1/\lambda_k^{(1)})(\mathbf{U}_k^{(1)}-\mathbf{U}_s^{(1)}e^{-i(\omega/c)\varphi\mathbf{N}\cdot\mathbf{r}})\mathbf{U}_k^{(1)}\cdot 4\pi\mathbf{P}\} \times e^{i(\omega/c)n'\mathbf{s}_k\cdot\mathbf{r}}, \quad (30)$$

where the second term in brackets is seen to be finite as $\lambda_k^{(1)}\rightarrow 0$, $\varphi\rightarrow 0$, $\mathbf{U}_k^{(1)}\rightarrow\mathbf{U}_s^{(1)}$. It may readily be shown that the limit approached by the second term is, leaving off the exponential factor,

$$\lim_{\lambda_k(1)\rightarrow 0}\{(1/\lambda_k^{(1)})[(\mathbf{U}_k^{(1)}-\mathbf{U}_s^{(1)}) +\mathbf{U}_s^{(1)}i(\omega/c)\varphi\mathbf{N}\cdot\mathbf{r}]\mathbf{U}_s^{(1)}\cdot 4\pi\mathbf{P}\}. \quad (31)$$

Clearly the first term in (30) is identical with the first term in (25) representing the nongrowing forced wave due to components of $\mathbf{P}$ which cannot produce a growing wave in the matching limit. The term containing $(\mathbf{U}_k^{(1)}-\mathbf{U}_s^{(1)})$ in (31) is to be identified with the term in (25) proportional to $E_1$. Finally the growing term containing $\mathbf{N}\cdot\mathbf{r}$ in (31) is to be identified with $E_1(\mathbf{N}\cdot\mathbf{r})$, with $E_1$ given by (24). If these identifications are correct $\mathbf{E}_k(\mathbf{r})$ in (30) approaches $\mathbf{E}_M(\mathbf{r})$ in (21) in the matching limit.

We now verify that (30) does, in fact, reduce to (21) in the limit $\mathbf{s}_k\rightarrow\mathbf{s}_M$. The identifications mentioned above will hold providing the following limits are true:

$$\lim(\varphi/\lambda_k^{(1)}) =-(2n_M)^{-1}[(\mathbf{N}\cdot\mathbf{U}_M^{(1)})(\mathbf{s}_M\cdot\mathbf{U}_M^{(1)})-\mathbf{N}\cdot\mathbf{s}_M]^{-1}, \quad (32)$$

$$\lim\left[\frac{\mathbf{U}_M{}^{(2)}\cdot(\mathbf{U}_k{}^{(1)}-\mathbf{U}_s{}^{(1)})}{\lambda_k{}^{(1)}}\right]$$

$$=-\frac{\mathbf{U}_M{}^{(2)}\cdot(\mathbf{N}\mathbf{s}_M+\mathbf{s}_M\mathbf{N})\cdot U_M{}^{(1)}}{2\lambda_M{}^{(2)}[(\mathbf{N}\cdot\mathbf{U}_M{}^{(1)})(\mathbf{s}_M\cdot\mathbf{U}_M{}^{(1)})-\mathbf{N}\cdot\mathbf{s}_M]}. \quad (33)$$

From (28)

$$n'^2-n^2=2n\varphi(\mathbf{N}\cdot\mathbf{s}) \quad (34)$$

to first order in $\varphi$. Therefore, from (7) and (15)

$$\alpha_k=\alpha_s+n\varphi(2\mathbf{N}\cdot\mathbf{s}I-\mathbf{s}\mathbf{N}-\mathbf{N}\mathbf{s}) \quad (35)$$

to first order in $\varphi$. From (17)

$$\lambda_k{}^{(1)}=\mathbf{U}_k{}^{(1)}\cdot\alpha_k\cdot\mathbf{U}_k{}^{(1)}, \quad (36)$$

while to first order in $\varphi$

$$\mathbf{U}_k{}^{(1)}\cdot\alpha_s\cdot\mathbf{U}_k{}^{(1)}=0. \quad (37)$$

Thus to first order we have

$$\lambda_k{}^{(1)}=2n\varphi[\mathbf{N}\cdot\mathbf{s}-(\mathbf{s}\cdot\mathbf{U}_k{}^{(1)})(\mathbf{N}\cdot\mathbf{U}_k{}^{(1)})]. \quad (38)$$

Now (32) follows immediately from (38) as $\mathbf{s}\rightarrow\mathbf{s}_M$. The eigenvector equation for $\mathbf{U}_k{}^{(1)}$ can be written from (17) and (35)

$$\alpha_s\cdot\mathbf{U}_k{}^{(1)}+n\varphi(2\mathbf{N}\cdot\mathbf{s}I-\mathbf{s}\mathbf{N}-\mathbf{N}\mathbf{s})\cdot\mathbf{U}_k{}^{(1)}=\lambda_k{}^{(1)}\mathbf{U}_k{}^{(1)} \quad (39)$$

to first order in $\varphi$. Taking the scalar product with $\mathbf{U}_s{}^{(2)}$ and retaining only first-order terms gives

$$\lambda_s{}^{(2)}\mathbf{U}_s{}^{(2)}\cdot\mathbf{U}_k{}^{(1)}=n\varphi\mathbf{U}_s{}^{(2)}\cdot(\mathbf{s}\mathbf{N}+\mathbf{N}\mathbf{s})\cdot\mathbf{U}_k{}^{(1)}. \quad (40)$$

The proof of (33) now follows upon noting that

$$\lim\left[\frac{\mathbf{U}_M{}^{(2)}\cdot(\mathbf{U}_k{}^{(1)}-\mathbf{U}_s{}^{(1)})}{\lambda_k{}^{(1)}}\right]=\lim\left[\frac{\mathbf{U}_s{}^{(2)}\cdot\mathbf{U}_k{}^{(1)}}{\lambda_k{}^{(1)}}\right] \quad (41)$$

and substituting for $\varphi/\lambda_k{}^{(1)}$ from (38). This completes the proof that (30) reduces to (21) in the matching limit. Therefore $\mathbf{E}_k(\mathbf{r})$ is a proper forced wave for the nearly matching case. It should be noted that no approximations have been made thus far and (30) is exact.

We now introduce approximations appropriate to the nearly matching case. The second term of (30) can be written in the form, leaving off the exponential factor,

$$(1/\lambda_k{}^{(1)})[(\mathbf{U}_k{}^{(1)}-\mathbf{U}_s{}^{(1)})+\mathbf{U}_s{}^{(1)}i(\omega/c)\varphi\mathbf{N}\cdot\mathbf{r}g(i\omega\varphi\mathbf{N}\cdot\mathbf{r}/c)]\cdot(\mathbf{U}_k{}^{(1)}\cdot 4\pi\mathbf{P}), \quad (42)$$

where $g$ is the function

$$g(z)=(1-e^{-z})/z. \quad (43)$$

If, now, we go to the matching limit $\mathbf{s}_k\rightarrow\mathbf{s}_M$ but retain $g$, we obtain

$$\mathbf{E}_k(\mathbf{r})=\mathbf{E}_1(\mathbf{N}\cdot\mathbf{r})g(2i\psi)e^{i(\omega/c)n_M\mathbf{s}_M\cdot\mathbf{r}}+\mathbf{E}_{2,3}e^{i(\omega/c)n_M\mathbf{s}_M\cdot\mathbf{r}}, \quad (44)$$

where $\mathbf{E}_1$, $\mathbf{E}_{2,3}$ are defined in (22), (24), and (25), and

$$\psi=(\omega/2c)\varphi\mathbf{N}\cdot\mathbf{r}. \quad (45)$$

All the essential effects of small deviations from the matching condition are contained in $g(2i\psi)$. The linearly growing amplitude of the $E_1$ wave in the matching limit is due to interference between the forced wave $n'$, $\mathbf{s}_k$ and the free wave $n$, $\mathbf{s}$. The distance over which this interference is effective is determined by $\varphi$ defined in (28). The *coherence length for the wave* $l_{\text{coh}}$ may be defined in terms of $\varphi$ as follows: Let $g(2i\psi)$ vanish for $\mathbf{r}=\mathbf{s}l_{\text{coh}}$; then

$$l_{\text{coh}}=|2\pi c/\omega\varphi|(\mathbf{N}\cdot\mathbf{s})^{-1}. \quad (46)$$

This definition is equivalent to that previously given in (3).

Consider the index surface and the effective-index surface in the region of their intersection. It is assumed that such an intersection exists and that the polarization $\mathbf{P}(\mathbf{r},t)$ is a narrow pencil of waves $n'$, $\mathbf{s}_k$ which includes part of the intersection. Let unit normal vectors be constructed to the two surfaces, $\mathbf{N}_s$ for the index surface and $\mathbf{N}_k$ for the effective-index surface. Then the vector $\mathbf{N}_k\times\mathbf{N}_s$ lies along the intersection, which, if the pencil is narrow, may be considered a straight line. Furthermore, let

$$\boldsymbol{\delta}=\mathbf{s}_k-\mathbf{s}_0, \quad |\delta|\ll 1, \quad (47)$$

where $\mathbf{s}_0$ is the central wave of the pencil, and let $\mathbf{s}_0$ be a matching direction $\mathbf{s}_M$. Then from (28) it can readily be shown that $\varphi$ may be written

$$\varphi=n_0(\mathbf{N}\cdot\mathbf{N}_s)^{-1}(\mathbf{s}_0\cdot\mathbf{N}_k)^{-1}\times[(\mathbf{s}_0\cdot\mathbf{N}_k)(\boldsymbol{\delta}\cdot\mathbf{N}_s)-(\mathbf{s}_0\cdot\mathbf{N}_s)(\boldsymbol{\delta}\cdot\mathbf{N}_k)], \quad (48)$$

where $n_0$ is the matching index at $\mathbf{s}_0$. It is evident that $\varphi$ given by (48) vanishes when $\boldsymbol{\delta}$ lies along $\mathbf{N}_k\times\mathbf{N}_s$. Clearly $\boldsymbol{\delta}$ is perpendicular to $\mathbf{s}_0$ and therefore may be expressed in terms of a pair of polar coordinates. Let $\delta=|\boldsymbol{\delta}|$ and let $\theta$ be the angle $\boldsymbol{\delta}$ makes with the matching line $\mathbf{N}_k\times\mathbf{N}_s$ projected onto a plane normal to $\mathbf{s}_0$. Then it can be shown that $\varphi$ is given by

$$\varphi=n_0(\mathbf{N}\cdot\mathbf{N}_s)^{-1}(1-2R\mathbf{N}_k\cdot\mathbf{N}_s+R^2)^{1/2}\delta\sin\theta, \quad (49)$$

where

$$R=(\mathbf{s}_0\cdot\mathbf{N}_s)/(\mathbf{s}_0\cdot\mathbf{N}_k). \quad (50)$$

If $\mathbf{N}_s$, $\mathbf{N}_k$, $\mathbf{s}_0$ lie in the same plane, as in uniaxial crystals,

$$(1-2R\mathbf{N}_k\cdot\mathbf{N}_s+R^2)^{1/2}=(\mathbf{s}_0\cdot\mathbf{N}_k)^{-1}\sin\rho, \quad (51)$$

where $\rho$ is the angle[16] between $\mathbf{N}_k$ and $\mathbf{N}_s$. In many cases $\mathbf{N}_s$ and $\mathbf{N}_k$ will be very nearly in the direction $\mathbf{s}_0$, and we can write

$$\begin{aligned}\mathbf{N}_s&=(1-\delta_s{}^2)^{1/2}\mathbf{s}_0+\boldsymbol{\delta}_s, \quad \mathbf{s}_0\cdot\boldsymbol{\delta}_s=0\\ \mathbf{N}_k&=(1-\delta_k{}^2)^{1/2}\mathbf{s}_0+\boldsymbol{\delta}_k, \quad \mathbf{s}_0\cdot\boldsymbol{\delta}_k=0\\ &|\delta_s|^2\ll 1, \quad |\delta_k|^2\ll 1.\end{aligned} \quad (52)$$

[16] The angle $\rho$ is related to the constant $K$ defined by Giordmaine (reference 4) as follows: $K=2\tan\rho$.

Then we have

$$(1-2R\mathbf{N}_k\cdot\mathbf{N}_s+R^2)^{1/2}=|\boldsymbol{\delta}_k-\boldsymbol{\delta}_s|. \quad (53)$$

We now consider the energy flow in the forced wave (44) with the aid of two reasonable assumptions which greatly simplify the discussion. We assume that conditions are very close to matching, so that

$$\varphi\ll n_M,$$
$$l_{\mathrm{coh}}\gg\lambda_M,$$

where $\lambda_M=(2\pi c/n_M\omega)$ is the wavelength. Then the $\mathbf{E}_1$ wave can be very large compared to the $\mathbf{E}_{2,3}$ wave; except near the zeros of $g(2i\psi)$ we have

$$E_1lg\gg E_2,\ E_3,$$

where $l=\mathbf{N}\cdot\mathbf{r}$. The second assumption is that the medium is not exceptionally anisotropic. The treatment up to this point has been completely general in regard to anisotropy. It is well known[10] that in an anisotropic medium the flow of energy is not in general along $\mathbf{s}$. The component of the time average Poynting vector along $\mathbf{s}$ is

$$S=(nc/8\pi)|E_t|^2, \quad (54)$$

where $E_t$ is the transverse component of the amplitude of the electric field. Since the anisotropies of most crystals are quite small, of the order of a few percent, we shall assume that $\mathbf{E}_1$ is nearly transverse, and that $S$ in (54) is essentially the total energy flow. Thus we write

$$S=(nc/8\pi)|E_1|^2l^2|g(2i\psi)|^2. \quad (55)$$

From (43) it follows that

$$|g(2i\psi)|^2=(\sin^2\psi)/\psi^2. \quad (56)$$

It is clear that $S$ varies over the narrow pencil of waves comprising the laser beam. An individual wave in this pencil is described by the coordinates $\delta$, $\theta$ in terms of which $\psi$ may be calculated from (45) and (49). Thus $S$ may be averaged over the pencil to give the mean Poynting vector

$$\langle S\rangle=(n_0c/8\pi)|E_1|^2l^2\langle(\sin^2\psi)/\psi^2\rangle. \quad (57)$$

The total SHG is the area of the laser beam multiplied by $\langle S\rangle$. This includes all the contributions due to mixing of slightly divergent waves in the pencil.

We may define the average in (57) as a function

$$F(\beta l\Delta)=\langle(\sin^2\psi)/\psi^2\rangle, \quad (58)$$

where $\Delta$ is a measure of the spread of the beam, and from (49) and (45)

$$\beta=n_0(\omega/2c)(\mathbf{N}\cdot\mathbf{N}_s)^{-1}(1-2R\mathbf{N}_k\cdot\mathbf{N}_s+R^2)^{1/2}. \quad (59)$$

The form of $F$ will depend upon the distribution of directions $\delta$, $\theta$ in the beam. The simplest assumption is that the distribution is independent of $\theta$ and constant for $\delta$ in the range $0\leq\delta\leq\Delta$ and vanishes for $\delta>\Delta$. Then

$$F(z)=\frac{4}{\pi}\int_0^1\frac{(1-\xi^2)^{\frac{1}{2}}\sin^2 z\xi}{z^2\xi^2}d\xi \quad (60)$$

$$=1-z^2/12+\cdots \quad (61)$$

$$\approx 2/z-4/\pi z^2+\cdots. \quad (62)$$

The two expansions cover the entire range of $z$ with sufficient accuracy for most purposes. At $z=2$, (61) and (62) give 0.667 and 0.682, respectively; (61) should be used for $z<2$. We see that there is a transition between two different behaviors at $z=\beta l\Delta=2$. It is convenient to define a *coherence length for the pencil*

$$l_{\mathrm{coh}}'=2/\beta\Delta. \quad (63)$$

This should not be confused with $l_{\mathrm{coh}}$ in (46), the coherence length for the wave. The pencil contains many waves with widely varying $l_{\mathrm{coh}}$; $l_{\mathrm{coh}}'$ is a suitable average over the distribution of $l_{\mathrm{coh}}$ in the pencil. We may distinguish two limiting cases:

(a) Thin crystal: $l\ll l_{\mathrm{coh}}'$.

In this case (57) becomes

$$\langle S\rangle=(n_0c/8\pi)|E_1|^2l^2. \quad (64)$$

(b) Thick crystal: $l\gg l_{\mathrm{coh}}'$.

$$\langle S\rangle=(n_0c/8\pi)|E_1|^2(2l/\beta\Delta). \quad (65)$$

This shows that $l_{\mathrm{coh}}'$ may be regarded as that crystal thickness at which $\langle S\rangle$ changes over from a quadratic dependence on $l$ to a linear dependence. Whether a crystal is thick or thin depends upon the beam spread $\Delta$, which ideally should be as small as possible. In the thick crystal case the distribution of the beam with respect to direction $\delta$ will be much different from that of the laser beam. There will be a marked concentration of intensity along the matching line $\theta=0$ with rapid falling off perpendicular to this direction. The second harmonic beam therefore should have a marked elongation along the matching line.

Ideally it would be desirable for the matching line to shrink to a single point of tangency between the index surface and the effective-index surface. The analysis given above would then not apply for a pencil centered on the matching point. A detailed analysis for this case will not be given here, but it is easy to see that $\varphi$ would vary quadratically with $\delta$. If $\varphi\sim a\delta^2$, the coherence length for the pencil would be $\sim(4c/\omega a\Delta^2)$, which for narrow pencils might be several orders of magnitude longer than (63).

In this section the concept of the effective-index surface has been used as a basis for discussing the analytic behavior of the forced wave as a function of its direction $\mathbf{s}_k$ for $\mathbf{s}_k$ near the matching curve (26). This behavior has then been used to evaluate the average Poynting vector (57) for a narrow pencil containing the matching curve. In this treatment we have assumed that the most important average is over the angular spread of the pencil. There will also be a spread in frequency, which may be written as a spread in $n'$ through (2). A spread in $n'$ can also come from the mixing of 2 beams making a finite angle with $\mathbf{s}_k$ and each having a finite directional spread. The effect of this spread is to smear out the effective-index surface into a thin shell, and the matching curve (26) becomes a thin ribbon on the index surface. The thickness of this ribbon is $(\Delta n_0)(\mathbf{s}_0\cdot\mathbf{N}_k)[1-(\mathbf{N}_s\cdot\mathbf{N}_k)^2]^{-1/2}$, where $(\Delta n_0)$ is a measure of the spread in $n_0=n'(\mathbf{s}_0)$. The treatment of this section is adequate if the ribbon subtends a smaller angle than the beam spread $\Delta$. We estimate that the subtended angle of the ribbon has the order of magnitude $100(\Delta n_0)\sim 100(\Delta\omega/\omega)\sim 10^{-3}$ rad in a typical experimental arrangement employing a cooled ruby laser. This may be compared with a full width $2\Delta\sim 30'$ $\sim 10^{-2}$ rad reported by Giordmaine[4] and by Maker *et al.*[5] A somewhat different situation is obtained, however, if $\mathbf{P}$ arises from the mixing of two beams meeting at a finite angle. Then the angular spread of each beam gives rise to a spread in $n'$ as well as a spread in the direction $\mathbf{s}_k$. The results of this section may not apply quantitatively in this case.

## IV. SURFACE WAVES

In this section the complete solution to (12) will be obtained for the case of a semi-infinite crystal having a single plane surface. The laser beam is incident on this surface and produces a polarization wave (11) in the crystal such that $\mathbf{N}\cdot\mathbf{s}_k>0$. The solution inside the crystal then consists of the forced wave and the two free waves $n$, $\mathbf{s}$ which satisfy (27). We shall denote these free waves by $n_\alpha$, $\mathbf{s}_\alpha$ and $n_\beta$, $\mathbf{s}_\beta$. There will also be a wave in the vacuum outside the crystal with wave direction $\mathbf{s}_v$ determined by (10). The field in the crystal may be written

$$\mathbf{E}(\mathbf{r})=\mathbf{F}e^{i(\omega/c)n'\mathbf{s}_k\cdot\mathbf{r}}+E_\alpha\mathbf{U}_\alpha^{(1)}e^{i(\omega/c)n_\alpha\mathbf{s}_\alpha\cdot\mathbf{r}}+E_\beta\mathbf{U}_\beta^{(1)}e^{i(\omega/c)n_\beta\mathbf{s}_\beta\cdot\mathbf{r}}, \quad (66)$$

where $\mathbf{F}$ represents the forced wave. In the nonmatching case one could choose $\mathbf{F}=\mathbf{E}'$ given by (16). In the nearly matching case $\mathbf{s}_k=\mathbf{s}_M=\mathbf{s}_\alpha\neq\mathbf{s}_\beta$ and $\mathbf{F}$ may be obtained from (44). The unit vectors $\mathbf{U}_\alpha^{(1)}$, $\mathbf{U}_\beta^{(1)}$ defined by (19) and (20) represent the electric field directions for the free waves. The field in the vacuum outside the surface is

$$\mathbf{E}_v(\mathbf{r})=\mathbf{E}_v e^{i(\omega/c)\mathbf{s}_v\cdot\mathbf{r}}. \quad (67)$$

The coefficients $E_\alpha$, $E_\beta$ and the vector $\mathbf{E}_v$ are to be determined by the condition that the tangential components of the electric and magnetic fields be continuous across the surface $\mathbf{N}\cdot\mathbf{r}=0$. It is sufficient to satisfy this condition at a single point $\mathbf{r}=0$, since (27) insures that the condition will then hold at every point on the surface. In addition the wave in vacuum must be the transverse:

$$\mathbf{E}_v\cdot\mathbf{s}_v=0. \quad (68)$$

The continuity conditions may be written

$$\begin{aligned}\mathbf{N}\times\mathbf{E}(\mathbf{r})&=\mathbf{N}\times\mathbf{E}_v(\mathbf{r}),\\ \mathbf{N}\times\nabla\times\mathbf{E}(\mathbf{r})&=\mathbf{N}\times\nabla\times\mathbf{E}_v(\mathbf{r})\quad\text{at}\quad\mathbf{r}=0.\end{aligned} \quad (69)$$

Therefore, from (66), (67), and (69),

$$\mathbf{N}\times\mathbf{E}_v=E_\alpha\mathbf{N}\times\mathbf{U}_\alpha^{(1)}+E_\beta\mathbf{N}\times\mathbf{U}_\beta^{(1)}+\mathbf{N}\times\mathbf{F}, \quad (70)$$

$$\begin{aligned}\mathbf{N}\times(\mathbf{s}_v\times\mathbf{E}_v)=&E_\alpha n_\alpha\mathbf{N}\times(\mathbf{s}_\alpha\times\mathbf{U}_\alpha^{(1)})+E_\beta n_\beta\mathbf{N}\times(\mathbf{s}_\beta\times\mathbf{U}_\beta^{(1)})\\ &+n'\mathbf{N}\times(\mathbf{s}_k\times\mathbf{F})-i(c/\omega)\mathbf{N}\times\nabla\times\mathbf{F},\end{aligned} \quad (71)$$

where $\mathbf{F}$ is to be evaluated at $\mathbf{r}=0$ if it depends on $\mathbf{r}$. These equations may be solved together with (68) as follows: Scalar multiplication of (70) with $\mathbf{U}_\alpha^{(1)}$ and $\mathbf{U}_\beta^{(1)}$ gives

$$\begin{aligned}E_\alpha&=(\mathbf{U}_\beta^{(1)}\times\mathbf{N}\cdot\mathbf{E}_v-\mathbf{U}_\beta^{(1)}\cdot\mathbf{N}\times\mathbf{F})/(\mathbf{U}_\beta^{(1)}\cdot\mathbf{N}\times\mathbf{U}_\alpha^{(1)}),\\ E_\beta&=(\mathbf{U}_\alpha^{(1)}\times\mathbf{N}\cdot\mathbf{E}_v-\mathbf{U}_\alpha^{(1)}\cdot\mathbf{N}\times\mathbf{F})/(\mathbf{U}_\alpha^{(1)}\cdot\mathbf{N}\times\mathbf{U}_\beta^{(1)});\end{aligned} \quad (72)$$

scalar multiplication of (71) with $\mathbf{s}_\alpha\times\mathbf{U}_\alpha^{(1)}$ and $\mathbf{s}_\beta\times\mathbf{U}_\beta^{(1)}$ followed by use of (72) to eliminate $E_\alpha$ and $E_\beta$ gives

$$\mathbf{Q}_\alpha\cdot\mathbf{E}_v=p_\beta,\quad \mathbf{Q}_\beta\cdot\mathbf{E}_v=p_\alpha, \quad (73)$$

where the following abbreviations have been used:

$$\begin{aligned}p_\alpha&=\mathbf{s}_\beta\times\mathbf{U}_\beta^{(1)}\cdot[n'\mathbf{N}\times(\mathbf{s}_k\times\mathbf{F})-i(c/\omega)\mathbf{N}\times\nabla\times\mathbf{F}]-n_\alpha\gamma\mathbf{U}_\beta^{(1)}\cdot\mathbf{N}\times\mathbf{F},\\ p_\beta&=\mathbf{s}_\alpha\times\mathbf{U}_\alpha^{(1)}\cdot[n'\mathbf{N}\times(\mathbf{s}_k\times\mathbf{F})-i(c/\omega)\mathbf{N}\times\nabla\times\mathbf{F}]-n_\beta\gamma\mathbf{U}_\alpha^{(1)}\cdot\mathbf{N}\times\mathbf{F},\\ \gamma&=\mathbf{s}_\alpha\times\mathbf{U}_\alpha^{(1)}\cdot\mathbf{N}\times(\mathbf{s}_\beta\times\mathbf{U}_\beta^{(1)})/(\mathbf{U}_\alpha^{(1)}\cdot\mathbf{N}\times\mathbf{U}_\beta^{(1)}),\\ \mathbf{Q}_\alpha&=[(\mathbf{s}_\alpha\times\mathbf{U}_\alpha^{(1)})\times\mathbf{N}]\times\mathbf{s}_v-n_\beta\gamma\mathbf{U}_\alpha^{(1)}\times\mathbf{N},\\ \mathbf{Q}_\beta&=[(\mathbf{s}_\beta\times\mathbf{U}_\beta^{(1)})\times\mathbf{N}]\times\mathbf{s}_v-n_\alpha\gamma\mathbf{U}_\beta^{(1)}\times\mathbf{N};\end{aligned} \quad (74)$$

from (68) and (73)

$$\mathbf{E}_v=\mathbf{s}_v\times(p_\beta\mathbf{Q}_\beta-p_\alpha\mathbf{Q}_\alpha)/(\mathbf{s}_v\cdot\mathbf{Q}_\beta\times\mathbf{Q}_\alpha). \quad (75)$$

The formal solution is now given by (72) and (75). The waves associated with $E_\alpha$, $E_\beta$ and $\mathbf{E}_v$ are conveniently called *surface waves*, since they arise from the continuity conditions required at the surface.

It is clear that $\mathbf{s}_v\cdot\mathbf{Q}_\beta\times\mathbf{Q}_\alpha\neq 0$, since otherwise one could set $\mathbf{F}=0$ and still find a nonvanishing $\mathbf{E}_v$. It is possible, however, to have $\mathbf{U}_\alpha^{(1)}\cdot\mathbf{N}\times\mathbf{U}_\beta^{(1)}\rightarrow 0$, in which case $\mathbf{E}_v$ is determined by (68) and (70). The correct answer is readily obtained from (75) by letting $\gamma\rightarrow\infty$. It can be shown that the numerators in (72) also vanish in this limit making these formulas indeterminate, but expressions for $E_\alpha$, $E_\beta$ in this case may be obtained from (71).

In the nonmatching case $\mathbf{F}$ can be chosen a constant so that $\nabla\times\mathbf{F}=0$. It is always permissible to add a free wave to the forced wave, which corresponds to the transformation

$$\mathbf{F}\rightarrow\mathbf{F}+F_\alpha\mathbf{U}_\alpha^{(1)}e^{i(\omega/c)(n_\alpha\mathbf{s}_\alpha-n'\mathbf{s}_k)\cdot\mathbf{r}}+F_\beta\mathbf{U}_\beta^{(1)}e^{i(\omega/c)(n_\beta\mathbf{s}_\beta-n'\mathbf{s}_k)\cdot\mathbf{r}}, \quad (76)$$

and $F$ then depends on $\mathbf{r}$. We should expect that such a transformation would merely change $E_\alpha$, $E_\beta$:

$$E_\alpha \to E_\alpha - F_\alpha, \quad E_\beta \to E_\beta - F_\beta, \tag{77}$$

so as to cancel the additional free waves in (66). This indeed follows from (72) when it is observed that $p_\alpha$, $p_\beta$, and hence $\mathbf{E}_v$, are invariant to the transformation (76). Although $\mathbf{F}$, as given by (16), has a singularity in the matching limit, there is no singularity in $\mathbf{E}_v$, since the procedure of the preceding section which eliminates the singularity in $\mathbf{F}$ is equivalent to a transformation of the form (76). This may be of experimental importance in connection with obtaining reliable measurements of the second-order polarization tensor $d$, since the vacuum wave should be quite insensitive to the coherence length. It is of interest to consider the free waves as matching conditions are approached. All singularities are, of course, avoided if $\mathbf{F}$ is taken to be

$$\mathbf{F}(\mathbf{r}) = \mathbf{E}_1(\mathbf{N}\cdot\mathbf{r})g(2i\psi) + \mathbf{E}_{2,3}, \tag{78}$$

according to (44), but it is more instructive to let $\mathbf{F}$ be given by (16). It may be recalled that in the derivation of (44) and in the discussion of the matching case (21) the presence of a surface is implied by the vector $\mathbf{N}$, although at that time no boundary conditions were explicitly considered. We may now ask whether the detailed theory of the surface waves can provide a physical explanation for the free wave needed in (29) to eliminate the singularity in (16) at the matching limit. Therefore, let $\mathbf{F}$ be given by (16), let $n'\mathbf{s}_k$ approach $n_\alpha\,\mathbf{s}_\alpha$, and consider the term in $E_\alpha$ which becomes singular; from (72) this term is

$$-\left[\frac{\mathbf{U}_\beta^{(1)}\cdot\mathbf{N}\times\mathbf{U}_k^{(1)}}{\mathbf{U}_\beta^{(1)}\cdot\mathbf{N}\times\mathbf{U}_\alpha^{(1)}}\right]\frac{4\pi\mathbf{P}\cdot\mathbf{U}_k^{(1)}}{\lambda_k^{(1)}}\mathbf{U}_\alpha^{(1)}. \tag{79}$$

The factor in brackets approaches unity in the limit $\lambda_k^{(1)} \to 0$. It follows immediately that the singular part of $E_\alpha$ is exactly the free wave assumed in (29), and furthermore, no singularity arises in $E_\beta$ since $\mathbf{U}_\alpha^{(1)}\times\mathbf{N}\cdot\mathbf{U}_\alpha^{(1)}=0$. The theories of this section and the preceding one are, therefore, consistent if $\mathbf{N}$ is interpreted as the surface normal.

The problem of this section is formally equivalent to the problem of finding the reflected and transmitted waves when a plane wave is incident on the plane surface of an anisotropic medium. The transmitted waves are given by (72) and the reflected wave by (75) for an incident wave

$$-\mathbf{F}e^{i(\omega/c)\mathbf{s}_k\cdot\mathbf{r}}, \tag{80}$$

with $n'=1$ in (73) corresponding to free space. The forced wave is therefore analogous to an incident wave in ordinary optics and produces an analogous result. Thus (72) and (75) constitute generalized Fresnel formulas[17] which reduce to the familiar form in the special case of an isotropic medium.

[17] Reference 10, p. 39.

In a practical experiment to observe SHG, the sample will ordinarily have two parallel surfaces. The second surface produces surface waves which can be treated by the methods of this section, although the general solution is quite complicated and will not be given here. Of primary interest is the vacuum wave $\mathbf{E}_v'$, $\mathbf{s}_v'$ emerging from the second surface. Its direction

$$\mathbf{s}_v' = \mathbf{s}_v - 2(\mathbf{N}\cdot\mathbf{s}_v)\mathbf{N} \tag{81}$$

is the reflection of $\mathbf{s}_v$ in the surface. The same vector $\mathbf{N}$ may be used as the normal for both surfaces since the theory is independent of the sign of $\mathbf{N}$. If the laser beam is a pencil containing the matching line $\varphi=0$ [see discussion following (47)], all waves but the growing wave may be neglected. Thus we may take

$$\mathbf{F} = E_1\mathbf{U}_\alpha^{(1)}(\mathbf{N}\cdot\mathbf{r})g(2i\psi), \tag{82}$$

an approximation already introduced in (53). It is also permissible to neglect $\nabla\times\mathbf{F}$ in (74), since this is of the same order as the neglected waves. The free waves $\mathbf{U}_\alpha^{(1)}$, $\mathbf{s}_\alpha^{(1)}$ and $\mathbf{U}_\beta^{(1)}$, $\mathbf{s}_\beta^{(1)}$ in (66) must be replaced with two other free waves satisfying (27) but propagating into the crystal from the second surface. Also $\mathbf{s}_v$ must be replaced with $\mathbf{s}_v'$. When all these replacements have been made, (75) gives $\mathbf{E}_v'$. It may be expected that if the beam is not too far from normal incidence ($\mathbf{s}_v\sim\mathbf{N}$), the transmission coefficient for the intensity will be close to unity.

## V. APERTURE EFFECTS

We shall call the diameter of the laser beam in the crystal the *aperture a* of the system. The finite aperture gives rise to two kinds of effects, the more familiar of which is diffraction. In so far as diffraction causes a spreading of the beam to form a pencil of waves it has already been taken into account in Sec. III. Diffraction will be considered again in the last section in connection with focussed beams. The other effect is due to the anisotropy of the medium, which causes the energy in a wave to be propagated in a slightly different direction than the wave direction $\mathbf{s}$.[18] This will be called the *aperture effect.*

When the coherence length is very long in a thick crystal the aperture effect can limit the intensity of the SHG. The laser beam enters the crystal as a pencil of waves of diameter $a$ propagating in the direction $\mathbf{s}_k$. The energy in this beam propagates in a slightly different direction $\mathbf{t}_k$. If $\mathbf{s}_k=\mathbf{s}$ is a matching direction, a linearly growing second harmonic amplitude is generated. The second harmonic energy propagates in the direction $\mathbf{t}$, and in general $\mathbf{t}\neq\mathbf{s}\neq\mathbf{t}_k$. It is physically evident that SHG occurs only within the pencil of $\mathbf{t}_k$. The effective coherence volume in the matching case is not infinite, but is no greater than the volume which is common to the pencils of $\mathbf{t}_k$ and $\mathbf{t}$. The maximum path

[18] Reference 10, p. 665.

of the ray $\mathbf{t}$ within the pencil of $\mathbf{t}_k$ is

$$l_{\max}=a[1-(\mathbf{t}\cdot\mathbf{t}_k)^2]^{-1/2}, \tag{83}$$

which may be taken as a measure of the maximum effective coherence length. If $l_{\max}$ is less than the crystal thickness the entire pencil of $\mathbf{t}_k$ generates second harmonic radiation, but the growing wave $E_1(\mathbf{N}\cdot\mathbf{r})$ in (21) never achieves an amplitude greater than $E_1 l_{\max}(\mathbf{N}\cdot\mathbf{s}_k)$.

We consider briefly the magnitude of the effect in potassium dihydrogen phosphate (KDP). In the experiments,[4,5] where matching has been observed, to be discussed more fully in Sec. VII, the laser beam passed through the crystal as an ordinary wave. This wave is transverse and hence $\mathbf{t}_k=\mathbf{s}$. The second harmonic is produced as an extraordinary ray, which is not transverse, and hence $\mathbf{t}\neq\mathbf{t}_k$. For *uniaxial* crystals a formula can readily be derived giving $\mathbf{t}\cdot\mathbf{s}$ for the extraordinary wave

$$\mathbf{t}\cdot\mathbf{s}=\frac{1+\xi^2\tan^2\psi}{(1+\tan^2\psi)^{\frac{1}{2}}(1+\xi^4\tan^2\psi)^{\frac{1}{2}}}, \tag{84}$$

where $\xi=n_\epsilon/n_\omega$ measures the anisotropy and $\psi$ is the angle between $\mathbf{s}$ and the optic axis. For most crystals the anisotropy is small and $\xi$ is very nearly unity. If $|\xi-1|\ll1$, the angle between $\mathbf{t}$ and $\mathbf{s}$ is given by

$$\angle(\mathbf{t},\mathbf{s})=2(\xi-1)\tan\psi/(1+\tan^2\psi)+\cdots, \tag{85}$$

which is maximum for $\psi=45°$. In KDP,[4] the matching condition is $\psi=50°$, and $\xi-1=0.0316$, so that (85) gives $\angle(\mathbf{t},\mathbf{s})=0.0312$ rad$=1.79°$. If $a=0.5$ cm and incidence is normal (83) gives $l_{\max}=16$ cm. This is to be compared with the coherence length for the laser beam $l_{\rm coh}'\sim0.1$ cm and the crystal thickness $\sim0.2$ cm. We may conclude that the aperture effect may someday limit SHG if laser beams can be produced which are narrower in angular spread than present beams by two orders of magnitude and correspondingly narrow in frequency.

## VI. NEARLY ISOTROPIC MEDIA

In this section we shall consider the simplifications in the theory which come from neglecting all the unimportant effects of the anisotropy. An approximation of this kind has already been introduced in connection with the Poynting vector (54). The justification for this approach is the small anisotropy found in most crystals. Typically the anisotropies of crystals are of the order of a few percent, as we noted for KDP in the preceding section. This does not mean that the anisotropy is entirely unimportant. It seems probable that SHG experiments giving large outputs will make essential use of anisotropy for index matching. This is because matching in an isotropic medium is only possible if there is anomalous dispersion somewhere in the range $\omega/2$ to $\omega$. The anomalous dispersion would imply the presence of a significant absorption at $\omega/2$, or $\omega$, or both. An absorption at $\omega/2$ would limit the laser power that could safely be used without burning the crystal, while absorption at $\omega$ would prevent observation of SHG. Therefore, we shall regard matching as primarily an anisotropic effect. Likewise the direction of polarization of the second harmonic waves will be determined by the anisotropy, even though it may be extremely small. These are the essential effects of anisotropy which must be retained in the theory, while at the same time treating the problem in all other respects as if the medium were isotropic. This nearly isotropic approximation should be valid for most crystals, including, of course, the isotropic crystals for which it becomes exact. Recently SHG has been observed[6] in the isotropic crystals $NaClO_3$ (symmetry $T$) and ZnS (Td).

First consider an isotropic medium for which

$$\epsilon=n^2I \tag{86}$$

$$\alpha_s=-n^2\mathbf{ss}, \tag{87}$$

where $n^2$ is a constant. The eigenvalues of $\alpha_s$ are

$$\lambda_s^{(1)}=\lambda_s^{(2)}=0,\quad \lambda_s^{(3)}=-n^2, \tag{88}$$

and the eigenvectors are described by

$$\mathbf{s}\cdot\mathbf{U}_s^{(1)}=\mathbf{s}\cdot\mathbf{U}_s^{(2)}=0,\quad \mathbf{U}_s^{(3)}=\mathbf{s}. \tag{89}$$

The index surface consists of a single sheet, a sphere of radius $n$; nevertheless, there are two free waves belonging to any direction $\mathbf{s}$, since both $\mathbf{U}_s^{(1)}$ and $\mathbf{U}_s^{(2)}$ are now allowed directions for $\mathbf{E}$. Similarly, the eigenvalues of $\alpha_k$ are

$$\lambda_k^{(1)}=\lambda_k^{(2)}=n'^2-n^2,\quad \lambda_k^{(3)}=-n^2, \tag{90}$$

and the eigenvectors satisfy (89) with $\mathbf{s}$ replaced by $\mathbf{s}_k$. The plane forced wave (16) is

$$\mathbf{E}'=4\pi\mathbf{P}_t/(n'^2-n^2)-4\pi\mathbf{P}_l/n^2, \tag{91}$$

where $\mathbf{P}_t$ and $\mathbf{P}_l$ are the transverse and longitudinal parts of $\mathbf{P}$, respectively. For most purposes it is more useful to have the forced wave in the form (44) with the singularity at $n'=n$ removed. It has been shown in Sec. IV that this is equivalent to combining $E'$ with the reflected wave from the surface of the medium.

We now suppose the medium has a small anisotropy

$$\epsilon=n^2I+\delta\epsilon, \tag{92}$$

where $\delta\epsilon$ is a small dyadic with zero diagonal sum. For any $\mathbf{s}$ there are now two free waves corresponding to the two sheets of the index surface. These waves may be designated $n_\alpha$, $\mathbf{s}$, $\mathbf{U}_\alpha^{(1)}(\mathbf{s})$ and $n_\beta$, $\mathbf{s}$, $\mathbf{U}_\beta^{(1)}(\mathbf{s})$, and the transverse parts of $\mathbf{U}_\alpha^{(1)}$, $\mathbf{U}_\beta^{(1)}$ may be written $\mathbf{U}_{\alpha t}^{(1)}$, $\mathbf{U}_{\beta t}^{(1)}$. Then it follows from (7), and the fact that $\epsilon$ is symmetric, that $\mathbf{U}_{\alpha t}^{(1)}$ and $\mathbf{U}_{\beta t}^{(1)}$ are orthogonal if $n_\alpha\neq n_\beta$,

$$(n_\alpha^2-n_\beta^2)\mathbf{U}_{\alpha t}^{(1)}\cdot\mathbf{U}_{\beta t}^{(1)}=0. \tag{93}$$

From (7) and (92) it may readily be shown that the longitudinal components of $\mathbf{U}_\alpha^{(1)}$ and $\mathbf{U}_\beta^{(1)}$ are given to first order in $\delta\epsilon$ by expressions of the form

$$U_{\alpha l}^{(1)}=-n^{-2}\mathbf{s}\cdot\delta\epsilon\cdot\mathbf{U}_{\alpha t}^{(1)}. \quad (94)$$

Similarly, it may be shown that the transverse components of $\mathbf{U}_\alpha^{(3)}$, $\mathbf{U}_\beta^{(3)}$ are small quantities of order $\delta\epsilon$, and the corresponding eigenvalues are given by

$$\lambda_\alpha^{(3)}=\lambda_\beta^{(3)}=-n^2-\mathbf{s}\cdot\delta\epsilon\cdot\mathbf{s} \quad (95)$$

to order $\delta\epsilon$. To the same order

$$\lambda_\alpha^{(2)}=-\lambda_\beta^{(2)}=n_\alpha^2-n_\beta^2. \quad (96)$$

It is apparent that it is unnecessary to deal with two sets of eigenvalues and eigenvectors if components of order $\delta\epsilon$ may be neglected. We introduce a single set of unit eigenvectors $\mathbf{U}_s^{(i)}$ defined by

$$\mathbf{U}_s^{(1)}=\mathbf{U}_{\alpha t}^{(1)}(\mathbf{s}),\quad \mathbf{U}_s^{(2)}=\mathbf{U}_{\beta t}^{(1)}(\mathbf{s}),\quad \mathbf{U}_s^{(3)}=\mathbf{s}, \quad (97)$$

where it is understood that $\mathbf{U}_{\alpha t}^{(1)}(\mathbf{s})$ and $\mathbf{U}_{\beta t}^{(1)}(\mathbf{s})$ are to be raised slightly to unit length. These vectors satisfy the conditions (89) for an isotropic medium. Light may propagate with its electric vector along $\mathbf{U}_s^{(1)}$ or $\mathbf{U}_s^{(2)}$ with refractive index $n_1=n_\alpha$ or $n_2=n_\beta$, respectively. Similarly the eigenvalues of $\alpha_k$ for the nearly isotropic medium are

$$\lambda_k^{(1)}=(n'^2-n_1^2),\quad \lambda_k^{(2)}=(n'^2-n_2^2),\quad \lambda_k^{(3)}=-n^2, \quad (98)$$

and the eigenvectors are

$$\mathbf{U}_k^{(1)}=\mathbf{U}_{\alpha t}^{(1)}(\mathbf{s}_k),\quad \mathbf{U}_k^{(2)}=\mathbf{U}_{\beta t}^{(1)}(\mathbf{s}_k),\quad \mathbf{U}_k^{(3)}=\mathbf{s}_k. \quad (99)$$

We now consider the nearly matching case defined in Sec. III, where it is shown that the singularity at $\lambda_k^{(1)}=0$ can be removed by adding a suitable free wave. It is evident from (98) that both $\lambda_k^{(1)}=0$ and $\lambda_k^{(2)}$ may be very small, and, therefore, it would be desirable to go one step further than was done in (30) and eliminate the term in $(1/\lambda_k^{(2)})$. This leads to considerable complications in the general case, but can be done very simply in the nearly isotropic approximation. This is because $E_2$ given in (25) does not depend on $E_1$ if the longitudinal components of $\mathbf{U}_M^{(1)}$ and $\mathbf{U}_M^{(2)}$ are neglected. The relation between $\mathbf{s}_k$ and $\mathbf{s}$ is determined by the index surface and the surface normal vector $\mathbf{N}$. Since we are taking into account both free waves instead of only one as in Sec. III, there will be two relations of the form (28)

$$\begin{aligned} n'\mathbf{s}_k-n_1\mathbf{s}_1&=\varphi_1\mathbf{N},\\ n'\mathbf{s}_k-n_2\mathbf{s}_2&=\varphi_2\mathbf{N}. \end{aligned} \quad (100)$$

For the nearly matching case,

$$|n'-n|\ll 1, \quad (101)$$

it follows from (100) that

$$\begin{aligned} \varphi_1&=(n'-n_1)(\mathbf{N}\cdot\mathbf{s}_k)^{-1},\\ \varphi_2&=(n'-n_2)(\mathbf{N}\cdot\mathbf{s}_k)^{-1}. \end{aligned} \quad (102)$$

Once $\varphi_1$ and $\varphi_2$ have been obtained, it is unnecessary to distinguish between $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_k$. The appropriate forced wave is

$$\begin{aligned} \mathbf{E}_k(\mathbf{r})=[&E_1(\mathbf{N}\cdot\mathbf{r})g(2i\psi_1)\mathbf{U}_k^{(1)}\\ &+E_2(\mathbf{N}\cdot\mathbf{r})g(2i\psi_2)\mathbf{U}_k^{(2)}+E_3\mathbf{U}_k^{(3)}]e^{i(\omega/c)n'\mathbf{s}_k\cdot\mathbf{r}}, \end{aligned} \quad (103)$$

where

$$\begin{aligned} E_1&=2\pi i(\omega/cn)(\mathbf{N}\cdot\mathbf{s}_k)^{-1}(\mathbf{P}\cdot\mathbf{U}_k^{(1)}),\\ E_2&=2\pi i(\omega/cn)(\mathbf{N}\cdot\mathbf{s}_k)^{-1}(\mathbf{P}\cdot\mathbf{U}_k^{(2)}),\\ E_3&=-(4\pi/n^2)(\mathbf{P}\cdot\mathbf{U}_k^{(3)}), \end{aligned} \quad (104)$$

and

$$\begin{aligned} \psi_1&=(\omega/2c)(\mathbf{N}\cdot\mathbf{s}_k)^{-1}(n'-n_1)\mathbf{N}\cdot\mathbf{r},\\ \psi_2&=(\omega/2c)(\mathbf{N}\cdot\mathbf{s}_k)^{-1}(n'-n_2)\mathbf{N}\cdot\mathbf{r}. \end{aligned} \quad (105)$$

Coherence lengths can be defined for the $E_1$ and $E_2$ waves according to (46):

$$\begin{aligned} l_{\text{coh}}^{(1)}&=(2\pi c/\omega)/|n'-n_1|,\\ l_{\text{coh}}^{(2)}&=(2\pi c/\omega)/|n'-n_2|. \end{aligned} \quad (106)$$

The energy flow in the $E_1$ and $E_2$ waves may be written according to (55)

$$\begin{aligned} S_1&=(\pi/2cn)(\mathbf{P}\cdot\mathbf{U}_k^{(1)})^2\omega^2l^2(\sin^2\psi_1)/\psi_1^2(\mathbf{N}\cdot\mathbf{s}_k)^2,\\ S_2&=(\pi/2cn)(\mathbf{P}\cdot\mathbf{U}_k^{(2)})^2\omega^2l^2(\sin^2\psi_2)/\psi_2^2(\mathbf{N}\cdot\mathbf{s}_k)^2, \end{aligned} \quad (107)$$

For the isotropic medium, (103) and (104) may be combined to yield

$$\mathbf{E}_k(\mathbf{r})=[2\pi i(\omega/cn)(\mathbf{N}\cdot\mathbf{s}_k)^{-1}(\mathbf{N}\cdot\mathbf{r})g(2i\psi)\mathbf{P}_t-4\pi\mathbf{P}_l/n^2]\times e^{i(\omega/c)n'\mathbf{s}_k\cdot\mathbf{r}}, \quad (108)$$

and (107) becomes

$$S=(\pi/2cn)P_t^2\omega^2l^2(\sin^2\psi)/\psi^2(\mathbf{N}\cdot\mathbf{s}_k)^2. \quad (109)$$

For the consideration of surface waves discussed in Sec. IV, it is sufficient to neglect the anisotropy completely. This is because matching has no effect on surface waves, and $\mathbf{F}$ of (66) may be taken as $E'$ of (91). It is convenient to choose for $\mathbf{U}_\alpha^{(1)}$, $\mathbf{U}_\beta^{(1)}$

$$\mathbf{U}_\alpha^{(1)}=\mathbf{U}_\perp,\quad \mathbf{U}_\beta^{(1)}=\mathbf{U}_\parallel,\quad \mathbf{U}_\perp\times\mathbf{U}_\parallel=\mathbf{s}, \quad (110)$$

where $\mathbf{U}_\perp$ and $\mathbf{U}_\parallel$ are perpendicular and parallel, respectively, to the plane of incidence, the plane of $\mathbf{N}$, $\mathbf{s}$, $\mathbf{s}_k$. The perpendicular and parallel components of the electric field of the surface wave $\mathbf{s}$ will be denoted $E_\perp$ and $E_\parallel$. Similarly we express $\mathbf{E}_v$ in terms of perpendicular and parallel components $E_{v\perp}$, $E_{v\parallel}$, where $E_{v\parallel}$ is measured along the direction $\mathbf{s}_v\times\mathbf{U}_\perp$. Finally $\mathbf{F}$ is written in terms of its components $F_\perp$, $F_\parallel$, $F_l$, where $F_\parallel$ is measured along $\mathbf{s}_k\times\mathbf{U}_\perp$ and $F_l$ is the longitudinal component. Let $\theta_s$, $\theta_k$, $\theta_v$ be the angles measured from $\mathbf{N}$ of $\mathbf{s}$, $\mathbf{s}_k$, $\mathbf{s}_v$, respectively. Then for an isotropic medium, (74) becomes

$$\begin{aligned} p_\alpha&=p_\perp=F_\perp(n'\cos\theta_k-n\cos\theta_s),\\ p_\beta&=p_\parallel=F_\parallel(n\cos\theta_k-n'\cos\theta_s)-F_ln\sin\theta_k,\\ \gamma&=1,\\ \mathbf{Q}_\alpha&=\mathbf{Q}_\parallel=\mathbf{U}_\perp\times\mathbf{N}(\cos\theta_v\cos\theta_s-n)-\mathbf{N}\sin\theta_v\cos\theta_s,\\ \mathbf{Q}_\beta&=\mathbf{Q}_\perp=\mathbf{U}_\perp(\cos\theta_v-n\cos\theta_s). \end{aligned} \quad (111)$$

Other useful relations are

$$\mathbf{U}_{\parallel}\cdot\mathbf{N}\times\mathbf{U}_{\perp}=\cos\theta_s, \tag{112}$$
$$\mathbf{s}_v\cdot\mathbf{Q}_{\perp}\times\mathbf{Q}_{\parallel}=(\cos\theta_v-n\cos\theta_s)(n\cos\theta_v-\cos\theta_s),$$

and from (10) and (27) Snell's law

$$n'\sin\theta_k=n\sin\theta_s=\sin\theta_v. \tag{113}$$

The vacuum wave (75) becomes

$$E_{v\perp}=F_{\perp}(n'\cos\theta_k-n\cos\theta_s)/(\cos\theta_v-n\cos\theta_s),$$
$$E_{v\parallel}=F_{\parallel}(n\cos\theta_k-n'\cos\theta_s)/(n\cos\theta_v-\cos\theta_s)$$
$$-F_l n\sin\theta_k/(n\cos\theta_v-\cos\theta_s), \tag{114}$$

and the free wave in the medium (72) becomes

$$E_{\perp}=F_{\perp}(n'\cos\theta_k-\cos\theta_v)/(\cos\theta_v-n\cos\theta_s),$$
$$E_{\parallel}=F_{\parallel}(\cos\theta_k-n'\cos\theta_v)/(n\cos\theta_v-\cos\theta_s)$$
$$-F_l\sin\theta_k/(n\cos\theta_v-\cos\theta_s). \tag{115}$$

The usual form of Fresnel's equations[17] is obtained from (114) and (115) by setting $F_l=0$, $n'=1$, and interpreting $-\mathbf{F}$ as the incident wave according to (80).

A more convenient form for (114) is obtained by taking the limit $n'\rightarrow n$, since $|n'-n|$ will ordinarily be small; from (91) and (114) the vacuum wave then becomes

$$E_{v\perp}=-2\pi P_{\perp}[(n\cos\theta_k-\cos\theta_v)n\cos\theta_k]^{-1},$$
$$E_{v\parallel}=-2\pi P_{\parallel}[(n\cos\theta_v-\cos\theta_k)n\cos\theta_k]^{-1}\cos2\theta_k$$
$$+4\pi P_l\sin\theta_k[(n\cos\theta_v-\cos\theta_k)n]^{-1}. \tag{116}$$

Since $\mathbf{E}_v$ is insensitive to the matching condition as pointed out in Sec. IV, and in practice $|n'-n|$ should be only a few percent, it is expected that (116) should be adequate for most situations. At normal incidence, $\cos\theta_k=-\cos\theta_v=1$, it reduces to

$$\mathbf{E}_v=-2\pi\mathbf{P}_t/(n^2+n), \tag{117}$$

which shows that the vacuum wave is of the same order of magnitude as the longitudinal wave in (91) or a badly mismatched wave in a highly anisotropic crystal.

## VII. APPLICATIONS TO KDP AND QUARTZ

Consider first the experiment of Giordmaine[4] on potassium dihydrogen phosphate (KDP). The laser beam passed through the crystal as an ordinary ray. The effective-index surface is then a sphere of radius[19] 1.506, so the matching index is $n_0=1.506$. If the ordinary and extraordinary refractive indices at the second harmonic frequency are denoted by $n_\omega$, $n_\epsilon$, and if $\theta$ is the angle between $\mathbf{s}$ and the optic axis, the equation for the index surface is

$$n^{-2}=n_\omega^{-2}\cos^2\theta+n_\epsilon^{-2}\sin^2\theta, \tag{118}$$

which represents an ellipsoid of revolution. For KDP $n_\omega=1.534$, $n_\epsilon=1.487$, and the matching angle for which $n=n_0$ is $\theta_M=49.9°$. The angle $\rho$ between the normals of the index surface and the effective-index surface is $\rho=1.8°$. From (51) and (59) with $\mathbf{N}\cdot\mathbf{N}_s=\mathbf{s}_0\cdot\mathbf{N}_k=1$ we obtain $\beta=4.2\times10^3$ cm$^{-1}$. The laser beam typically has a width $\Delta=\frac{1}{4}°=4.4\times10^{-3}$ rad. Thus the coherence length for the laser pencil in this experiment was

$$l_{\text{coh}}'\sim0.11\text{ cm}, \tag{119}$$

which may be compared with the crystal thickness $l=0.22$ cm. This corresponds roughly to the thick crystal case (65), but it is a little more accurate to use (58) and (62); from (62) $F(\beta l\Delta)=F(4.0)=0.42$. The second harmonic polarization may be written

$$P_z=2d_{36}E_xE_y,\quad P_x=P_y=0. \tag{120}$$

The notation here is similar to the usual piezoelectric notation[7]: $\mathbf{EE}$ in (1) is regarded as a column vector with components $E_x^2$, $E_y^2$, $E_z^2$, $2E_yE_z$, $2E_xE_z$, $2E_xE_y$, and $d$ is a $3\times6$ matrix, which for KDP has only the components $d_{14}$, $d_{25}=d_{14}$, and $d_{36}$. Since only the second harmonic extraordinary wave is matched, we regard $E_1$ in (104) as the extraordinary wave and ignore $E_2$ and $E_3$. Thus

$$\mathbf{P}\cdot\mathbf{U}_k^{(1)}=P_z\sin\theta_M, \tag{121}$$

and we assume $\mathbf{N}\cdot\mathbf{s}_k=1$. From (58) and (107)

$$S=E_x^2E_y^2(2\pi\omega^2l^2/cn_0)d_{36}^2\sin^2\theta_MF(\beta l\Delta). \tag{122}$$

In the optimum orientation $E_x=E_y$, so that $4E_x^2E_y^2=(8\pi/cn_0)^2S_L^2$, where $S_L$ is the laser intensity inside the crystal. Thus (122) becomes

$$S=S_L^2(32\pi^3\omega^2l^2/c^3n_0^3)d_{36}^2\sin^2\theta_MF(\beta l\Delta). \tag{123}$$

The intensities actually measured are outside the crystal. It would be possible to rewrite (123) to include a transmission factor so that $S$ and $S_L$ could be interpreted as outside intensities. For a crystal in air with smooth flat surfaces the appropriate factor would be $64n^3/(n+1)^6$. For $n=1.5$ this factor has the value 0.885, which is so close to unity, in view of the uncertainties in experiments of this kind, that $S$ and $S_L$ in (123) may be interpreted as measured intensities with negligible error. Giordmaine[20] has observed a second harmonic power of $\frac{1}{2}$ mW with a laser power of 3 kW and a beam area of 0.2 cm$^2$; thus, $S\sim2.5\times10^4$ esu and $S_L\sim1.5\times10^{11}$ esu. With the numbers given it is possible to solve (123) for $d_{36}$,

$$d_{36}\sim6\times10^{-10}\text{ esu}\quad\text{(KDP)}. \tag{124}$$

As a second example consider quartz, the crystal in which SHG was first discovered.[2] We suppose that the crystal is a $z$-cut plate (plane of plate normal to the optic axis); then the transverse component of polari-

[19] All the refractive index data for KDP quoted here was obtained by Giordmaine from extrapolations of published data in the visible region of the spectrum.

[20] J. A. Giordmaine (private communication).

zation is given by

$$P_t = d_{11}E^2 = d_{11}(8\pi/cn_\omega)S_L, \tag{125}$$

where $S_L$ is the laser intensity in the crystal, and $d_{11}$ is a second order polarization coefficient in the conventional notation. For the orientation assumed, the relevant optical constants[21] are those of the ordinary wave at the laser and second harmonic frequencies; thus $n'=1.541$ and $n=1.566$. Quartz is a good example of a nearly isotropic medium; in fact, the anisotropy is considerably less than the dispersion,

$$|n_\epsilon - n_\omega| \ll |n' - n_\omega|,$$

so that matching is impossible in this region of the spectrum ($\omega/2\pi c = 28{,}800$ cm$^{-1}$) for any orientation of the crystal. Therefore the medium may be considered isotropic for our present purposes, and $S$ is given by (109) with $P_t$ given by (125) and $\psi$ given by

$$\psi = (\omega/2c)(n'-n)l. \tag{126}$$

From (109), (125), and (126)

$$S = S_L^2(128\pi^3/(n'-n)^2cn)d_{11}^2 \sin^2\psi. \tag{127}$$

The periodic dependence of $S$ on $l$, predicted by (127), has been seen experimentally by Maker *et al.*,[5] who found that the maxima in $S$ are separated by $\Delta l = 14\ \mu$. This distance is to be identified with the coherence length for the wave $l_{\text{coh}}$; from (106), with $n_1 = n_2 = n$,

$$l_{\text{coh}} = 1.5\times10^{-4}\ \text{cm} = 15\mu, \quad \text{(quartz)} \tag{128}$$

in good agreement with the experimental value. With $\sin^2\psi = 1$ and $S_L \sim 1.5\times10^{11}$ esu it was observed[5] that one blue photon was produced for $5\times10^{12}$ red photons; thus $S/S_L = 4\times10^{-13}$. It is now possible to solve (127) for $d_{11}$,

$$d_{11} \sim 1\times10^{-10}\ \text{esu} \quad \text{(quartz)}. \tag{129}$$

We shall not enter into a discussion here of the values obtained for the second-order polarization coefficients, except to point out that they are several orders of magnitude smaller than predicted by Franken *et al.*[2] According to these authors $d$ should be of the order of an inverse atomic electric field, an example of which would be the energy gap of an insulator divided by the interatomic spacing; therefore one would expect $d \sim 10^{-8}$ esu. To explain the discrepancy between this estimate and the measured values (124) and (129), it would be necessary to consider the mechanism of the second-order polarization at optical frequencies. In this section we have analyzed two types of SHG experiments. In the case of KDP where the beam passed through the crystal in a matching direction, it was necessary to use (58) which takes into account the angular spread of the beam; this leads to (122) containing the function $F(z)$ defined in (60). In the case of quartz, on the other hand, it was unnecessary to consider the spread of the beam, since, in the experiment[5] referred to, interference effects are clearly resolved.

[21] *American Institute of Physics Handbook* (McGraw-Hill Book Company, Inc., New York, 1957), Vol. 6, p. 23.

## VIII. FOCUSED BEAMS

When the laser beam is focused within the crystal a very intense polarization $\mathbf{P}(\mathbf{r},t)$ is produced in the region of the focus. In this section we shall consider SHG by focused laser beams in an isotropic medium of refractive index $n''$ at the laser frequency and $n$ at the second harmonic frequency. If desired the main effects of a small anisotropy can be taken into account in the final formulas by allowing $n$ to be a function of the direction of the second harmonic radiation. The theory is beset by one major difficulty, which is to account in a reasonable way for the effects of partial coherence in the beam and aberrations in the lens. When both these effects are absent the focus is determined entirely by diffraction, and may be called an *ideal focus.*

The theory of the ideal focus has been given in great detail.[22] The nature of the focus is determined by the $f$/number of the system.

$$f = 1/\alpha, \tag{130}$$

where $\alpha$ is the full apex angle of the focal cone which is assumed small in the theory. If the focus is formed by a lens in air of focal length $l$, and if the beam diameter is $a$, the effective $f$/number is

$$f = n''l/a. \tag{131}$$

The intensity at the focus is

$$S_{\max} = \frac{\pi^2}{4}\frac{a^2}{\lambda''^2 f^2} S_L, \tag{132}$$

where $S_L$ is the laser intensity in the medium, and $\lambda''$ is the wavelength in the medium at the laser frequency. The region of high intensity is conveniently regarded as a cylinder of length $\xi''$ and diameter $\delta''$, where

$$\xi'' = 8\lambda'' f^2, \quad \delta'' = 2\lambda'' f. \tag{133}$$

Near the edges of this cylinder the intensity is quite small, and varies smoothly to $S_{\max}$ at the center. Since the SHG varies as the square of the fundamental intensity, the *effective radiating volume* is roughly a cylinder with dimensions half those of (133),

$$\xi = 8\lambda f^2, \quad \delta = 2\lambda f, \quad V = \tfrac{1}{4}\pi\xi\delta^2, \tag{134}$$

where $\lambda$ is the second harmonic wavelength $\lambda \approx \frac{1}{2}\lambda''$. Within this volume the focused beam is essentially a plane wave in the direction of the axis of the system and having a compressed wavelength $\lambda'$, where

$$\lambda' = \lambda''[1-(4f)^{-2}]. \tag{135}$$

The effective refractive index $n'$ describing the plane

[22] Reference 10, Sec. 8.8.

wave in the cylinder $\xi$, $\delta$ is, therefore,

$$n'=n''[1+(4f)^{-2}], \tag{136}$$

assuming $16f^2\gg 1$.

We shall assume that the conditions in an actual focus are somewhat as follows: The volume of the focus is larger than that given by (133) and the intensity is much less than $S_{\max}$. On the other hand the volume which can radiate coherently is still given by (134). We picture the actual focus as a bundle of cylinders (134) side by side radiating incoherently with each other. The total SHG will be proportional to the number of these cylinders and to the square of the volume of the cylinder.

As an example consider a quartz medium with $n''=1.57$, a laser beam of diameter $a=0.5$ cm., and a lens of focal length $l=1$ in. $=2.54$ cm. From (131) $f=8$, from (135) $\lambda'/\lambda''=0.999$, and from (134) $\xi=120\mu$, $\delta=5.6\mu$ ($1\mu=10^{-4}$ cm). It will be observed that $\xi$ is an order of magnitude larger than the coherence length in quartz[5] $l_{\text{coh}}\sim 14\mu$. We should therefore expect matching considerations to be important just as in the case of parallel beams. The wavelength compression effect (135) is very small in this example, but could be significant in a highly converging system $f\sim 1$. Since $\xi$ is so large, it is probably insensitive to small amounts of incoherence in the focus.

The volume (134) contains what we shall consider a plane polarization wave of the type (11) with constant amplitude inside the cylinder and zero amplitude outside. The wave vector is of the form (2) with $\mathbf{s}_k$ along the axis of the cylinder and $n'$ given by (136). The cylinder therefore acts as an antenna which radiates an electromagnetic field[23] at the second harmonic frequency,

$$\mathbf{E}(\mathbf{r})=\int_V [(\omega^2/c^2)\mathbf{P}(\mathbf{r}')\Phi(\mathbf{r}-\mathbf{r}') -n^{-2}(\nabla\cdot\mathbf{P}(\mathbf{r}'))\nabla\Phi(\mathbf{r}-\mathbf{r}')]d\mathbf{r}', \tag{137}$$

where $\nabla'$ operates on the coordinate $\mathbf{r}'$, and $\Phi$ is the Green's function

$$\Phi(\mathbf{r}-\mathbf{r}')=|\mathbf{r}-\mathbf{r}'|^{-1}e^{i(\omega/c)n'|\mathbf{r}-\mathbf{r}'|}. \tag{138}$$

The second term under the integral in (137) has the effect of subtracting off the longitudinal part of $\mathbf{P}(\mathbf{r})$. By means of an integration by parts, (137) can be written

$$\mathbf{E}(\mathbf{r})=(\omega^2/c^2)\int_V (I-\mathbf{ss})\cdot\mathbf{P}(\mathbf{r}')\Phi(\mathbf{r}-\mathbf{r}')d\mathbf{r}', \tag{139}$$

where $\mathbf{s}$ is a unit vector in the direction $\mathbf{r}-\mathbf{r}'$. The integration may readily be carried out for the radiation field at a great distance $r\gg r'$. The result is

$$r\mathbf{E}(\mathbf{r})=V(\omega^2/c^2)\mathbf{P}_t X(\xi,\theta)D(\delta,\theta)e^{i(\omega/c)nr}, \tag{140}$$

where $\mathbf{P}_t=(I-\mathbf{ss})\cdot\mathbf{P}$ is the transverse part of $\mathbf{P}$, $\theta$ is the angle between $\mathbf{r}$ and the axis, and

$$X(\xi,\theta)=\frac{\sin[((n'/n)-\cos\theta)\pi\xi/\lambda]}{[((n'/n)-\cos\theta)\pi\xi/\lambda]}, \tag{141}$$

$$D(\delta,\theta)=\frac{2J_1[(\pi\delta/\lambda)\sin\theta]}{[(\pi\delta/\lambda)\sin\theta]}. \tag{142}$$

In the limit of a small source $\xi\rightarrow 0$, $\delta\rightarrow 0$ we have $X=D=1$, and (140) gives the familiar formula[23] for the radiation field of an electric dipole $V\mathbf{P}$. The function $D(\delta,\theta)$ gives the diffraction pattern of the cylinder of diameter $\delta$. This function has a maximum value of unity when its argument $(\pi\delta/\lambda)\sin\theta=0$, and vanishes when the argument is 3.8. If we neglect the small amount of radiation outside the central peak, the full apex angle of the second harmonic radiation is essentially $\alpha$, the angle of the focused laser beam. In the function $X(\xi,\theta)$ we see a coherence behavior dependent on $n'$, $n$ very similar to that which occurs for parallel beams. In the spirit of (46) the coherence length may be defined as the smallest value of $\xi$ which causes $X(\xi,\theta)$ to vanish,

$$l_{\text{coh}}=\lambda/[(n'/n)-\cos\theta]. \tag{143}$$

For $\theta=0$, this agrees with (3) and (46), but the additional possibility now exists of choosing $\theta$ to satisfy the matching condition $\theta=\theta_M$, where

$$\cos\theta_M=n'/n, \tag{144}$$

providing $n'<n$. This will not lead to large SHG unless $2\theta_M$ lies within the apex angle $\alpha$ of the central peak of $D(\delta,\theta)$, the focal cone. Nevertheless, $D(\delta,\theta)$ does not vanish outside this cone, so that a weak but sharply defined ring may appear at angle $\theta_M$. If $\theta_M\ll 1$, (144) becomes

$$\theta_M=\{2[1-(n'/n)]\}^{1/2}. \tag{145}$$

Normal dispersion alone would be expected to give $\theta_M\sim 10°$, neglecting the compression effect (136) and assuming an isotropic medium. For angles $\theta$ close to $\theta_M$, (141) may be written

$$X(\xi,\theta)\approx\frac{\sin[(\theta-\theta_M)\xi\theta_M\pi/\lambda]}{[(\theta-\theta_M)\xi\theta_M\pi/\lambda]}. \tag{146}$$

The angular width of the matching ring is, therefore, (the width between zeros)

$$\Delta\theta=2\lambda/\xi\theta_M=1/4f^2\theta_M, \tag{147}$$

providing $\theta_M\neq 0$.

For quartz, and the other conditions of the example previously given in this section, (147) gives $\Delta\theta=0.04$ rad $\sim 2°$. At $\theta_M=0.18=10.5°$, for this example, $D(\delta,\theta_M)=0.054$, while at $\theta=0$, $X(\xi,0)=0.071$. There-

[23] Reference 14, Chap. VIII.

fore we should expect the spot in the forward direction due to $D$ and the ring of angle 10.5° due to $X$ to have roughly comparable intensity. Actually, (140) cannot be expected to give the radiation pattern in detail, since the assumption of a cylindrical focus containing a plane wave is only a convenient approximation.

No convenient formula for the total radiated power can be given which is valid for all cases. The usual case is that $\theta_M$ exists and lies well outside the focal cone. For this case, or for the case when $\theta_M$ does not exist, a simple approximate formula can be given for the total power in the focal cone:

$$\text{Power} \approx \frac{\pi^8 f^2}{n^3(n'-n)^2} \frac{c^3}{\omega^2} P_t^2, \tag{148}$$

which assumes an ideal focus. Since $P_t$ is proportional to $S_{\max}$, which according to (132) varies as $f^{-2}$, the power should vary as $f^{-2}$. It is difficult to estimate $d$ from focused beam experiments using this formula because of uncertainties about the nature of the actual focus. In any case (148) should be an upper bound on the power unless $\theta_M$ falls within the focal cone. For the matching case $n'=n$, $\theta_M=0$ the power is

$$\text{Power} \approx \frac{2\pi^4 f^2}{n^3} c \xi^2 P_t^2. \tag{149}$$

This result may be compared with (109) with $\psi=0$ for a parallel beam in an isotropic medium in the matching case. To obtain (109), with $\psi=0$, it is only necessary to divide (149) by the area $(\pi/4)\delta^2$ of the focal cylinder, and interpret $\xi$ as the path length $x$.

The formulas of this section remain valid for slightly anisotropic crystals, providing $n$ is considered a function of the direction of **r**. This takes the anisotropy into account as far as the radiation from the focal cylinder is concerned. There may, however, be an additional effect due to the distortion of the focus. We shall not attempt to treat the distortion problem, since it is only part of the much larger problem of characterizing an actual focus. Furthermore, there are two important cases in which the distortion would vanish: (a) the crystal is uniaxial with the axis along the beam axis, and (b) the crystal is uniaxial with the axis perpendicular to the beam axis and to the electric vector of the beam. In general, a focused beam experiment will produce in an anisotropic crystal both waves belonging to a direction **s**. In a uniaxial crystal both the ordinary and extraordinary waves will be produced. For certain orientations, however, crystal symmetry may require one or the other to vanish. For isotropic crystals (species $T$ or $T_d$) $P_t$ vanishes if the beam is directed along one of the crystal axes in the usual notation.[7] For uniaxial crystals in arrangement (a) defined above, $P_t$ vanishes except for types $C_3$, $C_{3v}$, $D_3$, and $D_{3h}$. In arrangement (b) only the ordinary wave is obtained from types $C_{3h}$, $D_3$, and $D_{3h}$, while only the extraordinary wave is obtained from types $C_4$, $S_4$, $C_{4v}$, $D_{2d}$, $C_6$, and $C_{6v}$. The output of $D_{2d}$, $D_4$, and $D_6$ vanishes when the beam is along any crystal axis. Thus it is a relatively simple matter to arrange an experiment which produces only one wave.

## SUMMARY

In this paper we have considered the electromagnetic field arising from a dielectric polarization of the kind that might be produced at the second harmonic frequency by a laser beam. The general formulation of Sec. II begins by postulating the existence of a plane polarization wave which acts as a source term in the inhomogeneous vector wave equation for the electric field. The plane wave solutions are determined by an extension of the methods familiar in the theory of crystal optics. A distinction is made between forced waves, which are particular solutions of the inhomogeneous equation, which vanish when the polarization vanishes, and free waves, which can propagate in the absence of sources. It is shown that the plane forced wave has a singularity in the matching case when the velocity of the polarization wave equals that of a free wave in the same direction. A separate treatment of the matching case is given which leads to a forced wave of linearly growing amplitude. Absorption and optical activity are neglected, but anisotropy is treated rigorously.

In Sec. III a forced wave is constructed suitable for the nearly matching case which remains finite and goes over into the correct solution in the matching limit. It is shown that the linearly growing wave in the matching case arises from interference between the plane forced wave and the matching free wave. In the nearly matching case the forced wave is periodic with a period which increases as the matching condition is approached. This behavior is described in terms of a coherence length defined as the distance between zeros of the amplitude of the forced wave. Then a beam is considered consisting of a narrow pencil of plane waves slightly differing in direction and centered on a matching direction. A general expression is given for the output intensity which is characterized in terms of two special cases, the thick crystal and the thin crystal. A coherence length for the beam is defined which depends on the beam spread. For crystals thinner than this beam coherence length the intensity varies as the square of the thickness, and for thicker crystals it varies linearly with the thickness.

In Sec. IV the complete solution is obtained for the electromagnetic field for the case of a semi-infinite medium having a single plane surface. This involves finding the appropriate combination of forced and free waves so as to satisfy boundary conditions on the surface. It is shown that the problem is analogous to the

Fresnel problem of finding the reflected and transmitted wave when a plane wave is incident on a plane surface. It is also shown that the forced wave constructed in Sec. III, from arguments based on the general properties that it should possess, can be derived directly from a detailed consideration of the surface. It is shown that the reflected wave in vacuum is insensitive to matching conditions in the crystal, which may provide a very good way of measuring the second-order polarization coefficients.

In Sec. V the effects of the finite size of the beam are discussed. The effect due to the anisotropy of the medium combined with the finite aperture is called the aperture effect. This effect can limit the effective coherence length in very thick crystals. Due to the small anisotropy of actual crystals the effect will not ordinarily be important. As an example, KDP is considered, and it is shown that the maximum coherence length is about 16 cm for a beam aperture of 0.5 cm.

In Sec. VI the general theory for anisotropic media is applied to the case of the nearly isotropic medium in such a way as to retain all the essential effects of the anisotropy and neglect all minor effects. It is pointed out that most anisotropic crystals may be considered nearly isotropic. The general theory of surface waves of Sec. IV is applied to an isotropic medium. It is shown that the familiar Fresnel formulas may be derived in this way.

In Sec. VII experimental results on KDP and quartz are analyzed according to the theory of Sec. VI. Values are obtained for the second-order polarization coefficients in these materials.

In Sec. VIII the theory is given for second harmonic generation by focused laser beams in an isotropic medium. The intense region of the focus is somewhat idealized as a cylinder containing a plane polarization wave. The cylinder then radiates like an antenna giving rise to the second harmonic output. The result contains two factors representing coherence effects, one of which is the diffraction pattern characteristic of the circular end of the cylinder, and the other depends upon the matching conditions very much like the forced wave for parallel beams. It is shown that the second harmonic radiation comes out in a cone of the same angle as the focussed laser beam. In addition, there may be a larger thin ring of radiation due to the matching condition. Expressions are given for the total radiated power in the main cone for the matching and nonmatching cases. The theory is based on an ideal focus determined entirely by diffraction. The effects of aberrations and partial coherence are described qualitatively but not treated in detail. The treatment can be extended to slightly anisotropic crystals if the dependence of refractive index on direction is taken into account.

## ACKNOWLEDGMENTS

The author wishes to acknowledge with appreciation the encouragement of W. S. Boyle which initiated this work, the many stimulating discussions with J. A. Giordmaine, and the help of D. E. McCumber on certain theoretical problems. The above mentioned were so kind as to read the manuscript and offer many profitable suggestions.