Theory of second-harmonic generation upon reflection of light from a medium with a center of inversion

V. M. Agranovich and S. A. Darmanyan
Institute of Spectroscopy, Academy of Sciences of the USSR

(Submitted 3 December 1981)

Pis'ma Zh. Eksp. Teor. Fiz. **35**, No. 2, 68–70 (20 January 1982)

New terms, in the same order in the field, are incorporated in the expression for the nonlinear polarization, which leads to second-harmonic generation, for inhomogeneous (bounded) media with a center of inversion. These new terms reconcile the theory with energy conservation. The new expression is used to find boundary conditions and the amplitudes of the second harmonic which arises upon the reflection of light from the boundary of an isotropic medium.

PACS numbers: 42.65.Cq, 42.20. – y

The recent expansion of research on surface physics has greatly intensified interest in second-harmonic generation upon reflection of light from the surfaces of media having a center of inversion.[1] In such media, the nonlinear part of the polarization, which causes the second-harmonic generation, can be described for an isotropic medium by

$$\mathbf{P}^{NL} = \frac{1}{4\pi} [\![\alpha\, \mathbf{E}\, \mathrm{div}\, \mathbf{E} + \beta (\mathbf{E}\vec{\nabla})\mathbf{E} + \gamma [\mathbf{E}\, \mathrm{rot}\, \mathbf{E}]]\!]. \tag{1}$$

Here, as in the theory of the gyrotropy of crystals, and more generally in theories for effects of spatial dispersion, the constitutive relation for the electric displacement contains not only the electric field but also its spatial derivatives. Although expression (1) has been used in many theoretical papers on second-harmonic generation (see Refs. 2-4 and the literature cited there), it is actually correct for only a homogeneous and infinite medium. If the medium is inhomogeneous or, in particular, bounded (in which case the coefficients α, β and γ depend on the coordinates), we must also take into account terms proportional to the gradients of the constitutive constants in the expression for the electric displacement which contains the spatial derivatives of the fields (this circumstance was emphasized in Ref. 5 in connection with gyrotropy). In this case the expression for the displacement $\mathbf{D}(\mathbf{r},t)$ can be written as follows (here and below, we are ignoring the spatial dispersion):

$$\mathbf{D} = \epsilon\, \mathbf{E} + \alpha\, \mathbf{E}\, \mathrm{div}\, \mathbf{E} + \beta (\mathbf{E}\vec{\nabla})\mathbf{E} + \gamma [\mathbf{E}\, \mathrm{rot}\, \mathbf{E}] + \mathbf{E}(\mathbf{E}\vec{\nabla}\rho) + \mathbf{E}^2 \vec{\nabla}\kappa, \tag{2}$$

where ρ and κ are new constitutive constants.

To determine the relationship between the constitutive constants in (2), we will use energy conservation, ignoring dissipative processes. In an unbounded homogeneous medium with the nonlinear polarization in (1), the Poynting equation

$$\frac{1}{4\pi}\left(\mathbf{E}\frac{\partial \mathbf{D}}{\partial t} + \mathbf{H}\frac{\partial \mathbf{H}}{\partial t}\right) = -\frac{c}{4\pi}\, \mathrm{div}\, [\mathbf{E}\,\mathbf{H}] \tag{3}$$

corresponds to a conservation law for the field energy if

$$a = -\beta = -\gamma. \tag{4}$$

If, on the other hand, the medium is inhomogeneous, then the substitution of (2) in (3) and the use of (4) lead to the equation

$$\frac{\partial W_0}{\partial t} + \operatorname{div} \mathbf{S}_0 + \frac{\partial E^2}{\partial t}\left(\mathbf{E}\vec{\nabla}\left(\kappa + \frac{\rho}{2} + \frac{a}{2}\right)\right) + E^2\left(\frac{\partial \mathbf{E}}{\partial t}\vec{\nabla}\rho\right) = 0, \tag{5}$$

where

$$W_0 = \frac{1}{8\pi}(E^2 + H^2 + 2aE^2 \operatorname{div}\mathbf{E}), \quad \mathbf{S}_0 = \frac{c}{4\pi}[\mathbf{E}\,\mathbf{H}] - \frac{aE}{8\pi}\frac{\partial E^2}{\partial t}.$$

Equation (5) has the form of an energy conservation law, $\partial W/\partial t + \operatorname{div}\mathbf{S} = 0$, only if $\rho = \alpha + 2\kappa$; in this case, we have $W = W_0 + W_1$ and $S = S_0$, where $W_1 = E^2(\mathbf{E}\vec{\nabla}\rho)/4\pi$. An interesting feature of the expression for W is the appearance of the surface energy density [at the sharp boundary at $z = 0$, the corresponding terms are $\sim\delta(z)$]. We might also note that if the terms proportional to $\nabla\rho$ and $\nabla\kappa$ in (2) are omitted, i.e., if we assume that Eq. (1) applies even to a bounded medium (as has been done elsewhere; see Ref. 2, for example), then we find from (5) that nonphysical energy sources $(1/2)(\partial E^2/\partial t)(\mathbf{E}\vec{\nabla}\alpha)$ appear at the surface.

The use of constitutive relation (2) instead of (1) significantly changes the boundary conditions for the fields at the doubled frequency. It can be shown that these conditions are

$$E_y(l) - E_y(0) = 0, \; E_x(l) - E_x(0) = \frac{2\omega i}{c}(\mu_1\epsilon_\infty^2 T_z^2 + \mu_2 T_t^2)\sin\theta,$$

$$\mathbf{H}_t(l) - \mathbf{H}_t(0) = \frac{2i\omega}{c}(a - 2\epsilon_\infty\mu_2)T_z[\mathbf{T}\,\mathbf{n}],$$

where

$$\mu_1 = \int_0^l \frac{dz}{\epsilon^3(z)}\frac{d}{dz}[a(z) + 3\kappa(z)] \quad \mu_2 = \int_0^l \frac{dz}{\epsilon(z)}\frac{d}{dz}\kappa(z)$$

is a characteristic of the transition layer, and n is a unit vector along the positive z direction. A nonlinear medium with a dielectric function $\epsilon(z)$ ($\epsilon_\infty \equiv \epsilon(\infty)$) occupies the region $z < l$, where l is the thickness of the transition layer. We also assume that light with a frequency θ is incident at an angle ω on the surface of the nonlinear medium; T is the amplitude of the transmitted wave, determined by the usual Fresnel formulas. It is easy to show that the resulting boundary conditions lead to the following expressions for the field at the doubled frequency in a linear medium (i.e., for $z > l$):

$$E_\perp^R(2\omega) = -\frac{2i\omega(a - 2\epsilon_\infty\mu_2)}{c(\cos\theta + \sqrt{\epsilon_\infty - \sin^2\theta})}T_y(\omega)T_z(\omega),$$

$$E_\parallel^R(2\omega) = \frac{2i\omega}{c(\sqrt{\epsilon_\infty}\cos\theta + \sqrt{\epsilon_\infty - \sin^2\theta})}\Big[\!\Big[\frac{a}{2}\sin\theta\, T^2(\omega) + \epsilon_\infty\sin\theta\,[\mu_1\epsilon_\infty^2 T_z^2(\omega) + \mu_2 T_t^2(\omega)]$$

$$+ (a - 2\epsilon_\infty\mu_2)T_x(\omega)T_z(\omega)\sqrt{\epsilon_\infty - \sin^2\theta}\Big]\!\Big]$$

where $E_\perp^R$ and $E_\parallel^R$ are the amplitudes of the waves polarized respectively perpendicular and parallel to the plane of incidence, $y=0$.

These results differ from the equations which have been used previously, and this difference should be kept in mind in interpreting experimental results.

1. C. K. Chen, A. R. B. de Castro, and Y. R. Shen, Phys. Rev. Lett. **46**, 145 (1981).
2. N. Bloembergen, R. K. Chang, S. S. Jha, and C. H. Lee, Phys. Rev. **174**, 813 (1968).
3. J. Rudnick and E. A. Stern, Phys. Rev. **B4**, 4274 (1971).
4. J. E. Sipe, V. C. Y. So, M. Fukui, and G. I. Stereman, Phys. Rev. **B 21**, 4389 (1980).
5. V. M. Agranovich and V. L. Ginzburg, Zh. Eksp. Teor. Fiz. **63**, 838 (1972) [Sov. Phys. JETP **36**, 44 (1973)]; V. M. Agranovich and V. I. Yudson, Opt. Commun. **9**, 58 (1973).

Translated by Dave Parsons
Edited by S. J. Amoretty