# Nonlinear polarization, light modulation and frequency conversion

Under “everyday” circumstances, the electric field of a light wave induces a dipole density called polarization that in a lossless and dispersionless medium is connected with the electric field of a light wave by the linear constitutive law given by (IV-15). For very strong field strengths this simply linear relationship between the applied field and the dipole moment induced by the field may not hold true and the term linear in E needs to be supplemented with an additional induced polarization vector $\mathrm{P}^{NL}$ that depends in a nonlinear manner upon the strength of the light field

$$\mathrm{P}(\mathrm{r},t) = \mathrm{P}^{L}(\mathrm{r},t) + \mathrm{P}^{NL}(\mathrm{r},t) \qquad \text{(IV-61)}$$

$$P_i(\mathrm{r},t) = P_i^{L}(\mathrm{r},t) + P_i^{NL}(\mathrm{r},t)$$

where the subscript *i* stands for one of the three Cartesian components of the field vector and $\mathrm{P}^{L}$ is the polarization vector increasing linearly with the electric field strength. In order to express the nonlinear polarization vector with the light electric field in a general manner, we assume that the electric field vector can be represented as a sum of a number of monochromatic waves

$$\mathrm{E}(\mathrm{r},t) = \sum_n \frac{1}{2}\mathrm{E}_n\, e^{-i\omega_n t} + c.c. = \sum_n \frac{1}{2}\mathrm{F}_n\, e^{i(\mathrm{k}_n\mathrm{r}-\omega_n t)} + c.c. \qquad \text{(IV-62)}$$

Where $\mathrm{E}_n$ is the complex amplitude and $\mathrm{F}_n$ is the complex envelope of the wave of frequency $\omega_n$ and wave vector $\mathrm{k}_n$. Both $\mathrm{E}_n$ and $\mathrm{F}_n$ are dependent on the spatial coordinate **r**, with $\mathrm{F}_n$ usually exhibiting only a small variation over a distance comparable to the wavelength (slowly-varying envelope approximation). For convenience, we introduce the notation

$$\frac{1}{2}\mathrm{E}_n = \mathrm{E}(\omega_n); \qquad \frac{1}{2}\mathrm{F}_n = \mathrm{F}(\omega_n) \qquad \text{(IV-63)}$$

to eliminate ½ and its powers in the subsequent expressions and extend the notation to negative frequency components

$$\mathrm{E}(-\omega_n) = \mathrm{E}^*(\omega_n); \qquad \mathrm{F}(-\omega_n) = \mathrm{F}^*(\omega_n) \qquad \text{(IV-64)}$$

which allow us to write (IV-62) in the more compact form

$$\mathrm{E}(\mathrm{r},t) = \sum_n \mathrm{E}(\omega_n) e^{-i\omega_n t} = \sum_n \mathrm{F}(\omega_n) e^{i(\mathrm{k}_n\mathrm{r}-\omega t)}, \qquad \text{(IV-65)}$$

$$E_i(\mathrm{r},t) = \sum_n E_i(\omega_n) e^{-i\omega_n t} = \sum_n F_i(\omega_n) e^{i(\mathrm{k}_n\mathrm{r}-\omega t)}.$$

where the summation is to be extended to all frequencies including the negative ones. This together with (IV-64) ensures that the sum results in a real quantity.
With this notation the nonlinear polarization can be expanded in a series of ascending powers of the applied field and the complex amplitude of its *i*-component be expressed with the applied fields in a similar manner as the linear component

$$P_i^L(\mathbf{r},t) = \varepsilon_0 \sum_{j,n} \chi_{ij}^{(1)}(\omega_n)\, E_j(\omega_n) e^{-i\omega_n t} \tag{IV-66a}$$

$$P_i^{NL}(\mathbf{r},t) = \varepsilon_0 \sum_{jk,mn} \chi_{ijk}^{(2)}(\omega_m,\omega_n)\, E_j(\omega_m) E_k(\omega_n) e^{-i(\omega_m+\omega_n)t} +$$

$$+\varepsilon_0 \sum_{jk\ell,mns} \chi_{ijk\ell}^{(3)}(\omega_m,\omega_n,\omega_s)\, E_j(\omega_m) E_k(\omega_n) E_\ell(\omega_s) e^{-i(\omega_m+\omega_n+\omega_s)t} + \ldots\ldots \tag{IV-66b}$$

where the coefficient $\chi^{(n)}$ is the $n^{th}$-order nonlinear susceptibility, an $(n+1)^{th}$-rank tensor, and $\chi^{(1)}$ is the second-rank linear susceptibility tensor, which was introduced in its simplest, frequency-independent form by Eq. (IV-15). The linear susceptibility tensor has – in the most general case – three independent components: in what is called the principal coordinate system of the propagation material only the three diagonal elements of $\chi^{(1)}$ are different from zero. The MKSA unit of the $n^{th}$-order nonlinear susceptibility is $(m/V)^{n-1}$. An obvious consequence of the nonlinear response of matter to strong light fields is the emergence of induced polarization components at new frequencies

$$\mathbf{P}(\mathbf{r},t) = \mathbf{P}^L + \mathbf{P}^{NL} = \mathbf{P}^{(1)} + \mathbf{P}^{(2)} + \mathbf{P}^{(3)} + \ldots\ldots = \tag{IV-67}$$

$$= \sum_p \mathbf{P}^{(1)}(\omega_p) e^{-i\omega_p t} + \sum_q \mathbf{P}^{(2)}(\omega_q) e^{-i\omega_q t} + \sum_r \mathbf{P}^{(3)}(\omega_r) e^{-i\omega_r t} + \ldots$$

where the new frequencies emerge at all possible sums of those of the applied fields (including the negative ones!)

$$\omega_p = \omega_n; \quad \omega_q = \omega_m + \omega_n; \quad \omega_r = \omega_m + \omega_n + \omega_s \tag{IV-68}$$

This Taylor expansion of the nonlinear constitutive law converges only if terms of increasing order decrease in magnitude. We shall now estimate under which conditions this is the case. One might expect that the magnitude of the lowest-order correction term $\mathbf{P}^{(2)}$ becomes comparable to the linear response $\mathbf{P}^L$ when the amplitude of the applied field E is of the order of the characteristic atomic electric field strength

$$E_{at} = \frac{1}{4\pi\varepsilon_0} \frac{e}{a_0^2} \tag{IV-69}$$

where $e$ is the electron charge and $a_0$ is the Bohr radius, which is the radius of the most-strongly-bound orbit of the electron in Bohr's model of the hydrogen atom. It can be expressed with Planck's constant
and the electron's mass and charge as

$$a_0 = \frac{\hbar^2}{me^2} \approx 0.5\text{Å} = 5\times 10^{-11}\text{m} \tag{IV-70}$$

yielding for the characteristic atomic electric field strength

$$E_{at} = \frac{m^2 e^5}{4\pi\varepsilon_0 \hbar^4} \approx 6\times 10^9 \frac{\text{V}}{\text{cm}} \tag{IV-71}$$

The ratio of the lowest-order nonlinear to the linear response, or that of two subsequent contributions to the nonlinear response can now be expressed as

$$\left|\frac{\mathrm{P}^{(2)}}{\mathrm{P}^{(1)}}\right| \approx \left|\frac{\mathrm{P}^{(n+1)}}{\mathrm{P}^{(n)}}\right| \approx \left|\frac{E}{E_{at}}\right| \tag{IV-72}$$

The power series in (IV-66b) can be approximated by the first few lowest order terms when the expansion parameter $\kappa$ fulfils the condition

$$\kappa = \left|\frac{E}{E_{at}}\right| \ll 1 \tag{IV-73}$$

Eq. (IV-73) establishes the range of validity of (IV-66b) and justifies the *perturbative* description of the nonlinear response of matter to an external electric field, hence defining the range of *perturbative nonlinear optics*.

The intensity of a light wave corresponding to the electric field strength can be obtained by using (IV-34) here for free space ($n$ = 1)

$$I_{at} = \frac{1}{2}\frac{\left|E_{at}\right|^2}{377\Omega} \approx 5\times 10^{16} \frac{\text{W}}{\text{cm}^2} \tag{IV-74}$$

For intensities substantially lower than $10^{15}$ W/cm$^2$, the above perturbative approach generally gives an accurate account of nonlinear optical phenomena in atoms with a binding energy comparable to or larger than that of the electron in the ground-state of hydrogen, approximately 15 eV, (which served as a basis for deriving the above criterion). In condensed matter, the binding energy of valence electrons is typically much lower, resulting in ionization and subsequent breakdown at intensity levels typically a couple of orders of magnitude lower. Hence the optical breakdown sets a practical limit to reversible nonlinear optics in solid materials at intensity levels of the order of $10^{15}$ W/cm$^2$.

Note that a 10-Watt laser beam delivered by a powerful continuous-wave laser gives rise to an intensity of $10^9$ W/cm$^2$ when focused down to a spot diameter of 1 micrometer. Hence the investigation and exploitation of nonlinear optical phenomena usually require not only spatial but also *temporal confinement of light energy*, i.e. short-pulsed radiation.

The implications of the induced nonlinear polarization become apparent by substituting the generalized constitutive law given by (IV-61) and (IV-66) into the wave equation. Here we shall use the wave equation derived for the electric field (IV-19) rather than those obtained for the potentials (IV-27) simply because the constitutive law relates the polarization vector directly to the electric field rather than to the latter auxiliary quantities. Decomposing the induced polarization P into a component scaling linearly and nonlinearly with the strength of the applied fields, $P^L$ and $P^{NL}$, respectively, and assuming – for the sake of simplicity – a lossless, dispersionless and isotropic propagation medium, so that the constitutive law for the linear component takes the form $P^L = \varepsilon_0\chi E$, substitution of (IV-61) into (IV-19) yields in the absence of dc current ($J_0 = 0$)

$$\nabla \times (\nabla \times \mathrm{E}) + \frac{1}{c^2}\frac{\partial^2 \mathrm{E}}{\partial t^2} = -\frac{\chi}{c^2}\frac{\partial^2 \mathrm{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \mathrm{P}^{NL}}{\partial t^2} \qquad \text{(IV-75)}$$

which, by making use of $\varepsilon_r = 1 + \chi$ can be simplified to

$$\nabla \times (\nabla \times \mathrm{E}) + \frac{\varepsilon_r}{c^2}\frac{\partial^2 \mathrm{E}}{\partial t^2} = -\mu_0 \frac{\partial^2 \mathrm{P}^{NL}}{\partial t^2} \qquad \text{(IV-76)}$$

When the relative permittivity $\varepsilon_r$ does not vary significantly over length scales comparable to the wavelength, the last term in (IV-21) can be neglected and we arrive at a more common form of the wave equation for the electric field in the presence of an induced nonlinear polarization

$$\nabla^2 \mathrm{E}\,(\mathrm{r},\mathrm{t}) - \frac{n^2}{c^2}\frac{\partial^2 \mathrm{E}(\mathrm{r},t)}{\partial t^2} = \mu_0 \frac{\partial^2 \mathrm{P}^{NL}}{\partial t^2} \qquad \text{(IV-77)}$$

This equation has the form of a driven (inhomogeneous) wave equation; the nonlinear polarization acts as a source term appears on the right-hand side of this equation. In the absence of this source term, the equation admits solutions of the form of free waves propagating with a phase velocity $c/n$ where $n = \varepsilon_r^{1/2}$ is the refractive index of the medium. The nonlinear wave equations (IV-76) or (IV-77) together with the constitutive law in the perturbative limit of nonlinear optics provide a powerful mathematical framework for the description of a wide range of nonlinear optical phenomena.

## Second-order nonlinear susceptibility $\chi^{(2)}$: second harmonic generation, sum- and difference-frequency mixing, optical parametric amplification

The second-order nonlinear susceptibility $\chi^{(2)}$ is nonvanishing if and only if the medium is noncentrosymmetric and gives rise to a number of important nonlinear optical effects.

### *Second-harmonic generation*

Consider an intense light wave of frequency $\omega_1$ propagating in the *z* direction and with its electric field polarized along the *x* axis

$$E_1(\mathbf{r},t) = -\frac{1}{2} F_1 \, e^{i(k_1 z - \omega_1 t)} + c.c. \tag{IV-78}$$

where the wave vector has the magnitude $k_1 = \omega_1 n(\omega_1)/c$. The second-order nonlinear polarization induced along the $x$ axis at a frequency $\omega_2 = 2\omega_1$, according to (IV-66b), given by

$$P_2(\mathbf{r},t) = \frac{1}{4} \varepsilon_0 \, \chi^{(2)}_{xxx}(\omega_1,\omega_1) \, F_1^2 \, e^{i(2k_1 z - 2\omega_1 t)} + c.c. \tag{IV-79}$$

Substituting (IV-79) into the wave equation introduces a polarization wave carried at a frequency $\omega_2$ and with a wave vector of magnitude $2k_1$. This polarization wave will generate an electric field wave

$$E_2(\mathbf{r},t) = \frac{1}{2} F_2 \, e^{i(k_2 z - \omega_2 t)} + c.c. \tag{IV-80}$$

with a wave vector of magnitude $k_2 = \omega_2 n(\omega_2)/c$ and at a frequency $\omega_2 = 2\omega_1$, i.e. at the *second harmonic* of the input light wave $E_1$. Because $\kappa << 1$ below the ionization (optical breakdown) threshold, the amplitude of the induced polarization wave is typically very small. Consequently, the electric field wave driven by this polarization wave can build up and attain a sizeable amplitude and intensity only over an extended propagation length $L >> \lambda$. The coherent buildup of the output wave is generally severely compromised by a phase mismatch between the harmonic wave source wave and the harmonic wave $\Delta k = k_2 - 2k_1 = \omega_2 [n(\omega_2) - n(\omega_1)]/c$, which is a direct consequence of the frequency dependence of the refractive index. For the growth of the output wave with the propagation distance $L$ the wave equation (IV-77) yields (Exercise)

$$F_2 \propto \chi^{(2)}_{xxx}(\omega_1,\omega_1) \, F_1^2 \, \frac{e^{i\Delta k L} - 1}{i\Delta k} \tag{IV-81}$$

with the output intensity scaling as

$$I(2\omega_1) \propto \left| \chi^{(2)}_{xxx}(\omega_1,\omega_1) \right|^2 I^2(\omega_1) \, L^2 \, \frac{\sin^2(\Delta k L/2)}{(\Delta k L/2)^2} \tag{IV-82}$$

The intensity of the harmonic wave stops growing at $L_c = \pi/\Delta k$, also referred to as the *coherence length* of the harmonic generation process. Beyond this propagation length the harmonic emission produced by the induced dipoles adds destructively to the accumulated harmonic wave and decreases its amplitude and intensity.

Due to the frequency-dependence of the refractive index – briefly, dispersion – $\Delta k$ is always nonzero if the electric wave driving the polarization wave and the new wave are polarized along the same direction. However, the anisotropy of crystals can be utilized to circumvent this problem. By inducing the polarization wave orthogonally to the electric field vector of the driving wave upon exploiting a component of the nonlinear susceptibility tensor $\chi_{jjj}^{(2)}$ so that $i \neq j$ the polarization wave and the harmonic wave will propagate with a wavevector $2k_1 = \omega_2 n_j(\omega_1)/c$ and $k_2 = \omega_2 n_i(\omega_2)/c$, respectively. Because the refractive indices $n_j(\omega_1)$ and $n_i(\omega_2)$ are to be taken for different polarization directions in an anisotropic crystal, they can be equal in spite of dispersion, resulting in $\Delta k = 0$, called *phase matching*.

Phase-matched second harmonic generation is an important technique for converting the coherent radiation of lasers to shorter wavelength radiation. With intense ultrashort (nano- to picosecond) light pulses conversion efficiencies up to 80% can be achieved.

*Sum-frequency generation*

For two different waves of frequencies $\omega_1$ and $\omega_2$ the second-order polarization response may also result in efficient generation of radiation at $\omega_3 = \omega_1 + \omega_2$ , if the wave vector $k_1 + k_2$ of the polarization wave

$$P_3(\mathbf{r},t) = \frac{1}{4}\,\varepsilon_0\,\chi_{eff}^{(2)}(\omega_1,\omega_2)F_1F_2\,e^{i(k_1+k_2)z-i(\omega_1+\omega_2)t} + c.c. \qquad \text{(IV-83)}$$

is matched to that of the output wave of $\omega_3$ by exploiting crystal anisotropy. Here $\chi_{eff}^{(2)}$ stands for the effective susceptibility for the selected input-wave and output-wave polarizations including also summations over permutations of the input fields according to (IV-66b).

*Difference-frequency generation and parametric amplification*

Consider the case of two input waves at frequencies $\omega_3$ and $\omega_1$ such that $\omega_3 > \omega_1$ and let us induce the polarization wave

$$P_2(\mathbf{r},t) = \frac{1}{4}\,\varepsilon_0\,\chi_{eff}^{(2)}(\omega_3,-\omega_1)F_3F_1^*\,e^{i(k_3-k_1)z-i(\omega_3-\omega_1)t} + c.c. \qquad \text{(IV-84)}$$

which drives a third wave at the difference frequency $\omega_2 = \omega_3 - \omega_1$

$$E_2(\mathbf{r},t) = \frac{1}{2}\,F_2\,e^{i(k_2z-\omega_2t)} + c.c. \qquad \text{(IV-85)}$$

We can now distinguish two distinctly different operation regimes.

$|F_1| \approx |F_3| \rightarrow$ *difference-frequency generation*
When the two input waves are comparable in strength, phase-matched difference-frequency mixing results in the efficient build-up of the wave at $\omega_2$ with energy being coupled from both input waves into the new driven wave in a balanced manner, similarly to the power coupling in sum-frequency generation.

$|F_1| << |F_3| \rightarrow$ *optical parametric amplification*
When the wave of $\omega_3$ is much more intense than that of $\omega_1$, after an initial buildup of the wave of $\omega_2$ this becomes comparable in strength to that of frequency $\omega_1$. As a consequence, together with the strong (pump) wave at $\omega_3$ it creates a polarization wave

$$P_1(\mathbf{r},t) = \frac{1}{4}\,\varepsilon_0\,\chi_{eff}^{(2)}(\omega_3,-\omega_2)F_3F_2^*\,e^{i(k_3-k_2)z-i(\omega_3-\omega_2)t} + c.c. \qquad \text{(IV-86)}$$

at $\omega_1 = \omega_3 - \omega_2$. If the wave mixing process is phase-matched this polarization wave drives and boosts the power of the $\omega_1$ wave in the same way as the polarization wave $P_2$ drives and builds up the $\omega_2$ wave. Consequently, the strong $\omega_3$ input wave serves as a pump for building up the $\omega_2$ wave and amplifying the weak input wave at $\omega_1$ simultaneously. The latter process is referred to as optical parametric amplification (often abbreviated as OPA) and is a powerful technique for the efficient amplification of either broadband (i.e. ultrashort-pulsed) or tunable laser radiation ($\omega_1$) by means of an energetic laser pulse ($\omega_3$).

### Second-order nonlinear susceptibility $\chi^{(2)}$: phase and amplitude modulation and switching of light by means of the electro-optic effect

A second-order nonlinear polarization induced by an optical (high-frequency, $\omega_1$) and a static (dc) electric field

$$P_1^{(2)}(\mathbf{r},t) = \frac{1}{4}\,\varepsilon_0\,\chi_{eff}^{(2)}(\omega_1,0)\,E_1\,E_{dc}\,e^{-i\omega_1 t} + c.c. \qquad \text{(IV-87)}$$

oscillates at the same frequency $\omega_1$ as the linear polarization, hence the two components can be combined to a single polarization wave

$$P_1 = P_1^{(1)} + P_1^{(2)} = \frac{1}{2}\,\varepsilon_0\left(\chi_{eff}^{(1)} + \frac{1}{2}\,\chi_{eff}^{(2)}\,E_{dc}\right)E_1\,e^{-i\omega_1 t} + c.c. \qquad \text{(IV-88)}$$

of frequency $\omega_1$ and the second-order contribution makes a small contribution to the refractive index

$$n_{i'} = \sqrt{1 + \chi_{eff}^{(1)} + \frac{1}{2}\,\chi_{eff}^{(2)}\,E_{dc}} \approx n_i + \frac{1}{2}\,n_i^3 r_{eff}\,E_{dc} \qquad \text{(IV-89)}$$

that is dependent on the static field $E_{dc}$. Here the coefficient $r_{eff}$ is the respective linear electro-optic coefficient (depending on the direction of $E_{dc}$), the defining equation of which is $\Delta\eta_{ij} = r_{ijk}\,E_k$, where the tensor $\eta_{ij} = (1 + \chi^{(1)})^{-1}{}_{ij}$ is called the impermeability tensor[1]. In the coordinate system, where the impermeability tensor is diagonal, its principal values are $\eta_i = 1/n_i^2 = 1/(1 + \chi^{(1)}{}_{ii})$, where $i = x$, $y$, or $z$. The impermeability tensor is very useful in describing light wave propagation through anisotropic media. The prime in (IV-89) indicates that the linear electro-optic (or Pockels) effect may – depending on the crystal symmetry – have rotated the principal axes and $i'$ denotes the new principal axis.

The electro-optic effect provides the most efficient and most widely-used means of modulating the phase and the amplitude of light. Let's consider the specific case of a KDP crystal with $n_x = n_y = n_o$ (called the *ordinary index*) and $n_z = n_e$ (called the *extraordinary index*). By applying a dc field $E_z$ along the $z$ axis the principal coordinate system of the crystal is rotated about this axis by 45 degree, yielding the refractive indices for light polarized along the new $x'$ and $y'$ axes

$$n_{x'} = n_\sigma - \frac{1}{2}\,n_\sigma^3\,r_{xyz}\,E_z$$

[1] For more details, see e.g. A. Yariv, P. Yeh, Optical waves in crystals, John Wiley & Sons, 1984.

$$n_{y'} = n_\sigma + \frac{1}{2} n_\sigma r_{xyz} E_z \tag{IV-90}$$

$$n_z = n_e$$

where for $r_{xyz} = r_{yxz}$ often the contracted notation $r_{63} = r_{xyz} = r_{yxz}$ is used. With this contracted notation, the phase shift induced by the applied voltage to a light wave propagating along the $z$ axis with its electric field vector polarized along the induced birefringence axes $x'$ and $y'$ can be written as

$$\Delta\varphi_{x'} = \Delta k_{x'} L = k_0 L \Delta n_{x'} = -\frac{k_0 n_o^3 r_{63}}{2} E_z L \tag{IV-91}$$

$$\Delta\varphi_{y'} = \Delta k_{y'} L = k_0 L \Delta n_{y'} = +\frac{k_0 n_o^3 r_{63}}{2} E_z L$$

*Phase modulation of light*

We can now exploit this electro-optic effect for modulating the phase of a light beam $E_{in} = F\cos\omega t$ polarized e.g. along the $x'$ axis as shown in Fig. IV-5 at the output as

$$E_{\text{out}} = F \cos\left[\omega t - k_0 n_o L + \frac{k_0 n_o^3 r_{63} L}{2} E_z(t)\right] \tag{IV-92}$$

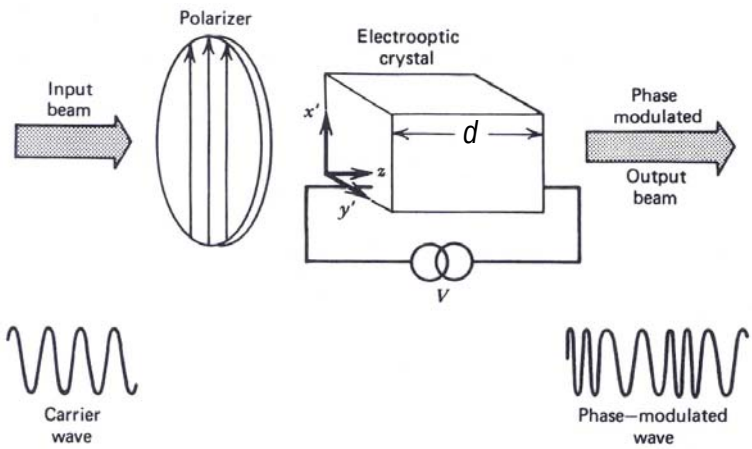


Fig. IV-5

*Amplitude modulation and switching of light*

If the electro-optic crystal is sandwiched between crossed polarizers parallel to the *x* and *y* axes, respectively (i.e. rotated by an angle of 45 degree with respect to the induced birefringence axes *x*' and *y*') the electro-optic effect can be used for amplitude modulation and of light transmitted through this apparatus (Fig. IV-6). In this case the input wave $E_{in} = F\cos\omega t$ is decomposed into two equally strong components polarized linearly along the *x*' and *y*' axes, the latter of which suffers a phase retardation with respect to the former:

$$\Gamma = \Delta\varphi_{y'} - \Delta\varphi_{x'} = k_0 n_o^3 r_{63} E_z L = \frac{2\pi}{\lambda_0} n_o^3 r_{63} V = \pi \frac{V}{V_\pi} \qquad \text{(IV-93)}$$

where $V = E_z L$ is the applied voltage and

$$V_\pi = \frac{\lambda_0}{2 n_\sigma^3 r_{63}} \qquad \text{(IV-94)}$$

is the voltage causing a phase retardation of π between the two waves in the crystal and is referred to as the *half-wave voltage*. The half-wave voltage for a *z*-cut KDP crystal at $\lambda_0$ = 800 nm is about 11.7 kV.
For Γ = π the input polarization is rotated by 90 degree and the wave is completely transmitted through the second polarizer. Switching on and off the half-wave voltage changes the transmittivity of the system between 100% and zero, resulting in *switching of a light wave*. For modulating the amplitude of the transmitted wave (instead of switching it on and off), we can apply a bias voltage $V_{bias} = V_\pi/2$ together with the modulation voltage $V(t)$. This results in modulation of the transmittivity and hence of the amplitude of the transmitted wave as shown in Fig. IV-7.

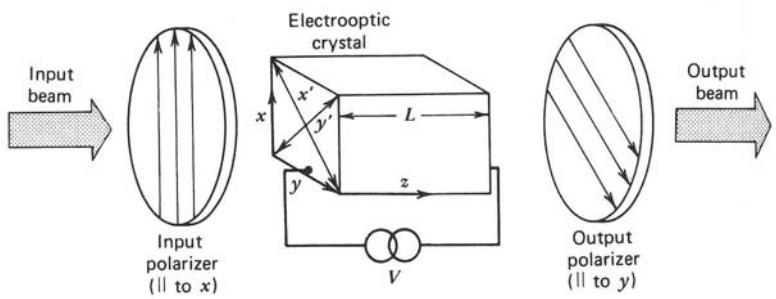


Fig. IV-6

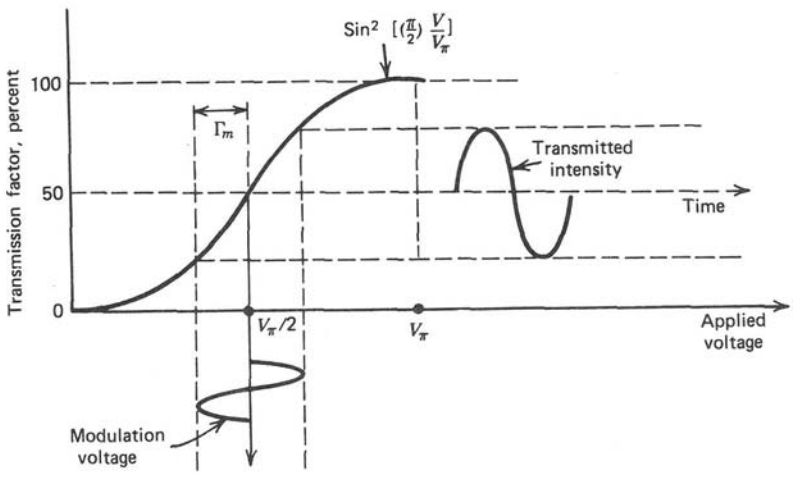


Fig. IV-7

Fig. IV-8 shows a completely different implementation of amplitude modulation and switching based on the electro-optic effect in an integrated-optic Mach-Zehnder interferometer.

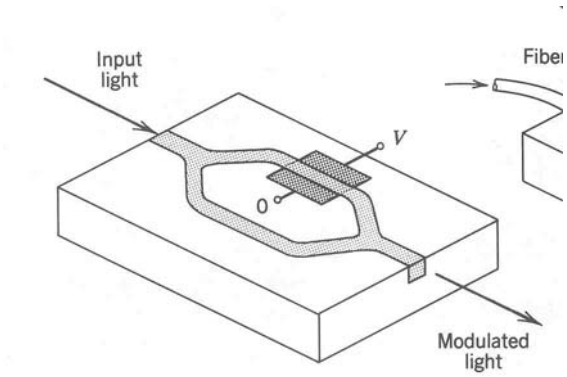


Fig. IV-8

Third-order susceptibility $\chi^{(3)}$: nonlinear index of refraction, self-phase modulation, self focusing of laser light

One of the most obvious implications of the third-order term in the Taylor expansion of the nonlinear polarization is the emergence of the third harmonic $\omega_3 = 3\omega_1$ of the input wave ($\omega_1$) driving the polarization wave. However, the practical importance of this effect is rather limited, because the same radiation can be much more efficiently produced by cascading two phase-matched $\chi^{(2)}$ processes: second harmonic generation to yield $\omega_2 = 2\omega_1$ and sum-frequency mixing of the $\omega_2$ and the remaining part of the fundamental at $\omega_1$ to result in $\omega_3 = \omega_2 + \omega_1 = 3\omega_1$. All the more important is another $\chi^{(3)}$-based nonlinear effect which gives rise to an intensity-induced change of the refractive index by mixing three waves of frequencies $\omega$, $\omega$, and $-\omega$.

Consider a light wave propagating along the *z* axis with its electric field vector polarized along the *x* axis as given by

$$\mathrm{E}(\mathrm{r},t) = \frac{1}{2}\hat{\mathrm{x}}\, E_x e^{i(kz-\omega t)} + c.c. = \hat{\mathrm{x}} E_x(\omega) e^{-i\omega t} + c.c. \qquad \text{(IV-95)}$$

Where, for the sake of simplicity, the coordinate system is aligned to be equivalent to the principal axes of the (possibly anisotropic) propagation medium. The constitutive law (IV-66b) along with the intrinsic permutation symmetry $\chi^{(3)}_{xxxx}(\omega,\omega,-\omega) = \chi^{(3)}_{xxxx}(\omega,-\omega,\omega) = \chi^{(3)}_{xxxx}(-\omega,\omega,\omega)$ yields for the polarization wave at the same frequency and polarized along the same (*x*) direction

$$P_x^{(3)}(\mathrm{r},t) = 3\varepsilon_0\, \chi^{(3)}_{xxxx}(\omega,\omega,-\omega)\, \left|E_x(\omega)\right|^2 E_x(\omega) \qquad \text{(IV-96)}$$

The total (linear + nonlinear) polarization density induced along the *x* axis is then given by

$$P_x(\mathrm{r},t) = P_x^{(1)}(\mathrm{r},t) + P_x^{(3)}(\mathrm{r},t) = \varepsilon_0 \left[\, \chi^{(1)}_{xx}(\omega) + 3\chi^{(3)}_{xxxx} \left|E_x(\omega)\right|^2 \right] E_x(\omega) e^{-i\omega t} + c.c. \qquad \text{(IV-97)}$$

The expression within the brackets may be regarded as a generalized (complex) susceptibility with the field-induced contribution giving rise to a small and generally complex correction of the refractive index, the real ($\Delta n$) and imaginary ($\Delta\alpha/2k_0$) part of which can be derived from

$$(n+\Delta n+i\Delta\alpha/2k_0)^2 \approx n^2+2n\Delta n+in\Delta\alpha/k_0 = 1+\chi_{xx}^{(1)}+\frac{3}{4}\chi_{xxxx}^{(3)}\left|E_x\right|^2 \qquad \text{(IV-98)}$$

as

$$\Delta n = \frac{3}{8n}\,\mathrm{Re}\left(\chi_{xxxx}^{(3)}\right)\left|E_x\right|^2 = \frac{n_2 n}{2Z_0}\left|E_x\right|^2 = n_2 I \qquad \text{(IV-99a)}$$

$$\Delta\alpha = \frac{3\omega}{4nc}\,Im\left(\chi_{xxxx}^{(3)}\right)\left|E_x\right|^2 = \frac{\alpha_2 n}{2Z_0}\left|E_x\right|^2 = \alpha_2 I \qquad \text{(IV-99b)}$$

and

$$n_2 = \frac{3Z_0}{4n^2}\,\mathrm{Re}\left(\chi_{xxxx}^{(3)}\right) = \frac{3}{4\varepsilon_0 n^2 c}\,\mathrm{Re}\left(\chi_{xxxx}^{(3)}\right) \qquad \text{(IV-100a)}$$

$$\alpha_2 = \frac{3\omega Z_0}{2n^2 c}\,Im\left(\chi_{xxxx}^{(3)}\right) = \frac{3\omega}{2\varepsilon_0 n^2 c^2}\,Im\left(\chi_{xxxx}^{(3)}\right) \qquad \text{(IV-100b)}$$

are referred to as the *nonlinear-index coefficient* (often also: *nonlinear refractive index*) and *two-photon absorption coefficient* of the propagation medium. Eqs. (IV-99) reveal that the real part of $\chi^{(3)}_{xxxx}(\omega,\omega,-\omega)$ introduces a small change of the refractive index that scales with the intensity of the light wave. This phenomenon is known as the *optical Kerr effect*. The imaginary part of $\chi^{(3)}_{xxxx}(\omega,\omega,-\omega)$ on the other hand gives rise to a small absorption coefficient that is proportional to the light intensity, consequently the amount of absorbed light in unit propagation length scales with the square of the light intensity, which is indicative of two photons inducing the respective atomic transitions. This latter process is likely only if the frequency of the incident radiation is close enough to bandgap of the transparent propagation medium so that this bandgap can be overcome by two photons. For large-gap materials such as quartz or fused silica (bandgap ~ 10 eV) two-photon absorption requires ultraviolet photon energies and hence the imaginary part of $\chi^{(3)}_{xxxx}(\omega,\omega,-\omega)$ vanishes in the visible spectral range.

For large-gap materials at intensity levels safely below the damage threshold

$$n_2 \approx 10^{-16}\,\frac{cm^2}{W},\; I = 10^{12}\,\frac{W}{cm^2} \;\Rightarrow\; \Delta n = n_2 I \approx 10^{-4} \qquad \text{(IV-101)}$$

The refractive index can only be changed in the fourth digit behind comma even with intense laser light.

Nevertheless even this small change results in the accumulation of a substantial phase shift over a relatively short propagation

$$\Delta\varphi_{\text{nonlinear}} = k_0 n_2 I L = 2\pi n_2 I \frac{L}{\lambda} \tag{IV-102}$$

e.g. $\Delta n = 10^{-4}$ implies a nonlinear phase shift $\Delta\varphi_{\text{nonlinear}} = 2\pi$ for a propagation distance as short as $L$ = 8 mm for $\lambda$ = 800 nm. In a pulsed laser beam, the intensity is dependent on both time and the radial coordinate, resulting in a temporally and radially dependent nonlinear phase shift. Former leads to *self-phase modulation* of the propagating laser pulse whereas the latter implies *self focusing*. These effects are often undesirable and – if uncontrolled – may lead to severe problems such as damage to optical components in high-power laser systems.

This nonlinearity is of electronic origin, occurs as the result of the nonlinear response of bound electrons to an applied optical field. Because the light frequency in a transparent wide-gap material is far from resonances, the response is small but extraordinarily fast. The characteristic response time of this process is the time it takes for the electron cloud around the nucleus to become slightly distorted (giving rise to a small dipole moment induced by the field) in response to an applied optical field. This response time can be estimated as the oscillation period of the electron density in an excited atom. This can be assessed by noting that the modulus square of the wavefunction of the electron occupying its ground state (of energy $E_0$) and with some finite probability its first excited state (of energy $E_1$) takes the form

$$\left|\psi^2\right| = \left| c_0\psi_0 + c_1\psi_1 e^{-\frac{i}{\hbar}(E_1 - E_0)t} \right|^2 \tag{IV-103}$$

where $\psi_0$ and $\psi_1$ are the wavefunctions of the electron in the ground and the first excited state. The oscillation period is determined by the energy gap $\Delta E = E_1 - E_0$ between the ground state and the first excited state as

$$\tau_{\text{osc}} = \frac{2\pi\hbar}{\Delta E} \approx \frac{4fs}{(\Delta E / 1\text{eV})} \tag{IV-104}$$

Where 1 femtosecond = 1 fs = $10^{-15}$ s. For quartz, with a bandgap of $\Delta E \approx 10$ eV we obtain a response time of 400 attoseconds (1 as = $10^{-18}$ s) for the nonlinear index of refraction (as well as for other nonlinearities of electronic origin). This response time amounts to a tiny fraction of the field oscillation period of visible light, hence much faster than the fastest changes in the cycle-averaged intensity of optical radiation. As a consequence, the nonlinear index of refraction of electronic origin has a virtually instantaneous response to even the fastest changes of optical radiation. This fact provides the basis for the operation of state-of-the-art ultrashort-pulse (femtosecond) laser technology (see Chapter VIII: Ultrafast optics, Photonics II).