Upper bound imposed on responsivity of optical modulators

Eyal Buks

Department of Electrical Engineering, Technion, Haifa 32000, Israel

Received October 14, 2005; revised January 20, 2006; accepted January 23, 2006; posted January 23, 2006 (Doc. ID 65385)

I study theoretically the responsivity of optical modulators. For the case of a linear response, by using perturbation theory I find an upper bound imposed on the responsivity. For the case of a two-mode modulator I find a lower bound imposed on the optical path required for achieving full modulation when the maximum birefringence strength is given. © 2006 Optical Society of America

OCIS codes: 060.4080, 230.4110, 350.1370.

Optical modulators are devices of great importance for telecommunications and other fields. These devices allow the transmission T $(0 \leqslant T \leqslant 1)$ to be controlled between input and output ports by applying some external perturbation. One of the key characterizations of optical modulators is their responsivity, namely, the dependence of T on the applied external perturbation. Enhancing the responsivity is highly desirable for many applications. This raises the question, what is the largest possible responsivity that can be achieved for a given perturbation mechanism? Here I consider this question for the case of linear modulators. I show that the linearity of such devices imposes an upper bound on the responsivity and, consequently, a lower bound on the length of the optical path required for achieving full modulation between $T=0$ and $T=1$. Such bounds cannot be exceeded unless a nonlinear response is being employed. The analysis presented here also provides some practical guidelines for optimizing the design of a modulator to enhance its responsivity.

Perturbation Theory. Consider an optical modulator consisting of an optical path of length $\Delta s = s_1 - s_0$. Let us consider the case where the light passes the optical path only once (in contrast to the case of a resonator where multiple reflections occur). At each point s along the optical path the field is expanded by using some local orthonormal basis. It is convenient to employ Dirac's notation,[1] (bra and ket) (even though no quantum effects are discussed in this paper). The field at point s is denoted $|\psi(s)\rangle$, which represents a column vector of amplitudes. The equation of motion is given by

$$\frac{\mathrm{d}}{\mathrm{d}s}|\psi\rangle = i\mathcal{K}|\psi\rangle, \tag{1}$$

where the Hermitian linear operator $\mathcal{K}$ is the Hamiltonian of the system. Consider the effect of adding a small perturbation $\epsilon\mathcal{K}_1(s)$ to the unperturbed Hamiltonian $\mathcal{K}_0$, namely,

$$\mathcal{K}(s) = \mathcal{K}_0(s) + \epsilon\mathcal{K}_1(s), \tag{2}$$

where $|\epsilon| \ll 1$ is a small real parameter. For any given ϵ the final state $|\psi_f\rangle = |\psi(s_1)\rangle$ is related to the initial state $|\psi_i\rangle = |\psi(s_0)\rangle$ by the relation

$$|\psi_f\rangle = U(\epsilon)|\psi_i\rangle, \tag{3}$$

where $U(\epsilon)$ is the s evolution operator for the Hamiltonian $\mathcal{K} = \mathcal{K}_0 + \epsilon\mathcal{K}_1$ from $s = s_0$ to $s = s_1$. The final state $|\psi_f\rangle$ is filtered by a polarizer having a normalized state $|\psi_p\rangle$. The transmission of the modulator is given by

$$T(\epsilon) = |\langle\psi_p|\psi_f(\epsilon)\rangle|^2. \tag{4}$$

Given the perturbed and unperturbed final states, $|\psi_f(\epsilon)\rangle$ and $|\psi_f(0)\rangle$, respectively, what is the optimum choice of a normalized $|\psi_p\rangle$ that will maximize $|\mathrm{d}T/\mathrm{d}\epsilon|$? Define the density operator

$$\rho(\epsilon) = |\psi_f(\epsilon)\rangle\langle\psi_f(\epsilon)| \tag{5}$$

and the operator

$$\Delta\rho = \rho(\epsilon) - \rho(0). \tag{6}$$

For a small ϵ one has

$$\frac{\mathrm{d}T}{\mathrm{d}\epsilon} = \frac{1}{\epsilon}\langle\psi_p|\Delta\rho|\psi_p\rangle. \tag{7}$$

The operator $\Delta\rho$ is Hermitian; thus the $|\psi_p\rangle$ that will maximize $|\mathrm{d}T/\mathrm{d}\epsilon|$ is the eigenvector of $\Delta\rho$ with the largest eigenvalue in absolute value. The nonvanishing eigenvalues of $\Delta\rho$ are given by $\pm[1-|\langle\psi_f(\epsilon)|\psi_f(0)\rangle|^2]^{1/2}$; thus

$$\left|\frac{\mathrm{d}T}{\mathrm{d}\epsilon}\right| \leqslant \frac{1}{|\epsilon|}[1-|\langle\psi_i|U^\dagger(0)U(\epsilon)|\psi_i\rangle|^2]^{1/2}. \tag{8}$$

Using perturbation expansion,[1] one finds, to second order in ϵ,

$$\langle\psi_i|U^\dagger(0)U(\epsilon)|\psi_i\rangle = 1 + i\epsilon\int_{s_0}^{s_1}\mathrm{d}s'\langle\mathcal{K}_1(s')\rangle - \epsilon^2\int_{s_0}^{s_1}\mathrm{d}s'\int_{s_0}^{s'}\mathrm{d}s''\langle\mathcal{K}_1(s')\mathcal{K}_1(s'')\rangle, \tag{9}$$

where the symbol $\langle\ \rangle$ represents the expectation value, namely, $\langle A\rangle = \langle\psi_i|A_H|\psi_i\rangle$ for a general operator A, where A_H is defined as

$$A_H(s) \equiv u_0^\dagger(s,s_0) A u_0(s,s_0), \tag{10}$$

and $u_0(s,s_0)$ is the s evolution operator from s_0 to s generated by $\mathcal{K}_0$. Since $\mathcal{K}_1(s)$ is Hermitian, one finds, to lowest order in ϵ,

$$|\langle\psi_i|U^\dagger(0)U(\epsilon)|\psi_i\rangle|^2 = 1 - \epsilon^2 \int_{s_0}^{s_1} \mathrm{d}s' \int_{s_0}^{s_1} \mathrm{d}s'' \times \langle \Delta\mathcal{K}_1(s')\Delta\mathcal{K}_1(s'')\rangle, \tag{11}$$

where $\Delta\mathcal{K}_1(s)=\mathcal{K}_1(s)-\langle\mathcal{K}_1(s)\rangle$. Thus

$$\left|\frac{\mathrm{d}T}{\mathrm{d}\epsilon}\right|^2 \leqslant \left|\int_{s_0}^{s_1} \mathrm{d}s' \int_{s_0}^{s_1} \mathrm{d}s'' \langle \Delta\mathcal{K}_1(s')\Delta\mathcal{K}_1(s'')\rangle\right|. \tag{12}$$

This upper bound imposed on the responsivity can be further simplified by employing the Schwartz inequality

$$\left|\frac{\mathrm{d}T}{\mathrm{d}\epsilon}\right|^2 \leqslant \int_{s_0}^{s_1} \mathrm{d}s' \{\langle[\Delta\mathcal{K}_1(s')]^2\rangle\}^{1/2}. \tag{13}$$

Thus the obtained upper bound is determined by integrating the standard deviation of the operator $\mathcal{K}_1$ with respect to the local unperturbed state along the optical path. This result suggests that responsivity can be enhanced by employing a perturbation $\mathcal{K}_1$ for which the standard deviation is maximized.

Two-Mode Case. Consider the case where the dimensionality of $|\psi(s)\rangle$ is two. Ignoring a common phase factor, the Hermitian operator $\mathcal{K}_1$ can be assumed to be traceless; thus it can be expressed as

$$\mathcal{K}_1 = \boldsymbol{\kappa}_1 \cdot \boldsymbol{\sigma}, \tag{14}$$

where $\boldsymbol{\kappa}_1=|\boldsymbol{\kappa}_1|\hat{\kappa}_1$ is a three-dimensional real vector with length $|\boldsymbol{\kappa}_1|$ ($\hat{\kappa}_1$ is a unit vector) and the components of the Pauli matrix vector[1] $\boldsymbol{\sigma}$ are given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{15}$$

It is straightforward to show that inequality (13) for the present case yields

$$\left|\frac{\mathrm{d}T}{\mathrm{d}\epsilon}\right| \leqslant \int_{s_0}^{s_1} \mathrm{d}s' |\boldsymbol{\kappa}_1(s')|. \tag{16}$$

A similar upper bound can be found for the angle θ between the polarization unit vectors $\mathbf{p}(\epsilon) = \langle\psi_i|U^\dagger(\epsilon)\boldsymbol{\sigma}U(\epsilon)|\psi_i\rangle$ and $\mathbf{p}(0)=\langle\psi_i|U^\dagger(0)\boldsymbol{\sigma}U(0)|\psi_i\rangle$ on the Bloch sphere. Using Eq. (11) and assuming the case $\theta \ll 1$, one finds

$$\theta \leqslant 2\int_{s_0}^{s_1} \mathrm{d}s' |\epsilon\boldsymbol{\kappa}_1(s')|. \tag{17}$$

Full modulation between $T=0$ and $T=1$ requires that the total change in θ exceed π (assuming $|\psi_p\rangle$ is kept fixed). Thus, if the applied birefringence strength is bounded by $|\epsilon\boldsymbol{\kappa}_1(s')| \leqslant \kappa_{\max}$, full modulation occurs only for

$$\Delta s \geqslant \pi/2\kappa_{\max}. \tag{18}$$

Thus the linearity of the system imposes a lower bound on the optical path length Δs required for achieving full modulation.

Examples. As a simple example, consider a modulator based on an optical fiber. Circularly polarized light is injected into the fiber, and a polarizer located at the fiber end allows transmission of only linearly polarized light. Modulation is achieved by applying linear birefringence along some section of the fiber of length Δs.

For the present example let us choose $|\psi\rangle=|+;\hat{2}\rangle$ at $s=0$ ($|\pm;\hat{u}\rangle$, with $\hat{u}$ being a unit vector, denotes an eigenvector of $\boldsymbol{\sigma}\cdot\hat{u}$ with an eigenvalue ± 1), and the polarizer state is $|\psi_p\rangle=|+;\hat{3}\rangle$. Moreover, $\mathcal{K}_0=0$ and $\mathcal{K}_1 = \boldsymbol{\kappa}\cdot\boldsymbol{\sigma}$, where $\boldsymbol{\kappa}=(1/2)\,(k_1,0,0)$. Integrating the equation of motion yields

$$T(\epsilon) = \sin^2\left(\frac{\epsilon k_1 s_1}{2} - \frac{\pi}{4}\right). \tag{19}$$

Thus at $\epsilon=0$ the derivative $|\mathrm{d}T/\mathrm{d}\epsilon|$ approaches the bound given by inequality (16). Moreover, full modulation is obtained for $s_0=-\pi/2\epsilon k_1$ and $s_1=\pi/2\epsilon k_1$; thus for this case the bound given by inequality (18) is also achieved.

The next example deals with a modulator based on a transition between adiabatic and nonadiabatic regimes, as in Ref. 2. Consider the case where $\mathcal{K}=\boldsymbol{\kappa}\cdot\boldsymbol{\sigma}$,

$$\boldsymbol{\kappa}(s) = \gamma(0,[\lambda^2-(\gamma s)^2]^{1/2},\gamma s), \tag{20}$$

where γ is a real constant with dimensionality of 1/length, λ is a nonnegative dimensionless real parameter, and $|\gamma s| \leqslant \lambda$.

Consider the case where for $s_0=-\lambda/\gamma$ the state of the system is a local eigenstate of $\mathcal{K}(s)$ with a positive eigenvalue, namely, $|\psi(s_0)\rangle=|-;\hat{3}\rangle$. When $\lambda \gg 1$, the state evolves adiabatically[3] and remains parallel to $\boldsymbol{\kappa}(s)$. The polarizer is located at $s_1=\lambda/\gamma$, and its state is given by $|\psi_p\rangle=|-;\hat{3}\rangle$. Thus in the adiabatic limit $T=0$. The transition between adiabatic to nonadiabatic regimes occurs near $\lambda=1$. The approximation solution for the case $\lambda \gtrsim 1$ can be found by considering the lowest-order correction to the adiabatic limit.[2,4] The value of T is the probability that the Zener transition will occur, which can be calculated to lowest order:

$$T \simeq \frac{\pi^2}{4} J_0^2(2\lambda^2) \qquad (\text{for } \lambda \lesssim 1), \tag{21}$$

where J_0 is the Bessel function of the first kind of order 0.

This approximation is compared with the calculated value of $T(\lambda)$ obtained from numerical integration of the equation of motion. The case $\lambda=5$ is presented in Fig. 1 as an example. Figure 2(a) shows the

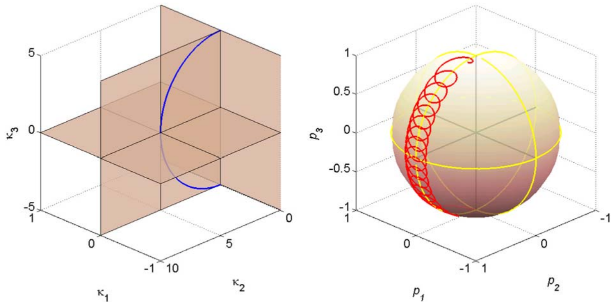

Fig. 1. (Color online) Example of numerical integration of the equation of motion for the case $\lambda = 5$. Left, curve $\kappa(s)$; right, evolution of the polarization vector $\mathbf{p}(s)$ on the Bloch sphere.

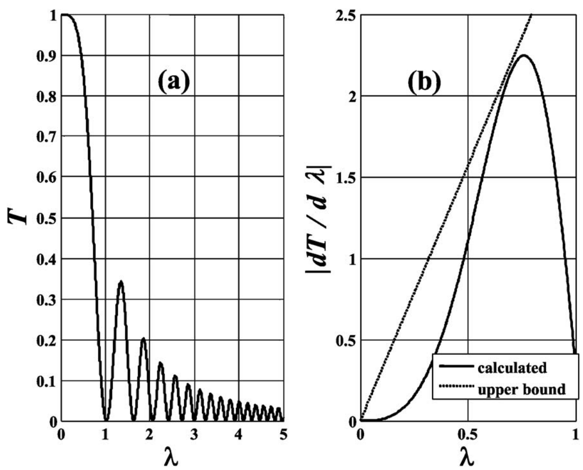

Fig. 2. Calculated and upper bound of responsivity. (a) Numerical calculation of T versus λ. (b) Comparison between the calculated $|\mathrm{d}T/\mathrm{d}\lambda|$ and the upper bound given by inequality (22).

calculated value of $T(\lambda)$ in the range $0 \leqslant \lambda \leqslant 5$. Comparing the approximate result in relation (21) with the numerical solution shows, as expected, good agreement for $\lambda \gtrsim 1$.

On the other hand, the upper bound given by inequality (16) for this case reads as

$$\left| \frac{\mathrm{d}T}{\mathrm{d}\lambda} \right| \leqslant \int_{\lambda/\gamma}^{\lambda/\gamma} \mathrm{d}s' \frac{\gamma\lambda}{[\lambda^2 - (\gamma s')^2]^{1/2}} = \pi\lambda. \qquad (22)$$

A comparison between the numerically calculated $|\mathrm{d}T/\mathrm{d}\lambda|$ and the above upper bound is seen in Fig. 2(b). Contrary to the previous example, in this case the upper bound is not reached for any value of λ. However, in the transition region, between the adiabatic and the nonadiabatic limits, near $\lambda = 0.695$ the responsivity is only some 2% below the upper bound. Similarly, for the modulator discussed in Ref. 2, it was found that the largest responsivity is obtained in the transition region between adiabatic and nonadiabatic limits.

Note that the bounds discussed in this Letter can be employed for other linear systems. For example, the same analysis may lead to a lower bound imposed on the time required for performing a given quantum gate on a system of quantum bits in a quantum computer, when the maximum perturbation strength is given.

The author thanks Avishai Eyal for very useful and stimulating discussions. E. Buks's e-mail address is eyal@ee.technion.ac.il.

References

1. M. Weissbluth, *Photon–Atom Interaction* (Academic, 1989).
2. E. Buks, J. Opt. Soc. Am. B **23**, 0000 (2006); arXiv: quant-ph/0412198.
3. M. V. Berry, Proc. R. Soc. London Ser. A **392**, 45 (1984).
4. A. B. Migdal, *Qualitative Methods in Quantum Theory* (Benjamin, 1977).