

Plane Wave Theory of Optical Bistability
in Reflection at a Nonlinear Interface

R. T. Deck, H. J. Simon, and G. M. Wysin

The University of Toledo
Toledo, OH 43606

Abstract

We consider the theory of optical bistability in reflection from a nonlinear interface. When an intense laser pulse is incident on a medium with an intensity dependent index of refraction near the critical angle for total internal reflection the reflected intensity may display hysteresis as a function of the incident intensity. We discuss Kaplan's plane wave theory for this phenomenon and derive a simple procedure to calculate the reflectivity. For a medium with a positive nonlinear coefficient we find that a significant region of bistable operation does exist. However for a medium with a negative nonlinear coefficient, limited bistability exists over only a small region. Estimates of the incident power required to achieve bistability are given.

I. Introduction

There is current interest in phenomena associated with optical bistability.⁽¹⁾ In a series of publications⁽²⁻⁶⁾ Kaplan has developed a plane wave theory of the optical bistability of light reflected from the surface of a nonlinear medium. An apparent observation of this phenomenon was first reported by Smith et.al.⁽⁷⁾ The experiment involved a ruby laser pulse incident at an angle exceeding the critical angle on a glass - CS₂ interface. By analysis of the shape of the reflected pulse in comparison with the input, the authors were able to infer the existence of a discontinuous transition between total internal reflection (TIR) and partial transmission as the interface responded to the changing field intensity in the incident pulse. Recently the same authors have reviewed the status of the comparison between experiment and theory.⁽⁸⁾

The explanation of the phenomenon seems to be straight-forward. In a nonlinear medium in which the dielectric constant depends on the amplitude of the electric field in the medium, the critical angle of incidence for TIR depends on the intensity of the incident light, and therefore as the intensity is increased for a fixed angle of incidence θ_i , it can happen that the critical angle is shifted to the opposite side of θ_i and that the amount of reflected light is changed from that corresponding to partial reflection to that corresponding to total reflection or vice versa. The resulting effect is interesting to the extent that the change occurs abruptly rather than gradually (i.e., discontinuously rather than continuously). The discontinuous type of change can occur only when the two-medium-interface system, for some range of incident intensities, simultaneously supports two stable modes of reflected or transmitted light, in which case the system is said to be "bistable".

We show that in order to investigate the possible optical bistability of a given two-medium-interface system it is necessary only to determine the

reflection coefficient R of the system as a function of the incident light intensity. In contrast to Kaplan's more formal approach to bistability theory, we here take this direct approach and present a simple procedure for calculating R . This procedure allows direct calculation of optical bistability in reflection for all regions of interest; e.g., for positive and negative nonlinearities, arbitrary angles of incidence, wave guide geometries, etc. The present authors have applied this method to the calculation of optical bistability with surface plasmons.⁽⁹⁾

We focus here exclusively on the case of s-polarization where the electric field vector of the incident wave is perpendicular to the plane of incidence. Since this case can be analyzed exactly, it allows the conditions for bistability to be exhibited clearly. In order to derive the equations for the reflectivity it is necessary to follow Kaplan's method of solution of the wave equation in the nonlinear medium. On the other hand, for our analysis it is sufficient to determine only the derivative and phase of the field at the interface. This is done in Section II(a). In Section II(b) particular cases corresponding to positive and negative nonlinearities are considered and the nature of the possible bistability is discussed. In addition we display here computed graphs of reflectivity versus incident intensity for certain representative cases. We conclude with some summary comments in Section III.

II. Theory

Consider a plane electromagnetic wave incident from a linear medium into a medium which exhibits an optical Kerr effect. The plane of propagation is defined to be the x-z plane with the boundary interface at $z=0$. The dielectric constant ϵ_t of the nonlinear medium is assumed to be of the form

$$\epsilon_t = \epsilon_t^0 + \alpha |E_t|^2, \quad (1)$$

where ϵ_t^0 is the dielectric constant of the nonlinear medium at zero intensity, E_t the field in the nonlinear medium, and α a nonlinearity constant which is connected to the optical Kerr dielectric constant n_2 via the relation

$$n_2 = \frac{4\pi}{c\epsilon_t^0} \alpha.$$

In the case of s-polarization the spatial dependence of the incident, reflected, and transmitted fields can be represented by the respective functions:⁽¹⁰⁾

$$\vec{E}_i = \hat{y} E_i e^{i(k_{ix}x + k_{iz}z)}, \quad \vec{E}_r = \hat{y} E_r e^{i(k_{ix}x - k_{iz}z)} \quad \text{and} \quad \vec{E}_t = \hat{y} E_t(x,z),$$

where k_{ix} and k_{iz} represent the components of the incident propagation vector in the linear medium with magnitude $k_i = \frac{\omega}{c} \sqrt{\epsilon_i}$, and \vec{E}_t corresponds to a (transverse wave) solution of Maxwell's equations in the nonlinear medium. For a (non-magnetic) medium described by the dielectric constant ϵ_t given above, and \vec{E}_t polarized perpendicular to the plane of propagation (as in s-polarization) the latter equations are equivalent to the vector wave equation

$$\nabla^2 \vec{E}_t + \frac{\omega^2}{c^2} [\epsilon_t^0 + \alpha |E_t|^2] \vec{E}_t = 0. \quad (2)$$

We are interested in determining the ratio of the reflected to the incident intensity, $|E_r|^2/|E_i|^2$. This ratio is strictly determined by Maxwell's equations (which require the above forms for \vec{E}_i and \vec{E}_r in the linear medium) and by the associated boundary conditions at the interface $z=0$. The latter boundary conditions require \vec{E}_t to have an x-dependence of the form

$$E_t(x, z) = \mathcal{E}(z) e^{ik_i x} , \quad (3)$$

and result in connections between the amplitudes E_i , E_r , and $\mathcal{E}(0)$,

$$E_i + E_r = \mathcal{E}(0) \quad (4.a)$$

$$k_{iz}(E_r - E_i) = i \frac{d\mathcal{E}(0)}{dz} , \quad (4.b)$$

which we choose to re-express in the form

$$E_r = \frac{1}{2} \left[\mathcal{E}(0) + \frac{i}{k_{iz}} \frac{d\mathcal{E}(0)}{dz} \right] \quad (5.a)$$

$$E_i = \frac{1}{2} \left[\mathcal{E}(0) - \frac{i}{k_{iz}} \frac{d\mathcal{E}(0)}{dz} \right] . \quad (5.b)$$

For a given incident amplitude E_i relations (4) or (5) provide two equations for the three unknown quantities E_r , $\mathcal{E}(0)$ and $\frac{d\mathcal{E}(0)}{dz}$. An additional connection between the last two quantities is provided by Eq. (2) which requires $\mathcal{E}(z)$ to satisfy the equation

$$\frac{d^2 \mathcal{E}}{dz^2} + \left[\frac{\omega^2}{c^2} (\epsilon_t^0 + \alpha |\mathcal{E}|^2) - k_{ix}^2 \right] \mathcal{E} = 0 . \quad (6)$$

Following Kaplan^(5,6) it is useful to re-express the complex function $\mathcal{E}(z)$ in terms of real amplitude and phase functions $U(z)$ and $\zeta(z)$ via the definition

$$\mathcal{E}(z) \equiv \sqrt{\frac{\epsilon_i}{|\alpha|}} U(z) e^{i\zeta(z)} , \quad (7)$$

where the factor $\sqrt{\frac{\epsilon_i}{|\alpha|}}$ is included in the amplitude so as to simplify certain later equations. With no loss of generality the phase function can be written in the form

$$\zeta(z) = k_i \int_0^z \xi(z') dz' + \zeta(0) , \quad (8)$$

with $k_i \xi(z) = \frac{d\zeta(z)}{dz}$. By use of Eqs. (7) and (8) in Eq. (6), the real and imaginary parts of that equation separate into the two equations

$$\frac{d^2U}{dz^2} = -k_i^2 \left(\frac{\epsilon_t^0}{\epsilon_i} - \sin^2 \theta_i + \frac{\alpha}{|\alpha|} U^2 - \xi^2 \right) U, \quad (9)$$

$$\frac{d}{dz} (U^2) = 0. \quad (10)$$

The ratio ϵ_t^0/ϵ_i in Eq. (9) defines the zero field critical angle of incidence θ_c^0 by the relation

$$\frac{\epsilon_t^0}{\epsilon_i} = \sin^2 \theta_c^0. \quad (11)$$

The subsequent analysis of the above equations can be made independent of the magnitude of α (and dependent only on the sign of α) by introduction of the dimensionless intensities

$$U_i \equiv \frac{|\alpha|}{\epsilon_i} |E_i|^2, \quad U_r \equiv \frac{|\alpha|}{\epsilon_i} |E_r|^2, \quad U_t \equiv \frac{|\alpha|}{\epsilon_i} |\mathcal{E}(0)|^2 = [U(0)]^2 \quad (12)$$

in terms of which the absolute squares of Eqs. (5) assume the form

$$U_r = \frac{1}{4} \left\{ \left[1 - \frac{k_i}{k_{iz}} \xi(0) \right]^2 U_t + \frac{1}{k_{iz}^2} \left[\frac{dU(0)}{dz} \right]^2 \right\} \quad (13.a)$$

$$U_i = \frac{1}{4} \left\{ \left[1 + \frac{k_i}{k_{iz}} \xi(0) \right]^2 U_t + \frac{1}{k_{iz}^2} \left[\frac{dU(0)}{dz} \right]^2 \right\}. \quad (13.b)$$

In addition it is convenient to introduce the notation

$$\Delta \equiv \sin^2 \theta_c^0 - \sin^2 \theta_i. \quad (14)$$

To determine the reflection coefficient $R = U_r/U_i$ from Eqs. (13) it is necessary only to have expressions for $\xi(0)$ and $\frac{dU(0)}{dz}$ in terms of the quantity U_t . In Section II(a) the required expressions are extracted from Eqs. (9) and (10). Since the resulting Eqs. (13) are in general nonlinear, these equations have no simple analytic solutions. Instead it is convenient to adopt an indirect method of solution of Eqs. (13) based on their re-interpretation

as equations for U_r and U_i in terms of the "parameter" U_t . By incrementing U_t over an appropriate range of values and computing corresponding values of U_r and U_i from Eqs. (13), we are then able to determine the variation of U_r with U_i quite simply.

It is worth noting on the basis of Eqs. (13) that solutions of Eqs. (9) and (10) corresponding to a zero value of $\xi(0)$ will result in equal values of U_r and U_i and will therefore correspond to the physical situation of TIR. In this case the amplitude $U(z)$ will be attenuated as a function of z and the field in the nonlinear medium will propagate as a surface wave in the approximate vicinity of the $z = 0$ boundary.

(a) Integration of Nonlinear Wave Equation

We are interested in solutions of Eqs. (9) and (10) which represent waves propagating along or away from the boundary in the nonlinear medium. (11)

In the case of such solutions the amplitude U at $z = \infty$ needs to approach a constant value U_∞ which (in the absence of damping) can be non-zero. The appropriate boundary conditions on U and its derivatives at infinity therefore can be expressed as

$$U^2(z) \xrightarrow{z \rightarrow \infty} U_\infty^2, \quad (15.a)$$

$$\lim_{z \rightarrow \infty} \frac{dU}{dz} = \lim_{z \rightarrow \infty} \frac{d^2U}{dz^2} = 0. \quad (15.b)$$

In terms of U_∞ and ξ_∞ Eq. (10) assumes the form

$$U^2 \xi = U_\infty^2 \xi_\infty \quad (10)'$$

where the value ξ_∞ is obtained for $U_\infty \neq 0$ by use of condition (15.b) in (9),

$$\xi_\infty^2 = \Delta + \frac{\alpha}{|\alpha|} U_\infty^2, \quad U_\infty^2 \neq 0. \quad (16)$$

In the special case that U_∞^2 vanishes, (10)' (for non-zero U^2) requires ξ to be identically zero, which value corresponds to the case of TIR. It follows in general that ξ^2 can be expressed in terms of U^2 as

$$\xi^2 = \left(\Delta + \frac{\alpha}{|\alpha|} U_\infty^2 \right) \frac{U_\infty^4}{U^4} \quad (17)$$

Making use of Eq. (17) and writing $\frac{d}{dz} = \frac{dU}{dz} \frac{d}{dU}$, $\frac{d^2U}{dz^2} = \frac{1}{2} \frac{d}{dU} \left(\frac{dU}{dz} \right)^2$, Eq. (12) can be integrated once to obtain an equation for $\frac{dU}{dz}$, which, subject to the boundary condition (15.b), reduces to

$$U^2 \left(\frac{dU}{dz} \right)^2 = -k_i^2 (U^2 - U_\infty^2)^2 \left[\Delta + \frac{\alpha}{|\alpha|} U_\infty^2 + \frac{1}{2} \frac{\alpha}{|\alpha|} U^2 \right] \quad (18)$$

Eqs. (17) and (18) provide the relations between ξ , $\frac{dU}{dz}$, and $U_t = U^2(0)$ required for evaluation of Eqs. (13). In what follows we analyze the implications of these equations in the separate cases of a positive and negative nonlinearity constant α . The interest in the analysis is in the possibility of a bistable variation in the reflected intensity as a function of the incident intensity. Such a bistable variation requires the existence of a transition between the transmission and TIR modes of the boundary interface which can occur in the case of a positive (negative) nonlinearity only where the angle of incidence θ_i is greater (less) than the zero field critical angle θ_0^C , or equivalently Δ is greater (less) than zero. The above conclusions can (also) be seen to follow directly from Eqs. (9) and (18). In particular, since in the case α and Δ positive, the quantities on the right and left hand sides of (18) have opposite signs when $U_\infty^2 = \xi = 0$, there can be no solution of (18) corresponding to the TIR mode. On the other hand since, in the case α and Δ negative, the factor multiplying U on the right hand side of (9) is positively definite, $\frac{d^2U}{dz^2}$ can vanish at infinity only if U vanishes at infinity, and therefore there can in this case be only a TIR mode (with $U_\infty^2 = \xi = 0$).

The possibility for a bistable transition between the TIR and transmission modes thus exists only in the two remaining cases, $\alpha > 0, \Delta < 0$, and $\alpha < 0, \Delta > 0$, which we consider in the following. For bistability to be exhibited in Eqs. (13) in either of these cases the intensity U_T for some range of U_i will need to be a multi-valued function of U_i .

(b) Analysis of Specific Cases

Case of Positive α and Negative Δ

In the case of $\alpha > 0, \Delta < 0$, relations (17) and (18) result in the equations

$$\xi^2 = (U_\infty^2 - |\Delta|) \frac{U_\infty^4}{U^4} \quad , \quad (19)$$

$$U^2 \left(\frac{dU}{dz} \right)^2 = k_i^2 (U^2 - U_\infty^2)^2 \left[|\Delta| - \frac{1}{2} U^2 - U_\infty^2 \right] \quad , \quad (20)$$

which, since ξ and U are required to be real, have allowed solutions only when their right hand sides are non-negative. Applied to Eq. (19), this requirement restricts U_∞^2 to the values

$$U_\infty^2 = 0, \quad \text{or} \quad U_\infty^2 \geq |\Delta| \quad ,$$

while, applied to Eq. (20), it restricts U_∞^2 to the values

$$U_\infty^2 = U^2, \quad \left(\frac{dU}{dz} = 0 \right), \quad \text{or} \quad U_\infty^2 < |\Delta| \quad .$$

Since two of the latter possibilities are mutually exclusive, there remain only the possibilities $U_\infty^2 = 0$ and $U_\infty^2 = U^2$. The first leads to an allowed solution of (19) and (20) (with $\xi = 0$) only if U^2 is less than or equal to $2|\Delta|$,

$$U_\infty^2 = \xi^2 = 0, \quad U^2 \leq 2|\Delta| \quad , \quad (21.a)$$

while the second leads to an allowed solution of (19) and (20) only if $U^2 (= U_\infty^2)$ is greater than or equal to $|\Delta|$,

$$U_{\infty}^2 = U^2, \quad \frac{dU}{dz} = 0, \quad U^2 \geq |\Delta| \quad (21.b)$$

In the case of a solution of type (21.a), (19) and (20) result in the equations

$$\begin{aligned} \xi &= 0 & U^2 &\leq 2|\Delta| & (22) \\ \left(\frac{dU}{dz}\right)^2 &= k_i^2 \left[|\Delta| - \frac{1}{2} U^2 \right] U^2, \end{aligned}$$

and Eqs. (13) reduce to the TIR relation

$$U_r = U_i = \frac{1}{4} (1 + \sec^2 \theta_i [|\Delta| - \frac{1}{2} U_t]) U_t, \quad U_t \leq 2|\Delta| \quad (23)$$

which is equivalent to the set of equations

$$\begin{aligned} U_r &= \frac{1}{4} \left| 1 - i \sec \theta_i \sqrt{\frac{1}{2} U_t - |\Delta|} \right|^2 U_t, \\ U_i &= \frac{1}{4} \left| 1 + i \sec \theta_i \sqrt{\frac{1}{2} U_t - |\Delta|} \right|^2 U_t. \end{aligned} \quad U_t \leq 2|\Delta| \quad (23)'$$

In the alternative case of a solution of type (21.b), (19) and (20) result in the equations

$$\begin{aligned} \xi^2 &= U^2(0) - |\Delta| = U_t - |\Delta|, \\ \frac{dU}{dz} &= 0, \end{aligned} \quad U^2 \geq |\Delta| \quad (24)$$

and Eqs. (13) reduce to the relations

$$\begin{aligned} U_r &= \frac{1}{4} (1 - \sec \theta_i \sqrt{U_t - |\Delta|})^2 U_t \\ U_i &= \frac{1}{4} (1 + \sec \theta_i \sqrt{U_t - |\Delta|})^2 U_t \end{aligned} \quad U_t \geq |\Delta| \quad (25)$$

Equations (25) are completely equivalent to the usual Fresnel relations connecting the amplitudes of incident, reflected and transmitted plane waves at a boundary between two media with dielectric constants ϵ_i and $\epsilon_t = \epsilon_t^0 + \alpha |\mathbf{E}|^2$. Similarly Eqs. (23)' are equivalent to the Fresnel relations at a boundary

between two media with dielectric constants ϵ_i and ϵ_t' with

$$\epsilon_t' = \epsilon_t^0 + \frac{1}{2} \alpha |\mathbf{E}|^2 . \quad (1)'$$

The nonlinear interface for $\alpha > 0$ therefore effectively supports two distinct modes of reflection corresponding to two nonlinear dielectric constants ϵ_t and ϵ_t' and associated with two effective critical angles θ_c and θ_c' given by

$$\sin^2 \theta_c = \frac{\epsilon_t}{\epsilon_i} = \sin^2 \theta_i + \Delta + U_t \quad (26.a)$$

$$\sin^2 \theta_c' = \frac{\epsilon_t'}{\epsilon_i} = \sin^2 \theta_i + \Delta + \frac{1}{2} U_t . \quad (26.b)$$

Examination of the restrictions on U_t associated with the Eqs. (23) and (25) shows that, for values of U_t within the range

$$|\Delta| \leq U_t \leq 2|\Delta| , \quad (27)$$

the two modes of reflection can coexist. In particular, for values of U_t within this range there exist two possible connections between U_r , U_i and U_t and two possible values of U_r for each value of U_i . The range of values of incident intensity U_i corresponding to the range (27) can be gotten directly from Eqs. (27) and (23) by setting U_t equal to the extreme values $|\Delta|$ and $2|\Delta|$ respectively, with the result

$$\frac{|\Delta|}{4} \leq U_i \leq \frac{|\Delta|}{2} . \quad (27)'$$

For values of U_i within this range the system can be expected to be bistable.

This can be confirmed for given values of the parameters Δ and θ_i by numerical evaluation of the reflectivity U_r/U_i determined by Eqs. (23) and (25), with U_t varied as a parameter in the manner discussed following Eqs. (13). Figure (1) shows the results of such an evaluation (for positive α) with the linear critical angle θ_c^0 chosen to be 88° and θ_i taken to be 88.5° . (12)

Here, between the values U_i^C and U_i^S defined by the extreme values in Eq. (27)', each value of U_i is associated with two distinct values of R , one corresponding to partial transmission and the other to TIR. For small values of incident intensity U_i , with the angle of incidence θ_i greater than θ_c^0 , the incident field must be totally reflected and the physical value of R lies along the R -equals-unity curve corresponding to the TIR solution (23). As the incident intensity is increased, however, the field dependence of the nonlinear dielectric constant ϵ_t' increases the critical angle θ_c' corresponding to this solution in the direction of θ_i until at the value U_i^S equal to $\frac{|\Delta|}{2}$, θ_c' becomes equal to θ_i . Beyond this value of U_i the only allowed value of R is that corresponding to the transmission mode solution defined by Eq. (25) and the reflectivity must therefore discontinuously switch to the transmission curve at $U_i = U_i^S$ and advance along this curve as U_i is further increased. On the other hand when U_i is then decreased from values above U_i^S to values below U_i^S the operating point can remain on this curve as R changes in accord with the transmission mode solution of Eq. (25), until at $U_i = U_i^C = \frac{|\Delta|}{4}$, θ_c' becomes equal to θ_i and the interface returns to the TIR mode. The multivaluedness of R in the region of U_i defined by (27)' therefore leads to hysteresis in the reflectivity as a function of U_i .

For given values of the parameters, bistability can be observed only if the dimensionless intensity U_i can be increased at least to the switching value U_i^S . In addition, the discontinuity at the switching value is maximally large only when $\epsilon_t(\epsilon_t')$ is close to ϵ_i , in which case the reflectivity shifts at the critical angle from a value equal to unity to a value near zero. But if ϵ_t is close to ϵ_i , the critical angle is near grazing incidence. On the other hand, since the nonlinear term $\alpha|\mathcal{E}|^2$ in ϵ_t is in practice quite small, the critical angle can be shifted from one side to the other of the angle of

incidence as $|\xi|^2$ varies only if the initial offset angle $|\theta_i - \theta_c^0|$ is also quite small and therefore θ_i also close to 90° . Therefore, although the present analysis shows that bistability can occur in principle for angles far removed from 90° , the switching intensity required at such angles to achieve a significant effect may be increased by an order of magnitude relative to that required for the grazing incidence case corresponding to Figure 1.

The numerical values of the critical angle and offset angle for Figure 1 are chosen to be those appropriate to the recent experiment of Smith et. al.⁽⁷⁾ For this experiment which used CS_2 as the nonlinear medium we calculate the incident critical switching power to be 7×10^9 watts/cm². Since the present plane wave analysis ignores the Gaussian shape of the incident beam, the excellent agreement between this calculated switching power and that observed in the experiment may be somewhat fortuitous.

Case of Negative α and Positive Δ

In the case $\alpha < 0$, $\Delta > 0$ relations (17) and (18) result in the equations

$$\xi^2 = (\Delta - U_\infty^2) \frac{U_\infty^4}{U^4} \quad (28)$$

$$U^2 \left(\frac{dU}{dz} \right)^2 = k_i^2 (U^2 - U_\infty^2)^2 \left[U_\infty^2 + \frac{1}{2} U^2 - \Delta \right], \quad (29)$$

which again have allowed solutions only when their right hand sides are non-negative. Since, whenever $U_\infty^2 < \frac{2}{3} \Delta$ (for $U^2 \neq U_\infty^2$), the right hand side of (29) must become negative as U^2 approaches U_∞^2 , Eq. (29) is inconsistent with a value of U_∞^2 less than $\frac{2}{3} \Delta$, unless $U^2 = U_\infty^2$. On the other hand, for all values of U^2 , Eq. (28) is inconsistent with a value of U_∞^2 greater than Δ . It follows that allowed solutions of Eqs. (28) and (29) are constrained to have values of U_∞^2 in the interval $\frac{2}{3} \Delta \leq U_\infty^2 \leq \Delta$, when $U^2 \neq U_\infty^2$, and to have values

