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 — CS_2 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

$$\epsilon_t = \epsilon_t^{\,o} + \alpha|E_t|^2 \quad , \tag{1}$$

where $\epsilon_t^{\,o}$ 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 $n_2 = \frac{4\pi}{c\epsilon_t^{\,o}}\,\alpha$.

In the case of s-polarization the spatial dependence of the incident, reflected, and transmitted fields can be represented by the respective functions:[10] $\vec{E}_i = \hat{y}\,E_i\,e^{i(k_{i_x}x + k_{i_z}z)}$, $\vec{E}_r = \hat{y}\,E_r\,e^{i(k_{i_x}x - k_{i_z}z)}$ and $\vec{E}_t = \hat{y}\,E_t(x,z)$, where k_{i_x} and k_{i_z} 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

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

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

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

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

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

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

which we choose to re-express in the form

$$E_r = \frac{1}{2}\,[\mathcal{E}(0) + \frac{i}{k_{i_z}}\,\frac{d\mathcal{E}(0)}{dz}] \tag{5.a}$$

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

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

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

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

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

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

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

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

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

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

The ratio $\varepsilon_t{}^0/\varepsilon_i$ in Eq. (9) defines the zero field critical angle of incidence $\theta_c{}^0$ by the relation

$$\frac{\varepsilon_t{}^0}{\varepsilon_i} = \sin^2\theta_c{}^0 \quad . \tag{11}$$

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

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

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

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

$$U_i = \frac{1}{4}\left\{[1 + \frac{k_i}{k_{i_z}}\,\xi(0)]^2 U_t + \frac{1}{k_{i_z}^2}\,\left[\frac{dU(0)}{dz}\right]^2\right\} \quad . \tag{13.6}$$

In addition it is convenient to introduce the notation

$$\Delta \equiv \sin^2\theta_c{}^0 - \sin^2\theta_i \quad . \tag{14}$$

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

$$U^2(z) \xrightarrow[z \to \infty]{} U_\infty^2 \quad , \tag{15.a}$$

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

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

$$U^2\xi = U_\infty^2 \xi_\infty \tag{10$'$}$$

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

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

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

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

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

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

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_r 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

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

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

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

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

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 $2|\Delta|$,

$$U_\infty^2 = \xi^2 = 0, \qquad U^2 \leq 2|\Delta| \quad , \tag{21.a}$$

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 $|\Delta|$,

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

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

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

and Eqs. (13) reduce to the TIR relation

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

which is equivalent to the set of equations

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

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

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

and Eqs. (13) reduce to the relations

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

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 $\varepsilon_t = \varepsilon_t^0 + \alpha|\mathcal{E}|^2$. Similarly Eqs. (23)$'$ are equivalent to the Fresnel relations at a boundary between two media with dielectric constants ε_i and ε_t' with

$$\varepsilon_t' \equiv \varepsilon_t^{o} + \frac{1}{2}\alpha\,|\mathcal{E}|^2 \quad . \tag{1'}$$

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

$$\sin^2\theta_c = \frac{\varepsilon_t}{\varepsilon_i} = \sin^2\theta_i + \Delta + U_t \tag{26.a}$$

$$\sin^2\theta_c' = \frac{\varepsilon_t'}{\varepsilon_i} = \sin^2\theta_i + \Delta + \frac{1}{2}U_t \quad . \tag{26.b}$$

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

$$|\Delta| \le U_t \le 2|\Delta| \quad , \tag{27}$$

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

$$\frac{|\Delta|}{4} \le U_i \le \frac{|\Delta|}{2} \quad . \tag{27'}$$

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^{o} 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^o, 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 $\varepsilon_t(\varepsilon_t')$ 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 $|\mathcal{E}|^2$ varies only if the initial offset angle $|\theta_i - \theta_c^o|$ 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

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

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

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 $\frac{2}{3}\Delta \le U_\infty^2 \le \Delta$, when $U^2 \neq U_\infty^2$, and to have values of U_∞^2 less than or equal to Δ when $U^2 = U_\infty^2$. Since the former condition effectively eliminates the TIR solution, $U_\infty^2 = \xi^2 = 0$, $U^2 \neq U_\infty^2$, bistable switching between transmission and TIR modes cannot occur in this case. On the other hand, because, with U_∞^2 in the interval $\frac{2}{3}\Delta \leq U_\infty \leq \Delta$, there exist more than one form of transmission mode solution (with $\xi^2 \neq 0$), there exists the possibility of a bistable switching between distinct transmission modes.

It is emphasized that the transmission modes with $U_\infty^2 \neq U^2$ (in the interval $\frac{2}{3}\Delta \leq U_\infty^2 \leq \Delta$) are quite distinct from the transmission modes previously encountered with $U_\infty^2 = U^2$. Since the fields associated with the latter modes have constant amplitude ($= U_\infty$), these modes can be referred to as homogenous plane waves (PW). In contrast, Kaplan refers to the transmission modes with $U_\infty^2 \neq U^2$ as "longitudinally in-homogeneous travelling waves" (LITW).(5,6) In these modes the amplitude, phase, and direction of propagation of the associated fields all vary with z.

Because the boundary conditions at $z = 0$ and at $z = \infty$ are insufficient to determine the exact form of the LITW modes, there is in general a continuum of possible such modes with U_∞^2 lying in the interval $\frac{2}{3}\Delta \leq U_\infty^2 \leq \Delta$. By analysis of these modes in Ref. 5, however, Kaplan has succeeded in demonstrating that only the LITW mode with $U_\infty^2 = \frac{2}{3}\Delta$ persists in the presence of an infinitesimal amount of damping. Moreover, by the same analysis, Kaplan has eliminated the plane wave modes with $U_\infty^2 = U^2 > \frac{2}{3}\Delta$. Since with $U_\infty^2 = \frac{2}{3}\Delta$, the right hand side of Eq. (29) can be nonnegative for $U^2 \neq U_\infty^2$ only if $U^2 \geq \frac{2}{3}\Delta$, the value of $U^2(0) = U_t$ corresponding to the LITW mode is restricted by the condition $U_t \geq \frac{2}{3}\Delta$. It therefore follows that the physically allowed solutions of Eqs. (28) and (29) are of only the two types.

$$U_\infty^2 = U^2 \neq 0, \qquad \frac{dU}{dz} = 0, \qquad U^2 \leq \frac{2}{3}\Delta \tag{30.a}$$

$$U_\infty^2 = \frac{2}{3}\Delta, \qquad\qquad U^2 \geq \frac{2}{3}\Delta \ . \tag{30.b}$$

In the case of a PW solution of type (30.a), (28) and (29) result in the equations

$$\begin{aligned} \xi^2 &= (\Delta - U^2(0)) = (\Delta - U_t) \quad , \\ \frac{dU}{dz} &= 0 \quad , \end{aligned} \qquad U_t \le \frac{2}{3}\Delta \ , \qquad (31)$$

and Eqs. (13) reduce to the relations

$$\begin{aligned} U_r &= \frac{1}{4}\,[1 - \sec\theta_i \sqrt{\Delta - U_t}]^2\, U_t \quad , \\ U_i &= \frac{1}{4}\,[1 + \sec\theta_i \sqrt{\Delta - U_t}]^2\, U_t \end{aligned} \qquad U_t \le \frac{2}{3}\Delta \ . \qquad (32)$$

In the alternative case of an LITW solution of type (30.b), (28) and (29) result in the equations

$$\begin{aligned} \xi &= 2(\tfrac{\Delta}{3})^{3/2}/U^2 \quad , \\ (\frac{dU}{dz})^2 &= \frac{1}{2}\,k_i{}^2(U^2 - \tfrac{2}{3}\Delta)^3/U^2 \quad , \end{aligned} \qquad U^2 \ge \frac{2}{3}\Delta \qquad (33)$$

and Eqs. (13) produce the relations

$$\begin{aligned} U_r &= \frac{1}{4}\left\{[1 - \sec\theta_i\, 2(\tfrac{\Delta}{3})^{3/2}/U_t]^2\, U_t + \frac{1}{2}\sec^2\theta_i\, \frac{(U_t - \frac{2}{3}\Delta)^3}{U_t}\right\} \ , \\ U_i &= \frac{1}{4}\left\{[1 + \sec\theta_i\, 2(\tfrac{\Delta}{3})^{3/2}/U_t]^2\, U_t + \frac{1}{2}\sec^2\theta_i\, \frac{(U_t - \frac{2}{3}\Delta)^3}{U_t}\right\} \ . \end{aligned} \qquad U_t \ge \frac{2}{3}\Delta \qquad (34)$$

For $U_t = \frac{2}{3}\Delta$, (where $U_i = \frac{\Delta}{6}\,(1 + \sqrt{\frac{\Delta}{3}}\sec\theta_i)^2$) the two sets of relations (32) and (34) coincide.

Since the phase function ξ in (33) varies with z (as does U^2), the direction of propagation of the LITW mode, determined by the gradient of the phase of $E_t(x,z)$, also varies with z. For large U_t, at the $z = 0$ boundary the direction of propagation with respect to the surface, given by the angle $\psi = \tan^{-1}\frac{\xi(z)}{\sin\theta_i}$, is approximately parallel to the boundary as in the case of a surface wave; but as z increases away from the boundary this direction of propagation changes so as to asymptotically approach that of the plane wave mode with $U_t = \frac{2}{3}\Delta$. This behavior of the LITW mode is consistent with its interpretation as a frustrated evanescent wave mode. As the incident intensity increases in the case of negative α and the critical angle approaches the angle of incidence θ_i from above, the field in the nonlinear medium tends to become evanescent; but the resulting fall-off of the field away from the $z = 0$ boundary prevents the TIR condition from being maintained for larger values of z, and the field reverts to that of a transmission mode.

By computing U_r and U_i from (32) and (34) with U_t varied as a parameter, the negative α reflectivity U_r/U_i can be determined as a function of U_i for arbitrary values of θ_i and Δ.[12] Figure 2 shows the resulting graph of R versus U_i (for $\alpha < 0$) with the zero field critical angle $\theta_c{}^o$ chosen to be 88° (as in Figure (1)) and θ_i set equal to 87.5°. The LITW mode, which replaces the TIR mode for $\alpha < 0$ and $\Delta > 0$, corresponds to the right hand branch of the curve. In this case the reflectivity is everywhere single-valued and there is no bistability.

Kaplan[6] has shown that bistability can be obtained in the case of negative α only if the parameter Δ is increased sufficiently such that, for an angle of incidence θ_i, $\Delta \geq 3\cos^2\theta_i$. Figure 3 shows a graph of R versus U_i for a case where $\Delta = 6\cos^2\theta_i$, with θ_i taken to be 88°.[13] Here R becomes a multivalued function of U_i in the vicinity of the U_i value corresponding to the transition between the PW mode and the LITW mode. The amplified insert shows the nature of the bistability which occurs in this case. We point out that in this region where bistability of the LITW mode exists with respect to the PW mode, the LITW mode itself also exhibits a multivaluedness of U_t with respect to U_i. This case has been discussed by Kaplan.[6] In comparison to the case of a positive nonlinearity, the discontinuous change in the reflectivity in this negative nonlinearity case is seen to be quite small. Moreover, by inspection of the scale of values of U_i in Figures 1 and 3, the incident intensity required to reach the bistable region in Figure 3 can be seen to be a factor of five times higher than that required to reach the bistable transition in Figure 1. Even with the availabilty of the large negative nonlinear susceptibilities of semiconductors such as InSb in the infrared,[14] the amount of bistable switching action would appear to be insignificant.

III. Conclusions

In summary we have presented here a simplified plane-wave analysis of the reflectivity as a function of incident intensity at the boundary between a linear medium and a Kerr-type nonlinear medium. Our method of calculating the nonlinear reflectivity should provide a simple procedure for predicting bistable behavior in future devices. Consistent with the predictions of Kaplan[2-6] for a single interface, the analysis shows in the case of a positive nonlinearity that there exists significant bistability associated with a discontinuous transition between the TIR and partial transmission modes of reflection. In the case of a negative nonlinearity, the present analysis shows the bistability to be much less significant. Because the change in reflectivity at the bistable transition is in this latter case quite small and the incident intensity required to reach it is larger than in the former case, this case would appear to be mainly of academic interest.

Acknowledgement

We wish to thank Dr. Kaplan for a preprint of Reference 6.

References

1. P. W. Smith, "Introduction to Special Issue on Optical Bistability," IEEE J. Quantum Electron. QE-17, 300 (1981). This special issue contains several papers on optical bistability.

2. A. E. Kaplan, "Hysteresis Reflection and Refraction by a Nonlinear Boundary—A New Class of Effects in Nonlinear Optics," JETP Lett. 24, 114 (1976).

3. A. E. Kaplan, "Theory of Hysteresis Reflection and Refraction of Light by a Boundary of a Nonlinear Medium," Sov. Phys. JETP 45, 896 (1977).

4. A. E. Kaplan, "Criterion of Existence of Longitudinally Inhomogeneous Traveling Waves in Nonlinear Electrodynamics," Sov. J. Quantum Electron. 8, 95 (1978).

5. A. E. Kaplan, "Longitudinally Inhomogeneous Running Waves and Their Role in the Nonlinear Reflection and Refraction of Light," Radio Phys. Quantum Electron. 22, 229 (1979).

6. A. E. Kaplan, "Conditions of Excitation of New Waves (LITW) at Nonlinear Interface and Diagram of Wave States of the System," IEEE J. Quantum Electron. QE-17, 336 (1981).

7. P. W. Smith, J. P. Hermann, W. J. Tomlinson, and P. J. Maloney, "Optical Bistability at a Nonlinear Interface," Appl. Phys. Lett. 35, 846 (1979).

8. P. W. Smith, J. P. Hermann, W. J. Tomlinson, and P. J. Maloney "Experimental Studies of a Nonlinear Interface," IEEE J. Quantum Electron. QE-17, 340 (1981).

9. G. M. Wysin, H. J. Simon, and R. T. Deck, "Optical Bistability with Surface Plasmons," Optics Lett. 6, 30 (1981).

10. Higher harmonic fields are neglected and the time dependence of all fields is assumed to be $e^{-i\omega t}$ and is suppressed.

11. The approach and the notation in this section closely follow that of Kaplan, Ref. 6.

12. The plot of R versus U_i is easily generated by computation of U_i and U_r versus U_t on a pocket calculator.

13. In this case $\sin^2\theta_c^0$ exceeds one, which means that the index of refraction of the linear medium is less than the zero field index of refraction of the nonlinear medium.

14. D. A. B. Miller, S. D. Smith, and C. T. Seaton, "Optical Bistability and Multistability in the Semiconductor in Sb," in Proc. Int. Conf. Opt. Bistability, C. M. Bowden, M. Ciftan, and H. R. Robl, eds., New York, Plenum, 1980.

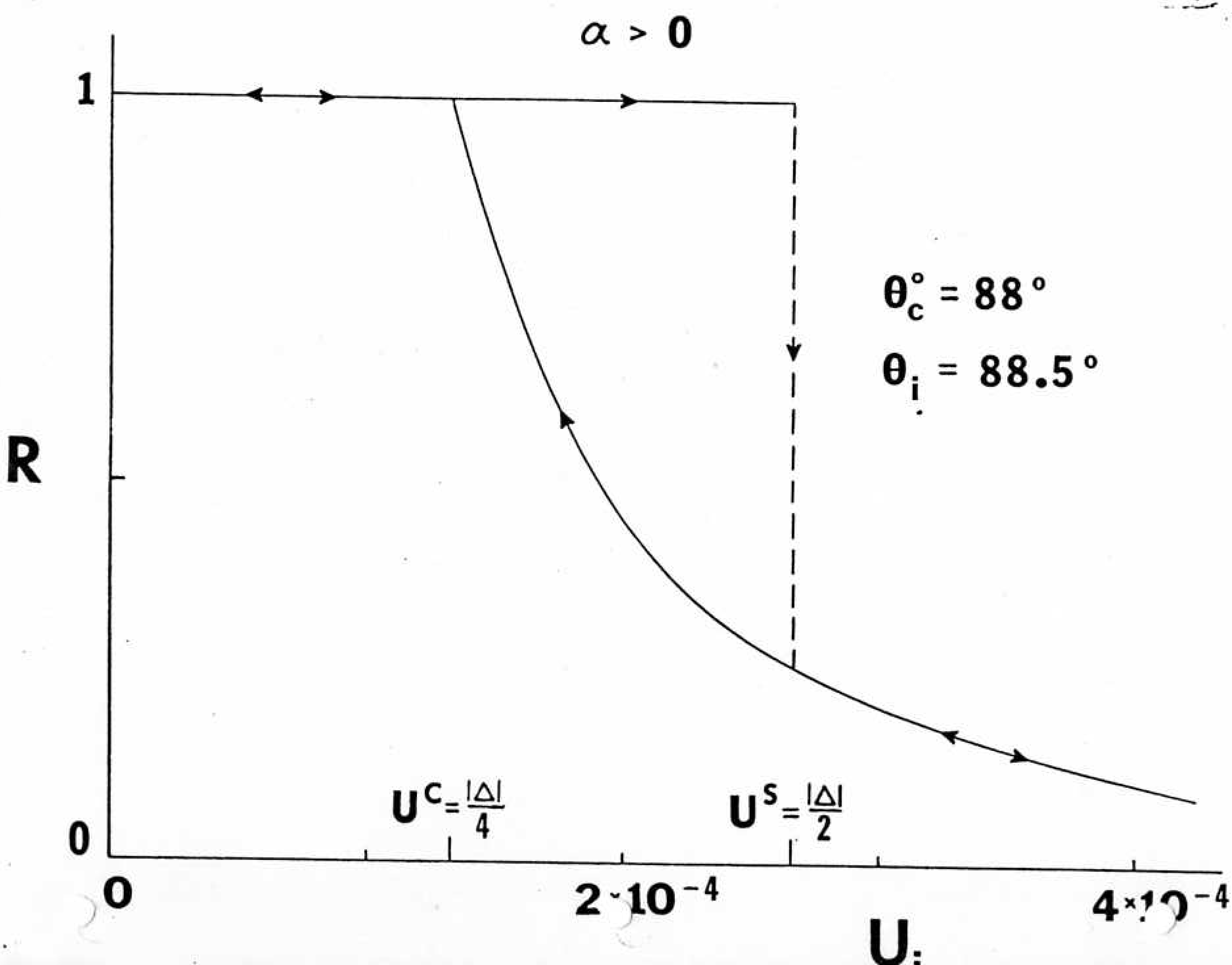
Figure Captions

Fig. 1: Graph of the reflectivity R versus the dimensionless incident intensity U_i for the case of positive α and $\Delta \equiv \sin^2\theta_c^o - \sin^2\theta_i \cong -0.28 \cos^2\theta_i$. The arrows indicate the direction in which the operating point moves when U_i is increased from zero to a value beyond U^s and is then returned to zero.

Fig. 2: Graph of the reflectivity R versus the dimensionless incident intensity U_i for the case of negative α and $\Delta \equiv 0.36 \cos^2\theta_i$. The transition point between the PW and LITW modes is indicated.

Fig. 3: Graph of reflectivity R versus the dimensionless incident intensity U_i for the case of negative α and $\Delta = 6 \cos^2\theta_i$. The insert shows an amplified view of the region of transition between the PW and LITW modes where bistability is possible.

Fig. 1



$\alpha > 0$
1
R
0
$\theta_c^{\circ} = 88^{\circ}$
$\theta_i = 88.5^{\circ}$
$U^c = \frac{|\Delta|}{4}$
$U^s = \frac{|\Delta|}{2}$
0
$2 \cdot 10^{-4}$
4×10^{-4}
U_i

Fig 2
Deck et. al

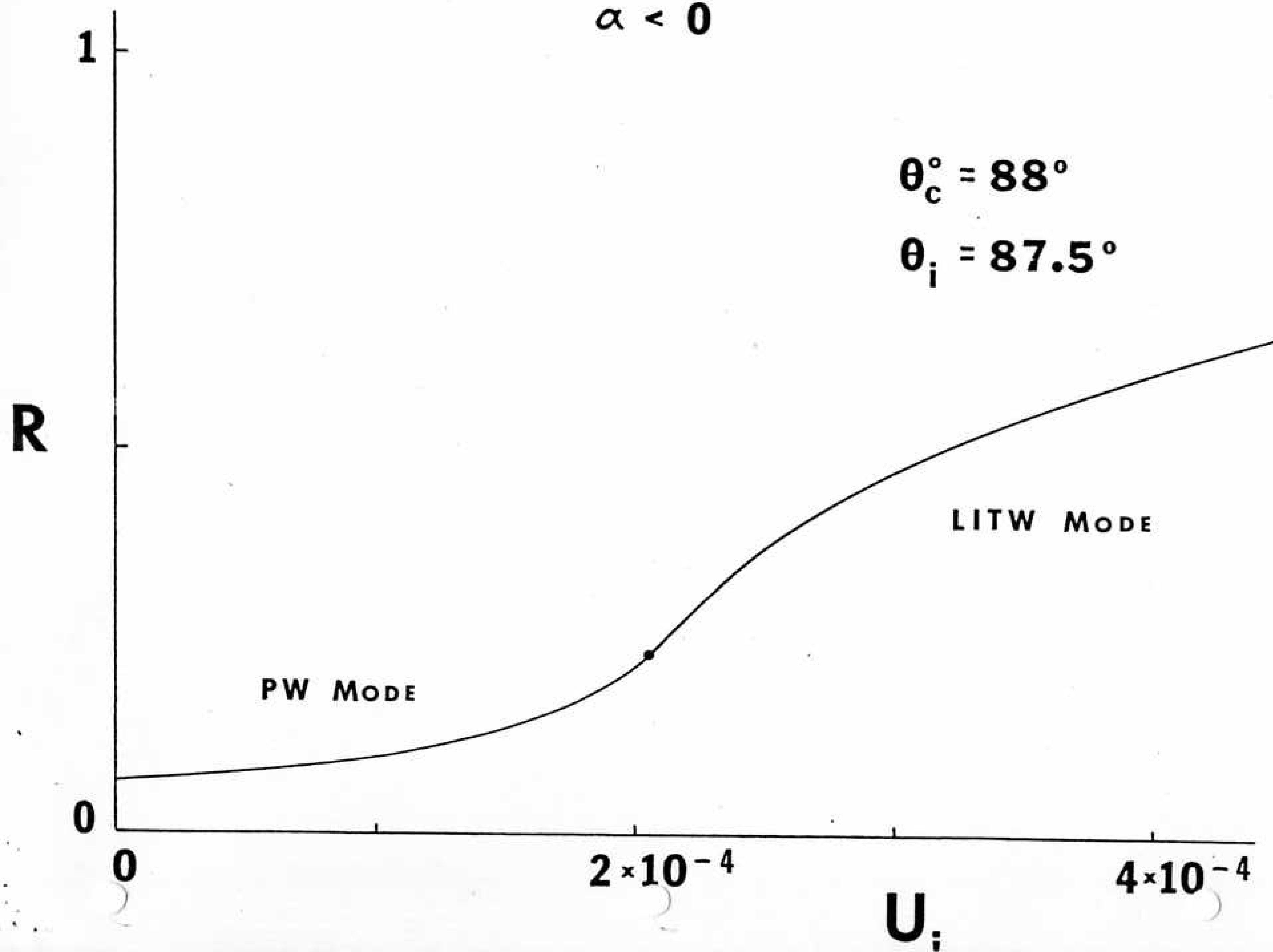


Fig 3
Deck et al