JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 60, NUMBER 10 OCTOBER 1970

Theory of Prism–Film Coupler and Thin-Film Light Guides

P. K. TIEN AND R. ULRICH
Bell Telephone Laboratories, Incorporated, Holmdel, New Jersey 07733
(Received 3 April 1970)

A prism–film coupler has been discussed recently by Tien, Ulrich, and Martin as a device to couple efficiently a laser beam into thin-film dielectric light guides. This coupler also allows an accurate measurement of the spectrum of propagating modes from which the refractive index and the thickness of the film can be determined. We present here a theory of the prism–film coupler. The physical principles involved are illustrated by a method that combines wave and ray optics. We study the modes in the thin-film light guide and their modification by the effect of coupling. We also calculate the field distributions in the prism and the film, the power transfer between the prism and the film, and derive a condition of optimum operation. In one example, 81% of the laser power can be fed into any desired mode of propagation in the film.
INDEX HEADINGS: Films; Lasers; Waveguides; Optics.

The need for integrated optics in optical communication systems has recently been discussed.[1] The possibility is foreseen of guiding laser beams into a thin-film structure and processing optical signals by modulation, switching, frequency conversion, and so on, entirely within the structure. Miniaturization of the system is expected to minimize the effects of ambient conditions and to cut the cost of production. Moreover, because the films are very thin, light waves propagating in thin-film waveguides possess enormous power densities which are attractive for both electro-optic and nonlinear optical devices.

The prism–film coupler recently discussed by Tien, Ulrich, and Martin[2] provides a convenient and efficient method of feeding a laser beam into a thin-film structure. It applies the principle of distributed coupling through evanescent fields to the modes of thin-film waveguides. In this coupler, the prism is placed above the thin-film guide and is separated from it by a small gap of low refractive index. The incident light is totally reflected at the base of the prism, and the waves in the prism and in the film are coupled through their evanescent fields in the gap. The coupler permits excitation of any one of the film modes by proper orientation of the direction of the incident beam.

The coupling phenomenon described is reminiscent of the tunneling of electrons through a barrier; hence the term optical tunneling has been used. In the prism–film coupler shown in Fig. 1, coupling between the prism and the film takes place along the entire width of the incident beam (from a to b) which typically covers hundreds of wavelengths. The coupling is strongest if the components of the wave vectors parallel to the gap are equal for the wave in the prism and the wave in the film. If they differ, the net coupling effect is zero. Let us divide the wave vector of the incident light in the prism into components parallel and normal to the gap. The parallel component is $kn_3 \sin\theta_3$, where $k=\omega/c$, ω is the laser angular frequency, c is the velocity of light in vacuum, n_3 is the refractive index of the prism, and θ_3 is the incident angle. As is discussed later, a thin-film waveguide can support a number of modes of propagation. Each mode propagates along the x direction with a different phase velocity, v_p. When the direction of the incident beam is such that $kn_3 \sin\theta_3$ is equal to the propagation constant, $(\beta=)\ \omega/v_p$, of one of the film modes, coupling becomes effective and optical energy can be transferred from the prism to the film and back from the film to the prism. The direction of the incident beam for which the above condition holds is called a synchronous direction. In one ZnO film, about 3 μm thick, we have excited more than 30 different TM and TE modes, by simply varying θ_3 and the polarization.[2] This possibility of exciting selectively any single mode of the film is important in many applications. For second-harmonic generation, for example, it is possible to accomplish phase matching in a thin-film guide by carrying the fundamental in the $m=0$ mode and the harmonic in the $m=2$ mode without using the birefringence of the crystal. As another example, thin-film lenses and prisms, which are mentioned briefly later, are constructed by depositing guides of varying thicknesses. They function properly only when waves are propagating in them as a single mode.

This paper develops the theory of the prism–film coupler. Previous work[2] has outlined some of the important results of the theory, which was, however, not derived nor thoroughly discussed there. During the preparation of this paper we found it necessary to

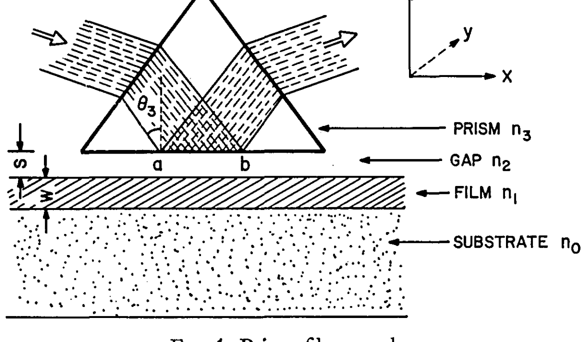

FIG. 1. Prism–film coupler.

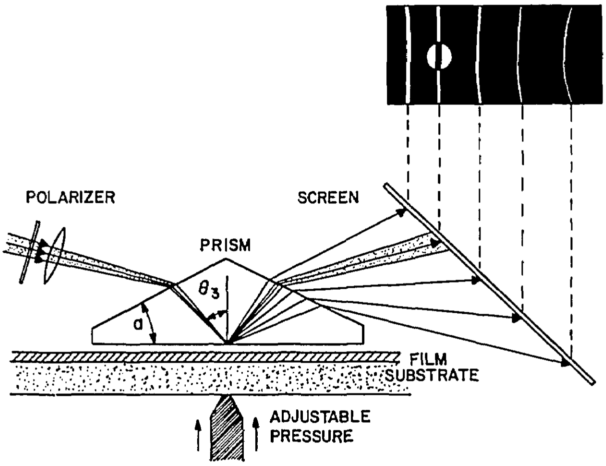

FIG. 2. Experimental arrangement for observation of the mode spectrum of a thin-film waveguide.

rewrite the theory of thin-film waveguides in order to illustrate some of their interesting properties. A theory of this sort is usually plagued by too many parameters. We have four coupled media: a prism, a gap, a film, and a substrate. If we solved Maxwell's equations in a straightforward manner, much of the physics would be lost in the mathematics. In order to illustrate fully the physical principles involved in this problem, we use a theory that combines wave and ray optics.

The thin-film (or dielectric planar) waveguide considered here is a dielectric film sandwiched between two media of refractive indices lower than that of the film. Note in Fig. 1 that when a wave propagates inside the film, one dimension of the beam cross section is guided by the thickness of the film, but in its other dimension, y, the wave can propagate freely. In this paper we give a simplified analysis of the coupler by assuming an incident light beam of infinite extent in y.

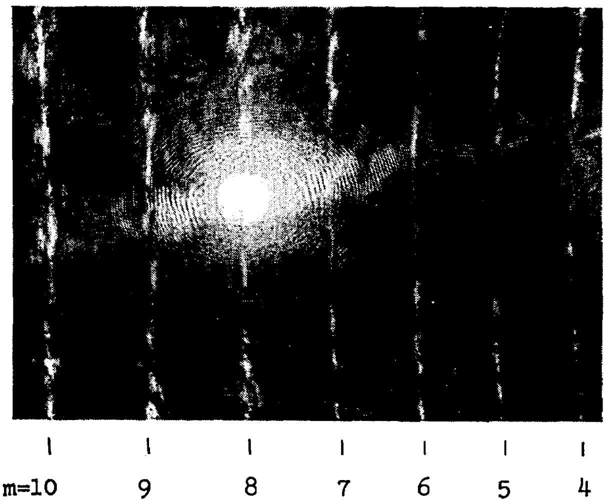

FIG. 3. Photograph of the m lines.

I. EXPERIMENTAL OBSERVATION

In the past year, several experiments have been carried out using the prism–film coupler. We summarize these results here to facilitate our later discussion because a part of the theory explains experimental facts.

In one experiment,[2] a symmetrical prism, as shown in Fig. 2, is used as coupler. The incident beam is focused on the prism base in a synchronous direction, and optical energy is fed into one of the waveguide modes of a ZnO film. Because of film inhomogeneities, the optical energy in the excited mode is rapidly scattered into other modes of the film and is then coupled back to the outside medium by the same prism. We thus see on the screen a series of bright lines with a bright spot on one of these lines. The bright spot is the beam totally reflected from the base of the prism. Inside this spot there is a dark line which represents the energy lost from the excited mode by scattering. We call bright lines m lines. Each line represents a mode of a different order m. Figure 3 is a photograph of these m lines. With increasing coupling (decreasing width of the gap between the prism and the film) the m lines first become brighter. The lines then broaden and their positions shift. Finally, they fade into the background. From the positions and the widths of the m lines it is possible to determine the mode spectra, the refractive index and the thickness of the film, the coupling strength, and to some extent the scattering and absorption properties of the film.

In the arrangement shown in Fig. 1, the amount of energy that is transferred from the incident beam in the prism into the film in the region between $x=a$ and $x=b$ is far greater than that transferred from the film to the prism in that region. The power flowing inside the film increases in the x direction. Because the incident beam stops at $x=b$ and the energy transfer from the film to the prism continues in $x>b$, the energy inside the film decreases rapidly to zero there. In order to retain this energy in the film it is therefore necessary to decouple the prism from the film beyond the point $x=b$. This can be accomplished by using a right-angle prism, as shown in Fig. 4. Here the right-angle corner of the prism is placed at the position corresponding to the point $x=b$ in Fig. 1. The waves in the film then continue to propagate inside the film as a guided wave. In a real film, the path of this wave can be observed because of reflections of the guided light from dust particles on the film surface and scattering caused by inhomogeneities of the film. This is shown in Fig. 5, where a narrow beam of light is guided in a Ta_2O_5 film, deposited on a microscope glass slide. The length of the observed streak is about 70 mm.

Figure 6 shows a thin-film prism. It is a ZnS film, the triangular portion of which is thicker than the background film. Figure 7 shows a photograph of the light scattered from a guided beam that is deflected by this prism. The step between the thicker portion and the background film is tapered and extends over many wavelengths. Because of this gradual transition, the mode order m is preserved for a light wave passing through the step. Here the wave is guided by the film and it propagates as $\exp(-i\omega t+i\beta x)$. We will show later that for a same mode order m, β is different in the thicker portion than in the background film. Consequently, the light wave is refracted at the step because it propagates at different speeds in the two portions of the film. The refraction follows the usual Snell law provided that the β's are used in the formula instead of the refractive indices.

The experiments described require films of low absorption and scattering losses. In the visible spectrum, at 6328 Å of the He–Ne laser, a streak of light wave approximately 25 mm long has been observed in our better ZnS films. In some Ta_2O_5 films, prepared by Hensler and Cuthbert,[3] longer streaks have been obtained at 4880, 5145, and 6328 Å. Goell and Standley[4] have reported a loss of less than 20%/cm in their sputtered-glass films at 6328 Å. Some organic films prepared by Smolinsky and Vasile[3] also have very low losses. In the infrared region, streaks 5–10 cm long have been observed in ZnS films at 1.06 μm.

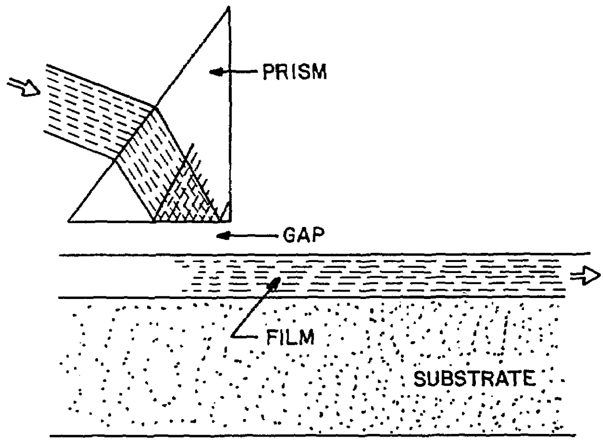

FIG. 4. Experimental arrangement for feeding a laser beam into a thin film.

II. NOTATION AND THE WAVE EQUATION

Throughout our equations we use the subscript 3 for quantities that belong to the prism, and the subscripts 2, 1, and 0 for quantities that belong to the gap, the film, and the substrate. Quantities at the interfaces are denoted by double subscripts—the subscript 12 denotes the interface between the film and the gap. All interfaces are parallel to the x–y plane. We follow Born and Wolf[5] and choose gaussian units for Maxwell's equations, so that formulas derived in their book can be quoted directly. For the two-dimensional analysis used here, $\partial/\partial y=0$, the wave equation for electric fields is

$$\partial^2E/\partial x^2+\partial^2E/\partial z^2=-(kn_j)^2E,\quad j=0,\,1,\,2,\text{ or }3,\qquad(1)$$

where n_j is the refractive index of medium j. For TE waves, we have field components E_y, H_x, and H_z only, and for TM waves, H_y, E_x, and E_z only. They correspond, respectively, to the waves of polarization normal and parallel to the plane of incidence. The two curl equations for the TE and TM waves are

$$H_x=(i/k)(\partial/\partial z)E_y \quad\text{and}\quad H_z=-(i/k)(\partial/\partial x)E_y \qquad (2)$$

$$E_x=-(i/kn_j^2)(\partial/\partial z)H_y$$

and

$$E_z=(i/kn_j^2)(\partial/\partial x)H_y. \qquad (3)$$

A time dependence $\exp(-i\omega t)$ is used, and $i=(-1)^{\frac{1}{2}}$. As a consequence of Eqs. (2) and (3) it is sufficient to consider only the E_y for TE waves and H_y for TM waves. Throughout this paper we denote the complex amplitudes of the incident and reflected beams by A_3 and B_3. Similarly, we denote the waves in the film by A_1 and B_1, as shown in Fig. 8. We use the convention that all the A_j waves propagate toward the lower right, and the B_j waves propagate toward the upper right. If the waves are coupled they all have the same phase constant β along the x axis.

III. THEORY OF THIN-FILM WAVEGUIDES

This section explains how, in a thin-film waveguide, plane waves interfere constructively in forming a propagating mode. A condition is then derived for the waves to add in phase, and this condition is shown to be equivalent to the equation of the modes. The method, which is simple, provides physical insight into the problem.

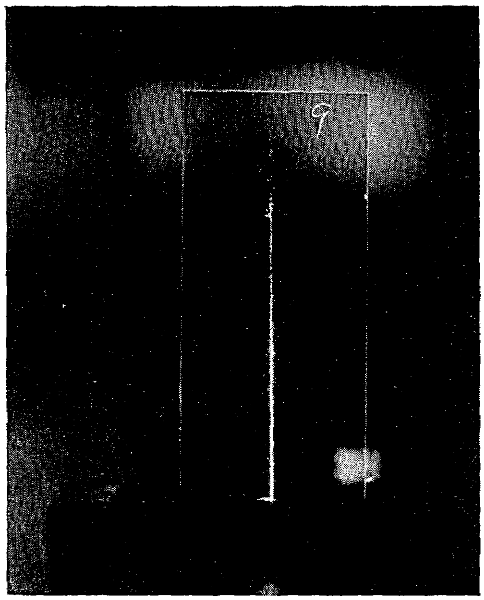

FIG. 5. Photograph of a streak of guided light excited in a semiconductor film.

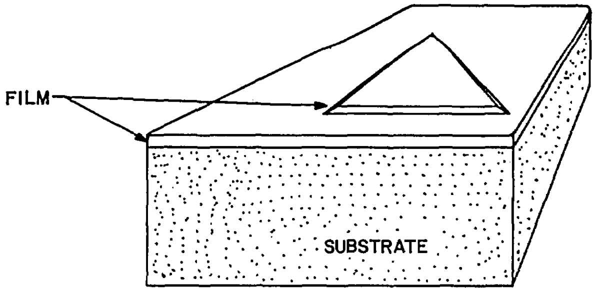

FIG. 6. Thin-film prism constructed by depositing a triangular portion thicker than the background film.

The thin film has a refractive index n_1 and a thickness W. It is sandwiched between two semi-infinite media of refractive indices n_0 and n_2. Here we assume $n_1>n_0>n_2$. We start by considering a plane wave, A_1, which propagates in the film toward its lower boundary, $z=0$, with an incident angle θ_1 [at (a) in Fig. 8] on the interface 0, 1. If θ_1 is larger than the critical angle between n_1 and n_0, the A_1 wave is totally reflected into the B_1 wave. Similarly, at (b) in Fig. 8, the B_1 wave is totally reflected into the A_1' wave at the upper film boundary. It is obvious that the A_1 and A_1' waves have the common propagation factor

$$\exp(-i\omega t-ib_1z+i\beta x),$$

where

$$b_1=kn_1\cos\theta_1 \quad \text{and} \quad \beta=kn_1\sin\theta_1.$$

Similarly, the B_1 wave has the propagation factor

$$\exp(i\omega t+ib_1z+i\beta x).$$

In the following, the common factor $\exp(-i\omega t+i\beta x)$ will be omitted in all expressions. All A_j waves therefore have the form $\exp(-ib_jz)$, and the B_j waves have the form $\exp(+ib_jz)$, where $j=1$ denotes the film and $j=3$ denotes the prism. For example, the A_1, A_1', and B_1 in the film have the forms $A_1\exp(-ib_1z)$, $A_1'\exp(-ib_1z)$, and $B_1\exp(ib_1z)$. The letters C_j and D_j denote the fields in the substrate or in the gap where $j=0$ or 2.

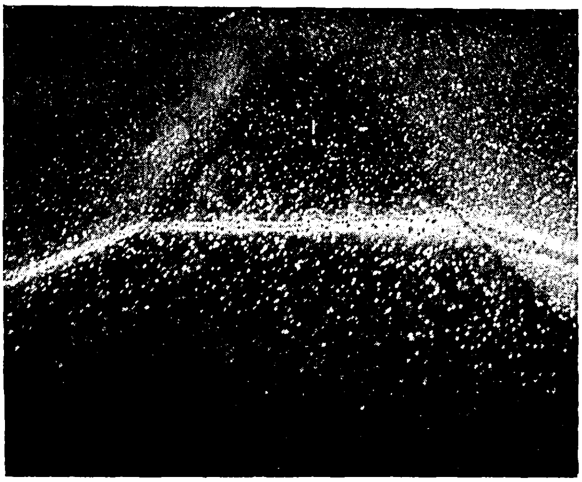

FIG. 7. Photograph of a light beam deflected by the thin-film prism.

All of the fields must satisfy the curl Eqs. (2) or (3). Taking TE waves, as an example, we have for the A_1 (or A_1') waves

$$E_y=A_1(\text{or } A_1')e^{-ib_1z};\quad H_x=n_1\cos\theta_1A_1(\text{or } A_1')e^{-ib_1z} \quad (4)$$

and for the B_1 waves

$$E_y=B_1e^{ib_1z};\quad H_x=n_1\cos\theta_1B_1e^{ib_1z}, \quad (5)$$

where $0<z<W$. Because of the total reflections discussed earlier, the fields in media n_0 and n_2 are exponentially decreasing functions. In the substrate, n_0, we have

$$E_y=C_0e^{p_0z};\quad H_x=\frac{ip_0}{k}C_0e^{p_0z};\quad z<0, \quad (6)$$

and in the gap, n_2,

$$E_y=D_2e^{-p_2(z-W)};\quad H_x=\frac{ip_2}{k}D_2e^{-p_2(z-W)};\quad z>w. \quad (7)$$

Substituting Eqs. (4)–(7) into the wave equation (1), one at a time, we obtain

$$\begin{aligned} \beta&=kn_1\sin\theta_1;\quad b_1=kn_1\cos\theta_1,\\ b_1^2&=(kn_1)^2-\beta^2,\\ p_0^2&=\beta^2-(kn_0)^2,\\ p_2^2&=\beta^2-(kn_2)^2. \end{aligned} \quad (8)$$

The quantities β, b_1, p_0, and p_2 are real and positive. Otherwise, the waves A_1 and B_1 are no longer totally reflected at the film boundaries and they form radiation modes[6] which will not be discussed here.

To match the boundary conditions at $z=0$ in Fig. 8, we add the E field (and also H field) of the A_1 wave in Eq. (4) to that of the B_1 waves in Eq. (5) and equate the sum to the E field (H field) of the evanescent wave in Eq. (6). We find

$$B_1/A_1=e^{-i2\Phi_{10}}. \quad (9)$$

Similarly, by matching boundary conditions at $z=-W$, we have

$$A_1'/B_1=e^{-i2\Phi_{12}}, \quad (10)$$

where

$$\tan\Phi_{10}=p_0/b_1,\quad \tan\Phi_{12}=p_2/b_1 \quad (11)$$

for the TE waves. Similarly, we can show, for the TM waves,

$$\tan\Phi_{10}=(n_1/n_0)^2p_0/b_1,\quad \tan\Phi_{12}=(n_1/n_2)^2p_2/b_1. \quad (12)$$

We choose those solutions Φ_{10} and Φ_{12} of Eqs. (11) and (12) for which $0\leqq\Phi\leqq\pi/2$. The A_1 wave in Eq. (9) [or the B_1 wave in Eq. (10)] suffers a phase change of

$-2\Phi_{10}$ (or $-2\Phi_{12}$) during the total reflection, which has an important effect upon the field distribution in the waveguide. If, for example, $\beta \rightarrow kn_1$, then $2\Phi_{12} \rightarrow \pi$ in Eq. (11). The incident and reflected waves differ by nearly a phase of π, and so they almost cancel at the boundary $z=W$. In accordance with this, p_2 is large and the fields penetrate only little into the medium, n_2.

Now we can combine the waves at (a) in Fig. 8 with those at (b) in Fig. 8 as shown at (c) in Fig. 8, where the A_1' wave follows the A_1 wave after one zigzag path. Because the reflections at both film surfaces are total, the amplitudes A_1' and A_1 can differ only by a phase Δ. After subsequent zigzags, the wave has phase differences 2Δ, 3Δ, 4Δ, $\cdots$ relative to A_1. In general, the superposition of such a set of plane waves is zero, except when $\Delta = 2m\pi$ with integer m. In that case the beams A_1, A_1' and all further reflections of this beam interfere constructively.

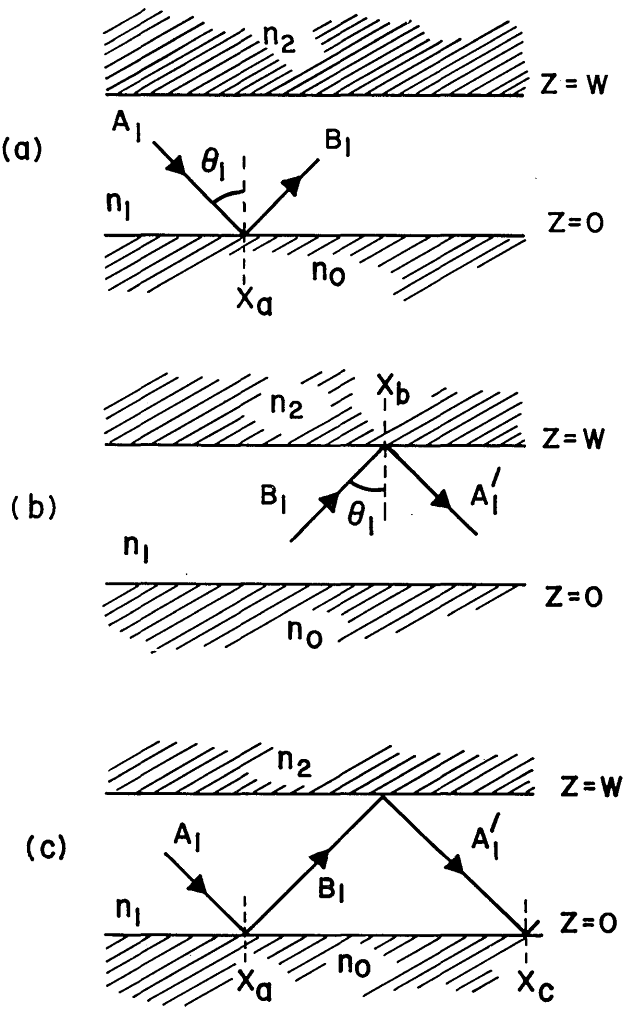

FIG. 8. (a) Total reflection from the lower film boundary, (b) total reflection from the upper film boundary, and (c) optics in a thin-film waveguide.

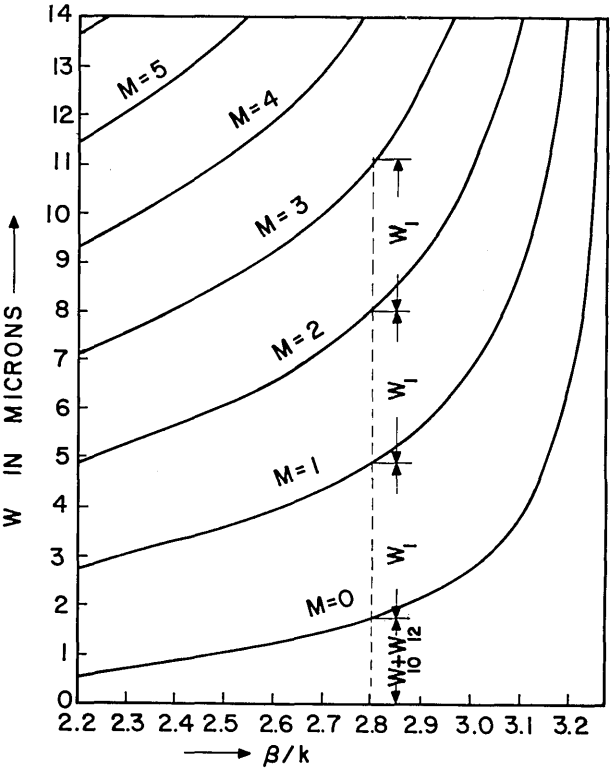

FIG. 9. W vs β/k for $n_0=2.190$, $n_1=3.275$, $n_2=1.00$, and laser wavelength $10.6\,\mu$m.

We can find the phase difference Δ directly. The phase of the A_1 wave at $x=x_c$ and $z=0$ is

$$-\omega t+\beta x_c. \tag{13}$$

The phase of the A_1' wave at the same point is the phase of the A_1 wave at $x=x_a$ and $z=0$ plus that of a zigzag. It is

$$-\omega t+\beta x_a+\beta(x_c-x_a)+2b_1W-2\Phi_{10}-2\Phi_{12}. \tag{14}$$

The difference of expressions (14) and (13) is $\Delta = 2m\pi$. Therefore

$$2b_1W-2\Phi_{10}-2\Phi_{12}=2m\pi. \tag{15}$$

This is the equation of the modes. Since b_1W is positive and both Φ_{10} and $\Phi_{12} \leqq \pi/2$, m cannot be negative. The integer m may then be 0, 1, 2, 3 up to a certain finite value, depending on W. This m specifies the order of the mode. Equation (15) is the same for both the TE and the TM waves, but the Φ_{ij} differ.

IV. PROPERTIES OF THE THIN-FILM WAVEGUIDES

Let $n_0 > n_2$ in the film waveguide shown in Fig. 8. Since β, b_1, p_0, and p_2 are all positive in Eq. (8), possible values of β range from kn_0 to kn_1. In the upper limit ($\beta \rightarrow kn_1$), we have $b_1 \rightarrow 0$. Thus, from Eq. (15), $W \rightarrow \infty$. This is understandable because in this case

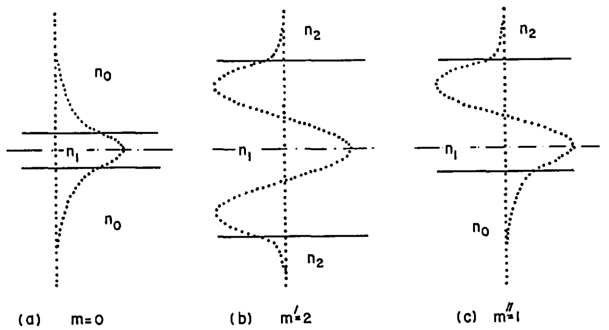

FIG. 10. (a) Symmetric waveguide having a phase constant β for a mode of order $m=0$, (b) symmetric waveguide having a phase constant β for a mode of order $m'=2$, and (c) asymmetric waveguide constructed by combining the lower half of (a) with the upper half of (b). The dotted curves are the field distributions, E_y for the TE waves and H_y for the TM waves.

the waves propagate as plane waves parallel to the x axis and for that, the boundaries of the film must be at $z=\pm\infty$. At the lower limit ($\beta\to kn_0$), we have

$$b_1\to k(n_1^2-n_0^2)^{\frac{1}{2}},\quad p_0\to 0,\quad \Phi_{10}\to 0,$$
$$p_2\to k(n_0^2-n_2^2)^{\frac{1}{2}}.$$

Therefore, the thickness of the film calculated from Eq. (15) is

$$W_{\min}=\frac{1}{k}\left[m\pi+\tan^{-1}\left(\frac{n_0^2-n_2^2}{n_1^2-n_0^2}\right)^{\frac{1}{2}}\right]\Big/(n_1^2-n_0^2)^{\frac{1}{2}} \quad (16)$$

for the TE waves. This is the minimum thickness required for a waveguide to support a mode of the order m. For a symmetric waveguide ($n_0=n_2$) and $m=0$, $W_{\min}\to 0$, the film can be infinitesimally thin.

Equation (15) cannot be solved explicitly in β because Φ_{10} and Φ_{12} involve transcendental functions. Conversely, however, if we assign a value to β then the quantities b_1, p_0, and p_2 can be calculated from Eqs. (8) and (11). In addition, if m is given, W can be calculated from Eq. (15). It is the thickness of the film required for a mode of the order m to propagate with a given phase constant β. As an example, in Fig. 9, the thickness W (in microns) is plotted vs β/k for a GaAs film deposited on an Irtran-II substrate. For a laser wavelength of 10.6 μm $n_0=2.190$, $n_1=3.275$, $n_2=1.000$. Note that W increases both with β/k and m.

Equation (15) can be rewritten

$$W=W_{10}+W_{12}+mW_1, \quad (17)$$

where

$$W_{10}=\Phi_{10}/b_1;\quad W_{12}=\Phi_{12}/b_1;\quad W_1=\pi/b_1.$$

Equation (17) indicates that we can construct an asymmetric waveguide ($n_0\neq n_2$) of the desired propagation characteristics from two symmetric ($n_0=n_2$ and $n_0'=n_2'$) waveguides that have films of identical refractive indices $n_1=n_1'$. For example (b) in Fig. 10 shows a symmetric waveguide with a phase constant β for the order $m'=2$. By combining the lower half of (a) in Fig. 10 with the upper half of (b) in Fig. 10, we obtain an asymmetric waveguide that has the same phase constant β for the mode of the order, $m''=(m+m')/2=1$. The same process can be performed, of course, with any other combination, m and m', both even. Also notice in Eq. (17) that, for a given β, the thickness of the film (for the mode order m) is simply that for $m=0$ plus mW_1. This is illustrated both in Figs. 9 and 11. The dotted curves in Fig. 10 are the field distribution across the waveguides (E_y for the TE waves and H_y for the TM waves). At the film boundaries, the field amplitudes are $2A_1\cos\Phi_{10}$ and $2A_1\cos\Phi_{12}$, respectively.

In Figs. 12 and 13, we have plotted, for the TE waves only, kn_1W_{10} (or kn_1W_{12}) and kn_1W_1 vs β/kn_1, using n_0/n_1 (or n_2/n_1) as the parameter. It is interesting that the kn_1W_1 curve is independent of the values of n_0/n_1 or n_2/n_1. To determine the thickness of the film for a waveguide of given β, we simply find the proper values of W_{10}, W_{12}, and W_1 from Figs. 12 and 13 and add them according to Eq. (17). Figures 12 and 13 are useful for the design of thin-film lenses and prisms.

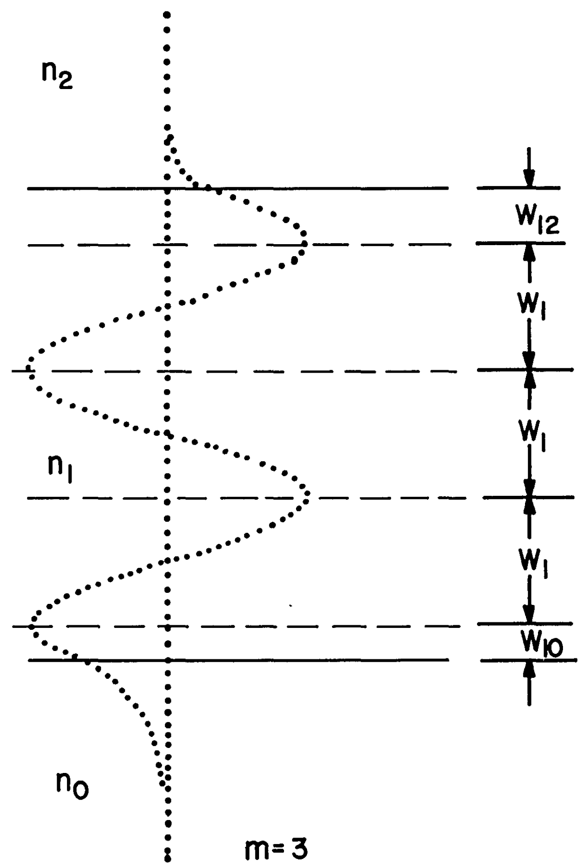

FIG. 11. The thickness of a thin-film waveguide may be considered as the sum of W_{10}, W_{12}, and mW_1.

Finally, we calculate, for TE waves only, the power carried in a waveguide, by integrating x component of the Poynting vector

$$(c/8\pi)\ \mathrm{Re}(E_y H_z^*)$$

of the total field (A_j and B_j waves) from $z=-\infty$ to $z=+\infty$. For a waveguide of unit width in y, this power is

$$P=(c/4\pi)A_1A_1^*n_1 \sin\theta_1(W+1/p_0+1/p_2). \quad (18)$$

Equation (18) has a simple interpretation. The quantity $(c/4\pi)A_1A_1^*n_1 \sin\theta_1$ is the Poynting vector along the x axis for the superposition of the A_1 and B_1 waves (at a in Fig. 8). The factor $(W+1/p_0+1/p_2)$ is then an equivalent thickness of the waveguide, W_{eq}, within which the energy of the waves is confined. It is larger than the actual thickness W of the film because the fields extend beyond its boundaries according to $\exp(p_0z)$ for $z<0$ and $\exp[-p_2(z-W)]$ for $z>W$. The power density in a film waveguide is inversely proportional to W_{eq} (not W). Hence, even though in a symmetric waveguide the film can be infinitesimally thin, the power density cannot approach infinity. For $m=0$, W_{eq} is approximately $\lambda/2n_1 \cos\theta_1$, where λ is the wavelength in a vacuum. The simple form of P in Eq. (18) does not apply to the TM waves.

V. THEORY OF THE PRISM–FILM COUPLER

To develop a theory for the prism–film coupler, we use essentially the same method as in Sec. III. The fields in the coupler are divided into groups of waves, and each group is studied in a self-consistent manner. Finally, all of the fields are assembled according to ray optics.

The problem treated here is illustrated in Fig. 14. A laser beam, A_3, enters into a prism of refractive index n_3 and propagates downward with an incident angle θ_3. Within the beam cross section, the A_3 wave is assumed to have a constant amplitude and hence it is uniformly distributed from $x=0$ to $x=l$ at the base of the prism. There a part of it is reflected as the B_3 wave and the remainder is coupled into the film as an A_1 wave. Because of coupling, we expect the amplitudes of the waves, except that of the A_3 wave, to vary with x. They are slowly varying functions because the rapidly changing phases of the waves are taken into account by the factor $\exp(-i\omega t+i\beta x\pm ib_jz)$. The purpose of this theory is to calculate their variation with x.

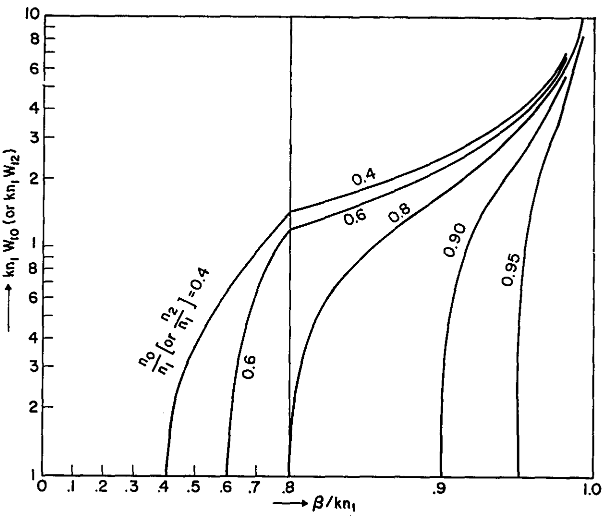

FIG. 12. kn_1W_{10} (or kn_1W_{12}) vs β/kn_1. The parameters on the curves are the values of n_0/n_1 (or n_2/n_1).

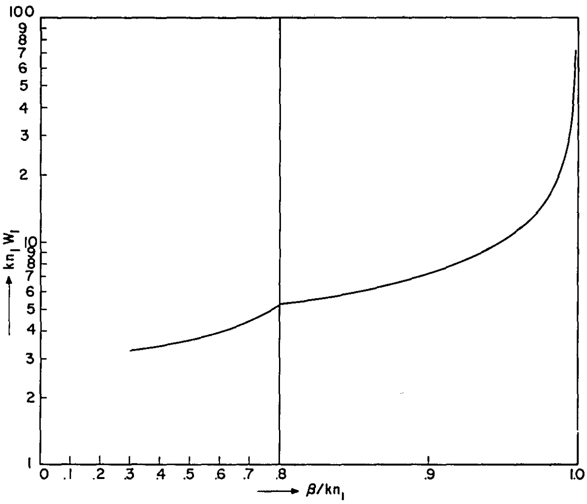

FIG. 13. kn_1W_1 β/kn_1. Parameters as in Fig. 12.

The prism is separated from the film by a gap of refractive index n_2 and of spacing S. The gap is assumed to be uniform, and all the interfaces are parallel to the x–y plane. For simplicity, only the TE waves will be studied here, although, the final results apply equally to TM waves. Solutions for the A_1 and B_1 waves in the film have been expressed by Eqs. (4) and (5). The A_3 and B_3 waves in the prism are similarly expressed by replacing the subscript 1 in Eq. (4) and (5) by 3. The evanescent fields in the substrate remain in the form of Eq. (7), but those in the gap now have two terms

$$E_y=C_2e^{p_2(z-W)}+D_2e^{-p_2(z-W)},$$
$$H_x=(ip_2/k)(-C_2e^{p_2(z-W)}+D_2e^{-p_2(z-W)}), \quad (19)$$

where $W\leq z\leq W+S$. Because the fields must satisfy the wave equation (1), we have in addition to Eq. (8),

$$\beta=kn_3 \sin\theta_3, \quad b_3=kn_3 \cos\theta_3,$$
$$b_3^2=(kn_3)^2-\beta^2. \quad (20)$$

For later convenience, we define

$$\tan\Phi_{32}=p_2/b_3; \quad \Phi_{32}\leq\pi/2 \quad (21)$$

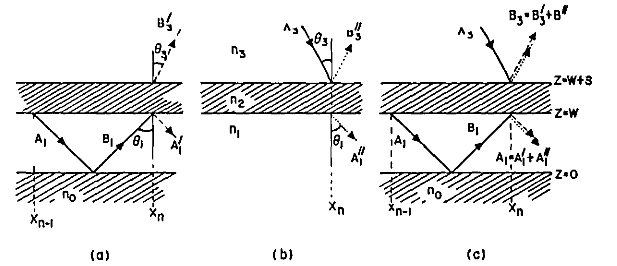

FIG. 14. (a) Thin film n_1 coupled to a semi-infinite medium n_3 through a gap S, (b) two semi-infinite media, n_3 and n_1, coupled through a gap S, and (c) waves in a prism–film coupler.

for the TE waves, and

$$\tan\Phi_{32}=(n_3/n_2)^2p_2/b_3;\quad \Phi_{32}\leq\pi/2 \tag{22}$$

for the TM waves. Following Eq. (9), we would expect $B_3/A_3=\exp(-i2\Phi_{32})$ at $z=W+S$, if the spacing S of the gap were infinitely large.

First, it is interesting to show the effect of the prism on the modes in the waveguide. Let us consider (a) in Fig. 14 where the medium n_2 is no longer semi-infinite, but separates the film from another semi-infinite medium n_3, which represents the prism. Figure 14 should be compared with Fig. 8(c). We notice that, in the presence of the n_3 medium, the B_1 wave is reflected into the A_1' wave, but at the same time it excites a B_3' wave in the prism, which constitutes a loss of energy to the thin-film waveguide. Because of its coupling to the prism, the waveguide has become leaky. The waves in the film thus will decrease with x. We assume that they decrease slowly and that they may be considered to be constant in the vicinity of a point $x=x_n$. In fact, we may consider the A_1, B_1, A_1', and B_3' waves as rays. The ray B_1 can be calculated from A_1 by Eq. (9). The rays B_1, A_1, and B_3' form a self-consistent set in the vicinity of $x=x_n$, and their amplitudes are related to each other by the boundary conditions at $z=W$ and $z=W+S$. To calculate this relation, we express the fields in the gap by Eq. (19). The calculation involves five amplitudes, B_1, A_1', B_3', C_2, and D_2, and four boundary conditions—two for E_y and H_x at $z=W$ and another two at $z=W+S$. It is possible to eliminate C_2 and D_2, and to solve A_1' and B_3' for A_1. We have, after considerable algebra,

$$\begin{gathered}A_1'/B_1=R_1,\\(n_3\cos\theta_3)^{\frac{1}{2}}B_3'/(n_1\cos\theta_1)^{\frac{1}{2}}B_1=T,\end{gathered} \tag{23}$$

$$R_1=r_1e^{-i2\Phi_{02}'}$$
$$=\frac{\tan\Phi_{32}-\tan\Phi_{12}-i(\tanh p_2S)(1+\tan\Phi_{32}\tan\Phi_{12})}{\tan\Phi_{32}+\tan\Phi_{12}-i(\tanh p_2S)(1-\tan\Phi_{32}\tan\Phi_{12})},$$

$$T=1/\cosh p_2S$$
$$\cdot\frac{2(\tan\Phi_{32}\tan\Phi_{12})^{\frac{1}{2}}}{\tan\Phi_{32}+\tan\Phi_{12}-i(\tanh p_2S)(1-\tan\Phi_{32}\tan\Phi_{12})}, \tag{24, 25}$$

where R_1 and T are complex quantities. Except for a constant factor, $R_1=r_{123}$ and $T=t_{321}$ according to Ref. (7). In Eq. (24) the magnitude and the phase of R_1 are denoted by r_1 and $-2\Phi_{12}'$, respectively. Following the argument given in Sec. III, the condition for a propagating mode is again that the phases of the A_1 and A_1' waves evaluated at $x=x_n$ differ by $2m\pi$. We remember from Fig. 8(c) that the phase of the A_1' wave at $x=x_c$ is calculated from the phase of the A_1 wave at $x=x_a$ plus that of a zigzag path. Consequently, we obtain here the condition

$$2b_1W-2\Phi_{10}-2\Phi_{12}'=2m\pi, \tag{26}$$

which is the new equation of the modes for a film waveguide coupled to a prism. The change from Φ_{12} to Φ_{12}' in the mode equation causes a change in θ_1 and a change in the synchronous direction. We have observed this in our experiments.[2] For an infinitely large gap, we find from Eq. (24) that $\Phi_{12}'\rightarrow\Phi_{12}$, and Eq. (26) is reduced to Eq. (15). For wide but finite gap, $\exp(p_2S)\gg1$, Eqs. (23) and (24) may be approximated by

$$|T|^2\cong4e^{-2p_2S}\sin2\Phi_{12}\sin2\Phi_{32} \tag{27}$$

and

$$\Phi_{12}'\cong\Phi_{12}+e^{-2p_2S}\sin2\Phi_{12}\cos2\Phi_{32}. \tag{28}$$

This correction of Φ_{12} for Φ_{12}' is small, of the order of $\exp(-2p_2S)$. The Poynting vector of the B_3' wave in the z direction is $(c/8\pi)B_3'B_c^*n_1\cos\theta_3$ and that of the B_1 wave is $(c/8\pi)B_1B_1^*n_1\cos\theta_1$. The quantity $B_3'B_3^{*\prime}n_3\times\cos\theta_3/B_1B_1^*n_1\cos\theta_1$ is then the fraction of the power transferred from the B_1 wave to the B_3' wave. As given by Eq. (24), this fraction is $|T|^2$, which is on the order of $\exp(-2p_2S)$ in Eq. (27). Similarly, the quantity

$$\frac{A_1'A_1'^*n_1\cos\theta_1}{B_1B_1^*n_1\cos\theta_1}=\frac{A_1'A_1'^*}{B_1B_1^*}=|R_1|^2=r_1^2$$

is the power reflectance at the top surface of the film. For conservation of energy, the sum of the ratio of the power transfer and that of the power reflectance must be equal to unity. We find indeed from Eqs. (23) and (24) that this is true; that is,

$$|R_1|^2+|T|^2=1. \tag{29}$$

Next, we remove the lower boundary of the film and consider Fig. 14, in which two semi-infinite media n_1 and n_3 are coupled through a gap of spacing S. We have a source of excitation, the A_3 wave, which is partially reflected at $z=W+S$ as the B_3'' wave and partially transmitted into the film as the A_1', wave.

After a close examination, we find that the set of the A_3, B_3'', and A_1'' waves in Fig. 14 (b) is analogous to the set of the B_1, A_1', and B_3' waves in Fig. 14(a), except that the media n_1 and n_3 are switched. Because of this, we obtain from Eq. (23) and (24), by interchanging B's with A's, the subscript 1 with the subscript 3, and the prime with the double prime,

$$B_\circ''/A_3=R_3, \tag{30}$$

$$(n_1\cos\theta_1)^{\frac{1}{2}}A_1''/(n_3\cos\theta_3)^{\frac{1}{2}}A_3=T, \tag{31}$$

where T is invariant in this operation and

$$R_3=r_3e^{-i2\Phi_{32}'}=\frac{\tan\Phi_{12}-\tan\Phi_{32}-i(\tanh p_2S)(1+\tan\Phi_{32}\tan\Phi_{12})}{\tan\Phi_{32}+\tan\Phi_{12}+i(\tanh p_2S)(1-\tan\Phi_{32}\tan\Phi_{12})}. \tag{32}$$

Again, because of conservation of energy, we have

$$|R_3|^2+|T|^2=1. \tag{33}$$

In the prism–film coupler, the optical energy is transferred into the film across the interfaces 3, 2 and 2, 1. It is therefore important to consider the z components of the Poynting vectors of the various waves. The equations become simpler if we normalize the wave amplitudes so that

$$\left.\begin{aligned}(A_j)(A_j)^*&=A_jA_j^*n_j\cos\theta_j\\(B_j)(B_j)^*&=B_jB_j^*n_j\cos\theta_j\end{aligned}\right\}\quad j=1,3 \tag{34}$$

for the TE waves, and

$$\left.\begin{aligned}(A_j)(A_j)^*&=A_jA_j^*\cos\theta_j/n_j\\(B_j)(B_j)^*&=B_jB_j^*\cos\theta_j/n_j\end{aligned}\right\}\quad j=1,3 \tag{35}$$

for the TM waves, where (A_j) and (B_j) denote the normalized amplitudes. The amplitudes A_j', B_j', etc. will also be normalized similarly.

Now we can assemble all of the fields in the prism–film coupler by simply superposing (a) and (b) in Fig. 14(c). Let us concentrate on an area near $x=x_n$ in Fig. 14(c). An A_3 wave, which is the incident laser beam in the prism, excites an A_1'' wave in the thin film. Also observed at $x=x_n$ is the A_1' wave, which was the A_1 wave at $x=x_{n-1}$ before progressing through a zigzag path. The resultant A_1 wave at $x=x_n$ is therefore the sum of the A_1' and A_1'' waves. Similarly, the resultant B_3 wave at $x=x_n$ is the sum of the B_3' and B_3'' waves. Using Eqs. (23), (24), (30), and (31) we have

$$(B_3)_n=(B_3')_n+(B_3'')_n=R_3(A_3)_n+T(B_1)_n \tag{36}$$

$$(A_1)_n=(A_1'')_n+(A_1')_n=T(A_3)_n+R_1(B_1)_n, \tag{37}$$

where the subscript n denotes the position $x=x_n$. In addition, the B_1 wave at $x=x_n$ can be calculated from the A_1 wave at $x=x_{n-1}$

$$(B_1)_n=(A_1)_{n-1}\exp(i2b_1W-i2\Phi_{10}). \tag{38}$$

Substituting Eq. (38) into Eq. (37) and using Eq. (24), $R_1=r_1\exp(-i2\Phi_{12}')$, we obtain

$$(A_1)_n=T(A_3)_n+r_1(A_1)_{n-1}\times\exp(i2b_1W-i2\Phi_{10}-i2\Phi_{12}'). \tag{39}$$

Because the incident beam in the prism is assumed to be in a synchronous direction, the equation of the modes (26) must be satisfied. We have then simply

$$(A_1)_n=T(A_3)_n+r_1(A_1)_{n-1}, \tag{40}$$

which is, after $(A_1)_{n-1}$ is subtracted from both sides,

$$(A_1)_n-(A_1)_{n-1}=T(A_3)_n+(1-r_1)(A_1)_{n-1}. \tag{41}$$

The distance between x_n and x_{n-1} is $2W\tan\theta_1$. If this distance is small, we may approximate Eq. (41) by a differential equation,

$$d[A_1(x)]/dx=(1/2W\tan\theta_1)\times\{T(A_3)-(1-r_1)[A_1(x)]\}. \tag{42}$$

Equation (42) predicts how the waves excited in the film progress. It is the same for both TE and TM waves and it also applies to all possible modes. If the gap of the coupler and the amplitude of the incident beam are not uniform, T, r_1, and A_3 vary in x and the solution of Eq. (42) can be obtained by numerical integration. In the following, we consider only the case in which the gap spacing S and the spatial distribution of the wave amplitude A_3 are uniform, thus making an analytical solution of Eq. (42) possible.

Before solving Eq. (42), it is worthwhile to examine it more closely. We refer back to Fig. 14(a) which represents a leaky waveguide. For this case, the term involving (A_3) in Eq. (42) vanishes and the solution for $[A_1(x)]$ becomes an exponentially decreasing function with an exponent, $-(1-r_1)x/2W\tan\theta_1$. Next, we keep the term involving (A_3) in Eq. (42) but omit the term involving $[A_1(x)]$ at the right-hand side of the equation. The solution for $[A_1(x)]$ then increases linearly in x with a slope $T(A_3)/2W\tan\theta_1$. It is thus clear that among the two terms at the right-hand side of the equation, the first term involving (A_3) represents an increase of A_1 due to the energy transfer from the prism into the film and the other term involving $[A_1(x)]$ represents a decrease of A_1 due to the energy transfer from the film to the prism. Because both terms are present in the actual prism–film coupler, as shown by Fig. 14(c), their relative magnitudes determine the form of the solution. At $x\cong0$, where $[A_1(x)]\cong0$, the solution of the equation must thus increase linearly with x. As x increases, the first term, which is proportional to $[A_3]$, remains constant, whereas the other term increases continuously with $[A_1(x)]$. If $d[A_1(x)]/dx$ does not grow indefinitely, the signs of these two terms must be opposite. They eventually cancel at a large x and then $d[A_1(x)]/dx \to 0$. Under this condition, we have from Eq. (42)

$$[A_1(x)/[A_3(x)] = T/(1-r_1),$$

which is the asymptotic solution of Eq. (46) discussed later. For weak coupling, we can show

$$|T/(1-r_1)| \to 2/|T|.$$

VI. PROPERTIES OF THE PRISM-FILM COUPLER

In this section, we discuss the amplitudes and phases of the waves in the film and in the prism, the amount of power transferred from the prism to the film, and the optimum condition of operation.

Let (A_3) be constant over the coupling length from $x=0$ to $x=l$, and zero elsewhere. Assuming $[A_1(x)]=0$ at $x=0$, we have for the solution of Eq. (42),

$$[A_1(x)] = \frac{T(A_3)}{(1-r_1)}\left\{1-\exp\left[-\frac{(1-r_1)x}{2W\tan\theta_1}\right]\right\}, \quad x<l. \quad (43)$$

From Eqs. (24), (25), (30), (31), and (29), we find

$$T^2/R_1R_3 = -TT^*/R_1R_1^* = -TT^*/(1-TT^*).$$

Using this relation and Eqs. (36), (37), and (43), we have

$$[B_3(x)] = R_3(A_3)\left\{1-\frac{TT^*}{1-TT^*}\frac{1}{1-r_1}\times\left[r-\exp\left(-\frac{(1-r_1)x}{2W\tan\theta_1}\right)\right]\right\}, \quad x<l. \quad (44)$$

For all practical circumstances, $\exp(2p_2S)\gg 1$; hence, T is a small quantity and $r_1 \cong 1$. Thus, for a weak coupling we have from Eq. (29)

$$1-r_1 = (1-r_1^2)/(1+r_1) \cong TT^*/2 = |T|^2/2. \quad (45)$$

By use of Eq. (45), Eqs. (43) and (44) become, respectively,

$$\left|\frac{[A_1(x)]}{(A_3)}\right| \cong \frac{2}{|T|}\left[1-\exp\left(-\frac{|T|^2x}{4W\tan\theta_1}\right)\right], \quad x<l, \quad (46)$$

$$\left|\frac{[B_3(x)]}{(A_3)}\right| \cong \left[-1+2\exp\left(-\frac{|T|^2x}{4W\tan\theta_1}\right)\right], \quad x<l. \quad (47)$$

In addition, from Eq. (38) we have

$$|B_1(x)| \cong |A_1(x)|. \quad (48)$$

According to Eqs. (43) and (46) the A_1 wave in the film has a phase equal to that of $T(A_3)$, and an amplitude that rises from 0 at $x=0$ and approaches $(2/|T|)[A_3]$ at large x (see (b) in Fig. 15 for $x<l$). Here $|T|$ is a small quantity. For example, if $|T|=0.1$, the amplitude of the A_1 (or B_1) wave can be 20 times, and thus its power density 400 times greater than that of the incident beam A_3. This large concentration of radiant power inside the film is particularly important for nonlinear and electro-optic devices. We also notice that the weaker the coupling, the smaller becomes $|T|$ and the longer is the distance l to achieve a significant amount of power transfer.

Now consider the reflected wave in the prism $[B_3(x)]$. It has a phase equal to that of $R_3(A_3)$ or $-2\Phi_{32}'$ in Eq. (44), and a magnitude [see Eq. (47)] never larger that that of (A_3) in the region $x<l$. There is, however a phase change of π in $[B_3(x)]$ at $x=0.693$ $(4W\tan\theta_1/|T|^2)$, which is explained in the following way.

We have shown in Eq. (36) and Fig. 14(c) that

$$[B_3(x)] = [B_3'(x)] + [B_3''(x)].$$

It turns out that (B_3') and (B_3'') are always 180° out of phase. The B_3'' wave is a partial reflection of the A_3 wave, and thus (B_3'') is constant in the region $0<x<l$. The B_3' wave is a partial transmission of the B_1 wave, and thus $[B_3'(x)]$ and $[B_1(x)]$ grow together with x. For comparison, $[B_3'(x)]$, $[B_3'(x)]$ and $[B_3(x)]$ are plotted vs x in Fig. 15 (c). The power transferred from the prism to the film between x and $x+\Delta x$ is

$$\Delta P = (c/8\pi)\{(A_3)(A_3)^* - [B_3(x)][B_3(x)]^*\}\Delta x. \quad (49)$$

At $x\cong 0$; $B_3'(x)\cong B_1(x)\cong 0$, we have $|[B_3(x)]|\cong|(A_3)|$ and $\Delta P\cong 0$, which is a minimum. At $x=0.693$ $(4W\tan\theta_1)/|T|^2$; $[B_3'(x)] = -[B_3''(x)]$, we have $[B_3(x)]=0$ and $\Delta P = (c/8\pi)(A_3)(A_3)^*$, which is the maximum. Beyond this point, $[B_3'(x)] > -[B_3''(x)]$ and consequently $[B_3(x)]$ reverses its sign. Finally, at large x (but for $x<l$); $[B_3'(x)]\cong -2[B_3''(x)]$, we again have $|[B_3(x)]| = |(A_3)|$ and $\Delta P\cong 0$. Thus the power in the film reaches its saturation.

In the region $x>l$ where $(A_3)=0$, the solution of Eq. (42) is

$$[A_1(x)] = [A_1(l)]\exp\left[-\frac{(1-r_1)(x-l)}{2W\tan\theta_1}\right] \quad (50)$$

$$\cong [A_1(l)]\exp\left[-\frac{|T|^2(x-l)}{4W\tan\theta_1}\right]; \quad x>l. \quad (51)$$

Using Eqs. (36), (50), and (51), we find

$$[B_3(x)] = \frac{T}{R_1}[A_1(x)] = -2R_3(A_3)\left[1-\exp\left(-\frac{|T|^2l}{4W\tan\theta_1}\right)\right] \times\exp\left[-\frac{|T|^2(x-l)}{4W\tan\theta_1}\right], \quad x>l. \quad (52)$$

Equations (51) and (52) are included at (b) and (c) in Fig. 15 for $x>l$. The amplitudes of the waves inside the film decrease exponentially in the region of $(A_3)=0$. We have discussed earlier, in connection with Fig. 4, that in order to retain optical energy inside the film, it is necessary to decouple the prism from the film at $x>l$. For the B_3 wave, we have shown before that $[B_3(x)]=[B_3'(x)]+[B_3''(x)]$. In the vicinity of $x=l$, $[B_3'(x)]$ varies little but $[B_3''(x)]$ drops to zero because of $(A_3)=0$. The magnitude of $[B_3(x)]$ therefore varies abruptly at $x=l$, and then decreases exponentially to zero at large x.

Finally, we calculate the total power transferred from the prism into the film. The power carried by the film at $x=l$ in the x direction is

$$(c/4\pi)[A_1(l)][A_1(l)]^*W\tan\theta_1.$$

The power supplied by the incident beam between $x=0$ and $x=l$ is

$$(cl/8\pi)(A_3)(A_3)^*.$$

We define an efficiency of the power transfer in the coupler as the ratio of these two quantities. We thus have

$$\text{coupling efficiency}=\frac{[A_1(l)][A_1(l)]^*}{(A_3)(A_3)^*}\frac{2W\tan\theta_1}{l}.$$

Using Eq. (46), we can also express this as

coupling efficiency

$$=\frac{8W\tan\theta_1}{l|T|^2}\left[1-\exp\left(-\frac{|T|^2l}{4W\tan\theta_1}\right)\right]^2. \quad (53)$$

By differentiating Eq. (53) with respect to the argument of the exponential and equating the result to zero, we find that the maximum power transfer occurs if

$$l|T|^2/4W\tan\theta_1\cong 1.25,$$

which is the optimum condition of operation for a prism–film coupler with uniform gap spacing and uniform (A_3). The maximum coupling efficiency computed from Eq. (53) under this condition is slightly greater than 81%.

VII. REFLECTANCES AND TRANSMITTANCES

We have frequently discussed the quantities R_1, R_3, and T; they are important parameters in our analysis. It is essential to attach some physical meaning to these quantities. In spite of a gap of spacing S between the prism and the film, we can consider R_1 and T as the reflectance and transmittance for a wave incident on the upper boundary of the film and R_3 and T as the reflectance and transmittance for a wave incident on the base of the prism. As the gap spacing S approaches zero, Eqs. (24), (25), and (32) may be reduced to

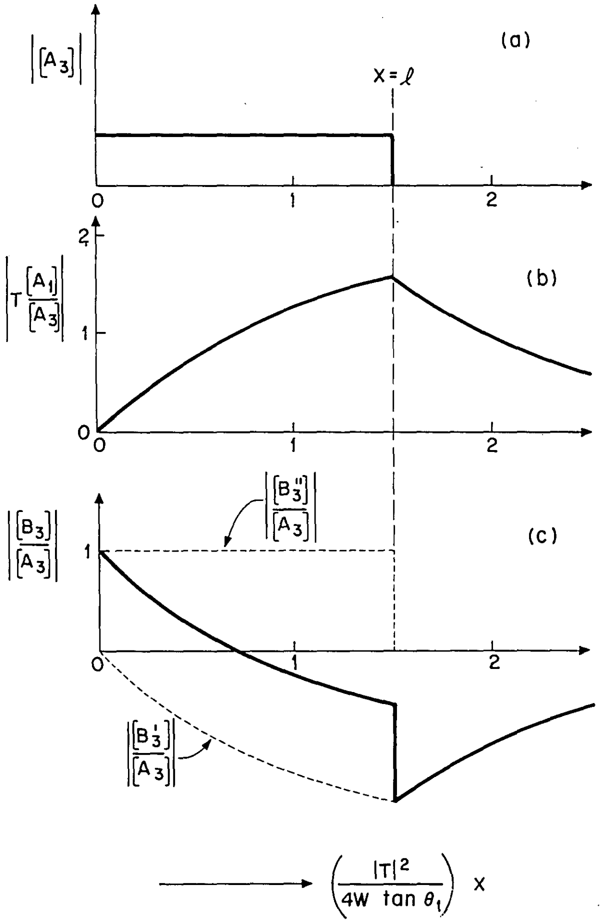

FIG. 15. Distributions of wave amplitudes in a prism–film coupler for (a) incident beam, (b) A_1 or B_1 wave in the film, and (c) reflected beam in the prism.

$$R_1\to(n_1\cos\theta_1-n_3\cos\theta_3)/(n_1\cos\theta_1+n_3\cos\theta_3),$$
$$R_3\to(n_3\cos\theta_3-n_1\cos\theta_1)/(n_1\cos\theta_1+n_3\cos\theta_3),$$
$$T\to[2(n_1n_3\cos\theta_1\cos\theta_3)^{\frac{1}{2}}]/(n_1\cos\theta_1+n_3\cos\theta_3). \quad (54)$$

These are the well-known Fresnel formulas for the reflectance and transmittance of the interface between media n_1 and n_3, provided that the amplitudes of the waves are normalized according to Eqs. (34) and (35). Under our scheme of normalization, the equations come out in more symmetric forms and may be applied to both the TE and TM waves.

VIII. A THIN-FILM INTERFEROMETER

Another way to study the problem of the prism–film coupler is to treat it as an interferometer. In particular, if we consider the quantities R_1, R_3, and T as the reflectance and transmittance, the prism–film coupler is simply a modified form of the Lummer–Gehrke parallel-plate interferometer.[5] Employing this

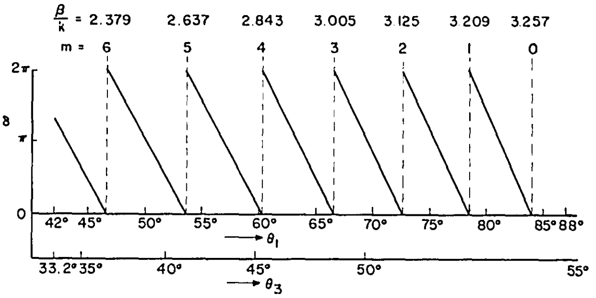

FIG. 16. δ vs θ_1 and θ_3.

concept, we can study the finesse and the line shape of the modes.

Let us consider, in general, that the mode equation (26) is not necessarily satisfied. We can then put

$$2b_1W-2\Phi_{10}-2\Phi_{12}'=2m\pi+\delta, \tag{55}$$

where δ is a function of θ_3 or θ_1. When $\delta=0$, the equation of the modes is satisfied and the incident laser beam is in a synchronous direction. Equation (39) is a recurrence formula for (A_1). By use of Eq. (55), Eq. (39) becomes

$$(A_1)_n=T(A_3)+r_1e^{i\delta}(A_1)_{n-1}. \tag{56}$$

Let $(A_1)_0=0$ at $x=0$; then $(A_1)_1$, at $x=x_1=2W\tan\theta_1$ is $T(A_3)$. By defining $x_q=2qW\tan\theta_1$ and by repeatedly applying the recursion formula (56), we have

$$\begin{aligned}[A_1(q)]&=T(A_3)[1+r_1e^{i\delta}+r_1^2e^{j2\delta}+\cdots+r_1^{q-1}e^{i(q-1)\delta}]\\&=T(A_3)\frac{(1-r_1^qe^{iq\delta})}{1-r_1e^{i\delta}}.\end{aligned} \tag{57}$$

Notice that Eq. (57) is reduced to Eq. (43), if $\delta=0$ and if the difference formula is replaced by a differential formula. We proceed by forming the ratio of the x components of the Poynting vectors,

$$\frac{[A_1(q)][A_1(q)]^*}{(A_3)(A_3)^*}=\frac{|T|^2(1-r_1^q)^2}{(1-r_1)^2}\frac{1+G_q\sin^2(q\delta/2)}{1+F\sin^2(\delta/2)}, \tag{58}$$

where

$$G_q=4r_1^q/(1-r_1^q)^2;\quad F=4r_1/(1-r_1)^2. \tag{59}$$

Equation (58) is in the form of the well-known Airy's formula. In fact, except for the coefficient of Eq. (58), Eqs. (58) and (59) are similar to Eqs. (75), (16), and (76) of Born and Wolf.[8] For the curves of $[A_1(q)][A_1(q)]^*$ vs δ, see Figs. 7.58 and 7.69 of Born and Wolf.[8] The ratio (58) is maximum at $\delta=0$. Note that, for a large q, the term that involves G_q approaches 0 and may be neglected. Thus, $[A_1(q)][A_1(q)]^*$ as a function of δ has the shape of a lorentzian line; the half-power points of Eq. (56) follow from

$$F\sin^2(\delta/2)=1.$$

For small δ and using the approximations $(1-r_1)\cong|T|^2/2$ and $r_1\cong1$, we find

$$\delta(\text{half-power})\cong|T|^2/2. \tag{60}$$

The full line width of any film mode is therefore $|T|^2$, an approximation of which is given in Eq. (27). It is thus possible to determine the coupling parameter, $\exp(-2p_2S)$, from a measurement of the half-power points of δ. Following tradition, we define

$$\text{finesse}=(\pi/2)(F)^{\frac{1}{2}}\cong\pi/|T|^2. \tag{61}$$

If the film considered has absorption loss, we can define a loss coefficient α_l such that the amplitude decreases by a factor $\exp(-\alpha_l)$ over a distance of $2W\tan\theta_1$. Then r_1 in Eq. (57) should be replaced by $r_1(1-\alpha_l)$. In addition, if both T and α_l are much smaller than unity, Eq. (60) becomes

$$\delta(\text{half-power})=|T|^2/2+\alpha_l. \tag{62}$$

In principle, it is thus possible to determine the loss coefficient of the film by observing the line width while reducing the coupling. Any residual line width must then be due to absorption. The line width can be determined experimentally from the width of the dark line observed in the reflected beam, as described in Sec. I. According to Eq. (26), it is also possible to determine the coupling parameter by measuring the shift of line position caused by the difference between Φ_{12}' and Φ_{12} in the mode equations (26) and (15).

Finally, we must correlate δ with θ_3 or θ_1. Unfortunately, an explicit analytical relation between the two quantities is not available. We have, therefore, plotted δ vs θ_3 and θ_1 for a specific example in Fig. 16. Here again we consider $n_0=2.190$, $n_1=3.275$, $n_2=1.000$, $n_3=4.000$, and $\lambda=10.6$ μm. The film thickness is 15 μm. It is interesting that δ varies almost linearly with θ_1 between any two neighboring modes. For all practical purposes, we can estimate δ from θ_1 by using the spacing between the modes for calibration. This is particularly convenient in the experiment for measuring the line widths of the modes.

IX. CONCLUSION

We have presented a theory for the prism–film coupler and thin-film waveguides. We utilized wave optics to calculate the fields in the vicinity of $x=x_n$ and then correlate the fields at different positions along the direction of propagation by ray optics. The method is mathematically simple, and shows clearly what conditions are involved for the waves to add together so as to emerge as a propagating mode. The equations derived can be applied to a prism–film coupler of an arbitrary gap spacing and with an incident beam of an arbitrary distribution of wave amplitudes. However, only the case of the uniform gap and uniform amplitude, A_3, is studied in detail.

To recapitulate, we started by considering a plane wave that bounces back and forth between two film boundaries. If the boundaries were metallic and if we consider a TE wave, E_y would have to be zero at the boundaries, and the mode equation would be simply $2b_1W(=2kn_1\cos\theta_1)=2m\pi$. Because we consider a dielectric film and the fields extend beyond the boundaries of the film, it is necessary to add $-2\Phi_{10}$ and $-2\Phi_{12}$ in the mode equation [see Eq. (15)] where $-2\Phi_{10}$ and $-2\Phi_{12}$ are the phase changes suffered by the waves during their reflections at the boundaries. It is understandable that coupling the film to a prism disturbs the fields and thus a corresponding shift in the mode spectrum occurs [Eq. (26)]. Next, we derived a difference equation for the field distributions in the film [Eq. (41)] and in the prism [Eq. (36)], and proceeded to calculate the power transfer between them [Eq. (53)]. Finally, we gave the condition for the most efficient power transfer.

As the gap spacing, S, between the film and the prism approaches infinity, the film and the prism are decoupled, and our solutions simply describe the phenomenon of total reflection. We have also treated the prism–film coupler as a thin-film interferometer and studied the film loss and the line widths of the modes.

REFERENCES

[1] S. E. Miller, Bell System Tech. J. 48, 2059 (1969).
[2] P. K. Tien, R. Ulrich, and R. J. Martin, Appl. Phys. Letters 14, 291 (1969).
[3] Bell Telephone Laboratories, Incorporated, Murray Hill, N. J.
[4] J. E. Goell and R. D. Standley, Bell System Tech. J. 48, 3445 (1969).
[5] M. Born and E. Wolf, *Principles of Optics, 3rd ed.* (Pergamon, New York, 1965).
[6] D. Marcuse, Bell System Tech. J. 48, 3187 (1969).
[7] W. N. Hansen, J. Opt. Soc. Am. 58, 380 (1967).
[8] Reference 5, pp. 327 and 344.