include metals and semiconductors at frequencies below the plasma frequency $\omega_p$. This fact has led to the study of metals through the simple, but surprisingly good, model in which the dielectric function of a free electron gas is used in the form

$$\varepsilon_2(\omega)=\varepsilon_L\left(1-\frac{\omega_p^2}{\omega(\omega+i\nu)}\right), \qquad \omega_p^2=\frac{Ne^2}{\varepsilon_L\varepsilon_0 m^*}, \tag{6}$$

where $\varepsilon_L$ is the high-frequency dielectric constant which, while it is unity for metals, is usually between 10 and 20 for semiconductors; $N$ is the free electron density; $e$ is the electronic charge; $m^*$ is the effective mass, $\varepsilon_0$ is the permittivity of free space; and $\nu$ is a damping term describing electron–photon collisions, etc. and is quite small, compared to $\omega$, in the optical range of frequencies. If $\nu$ is included the dielectric function of an active medium is a complex quantity with a large negative real part, provided $\omega<\omega_p$. $\varepsilon(\omega)$ is, approximately, real and negative for $\omega<\omega_p$. The plasma frequency for metals is typically $\sim 10^{15}$–$10^{16}$ and in semiconductors $\sim 10^{13}$–$10^{14}$, therefore surface plasmon-polaritons might be expected to be generated at infrared frequencies for semiconductors and at optical frequencies for metals.

For real materials $\nu$ ought to be included in the dielectric function and, as a consequence, equation (5) may then be interpreted in two different ways that correspond to a certain choice in experimental procedure. Firstly the wave number may be assumed to be real so that equation (5) may be solved for a complex frequency, corresponding to a wave that is temporally damped. Alternatively, the frequency may be taken as real so that equation (5) yields a complex wave-number, corresponding to spatial damping. In either case the surface wave is not a true, long-lived, normal mode and leaks energy into the active medium. This can be seen from the fact that $\beta_1$ and $\beta_2$ are complex and hence the fields may have a small propagating component normal to the surface.

The dispersion curves obtained by these two choices are radically different, as can be seen from Figure 2. The curve obtained for real wave number and complex frequency is, actually, almost indistinguishable from the curve that would be calculated with $\nu=0$. In the simple model used here, this curve has a non-dispersive ($d\omega/dk=0$) large wave number limit given by

$$\omega=\omega_p\left(\frac{\varepsilon_L}{\varepsilon_L+\varepsilon_1}\right)^{\frac{1}{2}}. \tag{7}$$

However, the curve obtained with real frequency and complex wave numbers, although similar at low frequencies, does not have such a limit, instead it exhibits a 'bend-back', at a critical frequency, to lower wave numbers. This difference is a fundamental property of 'leaky' or damped waves and, as will be shown later, is closely related to the experimental procedure used

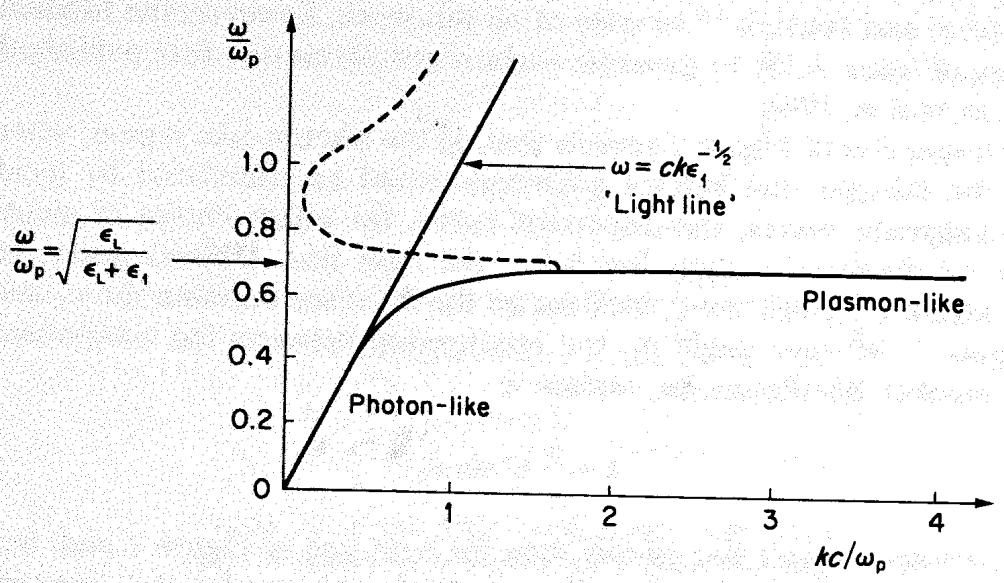


Figure 2. Surface plasmon-polariton dispersion curve for a free electron metal with finite damping plotted for: (a) real $k$, complex $\omega$ (solid curve), (b) real $\omega$, complex $k$ (broken curve). The curves are plotted in dimensionless units

in the measurement of the dispersion curve. Note that, for small $k$, both curves of Figure 2 are asymptotic to what is termed the 'light line', i.e. the straight-line dispersion relationship of electromagnetic waves in the passive medium. This is given by $\omega = ck\varepsilon_1^{-\frac{1}{2}}$ and represents the propagation of light in the passive dielectric.

Several techniques are available in the experimental study of surface plasmon-polaritons, and their use is dictated by the frequency and wave number range in which the excitation is being examined. For example, at short wavelengths (that is large wave number values) where the surface polariton is most 'plasmon-like', the most suitable probe of the dispersion equation is an experiment involving electrons, such as thin-film electron diffraction.[3] Conversely, at longer wavelengths (that is small wave number values) where the surface polariton is most 'photon-like', the most suitable probe is an experiment involving light of the appropriate frequency, and it is one such technique that is to be discussed here, from a computational point of view. It is known as attentuated total reflection (ATR).

ATR involves the use of the evanescent wave that is set up at a medium–air interface when light in a high refractive index medium, such as glass, suffers total internal reflection.[4] It is the reduction of the reflected wave due to absorption that is called ATR. If the weakening occurs by some other means it is usually called frustrated total internal reflection (FTIR). The names, however, have sometimes been interchanged. Rather remarkably, FTIR can be traced back to Newton,[5] and the knowledge and use of ATR has been widespread since the early 1960s due to pioneering work by

Fahrenfort and Harrick.[4,6] In spite of all this work, however, the fundamental idea of using ATR to generate surface polaritons was first published by Otto[7] as late as 1968.

An inspection of Figure 2 reveals that, in the photon-like region, where it might be thought that surface polaritons could be stimulated by incident electromagnetic waves, the dispersion curve, for either choice of solution, which we discussed earlier, lies *below* the light line. Unfortunately, for a plane electromagnetic wave, incident on the interface between the $\varepsilon_1$ and $\varepsilon_2$ of Figure 1, at some angle $\theta_1$, the relationship between the component of wave number parallel to the surface is

$$k = \frac{\omega}{c} \varepsilon_1^{\frac{1}{2}} \sin \theta_1. \tag{8}$$

This corresponds to a line *steeper* than the light line of Figure 2 and, hence, to the region *above* the light line. Indeed, only under the impractical condition of grazing incidence ($\theta_1 = 90°$) does $k$ approach a surface-wave value. The awkward conclusion is that surface polariton waves cannot be generated by propagating plane electromagnetic waves, in a semi-infinite passive medium, on to an interface separating it from a semi-infinite active medium. If surface modes are to be generated, then an ATR system, as first proposed by Otto (Figure 3a) and subsequently by Kretschmann[8] (Figure 3b), must be used.

The first system, proposed by Otto, and shown in Figure 3a will be called the prism–air–medium system (PAM). This consists of a prism (usually hemicylindrical) with a dielectric constant $\varepsilon_1$ separated from a thick sample of the active medium by a small air (vacuum) gap of dielectric constant $\varepsilon_2 < \varepsilon_1$. Light incident, through the prism on to the prism–air interface, at angles greater than the critical angle $\theta_c$ is normally totally reflected back out through the prism. However, under total internal reflection conditions, there will always be an exponentially decaying evanescent field extending into the air gap. The totally reflected wave has a component of wave number parallel to the surface given by equation (8). If the dispersion curve for surface polaritons on a metal–air interface is now considered, as in Figure 4, it can be seen that the curve, in the small $k$ region, lies only *below* the air or vacuum light line ($\omega = ck$). Now the evanescent wave, created by total reflection in the prism, actually exists in the range $\theta_c \leqslant \theta_1 \leqslant 90°$ for which

$$\frac{1}{\sqrt{\varepsilon_1}} < \sin \theta_1 < 1. \tag{9}$$

Hence the range of surface wave number associated with the evanescent wave is

$$\frac{\omega}{c} < k < \sqrt{\varepsilon_1}\frac{\omega}{c}, \tag{10}$$

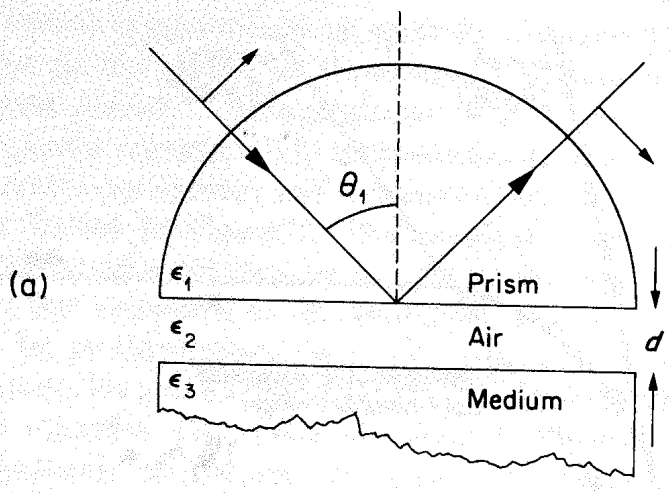


Figure 3a. PAM ATR configuration with an air gap of thickness $d$ separating a hemicylindrical prism from a bulk sample of the active medium. Plane-polarized light is incident at an angle $\theta_1$

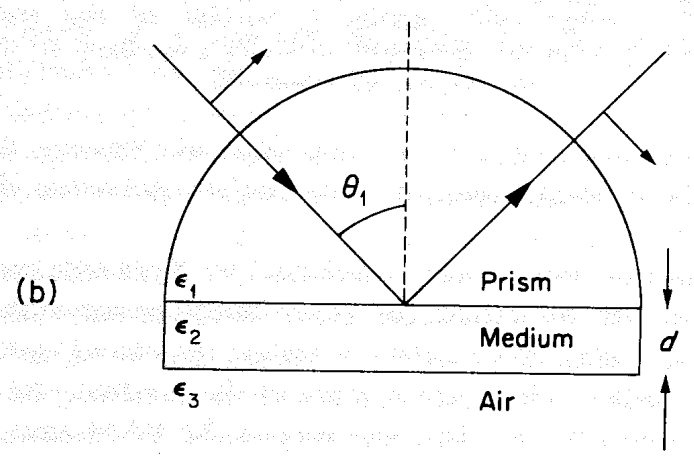


Figure 3b. PMA ATR configuration with an active film of thickness $d$ deposited on the base of the hemicylindrical prism

and corresponds to a region *above* the prism light line but *below* the air/vacuum light line. As indicated on Figure 4, it is clear that, because the evanescent wave has a larger $k$ than the corresponding vacuum wave vector, a range of $k$ of the air/metal surface polariton becomes directly accessible by using an ATR system.

From a practical point of view it is evident that, provided the air gap is sufficiently small, then the evanescent field from the prism can reach the air–medium interface and a surface plasmon-polariton can be excited. Energy from the incident wave is then used to stimulate a surface wave resulting in a weakening of the reflected intensity returned by the prism,

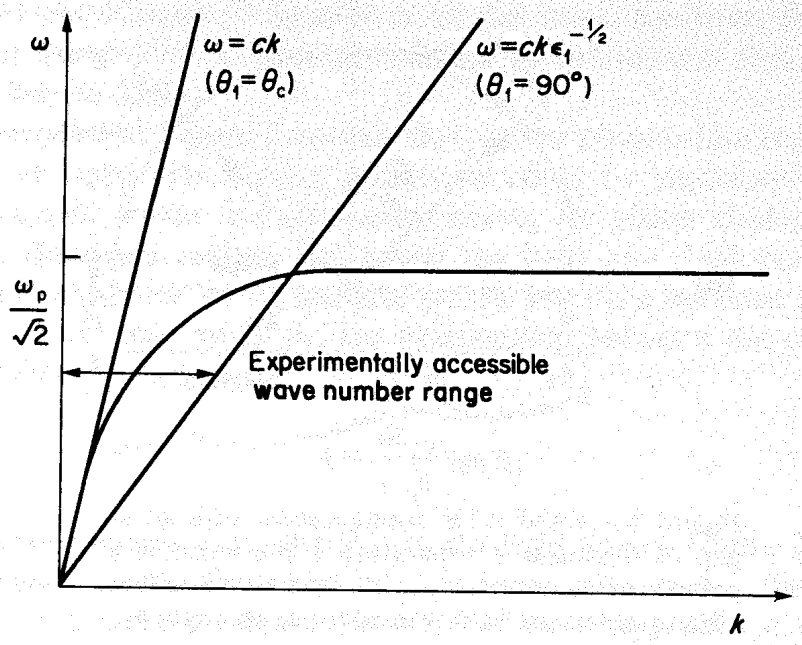


Figure 4. Schematic illustration of the way in which the ATR configuration enables a portion of the surface plasmon-polariton dispersion curve for a medium–air interface to be observed

hence the name 'attenuated total reflection'. By tracing the locus of the minimum in the reflected intensity, the surface polariton dispersion curve may be obtained.

There is also a second system, developed by Kretschmann, and shown in Figure 3b. This will be called the prism–medium–air system (PMA) and consists of a thin film of an active medium deposited onto the base of a hemicylindrical prism. The surface wave at the medium–air interface in the PMA case is excited in just the same way as the PAM case, except that the fields from the prism must now penetrate the film to reach the medium–air interface. In both cases the excitation of a surface plasmon-polariton occurs at the second interface and would be expected to be observed as a sharp minimum in the reflected intensity of (TM) p-polarised light. Also, in each case, (a) the prism light line corresponds to grazing incidence, (b) the air light line corresponds to the prism-air critical angle $\theta_c$, (c) intermediate angles correspond to straight lines of different slope, passing through the origin of Figure 4.

A typical ATR experiment would be to measure, as a function of real frequency, the locus of real angles at which the reflected intensity is a minimum, or to measure, as a function of a real angle, the locus of real frequencies at which a minimum occurs. The experimental dispersion curves, obtained by these approaches, turn out to be significantly different,[9] corresponding, essentially, to the fact that a choice is made as to whether the

resulting surface wave is temporally or spatially damped. This difference is exactly analogous to the choice of theoretical solutions of the dispersion equation described earlier.[10] Measurements with a fixed real angle and scanned real frequency correspond to a theoretical solution of the dispersion equation with a real wave number and a complex frequency, and produce the dispersion curve shown in Figure 2. The imaginary part of the eigenfrequency obtained from the solution of the dispersion equation has a direct interpretation as the linewidth in the frequency scan of the ATR experiment. It should be re-emphasized here that the ATR experiment clearly involves real frequencies and angles in contrast with the theoretical solutions of the dispersion equation. The choice of a fixed real frequency and scanned real angle corresponds to solving the dispersion equation with a real frequency and a complex wave number and results in the folded back curve. The imaginary part of the wave number obtained from the solution of the dispersion equation is interpreted as a linewidth in the angle scan of the ATR experiment. In either case it is clear that the air gap or film width is a critical factor since, if they are too large, there is insufficient coupling of energy into the surface wave and the resulting minimum in the reflected intensity is very weak, or, if they are too small, then the medium/air surface mode will be perturbed by the presence of the second interface (thin-film effects) and the measured dispersion curve will not be that of a free surface wave at a single air–medium interface.

A detailed picture of what is expected experimentally can be built up from a computational model and it is the purpose of this chapter to develop and exploit a model of this kind. For example, it is possible to find the optimum values of the variables (frequency, prism dielectric constant, gap, or thin-film width) for observation of the surface polariton. Moreover, since ATR, like many other optical problems, is adequately theoretically described by simple classical considerations (at least at the frequencies of interest) and leads in many cases to theoretical results which fit experiment very well,[11] a satisfying scope is afforded to the student to examine the behaviour of real systems in the form of computer experiments. In this regard the results of this chapter are best used graphically and also in an 'on-line' interactive fashion so that students may rapidly vary the parameters as they wish, in order to observe the behaviour of the system. The program is therefore designed for use on computer systems with visual display unit (VDU) graphical output capabilities, although simple modifications could be effected to make it suitable for running as a batch program leading to 'hard copy' graphs. In either case, a graph plotting package is really required in addition to this program. If this is not possible, listing the results file will enable the student to plot the graphs by hand.

The exact form of the graphical output in this case is a plot of the reflected intensity of incident p-polarized light, either as a function of

incident angle for a fixed frequency or vice-versa. If more sophisticated graphics software is available, some minor modifications to the program could enable the plotting of a 'reflectivity surface', that is, a three-dimensional plot of reflected intensity both as a function of frequency and angle. This gives very good insight into some of the processes involved in ATR experiments[9,12] since the dispersion curve can be seen directly as a curved valley in the surface.

## 2. ATR MODEL

Consider the simple model of a two-interface system shown in Figure 5. This model, with a suitable choice of the dielectric functions $\varepsilon_1$, $\varepsilon_2$, and $\varepsilon_3$, may be used to describe both PAM and PMA systems of Figures 3a and 3b. For p-polarised light, of unit electric field amplitude, incident through medium 1 at angle $\theta_1$ to the surface normal, the electric field in each medium may be expressed in the form

| | | | |
|---|---|---|---|
| Medium 1: | Incident | $\exp[\mathrm{i}(\alpha_1 \sin\theta_1 x - \alpha_1 \cos\theta_1 z - \omega t)]$ | (11) |
| | Reflected | $r \exp[\mathrm{i}(\alpha_1 \sin\theta_1 x + \alpha_1 \cos\theta_1 z - \omega t)]$, | (12) |
| Medium 2: | Incident | $E_2^t \exp[\mathrm{i}(\alpha_2 \sin\theta_2 x - \alpha_2 \cos\theta_2 z - \omega t)]$, | (13) |
| | Reflected | $E_2^r \exp[\mathrm{i}(\alpha_2 \sin\theta_2 x + \alpha_2 \cos\theta_2 z - \omega t)]$, | (14) |
| Medium 3: | Transmitted | $E_3^t \exp[\mathrm{i}(\alpha_3 \sin\theta_3 x - \alpha_3 \cos\theta_3 z - \omega t)]$, | (15) |

where $r$, $E_2^t$, $E_2^r$, and $E_3^t$ are reflected and transmitted electric field amplitudes, $\alpha_1$, $\alpha_2$, and $\alpha_3$ are the propagation constants in each medium, and $\theta_1$, $\theta_2$, and $\theta_3$ are the angles subtended by the incident and reflected waves to the surface normal.

The boundary conditions appropriate in this model are the continuity of $E_x$, the tangential component of the electric field in the plane of incidence, and the continuity of $D_z$, the normal component of the displacement vector **D**. $D_z$ is given in an isotropic medium by $D_z = \varepsilon_0 \varepsilon E_z$, where $\varepsilon$ is the dielectric function of the medium and $E_z$ is the normal component of the electric field. Application of these boundary conditions at each interface yields the four equations.

$$\cos\theta_1 + r\cos\theta_1 = E_2^t \cos\theta_2 + E_2^r \cos\theta_2, \tag{16}$$

$$\varepsilon_1 \sin\theta_1 - r\varepsilon_1 \sin\theta_1 = E_2^t \varepsilon_2 \sin\theta_2 - E_2^r \varepsilon_2 \sin\theta_2, \tag{17}$$

$$E_2^t \cos\theta_2 \exp(\mathrm{i}\alpha_2 \cos\theta_2 d) + E_2^r \cos\theta_2 \exp(-\mathrm{i}\alpha_2 \cos\theta_2 d) = E_3^t \cos\theta_3 \exp(\mathrm{i}\alpha_3 \cos\theta_3 d, \tag{18}$$

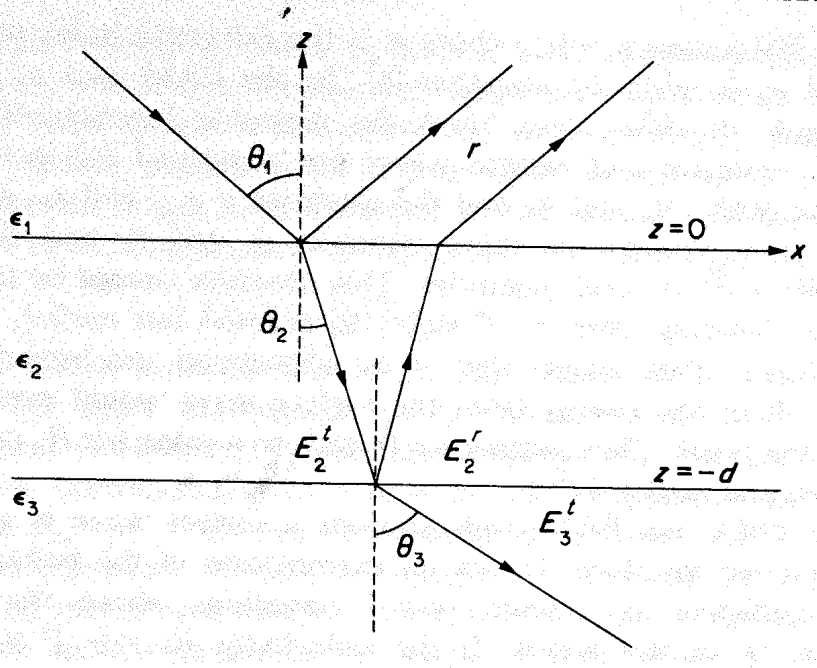


Figure 5. Illustration of the two-interface model for an ATR system. By suitably assigning $\varepsilon_2$ and $\varepsilon_3$ either the PMA or PAM configuration can be treated

$$E_2^t\varepsilon_2 \sin\theta_2 \exp(i\alpha_2 \cos\theta_2 d) - E_2^r\varepsilon_2 \sin\theta_2 \exp(-i\alpha_2 \cos\theta_2 d) = E_3^t\varepsilon_3 \sin\theta_3 \exp(i\alpha_3 \cos\theta_3 d), \quad (19)$$

where the exponential terms common to all the equations have been suppressed and $d$ is the thickness of medium 2.

These equations can be manipulated to give an expression for the reflected amplitude inside medium 1 (the prism) of the standard form

$$r = \frac{r_{12} + r_{23}\exp(2i\alpha_2 \cos\theta_2 d)}{1 + r_{12}r_{23}\exp(2i\alpha_2 \cos\theta_2 d)}, \quad (20)$$

where $r_{12}$ and $r_{23}$ are the Fresnel coefficients of the two interfaces and are given by

$$r_{12} = \frac{\varepsilon_2^{\frac{1}{2}}\cos\theta_1 - \varepsilon_1^{\frac{1}{2}}\cos\theta_2}{\varepsilon_2^{\frac{1}{2}}\cos\theta_1 + \varepsilon_1^{\frac{1}{2}}\cos\theta_2}, \quad (21)$$

$$r_{23} = \frac{\varepsilon_3^{\frac{1}{2}}\cos\theta_2 - \varepsilon_2^{\frac{1}{2}}\cos\theta_3}{\varepsilon_3^{\frac{1}{2}}\cos\theta_2 + \varepsilon_2^{\frac{1}{2}}\cos\theta_3}. \quad (22)$$

Here, $\alpha_2$ is given by $\varepsilon_2^{\frac{1}{2}}\omega/c$ and the angles, using Snell's law, satisfy

$$\cos\theta_2 = \left(1 - \frac{\varepsilon_1}{\varepsilon_2}\sin^2\theta_1\right)^{\frac{1}{2}}, \quad (23)$$

$$\cos\theta_3 = \left(1 - \frac{\varepsilon_1}{\varepsilon_3}\sin^2\theta_1\right)^{\frac{1}{2}}. \quad (24)$$

For the PMA case $\varepsilon_1 = n^2$, where $n$ is the refractive index of the prism, $\varepsilon_3 = 1$ and $\varepsilon_2$ is given by equation (6). In the PAM case $\varepsilon_2$ and $\varepsilon_3$ are interchanged. In either case, since the dielectric function of the active medium is complex and surface waves are generated for incident angles $\theta_1 > \theta_c$, the angles $\theta_2$ and $\theta_3$ and the quatities $r$, $r_{12}$, and $r_{23}$ are complex variables, even though the input variables $\omega$ and $\theta_1$ and the reflected intensity $R_p = |r|^2$ are real quantities. This situation cannot be simplified by setting the damping term $\nu = 0$ since, as pointed out earlier, the surface wave is leaky. This means that if an absorption mechanism were not provided, all of the energy from the surface wave would eventually leak back into the prism. Thus, setting $\nu = 0$ leads to a value for $R_p$ of unity at all frequencies and angles $\theta_1 > \theta_c$.

In both PMA and PAM configurations a surface wave is generated at the medium–air interface. Hence an examination of the behaviour of the Fresnel coefficient $r_{23}$ should reveal something about the generation mechanism of surface waves. If the reflectivity $R_{23} = |r_{23}|^2$ is considered further in, say, the PMA case then for $\theta_1 > \theta_c$

$$R_{23} = |r_{23}|^2 = \left| \frac{(n^2 \sin^2 \theta_1 - \varepsilon_2)^{\frac{1}{2}} - \varepsilon_2 (n^2 \sin^2 \theta_1 - 1)^{\frac{1}{2}}}{(n^2 \sin^2 \theta_1 - \varepsilon_2)^{\frac{1}{2}} + \varepsilon_2 (n^2 \sin^2 \theta_1 - 1)^{\frac{1}{2}}} \right|^2, \tag{25}$$

where it should be remembered that $\varepsilon_2$ for metals, e.g. sodium, aluminium, silver, or gold has, at optical frequencies, a large *negative* real part. A surface wave is generated at $\theta_1 = \theta_p$ with a wave number given by equation (8) so that, by equation (5),

$$n \sin \theta_p = \mathrm{Re} \left[ \frac{\varepsilon_2}{\varepsilon_2 + 1} \right]^{\frac{1}{2}}, \tag{26}$$

where Re denotes the real part. At $\theta_1 = \theta_p$ the denominator of $r_{23}$ tends to zero and $r_{23} \to \infty$. The physical meaning of this is clear because it is an example of a resonance in which a finite-amplitude electromagnetic wave can be created at the surface without any incident power.[13] If, for the moment, we consider a single interface between two semi-infinite media labelled 2 (metal) and 3 (air), or vice versa, then $r_{23}$ is the conventional reflection coefficient of an incident plane electromagnetic wave. Now $r_{23}$ can be regarded as a ratio $N/D$ so that $N = 0$ corresponds to the familiar Brewster effect $r_{23} = 0$ while, neglecting the imaginary part of $\varepsilon_2$ for the moment, $D = 0$ corresponds to the surface plasmon-polariton resonance under discussion here. This fascinating connection allows us to regard the Brewster effect as a surface-wave phenomenon. The $N = 0$ case corresponds to unbound, radiative surface waves (Brewster modes) while the $D = 0$ case corresponds to bound, non-radiative surface waves (called Fano modes in honour of Fano who published an early pioneering paper on this topic[14]).

Naturally, finite damping makes $\varepsilon_2$ have a small imaginary part and in a real system $r_{23}$ will simply become very large as $\theta_1$ approaches $\theta_p$.

Since the ATR system is a coupled two-surface system, the dispersion relation of surface polaritons on a free metal/air surface is not strictly observed. Instead it is a form disturbed by the presence of the prism and corresponds to the denominator of equation (20), being zero, i.e.

$$1 + r_{12} r_{23} \exp(2i\alpha_2 \cos\theta_2 d) = 0. \tag{27}$$

It should be noted, however, that in the limit, when the two interfaces are sufficiently far apart (uncoupled) to be treated separately, the single interface result is obtained. It may seem strange that the launching of a surface wave, corresponding to a pole in $r$ is expected to be seen as a minimum in $R_p = |r|^2$. This can be understood, however, by remembering that a direct computer solution of equation (27) would involve complex quantities and therefore a solution in terms of complex frequency and real wave number of vice versa must be sought. If instead of solving equation (27) the response in the $R_p$, of an ATR calculation, is observed by varying *real* frequency and *real* angle, it cannot be expected that a true solution of (27) will be found. In fact the ATR response actually corresponds to a projection, on to the real frequency-angle plane of a multidimensional complex space, of the solution of the dispersion equation, and is thus not observed as a pole in $R_p$ but as a sharp minimum.[15] The fact that a wave with real frequency, incident at a real angle, can be used to launch a leaky wave satisfying a complex dispersion relation can also be understood in terms of a 'forced' resonance[14,16] phenomenon.

## 3. ATR COMPUTER PROGRAM

The core of an ATR program of the kind indicated in section 1, is a very simple calculation. Once the choice of ATR configuration has been made, the system parameters such as $d$, $\varepsilon_1$, etc. specified, and the dielectric functions appropriately assigned to each of the three media in the model, the process consists simply of a loop, either over frequency or angle, in which $R_p = |r|^2$ is calculated from equation (20) using the same short subroutine for both loops. The results may then be written to a file for subsequent reading at a later stage, and also stored in arrays for the purposes of graphical output. When the loop is completed, a graph may then be plotted, of $R_p$ as a function of the chosen variable. On completion of this graph, the student may then wish to repeat the operation with different parameters to optimize the conditions for observation of the surface-wave minimum, or having done so, to carry out a set of scans to obtain the dispersion curve. The program, therefore, must return to some convenient point for recalculation.

Another important point in programs of this kind is that some students may need more complete prompting than others, to ensure that sensible parameters are chosen for experiment. Students should therefore have a choice of short prompting for their input of variables, or of more explanatory prompting that they can subsequently switch off as their experience grows.

There are also some more detailed points in the calculation that make the program more convenient for the student but, since they involve further choices of action, complicate the simple basic operation.

For example, when choosing a scan of frequency for fixed angle, the student is constrained to calculating the dielectric function of the active medium from equation (6), using the free electron model, since this quantity is frequency dependent. However, in the case of a scan over angle for fixed frequency, the dielectric function, being independent of angle in this simple model, is a constant across the scan. This means that it is possible either to calculate the dielectric function from (6) or to use directly experimental values of the optical constants $n_r$ and $n_i$, that is the real and imaginary parts of the refractive index, knowing the simple relation $\varepsilon = (n_r + in_i)^2$. If the free electron model is used, the required data are the high-frequency dielectric constant $\varepsilon_L$ (unity for all metals), the plasma frequency $\omega_p$, and the damping constant $\nu$. Although the value of this last parameter is not critical, it must be there. For most cases it is about 1 per cent of the plasma frequency.

To understand some of the other technical aspects of the program it is useful to examine the flow diagram of Figure 6 containing the whole program structure except for the choices in prompting options that would overcomplicate the diagram.

After the usual declarations, specification of the arrays and opening of input, output, and data-file channels, the program offers a simple explanation of the ATR technique in a few lines. The user is then asked if detailed input prompting is required. Depending on the answer, a parameter is set that controls the form of prompting for the rest of the program. At the start of each new calculation the program returns to this point so that the users can decide each time if their memories need refreshing about typical values for certain parameters. Following this, an identifying name for the active medium is requested, of up to 12 letters. This is merely for convenient identification of the data stored in the data file, and to label the graph.

The data for the program is put in at this stage, starting with the ATR configuration required. Depending on the response to this, either the air gap or film thickness in nanometres is requested, together with the refractive index of the prism. A further choice is then required of a fixed angle (varying frequency) or fixed frequency (varying angle) scan of the reflected intensity.

If a fixed angle scan is chosen, the model is constrained to the free

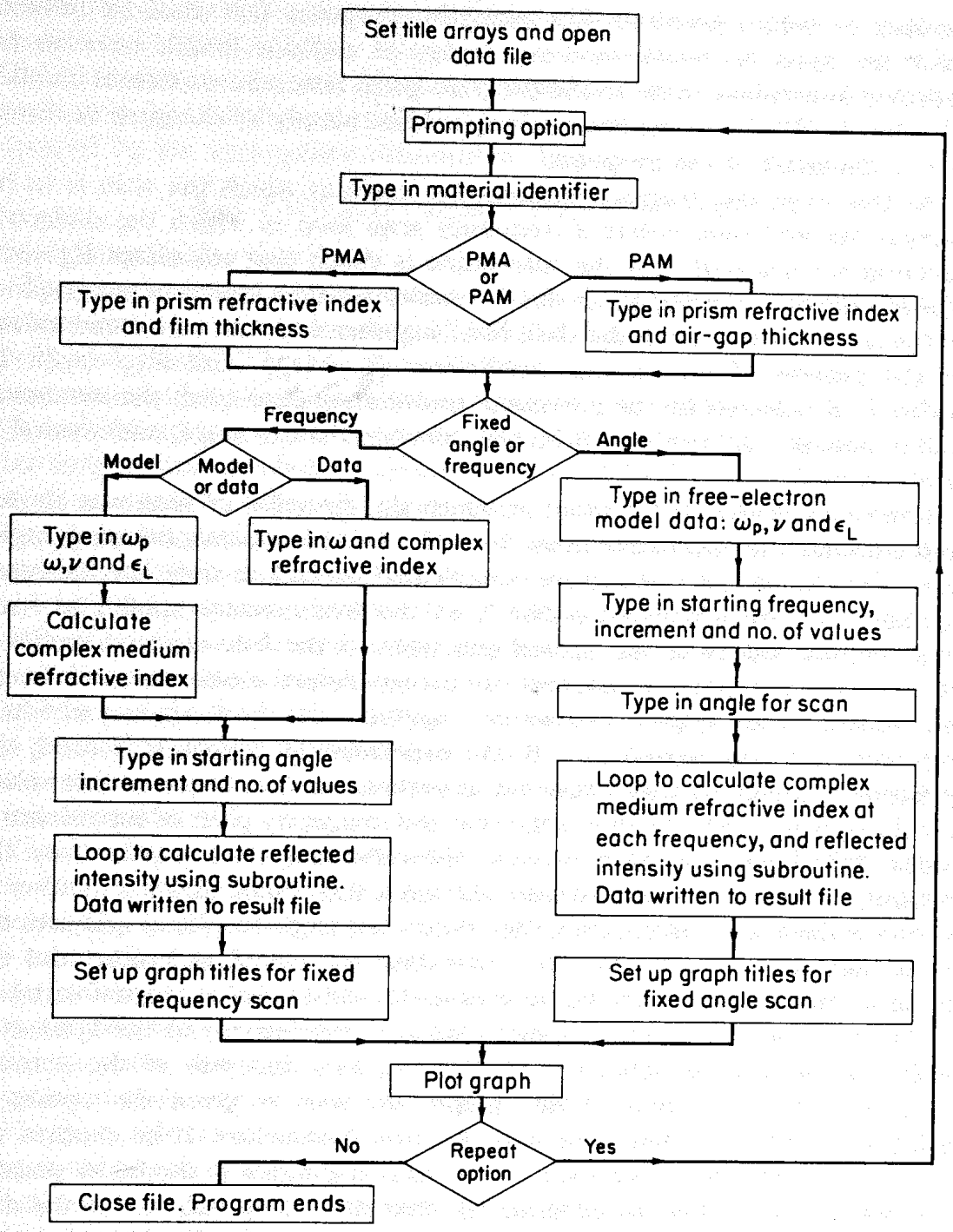


Figure 6. Flow diagram for the ATR program, illustrating the options available to a user

electron model (since the dielectric function of the active medium is frequency dependent). In this model, the plasma frequency, high-frequency dielectric constant, and damping parameter are required as constants for the scan. In addition the user then specifies the range and precision of the frequency scan in terms of a starting frequency, an increment, and the

number of values required. For graphical purposes this must be limited, since for most computer systems, arrays of definite length have to be specified in graphics work. In the program given here, the maximum number of points is 200; however, this can be increased simply by changing the array size at the head of the program.

At this stage the program requests the angle at which the scan is to be carried out and then enters a frequency scan loop in which the dielectric function is calculated, and the subroutine is called that calculates $R_p = |r|^2$. The frequency value and $R_p$ at this frequency are then stored in the graphics arrays and also output to the data files, together with the calculated values of the squares of the Fresnel coefficients $R_{12} = |r_{12}|^2$ and $R_{23} = |r_{23}|^2$. In addition, if required by the particular graphics package used, the maximum and minimum values of $\omega$ and $R_p$ are calculated at this stage, and a suitable title specified.

If we now return to the point at which the direction of scan was chosen and consider the alternative possibility of a fixed frequency (varying angle) scan, this choice leads to a more complicated procedure since the dielectric function may be calculated either from the free electron model or from experimental values of the optical constants. If the free electron model is chosen, the plasma frequency, high-frequency dielectric constant, and damping parameter are required as before, together with the frequency at which the scan is to be carried out. If the experimental approach is used, the program requires the scan frequency as before, but then requires only values for the optical constants, that is the real and imaginary parts of the refractive index. Whichever approach is used, the remaining data required are the starting value of angle for the scan, the angle increment, and the number of points in the scan. The program then enters the angle scan loop and uses the same subroutine to calculate $R_p$. The data are stored as before and the program then converges to the same point for either choice of fixed variable.

At this stage a suitable graphics package (appropriate to the system) is called to display the reflected intensity $R_p$ as a function of the scanned variable. On completion of this graph, the user is given the option of terminating the program or of selecting new parameters. If he chooses the former, the program closes the data file and terminates. It should be pointed out that at the start of the program the data file is emptied, so that the data file should be read, if required, immediately after terminating the program.

If the user wishes to perform further calculations, the program returns to the point at which the level of prompting is specified. Each subsequent calculation adds a new table to the data file so that the accumulated results are in a convenient neatly labelled form.

Several features of this program will vary between computer systems; however, these points could be quite simply modified. For example, input data in this program are handled mostly by a free-format subroutine FREAD. The advantage of this is that the use of free format reduces the

possibility of format errors by the user. Most systems have a package equivalent to FREAD, but if one is not available a conventional FORTRAN format is adequate. Similarly the input–output and data file assignments will be different. In this case the package FTNCMD should be replaced by the appropriate subroutines. In particular the command

CALL FTNCMD ('$EMPTY RESULT OK', 16)

simply empties the data file.

The length of the logical variables used in the character input will depend on the system, and the function EQUC(Y1, Y2) merely checks if the two variables Y1 and Y2 are the same. This function can clearly be replaced by a conventional statement or equivalent function. The logical variables Y1, Y2, Y3, and Y4 are test variables against which are checked the responses of the user to questions involving a choice of action. Y1 = Y(yes) is used to check yes or no answers. Y2 = A(angle) is used to check the response to a choice of angle or frequency scan. Y3 = D(data) checks the response to a choice of experimental data or the free electron model in the case of an angle scan. Y4 = M(PMA systems) identifies the ATR configuration. The variables Y5 and Y6 (2) are used to receive the responses to the questions and are then checked against these 'constants' for the choice of action. In most cases the first letter of the response is sufficient to delineate the choice. However, in the case of the choice between the PMA and PAM systems it is the second letter which is used and hence Y6 has two elements; the second of which is checked against Y4.

The remaining system-specific aspects of the program are graphical ones. The core of the graphics is the external package which plots the reflected intensity $R_p$ as a function of the chosen variable. In this case, this package is the subroutine CGPL, and the required parameters will be broadly the same in any package. However, the associated subroutines used in this program may be discarded if not required or replaced by equivalent packages. In order to make this task easier, there follows a brief explanation of the purpose of each routine in order of appearance in the program.

Firstly, a title array of some kind is usually required by most packages. This is contained in the array T. In addition, since the abscissa of the plot will depend on the choice of variable to be scanned, the title for this must be assigned after the data arrays X and Y are filled. Consequently, two alternative titles, contained in the variables TA1 to TA4 and TF1 to TF4, are specified at the beginning of the program and assigned to T at the appropriate stage.

Early in the program, the routine AUX093 is used to specify the device upon which the graph is to be plotted. This is a common requirement and most systems have equivalent routines. Note that by a suitable choice here, hard copies of graphs could be obtained although this has not been done here, since a visual display is the immediate purpose of this work.

The graph plotting package CGPL used here requires scale factors to define the units of the X and Y axes. These are calculated from SCX and SCY which in turn use the routine SLE to obtain a convenient scale. In many systems a scale factor is not required so that SCX, SCY, and the routine SLE or its equivalent may often be left out of the program. However, the maximum and minimum values of the data points, XMAX, YMAX, XMIN, and YMIN are usually required.

Finally the routine PLOT closes the graphics package data file. An equivalent is usually required to avoid 'overplotting' errors whereby some points from each graph are plotted at the next call to the graph plotting routine.

The remainder of the variables involved in CGPL are simply those involved in the choice of symbols, form of curve, etc. It has been found that the cross is usually the most suitable symbol, and since most curve-fitting procedures are inclined to 'overshoot' on rapidly varying functions such as $R_p$, it is generally better to plot only individual data points rather than a smooth curve.

The modifications outlined above should not present any serious problems to anyone who is reasonably familiar with their own graphics facilities.

## 4. TYPICAL RESULTS

As an example of the kind of results obtained from this program, data for sodium (a good free electron metal, which has a complex refractive index $0.044+i2.42$ at $\omega = 3.2\times10^{15}$ rad $s^{-1}$)* was used. The program was run for a fixed frequency scan, and a fixed angle scan. Table 1 contains a truncated copy of the data file showing the interesting parts of a fixed frequency scan using the PMA configuration. The graph obtained with this data is shown in Figure 7.

It can be seen that the reflected intensity in the fourth column of the table exhibits a sharp peak at 42° (the critical angle for a prism–air interface). This feature is typical of the PMA configuration and is due to the fact that the final medium in the system ($\varepsilon_3$) is air. An examination of the third column containing $R_{23} = |r_{23}|^2$ reveals a very large peak near 47° corresponding to the launching of a surface wave associated with the medium–air interface. Note that the reflected intensity shows a minimum very near to this angle, the shift being due to the fact that perturbation occurs because we really have a coupled system, even in an optimum parameter situation. The degree to which this perturbation occurs should now be studied as a function of the film thickness to determine the optimum value of $d$ consistent with a reasonable reduction from unity of $R_p$.[15]

* All frequencies used in the program are angular frequencies whose proper units are rad $s^{-1}$. In the program we have used the tidier, but not strictly correct, label Hz.

TABLE 1

SUBJECT MEDIUM IS SODIUM

PRISM-MEDIUM-AIR CALCULATION

PRISM REFRACTIVE INDEX= 0.1500E+01
FILM THICKNESS= 0.4000E+02 NANOMETRES

FIXED FREQUENCY CALCULATION

CALCULATION USING REFRACTIVE INDEX DATA

SCAN FREQUENCY= 0.3200E+16
GIVEN COMPLEX REFRACTIVE INDEX= 0.4400E-01 +J 0.2420E+01

| ANGLE(DEG) | R12 | R23 | R |
|---|---|---|---|
| 0.0 | 0.9680E+00 | 0.9747E+00 | 0.6787E+00 |
| 0.1000E+01 | 0.9680E+00 | 0.9747E+00 | 0.6787E+00 |
| 0.2000E+01 | 0.9679E+00 | 0.9746E+00 | 0.6785E+00 |
| . | . | . | . |
| . | . | . | . |
| 0.3600E+02 | 0.9610E+00 | 0.9603E+00 | 0.6260E+00 |
| 0.3700E+02 | 0.9607E+00 | 0.9600E+00 | 0.6282E+00 |
| 0.3800E+02 | 0.9604E+00 | 0.9602E+00 | 0.6339E+00 |
| 0.3900E+02 | 0.9601E+00 | 0.9613E+00 | 0.6460E+00 |
| 0.4000E+02 | 0.9599E+00 | 0.9643E+00 | 0.6715E+00 |
| 0.4100E+02 | 0.9596E+00 | 0.9719E+00 | 0.7318E+00 |
| 0.4200E+02 | 0.9594E+00 | 0.2179E+01 | 0.9776E+00 |
| 0.4300E+02 | 0.9591E+00 | 0.8131E+01 | 0.9760E+00 |
| 0.4400E+02 | 0.9589E+00 | 0.2232E+02 | 0.9573E+00 |
| 0.4500E+02 | 0.9587E+00 | 0.6757E+02 | 0.9022E+00 |
| 0.4600E+02 | 0.9585E+00 | 0.3204E+03 | 0.7930E+00 |
| 0.4700E+02 | 0.9583E+00 | 0.8271E+04 | 0.6940E+00 |
| 0.4800E+02 | 0.9582E+00 | 0.5893E+03 | 0.6898E+00 |
| 0.4900E+02 | 0.9581E+00 | 0.1718E+03 | 0.7333E+00 |
| 0.5000E+02 | 0.9580E+00 | 0.8778E+02 | 0.7764E+00 |
| 0.5100E+02 | 0.9579E+00 | 0.5627E+02 | 0.8092E+00 |
| 0.5200E+02 | 0.9578E+00 | 0.4070E+02 | 0.8333E+00 |
| . | . | . | . |
| . | . | . | . |
| 0.8500E+02 | 0.9891E+00 | 0.6518E+01 | 0.9850E+00 |
| 0.8600E+02 | 0.9912E+00 | 0.6488E+01 | 0.9880E+00 |
| 0.8700E+02 | 0.9934E+00 | 0.6465E+01 | 0.9909E+00 |
| 0.8800E+02 | 0.9956E+00 | 0.6449E+01 | 0.9939E+00 |
| 0.8900E+02 | 0.9978E+00 | 0.6439E+01 | 0.9970E+00 |
| 0.9000E+02 | 0.1000E+01 | 0.6436E+01 | 0.1000E+01 |

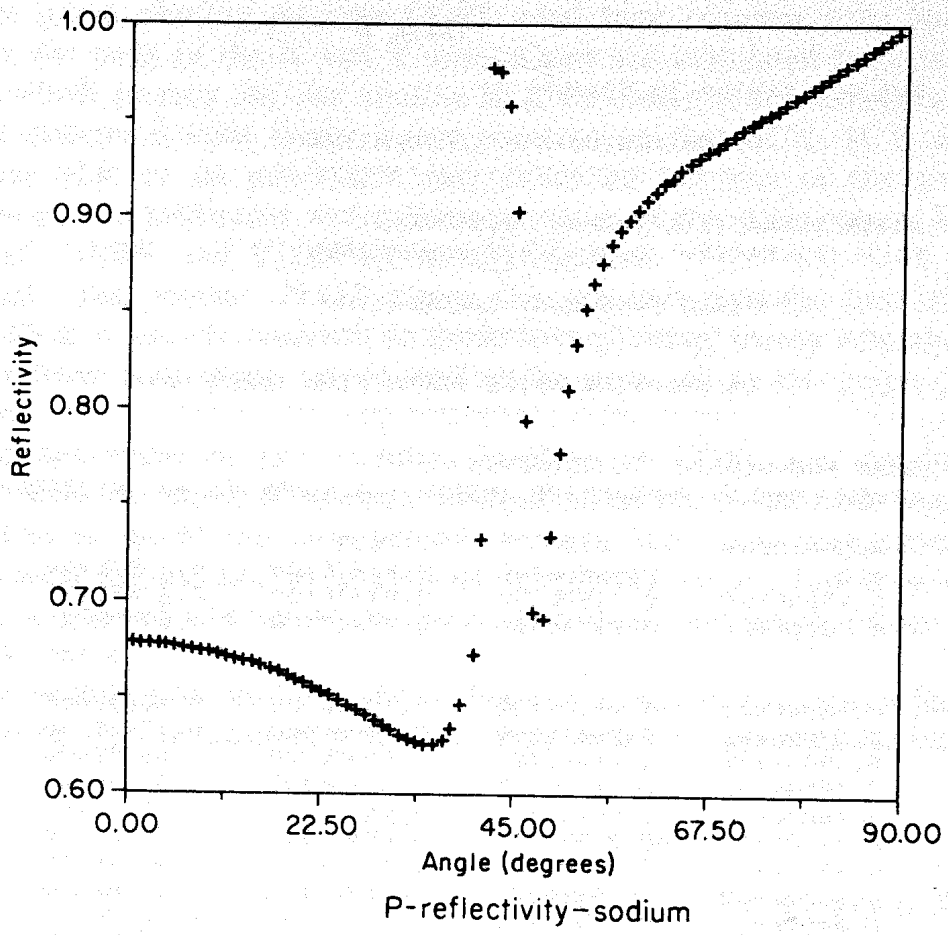


Figure 7. Plot of reflected intensity $R_p$ as a function of angle of incidence for sodium data using the PMA configuration. Angular frequency of scan is $3.2\times10^{15}$ rad s$^{-1}$ ($\approx0.4\times\omega_p$)

As a second example, a truncated set of results is displayed in Table 2 for a fixed angle scan in the PAM configuration using a free electron model with data appropriate to sodium. The graph obtained with this data and configuration is shown in Figure 8. This scan at 60° shows a deep minimum in the reflected intensity at a frequency a little above $\omega/\omega_p = 0.5$. This corresponds to a 'pole' in $R_{23}$ that is at a slightly higher frequency, the difference again being due to the perturbation by the prism–air interface. Note that $R_{12} = |r_{12}|^2$ is unity for all frequencies since the scan is made at an angle greater than $\theta_c = 42°$, the critical angle for the prism–air interface.

These examples vividly illustrate the principle of attenuated total reflection. Other metals can be considered, of course, e.g. silver with a complex dielectric constant $\varepsilon = -18.3 + i0.4$ at the He–Ne laser[17] frequency $\omega = 2.979\times10^{15}$ rad s$^{-1}$, or aluminium[18] with complex refractive index $\sqrt{\varepsilon} = 1.53 + i5.65$ at $\omega = 2.218\times10^{15}$ rad s$^{-1}$. Semiconductors can also be investigated such as GaAs or, more popularly, InSb[19] for which $\varepsilon_L = 15.68$. Computational experiments can be made with either experimental refractive

TABLE 2

SUBJECT MEDIUM IS SODIUM

PRISM-AIR-MEDIUM CALCULATION

PRISM REFRACTIVE INDEX= 0.1500E+01
AIR GAP THICKNESS= 0.2000E+03 NANOMETRES

FIXED ANGLE CALCULATION

FREE ELECTRON MODEL CALCULATION

PLASMA FREQUENCY= 0.8230E+16
DAMPING TERM= 0.1236E+15
HIGH FREQ. DIELECTRIC CONSTANT= 0.1000E+01
ANGLE FOR ATR SCAN= 0.6000E+02 DEGREES

| FREQ. | R12 | R23 | R |
|---|---|---|---|
| 0.3500E+16 | 0.1000E+01 | 0.2443E+02 | 0.9788E+00 |
| 0.3525E+16 | 0.1000E+01 | 0.2577E+02 | 0.9782E+00 |
| 0.3550E+16 | 0.1000E+01 | 0.2722E+02 | 0.9776E+00 |
| 0.3575E+16 | 0.1000E+01 | 0.2879E+02 | 0.9769E+00 |
| 0.3600E+16 | 0.1000E+01 | 0.3051E+02 | 0.9762E+00 |
| 0.3625E+16 | 0.1000E+01 | 0.3239E+02 | 0.9754E+00 |
| 0.3650E+16 | 0.1000E+01 | 0.3444E+02 | 0.9745E+00 |
| 0.3675E+16 | 0.1000E+01 | 0.3671E+02 | 0.9736E+00 |
| 0.3700E+16 | 0.1000E+01 | 0.3920E+02 | 0.9725E+00 |
| 0.3725E+16 | 0.1000E+01 | 0.4196E+02 | 0.9713E+00 |
| 0.3750E+16 | 0.1000E+01 | 0.4503E+02 | 0.9700E+00 |
| 0.3775E+16 | 0.1000E+01 | 0.4845E+02 | 0.9686E+00 |
| 0.3800E+16 | 0.1000E+01 | 0.5227E+02 | 0.9669E+00 |
| 0.3825E+16 | 0.1000E+01 | 0.5657E+02 | 0.9651E+00 |
| . | . | . | . |
| . | . | . | . |
| 0.4275E+16 | 0.1000E+01 | 0.7457E+03 | 0.7099E+00 |
| 0.4300E+16 | 0.1000E+01 | 0.1004E+04 | 0.6296E+00 |
| 0.4325E+16 | 0.1000E+01 | 0.1405E+04 | 0.5221E+00 |
| 0.4350E+16 | 0.1000E+01 | 0.2044E+04 | 0.3892E+00 |
| 0.4375E+16 | 0.1000E+01 | 0.3050E+04 | 0.2591E+00 |
| 0.4400E+16 | 0.1000E+01 | 0.4376E+04 | 0.1972E+00 |
| 0.4425E+16 | 0.1000E+01 | 0.5227E+04 | 0.2545E+00 |
| 0.4450E+16 | 0.1000E+01 | 0.4617E+04 | 0.3957E+00 |
| 0.4475E+16 | 0.1000E+01 | 0.3282E+04 | 0.5464E+00 |
| 0.4500E+16 | 0.1000E+01 | 0.2193E+04 | 0.6677E+00 |
| 0.4525E+16 | 0.1000E+01 | 0.1491E+04 | 0.7556E+00 |
| . | . | . | . |
| . | . | . | . |
| 0.4900E+16 | 0.1000E+01 | 0.8518E+02 | 0.9860E+00 |
| 0.4925E+16 | 0.1000E+01 | 0.7682E+02 | 0.9877E+00 |
| 0.4950E+16 | 0.1000E+01 | 0.6961E+02 | 0.9890E+00 |
| 0.4975E+16 | 0.1000E+01 | 0.6336E+02 | 0.9902E+00 |
| 0.5000E+16 | 0.1000E+01 | 0.5790E+02 | 0.9913E+00 |

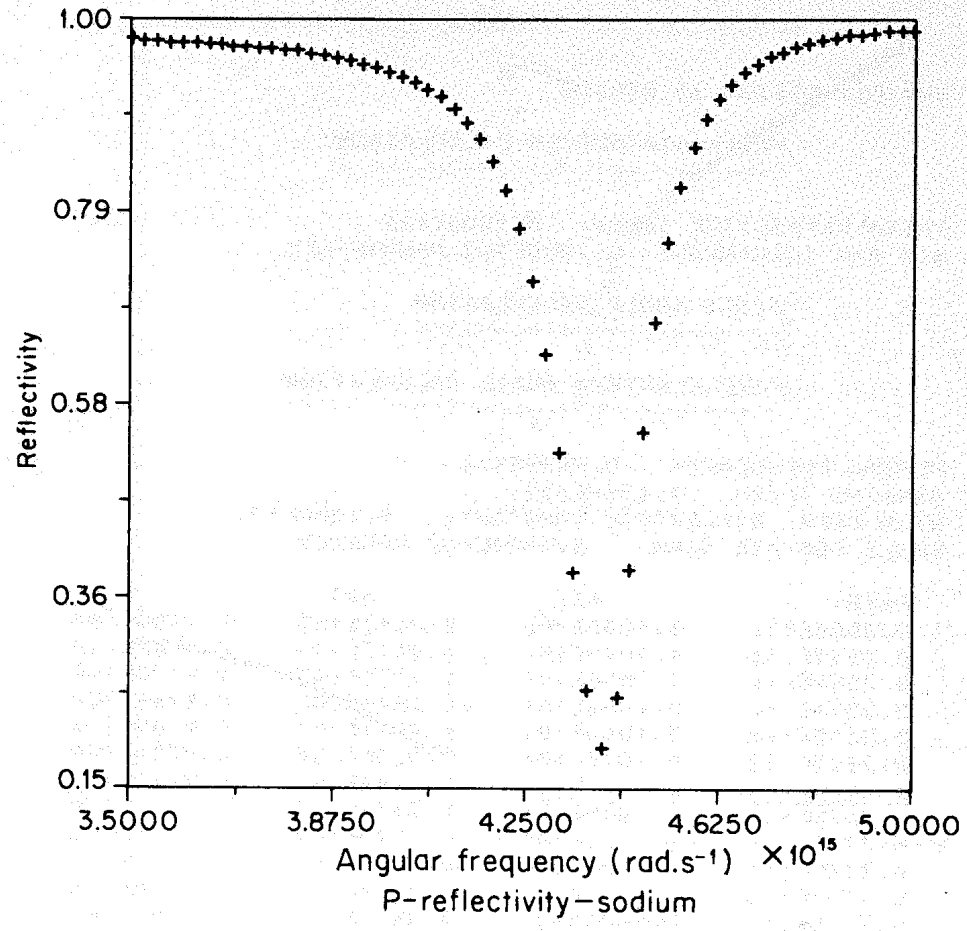


Figure 8. Plot of reflected intensity $R_p$ as a function of incident angular frequency for sodium using a free electron model for the PAM configuration. Angle of scan is 60°

index data or model parameters. For example, in the case of semiconductors, $\omega_p$ can be varied as the electron density is varied and the free electron model can be easily modified to incorporate phonon terms as $\omega_p$ is lowered. The variation of the results with the size of damping parameter is also an interesting one as is the variation of $\theta_{min}$, the angle of the minimum in $R_p$, with air gap thickness (PAM), or film thickness (PMA). Only for certain values of $d$ will $\theta_{min}$ become $\theta_p$ and, in the case of PMA, double-valued results will appear at larger $d$.

Finally, as can be appreciated, there is ample scope for experimentation and any results obtained will require interesting, and often subtle, physics for a satisfactory interpretation to be found.

## REFERENCES

1. E. Burstein, *Polaritons*, Eds. E. Burstein and F. de Martini (Pergamon, New York), 1 (1974).
2. E. Burstein, A. Hartstein, J. Schoenwald, A. A. Maradudin, D. L. Mills, and R. F. Wallis, *Polaritons*, Eds. E. Burstein and F. de Martini, (Pergamon, New