Contemporary Physics, 1991, volume 32, number 3, pages 173–183

Optical excitation of surface plasmons: an introduction

J. R. SAMBLES, G. W. BRADBERY and FUZI YANG

Beginning from low level concepts the basic understanding for the optical excitation of surface plasmons is developed. Prism coupling using the attenuated total reflection technique is discussed as well as the less traditional grating coupling technique. A brief discussion of some recent developments using twisted gratings is also presented. Finally a short summary of the potential device applications is given.

1. Introduction

The interaction of electromagnetic radiation with an interface can generate interesting surface excitations. There are, in electromagnetic terms, a range of interfaces of interest, for example dielectric to dielectric, dielectric to semiconductor and dielectric to metal. Here we shall be specially concerned with the dielectric to metal (strong reflector) interface where it is possible to have normal components of $\boldsymbol{E}$ fields in both media, which are both directed towards, or away from, the interface, ending or beginning at charges on the interface. If we make one of the two materials a dielectric such as vacuum or air and the other a metal, then as shown later, there may exist at the interface a trapped surface mode which has electro-magnetic fields decaying into both media but which, tied to the oscillating surface charge density, propagates along the interface. This is the surface plasmon which may be coupled to by optical radiation using suitable arrangements and which then may be used as an optical monitor of changes in the local environment. Thus it may be employed in studies of electrochemistry; catalysis; wetting; thin organic condensates; in biosensing; in gas or chemical sensing etc. It is also of interest, because it is a resonant phenonemon with strongly enhanced local optical fields, in non-linear optics particularly with non-centrosymmetric Langmuir–Blodgett film overlayers. Some of these potential applications for surface plasmons will be returned to at the end of this article. First however we develop some theory to show the existence of the surface plasmon, discuss methods by which it can be excited and show some typical experimental results.

Authors' address: Thin Film and Interface Group, Department of Physics, Exeter University, Stocker Road, Exeter, EX4 4QL, England

2. Simple theory

In order that we may study this interface and the interesting electromagnetic phenomenon which occurs there we need first to examine some relatively simple concepts of solid state physics and electromagnetism. Electromagnetic radiation in isotropic media consists of orthogonal oscillating electric and magnetic fields transverse to the direction of propagation. If, as is often the case, we pass such radiation through a linear polarizer then the radiation being transmitted will be plane polarized. This means there is a well specified plane in which $\boldsymbol{E}$ or $\boldsymbol{B}$ oscillate, this plane containing the appropriate electromagnetic field vector and the direction of propagation. Now if we consider such radiation falling at an incident angle θ_i upon a smooth planar interface then we have to consider two important situations.

In the first instance, illustrated in figure 1, the incident radiation has its electric vector in the plane of incidence (that plane perpendicular to the surface which contains both the incident and reflected wavevector). This is called p-polarized (parallel to the plane of incidence). For such radiation clearly the $\boldsymbol{B}$ vector has only one component, B_y, tangential to the interface (it is transverse magnetic or TM radiation) and its $\boldsymbol{E}$ vector has components E_z, normal and E_x tangential to the interface. In the second case the incident radiation is polarized so that its electric field vector is orthogonal to the plane of incidence. This is called s-polarized (from the german *senkrecht* meaning perpendicular). For such radiation clearly the $\boldsymbol{E}$ vector has only one component, E_y, tangential to the interface (it is transverse electric or TE radiation) and its $\boldsymbol{B}$ vector has components B_z normal, and B_x tangential, to the interface (this second component being orthogonal to the E_y component and in the plane of incidence). Any

0010-7514/91 $3.00

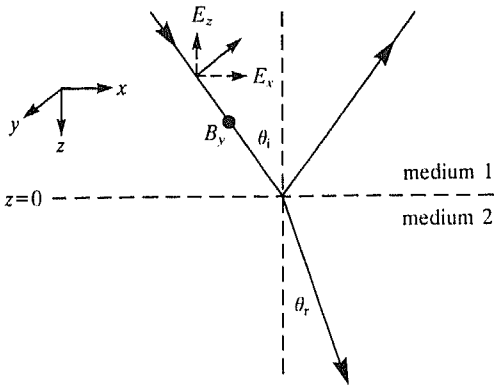

Figure 1. Representation of p-polarized electromagnetic radiation incident upon a planar interface between two media at an angle of incidence θ_i.

linearly polarized radiation may be readily represented by a sum of the above two cases.

Now consider that the second medium is a non-magnetic material, that is at the frequency of the incident radiation the relative permeability is unity. Then as far as the ***B*** part of the electromagnetic oscillation is concerned there is no discontinuity at the interface. In this case, which represents the majority of materials, it is the discontinuity in the dielectric constant which will govern the behaviour of the radiation on encountering the interface. For simplicity we shall throughout this article ignore optical activity, that is the property (the chirality) of a material which allows it to rotate the plane of polarization of an incident photon even if it is propagating along an axis of symmetry in the system.

Photons, with momentum $\hbar k$, when in a medium of refractive index n_1, are regarded as having momentum (strictly pseudomomentum) $\hbar k n_1 = \hbar k_1$ (where $k = 2\pi/\lambda$). If these arrive at a planar interface they may impart momentum in a direction normal to the interface and so there is no need to conserve the normal component of photon momentum, $\hbar k_{z1}$. For the reflected signal, since $\hbar|k_1|$ is conserved, unless the photon frequency is changed, and $\hbar k_{x1}$ is conserved for a smooth planar interface, then it follows that k_{z1} of the reflected signal is simply $-k_{z1}$, the usual law of reflection at a planar interface.

On the other hand inside the second medium the refractive index is n_2 so the radiation has a new wavelength, $\lambda_2 = \lambda/n_2$ and a new wavevector $k_2 = n_2 k$. In this medium the radiation propagates in a new direction, conserving k_x but allowing k_z to change. Now $k_{x1} = k_1 \sin\theta_i$ and $k_{x2} = k_2 \sin\theta_r$ where θ_r is the angle of refraction. Since the tangential momentum component is conserved, $k_{x1} = k_{x2}$ and $n_1 \sin\theta_i = n_2 \sin\theta_r$ which is Snell's law (resulting from the translational invariance of the system parallel to the interface).

Let us examine in detail an important limit of Snell's law. Suppose that the radiation is incident from a high index medium, $n_1 = \sqrt{\varepsilon_1}$, on to a low index medium, $n_2 = \sqrt{\varepsilon_2}$, (where ε_1, ε_2 are the relative permittivities) with $n_2 < n_1$. Then Snell's law, the conservation of in-surface-plane momentum condition, gives

$$\sqrt{\varepsilon_2}\sin\theta_2 = \sqrt{\varepsilon_1}\sin\theta_1. \tag{1}$$

Since the greatest in-surface-plane component available in medium 2 is when $\theta_2 = 90°$, there is a limiting angle of incidence, θ_c, given by

$$\sin\theta_c = \sqrt{\varepsilon_2}/\sqrt{\varepsilon_1}, \tag{2}$$

beyond which, for radiation incident from medium 1, there can be no propagating wave in medium 2. This limiting angle is called the critical angle. Radiation incident beyond the critical angle has more momentum along the surface plane than can be supported by medium 2. For such radiation incident from medium 1 the oscillating ***E*** field will cause the charges in medium 1, including those at the 1–2 interface, to oscillate. Thus even though the radiation is now totally reflected at the interface there are oscillating charges here which have associated radiation fields penetrating into medium 2. These fields cannot propagate, as we saw above, instead they are spatially decaying fields (evanescent) which oscillate in time, at the same frequency as the incident radiation, decaying in amplitude in medium 2 in a direction normal to the interface. At the critical angle the decay length is infinte but this falls rapidly to the order of the wavelength of light as the angle of incidence is further increased. This evanescent field for radiation incident beyond the critical angle is useful for coupling radiation to surface plasmons as we shall see later.

For the moment let us return to the boundary conditions on the ***E*** and ***B*** field components of our incident radiation. Since there is no boundary orthogonal to E_x this component is conserved across the boundary. However this is not the case for E_z, the normal component of ***E***. It is the normal component of ***D***, D_z, which is continuous (there is no free charge) and E_z is forced to change if ε is changed since $D_z = \varepsilon_1\varepsilon_0 E_{z1} = \varepsilon_2\varepsilon_0 E_{z2}$. This discontinuity in E_z results in polarization changes at the interface.

From these simple considerations it is obvious that while s-polarized incident radiation will not normally cause the creation of charge at a planar interface, p-polarized radiation will automatically create time-dependent polarization charge at the interface.

Suppose now we consider one of the two materials to be a metal. A metal may be regarded as a good conductor of electricity and heat and a reflector of radiation. This is a rather loose definition of a metal which relates to

the ability of the 'free' electrons in the metal to respond to the externally imposed fields. If the electrons are free (still of course constrained inside the metal) then they are able to respond with no scattering to the incident radiation giving an ideal metallic response. Such a material, which complete excludes $\boldsymbol{E}$ fields from itself, that is $\boldsymbol{E} = 0$ everywhere inside the metal, must therefore have $\varepsilon = \pm\infty$. An ideal metal in which the electrons respond perfectly to the applied external field, therefore cancelling it, is the limit $\varepsilon \to -\infty$. Such a material of course does not exist, for the free electrons inside a metal cannot respond infinitely quickly to an imposed oscillating field. The electrons have a finite mass and they suffer scattering with lattice vibrations (phonons), defects and the surface.

This means that as we increase the frequency of the incident radiation the free electrons progressively find it harder to respond. Ultimately at high enough frequencies, low enough wavelengths, the metal becomes transparent and behaves more like a dielectric.

3. More detailed theory

From this simplistic treatment of the free electrons in a metal it is easy to show that there is a limiting frequency, the plasma frequency, (for many metals in the ultra-violet) above which the metal is no longer metallic. In this article we shall concern ourselves only with frequencies below this limit, that is with long enough wavelengths so that ε is largely real and negative. As mentioned for real metals there is resistive scattering and hence damping of the oscillations created by the incident $\boldsymbol{E}$ field. This damping causes an imaginary component ε_i to ε. Before, however, concerning ourselves with the added complexity of ε_i let us examine the implications of having a dielectric with positive ε_r adjacent to the metal with negative ε_r.

Because of the requirement of the normal $\boldsymbol{E}$ fields to create surface charges we need only consider p-polarized electromagnetic waves. Further whatever form the surface wave takes it has to satisfy the electromagnetic wave equation in both media. If we take the x–y plane to be the interface plane and the positive z half space as medium 2, then for wave propagation in the x direction only, we have

$$\boldsymbol{E}_1 = (E_{x1}, 0, E_{z1}) \exp[i(k_x x - \omega t)] \exp(ik_{z1} z), \quad (3a)$$

$$\boldsymbol{H}_1 = (0, H_{y1}, 0) \exp[i(k_x x - \omega t)] \exp(ik_{z1} z), \quad (3b)$$

$$\boldsymbol{E}_2 = (E_{x2}, 0, E_{z2}) \exp[i(k_x x - \omega t)] \exp(ik_{z2} z) \quad (3c)$$

$$\boldsymbol{H}_2 = (0, H_{y2}, 0) \exp[i(k_x x - \omega t)] \exp(ik_{z2} z). \quad (3d)$$

If we now apply Maxwell's equation $\nabla \cdot \boldsymbol{E} = 0$ we find

$$E_{z1} = -E_{x1} \frac{k_x}{k_{z1}}, \quad (4a)$$

$$E_{z2} = -E_{x2} \frac{k_x}{k_{z2}}. \quad (4b)$$

Then to find the relationship between H_y and E_x we use Maxwell's equation $\nabla \wedge \boldsymbol{E} = -\mu \dfrac{\partial \boldsymbol{H}}{\partial t}$ (Faraday's law of electromagnetic induction) which with $\mu = \mu_0$ gives the following relationships between the field components, the permittivities and the normal component of the wavevectors in the two media:

$$H_{y1} = \omega E_{x1} \varepsilon_1 \varepsilon_0 / k_{z1}, \quad (5a)$$

$$H_{y2} = \omega E_{x2} \varepsilon_2 \varepsilon_0 / k_{z2}. \quad (5b)$$

Finally we need to apply the boundary conditions at $z = 0$. We know tangential $\boldsymbol{H}$ is continuous and so is tangential $\boldsymbol{E}$, thus $H_{y1} = H_{y2}$ and $E_{x1} = E_{x2}$ leading to the following simple relationship between the relative permittivities and the normal components of the wavevectors in both media:

$$\frac{\varepsilon_1}{k_{z1}} = \frac{\varepsilon_2}{k_{z2}}. \quad (6)$$

Also we have

$$k_{z1} = -i(k_x^2 - \varepsilon_1 k^2)^{1/2}, \quad \text{requiring } k_x^2 > \varepsilon_1 k^2 \quad (7a)$$

and

$$k_{z2} = i(k_x^2 - \varepsilon_2 k^2)^{1/2}, \quad \text{requiring } k_x^2 > \varepsilon_2 k^2, \quad (7b)$$

where $k = \omega/c$. If the wave is truly a trapped surface wave with exponential decays into both media then we need $ik_{z1} > 0$ and $ik_{z2} < 0$. Thus both k_zs are imaginary with opposite signs and so ε_1 and ε_2 are of opposite sign. The first condition, in the dielectric characterised by ε_1, tells us the surface mode wavevector k_x is greater than the maximum photon wavevector available in the dielectric, $\sqrt{\varepsilon_1} k$. The second condition, for the metal, is automatically satisfied with ε_2 negative.

We may substitute expressions (7) for k_{z1} and k_{z2} into (6) to give

$$k_x = k\left(\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}\right)^{1/2}. \quad (8)$$

And we then see for k_x to be real, the requirement for a propagating mode, with ε_2 negative, is that $|\varepsilon_2| > \varepsilon_1$.

Thus we now have satisfied Maxwell's equations and boundary conditions to give a trapped surface wave, with real k_x and appropriate k_z, provided $|\varepsilon_2| > \varepsilon_1$ and $\varepsilon_2 < 0$. Following the above analysis with purely real ε values leads to a surface wave having purely real k_x *which is*

larger than $\sqrt{\varepsilon_1}k$ the maximum value for the medium 1. It is also clear that this surface plasmon resonance is infinitely sharp and has an infinite propagation length.

As mentioned, for real metals there is resistive scattering and hence damping of the oscillations created by the incident $\boldsymbol{E}$ field. This damping causes an imaginary component to ε, ε_i. Then with, $\varepsilon_2 = \varepsilon_{2r} + i\varepsilon_{2i}$, we have

$$k_x = k\left[\frac{\varepsilon_1(\varepsilon_{2r} + i\varepsilon_{2i})}{\varepsilon_1 + \varepsilon_{2r} + i\varepsilon_{2i}}\right]^{1/2}, \tag{9}$$

which for $k_x = k_{xr} + ik_{xi}$ gives, provided $|k_{xi}| \ll k_{xr}$, with $|\varepsilon_{2r}| \gg \varepsilon_1$ and ε_{2i}

$$k_{xr} \sim k\varepsilon_1^{1/2}\left(1 - \frac{\varepsilon_1}{2\varepsilon_{2r}}\right), \tag{10a}$$

and

$$k_{xi} = \frac{1}{2}\, k\, \frac{\varepsilon_{2i}\varepsilon_1^{3/2}}{\varepsilon_{2r}^2} \tag{10b}$$

Hence we find the shift in wavevector, Δk_{xr}, of this surface plasmon resonance from the critical value, $\varepsilon_1^{1/2}k$, is given by

$$\Delta k_{xr} = k_{xr} - \varepsilon_1^{1/2}k \simeq -\frac{1}{2}\, k\, \frac{\varepsilon_1^{3/2}}{\varepsilon_{2r}}. \tag{11}$$

Thus the shift is inversely proportional to $|\varepsilon_{2r}|$ while the width of the resonance, which of course is proportional to k_{xi}, is proportional to ε_{2i} and inversely proportional to ε_{2r}^2. We therefore see that while at first sight it may appear beneficial to use small ε_{2i} to give a sharp resonance, this idea has to be balanced with the requirement that we need a large negative value of ε_{2r}. Indeed if we examine a range of metals it is clear that while ε_{2i} is generally smallest in the visible part of the spectrum becoming larger as we move to infra-red wavelengths there is an even more rapid increase in $|\varepsilon_{2r}|$. In figure 2 we illustrate the dependence of both the real and imaginary parts of the relative permittivities of silver and aluminium from the ultra-violet to the infra-red. This shows that both parameters increase in magnitude with wavelength. However, the width of the surface plasmon resonance is, remember, dictated by $\varepsilon_i/\varepsilon_r^2$ and since ε_r changes faster than ε_i there is, almost without exception, a narrowing of the resonance and consequential increase in observability as the wavelength increases. In figure 3 this ratio is shown for several metals over the visible and near infra-red region of the spectrum. A ratio of the order of 0.2 is the limit of sensible observability for a surface plasmon resonance. This leads to the general conclusion that while only a few metals such as Ag, Au, Al and Cu support a sharp surface plasmon resonance in the visible many more metals support a sharp resonance in the near infra-red. This is illustrated for nickel and platinum in figure 3.

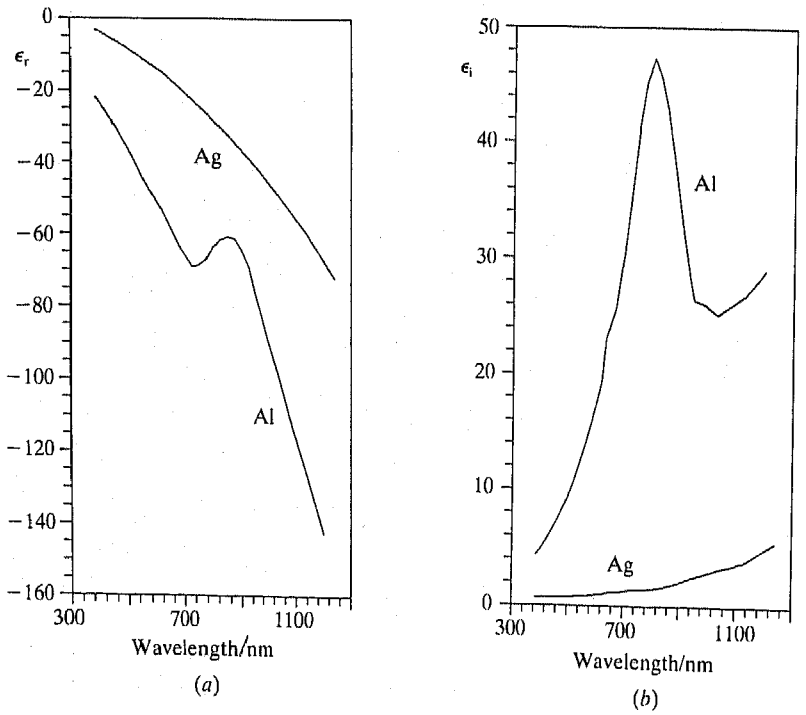

Figure 2. Wavelength dependence of (*a*) the real, ε_r, and (*b*) the imaginary, ε_i, parts of the relative permittivity ($\varepsilon = \varepsilon_r + i\varepsilon_i$) for gold and aluminium. Compiled from data in references [1] and [2].

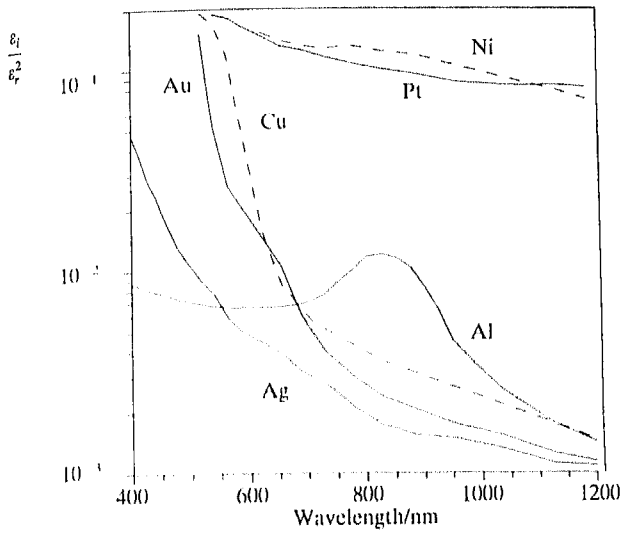

Figure 3. This shows how the surface plasmon resonance width ($\propto \varepsilon_i/\varepsilon_r^2$) varies with wavelength. Many other metals can support sharp resonances as the wavelength of the incident radiation is increased. Compiled from data in references [1] and [2].

4. Coupling to the surface

Before moving on to discuss some experimental results we need finally to examine how best to couple radiation to the surface plasmon resonance given that we have clearly established that its momentum is beyond that available in the dielectric medium adjacent to the metal.

Recall that, for our original two-dielectric system, beyond the critical angle of incidence there will be an evanescent field in the second half space. This evanescent field does not propagate in the z-direction, but it has momentum in the x-direction of $n_1\hbar k \sin\theta_i$. It is obvious that since $\sin\theta_i > \sin\theta_c (= n_2/n_1)$, then $n_1\hbar k \sin\theta_i > n_2\hbar k$. Hence we have an enhancement of the x-component of momentum in the second dielectric half space, above the limit value of $n_2\hbar k$ for a propagating wave.

This enhancement of momentum given by $n_1(\sin\theta_i - \sin\theta_c)\hbar k$ may be used to couple radiation to a surface plasmon provided it is possible to place the metal/dielectric interface which supports the surface plasmon close enough to the totally internally reflecting interface. An obvious geometry to consider is that shown in figure 4(*a*). This is conventionally called the Otto geometry, after Otto who first demonstrated this coupling technique in 1968[3]. An air gap (or a spacer of low index) less than a few radiation wavelengths thick (for visible $< 2\ \mu$m) provides the evanescent tunnel barrier across which the radiation couples, from the totally internally reflecting situation, to excite the surface plasmon at the air (dielectric) metal interface. By varying the angle of incidence of the p-polarized

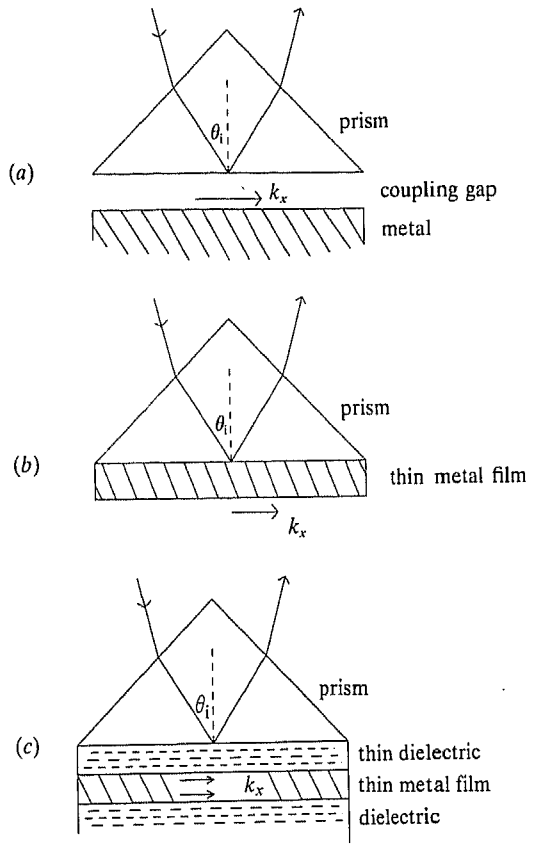

Figure 4. Geometries used to couple photons into a surface mode: (*a*) Otto, (*b*) Kretschmann–Raether, and (*c*) mixed hybrid arrangement.

radiation at the prism/dielectric interface we vary the momentum in the x-direction and this allows for simple tuning through the resonance. The form of the reflectivity curve for gold and silver at 632·8 nm is shown in figure 5, where we also show for comparison that for s-polarized light which is, of course, not capable of creating the surface plasmon. The position of the minimum of the resonance which is a measure of the surface plasmon momentum, is no longer dictated simply by the dielecric/metal boundary for it is additionally perturbed by the presence of the coupling prism. Likewise the linewidth, which is a measure of damping is also perturbed by the presence of the prism. As the coupling gap is increased so the perturbation by the prism diminishes and the resonance moves to the position corresponding to the two media surface plasmon and it also narrows. Of course in this process, illustrated for gold in figure 6, the resonance progressively shallows. If we wish to achieve 100% coupling then for visible radiation the gap has to be of the order of 0.5 μm which for an air gap demands extreme care in sample fabrication. This constraint is not so severe if we choose instead to work in the infra-red

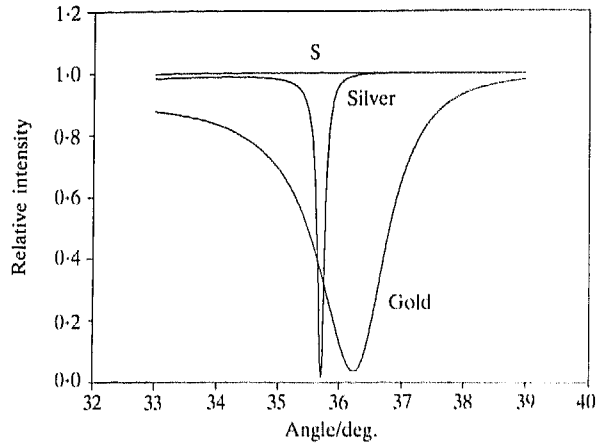

Figure 5. Form of the reflectivity curve for p-polarised and s-polarised radiation ($\lambda = 632{\cdot}8$ nm) from thick gold and silver films with a sapphire prism ($n = 1{\cdot}766$). Here the coupling gap is 0·5 μm for gold and 1·0 μm for silver.

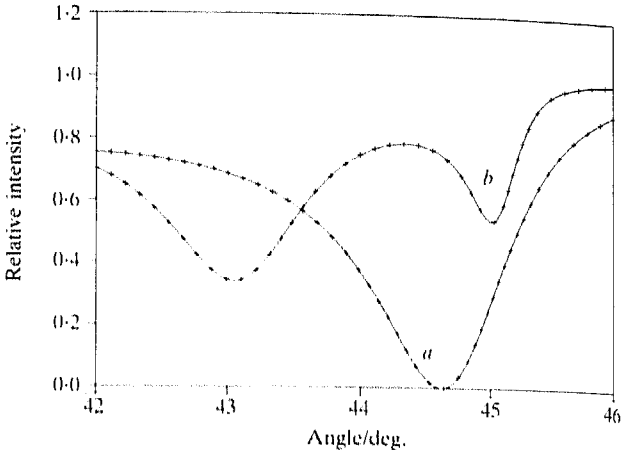

Figure 7. Reflectivity curves obtained from a palladium film on a sapphire prism ($n = 1{\cdot}699$) using 3·391 μm radiation, the surface plasmon occurs at about 45°. Coupling gap for (*a*) is 4·45 μm and for (*b*) is 9·0 μm.

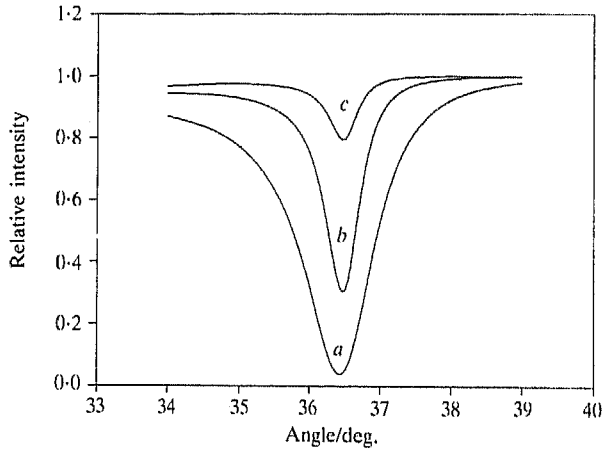

Figure 6. Variation of the surface plasmon resonance (at $\lambda = 632{\cdot}8$ nm) in gold for coupling gaps of (*a*) 0·5 μm, (*b*) 0·75 μm and (*c*) 1·0 μm. The strength of the resonance diminishes rapidly with increasing gap.

region of the spectrum, yet here surprisingly little experimental work has been conducted. We illustrate in figure 7 results for palladium in this region of the spectrum.

The spacing problem created by the Otto geometry may be addressed in quite a different manner by using, instead of an air gap, an evaporated dielectric spacing layer (or perhaps a spun polymer). This of course gives a non-adjustable gap but at least the fabrication is simple for now it is only necessary to evaporate, on top of the appropriate thickness dielectric, a thick metal layer to give the resonance. This particular procedure may indeed be very beneficial in the study of protected interfaces as for example in the case of magnesium or aluminium which normally oxidize[4,5]. Of course there is no simple manner in which it may be changed to use as a sensor or to optimize coupling at other wavelengths. It is for this reason and also because of the small air gap required for coupling in the visible that the Otto geometry has received rather limited attention over the years.

Fortunately this has not severely impeded progress in the area of the optical excitation of surface plasmons. This is because there is an alternative and much simpler geometry. Rather than use a dielectric spacer, Kretschmann and Raether[6] realized that the metal itself could be used as the evanescent tunnel barrier provided it was thin enough to allow radiation to penetrate to the other side. All that is now needed is a prism with a thin coating of some suitable metal. This is illustrated in figure 4(*b*). It is an easy matter to deposit a thin film (< 50 nm) of a metal such as silver or gold on to a prism and create a suitably smooth film which may support a very strong surface plasmon resonance. A typical result for silver in this geometry is shown in figure 8. The continuous line in this figure is, as in figure 7, the fit obtained using simple Fresnel reflectivity theory for a 2-interface system. For the Otto geometry the real and imaginary parts of the metal permittivity and the air gap thickness are unknown variables in the fitting procedure, while for the Kretschmann–Raether geometry the first two parameters are unknown plus the thickness of the metal film. By carefully recording data and comparing with theoretically generated curves it is possible to obtain useful information as regards the dielectric response of the metal which supports the surface plasmon. (In table 1 we list reported ε values for gold and silver using the Kretschmann–Raether technique).

Table 1 Values for the permittivity of gold and silver obtained using surface plasmon excitation in the Kretschmann–Raether geometry.

Wavelength (nm)	Silver			Gold		
	ε_r	ε_i	Ref.	ε_r	ε_i	Ref.
400	−4	0·30	7	−0·3	6·5	8
450	−6·6	0·31	7	−0·8	5·72	8
	−6·4	0·39	9			
500	−9·4	0·37	7	−2·3	3·45	8
	−8·0	0·40	9	−2·0	4·12	7
550	−12·4	0·45	7	−6·0	2·03	8
	−11·4	0·50	9	−5·5	1·98	7
600	−15·6	0·59	7	−9·5	1·62	8
	−14·7	0·60	9	−8·9	1·59	7
650	−19·0	0·74	7	−12·9	1·35	8
	−18·0	0·71	9	−12·1	1·42	7
700	−22·4	0·91	7	−16·8	1·32	8
	−21·3	0·81	9	−15·4	1·43	7
750	—	—		−20·6	1·57	8
	—	—		−18·6	1·62	7
800	—	—		−24·8	2·0	8
	—	—		−22·3	2·0	7

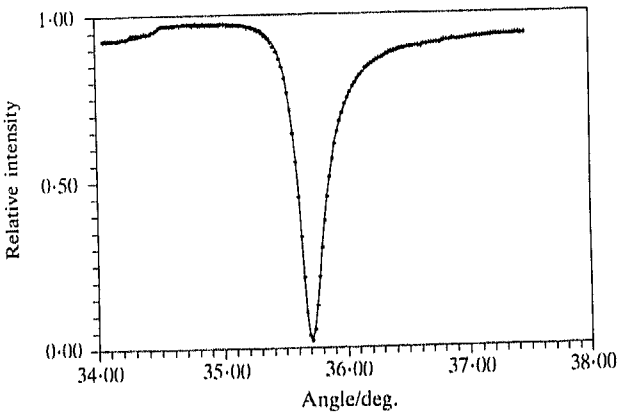

Figure 8. Surface plasmon resonance for a silver film using the Kretschmann–Raether geometry. Note that in this case the critical angle is clearly visible at ~34·5°. The solid line shows the quality of fit which can be obtained to Fresnel theory.

5. Experimental studies

These two types of attenuated total reflection arrangements have formed the basis of most of the studies of optically excited surface plasmons over the past twenty years although more intricate arrangements have also been devised. For example there are possible hybrid geometries in which a thin metal film is deposited on to a dielectric spacer layer as illustrated in figure 4(*c*). This gives an Otto type plasmon on the first (dielectric spacer/metal) interface and a Kretschmann–Raether type plasmon on the second. If now we add a final overcoating of dielectric with the same constants as the first (both of course lower than the prism) then the two surface modes have (in the absence of the prism) the same wavevector. This means that they couple together to give two coupled modes, one which is symmetric in surface charge and the other which is antisymmetric. The first of these modes has very weak electric fields in the metal and is the so-called long range surface plasmon (after Sarid[10]) while the second is labelled the short range surface plasmon since it has large fields in the metal and is therefore strongly attenuated through Joule heating. One could choose to examine even more elaborate multi-layered structures which support more complex coupled modes but this illustrates no particularly new physics and provides little extra potential for device development.

If we choose to have a non-planar interface then we are not necessarily restricted to the prism-coupled geometry. One possible technique, a somewhat unsatisfactory one, is to study a deliberately roughened interface. If we Fourier analyse the roughness there is likely to be a component that supplies the extra momentum needed to couple radiation directly to the surface plasmon. While this may yield a broad band response it is not a very satisfactory interface on which to perform carefully controlled scientific experiments. A better and more systematic approach is to use a diffraction grating with a well specified sinusoidal surface having known wavelength and groove depth. The grooves in the grating surface break the translational invariance of the interface and allow k_x of the outgoing wave to be different from that of the incoming wave. Conservation of momentum now gives in the x direction

$$k_x(\text{outgoing}) = k_x(\text{incoming}) \pm NG, \qquad (12)$$

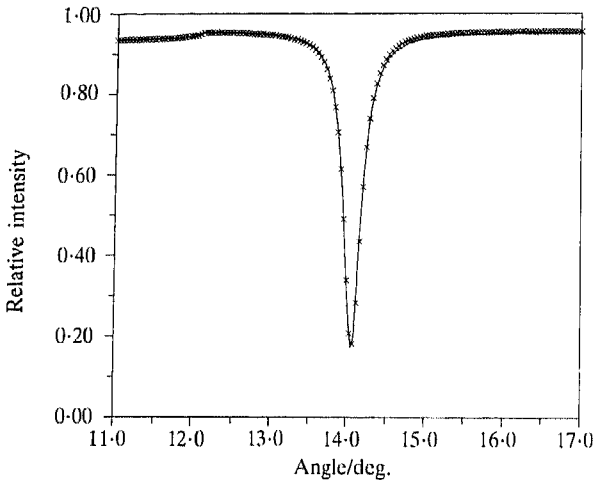

Figure 9. Fitted experimental surface plasmon data (λ = 632·8 nm), obtained from a silver coated grating of pitch 800·8 nm and depth 24·5 nm. The fitted silver film permittivity is $\varepsilon = -15{\cdot}98 + i0{\cdot}72$.

where $G = 2\pi/\lambda_g$, λ_g being the grating wavelength and N an integer. If the grating is relatively shallow (depth $< \lambda_g$) then k_{SP} on the grating surface will be little changed from k_{SP} on a planar surface. Thus all we need do to excite the surface plasmon on the grating surface is to satisfy the equation

$$k \sin\theta = k_{SP} \pm NG. \tag{13}$$

This then allows *direct* excitation of the surface plasmon from the dielectric half-space without imposing constraints upon film thickness or dielectric spacer thickness. However now the coupling strength is dictated by the groove depth and this is not as readily controlled as the air gap or the metal film thickness. Typical data for coupling radiation to a surface plasmon on a silver coated grating is given in figure 9. The smooth curve is a theoretical fit which is no longer trivial to generate. Because the interface is corrugated, simple planar interface Fresnel equations are no longer usable and a much more elaborate model is needed using a Fourier expansion description of the interface. In the results shown here we have used Chandezon's approach[11] where the sinusoidal surface is transformed into a new frame in which it is flat and in which all radiation fields are expressed in this new frame.

One entirely new aspect of surface plasmon excitation using grating coupling which is only just beginning to emerge is associated with rotating the grating so that the grooves are no longer perpendicular to the plane of incidence[12]. This breaks the symmetry of the system and has some very exciting implications for the use of these surface resonances in sensors.

In the twisted geometry shown in figure 10(*a*) the momentum conserving equation is now a two-dimensional vector equation of the form

$$k \sin\theta\hat{x} = \boldsymbol{k}_{SP} N\boldsymbol{G} \tag{14}$$

This is illustrated in figure 10(*b*) for the situation where $|G| < k_{SP}$ or since $|k_{SP}| \sim |1{\cdot}05k|$ then $\lambda_g > \lambda_o$. Now we note that $\boldsymbol{k}_{SP}$ is no longer collinear with $\boldsymbol{G}$ and so the surface plasmon $\boldsymbol{E}$ fields are no longer just in the plane of incidence since in propagating across the grooves a tilted component exists on rising up the side or dropping down the other side of a peak, which cannot be in the incident plane. Thus we have created 's' character in the radiation field associated with the surface plasmon. Indeed with the angle of twist, ϕ equal to 90° we have no 'p' coupling to a surface plasmon only 's'. We illustrate this fully in figure 11 where we show coupling to a silver surface plasmon for both p and s radiation at various angles of twist of the grating. The primary implication of this observation is that now we have p to

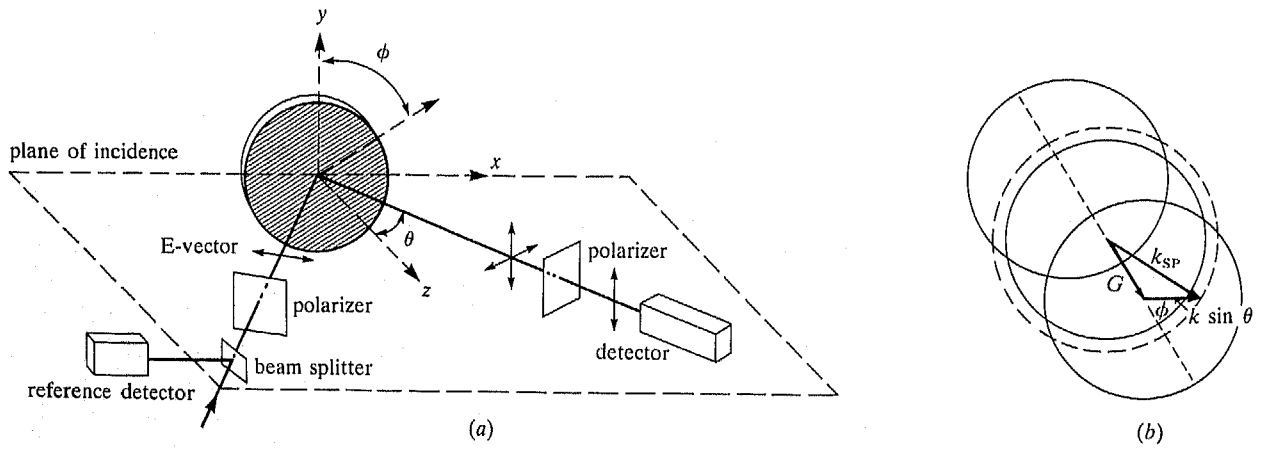

Figure 10. (*a*) Schematic diagram of grating coupling using a twisted geometry. (*b*) Vector representation of momentum conservation in grating coupling. The full circles are the maximum k values obtainable for the zero and ± 1 diffraction orders.

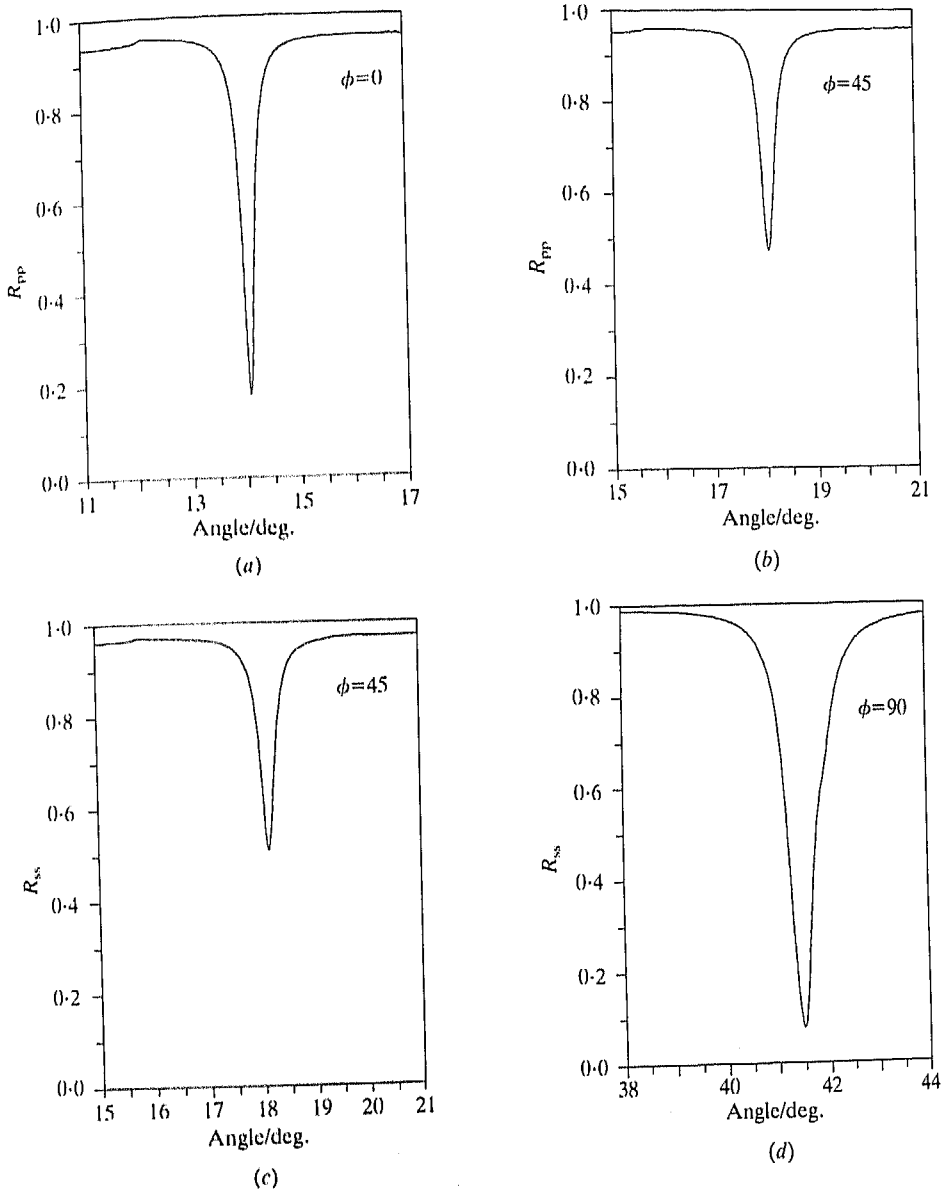

Figure 11. Surface plasmon resonance for silver coated grating. R_{pp} is the relative reflectivity for p-polarized incident and reflected radiation and R_{ss} the same for s-polarized radiation. (*a*) R_{pp} for $\phi = 0$, (b) R_{pp} for $\phi = 45°$, (*c*) R_{ss} for $\phi = 45°$ and (*d*) R_{ss} for $\phi = 90°$. For $\phi = 45$ both polarizations can couple equally well to the plasmon.

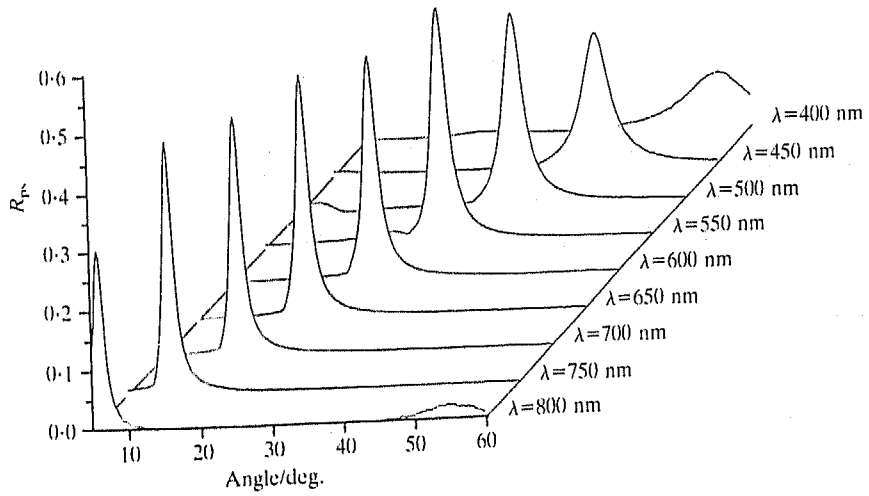

Figure 12. p to s conversion as a function of the incident wavelength. With this grating pitch on increasing the wavelength the maximum conversion occurs at smaller angles of incidence.

s conversion via the excitation of the surface plasmon hence we may record with suitable polarizers a surface plasmon resonance maximum. This has only very recently been examinated in detail and compared with theory by developing Chandezon's model further for the broken symmetry situation. The maximum p to s conversion occurs with a twist angle of 45° and with the maximum groove depth. Results for maximum p to s conversion using a surface plasmon on a silver coated grating are illustrated in figure 12. These observations are particularly exciting because prism coupling, at least for isotropic media, can never provide the required symmetry breaking, and this therefore opens up a new range of potential devices.

6. Applications

This then brings us to examine the possible exploitation of this novel surface mode in devices. Perhaps we need first address the question of why it is of interest at all. The basic answer is that the momentum of the surface plasmon, which is readily monitored by coupling incident radiation to it, is easily changed by thin layers of material deposited on the metal surface[13] or by small changes in the dielectric constant of the material adjacent to the metal.

One of the simplest studies that may be undertaken is that of the chemical contamination of the metal supporting the surface plasmon. For example it is simple, by monitoring the surface plasmon resonance of silver in the Kretschmann–Raether geometry to observe the progressive growth of silver sulphide on exposure to the atmosphere. Kovacs[14] performed this experiment and by monitoring the shift in resonance angle over many days found for his particular environment that 2 nm of silver sulphide formed after about thirty days of exposure. This is sufficiently slow to allow most experiments with silver to be conducted in air without undue concern over this overlayer formation. On the other hand if the same type of experiment were conducted with a thin film of aluminium the initial exposure to air results in the rapid formation of a relatively stable aluminium oxide layer some 3 to 4 nm thick[15]. While these studies are intrinsically interesting there is more interest in deliberately overcoating the surface plasmon supporting metal, the active medium, with other types of layers. For example studies have been performed with organic multilayers deposited using the Langmuir–Blodgett technique. By careful control of the deposition of these layers, well-defined stepped structures may be fabricated. For these the angular dependent reflectivity for the Kretschmann–Raether geometry shows a surface plasmon resonance which steps progressively to a higher angle as the layer thickness is stepped integrally[16]. This then allows determination of the assumed isotropic relative permittivities and thicknesses of the organic overlayers. A range of different experiments have also been performed with inorganic overlayers, again to study their dielectric properties. In a sense this is relatively unexciting; of more interest are studies where changes in the overlayer occur. For example the progressive laser-induced desorption of organic films, predeposited onto the active metal layer, may be studied [17], or the inverse, the condensation through the attractive potential between a volatile organic and a metal film may be readily observed[18]. In the latter case this leads, with careful experimentation, to the determination of the bonding potential of organics on to the metal layer[19]. A variant of monitoring changes in the thickness of overlayers is the study, again by measuring the shift in the resonance position, of the effective relative permittivity of the overlayer as a consequence of exposure to gas[20]. This technique, with the appropriate overlayers has applications in optical gas sensing. Extending this idea, solutions rather than gases may be placed adjacent to the metal film and then changes in the region of the surface plasmon decay length may once again be readily monitored. This opens up potential for immunoassay using antigen protein films on the metal layer which bind to specific antibodies in solution[21]. As the antibodies bind to the antigen so the surface plasmon resonance is shifted in angle and a direct optical measure of antigen concentration can be obtained. In the same context, of fluids adjacent to the active metal surface there are a large range of studies, from optical examination of electrochemical processes[22], to studies of liquid crystal alignment[23,24] and more complex processes such as the kinetics of adsoption of block copolymers from solution[25].

Currently there is much interest in the use of optical excitation of surface plasmons in these and related areas of physics, physical chemistry and biophysics. Added to this is a perceived potential for device application in areas other than just sensors. For instance, fibre polarizers with very high extinction ratios have already been fabricated in which thin metal layers provide the necessary surface plasmon resonance absorption thus destroying one polarization component[26]. There is also interest in the use of surface plasmon excitation in scanning surface microscopy[27]. Small variations in overlayers on an active metal film are easily converted into large differences in reflectivity by setting the system at the angle of the surface plasmon excitation and scanning across the sample.

Another area with potential for the use of surface plasmons is non-linear optics. The optical excitation of this surface-travelling wave resonance results in strong enhancement of the optical field at the surface supporting

the surface plasmon. This gives substantial potential for the study of strong optical field effects[28] as well as strong non-linear effects due to local heating[29,30]. One of the most striking experiments in this area is that by Nunzi and Ricard[31] who studied optical phase conjugation through pulsed laser excitation of thermal gratings in silver films. It would appear there is much more work to be undertaken in this area, particularly with non-centrosymmetric overlayers.

Finally we would like to highlight just one area of study where surface plasmon excitation is uniquely valuable. Recently Lajendijk and co-workers[32,33] have used pulsed excitation of surface plasmons to probe the dynamics of thin metal films. With some superb experimental techniques they have been able to examine electron relaxation rates as well as phonon decay times. Thus while there is much interest in the use of optical excitation of surface plasmons in devices it is also beginning to find a use in more fundamental physics, an area which we hope to see expand.

The authors acknowledge the support of SERC and the University of Exeter for this work and thank Dr S Elston, Mr G Bryan-Brown and Mr P Vukusic for providing the experimental data used to produce figures 8, 9, 11 and 12.

References

[1] Lynch, D. W., and Hunter, W. R., 1985, in *Handbook of Optical Constants of Solids* edited by E. D. Pallik, (Academic Press), 275.
[2] Smith, D. Y., Stiles, E., and Inokuti, M., 1985, *ibid*, 369.
[3] Otto, A., 1968, *Z. Phys*, **216**, 398.
[4] Tillin, M. D., and Sambles, J. R., 1988, *Thin solid Films*, **167**, 73.
[5] Tillin, M. D., and Sambles, J. R., 1989, *Thin solid Films*, **172**, 27.
[6] Kretschmann, E., and Raether, H., 1968, *Z. Naturf* A, **23**, 2135.
[7] Schröder, U., 1981, *Surf. Sci.*, **102**, 118.
[8] Innes, R. A., and Sambles, J. R., 1987, *J. Phys. F: Metal. Phys.*, **17**, 277.
[9] Lawrence, C., 1991, Private communication.
[10] Sarid, D., 1981, *Phys. Rev. Lett.*, **47**, 1927.
[11] Chandezon, J., Dupuis, M. T., Cornet, G., and Maystre, D., 1982, *J. opt. Soc. Amer.*, **72**, 839.
[12] Bryan-Brown, G. P., Sambles, J. R., and Hutley, M. C., 1990., *J. mod. Optics*, **37**, 1227.
[13] Pockrand, I., 1978, *Surf. Sci.*, **72**, 577.
[14] Kovacs, G. J., 1978, *Surf. Sci.*, **78**, L245.
[15] Dumas, P., Dubarry-Barbe, J. P. Rivière, D., Levy, Y., and Corset, J., 1983, *J. de Phys*, **44 C10**-205.
[16] Brown, C. A., Burns, F. C., Knoll, W., Swalen, J. D., and Fischer, A., 1983, *J. phys. Chem.*, **87**, 3616.
[17] Puderbach, S., Herminghaus, S., and Leiderer, P., 1988, *Phys. Lett.* A, **130**, 401.
[18] Pollard, J. D., and Sambles, J. R., 1987, *Optics Commun.*, **64**, 529.
[19] Sambles, J. R., Pollard, J. D., and Bradberry, G. W., 1990, *J. mod. Optics*, **37**, 841.
[20] Nylander, C., Liedberg, B., and Lind, T., 1982, *Sensors and Actuators*, **3**, 79.
[21] Fontana, E., Pantell, R. H., and Strober, S., 1990, *Appl. Optics*, **29**, 4694.
[22] Tadjeddine, A., Abraham, M., and Hadjadi, A., 1986, *J. electroanal. Chem.*, **204**, 229.
[23] Welford, K. R., and Sambles, J. R., 1987, *Appl. Phys. Lett.*, **50**, 871.
[24] Elston, S. J., and Sambles, J. R., 1989, *Appl. Phys. Lett.*, **55**, 1621.
[25] Tassin, J. F., Siemens, K. L., Tang, W. T., Hadziioannou, G., Swalen, J. D., and Smith, B. A., 1989, *J. phys. Chem.*, **93**, 2106.
[26] Johnston, W., Stewart, G., Culshaw, B., and Hart, T., 1988, *Electron. Lett.*, **24**, 866.
[27] Yeatman, E. M., and Ash, E. A., 1988, *SPIE Scanning Microscopy Technologies and Applications*, **897**, 100.
[28] Schmidlin, E. M., and Simon, H. J., 1988, *Applied Optics*, **28**, 3323.
[29] Martinot, P., Laval, S., and Koster, A., 1984, *J. de Phys.*, **45**, 597.
[30] Innes, R. A., and Sambles, J. R., 1989, *J. Phys: Condensed Matter*, **1**, 6231.
[31] Nunzi, J. M., and Ricard, D., 1984, *Appl. Phys.*, **B35**, 209.
[32] van Exter, M., and Lagendijk, A., 1988, *Phys. Rev. Lett.*, **60**, 49.
[33] Grooneveld, R. H. M., Sprik, R., and Lagendijk, A., 1990, *Phys. Rev. Lett.*, **64**, 784.

Roy Sambles obtained his PhD in electron microscopy from Imperial College in 1970. In 1972 he moved to a lecturership in the Physics Department at Exeter where he now leads a research team of 20 studying the physics of thin films. He has sat on a number of SERC committees and has recently been appointed to a personal Professorship in Experimental Physics.

Geoff Bradberry obtained his PhD in semiconductor physics from the University of Leicester in 1964 and was appointed lecturer in Department of Physics at Exeter in 1966. Subsequently his research covered various topics in solid state physics. Recently he has worked on Brillouin scattering in liquid crystals and the optical excitation of surface plasmons and other guided modes.

Fuzi Yang obtained his batchelors degree from Qinghua University in Beijing in 1968. In 1981 he was appointed Professor in the Applied Physics Department of Jiaotong University. Since November 1988 he has been a visiting researcher at the Department of Physics at Exeter. Recently he has been awarded a PhD by the University. At present he is studying the optics of thin liquid crystal layers.