The Optical Response of Short-Pitch Surface-Relief Gratings

Submitted by

Ian Richard Hooper

To the University of Exeter as a thesis for the degree of Doctor of Philosophy in Physics, May 2002

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

..... (signature)

Abstract

Controlling the optical response of surfaces by patterning the interface with a surface-relief periodic corrugation has been established for various applications for many years. The primary use concerns the creation of diffracted orders produced by the periodic corrugation. In addition, this surface patterning may allow the excitation of electromagnetic modes of the systems (either waveguide or surface modes). In this thesis original investigations of several corrugated surfaces are presented, some of which possess strongly coupled surface modes, and others which show only weak diffraction effects. Studies are presented of the influence of surface profile, depth, and phase on the optical responses.

There are two main sections to the work presented. In the first section the optical response of two-interface corrugated dielectric and metal slabs is investigated, with particular attention paid to the effect of altering the phase between the corrugations on the two surfaces. For a non-zero phase difference the dielectric structures exhibit effective 'blazing' behaviour in the reflected orders. Also, an extraordinary distribution of energy between the diffracted orders is found to occur for certain grating profiles, with almost all of the energy of the incident light channelled into the +1 and -1 transmitted diffracted orders. On the two-interface metal structures SPPs can be excited on both interfaces, or coupled SPPs may be excited. It will be shown that, by changing the phase difference between the corrugations, the optical response of the system can be precisely controlled, with it being possible to either enhance or suppress the transmission through the structure.

In the second section the effect of grating profile on the optical response of short-pitch single-interface metal gratings is explored. For deep gratings SPPs are excited even in the zero-order region of the spectrum, which was previously not thought possible. These SPP modes are broad and extremely flat-banded for gratings consisting of narrow grooves, but are relatively sharp features, with complex dispersion, if the grating peaks are narrow. These SPP modes are also observed experimentally for the first time.

Contents

List of Figures

inside the light lines and may be coupled to by incident radiation.

Figure 2.3.1.3 The dispersion curve of a SPP on a grating structure in the ω-k_z plane with $k_x = 0$. The SPP dispersion curve centred at the origin is shown (always outside of the lightline) as is the SPP dispersion curve scattered from $+k_g$. The form of this scattered SPP dispersion curve arises from taking a slice through the SPP 'horn' shape centred at $k_x = k_g$.

Figure 2.3.1.4 The thickly drawn circle of radius k_0 centred upon the origin describes the maximum wavevector available to a photon in the plane of the grating surface. At a slightly greater radius than this is the SPP circle (dotted line) corresponding to the dispersion curve described in the text. Both of these circles have been scattered by k_g, which results in the two circles occurring at the reciprocal lattice points. The angle ψ is the propagation angle of the SPP with respect to the k_x direction.

Figure 2.3.2.1 The zero-order TM reflectivity as a function of frequency for a 1μm pitch 30nm amplitude sinusoidal silver grating (with the dielectric function of the silver described with a Drude model with $\omega_p = 1.32\text{x}10^{16}\text{s}^{-1}$, and $\tau = 1.4\text{x}10^{14}$s), for normal incidence and $\phi = 0°$. Three features are evident; at $f = 0.36\text{x}10^{15}$Hz a SPP minimum due to the first order SPP, at $f = 0.37\text{x}10^{15}$Hz a pseudo critical edge due to the 1st diffracted order becoming evanescent, and at $f = 0.735\text{x}10^{15}$Hz a small reflectivity maximum due to the second order SPP.

Figure 2.3.2.2 The zero-order TM reflectivity for the same system as in figure 3.3.2.1, but with changing amplitude of the grating.

Figure 2.3.3.1 A sketch of the standing waves in a dielectric stack. The shaded regions correspond to a dielectric with a higher refractive index

than the unshaded regions. The standing wave with field extrema concentrated in the high index medium has a lower energy than the standing wave with field extrema in the low index medium

Figure 2.3.3.2 Numerically modelled zero-order TM reflectivity of a grating structure comprising two Fourier components. The first harmonic component produces a band-gap at the intersection of the +1 and-1 SPP branches (at normal incidence). The parameters used in the modelling were λ_g = 634nm, a_0 = 5nm, a_1 = 2nm and the permittivity of the metal is ε_r = -17.5 and ε_i = 0.7. The two components of the grating profile are in phase with each other.

Figure 2.3.3.3 The electric field and surface charge distribution for the two standing wave solutions on the $2k_g$ component of the grating profile. The field lines in the lower sketch are more distorted, and therefore this is the higher energy mode.

Figure 2.3.3.4 Numerical modelled zero-order TM reflectivity plots showing the influence of the phase difference between the fundamental and first harmonic components of the grating profile. The grating parameters are otherwise the same as for figure 4.3.3.2.

Figure 2.3.3.5 The k_g and $2k_g$ components of a distorted sinusoidal grating with relative phase between the two components of –90°, +90° and 0°.

Figure 2.3.4.1 Numerically modelled polarisation conserved and polarisation converted reflectivities for TM and TE polarised light of frequency 0.474 x 10^{15}Hz incident upon a 1μm pitch, 40nm amplitude silver grating as a function of k_x and k_z. a) Polarisation conserved for TM polarised incident light, b) Polarisation converted for TM incident light, c) Polarisation conserved for TE polarised incident light, and d) Polarisation converted for TE polarised incident light.

Figure 3.4.2.1 The multi-layer system.

Figure 3.4.4.1 The multi-shape grating system.

Figure 3.4.4.2 Orientation of the field components at the j+1st interface.

Figure 3.5.1 A comparison between the codes based upon the ISS method and the Chandezon method. TM reflectivity from a 7.5 nm amplitude, 600nm pitch, sinusoidal silver grating as a function of frequency (normal incidence, classical mount). a) the specularly reflected order, and b) the +1 diffracted order.

Figure 3.5.2 A comparison between the codes based upon the ISS method and the Chandezon method. TM reflectivity from a 7.5 nm amplitude 600nm pitch sinusoidal silver grating as a function of the in-plane wavevector for light of frequency.

Figure 4.2.1.1 The system under consideration in this section. TM polarised light incident upon an interface between two dielectrics with dielectric functions ε_1 for the upper (incident) medium, and ε_2 for the lower (transmitted) medium. The light is normally incident in the classical mount (the azimuthal angle is 0°).

Figure 4.2.1.2 The optical response of an air / SiO_2 grating interface. The four orders presented are: a) the zeroth reflected, b) the zeroth transmitted, c) the +1 reflected diffracted, and d) the +1 transmitted diffracted.

Figure 4.2.1.3 Time averaged $|\mathbf{H}_z|$ component of the fields for the air / SiO_2 interface used for figure 3.2.1.2 for six different frequencies: a) f = 0.4 x 10^{15}Hz, b) f = 0.514 x 10^{15}Hz, c) f = 0.6 x 10^{15}Hz, d) f = 0.7 x 10^{15}Hz, e) f = 0.75 x 10^{15}Hz, and f) f = 0.9 x 10^{15}Hz.

Figure 4.2.1.4 The optical response of a SiO_2 / Air grating interface, with the same parameters as for the air / SiO_2 interface considered above. The four orders presented are: a) the zeroth reflected, b) the zeroth transmitted, c) the +1 reflected diffracted, and d) the +1 transmitted diffracted.

Figure 4.2.1.5 Time averaged $|\mathbf{H}_z|$ component of the fields for the SiO_2 / Air interface used for figure 3.2.1.4 for six different frequencies: a) f = 0.4 x 10^{15}Hz, b) f = 0.514 x 10^{15}Hz, c) f = 0.6 x 10^{15}Hz, d) f = 0.7 x 10^{15}Hz, e) f = 0.75 x 10^{15}Hz, and f) f = 0.9 x 10^{15}Hz.

Figure 4.3.1.1 The planar dielectric waveguide.

Figure 4.3.2.1 A schematic of the system under consideration.

Figure 4.3.2.2 The optical response of a conformally corrugated thin SiO_2 slab, with air as the bounding media (λ_g = 400nm, a = 25nm), as a function of the frequency of incident TM polarised light and of the slab thickness. a) the zeroth order reflectivity, b) the zeroth order transmisivitty, c) the +1 order reflectivity, and d) the +1 order transmissivity.

Figure 4.3.2.3 Schematic showing that the transmitted diffracted order from the top interface propagates further than the zeroth transmitted order from the top interface.

Figure 4.3.2.4 Time averaged $|\mathbf{H}_z|$ component of the fields for the system shown in figure 3.3.2.1 for f = 1.0 x 10^{15}Hz with different slab thickness. a) d = 25nm, b) d = 75nm, c) d = 150nm, and d) d = 300nm.

Figure 4.3.2.5 Time averaged $|\mathbf{H}_z|$ component of the fields for the system shown in figure 3.3.2.1 for a) f = 0.732 x 10^{15} Hz (diffractive in SiO_2 only), and b) f = 0.492 x 10^{15}Hz (non-diffractive in both media).

Figure 4.3.2.6 The magnitude and phase of the amplitude coefficient for the system shown in figure 3.3.2.1, obtained from the ISS method by considering the two diffraction processes separately (as described in the text), for a) diffraction from the top interface, b) diffraction from the bottom interface, and c) the total transmitted fields for the system obtained by combining a) and b).

Figure 4.3.3.1 A schematic of the system under consideration.

Figure 4.3.3.2 The optical response of a 60nm thick dielectric slab corrugated on both surfaces with gratings of λ_g = 400nm, and amplitude of 25nm, as a function of the phase between the corrugations on the two interfaces, and of the frequency of the incident light. a) the reflected zeroth order, b) the transmitted zeroth order, c) the reflected +1 diffracted order, d) the transmitted +1 diffracted order, e) the reflected −1 diffracted order, and f) the transmitted −1 diffracted order.

Figure 4.3.3.3 The intensities of the zeroth reflected and transmitted orders, and of the 1^{st} diffracted transmitted and reflected orders for a 60nm thick dielectric slab corrugated on both surfaces with antisymmetric sinusoidal corrugations on each surface of 400nm pitch and 25nm amplitude.

Figure 4.3.3.4 The intensities of the various orders from the anti-phase two interface system as a function of frequency and slab thickness. a) The zeroth order reflected, b) the zeroth order transmitted, c) the +1 diffracted reflected, and d) the +1 diffracted transmitted.

Figure 4.3.3.5 The intensities of the transmitted orders as a function of frequency and grating amplitude for a) a 60nm thick dielectric slab, and b) a 150nm thick dielectric slab.

Figure 4.3.3.6 Time averaged $|\mathbf{H}_z|$ field profiles for succesive maxima and minima in the 1st order transmitted difracted order for an incident frequency of 1.75 x 10^{15}Hz. a) a = 18.35nm, b) a=37.2nm, c) a = 53.75nm. A red line has been drawn onto the plots at the tops of the bottom surface corrugation in order to demonstrate the way the field maxima fit within the grating grooves.

Figure 4.3.3.7 The intensities as a function of frequency for a two interface corrugated system of thickness 60nm and amplitude 25nm for a) the reflected ±1 diffracted orders, and b) the transmitted ±1 diffracted orders.

Table 4.3.3.1 The phase effects which need to be considered when determining the total ±1 reflected and transmitted diffracted order fields from a thin corrugated dielectric slab (ignoring multiple reflections).

Figure 5.2.1 Reflectivity of TM polarised light for normal incidence in the classical mount as a function of frequency, for a 40nm thick silver film corrugated with a 600nm pitch sinusoidal grating of 10nm amplitude. a) corrugated on both sides, b) corrugated on the bottom surface only, and c) corrugated on the top surface only.

Figure 5.2.2 Transmissivity and absorption of TM polarised light for normal incidence in the classical mount for a thin silver film with the parameters used to obtain figure 5.2.1(a).

Figure 5.2.3 Schematic showing the origin of the possible transmitted orders due to diffraction from a thin metal slab corrugated on both surfaces. The dotted lines indicate the evanescently decaying fields due to diffraction at the top interface.

Figure 5.2.4 The magnitude and phase of the 1st order transmitted diffracted

field around the SPP excitation frequency from a conformally corrugated structure for a) diffraction from the top surface, b) diffraction from the bottom surface, and c) the total 1st order transmitted diffracted field.

Figure 5.2.5 Results from a dual corrugated silver film as a function of frequency and the phase between the corrugation on the two interfaces. a) zero-order reflection, b) zero-order transmission (log scale), and c) the magnitude of the first diffracted order (log scale).

Figure 5.2.6 The zeroth order transmission, and absorption in the metal slab, for the same structures as used for figure 5.2.1.

Figure 5.3.1.1 The charge distributions for a) the short range SPP, and b) the long range SPP.

Figure 5.3.1.2 Instantaneous electric field profiles (parallel and perpendicular components) for the LRSPP (a) and SRSPP (b)

Figure 5.3.2.1 The optical response as a function of frequency and slab thickness for $k_x = 0$ on a conformal, sinusoidally corrugated silver slab of 10nm amplitude and 400nm pitch, in the classical mount.

Figure 5.3.2.2 The optical response of the same system described for figure 5.3.2.1 as a function of frequency and in-plane wavevector. a) The dispersion of the modes obtained from the scattering matrices, b) the reflection, c) the transmission, and d) the absorption for the system.

Figure 5.3.2.3 The optical response of the same system as figure 5.3.2.1, but with an additional $2k_g$ component in the grating profile description of 5nm amplitude, as a function of frequency and slab thickness. a) the position of the modes obtained from the scattering matrices (z

axis units are arbitrary), b) the reflection and c) the transmission of the system.

Figure 5.3.2.4 The dispersion curves of the SPPs of the system for a conformally corrugated thin metal slab with a pitch of 400nm, and a grating profile described by a 10nm amplitude k_g component, and a 5nm $2k_g$ component, with the phase of the $2k_g$ component being +90° out of phase with the k_g component. a) d = 70nm, b) d = 50nm, and c) d = 30nm (z axis units are arbitrary).

Figure 5.3.3.1 The zeroth order reflectivity (a), transmissivity (b), and absorption of the system (c), for a 30nm thick silver slab corrugated on both surfaces with a 400nm pitch, 10nm amplitude, sinusoid as a function of frequency and phase between the corrugations on the two surfaces.

Figure 5.3.3.2 The reflectivity, transmissivity, and absorption of the system, as a function of frequency and in-plane wavevector for the case where the two gratings are in anti-phase with each other for the same grating parameters as the system described previously.

Figure 5.3.3.3 The band structure (a), reflectivity (b), and transmissivity (c) of the zeroth order, for normally incident light as a function of frequency and slab thickness for a metal slab corrugated on both sides when the corrugations on the two surfaces are in anti-phase (the gratings have a 10nm amplitude k_g component, a 5nm $2k_g$ component, and a pitch of 400nm).

Figure 6.2.1 Reflectivity as a function of frequency and in-plane wavevector for TM polarised light incident on a 300nm deep, 50nm wide Gaussian-grooved, 200nm pitch silver grating held at a zero degree azimuthal angle. The bands are very flat for a large range of incident wavevectors

Figure 6.3.1 An example of the type of grating structure investigated. The profile is made up of a series of Gaussian grooves so that the depth, width and pitch of the grating may be altered independently, and so that in deep structures the SPP modes in neighbouring grooves will be largely isolated from each other.

Figure 6.3.2 Dispersion curves for 200nm pitch, 50nm wide, Gaussian-grooved silver gratings for $k_z = 0$, with a) d = 10nm, b) d = 25nm, c) d = 50nm, and d) d = 75nm.

Figure 6.3.3 $|\mathbf{H}_z|$ for the second order SPP resonance for a) d = 10nm, f = 1.36 x 10^{15}Hz, b) d = 25nm, f = 1.3 x 10^{15}Hz, and c) d = 50nm, f = 1.19 x 10^{15}Hz.

Figure 6.3.4 Dispersion curves for 200nm pitch, 50nm wide, Gaussian-grooved silver gratings for $k_z = 0$, with a) d = 100nm, b) d = 150nm, and c) d = 300nm.

Figure 6.3.5 $|\mathbf{H}_z|$ for the first four SCSPPs on a 300nm deep, 50nm wide, Gaussian grooved grating with $k_z = 0$. a) f = 0.19 x 10^{15}Hz, b) f = 0.44 x 10^{15}Hz, c) f = 0.66 x 10^{15}Hz, and d) f = 0.85 x 10^{15}Hz.

Figure 6.3.6 $|\mathbf{H}_z|$ for the second order SCSPP resonance on a 300nm deep, 50nm wide, Gaussian grooved grating with $2k_x / k_g = 0.5$ (f = 0.46x10^{15} Hz).

Figure 6.3.7 Mode frequency as a function of grating depth for 200nm pitch, 50nm wide, Gaussian-grooved gratings at $k_z = 0$. a) $2k_x / k_g = 0.0$, b) $2k_x / k_g = 0.1$, c) $2k_x / k_g = 0.5$, and d) $2k_x / k_g = 1.0$. Dotted lines are light-lines.

Figure 6.4.1 The dispersion curve for a 200nm pitch, 300nm deep, 50nm wide

a) f = 1.22×10^{15} Hz and $2k_x / k_g = 0$, and b) f = 1.16×10^{15} Hz and $2k_x / k_g = 1.0$.

Figure 7.2.6 Time averaged $|\mathbf{H}_z|$ component of the fields of the SPP mode excited on a 200nm pitch silver grating consisting of a series of 350nm high and 40nm wide (FWHM) Gaussian peaks, with radiation of f = 1.26×10^{15} Hz incident at a 0° azimuthal angle and at $2k_x / k_g = 1.0$.

Figure 7.2.7 TM reflectivity as a function of frequency and in-plane wavevector for a 200nm pitch silver grating, consisting of a series of 40nm wide (FWHM) Gaussian peaks, and peak heights of a) 10nm, b) 75nm, c) 150nm, and d) 250nm.

Figure 7.3.1 The zeroth order reflectivity from a 200nm pitch silver grating consisting of a series of 10nm deep 40nm wide Gaussian peaks, oriented at a 90° azimuthal angle, as a function of frequency and in-plane wavevector. a) TM polarised radiation, and b) TE polarised radiation.

Figure 7.3.2 The TM reflectivity from 200nm pitch silver gratings consisting of a series 40nm wide Gaussian peaks, oriented at a 90° azimuthal angle, as a function of frequency and in-plane wavevector. a) d = 50nm, b) d = 100nm, c) d = 200nm, and d) d = 300nm.

Figure 7.3.3 The $|\mathbf{H}_z|$ component of the fields (the component along the grooves) of the four lowest energy modes on the structure described for figure 7.3.2(c), and with $k_z = 0.4$. a) f = 0.721 x 10^{15}Hz, b) f = 0.833 x 10^{15} Hz, c) f = 1.008 x 10^{15}Hz, and d) f = 1.148 x 10^{15}Hz.

Figure 7.3.4 The $|\mathbf{H}_z|$ component of the fields in the x-z plane through the

middle field maximum shown in figure 7.3.3(c) (y = 110nm).

Figure 7.3.5 The band structure for a 200nm pitch silver grating consisting of a series of 40nm wide 50nm deep Gaussian peaks. The first and second order SPPs are shown as well as the light line and the SPP dispersion curve centred at the origin.

Figure 7.3.6 Zeroth order TM reflectivities for 200nm pitch silver gratings consisting of a series of 40nm wide Gaussian grooves oriented at a 45° azimuthal angle as a function of frequency and in-plane wavevector. a) 10nm deep polarisation conserved, b) 10nm deep polarisation converted, c) 50nm deep polarisation conserved, d) 50nm deep polarisation converted, e) 100nm deep polarisation conserved, f) 100nm deep polarisation converted, g) 200nm deep polarisation conserved, h) 200nm deep polarisation converted.

Figure 7.4.1 Reflectivity from a 200nm pitch silver grating consisting of a series of 60nm wide Gaussian peaks at a 45° azimuthal angle as a function of frequency and peak height. a) Polarisation conserved, and b) polarisation converted.

Figure 7.4.2 Reflectivity from a 200nm pitch silver grating consisting of a series of 40nm wide Gaussian peaks at a 45° azimuthal angle as a function of frequency and peak height. a) Polarisation conserved, and b) polarisation converted.

Figure 7.4.3 x and z components of the E fields at three points on the plot in figure 7.3.2(a). a) d = 100nm, f = 1.37 x 10^{15}Hz, b) d = 250nm, f = 0.97 x 10^{15}Hz, and c) d = 250nm, f = 0.41 x 10^{15}Hz.

Figure 7.4.4 Reflectivity from a 300nm pitch silver grating consisting of a series of 60nm wide Gaussian peaks at a 45° azimuthal angle. a) Polarisation conserved reflectivity as a function of frequency and

peak height, b) polarisation converted reflectivity as a function of frequency and peak height, and c) reflectivity as a function of wavelength for d=265nm.

Figure 7.4.5 Reflectivity from a 250nm pitch silver grating consisting of a series of 50nm wide Gaussian peaks at a 45° azimuthal angle. Silver permittivity described by a polynomial fitted to experimentally determined values. a) Polarisation conserved reflectivity as a function of frequency and peak height, b) polarisation converted reflectivity as a function of frequency and peak height, and c) reflectivity as a function of wavelength for d=240nm.

Figure 7.4.6 Reflectivity from a 300nm pitch aluminium grating consisting of a series of 232.5nm high, 60nm wide Gaussian peaks at a 45 degree azimuthal angle of 45°.

Figure 8.2.1.1 A schematic of the interferometer system used to manufacture the surface relief gratings.

Figure 8.2.1.2 SEM image of a typical photoresist grating produced by the method described in this section.

Figure 8.2.1.3 A schematic of the Reactive Ion Etcher (RIE) system

Figure 8.2.1.4 SEM image of the etched grating produced by reactive ion etching of a photoresist grating into the SiO_2 substrate.

Figure 8.2.1.5 A schematic of the vacuum deposition equipment used to coat the grating in silver.

Figure 8.2.2.1 A schematic of the system used to measure the wavelength dependent reflectivities of the samples.

Figure 8.3.1.1 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep 170nm pitch silver grating oriented at an azimuthal angle of 0° with TM polarised incident light. The light is incident through the SiO_2 substrate.

Figure 8.3.1.2 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep 170nm pitch silver grating oriented at an azimuthal angle of 0° with TM polarised incident light. The light is incident through a 45° SiO_2 prism which is index matched to the SiO_2 substrate of the sample.

Figure 8.3.1.3 The dispersion of the reflectivity minima obtained from figure 8.3.1.1 and 8.3.1.2. The light line and diffracted order light line are also shown.

Figure 8.3.2.1 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep 170nm pitch silver grating oriented at an azimuthal angle of 90° with TE polarised incident light. The light is incident through the SiO_2 substrate.

Figure 8.3.2.2 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep 170nm pitch silver grating oriented at an azimuthal angle of 90° with TE polarised incident light. The light is incident through a 45° SiO_2 prism which is index matched to the SiO_2 substrate of the sample.

Figure 8.3.2.3 The dispersion of the reflectivity minima obtained from figure 8.3.2.1 and 8.3.2.2

Figure 8.3.3.1 The polarisation conserved and converted wavelength dependent reflectivities for various angles of incidence of a 150 deep 170nm pitch silver grating oriented at an azimuthal angle of 45° with TM polarised incident light. For a) to c) the light is incident through the

SiO_2 substrate, whereas for d) to h) it is incident through a 45° SiO_2 prism which is index matched to the SiO_2 substrate of the sample. a) $\sin\theta = 0.1$, b) $\sin\theta = 0.2$, c) $\sin\theta = 0.3$, d) $\sin\theta = 0.4$, e) $\sin\theta = 0.5$, f) $\sin\theta = 0.6$, g) $\sin\theta = 0.7$, h) $\sin\theta = 0.8$.

Figure 8.3.3.2 The dispersion of the reflectivity minima obtained from the total reflectivity plots of figure 8.3.3.1

Acknowledgements

My time working in the Thin Film Photonics group has been a very enjoyable experience, and this is mainly due to the characters who over the years have made it the place it is. Therefore, I would like to take this opportunity to thank all of my colleagues, as well as some other people in my life, without whom this thesis may have never have seen the light of day.

Firstly I would like to thank my supervisor Roy Sambles, without whom I would never have been doing a PhD in the first place, and who always had words of encouragement even when things weren't exactly going to plan. Also many thanks to Bill Barnes, who was always ready to listen to problems and try to help out, even when he must have occasionally felt like hitting his head against the desk when I knocked at his door. And thanks to Chris Lawrence, my QinetiQ CASE sponsor, for his sense of humour and help with the dreaded QinetiQ bureaucracy, and otherwise leaving me to get on with things.

Next I have to thank all of the members of the group, both past and present, who helped out throughout the last four years. Firstly to the present members of group 20. The senior boys, Pete V and Mike Jory, for maintaining the folk lore of the basement, and especially Mike for demonstrating why the harsh world of business is probably best kept away from, and his enthusiasm for all things Star Wars (Oh, and those quite wonderful diagrams… Where does the sphere go again?!) To Weicao for introducing me to the wonderful world of modelling. To the hair bear duo of Al and Jimmy, the first of which has shown that there is still room in this world for the mullet, and Jimmy whose hair may possibly be the largest free-standing man made structure in the world! Also, thanks to Jimmy for being a most able partner in the sport of Shazza baiting. Talking of which, thanks to Shazz-cakes for not beating us over the head when we got a little too carried away with it! Thanks to Haze for being beside me (quite literally!) throughout the PhD ordeal, Nobu for being an endless source of amusement, and Matt for possibly the dirtiest laugh in existence! Then there are the old boys, Benny and Nate, who introduced me to the wonder of Monday lunchtime footie, where you can occasionally kick a lecturer with complete impunity!

Next comes the members of group 13, who, though I may have had less contact with them, are all integral parts of the ambience of the place. Piers, and Phil (now sadly

departed), who were always willing to help with problems. Mart for his invaluable assistance with the interferometer during the early days, Hobbit for a well rounded sense of humour, and Simon for camping things up in the basement quite unbelievably after his arrival.

Other people from the basement who I haven't mentioned so far, such as Steve, Jon, Andy, Ralph, Fuzi, and Lizhen have all helped at times, and therefore also deserve many thanks.

Then there's Pete Cann. A marvel with gadgets and gizmos, who was always willing to help with the next urgent job (in other words you ask at 4pm for something by the end of the day!), and somehow always managed to produce the goods, and to the highest quality. And also thanks to the other boys in the workshop, because Pete couldn't always be there to bail me out! Jon in stores for our chats about football, as well as giving me the bits and pieces I needed, and Dave Jarvis for innumerable evaporations.

Finally I would like to thank those people who, though they weren't directly involved with the work, have supported me over the last few years. To all of the people I've lived with, especially Boz, Derek, Pip, Stu, and Twinny, who have put up with me being just a tad stressed on occasions, and have cheered me up when it was needed. And finally, though by no means least, to my Mum and Dad for their support over all of the years of my education, and for not being driven mad by another son deciding to do a PhD.

Chapter 1

Introduction

1.1 Aim of Thesis

When this project first started the aim was to experimentally investigate high aspect ratio zero-order metal mono-gratings, zero-order hole arrays (in metal films) and metal island arrays. It was believed that the mono-grating structures would be of interest due to previous work by Sobnack *et al* (1998) and Tan *et al* (1999) in which very flat absorption bands were observed even in the zero-order region of the spectrum, and this work showed that these absorption bands were due to the excitation of Surface Plasmon Polaritons (SPPs). This result was somewhat surprising since it had previously been thought that SPPs could not be excited on zero-order gratings. However, after an extensive effort was made to manufacture such samples it became clear that, with the equipment available, it was not possible to make these structures of a sufficiently high quality, and it was at this point that the direction of the work changed somewhat. Similar problems were encountered when attempts were made to manufacture the hole and island arrays, and due to these problems this area of the project was not pursued.

During the experimental efforts the optical response of the mono-grating structures was modelled extensively for both narrow grooved (similar to the structure previously investigated) and narrow ridged structures, and it became clear that understanding the results obtained was, in fact, a significant piece of work in itself. Therefore, the modelling obtained, and the analysis, make up a large part of the work contained within this thesis, along with a chapter describing the experimental results which were obtained.

The remaining work for this thesis arose due to a paper by Schröter and Heitmann (1999) in which they observed that the SPP on the transmission side of a conformally corrugated thin metal slab could not be excited. During attempts to understand this work it was realised that their explanation of the results was incomplete, and it was in relation to the formulation of a more complete understanding of the optical response of these structures that the remaining work in this thesis was performed. It was noticed that, on these thin metal slabs, the role of the phase between the two corrugations on the two interfaces had a dramatic effect on their optical responses, and

that this subject had never before been investigated. This logically resulted in the question of what effect this phase issue would have on coupled SPPs, and on simple dielectric slab structures. It is the results of this work which makes up the remainder of this thesis.

It is clear that the aims of this thesis are dramatically different to those envisaged at the beginning of this project, but the general intention to increase the understanding of the role of periodic microstructures in the optical response of surfaces has been maintained.

1.2 Outline of Thesis

The work in this thesis can be generally categorised into three main sections. The first section (chapters 2 and 3) deals with the background theory needed to understand the remaining work and gives an extensive discussion on the nature of SPPs, and on the modelling methods used throughout this thesis to calculate the optical responses of the structures investigated. The second section (chapters 4 and 5) investigates the optical response of two-interface grating structures for both metals and dielectrics when both interfaces are corrugated, and the final section (chapters 6, 7 and 8) proceeds to investigate the effect of increasing the corrugation depth on the nature of the SPPs excited at a single air / metal interface. A brief outline of the contents of each chapter is given below.

In chapter 2 the surface plasmon polariton (SPP) is introduced, and its nature on shallow metal gratings is described. This chapter provides the basis for all of the following chapters which involve metal layers, and describes such phenomena as band gaps in the SPP dispersion curves, and polarisation conversion.

In chapter 3 the methods by which the optical response of grating structures may be calculated are described. A brief outline of the various possible methods is given, with detailed descriptions of the two methods used throughout the remainder of the thesis: the iterative series solution, and the differential method of Chandezon *et al*. This chapter is not needed to understand the remaining work, but is presented to describe the methods used extensively throughout this thesis.

In chapter 4 the optical response of shallow dielectric grating structures is described. Initially the case of a single air / dielectric interface is explored, before proceeding to discuss the optical response thin slab dielectric structures which are bounded by air and corrugated on both surfaces. The effect of the phase between the two corrugations is investigated, with several unexpected results occurring.

In chapter 5 the same type of structures as in chapter 4 are explored, but with the slab material consisting of a metal rather than a dielectric. For this reason SPPs may be excited on both surfaces when the thickness of the slab is very small. Two situations are studied. Firstly, the case when the two bounding dielectrics have different refractive indices, so that the SPPs on the two surfaces are excited at different frequencies of the incident light. And secondly, the case where the two bounding dielectrics have identical refractive indices, with the result of coupled SPPs. Both conformal geometries, and the effect of the phase difference between the corrugations on the two surfaces, are studied for both cases.

In chapter 6 there is a change of focus from shallow two interface structures, concentrating upon the effect on the optical response of single interface metal gratings of making the corrugation depth very large. In this chapter the case of a grating consisting of a series of narrow Gaussian grooves is examined, and it is shown that there are a family of flat-banded standing wave SPP resonances, which can be excited even in the zero-order region of the spectrum (in other words, when the grating is non diffractive). The mechanism by which these SPP resonances occur is described.

In chapter 7 once again deep, single interface metal gratings are explored, but in this chapter the gratings consist of a series of narrow Gaussian peaks. The dispersion of the modes with changing in-plane wavevector is shown to be considerably different to those of the Gaussian grooved structures described in chapter 6, and the reason for this is described.

In chapter 8 an experimental study of deep single interface metal gratings is reported. Reflectivity minima are recorded which are probably due to the excitation of

the flat banded SPP resonances described in chapter 6. The methods of fabrication and data acquisition are described before the results for varying wavelength and in-plane wavevector.

Finally, in chapter 9 the general conclusions from the thesis, possible applications for the structures investigated, and ideas for future studies are presented.

Chapter 2

Surface Plasmon Polaritons on Shallow Gratings

2.1 Introduction

The study of surface plasmon polaritons (SPPs) has been of interest for the last century. In fact it was Wood [1902] who, investigating the optical properties of ruled metallic gratings, first observed anomalies (a series of light and dark bands) in their wavelength dependent reflectivities (some of which were, many years later, attributed to the excitation of SPPs). These anomalies consisted of high and low regions of intensity in the spectrum obtained when illuminated with a continuous light source. At the time there was no theory which could account for these 'Wood's anomalies'.

It was not until Rayleigh [1907] published his Dynamical Theory of Diffraction Gratings that the bright bands were explained. He showed that these were pseudo critical edges caused by a diffracted order becoming evanescent and its energy becoming redistributed among the other propagating orders. However, Strong [1935] demonstrated that the angular position of the reflection minima depended upon the metal of which the grating was made, and therefore they could not be solely due to the geometry of the system as were the pseudo-critical edges. It was not until Fano [1941] published a seminal paper on the subject that the nature of the dark bands was explained. He described these dark bands in terms of the diffracted order becoming evanescent and producing a pair of 'superficial' waves which were unable to leave the surface.

Due to further understanding of metals, and the development of the plasma concept for describing the free electrons, Ferrell [1958] predicted that a beam of electrons incident upon a metal film would emit radiation at the plasma frequency. Steinmann [1960] observed this to be the case, and Ritchie and Eldridge [1961] showed that the emitted radiation had characteristic energies of $h\omega_p / 2\pi$ and $h\omega_p / 2\pi\sqrt{2}$, where ω_p is the plasma frequency, and $\omega_p / \sqrt{2}$ is the surface plasma frequency which will be described later in the chapter. It was this work which lead to the superficial waves described by Fano becoming known as surface plasmon polaritons.

In this chapter we shall introduce the SPP for a planar surface and briefly describe its properties. We shall then explain how a shallow grating may be used to couple light to this surface mode producing resonant absorption. We shall also explain two of the phenomena of SPPs on shallow gratings which shall be needed in order to understand some of the work later in this thesis: band gaps, and polarisation conversion.

Some of the plots contained within this chapter are obtained using the modelling methods described in the next chapter.

2.2 The Surface Plasmon Polariton on a Planar Surface

Fano [1941] investigated the Wood's anomalies by considering the case in which there is a glass overlayer on a metal substrate through which a trapped EM wave is travelling by total internal reflection. He produced an equation which connected the thickness of the glass overlayer to the normal and tangential components of the wavevector of the trapped wave solutions. He found that if the thickness of the glass is reduced to the point where it vanishes there is only one solution to this equation. This solution depends upon the trapped EM wave being TM polarised, and upon the real part of the permittivity of the substrate being negative (i.e. a metal) giving a large concentration of surface charge from the inversion of the direction of the normal component of the electric field. From this start the study of SPPs has progressed to the point where the phenomenon is now well understood.

The Surface Plasmon (SP) is a longitudinal oscillation of the surface charge at the interface between a metal and a dielectric. Typically it is excited with EM radiation and when this is coupled to the surface charge oscillation the resultant surface wave is known as a Surface Plasmon Polariton (SPP), and has mixed transverse and longitudinal character. This mode is a 'trapped' surface wave which decays exponentially away from the interface into both bounding media. If TM polarised light is incident upon an interface between a metal and a dielectric with a non-zero incident angle there is an **E** field component normal to the surface. This field is not continuous across the interface, and therefore a surface charge is induced upon it. For TE polarised light, however, there is no component of **E** normal to the surface for any incident angle, and no charge is induced at the surface. Therefore, since the SPP is the result of the surface charges, it may, on flat surfaces, only be excited by TM polarised light. In this section we shall

derive the dispersion relation for the SPP mode on a planar surface, and also discuss its properties.

2.2.1 The Dispersion Relation

We shall now derive the dispersion relation for the SPP surface mode from Maxwell's equations (Raether [1988]), and also show that it is equivalent to the Brewster mode for a metal / dielectric boundary.

For TM polarised light propagating with wavevector $\mathbf{k} = (k_x, k_y, 0)$ the electric and magnetic fields have the forms:

$$\mathbf{E} = [E_x, E_y, 0]\exp\left(i\left(k_x\hat{x} + k_y\hat{y} - \omega t\right)\right) \qquad 2.2.1.1$$

$$\mathbf{H} = [0, 0, H_z]\exp\left(i\left(k_x\hat{x} + k_y\hat{y} - \omega t\right)\right) \qquad 2.2.1.2$$

For light incident upon an interface a reflected and a transmitted beam are produced as in figure 2.2.1.1 (z direction is into the page).

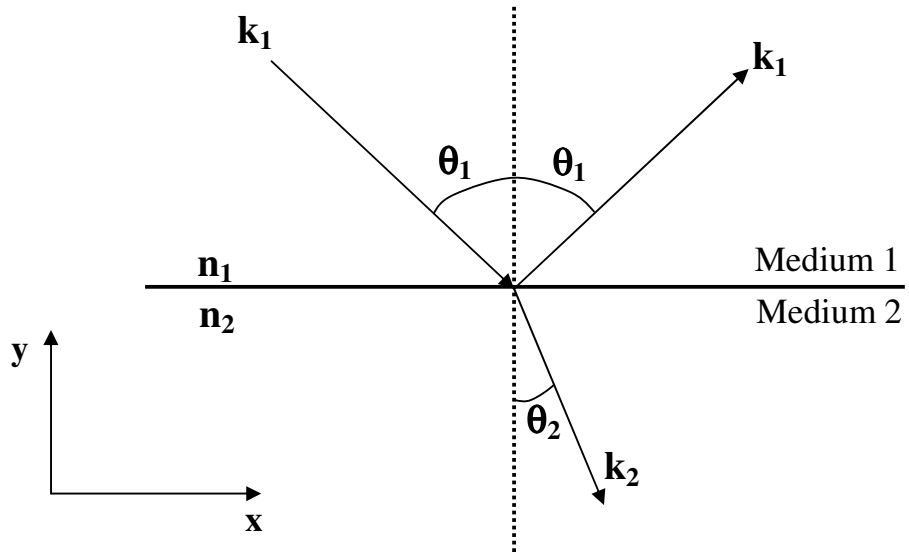

Figure 2.2.1.1 A schematic diagram of light with incident wavevector k_1 incident on an interface bounded by two media described by complex refractive indices n_1 and n_2.

By applying Maxwell's equation:

$$\nabla \times \mathbf{H} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} \qquad 2.2.1.3$$

we obtain

$$H_z = \frac{\varepsilon\omega E_x}{k_y} = -\frac{\varepsilon\omega E_y}{k_x} \qquad 2.2.1.4$$

and by combining equations 2.2.1.1,2 with equation 2.2.1.4 we obtain the following expressions for the incident, reflected and transmitted fields:

$$\mathbf{E}_1^+ = E_{x1}^+\left[1,-\frac{k_x}{k_{y1}},0\right]\exp\left(i\left(k_x\hat{x}+k_{y1}\hat{y}-\omega t\right)\right)$$

$$\mathbf{E}_1^- = E_{x1}^-\left[1,+\frac{k_x}{k_{y1}},0\right]\exp\left(i\left(k_x\hat{x}-k_{y1}\hat{y}-\omega t\right)\right)$$

$$\mathbf{E}_2^+ = E_{x2}^+\left[1,-\frac{k_x}{k_{y2}},0\right]\exp\left(i\left(k_x\hat{x}+k_{y2}\hat{y}-\omega t\right)\right)$$

$$\mathbf{H}_1^+ = E_{x1}^+\left[0,0,+\frac{\omega\varepsilon_1}{k_{y1}}\right]\exp\left(i\left(k_x\hat{x}+k_{y1}\hat{y}-\omega t\right)\right)$$

$$\mathbf{H}_1^- = E_{x1}^-\left[0,0,-\frac{\omega\varepsilon_1}{k_{y1}}\right]\exp\left(i\left(k_x\hat{x}-k_{y1}\hat{y}-\omega t\right)\right)$$

$$\mathbf{H}_2^+ = E_{x2}^+\left[0,0,+\frac{\omega\varepsilon_2}{k_{y2}}\right]\exp\left(i\left(k_x\hat{x}+k_{y2}\hat{y}-\omega t\right)\right) \qquad 2.2.1.5$$

Since the SPP is a trapped surface wave we are able to set either the incident or reflected field to zero (we set the incident field to zero), and we also have the boundary condition at a dielectric / metal interface that the tangential components (x and z directions) of the electric and magnetic fields must be continuous across the interface. Therefore,

$$E_{x1} = E_{x2} \qquad 2.2.1.6$$

and,

$$H_z = -E_{x1}^-\frac{\varepsilon_1\omega}{k_{y1}} = E_{x2}^+\frac{\varepsilon_2\omega}{k_{y2}} \qquad 2.2.1.7$$

By combining equations 2.2.1.6 and 2.2.1.7 we then obtain:

$$\frac{\varepsilon_1}{k_{y1}} = -\frac{\varepsilon_2}{k_{y2}} \qquad 2.2.1.8$$

We may also obtain the y component of the wavevector of the reflected and transmitted fields in terms of the total wavevector and the in-plane wavevector of the fields:

$$k_{yj} = \left(\varepsilon_j k_0^2 - k_x^2\right)^{1/2} \qquad 2.2.1.9$$

where the j subscript corresponds to the medium in which the field is propagating. By substituting equation 2.2.1.9 into equation 2.2.1.8 the SPP dispersion relation is obtained:

$$k_x = k_0\left(\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2}\right)^{1/2} \qquad 2.2.1.10$$

We shall now show that the SPP is, in fact, the manifestation of the Brewster mode on a metal dielectric interface.

Consider light incident upon a planar interface between two media with refractive indices n_1 and n_2 respectively. The tangential component of the incident wavevector must be conserved upon transmission through the interface:

$$k_{x1} = k_{x2} = n_j k_0 \sin\theta_j \qquad 2.2.1.11$$

where the j subscript corresponds to the medium in which the light is propagating. This is a form of Snell's law, which is more frequently written as:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \qquad 2.2.1.12$$

The Brewster angle is the angle at which TM polarised light is totally transmitted through an interface between two dielectrics. In the reflecting medium the plane of oscillation for the charges form an angle θ_2 with the normal to the surface. The proportion of light reflected from the surface depends upon the angle between the direction of propagation of the reflected light and that of the plane of oscillation of the charges. If this angle is 90° then there will be no reflected light since there can be no re-radiation from an oscillator in the direction orthogonal to the plane of oscillation, and therefore all of the light will be transmitted. This occurs at an angle $\theta_B = 90° - \theta_2$ and equation 2.2.1.12 then becomes:

$$n_1 \sin\theta_B = n_2 \sin\left(90 - \theta_B\right) = n_2 \cos\theta_B \qquad 2.2.1.13$$

which can be simplified to the more common form for the Brewster angle:

$$\tan\theta_B = \frac{n_2}{n_1} \qquad 2.2.1.14$$

Equation 2.2.1.13 can be used in order to determine the dispersion relation for the Brewster mode:

$$\sin\theta_B = \frac{n_2}{\left(n_1^2 + n_2^2\right)^{1/2}} = \left(\frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2}\right)^{1/2} \qquad 2.2.1.15$$

and combing this with equation 2.2.1.11 we then obtain:

$$k_x = n_1 k_0 \sin\theta_B = k_0\left(\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1 + \varepsilon_2}\right)^{1/2} \qquad 2.2.1.16$$

Equation 2.2.1.16 is identical to equation 2.2.1.10 and therefore it is clear that the SPP represents the Brewster mode on a metal / dielectric interface.

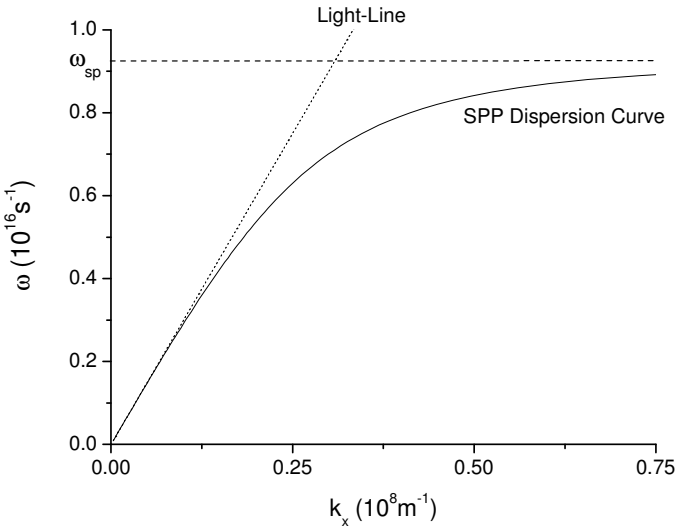

Figure 2.2.1.2 The SPP dispersion curve for a planar silver / vacuum interface using a Drude model to describe the frequency dependent dielectric function of the silver with ω_p = 1.36 x 10^{14}Hz, and τ = 1.4 x 10^{14}Hz. Also shown are the light-line, and the high k_x asymptotic limit of ω_{sp}.

The dispersion relation is shown in figure 2.2.1.2 for an air / silver planar interface with the dielectric function of the silver described by a Drude model of the form:

$$\varepsilon_r = 1 - \frac{\omega_p^2\tau^2}{(1+\omega^2\tau^2)}$$

$$\varepsilon_i = \frac{\omega_p^2\tau}{\omega(1+\omega^2\tau^2)}$$

The SPP dispersion curve is shown, along with the light-line (which corresponds to the maximum in-plane wavevector an incident photon may have for a particular frequency - in other words, the wavevector of a grazing photon as a function of ω).

The SPP dispersion curve has two asymptotic limits. For high frequencies the SPP dispersion curve becomes asymptotic to the surface plasmon frequency at $\omega_{sp} = \omega_p / (1 + \varepsilon_1)^2$ which is the frequency at which the real part of the dielectric function of the silver is equal, but opposite in sign, to the real part of the dielectric function of the dielectric. This limit arises since, from equations 2.2.2.10,16, it is clear that at frequencies above ω_{SP} and below ω_p, k_x is purely imaginary, and above ω_p, k_y (given by $k_y^2 = \varepsilon_1 k_0^2 - k_x^2$) is real indicating that it is no longer a trapped surface wave. In the region of the dispersion curve where it is close to ω_{SP} the SPP is 'plasmon like'. The second asymptote is the low frequency limit where the SPP dispersion curve approaches the light line. At this limit the SPP is like a grazing photon. These two limits will be discussed further in section 2.2.2.

We have assumed here that the incident radiation is incident in the x-y plane. Since the surface is planar the dispersion curve obtained above is valid for radiation incident in any plane since the polarisation of the incident radiation is defined relative to the plane of incidence. Therefore, the full dispersion curve, which takes into account the z component of the wavevector of the incident radiation, has a horn shape, formed by rotating the dispersion curve shown in figure 2.2.1.2 around the y-axis.

It is clear that the SPP dispersion curve is always at higher values of the in-plane wavevector than that obtainable by incident photons. For this reason the SPP may not be excited on a planar dielectric / metal interface without some coupling mechanism to enable the incident photon to gain enough in-plane wavevector for the wavevector matching condition to be satisfied.

Coupling to a SPP may be achieved optically by two main methods: prism coupling, and grating (roughness) coupling. Otto [1968] first used a prism to couple to a SPP on a smooth planar surface by placing a flat metal surface a small distance from the prism. Light is incident in the prism at past the critical angle so that it is totally internally reflected. An evanescent wave is created at the reflecting surface which decays into the dielectric gap between the prism and metal surface. This evanescent wave may have increased wavevector when compared to that available to light in the air gap, and therefore the wavevector matching condition between the exciting fields and the SPP may be satisfied, and the SPP excited. The coupling strength to the SPP depends upon the width of the air gap between the prism and metal surface, with ideal

coupling occurring when the air gap is of the order of the incident radiation wavelength. Due to the fact that this gap is very small for visible radiation it is difficult to use this method in practice.

A second prism coupling technique was developed by Kretschman and Raether [1968]. This method utilises a thin metal film evaporated directly on the reflecting prism surface. Since the refractive index of the metal is lower than that of the prism an evanescent wave is created when the radiation is incident at past the critical angle. This evanescent wave decays through the metal film and excites the SPP on the lower surface of the metal film. For ideal coupling to the SPP the metal film thickness should be of the order of 45nm at visible wavelengths.

The method of interest throughout the remainder of this thesis is that of grating coupling, and this will be described in some detail in section 2.3.

2.2.2 Penetration Depth

When a SPP is excited at the surface between a dielectric and a metal its fields decay exponentially away from the surface into both media. A useful measure of this decay is the penetration length L_y, which is the distance at which the field strength has decayed to 1/e of its maximum value at the surface. In order to obtain this it is necessary to have expressions of the form $E_0 \exp(i(a+ib))$ where the decay in the y direction is due to the [ib] term introduced by the complex part of the wavevector k_y. Therefore it is necessary to obtain an expression for the imaginary part of k_y:

$$k_{yj} = \pm k_0 \left[\varepsilon_j - \left(\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2} \right) \right]^{1/2} = \pm k_0 \left(\frac{\varepsilon_j^2}{\varepsilon_1 + \varepsilon_2} \right)^{1/2} \qquad 2.2.2.1$$

where the subscript j denotes the medium in which k_y is being determined ($j = 1,2$).

The upper (incident) medium is a dielectric with complex dielectric function such that $\varepsilon_{1r} > 0$ and $\varepsilon_{1i} = 0$, and the lower medium has complex dielectric function such that $\varepsilon_{2r} < 0$, and $\varepsilon_{2i} > 0$. If medium 2 is a good metal, ($|\varepsilon_{2r}| >> 1$ and $|\varepsilon_{2r}| >> \varepsilon_{2i}$) equation 2.2.2.1 becomes:

$$k_{yj} = \pm k_0 \left(\frac{\varepsilon_{jr}^2}{\varepsilon_{2r}} \right)^{1/2} \qquad 2.2.2.2$$

From equation 2.2.2.2 it is clear, for the condition on the dielectric functions described above, that k_{yj} is purely imaginary, and therefore the penetration length L_y is given by:

$$L_{yj} = \frac{1}{k_{yj}} = \frac{\lambda_0}{2\pi}\left(\left|\frac{\varepsilon_{2r}}{\varepsilon_{jr}^2}\right|\right)^{1/2} \qquad 2.2.2.3$$

Figure 2.2.2.1 shows the penetration depths for a planar silver / vacuum interface with the frequency dependent dielectric function described by a Drude model as described in section 2.2.1 with a plasma frequency of 1.32 x 10^{16}Hz, and a relaxation time of 1.4 x 10^{-14}s. The frequency range shown is for a region where the condition that $|\varepsilon_{2r}| >> \varepsilon_{2i}$, used to derive the penetration depth in equation 2.2.2.3, is obeyed for the Drude model.

(a)

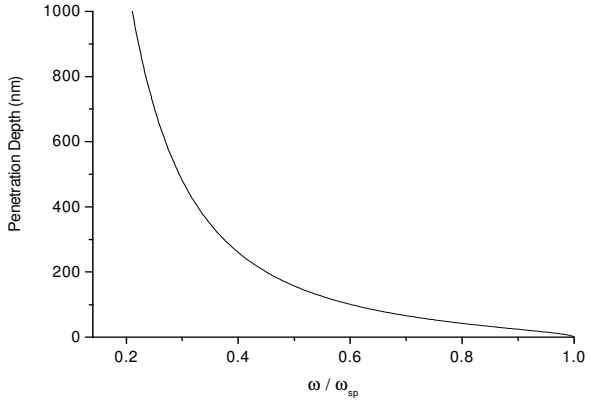

Figure 2.2.2.1 The penetration depth of the SPP fields into the vacuum for a planar silver / vacuum system. The frequency dependent dielectric function for the silver is modelled using the Drude model with ω_p =1.36 x 10^{16}Hz, and τ = 1.4 x 10^{14}s.

(b)

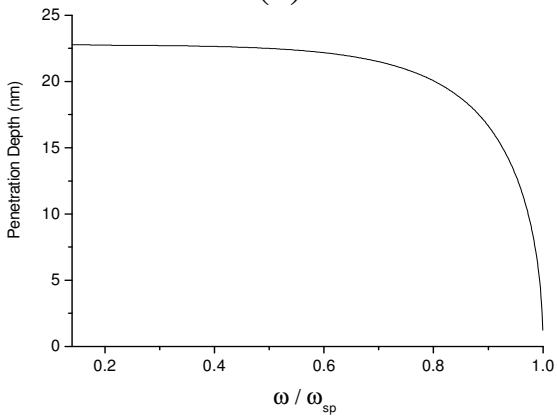

Figure 2.2.2.2 The penetration depth of the SPP fields into the silver for a planar silver / vacuum system. The frequency dependent dielectric function for the silver is modelled using the Drude model with ω_p =1.36 x 10^{16}Hz, and τ = 1.4 x 10^{14}s.

It is clear from equation 2.2.2.4 and from figure 2.2.2.1 that in the limit $\varepsilon_{2r} \to -\infty$, in other words a perfect metal, the decay depth in the dielectric becomes infinite. This agrees with the discussion in section 2.2.1 about the nature of the SPP for different regions of the dispersion curve. The conditions required for this situation correspond to those of the dispersion curve at very low frequencies where most metals can be considered as near perfect metals. At these low frequencies the SPP dispersion curve asymptotically approaches the lightline, indicating that the SPP is very much like a grazing photon. At the other limit of $|\varepsilon_{2r}| = \varepsilon_{1r}$ (at ω_{sp}) figure 2.2.2.1 shows that the decay length of the fields in both media have reduced to zero, indicating that there are no longer exponentially decaying fields present. The dispersion curve at this point has infinite k_x (k_{SPP}), and since $k_0^2 = k_{SPP}^2 + k_y^2$, k_y must equal minus infinity which corresponds to a decay length of zero. Other effects involving retardation then need to be taken into account.

2.2.3 Propagation Length

The propagation length is defined in a similar manner to the penetration depth in the last section. The imaginary part of k_x is given by:

$$k_{xi} = \frac{k_0}{2}\left(\frac{\varepsilon_1\varepsilon_{2r}}{\varepsilon_1 + \varepsilon_{2r}}\right)^{3/2}\frac{\varepsilon_{2i}}{\varepsilon_{2r}^2} \qquad 2.2.3.1$$

and therefore the propagation length is given by:

$$L_x = \frac{1}{2k_{xi}} = \frac{\lambda_0}{2\pi}\left(\frac{\varepsilon_1 + \varepsilon_{2r}}{\varepsilon_1\varepsilon_{2r}}\right)^{3/2}\frac{\varepsilon_{2r}^2}{\varepsilon_{2i}} \qquad 2.2.3.2$$

For silver at 632.8nm ($\varepsilon_{ag} = (-17.6 + 0.67i)$) the propagation length is 42.6μm.

From these results we now have a picture of the SPP on a planar metal surface. It is a longitudinal surface charge oscillation coupled to incident EM radiation which, at optical frequencies, is propagating along the surface. This surface charge oscillation causes field loops in the two media and the SPP is trapped at the surface since the total

fields of the SPP decay exponentially away from the surface into both media. This is shown schematically in figure 2.2.3.1.

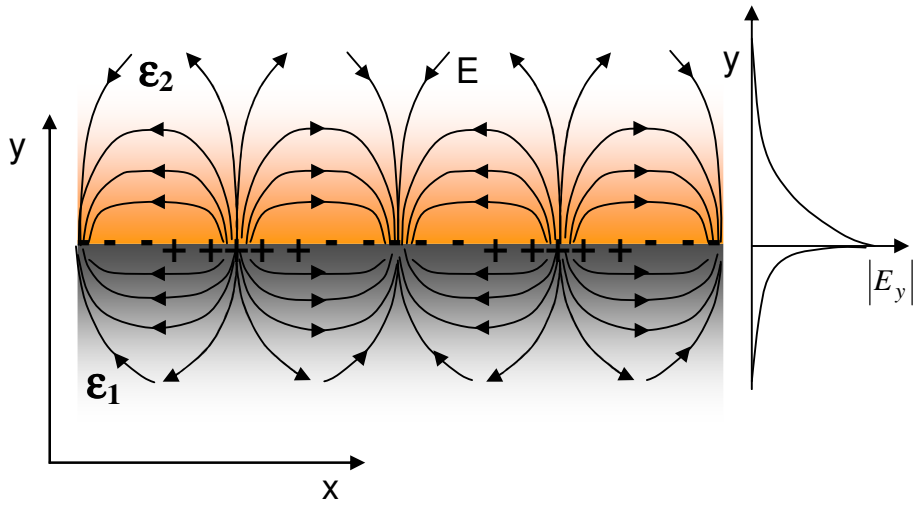

Figure 2.2.3.1 A representation of the polarisation surface charge density and associated electric field for the SPP mode. The electric field decays exponentially into both the metal (ε_1) and dielectric (ε_2)

The localised nature of the fields associated with SPPs means they are ideally suited for probing surface properties. They may be used to determine the optical constants of metals and overlayers (Tillin and Sambles [1988], Watson and Sambles [1991], Hibbins, Sambles and Lawrence [1998]). They may also be used to determine the profile of gratings (section 2.3) by fitting experimentally obtained reflectivity data to grating theories, and for sensor applications (Jory, Bradberry, Cann and Sambles [1995], Nylander, Liedberg and Lind [1982]).

2.3 Grating Coupling to Surface Plasmon Polaritons

In this section we shall described the method of coupling to, and the properties of, SPPs on metal diffraction gratings.

2.3.1 The Dispersion Relation

A schematic of the system discussed and the terms used in this and following sections is shown in figure 2.3.1.1

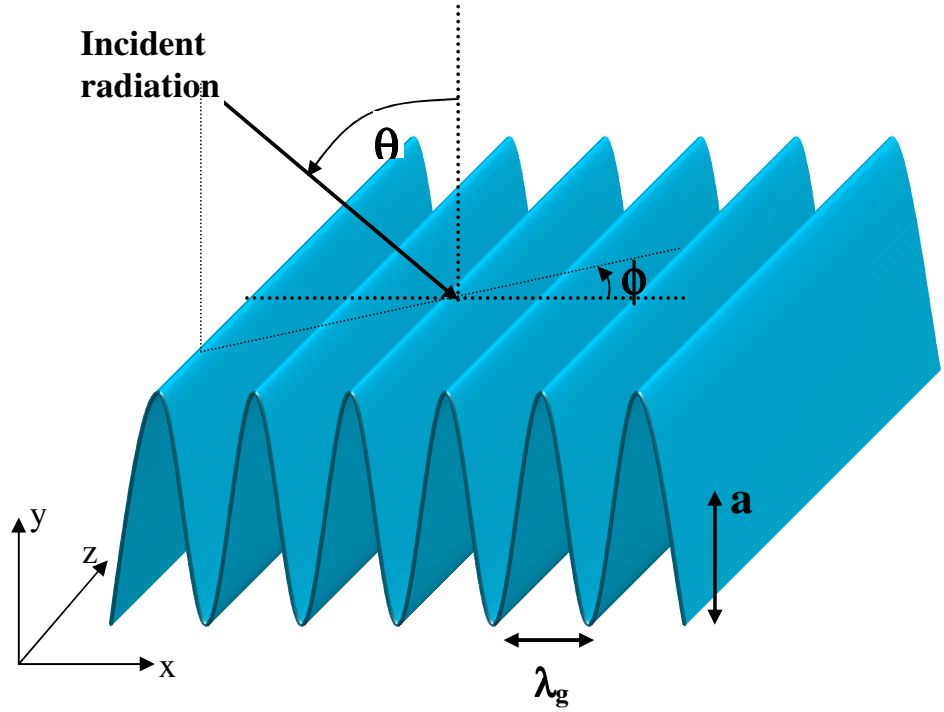

Figure 2.3.1.1 A Schematic showing the grating system under consideration and the terms used in order to define it.

A photon incident upon a grating surface can be scattered by subtracting or adding an integer multiple of the grating vector $\mathbf{k}_g$ ($k_g = 2\pi/\lambda_g$, where λ_g is the grating pitch). This is the origin of the diffracted orders produced by such a surface. If the grating is oriented such that it is in the classical mount (where the plane of incidence contains the grating vector - the azimuthal angle $\phi = 0$) and the frequency of the incident light and the incident polar angle are such that the diffracted order is evanescent, then their enhanced momentum is available to couple radiatively to SPPs according to the equation:

$$k_{SPP} = k_0 \sin\theta \pm Nk_g \qquad 2.3.1.1$$

The mechanism for grating-coupling to the SPP is best described by considering the effect upon the SPP dispersion curve (equation 2.2.1.10, figure 2.2.1.2) of corrugating the interface. In reciprocal space the periodicity of the surface is represented

by a line of points separated by the grating vector k_g, and the SPP dispersion curve, as well as the incident light, may be scattered from these lattice points. The resulting dispersion curve is shown in figure 2.3.1.2. Since the dispersion curve of the diffracted SPP now lies between the light lines (the region of ω - k space which is available to incident radiation – the shaded region in figure 2.3.1.2), the SPP may be directly radiatively coupled to the incident radiation.

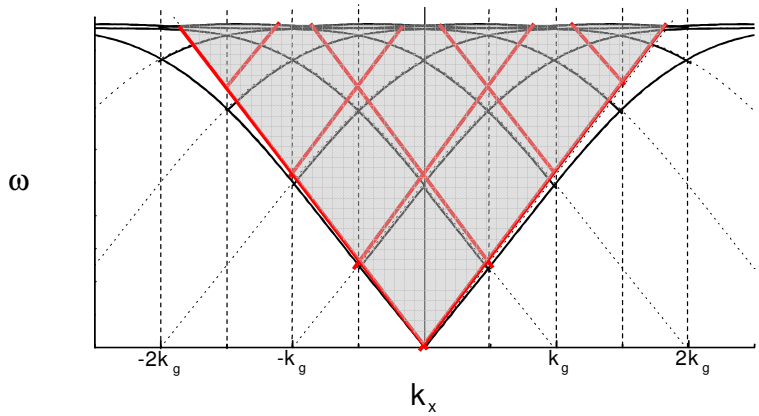

Figure 2.3.1.2. The SPP dispersion curve for a shallow monograting. The full black lines are the SPP dispersion curves, and the faint black lines are the diffracted and non-diffracted light lines. Due to the scattering from the grating vector (k_g) the SPP dispersion curve may be folded inside the light lines and may be coupled to by incident radiation.

If the frequency of the incident radiation is chosen such that the SPP can be excited at normal incidence then the resulting SPP is a combination of the two SPPs scattered from $\pm k_g$ and since the wavevectors of the SPPs are in opposite directions the resulting SPP is a standing wave (in fact there are two possible standing wave solutions which produce a band-gap, but this will be discussed in more detail in section 2.3.2)

Unlike the planar system case the dispersion curve is not identical in all directions in reciprocal space since the periodicity is in the x direction only, and therefore the scattering lattice points only occur in the x direction. When the light is incident at non-zero azimuthal angles (the conical mount) the SPP is still coupled to by diffraction caused by scattering from the periodicity in the x direction. Therefore the SPP propagates at close to the direction of the evanescent diffracted order (when the SPP wavevector is close to that of the maximum wavevector available to the diffracted order), and the dispersion curve of the SPP is given by taking a slice through the scattered SPP dispersion curve in the (k_x, k_z)–ω plane, with the values of k_x and k_z given by:

$$k_x = k_0 \sin\theta \cos\phi \qquad 2.3.1.2$$

$$k_z = k_0 \sin\theta \sin\phi \qquad 2.3.1.3$$

An example of how the SPP dispersion curve appears in one of these planes is shown in figure 2.3.1.3, for the k_z-ω plane with $k_x = 0$ (a 90° azimuthal angle).

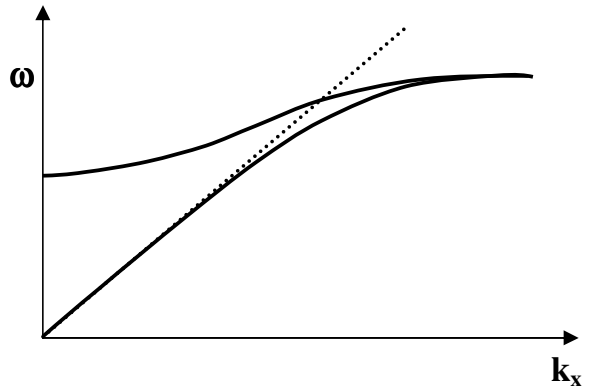

Figure 2.3.1.3 The dispersion curve of a SPP on a grating structure in the ω-k_z plane with $k_x = 0$. The SPP dispersion curve centred at the origin is shown (always outside of the lightline) as is the SPP dispersion curve scattered from $+k_g$. The form of this scattered SPP dispersion curve arises from taking a slice through the SPP 'horn' shape centred at $k_x = k_g$.

Previously we mentioned that the SPP propagates at close to the direction of propagation of the evanescent order which excites it. In order to show this it is useful to plot out the dispersion curve in the k_x-k_z plane for constant ω (figure 2.3.1.4).

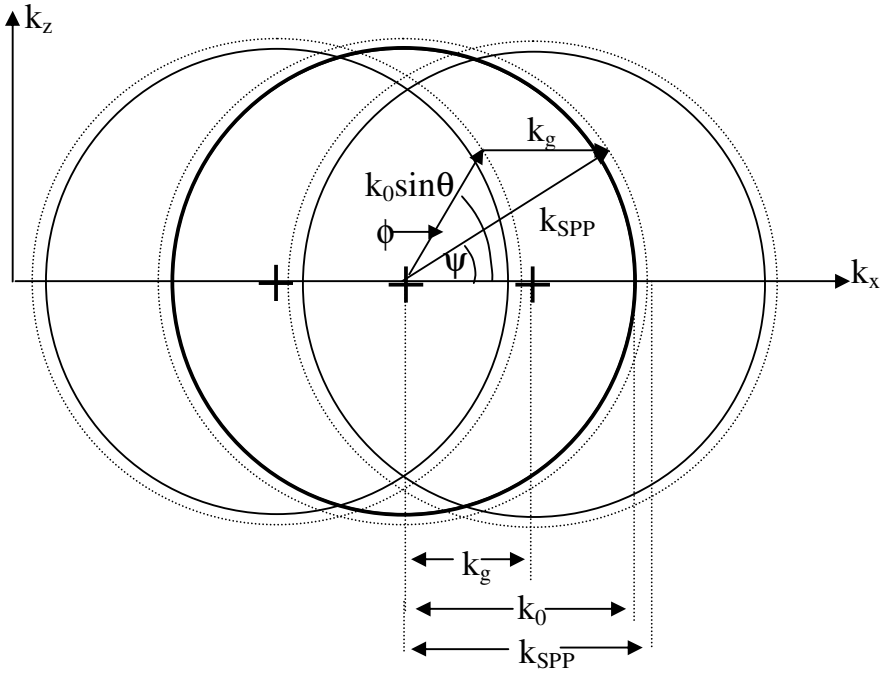

Figure 2.3.1.4 The thickly drawn circle of radius k_0 centred upon the origin describes the maximum wavevector available to a photon in the plane of the grating surface. At a slightly greater radius than this is

the SPP circle (dotted line) corresponding to the dispersion curve described in the text. Both of these circles have been scattered by k_g, which results in the two circles occurring at the reciprocal lattice points. The angle ψ is the propagation angle of the SPP with respect to the k_x direction.

The light circle and SPP dispersion curve originating at the origin, and those scattered by the lattice point at $\pm k_g$, are shown, including the mechanism for coupling to the SPP for a non-zero azimuthal angle. The SPP circle scattered from $-k_g$ is inside the light circle originating at the origin, which corresponds to the maximum wavevector available to the incident radiation. $k_0 \sin\theta$ is the in-plane wavevector of the incident light for the conditions required for SPP excitation from the $-k_g$ scattering process, where the coupling condition (previously equation 2.3.1.1 for the classical mount) is, by simple trigonometry, now:

$$k_{SPP}^2 - k_g^2 = k_0^2 \sin^2\theta - 2k_0 k_g \sin\theta\cos\phi \qquad 2.3.1.4$$

The scattering by k_g then gives the wavevector of the SPP at the equivalent point on the SPP circle originating at the origin. With the propagation angle of the SPP with respect to the k_x axis, ψ, given by

$$\psi = 90 - \cos^{-1}\left(\frac{k_0 \sin\theta \sin\phi}{k_{SPP}}\right) \qquad 2.3.1.5$$

For an example, if we use equations 2.3.1.4 and 5 to determine the propagation angle for 400nm wavelength light incident at a 45° azimuthal angle on a 2μm pitch silver grating the propagation angle is $\psi = 40.6°$ with respect to the k_x axis.

It is clear from figure 2.3.1.4 that if we use the same method for determining the propagation angle for light incident in the k_z-ω plane with $k_x = 0$ (a 90° azimuthal angle) the resulting SPP is a product of the two SPPs scattered from $\pm k_g$. The two contributing SPPs are propagating at:

$$\pm\psi = \tan^{-1}\left(\frac{k_g}{k_y}\right) \qquad 2.3.1.6$$

which is clearly not in the plane of the incident light. The resulting SPP is, therefore, a standing wave in the x direction (produced since the x-component of the wavevectors of the two SPPs are of equal magnitude and in opposite directions) which is propagating in the z direction.

Previously, for the planar interface case (section 2.2), the polarisation of the incident light was defined with respect to the plane of incidence of the light, and only

TM polarised light was capable of exciting the SPP since a component of the electric field normal to the surface was necessary. Now that the surface of the metal has been patterned with a grating structure the symmetry of the surface has been broken, and the incident light may now have components of its electric field normal to the surface for both linear polarisations depending upon the orientation of the grating grooves with respect to the plane of incidence. If the light is incident in the classical mount only TM polarised light may excite the SPP since TE polarised light has its electric field parallel to the grating grooves, and is therefore never normal to the surface for any polar angle. However, at a 90° azimuthal angle TE polarised light has an electric field component perpendicular to part of the surface and may be used to excite the SPP, whereas TM polarised light may not under most circumstances. (TM polarised light does have a component of its electric field normal to the surface for non-zero polar angles, but in general may no excite a SPP. However, there is a special case where TM polarised light may excite the SPP, but it requires a very deep grating with a specific type of profile (Watts, Preist and Sambles [1997]) and this will be discussed in chapter 7).

The SPP dispersion curves for these two planes in reciprocal space (shown in figures 2.3.1.2 and 3) can only be excited with TM and TE polarised light respectively. For azimuthal angles between 0° and 90° both TM and TE polarised light may excite the SPP, though due to the difference in the magnitude of the electric field component normal to the surface for the two polarisations the coupling strength to the SPP will be different for each. In general, the maximum coupling possible by TM polarised light is at $\phi = 0°$ and reduces monotonically as ϕ is increased until it disappears at $\phi = 90°$. For TE polarised light this behaviour is reversed.

2.3.2 Coupling Strength to SPPs

A plot of the zero-order TM reflectivity as a function of frequency for a 1μm pitch 30nm amplitude sinusoidal silver grating (with the dielectric function of silver described with a Drude model (as described in section 2.2) with $\omega_p = 1.32 \times 10^{16} s^{-1}$, and $\tau = 1.4 \times 10^{14} s$), for normal incidence and $\phi = 0°$ is shown in figure 2.3.2.1 (obtained numerically using the computer code based upon the method of Chandezon described in Chapter 2 – as will be the rest of the modelling in this chapter).

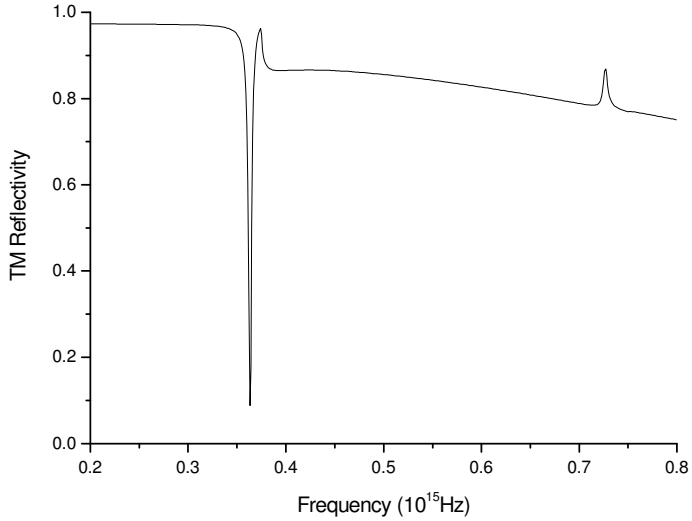

Figure 2.3.2.1 The zero-order TM reflectivity as a function of frequency for a 1μm pitch 30nm amplitude sinusoidal silver grating (with the dielectric function of the silver described with a Drude model with $\omega_p = 1.32\text{x}10^{16}\text{s}^{-1}$, and $\tau = 1.4\text{x}10^{14}$s), for normal incidence and $\phi = 0°$. Three features are evident; at f = $0.36\text{x}10^{15}$Hz a SPP minimum due to the first order SPP, at f = $0.37\text{x}10^{15}$Hz a pseudo critical edge due to the 1st diffracted order becoming evanescent, and at f = $0.735\text{x}10^{15}$Hz a small reflectivity maximum due to the second order SPP.

There are three main features to be seen in figure 2.3.2.1. Firstly there is a critical edge at f = $0.37\text{x}10^{15}$Hz which correspond to the frequency at which the 1st diffracted order has become evanescent and redistributed its energy to the remaining propagating orders. Secondly there are two features at f = $0.36\text{x}10^{15}$Hz and f = $0.735\text{x}10^{15}$Hz which are the frequencies at which the SPPs corresponding to the 1st and 2nd order diffraction processes occur. First we shall discuss the form of the reflectivity feature for the 1st order SPP.

The reflectivity feature for the SPP corresponding to the 1st order diffraction is clearly manifested as a reflectivity minimum in the zero-order reflection from the structure. This reflection minimum arises due to a combination of the specularly reflected light, and the re-radiated light from the SPP.

The incident light is scattered into a diffracted order which, when the coupling condition is satisfied (equation 2.3.1.1), excites the SPP which propagates along the surface. The SPP may then undergo a second scattering process which results in it being

re-radiated into the specularly reflected order, having had some of its energy lost due to Joule heating in the metal (caused by the oscillatory motion of the surface charges). The phase of the evanescent 1st diffracted order is 90° out of phase with respect to the incident light and, since the SPP is resonant as a forced damped oscillator, there is a second 90° phase change into the SPP. Therefore, the SPP is 180° out of phase with the incident light. Upon coupling of energy out of the SPP back into the specularly reflected order this process is repeated so that the re-radiated light is in phase with the incident light. However, the specularly reflected light is 180° out of phase with the incident light since the electric field is reversed upon reflection from a metal surface. Therefore, since the specularly reflected light and the re-radiated light are 180° out of phase, they cancel to produce a reflectivity minimum.

The shape of the resonance is given by the combination of the strength of coupling into and out of the SPP, and the absorption of energy from the SPP into the metal. The re-radiation and absorption of energy are the damping terms and, as with other resonance phenomena, the width of the resonance is due to the total damping of the oscillator. The absorptive part of the damping is relatively independent of the depth of the grating, whereas the re-radiation term increases as approximately a^2 (similar to the intensity of a diffracted order). Also, the depth of the resonance is given by the ratio of the radiative part of the damping to the absorptive part of the damping, with a maximum depth ($R = 0$) when $\Delta\omega_{rad} = \Delta\omega_{abs}$ (where $\Delta\omega_{rad}$ is the broadening due to the radiative part of the damping, and $\Delta\omega_{abs}$ is the broadening due to the absorptive part of the damping). Therefore, for a very shallow grating the SPP resonance is manifested as a narrow, but shallow resonance and, with increasing depth, the resonance widens and deepens until, when the radiative part of the damping is equal to the absorptive part of the damping, the minimum of the resonance is at a reflectivity of zero. On increasing the depth further the width of the resonance continues to increase, but the depth of the minimum decreases so that, for very deep gratings, a very broad and shallow minimum results. This is shown in figure 2.3.2.2, which shows the change in the zeroth order reflection from a grating around the frequency at which the 1st order SPP is excited with changing amplitude of the grating, with the grating described by the same parameters as those for figure 2.3.2.1 (for a fuller description of the effect of the grating profile on the optical response of gratings see Watts, Hibbins and Sambles [1999]).

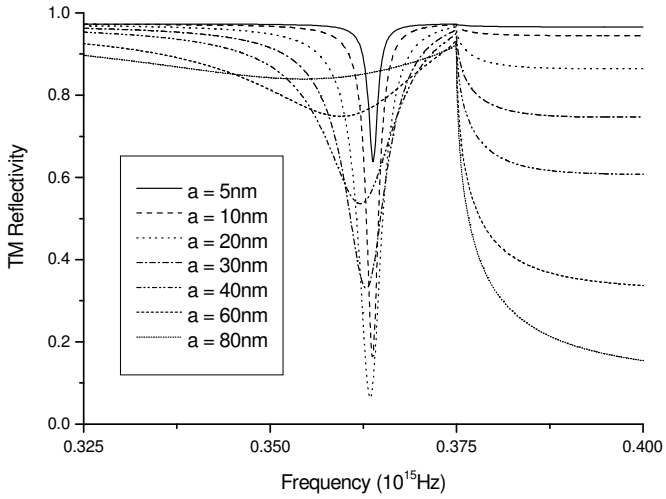

Figure 2.3.2.2 The zero-order TM reflectivity for the same system as in figure 2.3.2.1, but with changing amplitude of the grating.

The SPP corresponding to 2nd order diffraction from the grating shows as a small maximum in figure 2.3.2.1. The reason that this feature is a maximum, rather than a minimum as for the SPP corresponding to the 1st order diffraction, is that the phase of the re-radiated light is now in-phase with the specularly reflected order, rather than in anti-phase with it. This is because the scattering processes which lead both to the excitation of the SPP, and the re-radiation out of the SPP, is a $2k_g$ process, which on a purely sinusoidal grating can only arise through two scattering processes. Therefore, there is an extra 90° phase shift with respect to the incident light into the SPP, and upon re-radiation, which leads to the re-radiated light being 180° out of phase with the incident light as opposed to in-phase from the single scattering process required for the 1st order SPP. Therefore it is in-phase with the specularly reflected light and adds constructively to it.

The fact that it is a two scatter process is also the reason that the feature is much smaller than that for the 1st order SPP, since a two scatter process has a much smaller probability of occurring than does a one scatter process.

For gratings which are not purely sinusoidal, but which have higher harmonic components, the processes which lead to higher order SPP excitation can occur due to single scattering processes which will then mean that the features are a) much larger since single scattering processes are more likely to occur, and b) appear as minima since the single scattering process leads to the re-radiated light being in anti-phase with the specularly reflected light.

The addition of higher harmonics in the grating profile also affects the lower order SPP features. A two scatter process comprising of scattering from the $2k_g$ component in the $+k_x$ direction, followed by a scattering from the k_g component in the $-k_x$ direction, produces an effect in the 1st order SPP reflectivity feature in the specularly reflected order. However, since this is a multiple scattering process it is much weaker than the single scattering process from the fundamental component of the grating profile, and therefore the effect is relatively small.

This sensitivity of the optical response of gratings to the grating profile, and the accuracy of theoretical models, allows the profile of a grating described by a Fourier series to be accurately determined (Wood, Sambles, Cotter, and Kitson [1995], Pockrand [1974], Pipino and Schatz [1994]).

2.3.3 Band-Gaps in the Dispersion of SPPs on Gratings

In order to introduce the concept of band gaps in the SPP dispersion relation we shall first consider the simplest optical system which exhibits band-gaps: The quarter-wave multilayer dielectric stack consisting of alternate layers of two dielectrics of refractive indices n_1 and n_2 (figure 2.3.3.1). The thickness of the two differing slabs are chosen to be:

$$t_1 = \frac{\lambda_0}{4n_1} \qquad\qquad t_2 = \frac{\lambda_0}{4n_2}.$$

where λ_0 is the wavelength of light in vacuum.

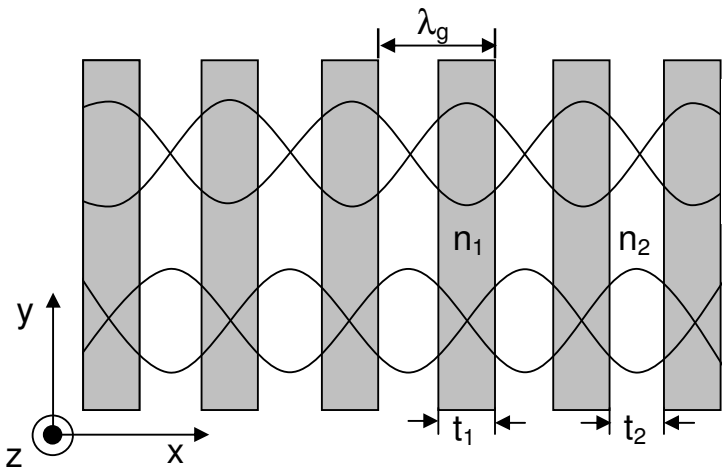

Figure 2.3.3.1 A sketch of the standing waves in a dielectric stack. The shaded regions correspond to a dielectric with a higher refractive index than the unshaded regions. The standing wave with field extrema concentrated in the high index medium has a lower energy than the standing wave with field extrema in the low index medium

Consider light propagating at normal incidence onto a dielectric stack such that its optical wavevector is equal to half of the Bragg vector (corresponding to the stack periodicity). If this is the case, Bragg scattering occurs resulting in both forward and backward propagating waves, which may interact to produce a standing wave. By simple symmetry arguments the standing wave produced must have its maxima and minima at the mid-points of the high index dielectric medium, or at the mid-points of the low index medium. These two possible standing wave solutions have different energies since the optical field at the midpoints of the media will be altered due to the refractive index of the medium in which it occurs. Therefore, the two possible standing wave modes have different energies (frequencies) but the same wavevector, and a band gap in the dispersion of the mode will occur.

The situation of a SPP propagating in the direction normal to the grating grooves is similar (Barnes, Preist, Kitson and Sambles [1996], Barnes, Kitson, Preist and Sambles [1997]). When the wavevector of the SPP is equal to half that of the grating vector a band gap occurs (at $k_x = k_g/2$) with the high and low energy solutions corresponding to different field distributions with respect to the grating grooves. If a

longer pitch component is added to the grating profile this band-gap is scattered into the region of ω-k_x space in which it may be coupled. In fact if the longer pitch component is exactly twice that of the original grating the band gap will occur at normal incidence. Conversely, therefore, if a $2k_g$ component is added to a grating profile then a band-gap at the crossing point of the SPP at normal incidence is opened. Another way of looking at this is that the k_g component of the grating scatters two SPPs (from $+k_g$ and from $-k_g$) so that they cross at normal incidence and are counter propagating. The $2k_g$ component then couples these two SPPs together forming the band-gap (figure 2.3.3.2).

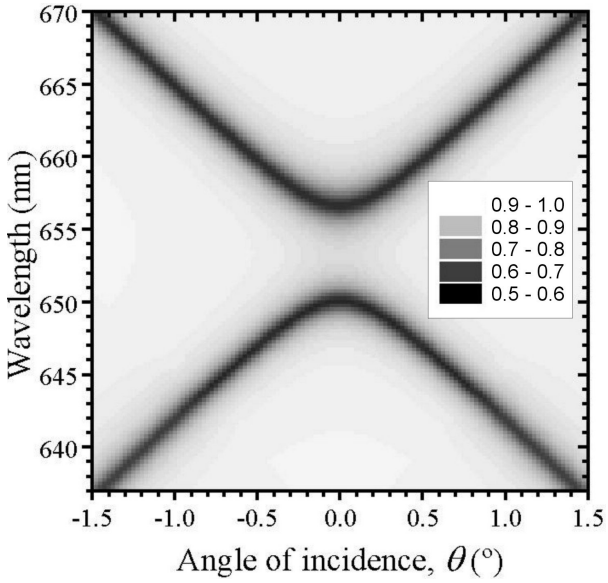

Figure 2.3.3.2 Numerically modelled zero-order TM reflectivity of a grating structure comprising two Fourier components. The first harmonic component produces a band-gap at the intersection of the +1 and-1 SPP branches (at normal incidence). The parameters used in the modelling were $\lambda_g = 634$nm, $a_0 = 5$nm, $a_1 = 2$nm and the permittivity of the metal is $\varepsilon_r = -17.5$ and $\varepsilon_i = 0.7$. The two components of the grating profile are in phase with each other.

In fact it is possible for the k_g component to produce both these effects (scattering of the two SPPs, and coupling them together). However, the coupling together of the two SPPs requires a two scatter process which is far weaker than the one scatter process from a $2k_g$ component of the grating profile. Therefore the $2k_g$ scattering process is the dominant one.

One of the main methods used to create diffraction gratings for the optical region of the spectrum is through the use of holographic lithography. Due to non-linearities in the exposure and development characteristics of photoresist (the material used in which the grating profile is recorded) higher harmonic components tend to

appear in the resulting grating profile. Therefore, band-gaps are frequently observed in the dispersion of SPPs on gratings.

So far we have only considered the case of the band gap caused by the $2k_g$ component of the grating which is scattered by the k_g component into the region of ω-k_x space where it may be excited by incident radiation. Of course, similar arguments hold for the other SPP crossing points at normal incidence. The two counter propagating SPPs at the crossing point which results from the $2k_g$ scattering process must be coupled together by a $4k_g$ component of the grating profile and so on. There are also SPP crossing points which occur at all half integer values of k_g (see figure 2.3.1.2). In order to couple these two SPPs together to form a band-gap a component of the grating profile with $(n+m)k_g$ is required (where n and m correspond to the k_g components of the grating producing the two interacting SPPs). The case of band-gaps at these half integer values of the grating vector will be discussed in more detail in chapter 7.

We shall now investigate the origin of the energy gap in terms of the field distributions between the two standing wave solutions. For the bandgap corresponding to the crossing of the SPPs scattered from $+k_g$ and $-k_g$ the two possible standing waves on either side of the band-gap are, by simple symmetry arguments:

$$\Psi_1 = \exp(ik_g x) + \exp(-ik_g x) = 2\cos(k_g x) \qquad 2.3.3.1$$

$$\Psi_2 = \exp(ik_g x) - \exp(-ik_g x) = 2i\sin(k_g x) \qquad 2.3.3.2$$

One of these solutions has its nodes at the peaks of the $2k_g$ component of the grating, and the other at the troughs of the $2k_g$ component of the grating (figure 2.3.3.3). The high energy solution has its extrema of the normal field component and surface charge distribution in the troughs of the $2k_g$ component, whereas in the low energy solution they occur on the peaks of the $2k_g$ component. This is due to the increased compression of the field lines leading to a higher stored energy for the high energy solution.

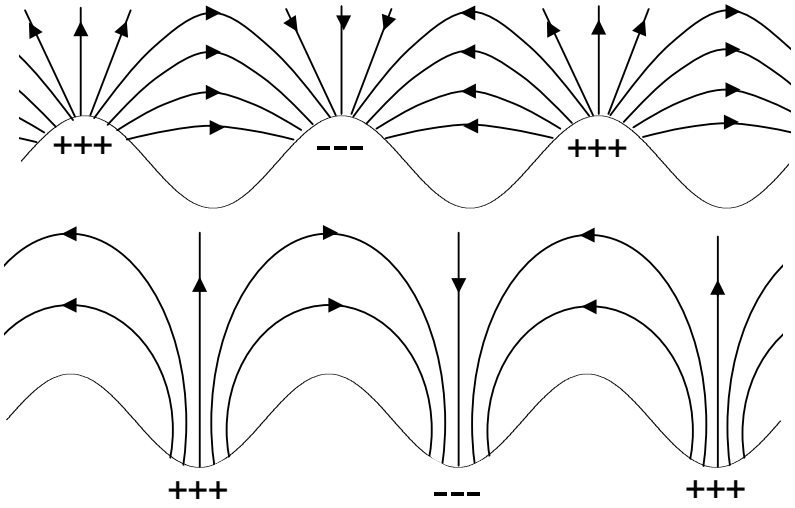

Figure 2.3.3.3 The electric field and surface charge distribution for the two standing wave solutions on the $2k_g$ component of the grating profile. The field lines in the lower sketch are more distorted, and therefore this is the higher energy mode.

An analytical form for the size and central frequency of the band gap has been sought by, for example, Mills [1977], Seshadri [1985] and Barnes, Preist, Kitson and Sambles [1996]. All of these analytical methods can only describe the band gap for shallow grating structures. The basic results obtained by Barnes *et al* using the method of Chandezon are that the gap width $\Delta\omega$ is, to first order, given by:

$$\frac{\Delta\omega}{\omega_0} = \frac{2k_g a_2}{\sqrt{-\varepsilon_1\varepsilon_2}} \qquad 2.3.3.3$$

where a_2 is the amplitude of the $2k_g$ component, and ω_0 is the frequency at which the SPP would have been excited were it not for the opening of the band gap.

The central frequency of the bandgap $\overline{\omega}$ is given by:

$$\frac{\overline{\omega}^2}{\omega_0^2} = \left(\frac{\omega_0}{c}\right)^2 \left[1 - (k_g a_2)^2\right] \qquad 2.3.3.4$$

The bandgap in the classical mount always occurs at this central frequency since the SPP is propagating in the direction of k_g, however if the SPP is propagating at an angle ψ with respect to k_g (the azimuthal angle is non-zero) this is no longer the case and the gap will instead occur when

$$k_{SPP}(\psi) = \frac{k_g}{\cos\psi} \qquad 2.3.3.5$$

(from consideration of the requirements to excite the SPP (see figure 2.3.1.4)), and therefore the frequency of the centre of the gap increases as $1/\cos\psi$.

As mentioned earlier this analytical form for the bandgap is only valid for shallow gratings and results for the case where the amplitude of the $2k_g$ component is large will be discussed in chapters 6 and 7.

Finally we shall discuss the effect of the relative phase between the k_g and $2k_g$ components in the description of the grating. The relative phase between the fundamental and first harmonic determines the coupling strength to the band edges

around the Brillouin zone boundaries (occurring at half integer values of the grating vector, and set up due to the periodicity of the structure), which in the case considered here is for normal incidence. If the phase of the $2k_g$ component is in phase with the k_g component then both band edges are coupled to equally (in this case the profile is blazed, or non-symmetric). However, if the phase between them is ±90° (the grating profile is symmetric) then only one of the band edges is coupled to (Weber and Mills[1985], Nash,Cotter, Wood, Bradberry and Sambles[1995]). This is shown in figure 2.3.3.4.

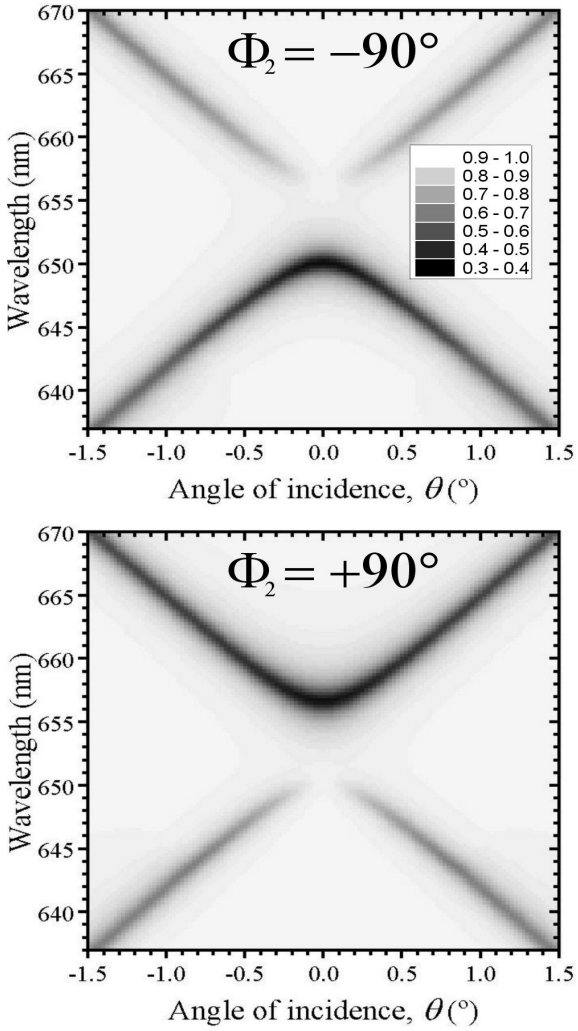

Figure 2.3.3.4 Numerical modelled zero-order TM reflectivity plots showing the influence of the phase difference between the fundamental and first harmonic components of the grating profile. The grating parameters are otherwise the same as for figure 2.3.3.2.

As mentioned previously a component of the incident electric field must be normal to the surface for excitation of a SPP to occur. For the standing waves which form the band edges of the bandgap the electric field must be normal to the surface at

the point on the grating profile where the surface charges are oscillating. The k_g and $2k_g$ components of the surface profile for the three cases considered in this section (the two components are 0°, +90°, -90° out of phase with each other) are shown schematically in figure 2.3.3.5.

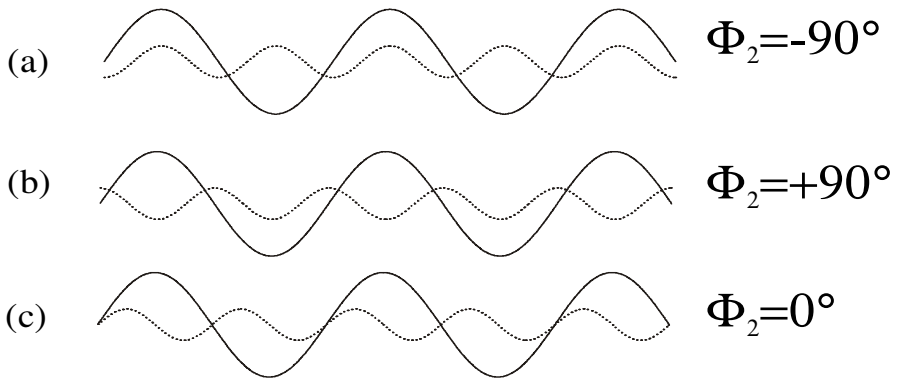

Figure 2.3.3.5 The k_g and $2k_g$ components of a distorted sinusoidal grating with relative phase between the two components of –90°, +90° and 0°.

We shall first consider the case where the two components are –90° out of phase with each other (figure 2.3.3.5(a)). For normally incident radiation the electric field has no normal components at the peaks and troughs of the k_g component of the grating, which is the component which enables the incident radiation to couple to the SPP. The charge distribution for the two standing waves relative to the $2k_g$ component are shown in figure 2.3.3.3. For the low energy mode the charges are located on the peaks of the $2k_g$ component, and from figure 2.3.3.5(a) these occur at the peaks and troughs of the k_g component. Since there is no normal component of the electric field at these points there can be no coupling to the SPP. However, the high energy solution has its charges located at the troughs of the $2k_g$ component, and these correspond to the sloping sides of the grating peaks of the k_g component. Therefore, there is a normal component of the incident electric field in this case, and coupling can occur. The reverse is true for the case where the $2k_g$ component of the grating profile is +90° out of phase with the k_g component since the peaks of the $2k_g$ component now occur on the sides of the peaks of the k_g component. If the phase between the two components is 0° then equal coupling to the two modes occurs. These results agree with the reflectivity plots of figures 2.3.3.2 and 2.3.3.4.

2.3.4 Polarisation Conversion from Gratings

There are two mechanisms by which linearly polarised incident radiation upon a grating may be converted to the orthogonal linear polarisation state. The first of these is mediated by SPP excitation (Inagaki, Goudonnet, and Arakawa [1986], Bryan-Brown, Sambles and Hutley [1990], Elston, Bryan-Brown and Sambles [1991], Depine and Lester [2001]). As discussed in section 2.3.1, when the orientation of the grating is such that the azimuthal angle is non-zero both TM and TE polarised light may be used to couple to SPPs. Since this is the case it is clear that when light is re-radiated out of the SPP both TM and TE polarised light may be emitted. If TM polarised light is used to excite the SPP then in this situation both TM and TE polarised light will be re-radiated so that when the specularly reflected light is investigated it consists of both polarisations. Therefore, some of the incident light has been polarisation converted. For shallow gratings the maximum in this polarisation converted signal occurs when the grating is oriented at an azimuthal angle of 45°, since it is at this orientation that the coupling strength to SPPs by TM and TE polarised light for a particular polar angle is most nearly equal. (Remember that the coupling strength to the SPP reduces monotonically from a maximum at a $\phi = 0°$ to zero at $\phi = 90°$ for TM polarised light, and vice versa for TE polarised light). For shallow gratings the ϕ dependence of the polarisation converted signal goes approximately as $\sin^2(2\phi)$, which can be shown by considering the electric field components relative to the grating profile (see Bryan-Brown *et al* (1990)). It should also be noted that there are a second set of polarisation conversion maxima which occur when the propagation direction of the SPPs is at 45° to the direction of the grating vector in the plane containing the grating (Depine and Lester [2001]). The polarisation conserved and converted reflectivities as a function of k_x and k_z for TM and TE polarised light incident upon a 1μm pitch 40nm amplitude silver grating for incident light of frequency 0.474×10^{15}Hz is shown in figure 2.3.4.1.

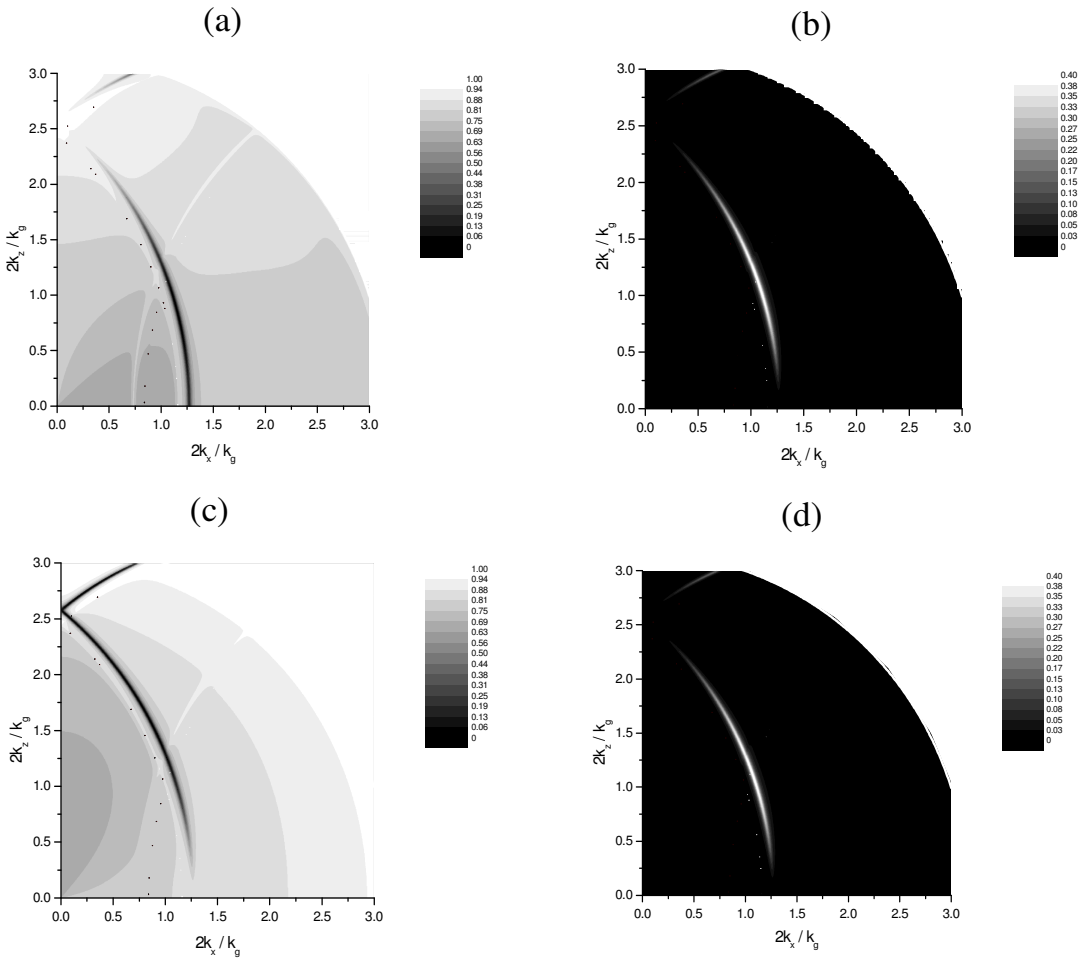

Figure 2.3.4.1 Numerically modelled polarisation conserved and polarisation converted reflectivities for TM and TE polarised light of frequency 0.474 x 10^{15}Hz incident upon a 1μm pitch, 40nm amplitude silver grating as a function of k_x and k_z. a) Polarisation conserved for TM polarised incident light, b) Polarisation converted for TM incident light, c) Polarisation conserved for TE polarised incident light, and d) Polarisation converted for TE polarised incident light.

The second method by which a grating produces polarisation conversion also involves the rotation of the grating by some azimuthal angle ϕ, and is cyclically dependent upon the grating depth (Watts and Sambles [1997]). It may cause polarisation conversion even when the grating pitch is so short that the structure is no longer

diffractive (it is zero-order). The mechanism by which this polarisation conversion occurs is as follows.

The electric field of the incident light may be separated into two components: one parallel to the grating vector (E_x) and one parallel to the grating grooves (E_z). Assuming that the light is incident in a region of ω-k space in which the structure is non diffractive the E_z component, since it does not cut across the grating grooves, will reflect from the grating as if it were reflected from a planar surface at the average plane of the grating. However, the E_x component does cut across the grating grooves and may produce circulating fields within the grating grooves, which alters the effective average plane of the grating for the E_x component. The difference in the mean positions of these 'effective mirrors' for the two electric field components produces a phase difference between them and this rotates the plane of polarisation. The phase difference between the two components depends upon the difference between the mean positions of these effective mirrors, and therefore the polarisation converted signal depends upon the depth of the grating with a maximum occurring when the phase difference between them is 180°. It is cyclic as a function of depth since it is possible to produce more than one circulating field loop within the grating grooves, it is also dependent upon ϕ since the polarisation conversion will depend upon the ratio of E_x to E_z, with a maximum occurring when they are of equal magnitude ($\phi = 45°$).

2.4 Summary

In this chapter we have introduced the surface plasmon polariton as an oscillation of the surface charge density at the boundary between a metal and a dielectric with mixed longitudinal and transverse character. We have derived the dispersion of the SPP mode for a planar interface, and shown that it propagates along the surface with fields which decay exponentially away from the interface into both bounding media. It is not possible to couple to the SPP on a planar surface since its momentum is greater than that available to incident EM radiation and, therefore, mechanisms to enhance the momentum of incident photons are required. We have described in some detail the mechanism by which a grating allows coupling to the SPP by adding or subtracting integer values of the grating vector to the wavevector of the incident light. We have then

determined the dispersion of the SPP on gratings, including a discussion on the formation of band-gaps at the points where two branches of the SPP dispersion curve cross. Finally, the phenomenon of polarisation conversion from gratings has been explained.

Chapter 3

Calculating the Optical Response of Grating Structures

3.1 Introduction

In this chapter the ways in which the optical response of grating structures may be calculated is described. This problem has been studied for nearly a century, ever since Rayleigh [1907] first treated the interface as a perturbation of a planar surface and hypothesised that the reflected fields could be considered as a superposition of plane waves. Implementation of these methods are inadequate for describing the optical response of all but very shallow gratings as they do not accurately describe the fields within the grooves, only in the half spaces above and below the grating profile maxima and minima.

The modelling of the optical response of grating structures tends to be computationally very intensive, and the majority of methods to calculate this response have therefore been developed relatively recently. A brief, non-mathematical, description of the most common methods will be presented in this chapter including their main advantages and disadvantages, followed by more detailed descriptions of the two methods utilised throughout this thesis. These two methods are the Iterative Series Solution (ISS) (Greffet and Maassarani [1990]), which is based upon the Rayleigh method, and a co-ordinate transformation method based upon the differential formalism of Chandezon Dupuis, Cornet and Maystre [1982].

The contents of this chapter are not necessary for understanding the remainder of this thesis, but are presented to describe the methods which have been extensively used.

3.2 Overview of Methods

3.2.1 The Perturbation Methods

The Rayleigh method treats the grating surface as a perturbation of a planar surface, and the fields reflected or transmitted from the structure as a superposition of the plane waves created via the interaction of the incident light with the grating (the Rayleigh hypothesis).

The total field is required to be quasi-periodic (from the Floquet theorem), and therefore it may be expanded as a Fourier series. In the two homogeneous regions outside of the grating grooves Maxwell's equations lead to a Helmholtz equation which is then used as a boundary condition for the problem. By combining the expanded representation of the fields with the Helmholtz equation an analytical expression for the field components is obtained, and therefore the optical response of the system may be calculated.

The main problem with the method is that it assumes that the expression for the field components is valid within the grating groove region. This is not generally the case since the Helmholtz equation is valid only for homogeneous regions of space (it is not specified on the grating surface). However, for very shallow gratings it produces results which agree with experiments.

There are also other forms based upon this method which use different field matching conditions such as the point matching method, the Fourier series method, and the variational method. However, they tend to suffer from the same limitations.

The simple perturbative methods can cope with gratings of a wide variety of materials, both metal and dielectric, although since the Rayleigh assumption breaks down for deeper structures, its use is relatively limited. However, due to the fact that the method is very simple in its analytical form it is very easy to implement and is useful as an illustrative tool.

3.2.2 The Integral Methods

The integral methods are more complex than both the perturbative methods and the differential methods. They are rigorous and do not suffer from the same limitations inherent in the methods based upon the Rayleigh assumption, meaning that the optical response of most grating systems can be calculated.

The integral methods define the fields at any point in space in terms of integrals on the grating surface. A set of unknown functions P are contained within these integrals and are defined at the surface. These functions are of physical properties and must be chosen very carefully since they must be capable of defining the fields at any point in space. The problem then becomes one of determining the P functions. In order to solve the boundary conditions several mathematical techniques may be employed such as Green's functions (Wirgin [1964], Neureuther and Zaki [1969]), the theory of distributions (Schwartz [1965,1966]), and the formalism of potentials (Dumery and Filippi [1970]). The fields are then determined from these functions by using the integral expressions for the fields to express the limit values of the field and its normal derivative on both sides of the interface in terms of the P functions. The optical response of the grating may then be obtained.

These methods are extremely complicated for the case of metallic gratings or multi-layered gratings, but have been utilised since they are stable for most situations (e.g. Maystre [1978]). Due to their complexities the computation times can also become much longer than for the other methods.

3.2.3 The Differential Methods

Maxwell's equations may also be solved directly in order to calculate the fields from a grating. If this is performed in Cartesian co-ordinates a set of coupled partial differential equations is obtained, and by utilising a point matching method in order to solve these (Moaveni, Kalhor and Afrashteh [1975], Moaveni, [1988]) the optical response of the grating may be calculated. Though this method is simple to implement it may only be used for shallow gratings since it is inherently unstable.

By using a curvilinear co-ordinate transformation in order to map the grating surface onto a flat plane these problems can be overcome. It is far simpler to solve Maxwell's equations in the transformed co-ordinate frame, and, by applying the appropriate boundary conditions, a set of ordinary differential equations are obtained (Neviére, Cerruti-Maori, Cadilhac [1971]). By solving these the optical response of the structure can then be calculated. The Chandezon method described later in this chapter is of this class, and this method will be described in more detail then.

A final differential method involves the projection of the propagation equations for the system onto a suitable basis of functions. A set of coupled ordinary differential equations are then produced, which can be solved numerically (Petit [1966], Numata [1982]).

In general the differential methods are applicable to a wide variety of grating systems being able to handle metallic or dielectric materials (Neviére, Maystre, and Vincent [1977], Hutley, Verill, Mcphedran, Neviére and Vincent [1975]) and being able to model multi-layered and multi-shaped systems (Plumey, Granet and Chandezon [1995], Preist, Cotter, and Sambles [1995]). They have also been used to calculate the optical response of gratings in the conical mount (Vincent, Neviére, and Maystre [1978], Popov and Mashev [1986], Elston, Bryan-Brown and Sambles [1991]), and have even been used for bigrating structures (Vincent [1978], Harris, Preist, Sambles, Thorpe and Watts [1996]) (though the computation times in this case are very large).

3.2.4 The Modal and Coupled Wave Methods

The classical modal method was developed by Botten, Craig, McPhedran, Adams and Andrewartha [1981a,b] and is well-suited to step-like grating profiles (e.g. rectangular gratings). In this method it is not necessary to use a Fourier expansion to describe the fields since Maxwell's equations can be found in a closed form inside the grating grooves and peaks. The boundary conditions are then applied on the vertical groove walls, and a set of 'modal constants' are obtained. The total field is then represented as a sum over all the modes of the system, and the coefficients of the modal expansion are obtained through applying the appropriate boundary conditions between the region containing the grooves and the homogeneous media on either side.

The main disadvantage of this method is that it is generally highly specialised to particular grating profiles, however Li [1993] has proposed a generalisation to the method which enables it to be used for arbitrary profiles.

An adaptation of this method was proposed by Mohoram and Gaylord [1977,1982,1986], and is sometimes referred to as rigorous coupled-wave theory. They differentiate between the two homogeneous regions above and below the grating and the region containing the grating, and they then expand the dielectric function within the grating region as a Fourier series. The fields in the two homogeneous media are expanded as a sum of plane waves, and in the grating region they expand the fields in terms of their space harmonic components (where the space harmonic components are phase matched with the plane wave sums in the two homogeneous media). By the use of the Helmholtz equation with the Fourier series description of the dielectric function, an infinite set of second-order coupled differential equations are obtained. These are reduced to two infinite sets of first-order differential equations which, on application of the appropriate boundary conditions, may be solved using an eigenvalue/eigenvector technique.

Though originally developed for rectangular grating profiles the method has been generalised to arbitrary profiles which are described in a ‘staircase’ approximation, and it then appears to be very similar to the classical differential method. Rigorous coupled wave theory is widely used since it is very powerful and relatively simple to implement, however the extension to arbitrary profiles suffers the same limitations as the classical differential theory.

3.3 The Iterative Series Solution Method

The Iterative series solution is a simple method which utilises the Rayleigh hypothesis. Though very limited in its use due to the fundamental limitations inherent in methods using the Rayleigh hypothesis, it produces a very simple result for simple grating profiles. Therefore, it is easy to implement a very fast computer code to calculate the efficiencies and phases of the diffracted orders from gratings. The method was first developed by Maradudin [1983] for p-polarised light incident upon a sinusoidal grating in the classical mount, his work being in turn based upon the development of a method

derived by Lopez, Yndurain and Garcia [1978] for the scattering of atoms from a periodically corrugated hard wall. The method was then extended by Greffet [1988] in order to calculate the optical response of a rough interface for both s and p-polarised incident light for any orientation, and this was simplified by Greffet and Maassarani [1990] for the case of a sinusoidal grating. In this section we shall derive the general result for a rough interface using the method and notation of Greffet, and show the simplified results of Greffet and Maassarani.

Firstly we shall consider the case of an electromagnetic wave impinging upon a rough surface described by the equations,

$$z = S(\rho), \text{ and } \rho = (x, z) \qquad 3.3.1$$

where the medium in the upper half space is described by a dielectric function ε_1, and the medium in the lower half space is described by a dielectric function ε_2. The field in the lower medium will be designated $\mathbf{E}_t$, and that in the upper medium will be the sum of the incident field $\mathbf{E}_i$, and the reflected field $\mathbf{E}_r$. All of the fields have a temporal dependence exp(-iωt).

The transmitted field must satisfy the Helmholtz equation:

$$\nabla \mathbf{E}_t + \varepsilon_2 k_0^2 \mathbf{E}_t = 0 \qquad 3.3.2$$

where $k_0 = \omega / c$.

If the Rayleigh expansion is assumed valid then the transmitted field may be written in the following form:

$$\mathbf{E}_t(\mathbf{r}) = \int d\boldsymbol{\kappa} \mathbf{e}_t(\boldsymbol{\kappa}) \exp[i(\boldsymbol{\kappa} \cdot \boldsymbol{\rho} - \gamma_t y)] \qquad 3.3.3$$

where

$$\gamma_t^2 + \boldsymbol{\kappa}^2 = \varepsilon_2 k_0^2, \text{ and } \operatorname{Im}(\gamma_t) > 0 \qquad 3.3.4$$

with γ_t being the component of the transmitted field with wavevector $\boldsymbol{\kappa}$ in the direction perpendicular to the average plane of the surface (the y-direction).

Similarly the incident and reflected fields can be represented by:

$$\mathbf{E}_i(\mathbf{r}) = \int d\boldsymbol{\kappa} \mathbf{e}_i(\boldsymbol{\kappa}) \exp[i(\boldsymbol{\kappa} \cdot \boldsymbol{\rho} - \gamma_i y)] \qquad 3.3.5$$

with

$$\gamma_i^2 + \boldsymbol{\kappa}^2 = \varepsilon_1 k_0^2, \text{ and } \gamma_i = k_0 \cos\theta_i \qquad 3.3.6$$

and

$$\mathbf{E}_r(\mathbf{r}) = \int d\boldsymbol{\kappa}\mathbf{e}_r(\boldsymbol{\kappa})\exp[i(\boldsymbol{\kappa}\cdot\boldsymbol{\rho} - \gamma_r y)] \tag{3.3.7}$$

with

$$\gamma_r^2 + \boldsymbol{\kappa}^2 = \varepsilon_1 k_0^2, \text{ and } \operatorname{Im}(\gamma_r) > 0 \tag{3.3.8}$$

where γ_i and γ_r are the components of the incident and reflected fields wavevector in the y direction.

The extinction theorem appears as a boundary condition (also known as the Green's function method) and we use this in order to determine the coefficients $\mathbf{e}_t(\boldsymbol{\kappa})$:

$$\mathbf{E}_i(\mathbf{r}) = -\frac{1}{4\pi k_0^2}\nabla\times\nabla\times\int\left[\mathbf{E}_t(\mathbf{r}')\frac{\partial G(\mathbf{r},\mathbf{r}')}{\delta n'} - G(\mathbf{r},\mathbf{r}')\frac{\partial \mathbf{E}_t(\mathbf{r}')}{\delta n'}\right]dS' \tag{3.3.9}$$

where $\mathbf{r}$ lies in the medium, $\mathbf{r}'$ is a point on the surface S, $\partial/\partial n'$ denotes differentiation along the normal to the surface S in the outward direction, and

$$G(\mathbf{r},\mathbf{r}') = \frac{\exp(ik_0|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} \tag{3.3.10}$$

is the free space Green's function.

To determine the amplitudes $\mathbf{e}_t(\boldsymbol{\kappa})$ we substitute equation 3.3.3 into equation 3.3.9. To do this it is convenient to introduce the spectral representation of the free space Green's function into equation 3.3.9:

$$G(\mathbf{r},\mathbf{r}') = \frac{i}{2\pi}\int\frac{d\boldsymbol{\kappa}}{\gamma_r}\exp[i\boldsymbol{\kappa}\cdot(\boldsymbol{\rho}-\boldsymbol{\rho}') + i\gamma_r|y-y'|] \tag{3.3.11}$$

Also, the following relations allow the integral in equation 3.3.9 to be performed over the x-z plane:

$$\frac{\delta}{\delta n'} = (\hat{\mathbf{n}}'\cdot\nabla), \text{ with } \hat{n}' = \frac{1}{\left[1+(\nabla_{\parallel} S)^2\right]^{1/2}}(-\nabla_{\parallel} S, 1) \tag{3.3.12}$$

$$dS' = dx'dz'\left[1+(\nabla_{\parallel} S)^2\right]^{1/2} \tag{3.3.13}$$

Integrating by parts using the identity:

$$\begin{aligned}&\nabla_{\parallel}\{\exp[i(\boldsymbol{\kappa}'-\boldsymbol{\kappa})\cdot\boldsymbol{\rho}' + i(\gamma_r-\gamma_t')S(\boldsymbol{\rho}')] - 1\} \\ &= i\left[\boldsymbol{\kappa}'-\boldsymbol{\kappa}+(\gamma_r-\gamma_t')\nabla_{\parallel} S\right]\times\exp[i(\boldsymbol{\kappa}'-\boldsymbol{\kappa})\cdot\boldsymbol{\rho}' + i(\gamma_r-\gamma_t')S(\boldsymbol{\rho}')]\end{aligned} \tag{3.3.14}$$

an integral equation for the transmitted field is obtained

$$\mathbf{e}_i(\boldsymbol{\kappa}) = \frac{\varepsilon_2-\varepsilon_1}{4\pi}\frac{1}{2\pi\gamma_r}\mathbf{k}_r^-\times\mathbf{k}_r^-\times\int d\boldsymbol{\kappa}'\mathbf{e}_t(\boldsymbol{\kappa}')\mathbf{I}(\boldsymbol{\kappa},\boldsymbol{\kappa}') \tag{3.3.15}$$

where the scattering potential $\mathbf{I}(\boldsymbol{\kappa},\boldsymbol{\kappa}')$ is

$$\mathbf{I}(\boldsymbol{\kappa},\boldsymbol{\kappa}') = \int d\boldsymbol{\rho}' \exp[i(\boldsymbol{\kappa}' - \boldsymbol{\kappa})\cdot\boldsymbol{\rho}'] \times \frac{\exp[i(\gamma_r - \gamma_t')S(\boldsymbol{\rho}')]}{\gamma_r - \gamma_t'} \qquad 3.3.16$$

and γ is a function of $\boldsymbol{\kappa}$, γ' is a function of $\boldsymbol{\kappa}'$, and

$$\mathbf{k}_{r,t}^{\pm} = (\boldsymbol{\kappa}, \pm\gamma_{r,t}) \qquad 3.3.17$$

Before solving the integral equation 3.3.15 we will introduce a method for separating the s and p components of the field, and in order to do this a local basis is introduced which depends upon $\boldsymbol{\kappa}$. Due to the transverse nature of the field the spectral components may be written in the following form:

$$\mathbf{e}_t(\boldsymbol{\kappa}) = \mathbf{e}_{ts}(\boldsymbol{\kappa})\hat{\mathbf{a}}_s(\mathbf{k}_t^-) + \mathbf{e}_{tp}(\boldsymbol{\kappa})\hat{\mathbf{a}}_p(\mathbf{k}_t^-) \qquad 3.3.18$$

where $\hat{\mathbf{a}}_s(\mathbf{k}_t^-)$ and $\hat{\mathbf{a}}_p(\mathbf{k}_t^-)$ are two unit vectors perpendicular to $\mathbf{k}_t^-$, and are defined as:

$$\hat{\mathbf{a}}_s(\mathbf{k}_t^-) = \hat{\mathbf{y}} \times \hat{\boldsymbol{\kappa}} \qquad 3.3.19$$

$$\hat{\mathbf{a}}_p(\mathbf{k}_t^-) = \hat{\mathbf{a}}_s(\mathbf{k}_t^-) \times \frac{\mathbf{k}_t^-}{|\mathbf{k}_t^-|} \qquad 3.3.20$$

and we then introduce e_{ts} and e_{tp} as the s and p components of the field respectively.

Using this basis the product, which appears as a linear operator, may be written in matrix form as:

$$\mathbf{R}(\boldsymbol{\kappa},\boldsymbol{\kappa}') = -k_0^2 \begin{bmatrix} \hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}' & -\frac{\gamma_t'}{nk_0}\hat{\boldsymbol{\kappa}}'\cdot(\hat{\mathbf{y}}\times\hat{\boldsymbol{\kappa}}) \\ -\frac{\gamma_r}{k_0}\hat{\boldsymbol{\kappa}}\cdot(\hat{\mathbf{y}}\cdot\hat{\boldsymbol{\kappa}}') & \frac{\boldsymbol{\kappa}\boldsymbol{\kappa}' + (\hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}')\gamma_r\gamma_t'}{nk_0^2} \end{bmatrix} \qquad 3.3.21$$

and we obtain a form of equation 3.3.15 for the s and p components of the transmitted field:

$$\begin{bmatrix} e_{is}(\boldsymbol{\kappa}) \\ e_{ip}(\boldsymbol{\kappa}) \end{bmatrix} = \frac{\varepsilon_2 - \varepsilon_1}{4\pi}\frac{1}{2\pi\gamma_r}\int d\boldsymbol{\kappa}'\mathbf{I}(\boldsymbol{\kappa},\boldsymbol{\kappa}')\mathbf{R}(\boldsymbol{\kappa},\boldsymbol{\kappa}')\begin{bmatrix} e_{ts}(\boldsymbol{\kappa}') \\ e_{tp}(\boldsymbol{\kappa}') \end{bmatrix} \qquad 3.3.22$$

It is this result which allows us to obtain an iterative series in order to obtain the e_{ts} and e_{tp} coefficients. Solutions for the generalised case of a rough surface described by the function S (defined in equation 3.3.1) are initially described, followed by the simplified case of a grating described by a purely sinusoidal profile.

Firstly the fields and the function $\mathbf{I}(\boldsymbol{\kappa},\boldsymbol{\kappa}')$ are expanded as follows:

$$e_t(\boldsymbol{\kappa}) = \sum_{n=0}^{\infty} \frac{e_t^{(n)}(\boldsymbol{\kappa})}{n!} \qquad 3.3.23$$

$$\mathbf{I}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') = \sum_{n=0}^{\infty} \frac{(i)^n (\gamma_r - \gamma_t')^{n-1}}{n!} S^n(\boldsymbol{\kappa}' - \boldsymbol{\kappa}) \qquad 3.3.24$$

with

$$S^n(\boldsymbol{\kappa}' - \boldsymbol{\kappa}) = \int d\boldsymbol{\rho}' \exp[i(\boldsymbol{\kappa}' - \boldsymbol{\kappa}) \cdot \boldsymbol{\rho}'] S^n(\boldsymbol{\rho}') \qquad 3.3.25$$

and where n is a number corresponding to an order which contributes to $e_t(\boldsymbol{\kappa})$, so that the sum over all n takes into account all possible orders (with wavevectors $\boldsymbol{\kappa}$). The total $e_t(\boldsymbol{\kappa})$ is therefore a sum over all of the plane waves, corresponding to all possible scattering events from the surface, in the half-space being investigated. In the iterative series below this n is used as a truncation parameter in the calculation, where only a sufficient number of orders are used to achieve convergence.

By introducing equations 3.3.24 and 3.3.25 into equation 3.3.22 and equating terms of the same order one obtains (for n > 0):

$$\begin{bmatrix} e_{ts}^n(\boldsymbol{\kappa}) \\ e_{tp}^n(\boldsymbol{\kappa}) \end{bmatrix} = \frac{\gamma_t - \gamma_r}{4\pi^2} \mathbf{R}^{-1}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') \qquad 3.3.26$$

$$\times \int d\boldsymbol{\kappa}' \mathbf{R}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') \sum_{q=1}^{n} \binom{n}{q} (i)^q (\gamma_r - \gamma_t')^{q-1} \mathbf{S}^q(\boldsymbol{\kappa}' - \boldsymbol{\kappa}) \begin{bmatrix} e_{ts}^{(n-q)}(\boldsymbol{\kappa}') \\ e_{tp}^{(n-q)}(\boldsymbol{\kappa}') \end{bmatrix}$$

where

$$\binom{n}{q} = \frac{n!}{q!(n-q)!} \qquad 3.3.27$$

and the zero order contribution is given by

$$\begin{bmatrix} e_{ts}^{(0)}(\kappa) \\ e_{tp}^{(0)}(\kappa) \end{bmatrix} = \begin{bmatrix} \frac{2\gamma_r}{\gamma_r + \gamma_t} & 0 \\ 0 & \frac{2n\gamma_r}{\varepsilon_2\gamma_r + \varepsilon_1\gamma_t} \end{bmatrix} \begin{bmatrix} e_{is}^{(0)} \\ e_{ip}^{(0)} \end{bmatrix} \qquad 3.3.28$$

which are the Fresnel coefficients for transmission through a planar surface.

The scattered fields in reflection can also be obtained using equations 3.3.10-15:

$$e_r(\boldsymbol{\kappa}) = \frac{\varepsilon_2 - \varepsilon_1}{4\pi}\frac{1}{2\pi\gamma_r}\mathbf{k}_r^+ \times \mathbf{k}_r^+ \times \int d\boldsymbol{\kappa}' e_t(\boldsymbol{\kappa}') \int d\boldsymbol{\rho}' \exp[i(\boldsymbol{\kappa}' - \boldsymbol{\kappa})\cdot\boldsymbol{\rho}'] \frac{\exp[-i(\gamma_r + \gamma_t')\mathbf{S}(\boldsymbol{\rho}')]}{\gamma_r + \gamma_t'} \quad 3.3.29$$

Using the operator $\mathbf{P}(\boldsymbol{\kappa},\boldsymbol{\kappa}')$:

$$\mathbf{P}(\boldsymbol{\kappa},\boldsymbol{\kappa}') = -k_0^2 \begin{bmatrix} \hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}' & -\frac{\gamma_t'}{nk_0}\hat{\boldsymbol{\kappa}}'\cdot(\hat{\mathbf{y}}\times\hat{\boldsymbol{\kappa}}) \\ \frac{\gamma_r}{k_0}\hat{\boldsymbol{\kappa}}\cdot(\hat{\mathbf{y}}\cdot\hat{\boldsymbol{\kappa}}') & \frac{\boldsymbol{\kappa}\boldsymbol{\kappa}' - (\hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}')\gamma_r\gamma_t'}{nk_0^2} \end{bmatrix} \quad 3.3.30$$

which is the equivalent operator for the reflected fields to $\mathbf{R}(\boldsymbol{\kappa},\boldsymbol{\kappa}')$ for the transmitted fields, the reflected field coefficients are given by:

$$\begin{bmatrix} e_{rs}^n(\boldsymbol{\kappa}) \\ e_{rp}^n(\boldsymbol{\kappa}) \end{bmatrix} = \frac{\varepsilon_2 - \varepsilon_1}{4\pi}\frac{\mathbf{1}}{2\pi\gamma_r} \times \int d\boldsymbol{\kappa}\mathbf{P}(\boldsymbol{\kappa},\boldsymbol{\kappa}')\sum_{q=0}^{n-1}\frac{n}{q}(i)^{(n-q)}\mathbf{S}^{(n-q)}(\boldsymbol{\kappa}'-\boldsymbol{\kappa})(\gamma_r+\gamma_t')^{n-q-1}\begin{bmatrix} e_{rs}^{(n-q)}(\boldsymbol{\kappa}') \\ e_{rp}^{(n-q)}(\boldsymbol{\kappa}') \end{bmatrix} \quad 3.3.31$$

for all values of n.

If the profile of the surface is defined as a pure sinusoid these equations are simplified and the iterative solutions are then given by

$$\begin{bmatrix} e_{ts}^n(\mathrm{m}) \\ e_{tp}^n(\mathrm{m}) \end{bmatrix} = \frac{k_0^2(\varepsilon_2-\varepsilon_1)}{(\gamma_t+\gamma_r)}\sum_{q=1}^{n}\binom{n}{q}(i)^q\frac{h^q}{2^q}\sum_{k=0}^{q}\binom{q}{k}(\gamma_r-\gamma_t')^{q-1} \times \mathbf{M}(m,m+q-2k)\begin{bmatrix} e_{ts}^{(n-q)}(\mathrm{m+q-2k}) \\ e_{tp}^{(n-q)}(\mathrm{m+q-2k}) \end{bmatrix} \quad 3.3.32$$

for the transmitted fields, with h being the amplitude of the grating,

$$\mathbf{M}(\mathrm{m},\mathrm{m+q-2k}) = \begin{bmatrix} \hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}' & -\frac{\gamma_t'}{nk_0}\hat{\boldsymbol{\kappa}}'\cdot(\hat{\mathbf{y}}\times\hat{\boldsymbol{\kappa}}) \\ \frac{-\gamma_r nk_0}{\boldsymbol{\kappa}^2+\gamma_r\gamma_t}\hat{\boldsymbol{\kappa}}\cdot(\hat{\mathbf{y}}\cdot\hat{\boldsymbol{\kappa}}') & \frac{\boldsymbol{\kappa}\boldsymbol{\kappa}' + (\hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}')\gamma_r\gamma_t'}{\boldsymbol{\kappa}^2+\gamma_r\gamma_t} \end{bmatrix} \quad 3.3.33$$

and

$$\boldsymbol{\kappa} = (\kappa_{ix} + m\kappa_g, \kappa_{iz}), \text{ and } \boldsymbol{\kappa}' = [\kappa_{ix} + (m+q-2k)\kappa_g, \kappa_{iz}] \quad 3.3.34$$

where m corresponds to the diffracted order of interest and κ_g is the grating vector defined as

$$\kappa_g = \frac{2\pi}{\lambda_g}\ ,\ \text{and } \lambda_g \text{ is the grating pitch}$$

The corresponding iterative series for the reflected fields is:

$$\begin{bmatrix} e_{rs}^{n}(\mathrm{m}) \\ e_{rp}^{n}(\mathrm{m}) \end{bmatrix} = \frac{k_0^2(\varepsilon_1-\varepsilon_2)}{2\gamma_r}\sum_{q=1}^{n}\binom{n}{q}(-i)^q\frac{h^q}{2^q}\sum_{k=0}^{q}\binom{q}{k}(\gamma_r+\gamma_t')^{q-1} \qquad 3.3.35$$
$$\times\mathbf{N}(m,m+q-2k)\begin{bmatrix} e_{ts}^{(n-q)}(\mathrm{m}+\mathrm{q}-2\mathrm{k}) \\ e_{tp}^{(n-q)}(\mathrm{m}+\mathrm{q}-2\mathrm{k}) \end{bmatrix}$$

With

$$\mathbf{N}(\mathrm{m},\mathrm{m}+\mathrm{q}-2\mathrm{k}) = \begin{bmatrix} \hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}' & -\frac{\gamma_t'}{nk_0}\hat{\boldsymbol{\kappa}}'\cdot(\hat{\mathbf{y}}\times\hat{\boldsymbol{\kappa}}) \\ \frac{\gamma_r}{k_0}\hat{\boldsymbol{\kappa}}\cdot(\hat{\mathbf{y}}\cdot\hat{\boldsymbol{\kappa}}') & \frac{\boldsymbol{\kappa}\boldsymbol{\kappa}'-(\hat{\boldsymbol{\kappa}}\cdot\hat{\boldsymbol{\kappa}}')\gamma_r\gamma_t'}{nk_0^2} \end{bmatrix} \qquad 3.3.36$$

[Note in equations 3.3.32 and 3.3.35 the i^q and $(-i)^q$ factors respectively. These will be of importance to the work in following chapters]

From these simple results the amplitude coefficients for whichever diffracted order is required can be simply obtained by using:

$$e_t(\mathrm{m}) = \sum_{n=0}^{\infty}\frac{e_t^{(n)}(\mathrm{m})}{n!} \qquad 3.3.37$$

$$e_r(\mathrm{m}) = \sum_{n=0}^{\infty}\frac{e_r^{(n)}(\mathrm{m})}{n!} \qquad 3.3.38$$

and the efficiencies are then obtained from:

$$\alpha_m = \frac{|e_{rs}(m)|^2+|e_{rp}(m)^2|}{|e_i|^2}\frac{\mathrm{Re}(\gamma_{rm})}{\mathrm{Re}(\gamma_i)} \qquad 3.3.39$$

$$\beta_m = \frac{|e_{ts}(m)|^2+|e_{tp}(m)^2|}{|e_i|^2}\frac{\mathrm{Re}(\gamma_{tm})}{\mathrm{Re}(\gamma_i)} \qquad 3.3.41$$

Therefore, in calculating the efficiency of any order from a sinusoidal grating these simple steps must be followed:

1) The coefficients e_{rs}, e_{rp}, e_{ts}, and e_{tp} are obtained iteratively from equations 3.3.32 and 3.3.35 to some desired order n, where the order required for convergence is dependent upon the grating profile, the dielectric function of

the media on either side of the interface, the polarisation of the incident light, and the orientation of the grating with respect to the incident light.

2) The sum of these coefficients for the diffracted order m being investigated is then obtained from equations 3.3.37 and 3.3.38.
3) The efficiency of the order is obtained from equations 3.3.39 and 3.3.41.

It is also simple to obtain the phase of these orders with respect to the incident light by taking the inverse tan of the imaginary part of the amplitude coefficient divided by the real part.

This model has been developed into a computer code and will be used throughout chapters 4 and 5.

3.4 The Differential Method of Chandezon

3.4.1 Introduction

The method of Chandezon *et al* [1980, 1982] is a differential method whose basic feature is that it uses a co-ordinate transformation that maps the interfaces of a multi-layer system onto parallel planes, thereby enabling the expression of the boundary conditions to be made much simpler, though at the same time making the solutions of Maxwell's equations more complex. This method is extremely flexible and can be used to model complex grating profiles such as multi-shaped multi-layer systems (Plumey, Granet and Chandezon [1995], Preist, Cotter, and Sambles [1995]) and also bigrating structures (though the computation times for these are prohibitive for all but very shallow gratings).

This method was first demonstrated in 1980 (Chandezon, Maystre, and Raoult [1980]) where the system was a single corrugated interface separating vacuum and a perfectly conducting metal, and was extended in 1982 to a multi-layer system with absorbing media (Chandezon, Dupuis, Cornet, and Maystre[1982]). This multi-layer method was based upon the use of transfer matrices in order to associate the eigenmodes across the interfaces. However, this method can become numerically unstable when the layer thickness become large, and therefore the use of scattering matrices was

introduced by Preist, Cotter, and Sambles [1995] in order to overcome this limitation (simultaneously an R-matrix technique was developed by Li [1995], which is almost identical to the scattering matrix method).

In this section the scattering matrix approach to the method of Chandezon *et al* will be described. This method has been developed and used within the research group for a number of years and, though some minor changes have been made for the work in this thesis, the general method has remained unchanged. For this reason, and the fact that a complete description of the method is beyond the scope of this work, the description will not be exhaustive. The description will be limited to the general process needed to describe the case of light incident in the classical mount, and for single shaped multi-layer structures. Brief descriptions of the extensions to the theory needed to describe gratings in the conical mount, and those for describing multi-shaped systems, will also be given.

3.4.2 The Method

The system to be solved is shown below:

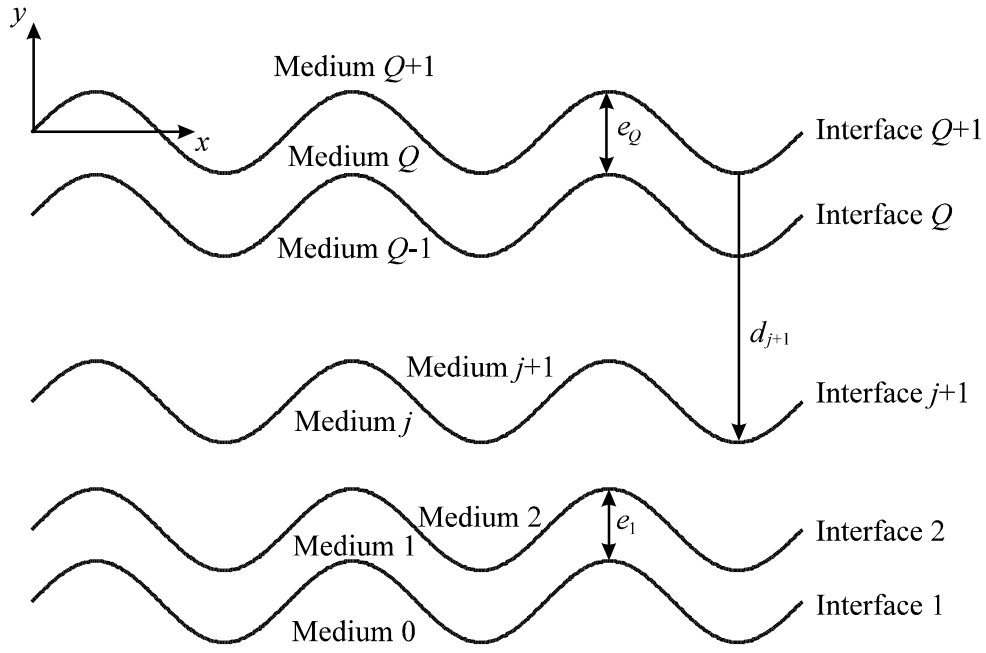

Figure 3.4.2.1 The multi-layer system

It consists of a series of Q layers (labelled j = 1,2…,Q) with relative permittivity, permeability, and mean thickness ε_r^j, μ_r^j, and e_j respectively. The lower semi-infinite medium has relative permittivity and permeability ε_r^0 and μ_r^0, and the corresponding values for the upper semi-infinite medium are denoted as ε_r^{Q+1} and μ_r^{Q+1}. By choosing the origin of y to be at the upper interface the equations for any interface j is given by

$$y_j = d_j + f(x) \qquad 3.4.2.1$$

and

$$d_j = -\sum_{n=j}^{Q} e_j \qquad 3.4.2.2$$

where the function $f(x)$ is the profile of all the interfaces.

A monochromatic polarised plane wave is incident on the system in the plane containing the grating vector (with polar angle θ, and azimuthal angle ϕ = 0º)

The co-ordinate frame used to solve the problem maps the grating profile onto a planar surface producing a new co-ordinate system (v,u,w) defined as:

$$v = x, \quad u = y - f(x), \quad w = z. \qquad 3.4.2.3$$

and two functions, C and D, are defined as

$$C = \frac{1}{1+f'(v)^2}$$

$$D = \frac{f'(v)}{1+f'(v)^2}. \qquad 3.4.2.4$$

Maxwell's equations are then obtained in the new co-ordinate system:

$$\frac{\partial E_3}{\partial u} - \frac{\partial\left(CE^2 + DE_1\right)}{\partial w} = ik_0\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\mu_r\left(CH_1 - DH^2\right) \qquad 3.4.2.5$$

$$\frac{\partial E_1}{\partial w} - \frac{\partial E_3}{\partial v} = ik_0\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\mu_r H^2 \qquad 3.4.2.6$$

$$\frac{\partial\left(CE^2 + DE_1\right)}{\partial v} - \frac{\partial E_1}{\partial u} = ik_0\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\mu_r H_3 \qquad 3.4.2.7$$

$$\frac{\partial H_3}{\partial u} - \frac{\partial\left(CH^2 + DH_1\right)}{\partial w} = -ik_0\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\varepsilon_r\left(CE_1 - DE^2\right) \qquad 3.4.2.8$$

$$\frac{\partial H_1}{\partial w} - \frac{\partial H_3}{\partial v} = -ik_0\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\varepsilon_r E^2 \qquad 3.4.2.9$$

$$\frac{\partial\left(CH^2 + DH_1\right)}{\partial v} - \frac{\partial H_1}{\partial u} = -ik_0\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\varepsilon_r E_3 . \qquad 3.4.2.10$$

where E_1, H_1, E_3, and H_3 are the tangential components of the fields, and E^2 and H^2 are the normal components of the fields.

Since only the tangential field components must be continuous across the interfaces the normal field components can be eliminated from these equations, producing four coupled equations:

$$\frac{\partial E_3}{\partial u} = \frac{\partial}{\partial w}\left[\frac{iC}{k_0\varepsilon_r}\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\left(\frac{\partial H_1}{\partial w} - \frac{\partial H_3}{\partial v}\right) + DE_1\right] \qquad 3.4.2.11$$
$$+ ik_0\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\mu_r\left[CH_1 + \frac{iD}{k_0\mu_r}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\left(\frac{\partial E_1}{\partial w} - \frac{\partial E_3}{\partial v}\right)\right]$$

$$\frac{\partial E_1}{\partial u} = \frac{\partial}{\partial v}\left[\frac{iC}{k_0\varepsilon_r}\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\left(\frac{\partial H_1}{\partial w} - \frac{\partial H_3}{\partial v}\right) + DE_1\right] - ik_0\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\mu_r H_3 \qquad 3.4.2.12$$

$$\frac{\partial H_3}{\partial u} = \frac{\partial}{\partial w}\left[\frac{-iC}{k_0\mu_r}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\left(\frac{\partial E_1}{\partial w} - \frac{\partial E_3}{\partial v}\right) + DH_1\right] \qquad 3.4.2.13$$
$$- ik_0\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\varepsilon_r\left[CE_1 - \frac{iD}{k_0\varepsilon_r}\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\left(\frac{\partial H_1}{\partial w} - \frac{\partial H_3}{\partial v}\right)\right]$$

$$\frac{\partial H_1}{\partial u} = \frac{\partial}{\partial v}\left[\frac{-iC}{k_0\mu_r}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\left(\frac{\partial E_1}{\partial w} - \frac{\partial E_3}{\partial v}\right) + DH_1\right] + ik_0\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\varepsilon_r E_3 . \qquad 3.4.2.14$$

Since only the system in the classical mount is being described there is no z-components of the fields meaning that the partial derivatives $\partial/\partial w$ are equal to zero. Therefore these equations become:

$$\frac{\partial E_3}{\partial u} = ik_0\mu_r C\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2} H_1 + D\frac{\partial E_3}{\partial v} \qquad 3.4.2.15$$

$$\frac{\partial E_1}{\partial u} = \frac{\partial}{\partial v}\left[\frac{-iC}{k_0\varepsilon_r}\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\frac{\partial H_3}{\partial v} + DE_1\right] - ik_0\mu_r\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2} H_3 \qquad 3.4.2.16$$

$$\frac{\partial H_3}{\partial u} = -ik_0\varepsilon_r C\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2} E_1 + D\frac{\partial H_3}{\partial v} \qquad 3.4.2.17$$

$$\frac{\partial H_1}{\partial u} = \frac{\partial}{\partial v}\left[\frac{iC}{k_0\varepsilon_r}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\frac{\partial E_3}{\partial v} + DH_1\right] + ik_0\varepsilon_r\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2} E_3 \qquad 3.4.2.18$$

where the surface shape is contained explicitly in the functions C and D.

By introducing the notation :

$$\text{TM: } F = H_3,\ G = -k_0\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}\varepsilon_r E_1 \qquad 3.4.2.19$$

$$\text{TE: } F = E_3,\ G = k_0\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}\mu_r H_1 \qquad 3.4.2.20$$

for the two polarisations (obtained due to the fact that $E_3 = 0$ and $H_1 = 0$ for TM polarisation, and $H_3 = 0$ and $E_1 = 0$ for TE polarisation) equations 3.4.2.15-18 may simplify to

$$\frac{\partial F}{\partial u} = D\frac{\partial F}{\partial v} + iCG \qquad 3.4.2.21$$

$$\frac{\partial G}{\partial u} = \frac{\partial}{\partial v}\left(DG + iC\frac{\partial F}{\partial v}\right) + ik_0{}^2\varepsilon_r\mu_r F\,. \qquad 3.4.2.22$$

which are valid for both polarisations.

It is possible to expand the fields by performing a Fourier expansion in the variable v due to the periodicity of the fields in this direction. This Bloch wave expansion takes the form:

$$F(v,u) = \sum_m F_m(u)\exp(i\alpha_m v) \qquad 3.4.2.23$$

$$G(v,u) = \sum_m G_m(u)\exp(i\alpha_m v) \qquad 3.4.2.24$$

where α_m is the in-plane wavevector given by:

$$\alpha_m = k\sin\theta + mK \qquad 3.4.2.25$$

The functions C and D (equation 3.2.4.2) are also periodic and can be expanded in a fourier series:

$$C(v) = \sum_p C_p \exp(ipKv) \qquad 3.4.2.26$$

$$D(v) = \sum_p D_p \exp(ipKv) \qquad 3.4.2.27$$

When these four expansions (equations 3.4.2.24-27) are substituted into equations 3.4.2.21,22 they become:

$$\sum_m \frac{dF_m}{du} e^{i\alpha_m v} = \sum_m \sum_p i\left(\alpha_m D_p F_m + C_p G_m\right) e^{i\alpha_{m+p} v} \qquad 3.4.2.28$$

$$\sum_m \frac{dG_m}{du} e^{i\alpha_m v} = \sum_m \sum_p i\alpha_{m+p}\left(D_p G_m - \alpha_m C_p F_m\right) e^{\alpha_{m+p} v} + \sum_m i k_0{}^2 \varepsilon_r \mu_r F_m e^{i\alpha_m v} . \qquad 3.4.2.29$$

These equations can then be simplified to a double infinite set of 1st order differential equations:

$$-i\frac{dF_n}{du} = \sum_m \left(\alpha_m D_{n-m} F_m + C_{n-m} G_m\right) \qquad 3.4.2.30$$

$$-i\frac{dG_n}{du} = \sum_m \left[\left(-\alpha_n \alpha_m C_{n-m} + k_0{}^2 \varepsilon_r \mu_r \delta_{nm}\right) F_m + \alpha_n D_{n-m} G_m\right]. \qquad 3.4.2.31$$

where the calculation is performed by truncating the infinite sums to a range ±N.

It is now desirable to find the field coefficients F_m and G_m in a form which lends itself to matching through the boundaries between layers. In order to do this a field vector ξ(u) is introduced, defined as:

$$\boldsymbol{\xi} = \left(F_{-N}, \ldots, F_N, \frac{G_{-N}}{\rho}, \ldots, \frac{G_N}{\rho}\right)^T \qquad 3.4.2.32$$

where ρ is given by $-k_0(\varepsilon_0/\mu_0)^{1/2}\varepsilon_r$ or $k_0(\mu_0/\varepsilon_0)^{1/2}\mu_r$ for TM or TE polarisation respectively. Using this field vector equations 3.4.2.30,31 can be written as an eigenvalue equation

$$-i\frac{d}{du}\boldsymbol{\xi}(u) = T\boldsymbol{\xi}(u) \qquad 3.4.2.33$$

where T is a square matrix constructed from four square sub-matrices:

$$T = \begin{pmatrix} \alpha_m D_{n-m} & C_{n-m} \\ -\alpha_n \alpha_m C_{n-m} + k_0{}^2 \varepsilon_r \mu_r \delta_{nm} & \alpha_n D_{n-m} \end{pmatrix}. \qquad 3.4.2.34$$

The truncated equations have eigensolutions of the form $\mathbf{v}_q e^{ir_q u}$, where r_q and $\mathbf{v}_q$ are the eigenvalues and eigenvectors of matrix T respectively. The field vector may then be expanded as a linear combination of these eigenmodes, and therefore:

$$\boldsymbol{\xi}(u) = M\phi(u)\mathbf{b} \qquad 3.4.2.35$$

where M is a square matrix composed of the eigenvectors $\mathbf{v}_q$

$$M = (\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{2(2N+1)}) \qquad 3.4.2.36$$

$\phi(u)$ is a diagonal matrix of components

$$\phi_{pq}(u) = \exp(ir_q u)\delta_{pq} \qquad 3.4.2.37$$

and $\mathbf{b}$ is a column vector of eigenmode amplitudes. For a given eigenmode, the corresponding eigenvector gives the relative electric and magnetic field strengths while the eigenvalue gives the field spatial dependence.

The corresponding equivalent to equation 3.4.2.35 for any j layer is:

$$\boldsymbol{\xi}^j(u) = M^j\phi^j(u-d)\mathbf{b}^j \qquad 3.4.2.36$$

The eigenmodes are then sorted as upward or downward going (either decaying or propagating) away from their interface of origin as follows:

$$M = (M_+ \quad M_-), \;\; \phi = \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix}, \;\; \mathbf{b} = \begin{pmatrix} \mathbf{b}_+ \\ \mathbf{b}_- \end{pmatrix} \qquad 3.4.2.37$$

where + denotes upward and – denotes downward propagation or decay.

The field coefficients F_m and G_m/ρ are now in a form to be matched across the interfaces since matching the field coefficients at the interfaces is equivalent to matching the field components tangential to the surface at the interfaces.

The concept of the scattering matrix is now introduced. In the original method proposed by Chandezon *et al* a transfer matrix approach was used in order to relate the fields in the upper and lower media by a recursive product of individual transfer matrices across each layer. The problem with a transfer matrix method is that there are a large number of evanescent decaying and growing fields across a layer, which leads to instabilities when the layer thickness becomes large. Cotter *et al* [1995] altered the method to use scattering matrices which are used in many fields of study which involve multi-layered systems.

The scattering matrix approach relates the incident fields to the output fields by a scattering matrix product, and avoids the instabilities inherent within the transfer matrix method, however the scattering matrix method does involve a few extra calculations and so the increase in stability is gained at a slight cost to the calculation time.

A scattering matrix S(0,Q+1) is introduced in order to relate the incident field eigenmodes amplitudes $\mathbf{b}_{-}^{Q+1}$ to the output field eigenmode amplitudes for the reflected and transmitted fields $\mathbf{b}_{+}^{Q+1}$ and $\mathbf{b}_{-}^{0}$:

$$\begin{pmatrix} \mathbf{b}_{+}^{Q+1} \\ \mathbf{b}_{-}^{0} \end{pmatrix} = S(0,Q+1) \begin{pmatrix} \mathbf{b}_{+}^{0} \\ \mathbf{b}_{-}^{Q+1} \end{pmatrix}. \qquad 3.4.2.38$$

and also a transfer matrix associated with the eigenmode amplitudes on either side of an interface bounded by the layer j, and j+1:

$$\begin{pmatrix} \mathbf{b}_{+}^{j} \\ \mathbf{b}_{-}^{j} \end{pmatrix} = I(j+1) \begin{pmatrix} \mathbf{b}_{+}^{j+1} \\ \mathbf{b}_{-}^{j+1} \end{pmatrix} \qquad 3.4.2.39$$

where I(j+1) is given by:

$$I(j+1) = \left[\phi^{j}(e_j)\right]^{-1} \left(M^{j}\right)^{-1} M^{j+1}. \qquad 3.4.2.40$$

which is obtained by matching the components of the fields tangential to the j+1st interface in the bordering media.

The definitions for the scattering matrix and the transfer matrix between the j+1st and jth media can be written as:

$$\begin{pmatrix} \mathbf{b}_{+}^{j} \\ \mathbf{b}_{-}^{j} \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} \begin{pmatrix} \mathbf{b}_{+}^{j+1} \\ \mathbf{b}_{-}^{j+1} \end{pmatrix} \qquad 3.4.2.41$$

$$\begin{pmatrix} \mathbf{b}_{+}^{j+1} \\ \mathbf{b}_{-}^{j} \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} \mathbf{b}_{+}^{j} \\ \mathbf{b}_{-}^{j+1} \end{pmatrix} \qquad 3.4.2.42$$

In fact the components of the scattering matrix S_{11}, S_{12} , S_{21}, S_{22} can be obtained from those of the transfer matrix I_{11}, I_{12}, I_{21}, I_{22} using the following relations:

$$S_{11} = (I_{11})^{-1} \qquad 3.4.2.43$$

$$S_{12} = -(I_{11})^{-1} I_{12}. \qquad 3.4.2.44$$

$$S_{21} = I_{21}(I_{11})^{-1} \qquad 3.4.2.45$$

$$S_{22} = I_{22} - I_{21}(I_{11})^{-1} I_{12}. \qquad 3.4.2.46$$

In a similar way the scattering matrix S(0,j+1) can be obtained from the scattering matrix S(0,j) and the transfer matrix I(j+1) using the following relations:

$$S_{11}(0, j+1) = \left[I_{11}(j+1) - S_{12}(0, j) I_{21}(j+1)\right]^{-1} S_{11}(0, j) \qquad 3.4.2.47$$

$$S_{12}(0,j+1)=\left[I_{11}(j+1)-S_{12}(0,j)I_{21}(j+1)\right]^{-1}\left[S_{12}(0,j)I_{22}(j+1)-I_{12}(j+1)\right]. \quad 3.4.2.48$$

$$S_{21}(0,j+1)=S_{22}(0,j)I_{21}(j+1)S_{11}(0,j+1)+S_{21}(0,j) \quad 3.4.2.49$$

$$S_{22}(0,j+1)=S_{22}(0,j)I_{21}(j+1)S_{12}(0,j+1)+S_{22}(0,j)I_{22}(j+1). \quad 3.4.2.50$$

Therefore, by induction, the scattering matrix for the whole system S(0,Q+1) may be generated from an initial unit matrix S(0,0). A method for relating the eigenmode amplitudes at each end of the system as in equation 3.4.2.28,42 has now been obtained, but this can actually be simplified further since $\mathbf{b}_{+}^{0}=0$ leaving only:

$$\mathbf{b}_{+}^{Q+1}=S_{12}\mathbf{b}_{-}^{Q+1} \quad 3.4.2.51$$

$$\mathbf{b}_{-}^{0}=S_{22}\mathbf{b}_{-}^{Q+1}. \quad 3.4.2.52$$

The reflection and transmission amplitude coefficients now need to be obtained. In order to do this the eigenmode expansion of the incident field is determined, which is needed to obtain the eigenmode amplitudes of the reflected and transmitted orders using equations 3.4.2.51,52. Only a general explanation of the method will be described here and the results needed for the calculation of the reflection and transmission amplitude coefficients as it is rather complex.

The incident field is defined as a plane propagating wave with a field of unit amplitude (H_z for TM polarised, and E_z for TE polarised) which is then transformed into the co-ordinate frame used to describe the grating system. Due to the periodicity of the incident field in the x-direction it may be expanded as a Fourier series and the eigenmode expansion of the incident field is obtained. A relationship between the eigenmode amplitudes and the plane wave amplitudes can then be found:

$$M_{-}^{Q+1}\mathbf{b}_{-}^{Q+1}=\mathbf{L}. \quad 3.4.2.53$$

where $\mathbf{L}$ is a column vector composed of two column vectors $\mathbf{L}$', and $\mathbf{L}$'' which themselves contain:

$$L_m(\beta) \quad 3.4.2.54$$

$$-\left[\beta-\left(\frac{\alpha}{\beta}\right)mK\right]L_m(\beta) \quad 3.4.2.55$$

The *propagating* reflected fields are also expanded in the same way and the relationship between the propagating reflected eigenmode amplitudes and the propagating reflected plane wave amplitudes are given by:

$$M_{a+}^{Q+1}\mathbf{b}_{a+}^{Q+1} = M'\mathbf{r}. \qquad 3.4.2.56$$

where the subscript a+ describes the fact that the fields are asymptotic (propagating) and that they are upward going. M' is a column vector which contains two column vectors M'', and M''' which contain:

$$L_{m-n}\left(-\beta_n^{Q+1}\right) \qquad 3.4.2.57$$

$$\left[\beta_n^{Q+1} - \left(\frac{\alpha_n}{\beta_n^{Q+1}}\right)(m-n)K\right] L_{m-n}\left(-\beta_n^{Q+1}\right) \qquad 3.4.2.58$$

The evanescent fields are then incorporated and the relationship between the total reflected field eigenmode amplitudes and their amplitudes is obtained:

$$M^{Q+1}\mathbf{b}^{Q+1} = M\begin{pmatrix}\mathbf{R}\\0\end{pmatrix} + \mathbf{L} \qquad 3.4.2.59$$

where

$$\mathbf{R} = \begin{pmatrix}\mathbf{r}\\ \mathbf{b}_{e+}^{Q+1}\end{pmatrix} \qquad 3.4.2.60$$

and $\mathbf{r}$ is a column vector containing the complex reflection amplitude coefficients of the propagating reflected orders, and $\mathbf{b}_{e+}^{Q+1}$ are the eigenmode amplitudes of the upward going evanescent orders.

Since equation 3.4.2.59 contains only square matrices it may be simply decomposed to give:

$$M_{11}^{Q+1}\mathbf{b}_{+}^{Q+1} + M_{12}^{Q+1}\mathbf{b}_{-}^{Q+1} = M_{11}\mathbf{R} + \mathbf{L}' \qquad 3.4.2.61$$

$$M_{21}^{Q+1}\mathbf{b}_{+}^{Q+1} + M_{22}^{Q+1}\mathbf{b}_{-}^{Q+1} = M_{21}\mathbf{R} + \mathbf{L}''. \qquad 3.4.2.62$$

and from equations 3.4.2.51,52 these may be described in terms of the scattering matrix element S_{12}:

$$A_1\mathbf{b}_{-}^{Q+1} = M_{11}\mathbf{R} + \mathbf{L}' \qquad 3.4.2.63$$

$$A_2\mathbf{b}_{-}^{Q+1} = M_{21}\mathbf{R} + \mathbf{L}'' \qquad 3.4.2.64$$

where

$$A_1 = M_{12}^{Q+1} + M_{11}^{Q+1}S_{12} \qquad 3.4.2.65$$

$$A_2 = M_{22}^{Q+1} + M_{21}^{Q+1}S_{12} \qquad 3.4.2.66$$

Elimination of $\mathbf{b}_{-}^{Q+1}$ in equations 3.4.2.63,64 leads to

$$\mathbf{R} = \begin{pmatrix} \mathbf{r} \\ \mathbf{b}_{e+}^{Q+1} \end{pmatrix} = \left(M_{11} - A_1 A_2^{-1} M_{21}\right)^{-1} \left(A_1 A_2^{-1} \mathbf{L}'' - \mathbf{L}'\right) \qquad 3.4.2.67$$

Therefore, the field amplitudes r can be calculated, and from these the reflectivity R is given by:

$$R^n = \left|r^n\right|^2 \frac{\cos\theta_n}{\cos\theta}. \qquad 3.4.2.68$$

where the n superscript corresponds to the reflected order being investigated, θ_n is the polar angle of the reflected order, and θ is the polar angle of the incident light.

The phase of the reflected order is given by:

$$\phi = \tan^{-1}\left(\frac{\mathrm{Im}(R)}{\mathrm{Re}(R)}\right) \qquad 3.4.2.69$$

The transmitted field is treated in a similar way and in this case a third A coefficient is found which is given by:

$$A_3 = S_{22}. \qquad 3.4.2.70$$

The equivalent result to equation 3.4.2.67 for the transmission amplitude coefficients is:

$$\mathbf{T} = \begin{pmatrix} \mathbf{t} \\ \mathbf{b}_{e-}^{0} \end{pmatrix} = \left(m_{12}\right)^{-1} M_{12}^0 A_3 \left(A_1\right)^{-1} \left(M_{11}\mathbf{R} + \mathbf{L}'\right). \qquad 3.4.2.71$$

where the transmissivities are given by:

$$T_{ss}^n = \left(\frac{\varepsilon_r^0 \mu_r^{Q+1}}{\varepsilon_r^{Q+1} \mu_r^0}\right)^{1/2} \left|t_{ss}^n\right|^2 \frac{\cos\theta_n^t}{\cos\theta} \qquad 3.4.2.72$$

$$T_{pp}^n = \left(\frac{\varepsilon_r^{Q+1} \mu_r^0}{\varepsilon_r^0 \mu_r^{Q+1}}\right)^{1/2} \left|t_{pp}^n\right|^2 \frac{\cos\theta_n^t}{\cos\theta}. \qquad 3.4.2.73$$

and the phases are given by the equivalent to equation 3.4.2.69 with only the transmission coefficient substituted for the reflection coefficient.

3.4.3 Extension to the Conical Mount

The extension of the theory necessary to convert the method described in section 3.4.2 to consider light incident upon a multi-layer grating system with non-zero azimuthal angles (the conical mount) is relatively simple, and has been described by

Popov and Mashev [1986], Elston, Bryan-Brown and Sambles [1991], and Plumey, Granet, and Chandezon [1995].

Firstly, Maxwell's equations in the transformed space are different to those in section 3.4.2 (equations 3.4.2.5–10). In the conical mount the partial derivatives $\partial / \partial w$ are set to zero, whereas in the conical mount they must be set to $i\gamma$, where γ is the component of the incident light wavevector in the z-direction. Once this is performed the method is analogous to that of the classical mount except that all four tangential field components are required (in the classical mount some components are zero depending upon the polarisation). This means that four coupled equations are obtained rather than the two necessary for the classical mount case (equations 3.4.2.21,22), and this means that all of the matrices involved in the solution have twice the dimensions of those used to describe gratings in the classical mount. The solution is then essentially the same as in section 3.4.2 with the reflection and transmission coefficients obtained from equations 3.4.2.67,71.

3.4.4 Extension to a Multi-Shape, Multi-Layer, Grating Theory

There are two main methods which have been developed in order to extend the differential theory of Chandezon to enable it to calculate the optical response of multi-layer systems comprising of layers with different surface shapes. The method of Granet Plumey, and Chandezon [1995] was the first to be published and involves matching the field components in the j^{th} layer to those in the $j+1^{st}$ layer at the $j+1^{st}$ interface where the eigensolutions of both layers are expressed in terms of the $j+1^{st}$ co-ordinate system (associated with their partitioning interface). Closely following this Preist, Cotter, and Sambles [1995] suggested a second method which involved a single direct step in the field matching. The method of Granet is more rigorous and has better convergence properties than that of Preist *et al*, however it is not as well suited for implementation in existing multi-layer codes. For this reason the method of Preist *et al* has been implemented in the modelling code used throughout this thesis and will be reviewed here. Only a general review of the method will be given followed by the results obtained since a full review is beyond the scope of this work.

Consider a multi-shaped multi-layer structure such as that in figure 3.4.4.1:

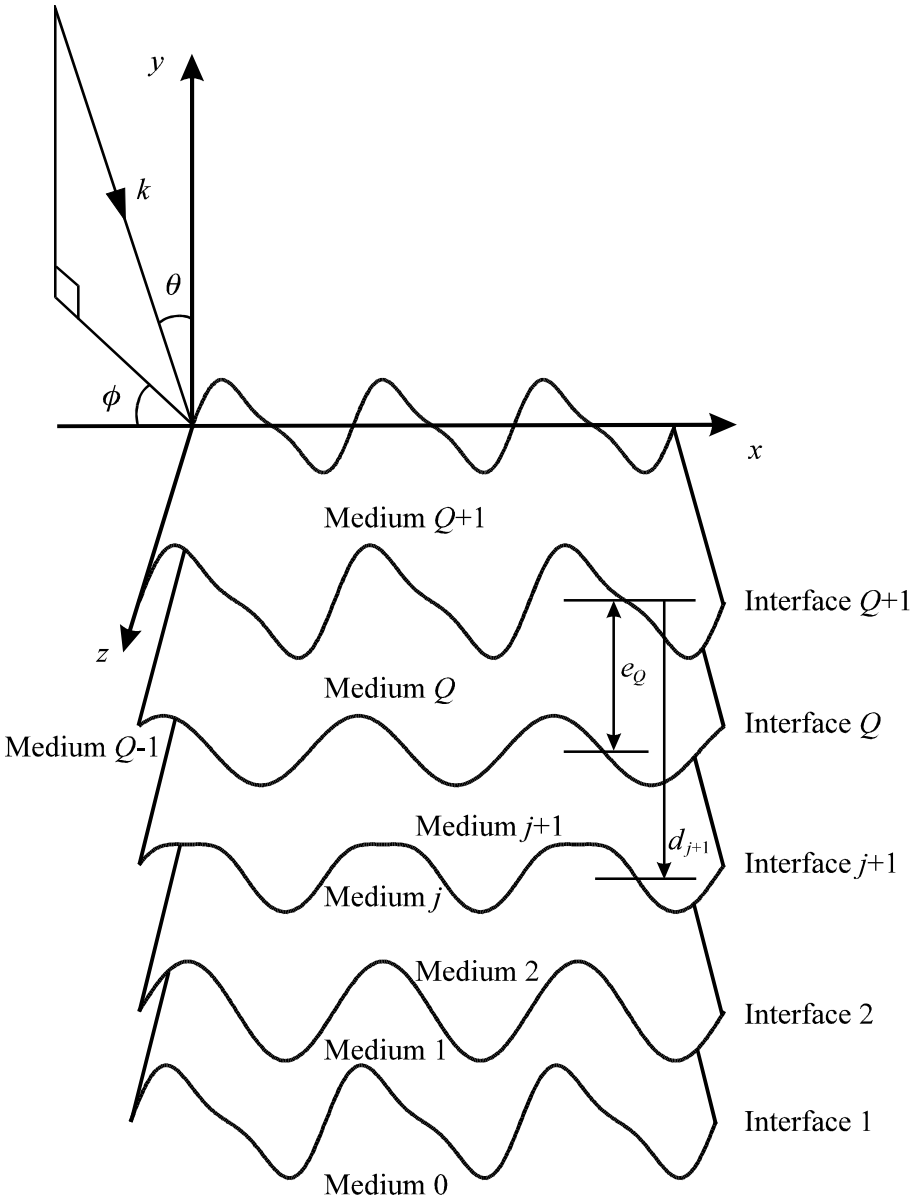

Figure 3.4.4.1 The multi-shape grating system.

For each layer j the calculation of the eigenvalues is performed in a co-ordinate system in which the jth interface is defined by:

$$v = x \qquad 3.4.4.1$$

$$u_j = y - f_j(x) \qquad 3.4.4.2$$

$$w = z \qquad 3.4.4.3$$

except for the 0th medium which is dealt with in the 1st interface co-ordinate system. The eigen solutions from each medium are expressed in the co-ordinate system in which its

lower interface is transformed into a planar surface. Therefore each co-ordinate system corresponds to an interface.

The multi-shape method of Preist *et al* involves matching the field components in an arbitrary layer j to those in the j+1st layer across the interface between them. Since the field components E_3 and H_3 (those tangential to the surface but orthogonal to the grating vector) are independent of the co-ordinate systems these are by far the easiest to match across the boundaries, and the matching condition is simply:

$$E_3^{j+1} = E_3^j \qquad 3.4.4.4$$

$$H_3^{j+1} = H_3^j . \qquad 3.4.4.5$$

The problem arises in matching the components E_1 and H_1 since these are the components tangential to the surface and in the plane containing the grating vector, and therefore are different for each interface. Therefore, it is necessary to include the components E^2 and H^2 (which are normal to the surface), which were discarded in our previous description where each interface had the same profile (section 3.4.2).

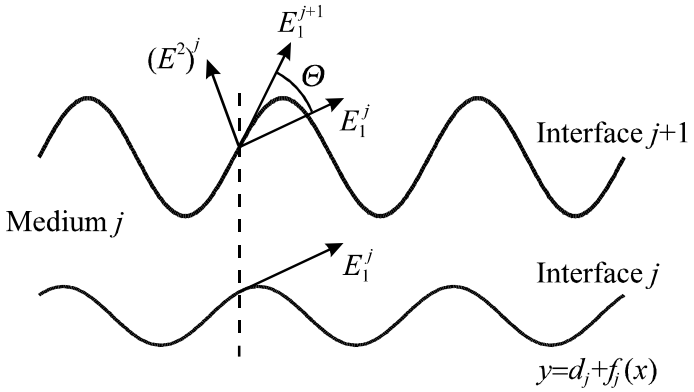

Figure 3.4.4.2 Orientation of the field components at the j+1st interface.

The θ in figure 3.4.4.2 is the angle between the tangents to the jth and j+1st interfaces. This can be accounted for in the first derivatives of the grating profiles, and if this is done the matching condition at the j+1st interface becomes:

$$E_1^{j+1} = E_1^j + \Delta_j' \left[C^j \left(E^2 \right)^j + D^j E_1^j \right] \qquad 3.4.4.6$$

where

$$\Delta_j(v) = f_{j+1}(v) - f_j(v) \qquad 3.4.4.7$$

$$\Delta_j'(v) = \frac{d}{dv}\Delta_j(v) = f_{j+1}'(v) - f_j'(v) \qquad 3.4.4.8$$

and

$$C^j(v) = \frac{1}{\left[1 + f_j'(v)^2\right]^{1/2}} \qquad 3.4.4.9$$

$$D^j(v) = \frac{f_j'(v)}{\left[1 + f_j'(v)^2\right]^{1/2}} \qquad 3.4.4.10$$

The H_1 component can be treated in exactly the same way giving:

$$H_1^{j+1} = H_1^j + \Delta_j'\left[C^j\left(H^2\right)^j + D^j H_1^j\right]. \qquad 3.4.4.11$$

Having obtained these it is necessary to obtain the matching conditions in terms of the eigenmode representation.

By substituting the Bloch wave expansions into the matching conditions of the field components E_3 and H_3, and substituting the eigenmode expansions of the field coefficients, the matching condition for these components is obtained:

$$\sum_q \left(M_2\right)_{mq}^{j+1} b_q^{j+1} = \sum_q \left(X_2\right)_{mq}^j \exp\left(ir_q^j e_j\right) b_q^j \qquad 3.4.4.12$$

$$\sum_q \left(M_4\right)_{mq}^{j+1} b_q^{j+1} = \sum_q \left(X_4\right)_{mq}^j \exp\left(ir_q^j e_j\right) b_q^j \qquad 3.4.4.13$$

where the matrix of the eigenvectors M^j is defined as :

$$M^j = \begin{pmatrix} \left(M_1\right)^j \\ \left(M_2\right)^j \\ \left(M_3\right)^j \\ \left(M_4\right)^j \end{pmatrix} \qquad 3.4.4.14$$

with $(M_1)^j$ corresponding to the field component E_1, $(M_2)^j$ corresponding to the field component E_3, $(M_3)^j$ corresponding to the field component H_1, $(M_4)^j$ corresponding to the field component H_3, and

$$\left(X_\eta\right)_{nq}^j = \sum_m \left(M_\eta\right)_{mq}^j \left[L_{n-m}^{j+1}\left(-r_q^j\right) + L_{n-m}^j\left(r_q^j\right)\right] \qquad 3.4.4.15$$

with η=2,4 and

$$L_n^j(s) = \frac{1}{\lambda_g}\int_0^{\lambda_g} \exp i\left[-sf_j(v) - nKv\right] dv. \qquad 3.4.4.16$$

with all other terms in equations 3.4.4.12,13,15,16 defined in the same way as in section 3.4.2.

The matching condition for the E_1 and H_1 field components must now be considered. In the same manner in which the matching condition for the field components E_3 and H_3 were performed, but incorporating the relations in equations 3.4.4.6,11:

$$\sum_q (M_1)_{mq}^{j+1} b_q^{j+1} = \sum_q (X_1)_{mq}^{j} \exp\left(ir_q^j e_j\right) b_q^j \qquad 3.4.4.17$$

for the E_1 component, and

$$\sum_q (M_3)_{mq}^{j+1} b_q^{j+1} = \sum_q (X_3)_{mq}^{j} \exp\left(ir_q^j e_j\right) b_q^j \qquad 3.4.4.18$$

for the H_1 component with

$$\begin{aligned}(X_1)_{nq}^{j} &= \sum_m (M_1)_{mq}^{j}\left[L_{n-m}^{j+1}\left(-r_q^j\right) + L_{n-m}^{j}\left(r_q^j\right)\right] \qquad 3.4.4.19\\ &+ \sum_m \sum_p (M_1)_{mq}^{j} \frac{D_p^j}{r_q^j}(n-m-p)K\left[L_{n-m-p}^{j+1}\left(-r_q^j\right) + L_{n-m-p}^{j}\left(r_q^j\right)\right]\\ &- \sum_m \sum_p (M_3)_{mq}^{j} \frac{\gamma C_p^j}{k_0 \varepsilon_r^0 r_q^j}\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}(n-m-p)K\left[L_{n-m-p}^{j+1}\left(-r_q^j\right) + L_{n-m-p}^{j}\left(r_q^j\right)\right]\\ &+ \sum_m \sum_p (M_4)_{mq}^{j} \frac{\alpha_m C_p^j}{k_0 \varepsilon_r^j r_q^j}\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2}(n-m-p)K\left[L_{n-m-p}^{j+1}\left(-r_q^j\right) + L_{n-m-p}^{j}\left(r_q^j\right)\right]\end{aligned}$$

and

$$\begin{aligned}(X_3)_{nq}^{j} &= \sum_m (M_3)_{mq}^{j}\left[L_{n-m}^{j+1}\left(-r_q^j\right) + L_{n-m}^{j}\left(r_q^j\right)\right] \qquad 3.4.4.20\\ &+ \sum_m \sum_p (M_3)_{mq}^{j} \frac{D_p^j}{r_q^j}(n-m-p)K\left[L_{n-m-p}^{j+1}\left(-r_q^j\right) + L_{n-m-p}^{j}\left(r_q^j\right)\right]\\ &+ \sum_m \sum_p (M_1)_{mq}^{j} \frac{\gamma C_p^j}{k_0 \mu_r^0 r_q^j}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}(n-m-p)K\left[L_{n-m-p}^{j+1}\left(-r_q^j\right) + L_{n-m-p}^{j}\left(r_q^j\right)\right]\\ &- \sum_m \sum_p (M_2)_{mq}^{j} \frac{\alpha_m C_p^j}{k_0 \mu_r^j r_q^j}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2}(n-m-p)K\left[L_{n-m-p}^{j+1}\left(-r_q^j\right) + L_{n-m-p}^{j}\left(r_q^j\right)\right]\end{aligned}$$

The four matching conditions for the four field components (equations 3.4.4.12,13,17,18) can be written as a single matrix equation:

$$M^{j+1}\mathbf{b}^{j+1} = X^j \phi^j\left(e_j\right)\mathbf{b}^j \qquad 3.4.4.21$$

where

$$X^j = \begin{pmatrix} (X_1)^j \\ (X_2)^j \\ (X_3)^j \\ (X_4)^j \end{pmatrix} \qquad 3.4.4.22$$

from which a transfer matrix $I(j+1)$ may be calculated:

$$I(j+1) = \left[\phi^j(e_j)\right]^{-1}\left(X^j\right)^{-1} M^{j+1} \qquad 3.4.4.23$$

This transfer matrix may then be used to obtain the scattering matrix for the system, and from this the reflectivities and transmissivities. The way in which this may be done is the same as that described in 3.4.2.

3.4.5 Calculating the Dispersion of the Modes of the System

In the previous sections the way in which the reflectivities and transmissivities of systems can be calculated was described. It is also useful to be able to calculate the dispersion of the various electromagnetic modes of the systems, and though the reflectivities and transmissivities can give information regarding these, they may only be calculated within the region of ω-k space available to the incident radiation.

Another way to obtain the dispersion curves of the optical modes is to directly investigate the scattering matrices of the system as a function of ω and k, since the information regarding the dispersion of the optical modes will be contained within them, and these will give the desired information for all values of ω and k, and not just those available to the incident radiation. If the sum of the scattering matrices is plotted for each point a peak will be observed when the optical modes are excited, and therefore by plotting these the dispersion of the modes may be obtained

3.5 Testing the Codes

The computer code based upon the differential method of Chandezon has been used for fitting experimental data obtained on metal grating structures in the optical region of the spectrum for a number of years, and has also been fully tested for a number of other systems (Watts, Sambles and Harris [1997], Watts, Sambles, Hutley, Preist and

Lawrence [1997], Watts, Sambles and Hutley [1998], Hibbins, Sambles and Lawrence [1998]). The parameters describing the system which are fitted to are; the real and imaginary parts of the dielectric functions for the layers, and the profiles (described by a truncated Fourier series) of the interfaces.

Since these tests have confirmed the validity of the Chandezon method it will not be tested here. However, the Iterative Series Solution (ISS) method has not been used in this way, and therefore some comparisons between the ISS code and that based on the Chandezon method will be performed to show that, within its region of convergence, the two methods produce the same results.

The most stringent test for shallow grating structures is to use a metal as the lower semi-infinite medium in the region where a surface plasmon polariton (see chapter 4) is excited. Therefore, to demonstrate the accuracy of the ISS method the optical response of single interface metal structures will be investigated. Figure 3.5.1 shows the zero-order and 1st diffracted order reflected intensities as a function of frequency for TM polarised light normally incident upon a 7.5nm amplitude sinusoidal grating with a grating pitch (λ_g) of 600nm (classical mount). The top semi-infinite medium is described as vacuum ($\varepsilon_r = 1.0$, $\varepsilon_i = 0.0$), and the lower semi-infinite medium is described as silver, where the frequency dependent dielectric function of the silver is described with separate polynomials for the real and imaginary parts, which were obtained by fitting to experimentally determined values (see Palik [1985]).

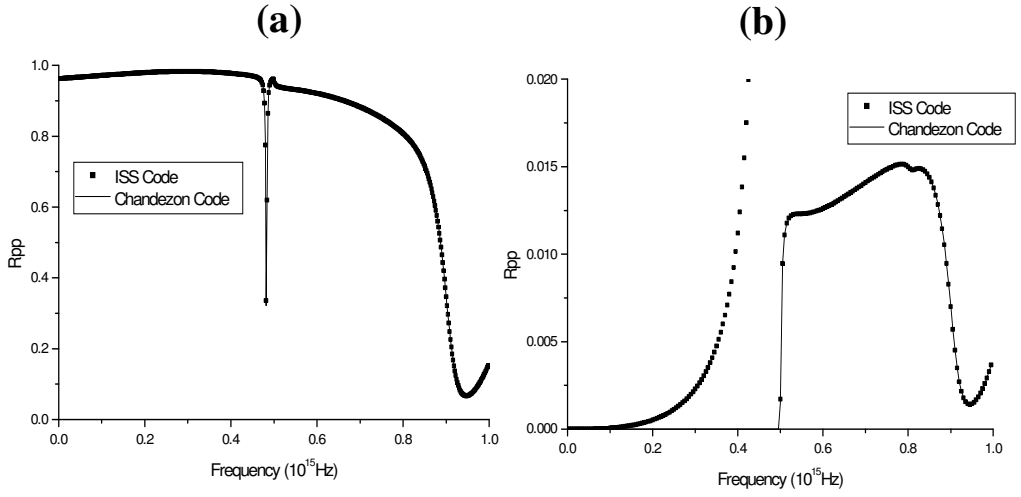

Figure 3.5.1 A comparison between the codes based upon the ISS method and the Chandezon method. TM reflectivity from a 7.5 nm amplitude 600nm pitch sinusoidal silver grating as a function of frequency (normal incidence, classical mount). a) the specularly reflected order, and b) the +1 diffracted order.

The zero-order reflectivity clearly shows that the two codes produce the same answers for this structure. However, the +1 diffracted order only agrees above a frequency of 0.5 x 10^{15}Hz. This is the frequency above which the diffracted order is a real propagating wave and not an evanescent order. Since the ISS code is based upon the Rayleigh hypothesis, which treats the evanescent orders in the same way as the propagating ones, intensities of the evanescent orders are found which have no physical meaning. However, the fact that the magnitudes and phases of the complex amplitude coefficient of the evanescent waves can also be determined from the ISS code will be used later in chapters 3 and 5.

A comparison of the zero-order reflectivity as a function of in-plane wavevector is shown in figure 3.5.2 for the same system as used in figure 3.5.1, and with incident light of frequency 0.4 x 10^{15}Hz. This shows that the two codes produce the same results for non-normal angles of incidence.

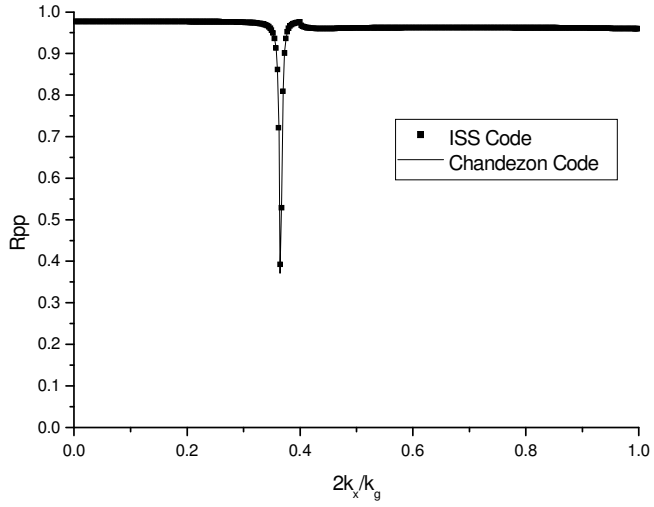

Figure 3.5.2 A comparison between the codes based upon the ISS method and the Chandezon method. TM reflectivity from a 7.5 nm amplitude 600nm pitch sinusoidal silver grating as a function of the in-plane wavevector for light of frequency 0.4 x 10^{15}Hz.

3.6 Summary

In this chapter the main methods used in calculating the optical response of grating structures have been described. The Iterative Series Solution (ISS) method based upon work by Greffet and Massaranni [1990] has then been described in some detail. This

method may only be used for single interface structures, and is limited to only very shallow corrugations due to its use of the Rayleigh method for describing the reflected and transmitted orders in terms of plane wave expansions. However, it is used due to its simplicity, and due to the fact that it is a powerful tool in understanding some of the phenomena observed on such systems.

This has been followed by a description of the differential method of Chandezon et al [1982] for multi-layer grating systems. This has been covered in some detail for the case of gratings oriented in the classical mount, and with identical profiles for all the interfaces. The way in which this method may be extended to the conical mount, and also how it may be extended to describe systems where the profiles of the interfaces in the multi-layer system are no longer all identical (multi-shape) has then been described in a general way.

The code based upon the method of Chandezon has been stringently tested previously, and therefore no additional testing of this method was performed. However the ISS method has not been tested before, and in order to do this it has been shown that, within the region of convergence of the ISS code, it produces the same answers as the Chandezon code.

Chapter 4

The Optical Response of Shallow Dielectric Gratings and of Thin Dielectric Grating Slabs

4.1 Introduction

Having described the methods for modelling the optical response of grating structures in chapter 3, in this chapter these computer codes will be used to calculate the optical response of dielectric grating structures. In the first section, the optical properties of the simplest surface relief grating structure possible, that of the shallow dielectric grating in the classical mount with normally incident light, will be investigated. The efficiencies and phases of the various reflected and transmitted diffracted orders (produced using the computer code based upon the ISS method) will be calculated for the case of TM polarised light incident from a lower index medium to a higher index medium, and from a higher index medium to a lower index medium. The analysis of non-blazed single interface dielectric structures for normally incident light appears to have been relatively ignored in the literature, presumably because the efficiencies of the diffracted orders are very small compared to those obtained from metallic gratings. The work which has been performed has tended to be based on developing computational methods to determine the efficiencies of the orders from such structures (see, for example, Knop [1978(a)], Pai and Awada [1991]), or on orientations whereby effective blazing can occur (Moharam and Gaylord [1982]). Also extensive work has been devoted to the possible applications of dielectric gratings, such as: beam-coupling to guided waves (described later in this chapter), in displays (Knop [1978(b)]), optical filtering (Tibuleac and Magnusson [1997], Kim and Fonstad (1979)), in monochromators (Flodstrom and Bachrach [1976]), and in laser applications such as Q-switching (Chesler, Karr and Geusic [1970]), and mode locking (Johnson [1973]). For a full review of the possible applications of dielectric gratings see Gaylord and Moharam [1985].

There are two main results from our analysis, which will be of use in the second section of this chapter, and also in chapter 5. Firstly, the phase of the ±1 diffracted order (either reflected or transmitted) with respect to the incident light may be either 0° or 180° (when the order is real and propagating) and that this depends upon whether the light is incident from the high index medium, or the low index medium, and that when both reflected and transmitted diffracted orders are evanescent their phase is the same as when both orders were real and propagating. Secondly, when a diffracted order in the semi-infinite medium which has a lower refractive index becomes evanescent, the intensity of the corresponding diffracted order in the higher refractive index medium (which is still a real propagating diffracted order) becomes zero. Correspondingly when the diffracted order in the higher index medium becomes evanescent the magnitude of the amplitude coefficient of the evanescent diffracted order in the lower refractive index medium becomes zero.

In the second section of the chapter thin dielectric grating slabs, which are corrugated on both interfaces, are investigated using the computer code based upon the method of Chandezon. For these structures, which are of major use in communications, the majority of previous work has involved the behaviour of waveguide modes in the dielectric material of which the slab is made (Dakss, Kuhn, Heidrich and Scott [1970], Tamir [1979], Avrutsky, Svakhin and Sychugov [1989], Li [1995], for a review of waveguide modes and their applications see Tien [1977]). Rather than investigate the waveguide modes of the system, the effect of the corrugated slab structure on the diffracted orders is investigated, which, though they are affected by the waveguide modes, will predominantly depend upon the interaction of the diffracted orders from the two interfaces. For this reason waveguide modes will only be described in a brief and general way.

It will be shown that, if the two surfaces are corrugated conformally, the transmitted diffracted order fields are reduced to almost zero. However, if the phase of the lower interface grating is changed relative to the grating on the top interface this is no longer the case, and the system may in fact be described as effectively blazed. Previous work has only considered the blazing effect on the waveguide modes (Yamasaki [1995], Peng and Tamir [1974]), and not on the diffraction efficiencies of the system. A further unexpected phenomenon occurs when the corrugations on the two surfaces are in anti-phase with each other, and under these conditions it is possible for

almost all of the energy to be distributed equally between the ±1 transmitted diffracted orders, with the intensity of the reflected and transmitted zeroth orders, and of the reflected diffracted orders, reduced to almost zero.

4.2 Single Interface Dielectric Gratings

In this section the optical response of single interface dielectric diffraction gratings will be investigated, the applications of which are briefly discussed in the introduction. Though it is the two interface systems described later in this chapter which are of most interest, the optical response of single interface dielectric gratings are of use in introducing the concepts which will be necessary for discussing the optical responses of these more complex grating systems.

4.2.1 The Efficiencies and Phases of the Diffracted orders from Single Interface Dielectric Gratings

When light is incident upon a grating it may gain (or lose) integer values of the grating vector ($k_g = 2\pi / \lambda_g$) in the x direction via scattering from the surface, and this is the origin of the diffracted orders. If light is incident at a polar angle θ_i in the classical mount (the incident light is in the plane containing the grating vector) the grating equation is

$$\lambda_g (\sin\theta_m - \sin\theta_i) = m\lambda_0 \qquad 4.2.1.1$$

where m is an integer which corresponds to the diffracted order of interest, θ_m is the angle at which the diffracted order propagates and λ_0 is the wavelength of the incident light. In this chapter the study will be limited to the case of normally incident light, where $\theta_i = 0$. The system is shown in figure 4.2.1.1.

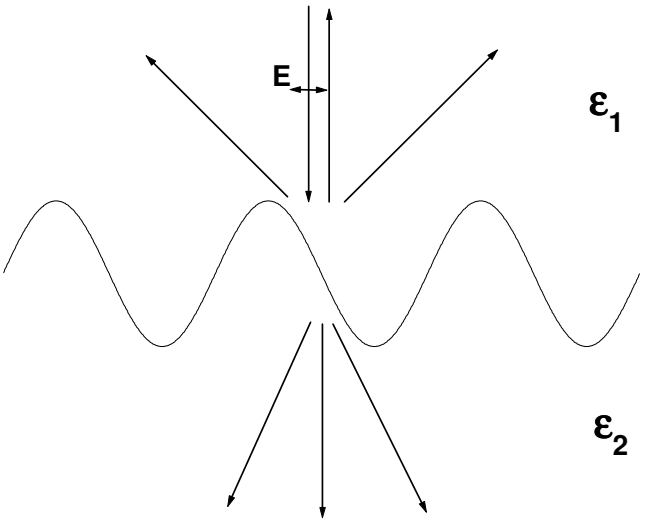

Figure 4.2.1.1 The system under consideration in this section. TM polarised light incident upon an interface between two dielectrics with dielectric functions ε_1 for the upper (incident) medium, and ε_2 for the lower (transmitted) medium. The light is normally incident in the classical mount (the azimuthal angle is 0°)

The efficiencies and phases of the various orders may be calculated using the methods described in chapter 3, although in this section only the zeroth and +1 orders (the –1 diffracted orders are identical to the +1 diffracted orders since we are considering normal incidence) will be considered. The ISS method is used since, in addition to the efficiencies and phases of the real propagating orders, it can also give us information about the evanescent orders (due to the fact that it utilises the Rayleigh hypothesis). A 400nm pitch (λ_g), 25nm amplitude, sinusoidal grating, will be studied, with the two dielectrics on either side of the grating interface described as air and SiO_2 (which is considered to be non-absorbing), with the frequency dependent dielectric function of the SiO_2 described by a polynomial fitted to experimentally determined values. The efficiencies and phases of the zero and +1 orders for reflection and transmission for light incident from the air side of the interface are shown in figure 4.2.1.2.

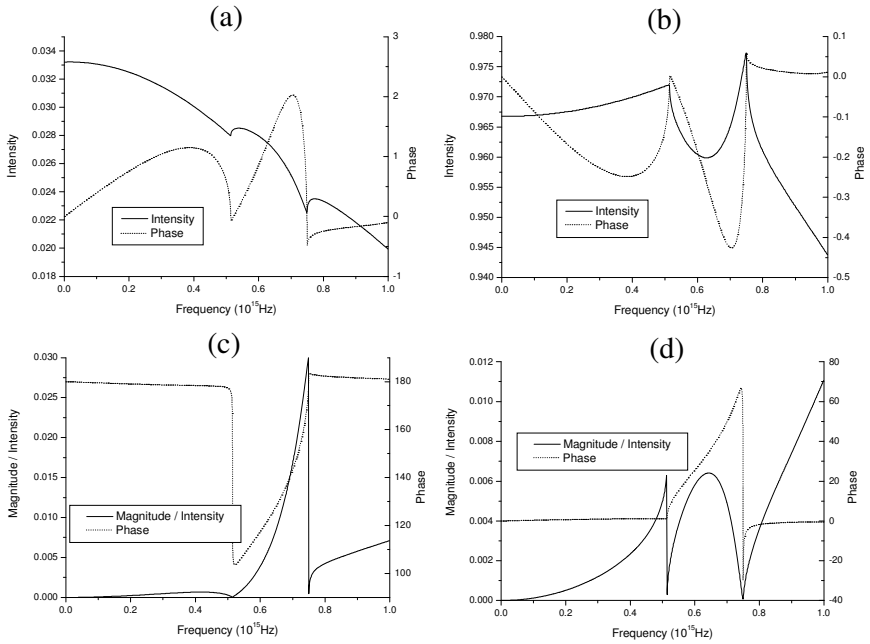

Figure 4.2.1.2 The optical response of an air / SiO_2 grating interface. The four orders presented are: a) the zeroth reflected, b) the zeroth transmitted, c) the +1 reflected diffracted, and d) the +1 transmitted diffracted. When the order is real and propagating the intensity of the order is shown, whereas when it is evanescent the magnitude ($\sqrt{\mathrm{Re}(\mathbf{r_p})^2 + \mathrm{Im}(\mathbf{r_p})^2}$) is shown.

The first points to note from figure 4.2.1.2 are the phases of the diffracted orders. There are three main regions: above f ≈ 0.75 x 10^{15}Hz the diffracted orders in both the air and the silica are real propagating orders, between f ≈ 0.5 x 10^{15} Hz the diffracted order in the SiO_2 is real and propagating, whereas in the air the diffracted order is evanescent, and for below f ≈ 0.5 x 10^{15} Hz both the diffracted orders in the air and the SiO_2 are evanescent. In the region where both diffracted orders are real and propagating the reflected diffracted order is 180° out of phase with the incident light, but the transmitted diffracted order is in phase with the incident light. The origin of this difference can be seen by examination of the expressions used in the ISS method for the diffracted order amplitude coefficients which, to first order, are,

$$e_{tp}^{1}(1) = \frac{i(\varepsilon_2 - \varepsilon_1)ak_0^3\gamma_r\sqrt{\varepsilon_2}e_{tp}^{0}(0)}{2(\gamma_t + \gamma_r)(k_g^2 + \gamma_r\gamma_t)} \qquad 4.2.1.2$$

for the +1 transmitted diffracted order, where a is the amplitude of the grating, ε_1 is the dielectric function of the incident medium, and ε_2 is the dielectric function of the transmission medium. The remaining variables are defined in the same way as in chapter 3. The corresponding equation for the reflected 1st diffracted order amplitude coefficient is:

$$e_{rp}^{1}(1) = \frac{i(\varepsilon_1 - \varepsilon_2)k_0 a e_{tp}^{0}(0)}{4}\left[\frac{(k_g^2 - \gamma_r\gamma_t)(\gamma_t - \gamma_r)}{(k_g^2 + \gamma_r\gamma_t)(\gamma_t + \gamma_r)} + 1\right] \qquad 4.2.1.3$$

(Note: The first term inside the square brackets is always positive, or negative and less than 1, and therefore the value of the expression inside the square brackets is always positive).

These expressions can be used to explain all of the features in the calculations of the efficiencies of the diffracted orders shown in figure 4.2.1.2. We shall first discuss the transmitted diffracted order. In the region in which both transmitted and reflected diffracted orders are real and propagating all of the variables in equation 4.2.1.2 are real and positive. Therefore the amplitude coefficient is purely real with its sign given by the $(\varepsilon_2 - \varepsilon_1)$ factor, which in the case of an air / SiO_2 interface is positive, so it appears that the phase of the transmitted diffracted order should be +90° due to the i factor.

This is not, in fact, the case since these expressions are for a grating profile described by a cosine, and therefore at y = 0, x = π/2. If the phase of the incident light is defined as zero at y = 0 then the phase of the diffracted orders requires a correction of 90° in order for it to be with respect to the incident light (in practice the i factor in equation 4.2.1.2 can be ignored). However, this is only true of the ±1 diffracted orders. For higher orders the i factor becomes important in calculating the phase, and it is also important when considering the effect of higher order contributions in the iterative series. This is all taken into account in the modelling code, and therefore the phases shown in figure 4.2.1.2 are correct. Later in this chapter, and also in chapter 5, the expressions for the diffracted order intensities shown above will be used in order to describe the diffraction processes, and in all these cases this 90° change is necessary.

At the low frequency limit of this region the reflected diffracted order becomes evanescent, and therefore $\gamma_r = 0$. It is clear from equation 4.2.1.2 that at this frequency the amplitude coefficient for the transmitted diffracted order reduces to zero. At lower frequencies we enter the region where the reflected order is evanescent, but the transmitted diffracted order is real and propagating. When this occurs γ_r becomes purely imaginary which results in a discontinuous shift in the phase of the transmitted diffracted order, with the amplitude coefficient being almost totally imaginary. However, the amplitude coefficient is no longer zero, meaning that the transmitted diffracted order increases in intensity once more. As the frequency is lowered further

the amplitude coefficient becomes progressively more real, since γ_r becomes larger and γ_t becomes smaller as the transmitted diffracted order nears the horizon. The phase lowers towards 0° again as a result of this. The intensity of the diffracted order reduces to zero as γ_t becomes zero, however this is not due to the amplitude coefficient becoming zero, but is rather due to the area factor needed to correct the intensity (which is given by the ratio of the z components of the wavevector of the diffracted and incident orders).

For lower frequencies this area factor is no longer required as we can no longer discuss the intensity of the diffracted order. In figure 4.2.1.2 we show the magnitude of the amplitude coefficient for these lower frequencies and this is the reason that there is a discontinuity when the transmitted order becomes evanescent. γ_t becomes imaginary in this final region where both diffracted orders are evanescent. Therefore the denominator in equation 4.2.1.2 becomes purely imaginary, and since the numerator is also purely imaginary the result is that the diffracted order is purely real and positive, and the phase of the diffracted order is 0° with respect to the incident light.

The reflected diffracted order case is somewhat different. The first term in the square brackets in equation 4.2.1.4 is always either positive, or negative and less than one (depending upon the ratio of γ_r to γ_t, which in turn depends upon the ratio of ε_2 to ε_1), and therefore the total expression inside the square brackets is always positive. The phase of the reflected diffracted order when both γ_r and γ_t are positive (when the transmitted and reflected diffracted orders are real and propagating) is therefore dependent upon the $(\varepsilon_1 - \varepsilon_2)$ term since all of the other parameters are real and positive. For the case of an air / SiO_2 interface this term is negative, and therefore the reflected diffracted order is 180° out of phase with the incident light.

We shall now consider the region where both diffracted orders are evanescent. The γ_t and γ_r factors are then purely imaginary, and since both the denominator and numerator are purely imaginary this produces a purely real amplitude coefficient. The same discussion as before means that the phases of the diffracted orders in this region are the same as in the region where both diffracted orders are real and propagating.

Finally, there is the region where the diffracted order in the air is evanescent, but the diffracted order in the SiO_2 is real and propagating. The phase of the reflected

diffracted order changes throughout the frequency region in which this is the case, from the amplitude coefficient being purely imaginary (the diffracted order is 90° out of phase with the incident), to the case where it is purely real and gives the same result as in the other two regions where the two diffracted orders have the same character. This is because the factors $\gamma_t + \gamma_r$ and $\gamma_t - \gamma_r$ are complex with both real and imaginary parts, and therefore the amplitude coefficient changes from being purely imaginary at the point in which the reflected order becomes evanescent (γ_r is imaginary, and γ_t is zero), to purely real when both orders become evanescent. Having discussed the phases of the diffracted orders it is important to note that the discussion using equations 4.2.1.2, 4 is only to first order and neglects the contribution to the diffracted order fields due to the higher orders. This is the reason that the phases shown in figure 4.2.1.2 are not exactly the values described in the discussion, but are slightly altered due to these higher order contributions.

We shall now proceed to discuss the magnitude of the complex amplitude coefficient (which in the case of the order being real and propagating corresponds to the intensity of the diffracted order). Again the plots can be divided into the three main regions described above. In the case of the reflected diffracted order it is only a real propagating diffracted order above f ≈ 0.75 x 10^{15}Hz, and below this the order is evanescent. For the transmitted diffracted order the mode is real and propagating above f ≈ 0.5 x 10^{15} Hz, and for lower frequencies it is evanescent. The surprising result from figure 4.2.1.2 is that the nature of the diffracted order in one of the media has a strong effect on the diffracted order in the other medium. The intensity of the reflected diffracted order reduces to zero at the point at which it becomes evanescent, however, the magnitude of the amplitude coefficient is not zero (as described above). The interesting point to note is that at the frequency at which the reflected diffracted order becomes evanescent the magnitude of the transmitted diffracted order, which is still real and propagating, reduces to zero. Correspondingly the intensity of the transmitted diffracted order reduces to zero when it becomes evanescent, and at this frequency the magnitude of the complex amplitude coefficient of the evanescent reflected diffracted order also reduces to zero.

This behaviour is clear from equations 4.2.1.2,3 but these do not give a physical interpretation for why this phenomenon occurs. The reason is relatively simple. We

previously described the process of diffraction in terms of an integer multiple of k_g being added to, or subtracted from, the wavevector of the incident light. Another way of looking at this is the way in which the diffraction grating problem is solved computationally. The boundary conditions at the surface (that the relevant field components are continuous across the boundary) demand that diffracted orders exist in order for them to be satisfied. The diffracted order whose intensity / magnitude reduces to zero when the diffracted order in the other half space becomes evanescent is no longer needed to satisfy the boundary condition, and this is the reason this phenomenon occurs. In fact it is clear from figure 4.2.1.2 that this is the case. If we consider the phase of the various orders relative to the incident light at the frequency at which one of the diffracted order become evanescent we see that the zeroth orders and the order which becomes evanescent are all in phase or out of phase with the incident light. In other words they all have their maximum field magnitude at the same point at the surface. The diffracted order in the other half space at this frequency is 90° out of phase with all of these components at this frequency, and therefore would have no field magnitude at the point at which the others go through a maximum even if the amplitude coefficient were not zero at this point.

In figure 4.2.1.3 we show a series of field profiles for the system described above for different frequencies. We show the time averaged H_z component of the fields (the z direction is into the page) since the z component of the H-fields is the only component for TM polarised light. Since we are using the time averaged fields the propagating fields are averaged out of the plots. This leaves the beating between the incident and reflected zeroth order, and also the diffracted orders which are still observed since they form a beating with the propagating zeroth and incident / reflected orders.

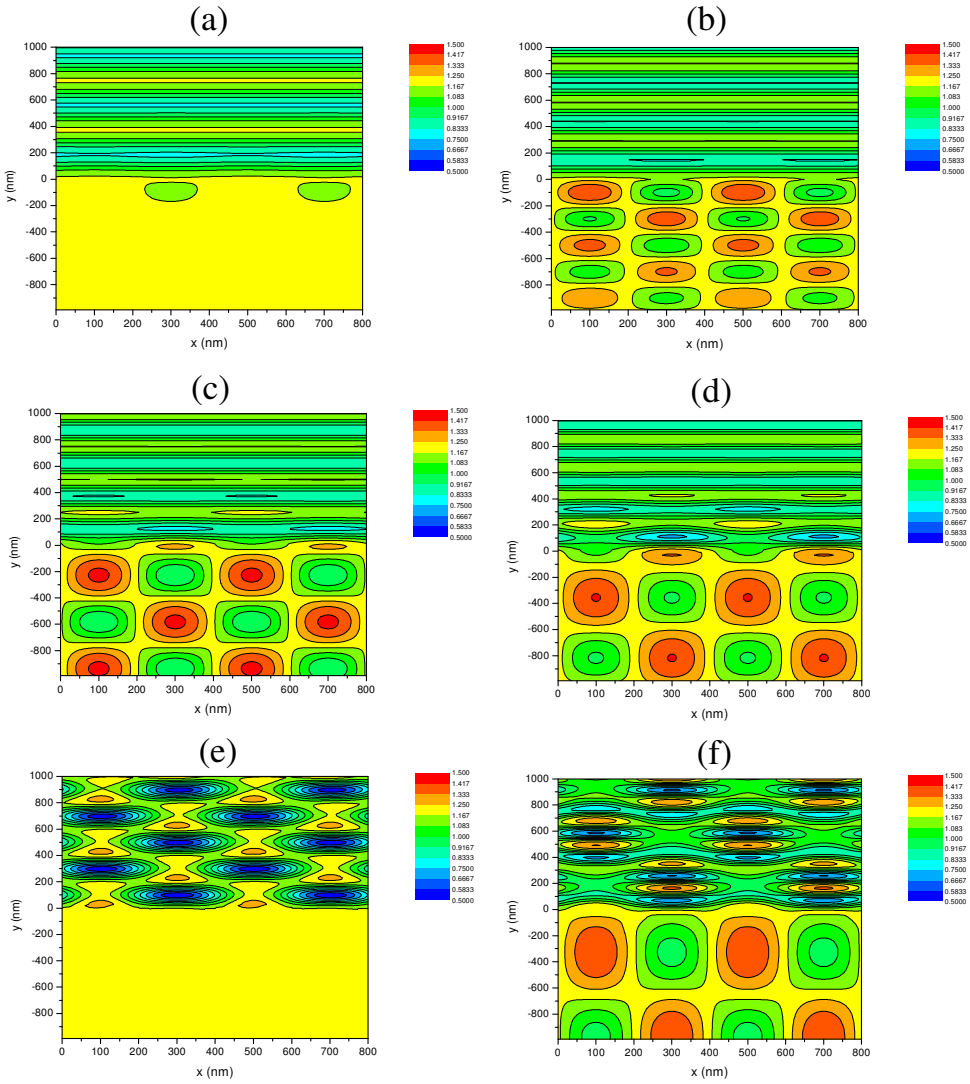

Figure 4.2.1.3 Time averaged H_z component of the fields for the air / SiO_2 interface used for figure 4.2.1.2 for six different frequencies: a) f = 0.4 x 10^{15}Hz, b) f = 0.514 x 10^{15}Hz, c) f = 0.6 x 10^{15}Hz, d) f = 0.7 x 10^{15}Hz, e) f = 0.75 x 10^{15}Hz, and f) f = 0.9 x 10^{15}Hz.

We shall first discuss the reflected fields, starting with figure 4.2.1.3(f). At this frequency (f = 0.9 x 10^{15}Hz) the reflected diffracted order is real and propagating, and the beating between the diffracted order and the incident and reflected zeroth orders is clear in the plot. When the frequency is reduced (figure 4.2.1.3(e), f = 0.75 x 10^{15}Hz), the diffracted order becomes evanescent. However, the magnitude of the amplitude coefficient does not reduce to zero, even though the intensity of the diffracted order does, and for this reason the fields due to the diffracted order are still present. Below this frequency (figures 4.2.1.3 (d), (c), (b), and (a), f = 0.7, 0.6, 0.514, 0.4 x 10^{15}Hz) the reflected diffracted order is evanescent, and this can be seen in the fact that the

magnitude of the fields due to diffraction reduce with distance from the surface. In fact, as the frequency is reduced the maximum magnitude of the evanescent diffracted order fields (at the surface) reduces (as is obvious from figure 4.2.1.2), but also the rate at which the evanescent order decays away from the surface increases. This is because the distance at which the evanescent orders have reduced to 1/e of their original value is proportional to $1/\gamma_r$, and as γ_r becomes larger and more imaginary as this distance reduces.

We shall now consider the case of light incident upon an identical profile to that considered above, but with light incident from the SiO_2 side of the system. (figure 4.2.1.4)

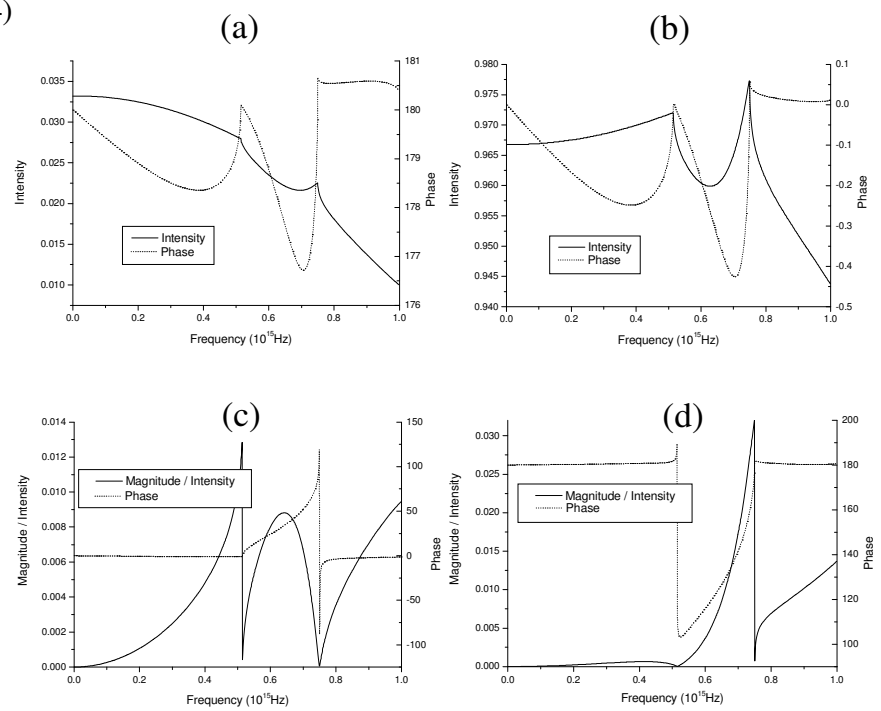

Figure 4.2.1.4 The optical response of a SiO_2 / Air grating interface, with the same parameters as for the air / SiO_2 interface considered above. The four orders presented are: a) the zeroth reflected, b) the zeroth transmitted, c) the +1 reflected diffracted, and d) the +1 transmitted diffracted

The results for this system are very similar to those for the air / SiO_2 interface. There are two main differences to note. Firstly, the behaviour of the magnitude / intensity of the diffracted orders are reversed. (In other words, the frequencies at which the diffracted orders for reflection and transmission become evanescent are reversed due to the reversal of the media). Otherwise the behaviour is the same as before in that when one of the diffracted orders becomes evanescent the amplitude coefficient of the other diffracted order reduces to zero. The other point to note is that the phases of the

diffracted orders are also reversed. This is due to the changing in sign of the $(\varepsilon_1 - \varepsilon_2)$ and $(\varepsilon_2 - \varepsilon_1)$ factors in equations 4.2.1.2 and 4.2.1.3 respectively.

The phase of the reflected zeroth order has also changed from 0° to 180°. However, the transmitted zeroth order is identical to the air / SiO_2 case. For completeness the fields of the same points determined for the air / SiO_2 interface case are shown for the SiO_2 / air case in figure 4.2.1.5.

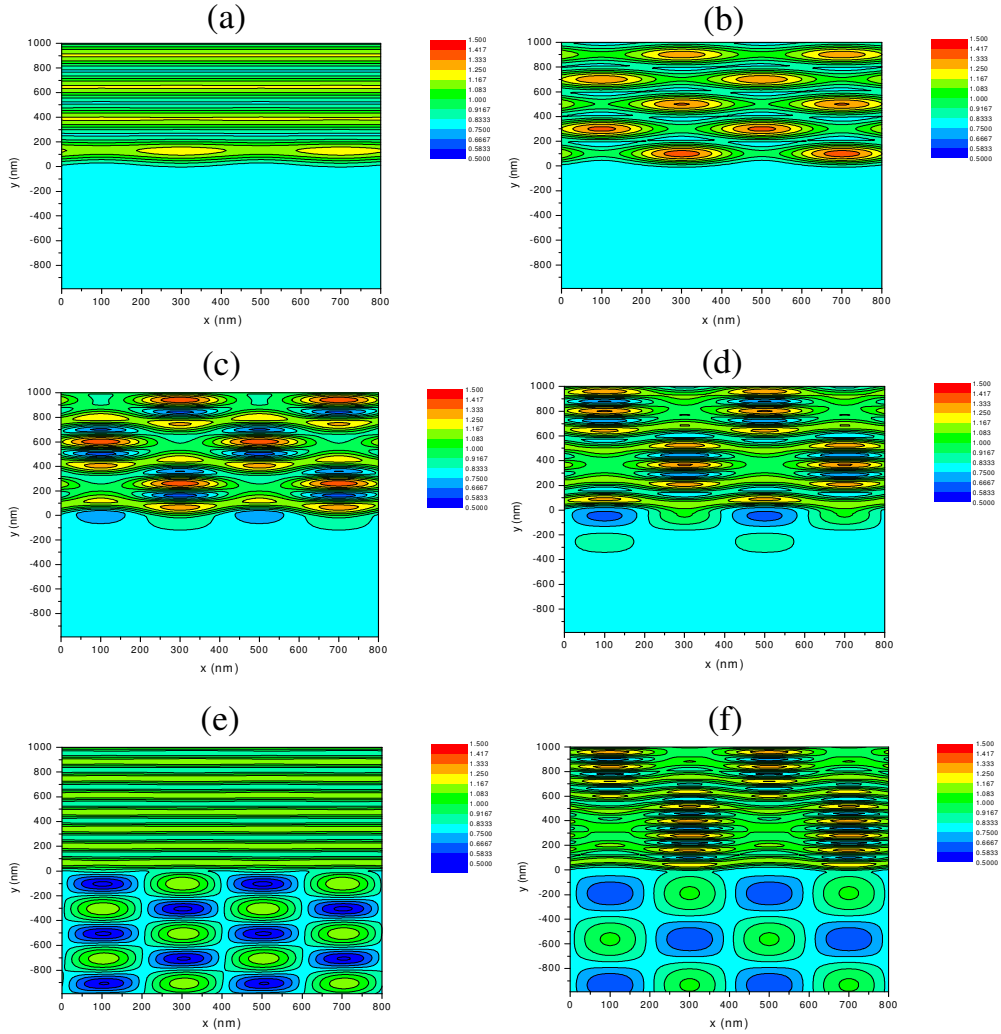

Figure 4.2.1.5 Time averaged H_z component of the fields for the SiO_2 / Air interface used for figure 4.2.1.4 for six different frequencies: a) f = 0.4 x 10^{15}Hz, b) f = 0.514 x 10^{15}Hz, c) f = 0.6 x 10^{15}Hz, d) f = 0.7 x 10^{15}Hz, e) f = 0.75 x 10^{15}Hz, and f) f = 0.9 x 10^{15}Hz.

The results shown in figure 4.2.1.5 mirror those of the air / SiO_2 case shown in figure 4.2.1.3. The differences are all due to the reversing of the media, and therefore the reversing of the frequencies at which the reflected and transmitted diffracted orders become evanescent.

Having investigated the optical response of simple single interface dielectric gratings, we shall now proceed to investigate thin corrugated dielectric slabs, where the results obtained in this section will be used to explain the optical responses obtained from such systems.

4.3 Thin Corrugated Dielectric Slabs

In this section we shall describe the optical response of thin corrugated dielectric slabs, and some surprising results which, we believe, have not been reported in the literature before.

Typically, thin dielectric slabs are used because of their waveguiding properties, and are especially useful for communications applications (for further applications and references see the introduction to this chapter). However, we will not be describing the waveguiding properties of these structures in any detail, but will rather discuss the effect of having the two interfaces close together upon the diffracted orders of the system. Therefore, we shall only give a brief introduction to waveguide modes, since the effects of these modes on the optical response of the structure will be evident in our results.

Following this we shall describe the case of the corrugated dielectric slab in a conformal geometry (both interfaces are corrugated identically), before discussing the asymmetric corrugated slab, where both interfaces are corrugated, but the corrugations are phase shifted in the x-direction with respect to each other.

Throughout this section we shall investigate the diffraction efficiencies by considering the two grating interfaces separately. The phases for the diffracted orders from the two interfaces are given in figures 4.2.1.2,4.

4.3.1 Waveguide modes

A waveguide is a structure which confines electromagnetic waves to a certain region of space. The simplest waveguide is that of the planar dielectric slab, and in this section we shall briefly describe this waveguide system followed by the way in which a grating allows coupling to the waveguide modes of a thin corrugated dielectric slab for normally incident light.

The simplest model of guided waves in a planar dielectric slab is that of the 'Ray model' as first described by Tien, Ulrich and Martin [1969]. The system under consideration is shown in figure 4.3.1.1.

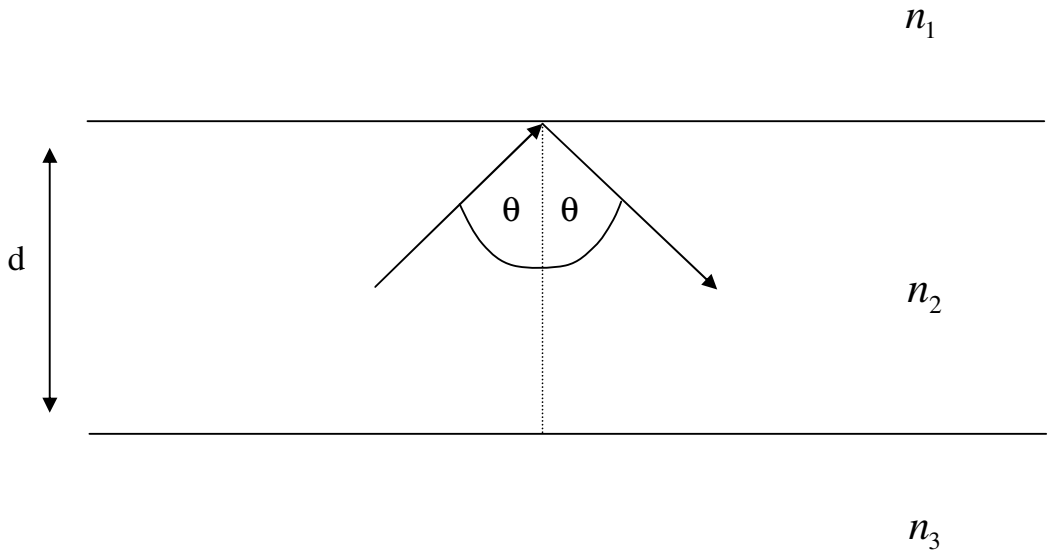

Figure 4.3.1.1 The planar dielectric waveguide.

A light ray inside the guiding medium is incident upon the upper planar interface where the dielectric above the interface has a refractive index lower than that of the guiding medium If it is incident at an angle beyond the critical angle (given by $\sin\theta_c = n_1 / n_2$) it is totally internally reflected. This reflected ray is then totally internally reflected at the lower interface if it is incident at beyond the critical angle determined from the refractive index of the lower medium. Repeated reflections lead to the wave propagating along the waveguiding medium.

This description is very much simplified, since there are other factors to be considered. The fields of the wave propagating in the positive y-direction within the guiding medium will interfere with the fields propagating in the negative y-direction, and these may interfere either constructively or destructively with each other. In the ray model constructive interference will occur when an integer number of wavelengths are traversed between the reflection from one interface and the reflection from the other interface, and it is under these circumstances that power may resonantly propagate along the waveguide. For this reason a propagating mode will only occur at certain values of the incident angle θ beyond the critical angle. The values which θ can take are

given by a phase matching condition which can be shown to be (e.g. Lorrain, Corson and Lorrain [1988]):

$$k_0 n_2 \cos\theta - \phi_{21} - \phi_{23} = m\pi \qquad 4.3.1.1$$

where k_0 is the wavevector of the light, m is an integer, and the two ϕ factors are phase changes upon reflection at the interfaces.

Since light may not escape from this mode it is also true that light may not couple into the mode. This is essentially due to a momentum mismatch between the incident light and the waveguide mode. There are various methods for overcoming this momentum mismatch, but the one of interest to the work in this chapter is that of grating coupling.

By periodically corrugating one of the surfaces, normally incident light from above the structure may be scattered directly into the waveguide mode if the corrugation period is chosen such that the momentum matching condition is obeyed for the transmitted scattered (diffracted) order. The diffracted order which gives rise to the guided mode is, in fact, evanescent, since this is the way in which the momentum is sufficiently increased in order for the excitation criteria to be satisfied. However, the light may, of course, also couple out via the reverse process. For this reason, if we observe the zeroth order reflection and transmission from the structure we observe resonant features which are the result of interference between the light which has been directly reflected from, or transmitted through, the structure, and that of the light which has been re-scattered from the waveguide mode back into the reflected and transmitted zeroth orders. The shape of these features depend upon the phase difference between the directly reflected / transmitted light and that of the re-radiated light which has coupled out of the waveguide, and also the intensity of the light which is coupled out of the waveguide. It is these features which will be observed in the results obtained in the following sections.

4.3.2 The Conformal Thin Dielectric Slab

The system which will be studied in this section is that of a thin corrugated wave guide of thickness d, corrugated identically on both interfaces with a grating of pitch λ_g and amplitude a. The refractive indices of the media are such that all the media are considered as non-absorbing with the refractive index of the slab having a higher

refractive index than that of the bounding media. We shall consider only TM polarised light normally incident upon the structure (figure 4.3.2.1).

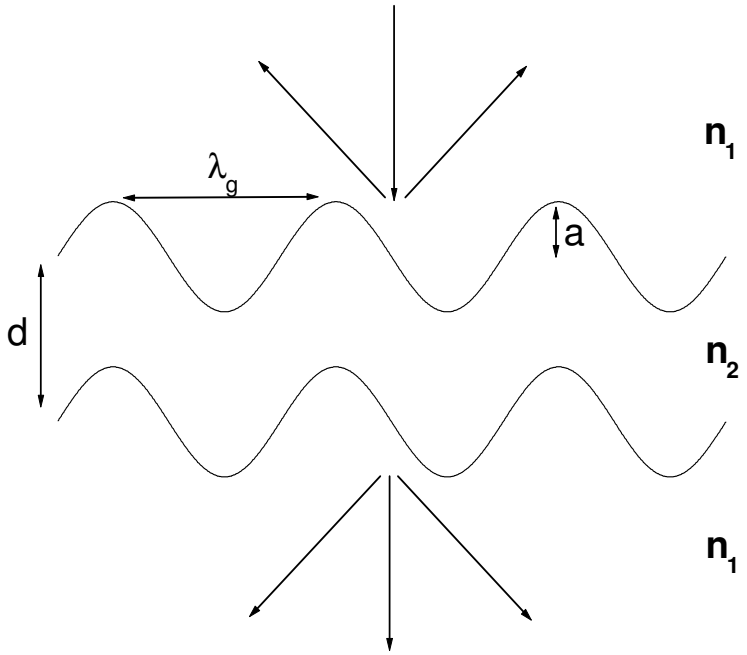

Figure 4.3.2.1 A schematic of the system under consideration.

In figure 4.3.2.2 we show the intensities of the reflected and transmitted zeroth orders, as well as the reflected and transmitted +1 orders (the –1 order is identical since we are only considering normal incidence), as a function of the thickness of the dielectric slab, and of the frequency of the incident light. The grating pitch is 400nm, the amplitude of the gratings is 25nm, and the bounding dielectrics are considered as air, with the slab described as SiO_2 (whose frequency dependent dielectric function is described by a polynomial fitted to experimentally derived values).

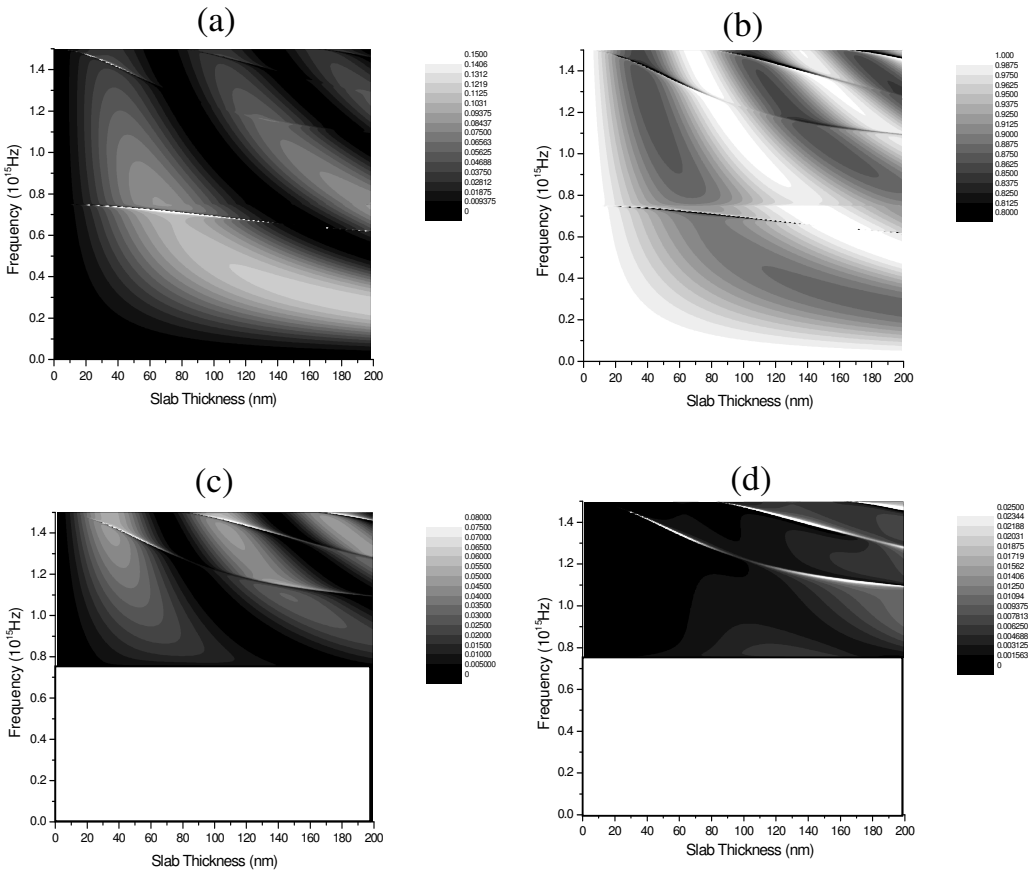

Figure 4.3.2.2 The optical response of a conformally corrugated thin SiO_2 slab, with air as the bounding media (λ_g = 400nm, a = 25nm), as a function of the frequency of incident TM polarised light and of the slab thickness. a) the zeroth order reflectivity, b) the zeroth order transmisivitty, c) the +1 order reflectivity, and d) the +1 order transmissivity.

We shall first describe the two different types of features evident in figure 4.3.2.2 before discussing in more detail some peculiarities of the optical response of the structure.

The sharp features (for example that which originates at f ≈ 0.75 x 10^{15} Hz for a slab thickness of zero) are results of the waveguide modes in the system redistributing energy between the different propagating orders.

The broader modes which disperse to lower frequencies as the slab thickness is increased are a result of interference. For reflection, the light which is reflected from the first interface may interfere with the light reflected from the second interface, and this interference may be either constructive or destructive (depending upon the thickness of the slab). In reflection the condition for constructive interference to occur is that the light reflected from the bottom interface must be in phase with the light reflected from the top interface. Since (from section 4.2) the reflection from the top air / SiO_2 interface

is 180° out of phase with the incident light, and that reflected from the bottom interface is in phase with the incident light, a phase change due to propagation of the light in the slab is required for the two reflected zeroth orders from the two interfaces to be in phase with each other and interfere constructively. This occurs when the slab thickness is an integer value of a quarter of the incident wavelength, since the light reflected from the bottom interface will then have to traverse half of the incident wavelength before it exits the system (in other words $f = nc/4d$, where f is the frequency of the incident light, n is an integer, c is the speed of light, and d is the slab thickness). The same condition is required for destructive interference between the light which is directly transmitted through the structure, and the light which is reflected from the bottom interface, and then reflected by the top interface before propagating through the bottom interface. It is destructive interference in this case since there is a 0° phase change with respect to the incident light for transmission through both interfaces, and a 180° phase change for the light reflected from the bottom interface and then reflected from the top interface due to its additional propagation length of twice the slab thickness. Therefore, these two contributions to the transmitted zeroth order are out of phase with each other resulting in destructive interference.

The depth / height of these features is surprisingly large. However, this can be simply attributed to the possibility of multiple reflections within the grating slab. The criteria necessary for these features to occur is that the reflected zeroth order from the bottom interface is in phase with the reflected zeroth order from the top interface. Therefore, the fields will be at a maximum at the same point on the top interface, which will in turn change the boundary conditions for the creation of all possible orders at the top interface. This alters the magnitude of the various orders, which will in turn mean that the reflection from the bottom interface is different, changing the boundary conditions once again. This is analogous to a feedback loop, and therefore our simple model which only considers the two corrugated surfaces separately is obviously flawed. However, it is still able to give some physical insight into the processes which give rise to the effect for these two interface grating systems, even though it cannot account for the intensities of the different orders.

These interference effects can also be observed in the reflected diffracted order plot, since the reflected diffracted orders created at the top interface may undergo interference with the reflected diffracted order created at the bottom interface in the

same way as the reflected zeroth order (even though the phase changes upon diffraction for the two interfaces is the opposite to that of the zeroth order reflection the result is the same). However, these features will disperse differently with changing frequency, since they depend upon the incident angle of the diffracted order upon the bottom interface which will define the distance it traverses within the dielectric slab before exiting the system. Therefore, the slab thickness at which constructive interference may occur is different to that for the zeroth orders (particularly noticeable close to the critical edge) which are always propagating normal to the average plane of the surface.

The transmitted diffracted order is slightly different. In this case interference may occur between the transmitted diffracted order created at the top interface, and that created at the bottom interface (to first order, disregarding the possibilities of multiple reflections within the SiO_2 slab). From section 4.2 we know that the transmitted diffracted order created at the top interface is in phase with the incident light, whereas that created at the bottom interface is 180° out of phase with the incident light. Therefore, destructive interference will occur for very thin slab thickness. The condition for constructive interference to occur is that the light diffracted at the top interface must travel half the wavelength of the incident light further than the light diffracted at the bottom interface. However, the only difference in the distance travelled by the two orders is due to the diffraction angle for the light diffracted at the top interface (figure 4.3.2.3).

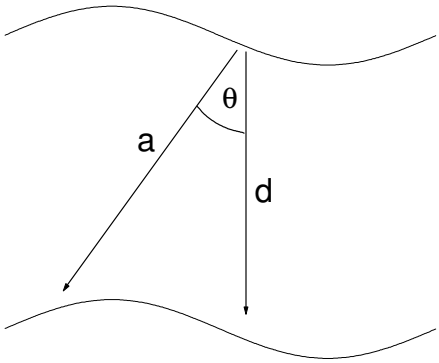

Figure 4.3.2.3 Schematic showing the transmitted diffracted order from the top interface propagates further than the zeroth transmitted order from the top interface.

So the condition we have is that the distance a must be equal to the slab thickness d +λ/2, where λ is the wavelength of the incident light. This can clearly only occur when the diffraction angle, θ, is very large, or when the frequency is very high,

for a slab of thickness of the order investigated here. This is the reason why there is no real propagating diffracted order for the conformally corrugated system at low slab thickness. Of course, the phase difference between the light diffracted at the top interface and that diffracted at the bottom interface will change continuously as the slab thickness is increased and therefore there will be increasing diffraction from the system when this is the case. This can be clearly seen in figure 4.3.2.4, in which we plot the time averaged fields for the system for incident light of frequency 1 x 10^{15} Hz for various slab thickness.

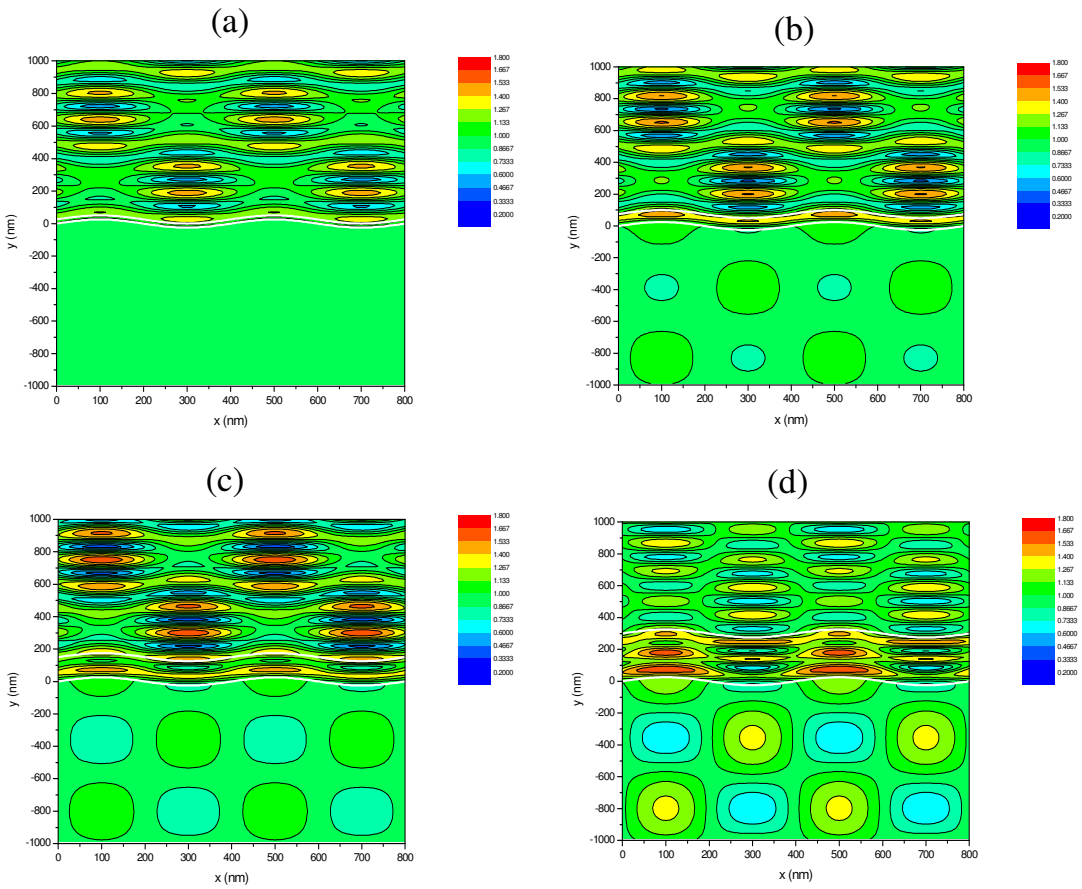

Figure 4.3.2.4 Time averaged H_z component of the fields for the system shown in figure 4.3.2.1 for f = 1.0 x 10^{15}Hz with different slab thickness. a) d = 25nm, b) d = 75nm, c) d = 150nm, and d) d = 300nm.

The analysis presented above is only for the region in which the diffracted orders are real and propagating. We shall now consider the other two possibilities: the case where diffraction in the SiO_2 slab is real and propagating, but the diffracted order in the SiO_2 is evanescent, and the case where the diffracted order in both media are evanescent. In figure 4.3.2.5 we show field distributions for these two cases.

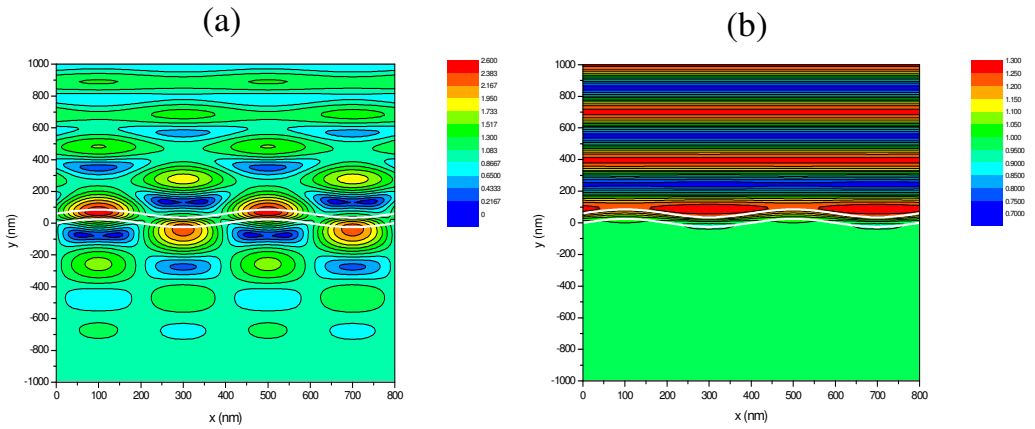

Figure 4.3.2.5 Time averaged H_z component of the fields for the system shown in figure 4.3.2.1 for a) f = 0.732 x 10^{15} Hz (diffractive in SiO_2 only), and b) f = 0.492 x 10^{15}Hz (non-diffractive in both media).

It is clear from figure 4.3.2.5 that when both diffractive orders are evanescent there are no transmitted diffracted fields, whereas when the transmitted diffracted order is real and propagating in the SiO_2 only there is an evanescently decaying transmitted field. We shall now use the same analysis method which we used before (considering the two diffraction processes from the two interfaces individually and seeing how they interfere with each other). We shall do this by using the ISS method used in section 4.2 and considering the two processes as: 1) the light is diffracted at the top interface and then transmitted through the bottom interface, and 2) the light is transmitted through the top interface and then diffracted by the bottom interface. We shall do this by using the ISS method twice for each process. For the 1st process we shall calculate the real and imaginary parts of the amplitude coefficient for the transmitted diffracted order created from the top interface, and also for light incident upon the bottom interface with $k_x = k_g$ (since the light is normally incident this will be the wavevector of the diffracted order for any frequency). By combining these two results in the appropriate way we can then obtain the total diffracted field in the exit medium of the system due to the two interfaces for light diffracted from the top interface. We will also calculate the transmitted diffracted order fields due to diffraction from the bottom interface in a similar way except that we shall calculate the zeroth order amplitude coefficient for light propagating through the top interface and combine this with the amplitude coefficient for the diffracted order created at the bottom interface. We shall then combine these two results which will give us the total diffracted order field for the two interface system, ignoring multiple reflection within the dielectric slab, which will assume that the slab thickness is zero since no account of the phase change due to

propagation of the diffracted order within the slab thickness is incorporated within the model. The results of this are shown in figure 4.3.2.6.

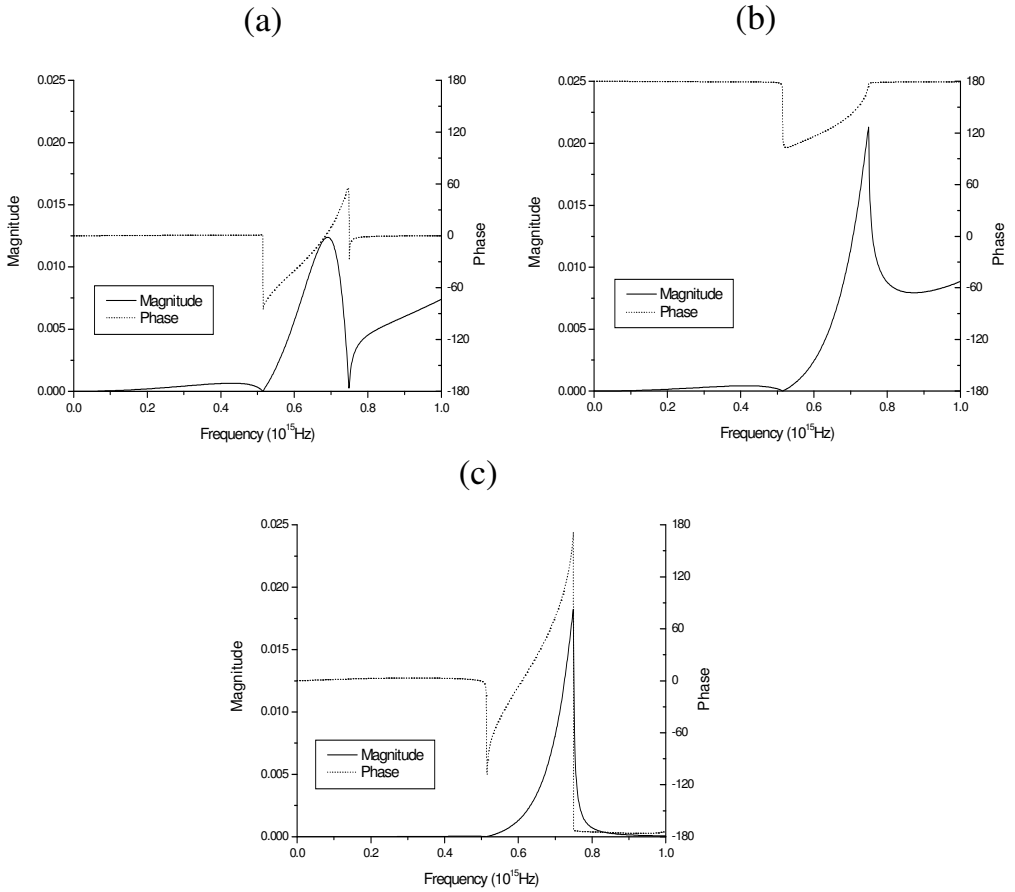

Figure 4.3.2.6 The magnitude and phase of the amplitude coefficient for the system shown in figure 4.3.2.1 obtained from the ISS method by considering the two diffraction processes separately (as described in the text), for a) diffraction from the top interface, b) diffraction from the bottom interface, and c) the total transmitted fields for the system obtained by combining a) and b).

It is clear from the figure 4.3.2.6(c) that there will be no transmitted diffracted order fields when the diffracted orders in the two media are real and propagating, and when they are evanescent. However, in the middle region, where the diffracted order is real and propagating in the SiO_2 only, evanescent diffraction in the exit medium of the system is evident. This result agrees with the results of the field distributions shown in figure 4.3.2.5. It is also evident that the lack of diffracted fields when the diffracted orders are both real and propagating, or both evanescent, is due to the fact that in these regions the phases of the two diffraction processes are 180° out of phase with each other (see figures 4.3.2.6(a) and (b)). This is not the case in the region where only the

diffraction in the SiO_2 is real and propagating, and therefore transmitted diffracted order fields are observed.

4.3.3 The Asymmetric Thin Dielectric Slab

Having discussed the thin dielectric slab in a conformal geometry we shall now consider the case where the two gratings have identical modulations, but where the bottom corrugation is phase shifted with respect to the corrugation on the top interface. The system is shown in figure 4.3.3.1.

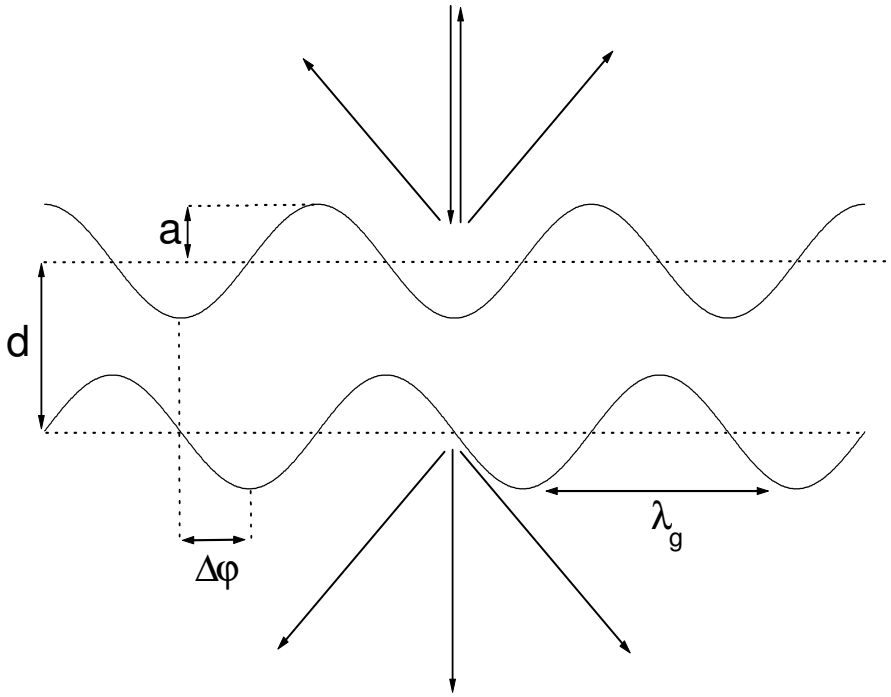

Figure 4.3.3.1. A schematic of the system under consideration

When the optical response of this structure is investigated we must now investigate the +1 and –1 diffracted orders in both reflection and transmission since the phase shift of the lower interface with respect to the top interface has broken the symmetry of the system. In figure 4.3.3.2 we show the optical response of a 60nm thick SiO_2 slab bounded by air on both sides, and with both surfaces having a corrugation of 400nm pitch and 25nm amplitude, as a function of the phase difference between the corrugations on the top and bottom surfaces, and of the frequency of the incident light.

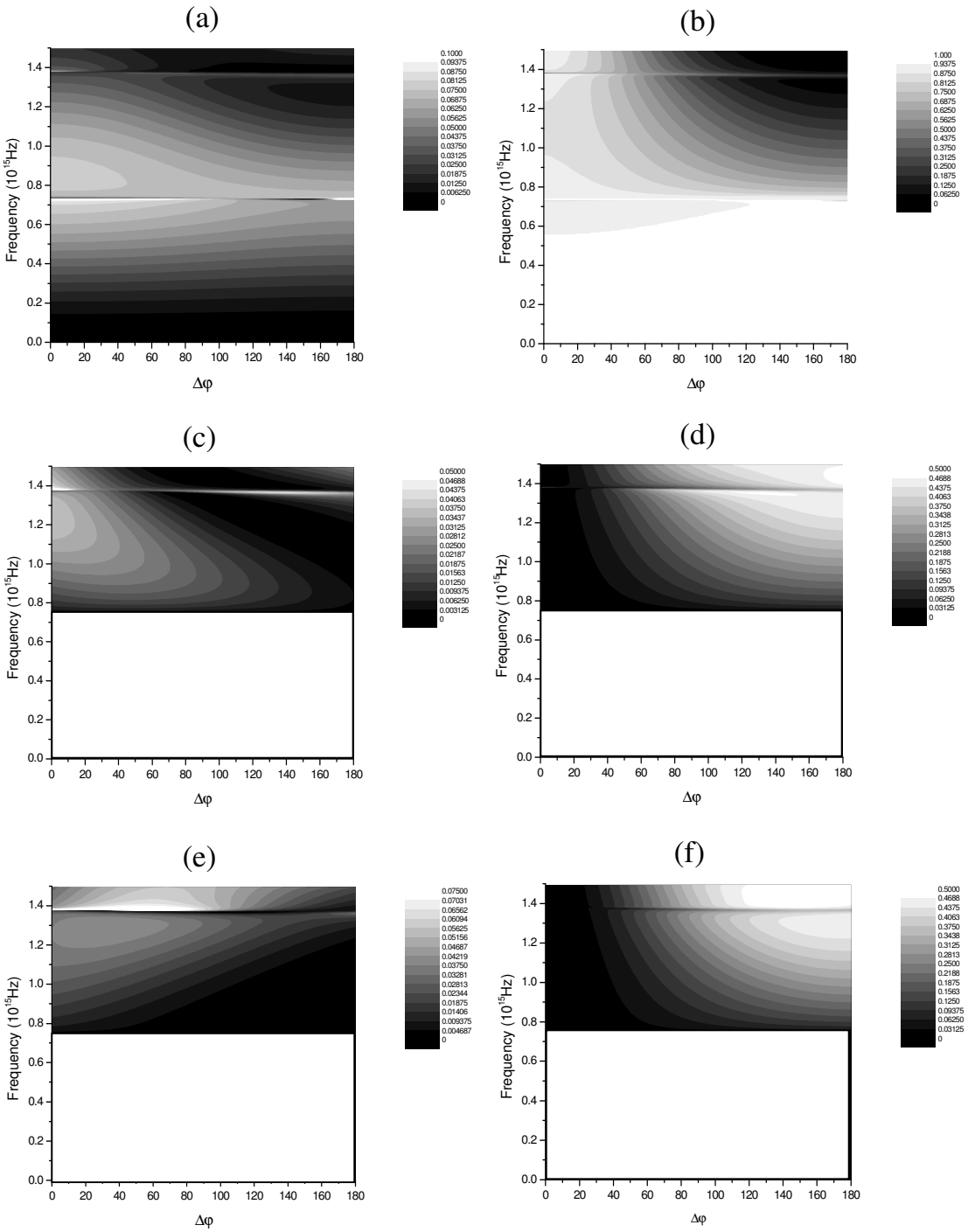

Figure 4.3.3.2 The optical response of a 60nm thick dielectric slab corrugated on both surfaces with gratings of λ_g = 400nm, and amplitude of 25nm, as a function of the phase between the corrugations on the two interfaces, and of the frequency of the incident light. a) the reflected zeroth order, b) the transmitted zeroth order, c) the reflected +1 diffracted order, d) the transmitted +1 diffracted order, e) the reflected –1 diffracted order, and f) the transmitted –1 diffracted order.

We shall begin by describing the optical response when the two corrugations are in anti-phase. (When this is the case the structure is symmetric and the +1 and –1

diffracted orders will therefore be identical). The diffracted order intensities as a function of frequency for this case are shown in figure 4.3.3.3 (taken from figure 4.3.3.2)

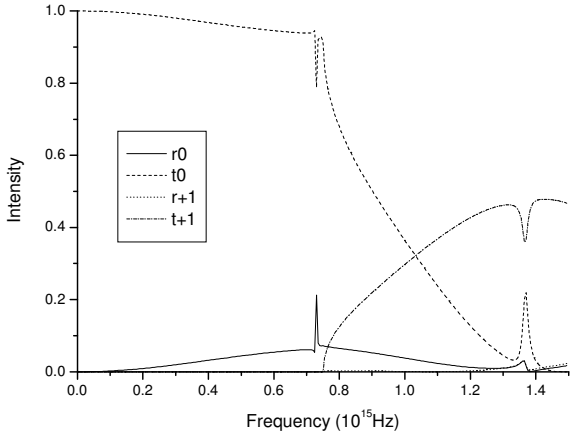

Figure 4.3.3.3 The intensities of the zeroth reflected and transmitted orders, and of the 1st order diffracted transmitted and reflected orders for a 60nm thick dielectric slab corrugated on both surfaces with antisymmetric sinusoidal corrugations on each surface of 400nm pitch and 25nm amplitude.

It is clear that there is a very peculiar phenomenon occurring for this structure. For high frequencies almost all of the energy of the incident light is being transferred in to the +1 (and, therefore, also the −1) transmitted diffracted orders. The fact that the reflected zeroth and diffracted orders are very small is no surprise since this will, in general, be the case from a dielectric grating structure. However, the reduction of the transmitted zeroth order to near zero intensity is somewhat unexpected.

In fact the reflected diffracted orders are very small indeed for the thickness slab investigated here. In section 4.3.2 we discussed the way in which the transmitted diffracted fields created from the top and bottom interfaces of a conformally modulated thin dielectric slab were out of phase with each other, and that due to this the total transmitted diffracted field from the structure was near zero (depending upon the slab thickness). When the two gratings are in anti-phase the diffracted orders created from the bottom surface are 180° out of phase when compared to those in a conformal geometry. This is because the diffracted fields of the light propagating in the dielectric slab (ignoring multiple reflections once again) are defined by the grating which they have already passed through. In other words, the periodicity (in the x direction) of these fields is defined by the first grating, and therefore the phase of the diffracted orders

created from the bottom interface will be altered when the phase of that grating is changed relative to the top surface corrugation. In fact, when the two corrugated interfaces are in anti-phase the phase of the diffracted orders created from the bottom surface are changed by 180° when compared to the conformal case.

For the reflected diffracted order in a conformal geometry, for the slab thickness considered here, the diffracted orders created from the two interfaces were interfering constructively. For the anti-phase structure the opposite is true, and therefore the diffracted orders created at the top and bottom interfaces for this slab thickness cancel each other reducing the total reflected diffracted order fields from the system. However, the zeroth order reflected light is unaffected by the change in phase of the bottom surface corrugation, but since the zeroth order reflection is always going to be small (away from any possible waveguide modes of the system) the majority of the energy will be contained within the transmitted orders.

By the same arguments above, the transmitted diffracted order created at the bottom interface is also changed by 180° when compared to that created at the bottom surface for the conformally corrugated system. For the conformally modulated system the transmitted diffracted order fields from the top and bottom surfaces cancelled, whereas for the anti-phase structure they will interfere constructively. As we have previously mentioned the intensity of the diffracted order cannot be considered as the addition of the two transmitted diffracted orders created at the top and bottom surfaces, and therefore it is possible that this high level of intensity of the diffracted order is simply the result of the combination of these two diffraction processes.

It is interesting to consider the effect of the slab thickness, and of the grating amplitude, for an anti-phase grating in order to determine whether this property of almost all of the incident energy being transferred to the transmitted diffracted orders is a general property of anti-phase corrugated thin dielectric slabs, or whether it only happens to occur for the parameters which we are investigating. Therefore, we have modelled the intensities of the various orders from the system as a function of frequency and slab thickness (figure 4.3.3.4), and also of the transmitted orders as a function of frequency and grating amplitude for a 60nm thick dielectric slab, and also of a 150nm thick dielectric slab (figure 4.3.3.5). (Note the different frequency range used for the 150nm thick dielectric slab modelling. The reason for this will be made clear later).

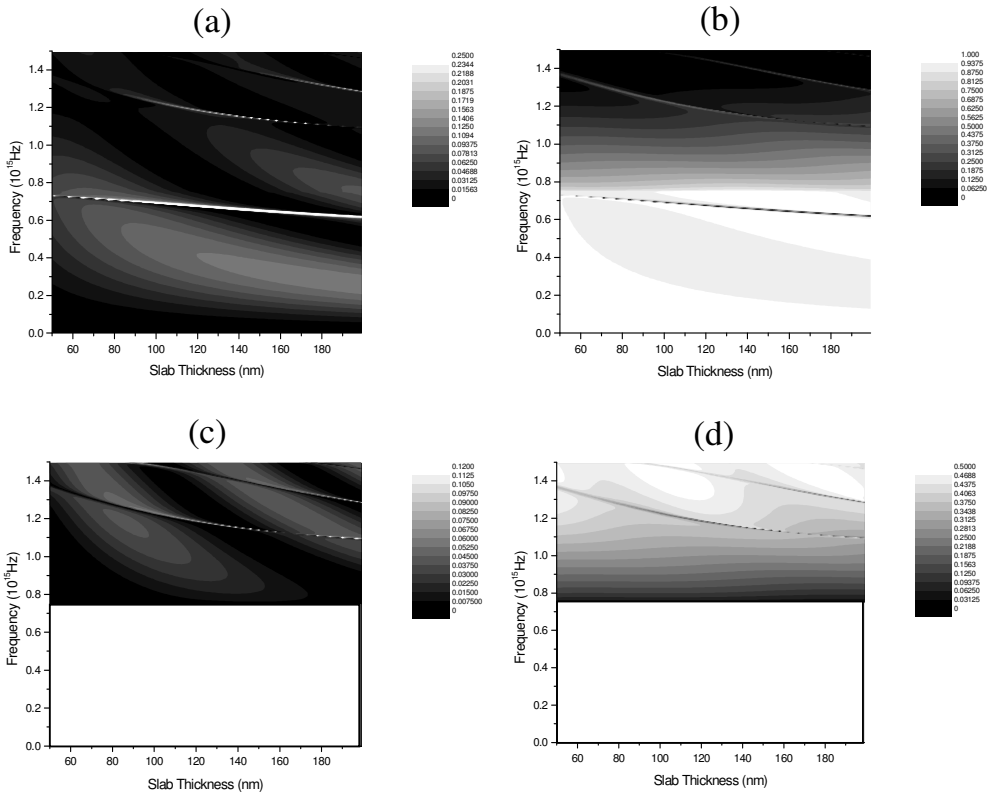

Figure 4.3.3.4. The intensities of the various orders from the anti-phase two interface system as a function of frequency and slab thickness. a) the zeroth order reflected, b) the zeroth order transmitted, c) the +1 diffracted reflected, and d) the +1 diffracted transmitted.

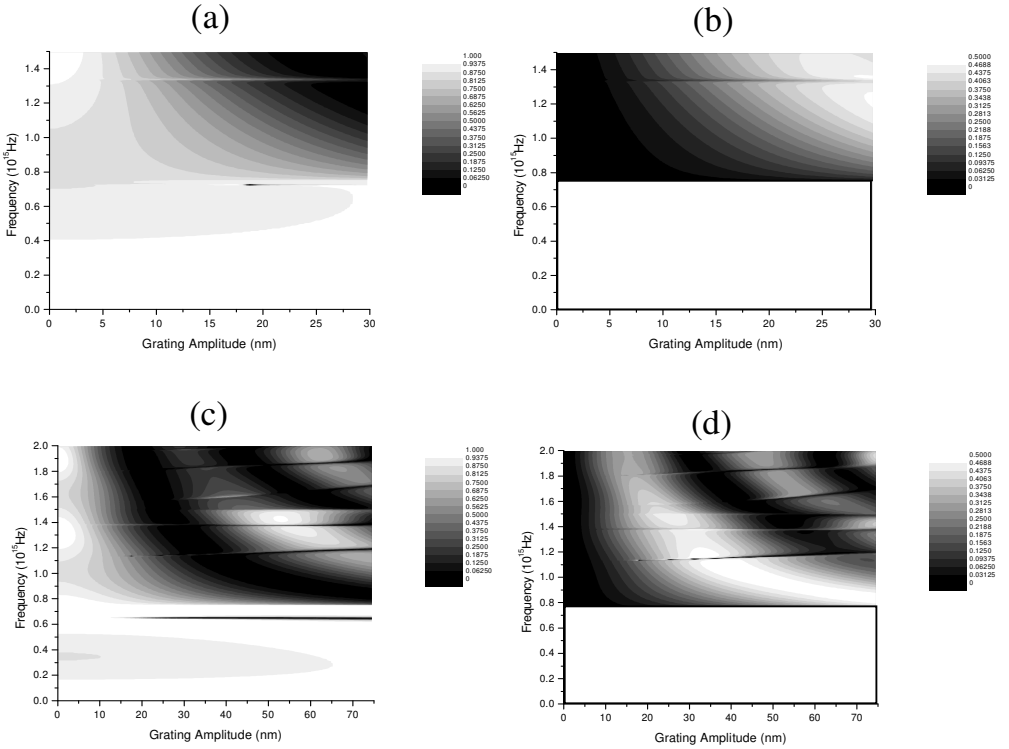

Figure 4.3.3.5. The intensities of the transmitted orders as a function of frequency and grating amplitude for a) and b) a 60nm thick dielectric slab, and c) and d) a 150nm thick dielectric slab. (a and c are the zeroth orders, and b and d are the diffracted orders).

Firstly, it is clear from figure 4.3.3.4 that the slab thickness has relatively little effect on the intensity of the transmitted diffracted order, except for a small periodicity as a function of slab thickness. This is caused by an increased intensity in the reflected orders due to the change in phase of the reflected orders created from the bottom interface, arising from the increased propagation length within the dielectric slab. Therefore, this phenomenon seems to be relatively independent of the thickness of the grating slab.

The effect of grating amplitude for a 30nm thick slab (figure 4.3.3.5) indicates that for lower grating amplitudes the diffraction efficiency of the +1 transmitted diffracted order increases with amplitude of the grating, which, of course, must be true since for a very shallow grating there is far less diffracted order intensity occurring. However, for the 150nm thick grating slab the amplitude of the grating can be increased to 75nm (simply because, for an anti-phase grating structure, the amplitude of the grating is limited to half the slab thickness). In this case it is clear that the grating efficiency is periodic with increasing grating amplitude, and in fact this periodicity is given by:

$$a = \frac{n\lambda_z}{2} \qquad 4.3.3.1$$

where λ_z is the z component of the wavelength of the diffracted order created at the top interface, and n is the refractive index of the slab dielectric, with a minimum in the diffracted order fields occurring when this condition is met. Two possible explanations for this periodicity will now be described.

Firstly, the interaction of the transmitted diffracted order created at the top interface with the reflected diffracted order created at the bottom interface creates field maxima and minima which are located within the grating grooves of the corrugation on the bottom interface (figure 4.3.3.6). It is possible that when these fields 'fill' the grating grooves in the bottom interface the zeroth order transmitted light from the top interface no longer 'sees' the corrugation on the bottom interface due to the occurrence of these fields. This would occur when the beating between the reflected and transmitted diffracted orders in the z direction has a wavelength of integer multiples of λ/2 which is the condition noted in equation 4.3.3.1.

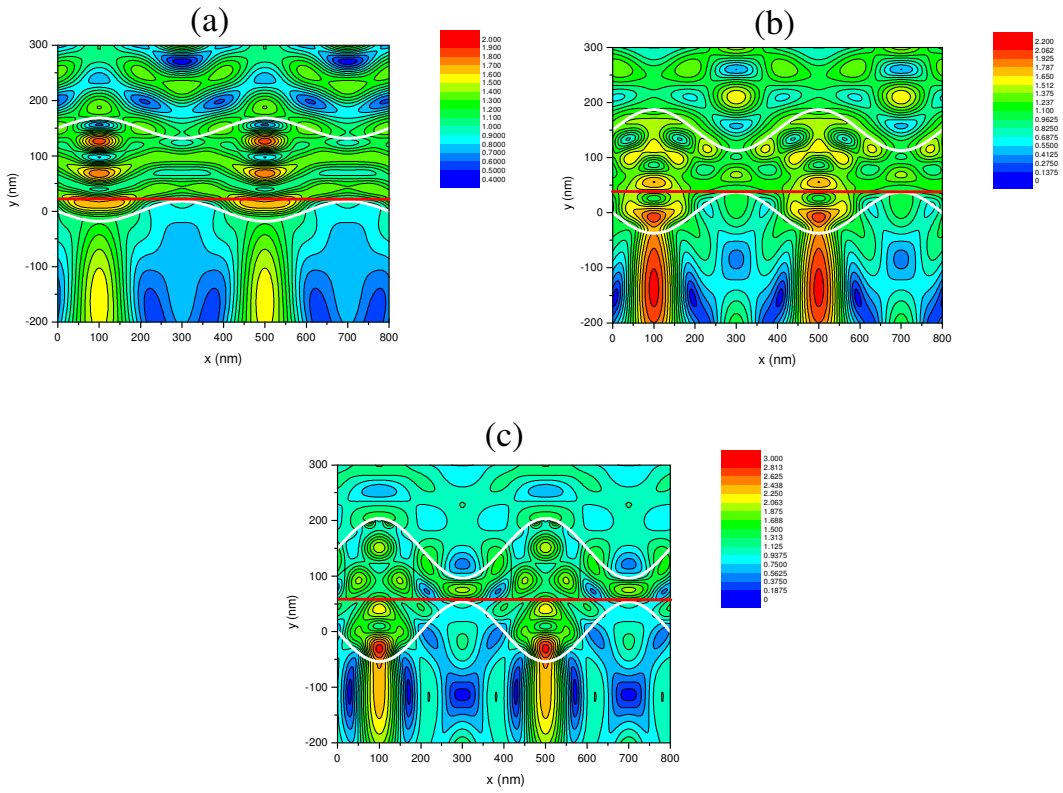

Figure 4.3.3.6 Time averaged field profiles for successive maxima and minima in the 1st order transmitted diffracted order for an incident frequency of 1.75 x 10^{15}Hz. a) a = 18.35nm, b) a = 37.2nm, c) a = 53.75nm. A red line has been drawn on to the plots at the tops of the bottom surface corrugation in order to demonstrate the way the field maxima fit within the grating grooves.

A second explanation relies on the fact that, if light is incident upon a single interface dielectric grating in the Littrow mount when only the zeroth and –1 diffracted orders are propagating, the diffracted order intensity of the transmitted diffracted order is periodic with grating amplitude, oscillating between an intensity of zero, and an intensity of 1 (the transmitted zeroth order oscillates in anti-phase with this) – Moharam and Gaylord [1982]. In the two interface system described above, the transmitted diffracted orders created at the top interface are incident upon the bottom interface with $k_x = k_g$ (since we are incident upon the structure at normal incidence). Therefore, the $-2k_g$ scattering of this field produces diffraction back in the direction of incidence, which is the condition for this phenomenon to occur. However, the periodicity of this effect for a single dielectric interface is approximately $a = \lambda_z$, which is not the periodicity noted from figure 4.3.3.4. Whether the fact that we have real propagating +1 and –1 transmitted diffracted orders in our system, or the fact that the transmitted diffracted order fields from the top surface corrugation are incident in the 2nd order Littrow mount

rather than the first, can account for this discrepancy is unknown, and therefore we cannot say whether this phenomenon is due to this mechanism or not.

We shall now consider the case where the phase of the bottom surface corrugation is out of phase with the top surface corrugation, but not in anti-phase with it. From figure 4.3.3.2 it is clear that, for a particular frequency and phase difference, the intensity of the +1 and –1 reflected diffracted orders are different, even though the light is normally incident upon the structure. Therefore, the structure exhibits the same type of behaviour as a blazed grating. However, the transmitted diffracted order intensities for the +1 and –1 orders are almost the same, which is not as expected for a blazed grating structure. This is shown in figure 4.3.3.7. Previous work has only considered a blazing effect on the waveguide modes (Yamasaki [1995], Peng and Tamir [1974]), and not on the diffracted orders.

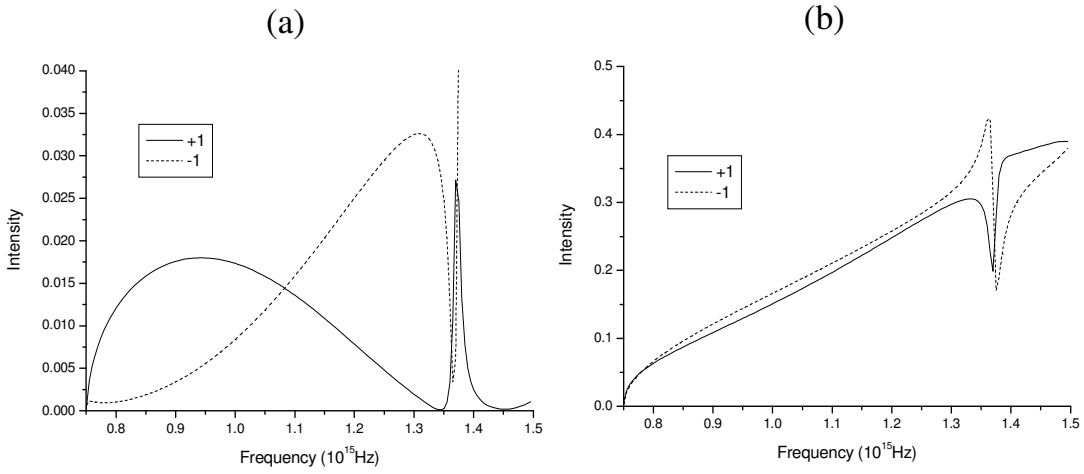

Figure 4.3.3.7 The intensities as a function of frequency for a two interface corrugated system of thickness 60nm and amplitude 25nm for a) the reflected ±1 diffracted orders, and b) the transmitted ±1 diffracted orders.

The phase of the transmitted diffracted orders created at the bottom interface depend upon the phase of the bottom surface corrugation with respect to the top surface corrugation. The +1 transmitted diffracted order will have phase $\phi^b + \Delta\phi^b$, and the –1 transmitted diffracted order will have phase $\phi^b - \Delta\phi^b$ (where ϕ is the phase change due to the refractive index change - 180° in this case – and $\Delta\phi$ is the phase shift of the bottom surface corrugation with respect to the top surface corrugation). However, when these are combined with the transmitted diffracted orders from the top interface the intensities

of the resultant transmitted ±1 diffracted order fields will be equal since $\phi^b + \Delta\phi^b$ and $\phi^b - \Delta\phi^b$ are symmetric about ϕ^t. The same is true of the reflected ±1 diffracted orders. Therefore, the blazing effect in the reflected diffracted orders must be due to a different effect.

If we continue to consider our simple ray model for the diffraction properties of these two interface systems with no multiple reflections within the dielectric slab this phenomenon can be understood. The phases of the transmitted and reflected diffracted orders created at the top and bottom surfaces are shown in the following table:

	Diffraction From	**+1**	**-1**
Reflected	Top Surface	ϕ^t	ϕ^t
	Bottom Surface	$\phi^b_{+1} + \Delta\phi^d + \Delta\phi^u_{+1}$	$\phi^b_{-1} + \Delta\phi^d + \Delta\phi^u_{-1}$
Transmitted	Top Surface	$\phi^t + \Delta\phi^d_{+1}$	$\phi^t + \Delta\phi^d_{-1}$
	Bottom Surface	$\phi^b_{+1} + \Delta\phi^d$	$\phi^b_{-1} + \Delta\phi^d$

Table 4.3.3.1 The phase effects which need to be considered when determining the total ±1 reflected and transmitted diffracted order fields from a thin corrugated dielectric slab (ignoring multiple reflections).

In table 4.3.3.1 the following notation has been used. If there is any difference in the phase between the +1 and –1 diffracted orders then the diffracted order has been added as a subscript. The t superscript denotes diffraction from the top surface, the b superscript denotes diffraction from the bottom surface, and the u and d superscripts denote any phase change due to propagation in the dielectric slab in the upward and downward going directions.

We shall first consider the transmitted diffracted orders. As described earlier the ϕ^b_{+1} and ϕ^b_{-1} terms are symmetric about ϕ^t, and therefore these will have no effect on any difference in intensity between the +1 and –1 transmitted diffracted orders. Also, the ϕ^d_{+1} and ϕ^d_{-1} terms (which are due to the differences in propagation length within the dielectric slab of the ±1 transmitted diffracted orders created at the top interface) will be approximately symmetric about ϕ_d (the phase change due to propagation of the zeroth order transmitted fields through the top interface). Therefore these will also have very little effect on any difference between the +1 and -1 transmitted diffracted order

intensities. It would therefore be expected that the +1 and –1 transmitted diffracted order intensities would be very similar even for a non-symmetric grating system.

In fact this can be more clearly demonstrated by using some test numbers. Let us say that $\phi^t = 0°$, and that the bottom surface corrugation is 90° out of phase with the top surface corrugation so that $\phi^b_{+1} = -90°$ and $\phi^b_{-1} = 90°$. We shall also introduce some slab thickness such that $\phi^d = 30°$, and since ϕ^d_{+1} and ϕ^d_{-1} are symmetric about this we shall give them values of 20° and 40° respectively. The phase of the +1 diffracted order created at the top interface is then 20° and that at the bottom interface is -60° giving an 80° phase difference between them. The corresponding values for the –1 diffracted order are 40° for the top surface diffraction and 120° for the bottom surface diffraction, giving an 80° phase difference between them. Since the phase differences for the two diffracted orders are the same it would be expected that the intensities of the two orders would be the same. This agrees with the modelling shown in figure 4.3.3.7(b).

In the reflection case the ϕ^u_{+1} and ϕ^u_{-1} terms are symmetric about ϕ_d (approximately). We shall again use some test numbers to demonstrate what happens with the phases in this case. $\phi^t = 180°$, $\phi^d = 30°$, with ϕ^u_{+1} and ϕ^u_{-1} taking values of 40° and 20° respectively, and also $\phi^b_{+1} = -90°$ and $\phi^b_{-1} = +90°$. These values then give the phase of the +1 reflected diffracted orders as 180° for diffraction from the top surface, and -20° for diffraction from the bottom surface, giving a phase difference between these of 160°. For the –1 reflected diffracted orders the corresponding results are 180° for diffraction from the top interface, and 140° for diffraction from the bottom interface, giving a phase difference between them of 40°. The phase differences for the +1 and –1 diffracted orders are clearly different due to the propagation within the dielectric slab, and this will lead to a different intensity in the two, and hence a blazing effect. Also, since this difference is a result of the propagation of the light in the dielectric slab for the bottom surface diffraction process, it would be expected that it would be frequency dependent. This agrees with the results in figure 4.3.3.7(a).

4.4 Summary

In this chapter we have investigated the optical response of single interface and two interface corrugated dielectric structures. In the first section we have described the diffracted order intensities / magnitudes and phases for the diffracted orders when they are propagating, and when they are evanescent. We have found that when either the reflected or transmitted diffracted order becomes evanescent the amplitude coefficient of the other diffracted order reduces to zero even though the amplitude coefficient of the diffracted order which has become evanescent still has some magnitude. We have also found that the phase of the diffracted orders depend upon whether the light is incident upon the air side of the structure, or the SiO_2 side of the structure.

In the second section the results obtained on single interface structures have been used to describe the way in which the transmitted diffracted orders created at the top and bottom corrugated surfaces of a conformally modulated thin dielectric slab cancel, producing no transmitted diffracted order fields when the dielectric slab thickness is very small. The exception to this occurs when real propagating diffracted orders are allowed within the SiO_2 slab only. In this case evanescently decaying transmitted diffracted fields are evident from the structure.

Following this we have discussed the optical response of thin dielectric slabs corrugated on both interfaces when there is some phase difference between the two corrugations. We have found that it is possible to have almost the entire energy of the incident light distributed between the ± 1 diffracted orders when the two corrugations are in anti-phase with each other. We have also described the way in which a blazing effect occurs in the reflected diffracted order when the phase between the two corrugations is between 0° and 180°.

Chapter 5

Surface Plasmon Polaritons on Thin Slab Metal Gratings

5.1 Introduction

In chapter 4 we investigated the case of a thin dielectric slab corrugated on both surfaces. In this chapter we will extend this to the case where the thin slab consists of a metal, upon which SPPs may be excited on both interfaces. There will be two main sections to this chapter. The first section will consider the case where the refractive indices of the two bounding dielectrics are different, and therefore the excitation frequencies of the SPPs on the two interfaces will also be different. This follows the work of Schröter and Heitmann [1999] where they observed that the SPP on the transmission side of the structure could not be excited when the metal slab was conformally modulated, but could be if the system was asymmeterised by flattening one of the interfaces, or by changing the phase of the corrugations with respect to each other. We will extend this work and explain that this is not entirely true, and give an explanation of their results. We shall also explain why, in this geometry, the transmission features due to SPP excitation may be observed as maxima, minima, of Fano shaped resonances.

In the second section we shall consider the case where the refractive indices of the two bounding dielectrics are identical. In this case coupled SPPs may be excited, and we shall begin the section by describing the long rang and short range SPPs (LRSPPs, SRSPPs) which may be excited. We shall then describe the effect on the dispersion and coupling of the coupled SPPs of changing the phase between the two corrugations, for both purely sinusoidal gratings, and those with a small $2k_g$ component in the grating profile.

5.2 SPPs on Thin Corrugated Metal Slabs Bounded by two Dielectrics with Different Refractive Indices

In a recently published paper (Schröter and Heitmann [1999]) an unexpected result when light was incident upon a periodically corrugated thin metal film bounded by dielectrics with different refractive indices was observed when the corrugations on the two interfaces were identical, and in phase with each other. They observed that it was not possible to excite the SPP on the metal surface not facing the incoming light, and they ascribed this to the lack of a thickness variation within the metal. In this section we present a somewhat different interpretation of their results, and show that the SPP is in fact excited on the transmission side of such structures, although only very weakly. We explain why this coupling is so weak in terms of cancellation of the evanescent diffracted orders from the two diffractive surfaces, and how, by changing the phase between the corrugations on either surface, this coupling becomes much stronger. We also present an explanation for the observation that SPP excitation on such structures may lead to transmission maxima, minima, or Fano shaped resonances (where the zeroth order transmission shows both maxima and minima through the SPP excitation frequency).

If an optically thin corrugated metal film bounded by dielectrics is investigated, then it is possible that SPPs may be excited at both metal / dielectric interfaces, and in each case it is the evanescent diffracted orders (corresponding to diffraction in each bounding dielectric medium) which excite the SPPs.

The samples Schröter and Heitmann manufactured (and also modelled) consisted of thin metal films approximately 80nm thick, which were corrugated on one, or both, surfaces. These films were produced on quartz substrates, with air as the other bounding medium. They then illuminated these samples from the quartz side with a multi-wavelength source at angles near normal incidence. They found that when only one surface was corrugated (either surface) SPPs could be excited at both interfaces, whereas if both surfaces were corrugated with identical gratings (in a conformal geometry) only the SPP on the surface irradiated could be excited. If one of the gratings on these doubly corrugated films was phase shifted with respect to the other then excitation of SPPs on both surfaces was again found to occur.

We have repeated their modelling using the method of Chandezon (see Chapter 3), but using different grating parameters in order to replicate their results as closely as possible (unlike Schröter and Heitmann we use a sinusoidal grating profile in order to simplify the problem since it removes the possibility of scattering processes from the higher harmonics of the surface profile). We use a polynomial fitted to experimentally determined values to describe the frequency dependent dielectric function of the silver (separate polynomials for the real and imaginary parts). The results of these calculations for similar structures to those investigated by Schröter and Heitmann with normally incident TM polarised light as a function of frequency are shown in figure 5.2.1. Our results confirm their obsrevations in that the SPP at the air / metal interface (which would be expected to occur at f ~ 0.475 x 10^{15}Hz) does not appear to be excited on the thin metal slab in a conformal geometry.

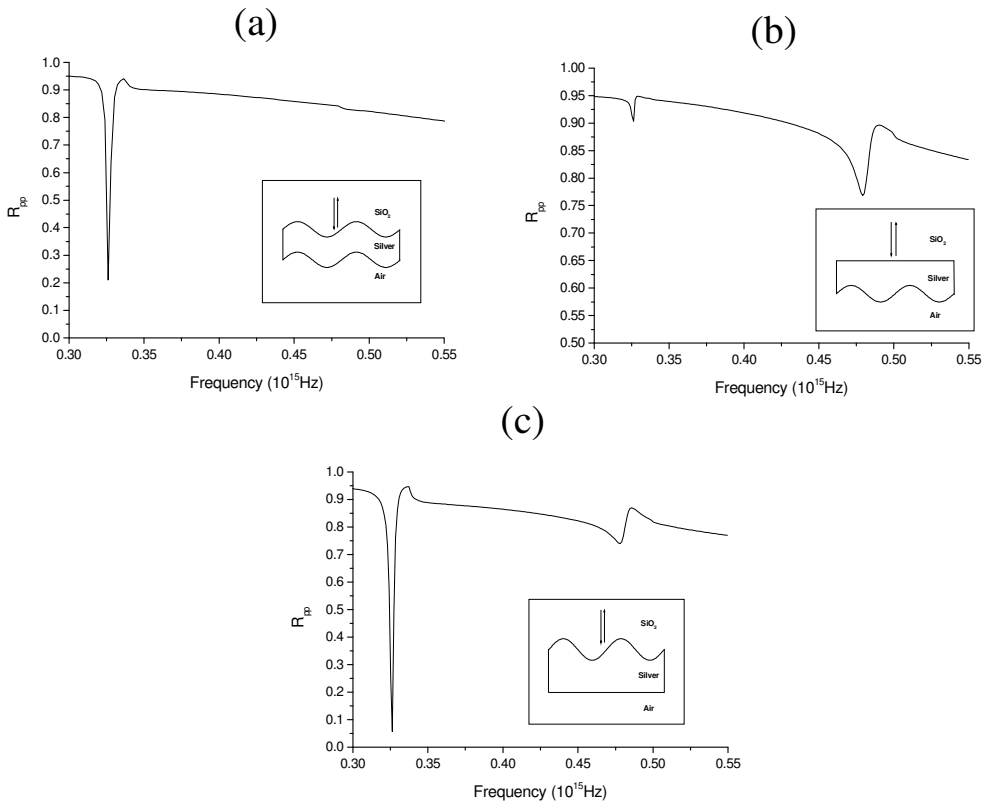

Figure 5.2.1. Reflectivity of TM polarised light for normal incidence in the classical mount as a function of frequency, for a 40nm thick silver film corrugated with a 600nm pitch sinusoidal grating of 10nm amplitude. a) corrugated on both sides, b) corrugated on the bottom surface only, and c) corrugated on the top surface only.

The explanation given by Schröter and Heitmann for the lack of coupling to the SPP on the second interface was that a thickness variation in the metal film was necessary for coupling to occur. However, while this is in essence true, they do not

explain why it is necessary. We present below an extension of their results, and show that the SPP on the non-incident side of the structure is in fact excited, but only very weakly. We explain why this is the case, and also describe a second phenomenon: the fact that the feature seen in the transmitted zero-order can be observed as a transmission maximum, minimum, or as a Fano shaped resonance.

Firstly we shall show that the SPP on the transmission side of the structure is excited even with a conformal geometry. Figure 5.2.2(a) shows the zero-order transmission, and total absorption, of the structure.

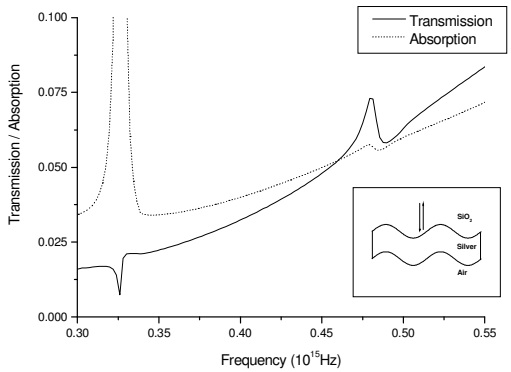

Figure 5.2.2. Transmissivity and absorption of TM polarised light for normal incidence in the classical mount for a thin silver film with the parameters used to obtain figure 5.2.1(a).

It is clear from these results that there is a very small transmission maximum at the frequency at which the SPP on the transmission side of the grating would be expected to occur ($f \approx 0.475 \times 10^{15}$Hz), and that this corresponds to some small resonant absorption of the incident light. Due to the fact that the coupling to this SPP is so weak it is not surprising that there is no feature in the corresponding reflectivity plot of figure 5.2.1(a). Since the fields due to the SPP will decay exponentially away from the surface they will be weak enough at the incident side that any scattering from that surface into the specularly reflected order will be too weak to have a noticeable effect on the reflectivity response of the structure. However, to understand why the coupling to the SPP is so weak it is necessary to consider the magnitude of the complex amplitude coefficients, and phases, of the possible contributions to the transmitted diffracted field which excites it (similar to the analysis performed for dielectric slabs in chapter 4).

Consider a thin metal slab with complex refractive index n_2, corrugated on both surfaces conformally with a 600nm pitch grating, and bounded on either side with dielectrics of refractive indices n_1 and n_3 (n_1 on the incident side, and n_3 on the transmission side). Then, over a certain wavelength range at normal incidence, light will

be diffracted at the first interface and produce three transmitted orders: the zeroth and ±1 diffracted (which will be evanescently decaying in the y-direction). (For this grating pitch, and with normally incident light of the frequency range investigated, the other evanescent diffracted orders are extremely small and will be ignored). To first order we have nine possible scattered fields transmitted through the slab: the zeroth transmitted order from the first interface may pass through the second interface, or be diffracted by it, and the evanescent diffracted orders from the first interface may pass through the second interface or be diffracted by it (figure 5.2.3). Therefore, the resultant transmitted diffracted field from such a system will be a combination of the diffraction from the two interfaces. The SPP on the bottom interface may then be excited when the resultant transmitted diffracted field becomes evanescent and matches the coupling condition for the SPP.

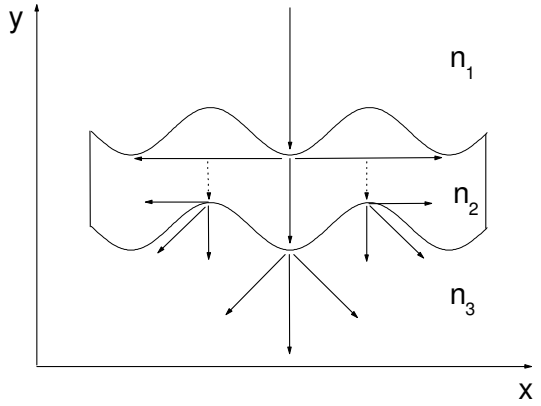

Figure 5.2.3. Schematic showing the origin of the possible transmitted orders due to diffraction from a thin metal slab corrugated on both surfaces. The dotted lines indicate the evanescently decaying fields due to diffraction at the top interface.

By performing calculations of the efficiencies and phases of a single interface grating (both SiO_2/metal, and metal/air) using the ISS method described in chapter 3, and combining the results from the evanescent diffracted order created at the incident interface which is not diffracted at the second interface, and the zero-order transmitted field from the incident interface which is scattered into an evanescent diffracted order at the second interface, the total transmitted +1 diffracted field around the SPP excitation frequency can be calculated.

It must be noted that the ISS method only achieves convergence for gratings with a lower aspect ratio than those studied by Schröter and Heitmann, and that the method is very limited in that it may only be used to calculate the efficiencies and phases from single interface structures, unlike the Chandezon method. Therefore,

combining the diffracted order efficiencies in the way described above will not give the total transmitted diffracted field for these thin slab structures since it does not account for the multiple reflections / scattering processes within the thin film. However, since all orders within the metal film are exponentially decaying, and therefore higher order contributions should be small, we can use this method as a close approximation since it's simple analytical form may enable a better physical understanding of the processes involved than does the Chandezon method. It also allows the various scattering processes to be calculated individually which may facilitate a better understanding of the phenomena investigated.

Figure 5.2.4(a and b) shows the magnitudes and phases of the +1 diffracted orders for the two cases described above, and also shows the magnitude of the total diffracted order for a thin film structure, obtained by combining these two results (c).

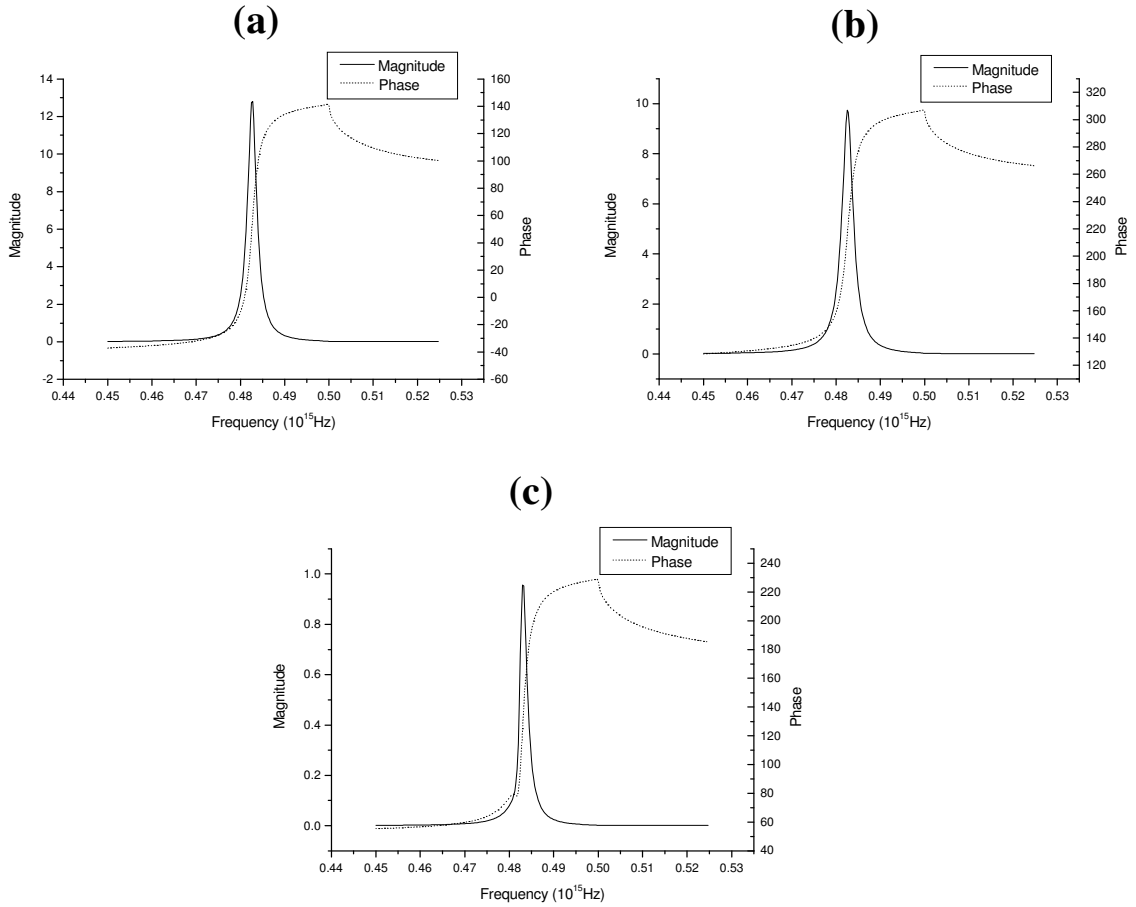

Figure 5.2.4. The magnitude and phase of the 1st order transmitted diffracted field around the SPP excitation frequency from a conformally corrugated structure for a) diffraction from the top surface, b) diffraction from the bottom surface, and c) the total 1st order transmitted diffracted field.

These figures show a peak in the magnitude of the complex amplitude coefficient which corresponds to the excitation of the SPP on the lower interface. The

phase also undergoes a 180° phase change through this point due to the fact that the SPP is being resonantly driven by the evanescent diffracted order (See Chapter 2).

It is clear that the two contributions to the total +1 transmitted diffracted order fields are approximately 180° out of phase with each other throughout the frequency range investigated. Therefore, when these two contributions are combined they cancel each other leaving only a small transmitted diffracted order field due to the difference in the magnitude of the two contributing diffraction processes. Since the transmitted diffracted field is very small only very weak coupling to the SPP occurs, which in turn produces only small features in the zero-order transmission from the structure, and nearly no feature in the zero-order reflection.

To first order the amplitude coefficient of the +1 transmitted diffracted order is given by (from Greffet and Massaranni [1990]):

$$e_{tp}^{1}(1) = \frac{i(\varepsilon_2 - \varepsilon_1)k_0^3 a \gamma_r \sqrt{\varepsilon_2}}{2(\gamma_t + \gamma_r)(k_g^2 + \gamma_r \gamma_t)} \qquad 5.2.1$$

(The definitions of the variables and the properties of this expression are discussed in detail for the case of dielectric interfaces in section 4.2 in chapter 4)

The 180° phase difference between the two diffraction processes can be seen from equation 5.2.1 where the $\varepsilon_2 - \varepsilon_1$ factor changes sign for diffraction at the SiO_2/metal boundary compared to that at the metal/air boundary. In fact it is clear that this must be the case since, as the thickness of the thin metal film tends towards zero, the resultant diffraction may only be due to any difference between n_1 and n_3 (this difference is also the reason that the results based upon the ISS method for the two diffraction processes which lead to the total transmitted diffracted field, shown in figure 5.2.4, are not exactly out of phase with each other)

For the other structures investigated by Schröter and Heitmann this cancellation does not occur. The structures which are only corrugated on one surface have stronger transmitted diffracted order fields because there is no secondary corrugation producing cancelling fields out of phase with them. However, for the case where both surfaces are corrugated, but phase shifted with respect to each other, the case is a little more complex.

Since the SPP is excited by the evanescent diffracted order the fields have a periodicity in the x-direction (parallel to the grating vector) only. This periodicity is caused by the diffraction from the grating surface, and the fields produced have their

maxima and minima at the maxima and minima of the grating profile. The diffraction occurring at the SiO_2/metal boundary at some instant in time may have field maxima at the maxima of the grating profile from which it is diffracted. If this is the case, then at the same instant in time the diffraction occurring at the metal/air boundary will have field minima at the maxima of the grating from which it is diffracted. This is the 180° phase difference described earlier. If the phase of the grating on the transmission side is altered with respect to that on the incident side by some phase ϕ then, since the field maxima and minima are locked to the grating profile from which it is diffracted, the phase difference between the two diffracted orders is now equal to 180°-ϕ. Since the phase difference is no longer 180° the two diffracted orders will not cancel in the same way and stronger coupling to the SPP can occur. This is shown in figure 5.2.5, where the zero-order reflectivity and transmissivity are shown as a function of frequency, and phase between the two gratings. Also shown is the magnitude of the amplitude coefficient for the +1 diffracted order.

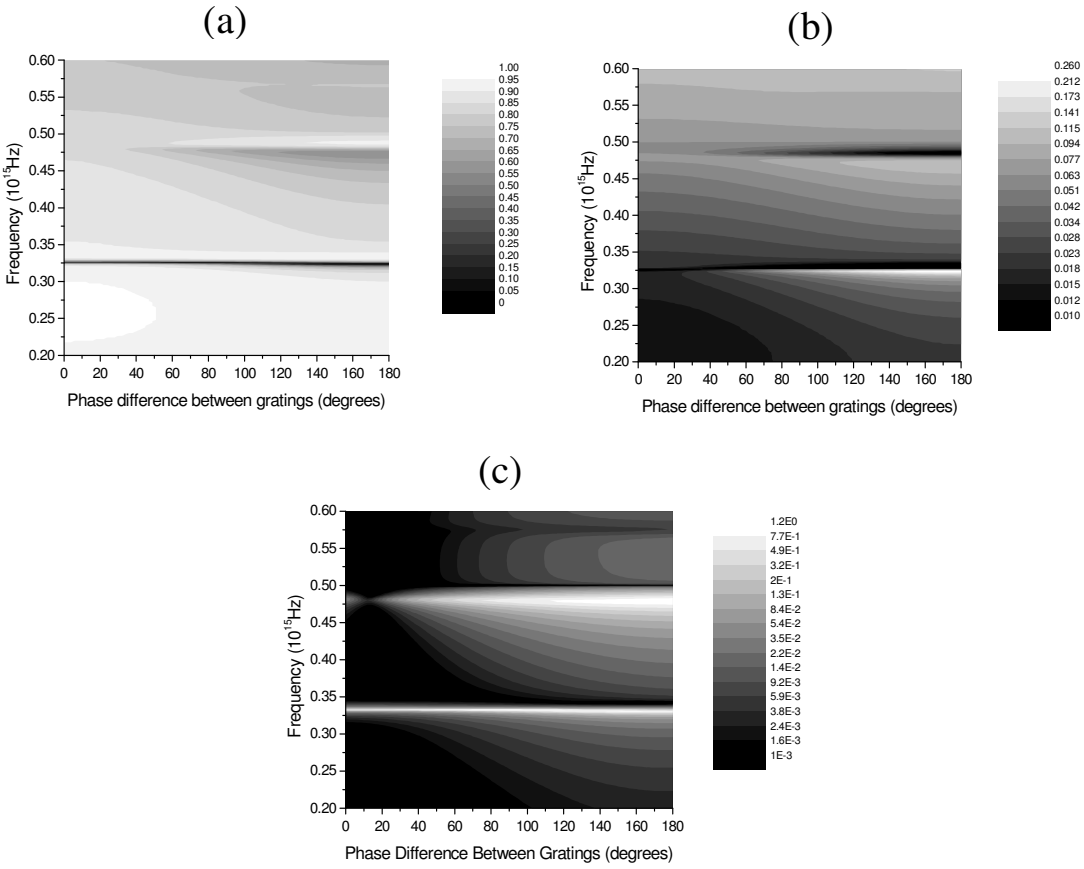

Figure 5.2.5. Results from a dual corrugated silver film as a function of frequency and the phase between the corrugation on the two interfaces. a) zero-order reflection, b) zero-order transmission (log scale), and c) the magnitude of the first diffracted order (log scale).

It is clear that the coupling to the SPP on the transmission side (at f ≈ 0.48×10^{15}Hz) generally increases as the phase between the two grating surfaces is increased, and that this corresponds to an increase in the magnitude of the complex amplitude coefficient of the 1st transmitted diffracted order. Also to be noted is that the minimum in the magnitude of the amplitude coefficient for the first diffracted order actually occurs when the phase between the two gratings is not 0°. This is due to additional small phase and amplitude differences between the fields created by the two diffraction processes caused by the thickness of the silver film.

Finally we shall discuss the shape of the resonance features observed in the transmitted zero-order. Figure 5.2.6 shows the zeroth order transmission and absorption for a conformally corrugated metal slab, a metal slab with a corrugation on the top surface only, and a metal slab with a corrugation on the bottom surface only. From these it is clear that transmission maxima or minima may occur on resonance, or that a Fano shaped resonance may occur (one which shows a maximum and minimum on either side of the resonance frequency). In order to understand this it is first necessary to understand how the features due to SPP excitation originate in the reflected and transmitted zero-orders (we shall only consider the 1st diffracted order processes)

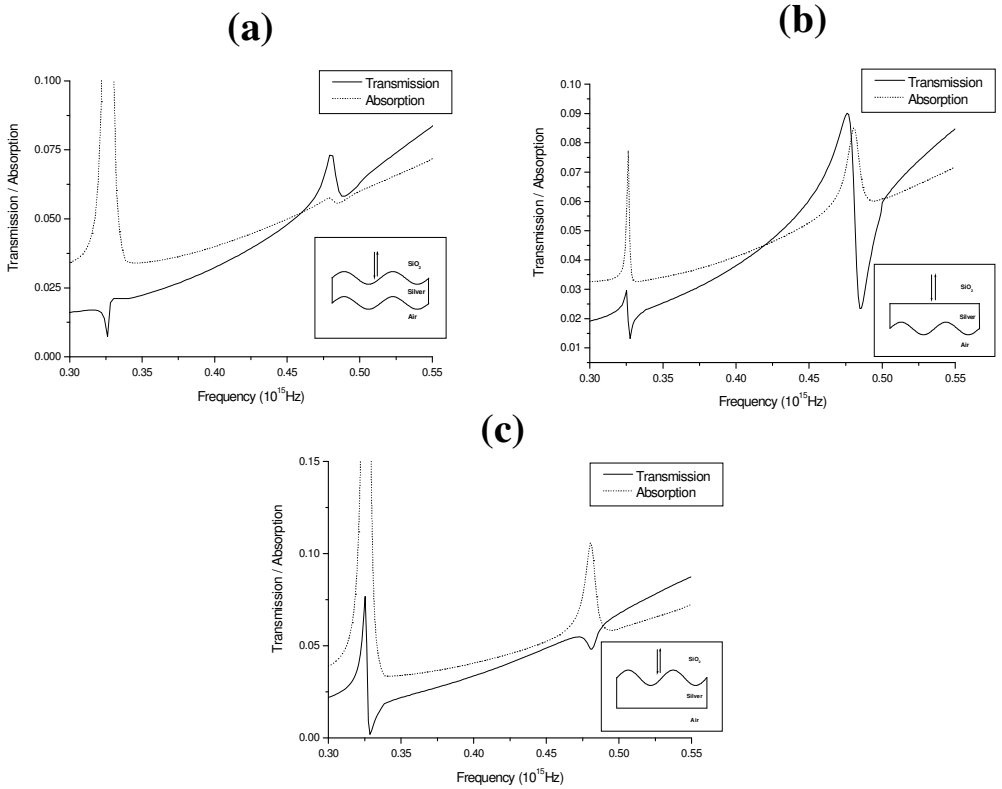

Figure 5.2.6. The zeroth order transmission, and absorption in the metal slab, for the same structures as used for figure 5.2.1.

For light incident upon a semi-infinite metal grating the phase of the E-field of the specularly reflected order is changed by 180° with respect to the incident light. Also, the evanescent diffracted orders are 90° out of phase with the incident light and, due to the fact that the SPP is resonantly driven by the evanescent order (producing another 90° phase shift), the SPP is in anti-phase with the incident light. When light is re-radiated from the SPP this process is repeated so that the re-radiated light is in anti-phase with the specularly reflected order, thereby cancelling and producing a reflectivity minimum.

When an optically thin metal film is investigated there is similarly a 180° phase change from the SPP when it is re-radiated into the transmitted zero-order. Figure 5.2.4(c) shows the resultant diffracted field phase and magnitude for the combined diffraction processes from the top and bottom surfaces of a thin silver film in a conformal geometry, and this shows that the phase on resonance is approximately 140° with respect to the incident light. Due to the metal film thickness the phase of the zero-order for a planar film with the same average thickness is approximately -40°, and therefore when the 180° phase change to the SPP upon re-radiation is taken into account these are in phase with each other, and result in a transmission maximum. This is as observed in figure 5.2.6(a).

This same type of analysis can be used to understand the other transmission features observed on the structures investigated in this chapter. The transmission minima occur when the re-radiated light is in anti-phase with the zero-order transmitted light, and the Fano shaped resonances occur when the re-radiated light is ±90° out of phase with the zero-order transmitted light. Of course, this is very much simplified here as in most real cases (for example when the phase between the two gratings on either surface is not a simple multiple of $\pi/2$) the phase between the re-radiated light and the transmitted zero-order will not be 0°, 90°, or 180°, but rather somewhere between. Also, the shape of the resonance will depend strongly upon the relative intensities between the re-radiated light and the transmitted zero-order, therefore being strongly dependent upon the film thickness and grating profile.

5.3 Coupled SPPs on Thin Metal Slabs Corrugated on Both Surfaces

In the previous section we considered the case of a thin metal slab corrugated on both surfaces, but bounded by dielectrics with different refractive indices in order to separate the SPP excitation frequencies for the two interfaces. In this section we shall consider the case where the refractive indices of the two bounding dielectrics are identical (we shall consider them to both be air). In this situation we can no longer consider the SPPs on the two surfaces separately as they couple together and form two different types of SPP, the long range SPP (LRSPP), and the short range SPP (SRSPP). We shall first describe these SPP modes for a planar interface, before considering the situation of a thin metal slab corrugated on both surfaces in either a conformal geometry, or with the two corrugations phase shifted with respect to each other.

5.3.1 Coupled SPPs on Planar Metal Slabs

As described in the previous section it is possible to excite the SPP on both surfaces of an optically thin corrugated metal slab. The corrugations are needed to allow coupling to the SPP modes, but the SPPs may be excited on a planar surface if, for example, a prism coupling technique is utilised (see chapter 2). The SPP on a planar surface is simpler to explain, and therefore in introducing the concept of coupled SPPs we shall again begin by describing the coupled SPP in this situation.

Consider an optically thin metal slab bounded by dielectrics with identical dielectric functions with an SPP excited on the top interface. The exponentially decaying fields of this SPP into the metal may excite another SPP on the bottom interface if the metal film thickness is small enough. These SPPs will then couple together, and in doing so they form two different types of coupled SPP modes, known as the long range SPP, and the short range SPP. These two SPPs have different charge distributions and these are shown in figure 5.3.1.1.

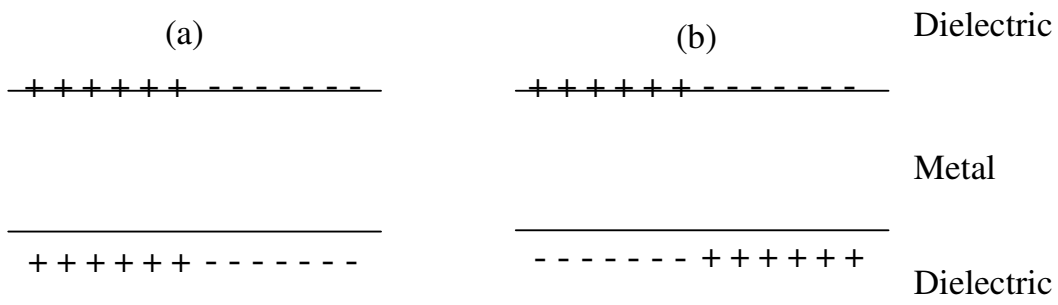

Figure 5.3.1.1 The charge distributions for a) the short range SPP, and b) the long range SPP.

The short range SPP has a symmetric charge distribution, and the long range SPP has an anti-symmetric charge distribution. These coupled SPPs are described as long range and short range simply because their propagation distances are very different from each other. These different propagation lengths arise due to a different level of damping for the two modes, and this difference in damping is caused by the different field distributions associated with each mode. The instantaneous electric field distributions (parallel and perpendicular to the metal slab) for the two modes are shown in figure 5.3.1.2. These plots were obtained for a planar metal film by modelling the system with prism coupling to the modes (a prism / air / metal / air structure), and the distances on the x-axis are measured from the prism surface. The metal slab begins 500nm from the prism, and is 30nm thick, with the permittivities of the silver $\varepsilon_r = -17.0$, and $\varepsilon_i = 0.6$.

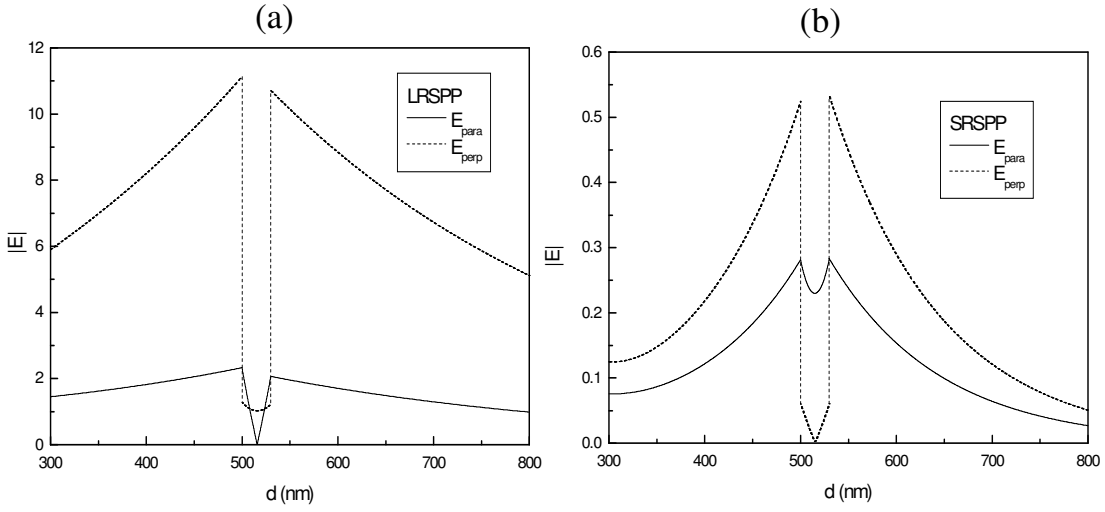

Figure 5.3.1.2 Instantaneous electric field profiles (parallel and perpendicular components) for the LRSPP (a) and SRSPP (b)

From figure 5.3.1.2 it is clear that the LRSPP has a lower proportion of its fields within the metal film than has the SRSPP, and hence the attenuation of the mode is lower than that of the SRSPP. As the film thickness is decreased the proportion of the fields within the metal film is further diminished, leading to a LRSPP with an increased range. For a 15nm silver film the range is approximately 0.5mm, compared to approximately 43μm for a single interface SPP (see chapter 2).

This diminishing of the proportion of the fields within the metal for the thinner slabs for the LRSPP also corresponds to an increase in the field enhancement at the metal surface. For a single interface SPP on a silver film at 632.8nm instantaneous field enhancement factors of the order of 30-40 can be achieved, however for the LRSPP on a 10nm slab with the same permittivities (which will not in fact be the case for very thin films, but can be used as an approximation) the instantaneous field enhancements are of the order of 600 (Sarid [1982]). These high field enhancements have meant interest in the LRSPP for non-linear applications (see, for example, Stegeman, Burke and Hall [1982], and Hickernell and Sarid [1986]).

In common with the SPP on a planar single interface, coupled SPPs may not be optically excited on planar systems without some method of increasing the wavevector of the incident light. In this section we shall consider grating coupling with both interfaces corrugated.

5.3.2 Coupled SPPs on Conformally Modulated Thin Metal Slabs

The dispersion relation of coupled SPPs on a planar metal slab involves a splitting of the single planar interface SPP dispersion curve into two separate dispersion curves, which take the same form as the single interface case, but with the LRSPP curve shifted up in frequency (for a given in-plane wavevector), and the SRSPP curve shifted down in frequency (for a given in-plane wavevector). When a corrugation is added to the surfaces these two curves are scattered back into the region of ω-k space available to incident radiation (in the same way as described for the single interface system in chapter 2), and these ‘coupled SPPs’ may then be coupled to by incident light (Dupta-Gupta, Varada, and Agarwal [1987], Inagaki, Motosuga, Arakawa, and Goudonet [1985]).

The size of the splitting between the LRSPP and the SRSPP is predominantly determined by the thickness of the metal slab. This is shown in figure 5.3.2.1, in which the zeroth order reflection, transmission (log scale), and absorption of the system has been calculated as a function of frequency and slab thickness for a thin metal slab described as silver (with separate polynomials describing the frequency dependent real and imaginary parts of its dielectric function), and with both surfaces corrugated identically with a 10nm amplitude, 400nm pitch, sinusoidal grating.

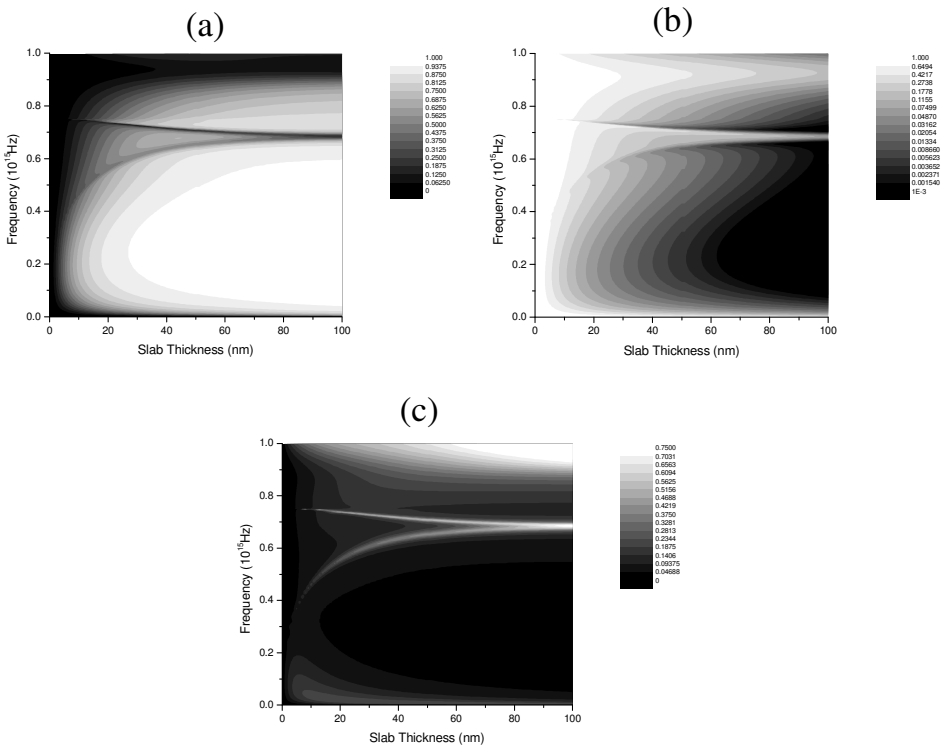

Figure 5.3.2.1 The optical response as a function of frequency and slab thickness for $k_x = 0$ on a conformal, sinusoidally corrugated silver slab of 10nm amplitude and 400nm pitch, in the classical mount.

At large slab thickness the two branches of the coupled SPPs converge to the frequency at which the SPP would occur on a single interface silver grating of the same parameters. However, as the slab thickness is reduced it is clear that the LRSPP approaches the diffracted order light line (at $f = 0.75 \times 10^{15}$Hz) asymptotically, and that the SRSPP reduces in frequency and rapidly approaches 0 for very thin slabs.

The cause of this asymptotic limit for the LRSPP is clear when the charge distributions and fields of the mode are considered. In figure 5.3.1.1 we have shown that the LRSPP has an anti-symmetric charge distribution on either side of the metal slab. Therefore, the fields of the system outside of the metal slab consist of field lines normal

to the surface pointing towards the metal slab at the negative charges on one side, and away from the metal slab at the positive charges on the opposite side. Therefore, at any instant in time, the fields outside of the metal slab resemble a photon propagating parallel to the average plane of the surface. Inside the metal slab the field lines are pointing in the opposite direction to those outside of the metal slab, and therefore the energy of the mode is not directly on the diffracted order light line. However, as the thickness of the slab is reduced the mode increasingly resembles a photon since the proportion of the fields of the SPP contained within the metal slab is reduced.

The opposite is true of the SRSPP. From the charge distribution it is clear that the field lines outside of the metal slab are pointing in opposite directions on either side. Therefore, as the grating thickness is reduced, its resemblance to a photon decreases and it has to move away from the diffracted order light line.

We have described the form of the dispersion of the mode as a function of slab thickness for normal incidence, but we have not considered the coupling to the SRSPP and LRSPP, or the nature of the resonant features in the zeroth order reflection and transmission obtained from the structure.

In reflection, both the SRSPP and the LRSPP are shown as reflectivity minima. This is not surprising since the predominant mechanism by which this feature will occur is from scattering of the SPPs from the corrugation on the top surface (if it is scattered from the bottom surface it will be attenuated due to propagation through the silver slab), which will then interfere with the specularly reflected light in the same way as described for the single interface metal grating considered in chapter 2. The field distributions at the top surface for the SRSPP and LRSPP are very similar since they have the same charge distribution. Therefore they both result in reflectivity minima since the re-radiated light is in anti-phase with the specularly reflected light.

In transmission, the resultant zeroth order transmitted fields will be predominantly due to scattering from the bottom surface, and the charge distributions for the LRSPP and SRSPP are 180° out of phase with each other. Therefore, the two modes will show the opposite features in the zeroth order transmitted intensities. This can be seen from figure 5.3.2.1 (b) where the feature due to the SRSPP is evident as a transmission minimum followed by a maximum, and that due to the LRSPP shows a transmission maximum followed by a minimum. This implies that the phase of the re-radiated light is approximately +90° out of phase with the directly transmitted light in

the case of the LRSPP, and approximately -90° in the case of the SRSPP. It should also be noted that the absorption due to the two coupled SPP modes are similar to each other, though the absorption due to the LRSPP for lower slab thickness is slightly higher than that of the SRSPP.

We shall now consider the case of changing in-plane wavevector (incident angle) on the nature of the coupled SPPs. In figure 5.3.2.2 we show the dispersion relation obtained from the scattering matrices, and the zeroth order reflection, transmission, and absorption of the system. The slab thickness is 30nm, with the rest of the parameters describing the system the same as for figure 5.3.3.1.

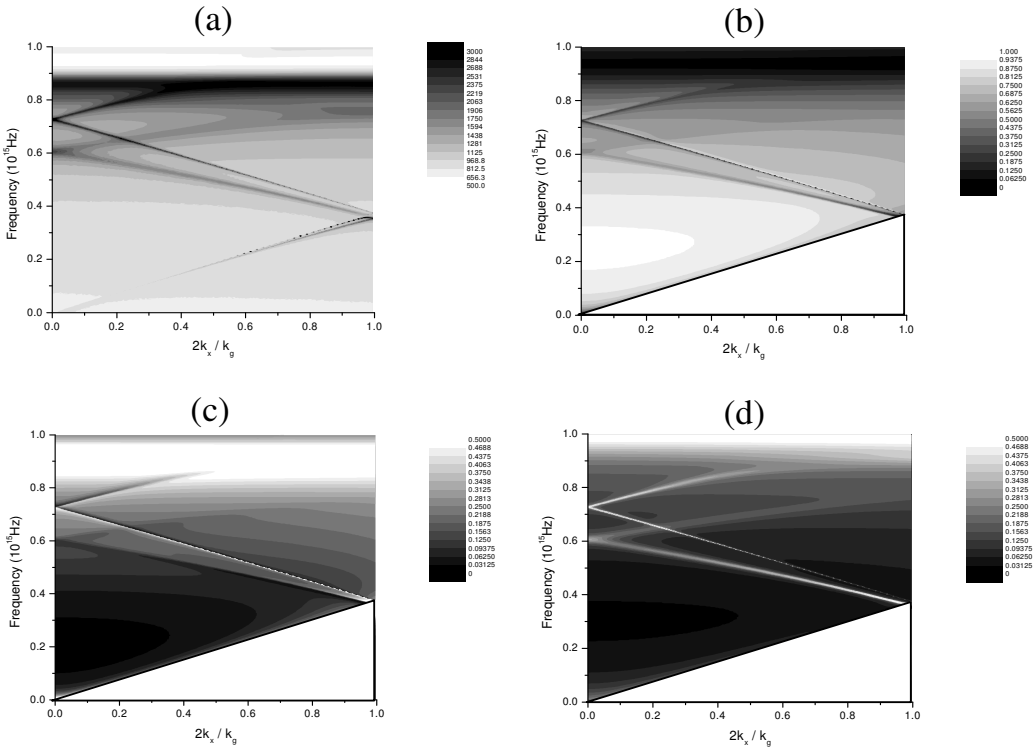

Figure 5.3.2.2 The optical response of the same system described for figure 5.3.2.1 as a function of frequency and in-plane wavevector. a) The dispersion of the modes obtained from the scattering matrices, b) the reflection, c) the transmission, and d) the absorption for the system.

Figure 5.3.2.2(a) confirms the previous comments regarding the form of the dispersion curve of the long range and short range SPPs on a conformally corrugated thin metal slab. Also to be noted is the change in coupling condition for the two modes as a function of the in-plane wavevector. For high values of k_x the coupling to the SRSPP is stronger than for low values of k_x, whereas the reverse appears to be the case

for the LRSPP. However, the nature of the reflection and transmission features due to the excitation of the coupled SPPs appears to remain unchanged.

We shall now consider the more complicated case of adding a higher harmonic to the grating profile. Higher harmonics are likely to occur in the manufacture of thin slab metal gratings if the corrugations are formed by interferometric lithography due to non-linearities in the exposure and development processes (to be described in chapter 8) with the metal film then deposited via vacuum evaporation or sputtering. In the single interface case (chapter 2) we described the formation of band gaps in the dispersion curve of the SPP due to these higher harmonics, and therefore it is clear that they can have a significant effect on the optical response of gratings. In figure 5.3.2.3 we show similar plots to those in figure 5.3.2.1, but in this case there is a 5nm amplitude $2k_g$ component +90° out of phase with the fundamental k_g component (of amplitude 10nm) in the grating profiles.

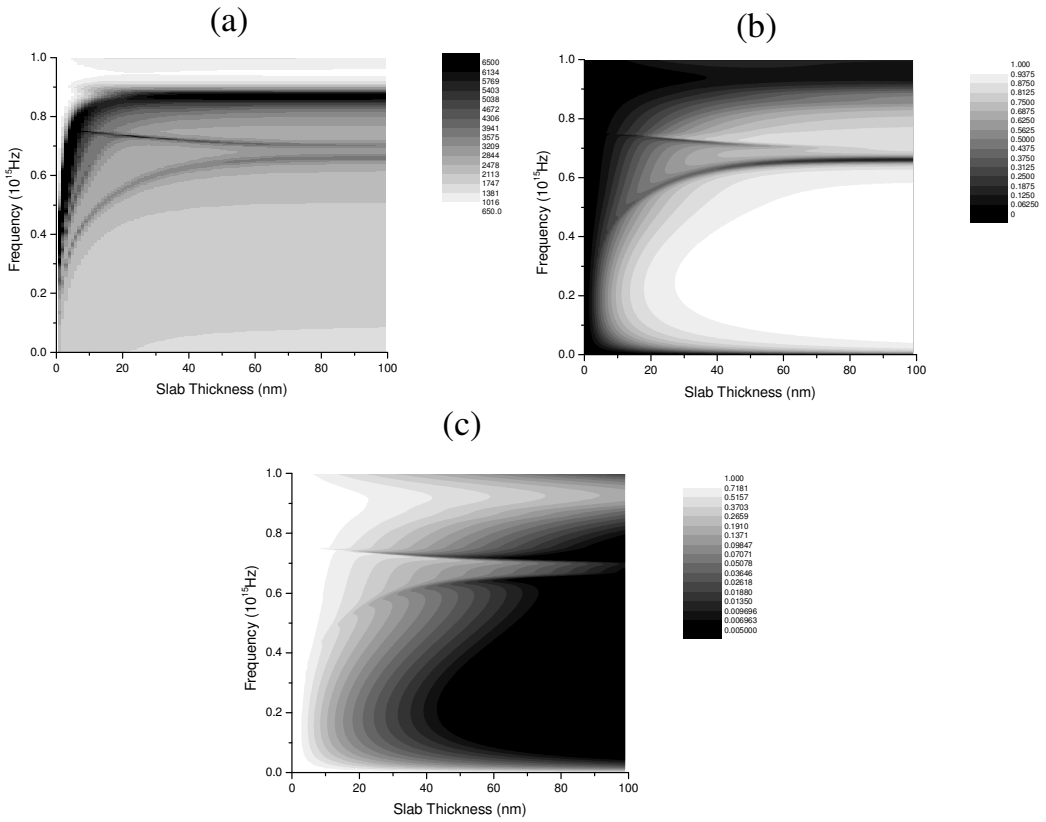

Figure 5.3.2.3 The optical response of the same system as figure 5.3.2.1, but with an additional $2k_g$ component in the grating profile description of 5nm amplitude, as a function of frequency and slab thickness. a) the position of the modes obtained from the scattering matrices (z axis units are arbitrary), b) the reflection and c) the transmission (log scale) of the system.

In the case of a purely sinusoidal grating profile when the slab thickness was large there was a single SPP mode, which was identical to the single interface SPP. When a $2k_g$ component is added to the grating profile a band gap opens (as described in chapter 2). This is the reason that there are two bands for large slab thickness in figure 5.3.2.3(a). For lower slab thickness the LRSPP and SRSPP form directly from the high and low energy edges of the bandgap respectively. This is somewhat surprising since the symmetry of the charges on the upper surface for the SRSPP and LRSPP are identical, whereas those of the high and low energy band edges are not. In fact, for the upper energy band edges, when the phase of the $2k_g$ component is +90° out of phase with the k_g component, the maximum surface charge densities occur on the grating peaks and troughs, whereas for the LRSPP which forms from it they are on the midpoints between the peaks and grooves.

In order to understand this it is useful to investigate the dispersion curves of the SPPs of the system for different slab thickness. This is shown in figure 5.3.2.4, where the dispersion curves are shown for slab thickness of 70nm, 50nm, and 30nm.

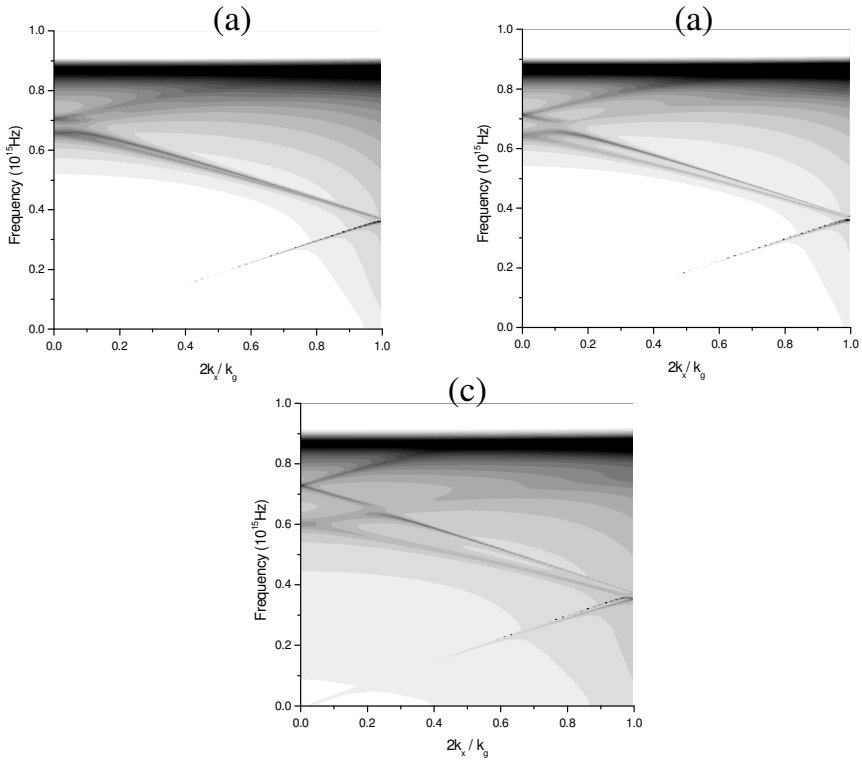

Figure 5.3.2.4 The dispersion curves of the SPPs of the system for a conformally corrugated thin metal slab with a pitch of 400nm, and a grating profile described by a 10nm amplitude k_g component, and a 5nm $2k_g$ component, with the phase of the $2k_g$ component being +90° out of phase with the k_g component. a) d = 70nm, b) d = 50nm, and c) d = 30nm (z axis units are arbitrary).

From figure 5.3.2.4 it is clear that, for a 70nm thick metal slab, there is anti-crossing between the LRSPP and the SRSPP very close to the position of the band gap. In fact, due to the fact that this anti-crossing is so close to the band gap, the modes around this point must be considered as having mixed SRSPP and LRSPP character. For decreasing slab thickness the anti-crossing occurs further from the $k_x = 0$ axis, so that the modes at $k_x = 0$ becomes more LRSPP and SRSPP in character the thinner the slab becomes. This leads us to the conclusion that when the slab thickness is made large enough that the SRSPP and LRSPP overlie each other the anti-crossing occurs at the position of the band gap, and that in this case the band gap and anti-crossing processes are the same. The charge distributions of the two band edges will then have the same form for both the LRSPP and SRSPP, since they are the same mode.

For a single interface grating with a $2k_g$ component the coupling to the modes for normally incident light depends upon the phase of the $2k_g$ component with respect to the k_g component, with coupling only possible to the mode with an anti-symmetric charge distribution on either side of a grating peak (see chapter 2). The same is true of the coupled SPP modes, and therefore, when the grating slab thickness is relatively large and the coupled SPP modes are largely similar to the single interface SPP, only one band edge is coupled to. As the layer thickness is decreased, so that the modes become more LRSPP and SRSPP like, coupling to both modes is possible since both possible charge distributions for the LRSPP and SRSPP modes overly each other in energy (and therefore frequency), and there is a solution with an asymmetric charge distribution on either side of a grating peak for both modes.

The reason for the fact that the two solutions for the two coupled modes overlie each other is that the light on the incident side 'sees' a $2k_g$ component with a +90° phase difference with respect to the incident light, whereas the light transmitted through the structure 'sees' a $2k_g$ component with a -90° phase difference with respect to the k_g component. The charge distributions on the two sides of the structure then occur in the troughs of the $2k_g$ component on one surface, and on the peaks of the $2k_g$ component on the other surface. This is true for both possible solutions for the two coupled modes, and since any energy difference between them will be due to the $2k_g$ component, both LRSPP modes and both SRSPP modes will have the same energy. The higher energy mode in figure 5.3.2.3 would, therefore, be expected to show no coupling for large slab thickness, but would show coupling for lower slab thickness. The opposite is true if the

phase of the $2k_g$ component is -90° out of phase with the k_g component, and the LRSPP could be coupled to for all slab thickness, whereas the SRSPP could only be coupled to for lower slab thickness

5.3.3 Coupled SPPs on Thin Metal Slabs in a Non-Conformal Geometry

In this section we shall consider the case of a thin metal slab bounded by dielectrics with identical dielectric functions (in this case air), and corrugated with identical grating structures on both surfaces, but with a phase difference between the two gratings. We will pay particular attention to the case where the system is anti-symmetric about the average plane of the structure.

We shall begin by considering the case where the two corrugations are perfectly sinusoidal. In figure 5.3.3.1 we show the zeroth order reflectivity, transmissivity, and absorption of the system for a 30nm thick silver slab (with the dielectric function described by polynomials) corrugated on both surfaces by a 400nm pitch, 10nm amplitude, sinusoidal grating, as a function of frequency and phase difference between the corrugations on the two interfaces.

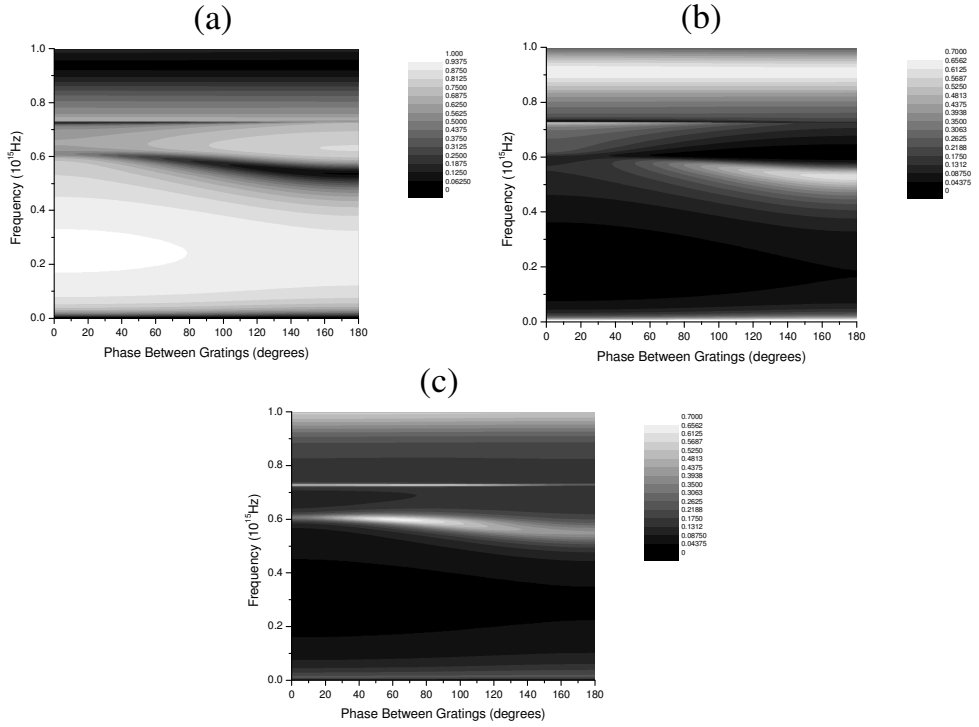

Figure 5.3.3.1 The zeroth order normal incidence reflectivity (a), transmissivity (b), and absorption of the system (c), for a 30nm thick silver slab corrugated on both surfaces with a 400nm pitch, 10nm amplitude, sinusoid as a function of frequency and phase between the corrugations on the two surfaces.

Before we consider the optical response of the system shown in figure 5.3.3.1 we shall consider the absorption. The maximum in absorption occurs at the frequency at which the coupled SPPs are excited, and it is clear from figure 5.3.3.1(c) that the excitation frequency is reduced when the phase difference between the corrugations on each surface is increased. Also, the coupling strength to the mode is altered as the phase difference is changed.

In order to understand this it is useful to consider an analogy to the coupled SPPs on these systems. If we have two pendulae connected together with a spring and we drive one harmonically then there are two possible steady state solutions: one in which the two pendulae oscillate in phase, and one in which they oscillate out of phase. These two solutions are analogous to the SRSPP and LRSPP on the thin metal slab systems, except that in this case the system is more complicated since the charge distributions on either side (the two pendulae in our analogy) can both be driven (by the reflected and transmitted diffracted orders). In section 5.2 we showed that the total transmitted diffracted order field of a doubly corrugated thin metal slab in a conformal geometry is nearly zero since the transmitted diffracted orders created at the two interfaces cancel. Therefore, in a conformal geometry we have the case where the charge distribution on only the incident interface is driven. When the phase of the bottom surface corrugation is changed with respect to the corrugation on the top (incident) surface the total transmitted diffracted order fields are no longer zero, resulting in the charge distributions on both interfaces being driven. The total transmitted diffracted order fields are at a maximum at a phase difference between the two corrugation of approximately 180° with some small change due to the thickness of the metal slab. When both charge distributions are being driven the frequency of the resonance is not necessarily going to be the same, and this is the reason that the SRSPP excitation frequency is reduced in figure 5.3.3.1.

The coupling strength of the SRSPP increases with increasing phase difference between the two corrugations, while the coupling strength to the LRSPP reduces, and this can be explained by considering the phase of the transmitted diffracted order with respect to the charge oscillations of the coupled SPPs on the bottom interface. This is analogous to the phase of the driving force on the second pendulum being different to the phase of oscillation of the second pendulum. If the driving force on the second

pendulum is in phase with the oscillation the coupling strength will be increased, whereas if it is out of phase the coupling strength will be decreased. Since the charge distributions on the bottom interface of the thin metal slab for the SRSPP and LRSPP are out of phase with each other this means that the coupling strength to one of the modes will be increased with an increase in the transmitted diffracted order, whereas the coupling strength to the other mode will be reduced. This is observed in figure 5.3.3.1.

The other interesting feature to note in figure 5.3.3.1 is the change in the features in the reflectivity and transmissivity plots due to the change in the phase between the gratings. In the previous section on conformally corrugated structures we stated that the features observed in the zeroth order reflectivity were predominantly due to the scattering of the coupled SPPs from the corrugation on the top surface of the structure, and that since the charge distribution of the SRSPP and LRSPP are the same at the top interface the features observed due to re-radiation from both modes would be the same. This is still true for this system, although it is now clear that there is a very small maximum occurring at a slightly higher frequency than the minimum due to the excitation of the coupled SPPs, and we believe that this is due to the small influence of the coupled SPPs being scattered from the corrugation on the bottom interface. For this reason there is also a small discrepancy between the frequency at which the maximum absorption occurs (which is at the exact frequency at which the coupled SPPs are excited) and the minimum in the reflectivities.

In transmission the predominant scattering process which affects the zeroth order transmissivity occurs at the bottom interface. The charge distributions of the LRSPP and SRSPP are 180° out of phase with each other meaning that the features in the zeroth order transmission are the opposite of each other. For the conformal geometry the feature due to the SRSPP is a maximum followed by a minimum, whereas the feature due to the LRSPP is a minimum followed by a maximum. Even if the phase of the bottom surface corrugation is changed with respect to the top surface corrugation the charge distribution of the two coupled modes remains unchanged, and therefore the nature of the features in the zero order transmission remain unchanged. The coupling strength changes due to the process described above.

The reflectivity, transmissivity, and absorption of the system, as a function of frequency and in-plane wavevector for the case where the two gratings are in anti-phase with each other is shown in figure 5.3.3.2.

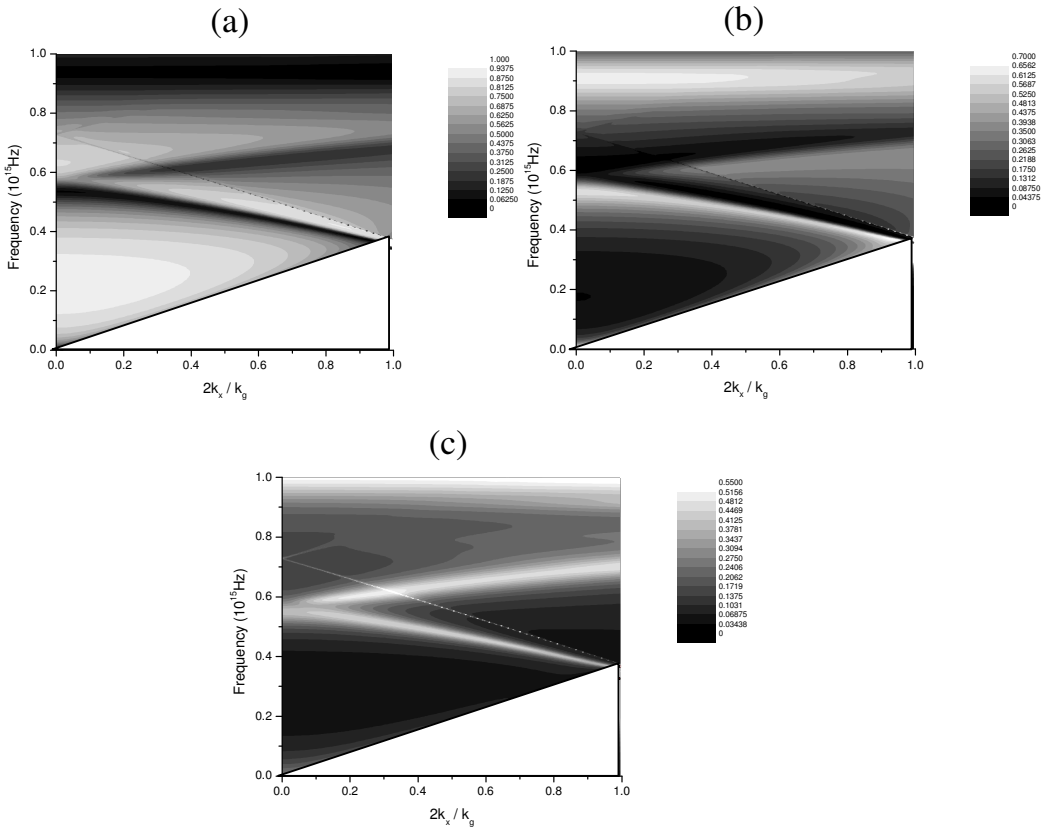

Figure 5.3.3.2 The reflectivity, transmissivity, and absorption of the system (a, b, and c), as a function of frequency and in-plane wavevector for the case where the two gratings are in anti-phase with each other for the same grating parameters as the system described previously.

The features in figure 5.3.3.2 for all values of the in-plane wavevector are the same as those described for the normal incidence case shown in figure 5.3.3.1. The LRSPP is observed to be a very weakly coupled mode, whereas the SRSPP is a very broad but well coupled mode. The form of the absorption of the system plot closely resembles that of the reflectivity, and the lack of absorption close to the point at which the SRSPP is 'reflected' from the $k_x = 0$ axis is due to the form of the features in the reflectivity plot. Since the resonance undergoes a reflection minimum followed by a small maximum, the absorption in the region of the small maximum is reduced.

We shall now consider the effect of adding a $2k_g$ component to the grating profile. The modelling presented is for the case where the $2k_g$ component on the top surface is +90° out of phase with the k_g component, although we shall also describe the differences in the optical response of the structure when the phase is other than this. Since we are particularly interested in the case where the two corrugations are in anti-

phase with each other it is not possible to plot the optical response of the system as a function of the phase between the two gratings. This is due to the fact that the sign of the $2k_g$ component on the bottom surface is reversed for the anti-phase case when compared to the conformal case, so that if we plotted the optical response as a function of the phase between the two corrugations we would not be investigating the change from a conformal geometry to that of an anti-phase geometry. Therefore, we shall only consider the case of the anti-phase structure here. A plot of the band structure, and of the reflectivity and transmissivity of the zeroth order for normally incident light, as a function of frequency and slab thickness is shown in figure 5.3.3.3 for a metal slab corrugated on both sides with the corrugations on the two surfaces in anti-phase with each other (the gratings have a 10nm amplitude k_g component, a 5nm $2k_g$ component, and a pitch of 400nm).

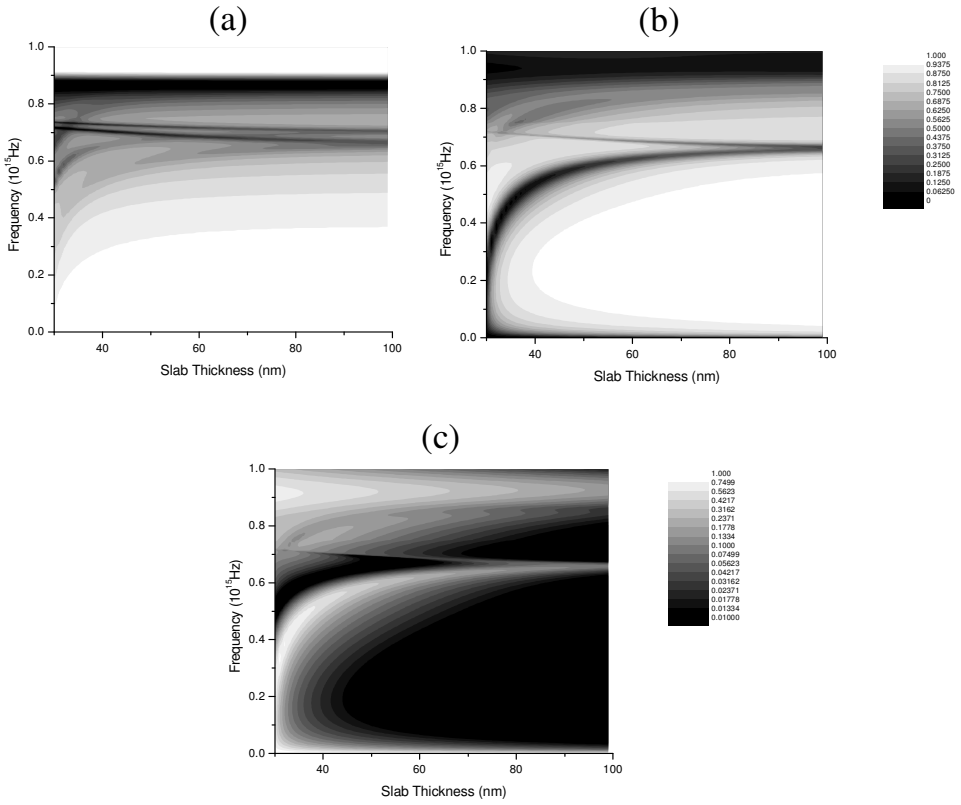

Figure 5.3.3.3 The band structure (a), reflectivity (b), and transmissivity (c) of the zeroth order, for normally incident light as a function of frequency and slab thickness for a metal slab corrugated on both sides when the corrugations on the two surfaces are in anti-phase (the gratings have a 10nm amplitude k_g component, a 5nm $2k_g$ component, and a pitch of 400nm).

It is clear from figure 5.3.3.3(a) that the dispersion of the modes is very different to that of the conformal case with a $2k_g$ component (figure 5.3.2.4). In that case there were two modes evident: one which corresponded to the LRSPP (which developed from the high energy branch of the SPP dispersion curve for large slab thickness), and one which corresponded to the SRSPP (which developed form the low energy branch of the SPP dispersion curve for large slab thickness). In figure 5.3.3.3, however, there are now four different modes evident. (In fact for low slab thickness there are five since the second order SRSPP had reduced in frequency to just below the two first order LRSPPs). Two of these correspond to LRSPPs and two to SRSPPs, and these develop from both the high and low energy band edges of the band gap of the SPP dispersion curve at large slab thickness.

In the conformal case the two LRSPP solutions, and the two SRSPP solutions had the same energy, and therefore they overlayed each other in the plots of figure 5.3.2.4. In the anti-phase case this is no longer true since the charge distributions of the LRSPP and SRSPP for the solution which has its maximum surface charge densities at the maxima and minima of the k_g component of the grating profile have a different thickness of metal between them than does the case where the maximum surface charge densities occur at the midpoints between maxima and minima of the k_g component. Therefore, the two LRSPP and SRSPP modes have different energies and occur at different frequencies for low slab thickness.

Another way of looking at this is to consider the charge distributions with respect to the $2k_g$ component of the grating profile on each surface. The surface charge density of the coupled SPP mode on the top (incident) surface corrugation is excited by the reflected diffracted order and 'sees' a $2k_g$ component which is +90° out of phase with the k_g component. The bottom surface charge density of the coupled SPP mode is excited by the transmitted diffracted order which, even though the phase of the $2k_g$ component of the grating is -90° out of phase with the k_g component, 'sees' the $2k_g$ component as being +90° out of phase. The LRSPP with surface charge density maxima on both surfaces occurring at the peaks and troughs of the k_g component has its electric field maxima at the peaks of the $2k_g$ component on both surfaces. For a single interface grating this would correspond to being the low energy band edge of the band gap at normal incidence. Correspondingly the LRSPP with surface charge density maxima on both surfaces at the midpoints between the maxima and minima of the k_g component has

its electric field maxima at the troughs of the $2k_g$ component on both surfaces, and this would correspond to the high energy band edge of the band gap at normal incidence for a single interface grating. Therefore, it is clear that the two possible LRSPP modes have different energies, with one resulting from the low energy band edge for large slab thickness, and one developing from the high energy band edge for large slab thickness. The same arguments are true for the two SRSPP modes.

It is clear from the reflectivity and transmissivity plots in figure 5.3.3.3 that coupling only occurs to the LRSPP and SRSPP developing from the low energy band edge at large slab thickness (with the shape of the resonance features having the same forms as those described in the previous section for anti-phase sinusoidally corrugated systems). This is because it is only the modes which have their surface charge density maxima at the midpoints between the maxima and minima of the k_g component of the grating profile which may be coupled to (for reasons described in chapter 2).

If the phase of the $2k_g$ component on the top surface (remembering that the $2k_g$ component on the bottom surface has the opposite sign) is -90° rather than +90° then the LRSPP and SRSPP modes developing from the high energy band edge at large slab thickness will be coupled to rather than those from the low energy band edge. If the phase is 0° then all four modes will be coupled to.

5.4 Summary

In this chapter we have initially considered the optical response of thin metal slabs where both surfaces are corrugated conformally, and where the thin metal slab is bounded by dielectrics with different dielectric functions for the incident and transmitted media. This has been performed in order to explain the results of Schröter and Heitmann (1999). When this system is conformally corrugated there is only very weak coupling to the SPP on the metal dielectric boundary on the transmission side of the system due to the fact that the transmitted diffracted orders created from the top and bottom surfaces of the metal slab cancel. When the phase of the corrugation on the bottom surface is changed with respect to the corrugation on the top surface then this cancellation effect is reduced so that the coupling to the bottom surface SPP increases. If the two corrugations are in anti-phase with each other the transmitted diffracted orders from the two interfaces interfere constructively so that the coupling to the bottom

surface SPP is at its maximum. We also described the shape of the features evident in the zeroth order transmission from the structure in terms of the phase between the re-radiated light from the SPPs and the zeroth order transmission which has propagated directly through the structure.

In the second section we proceeded to consider the case where the thin metal slab is bounded by dielectrics with identical dielectric functions, and in this system coupled SPP modes (the LRSPP and SRSPP) can be excited. We described the two coupled modes and the effect of the metal slab thickness and incident angle upon the features in the zeroth order reflectivity and transmissivity of the system when identical sinusoidal corrugations on the two surfaces are in phase with each other. The more interesting results occur when a small $2k_g$ component is added to the grating profile. In this case anti-crossing is evident between the LRSPP and SRSPP modes in the dispersion curves due to an interaction between them, and for large slab thickness this anti-crossing becomes the same as the band gap for normally incident light observed on single interface gratings. Due to this the LRSPP is observed to develop from the high energy band edge for reducing slab thickness, and the SRSPP from the low energy band edge.

Finally we considered the case where the two corrugations on each surface are non conformal, and we investigated the effect this has on the LRSPP and SRSPP modes. For purely sinusoidal corrugations the coupling strengths to the two coupled SPP modes alter as a function of the phase between the two corrugations due to the change in the magnitude of the transmitted diffracted order. The SRSPP mode couples more strongly with increasing phase difference between the two corrugations, whereas the LRSPP couples more weakly. This is due to the fact that the charge distributions of the SRSPP and LRSPP are 180° out of phase on the bottom surface.

The effect of adding a $2k_g$ component to the grating surface when the two corrugations are in anti-phase with each other is that the two possible LRSPP and SRSPP modes have different energies resulting in four different coupled SPP modes existing, with both LRSPP and SRSPP modes developing from the high energy and low energy band edges at large slab thickness. How many of these modes are evident in the optical response of the grating structure is dependent upon the phase of the $2k_g$ component with respect to the k_g component.

Chapter 6

Dispersion of Surface Plasmon Polaritons on Narrow-Grooved Short-Pitch Metal Gratings

6.1 Introduction

In the previous chapters we have considered the optical response of relatively shallow grating structures. In this, and subsequent, chapters we shall consider the optical response of single interface metal gratings when the grating depth is very large, with this chapter being concerned with arrays of Gaussian grooves. We shall pay particular attention to the zero order region of the spectrum, where, until recently, it was believed that SPPs could not be excited. In this region, when the grating grooves are deep enough, a family of self-coupled SPPs (to be defined later) have been shown to exist, which are very flat banded, and therefore lead to resonant absorption of light of a particular frequency over a wide range of incident angles.

In the first section of this chapter we shall briefly discuss the previous work which has been performed on deep metal gratings. Following this we shall show how the flat-banded SPP dispersion curves form from the well known shallow grating dispersion curve (described in chapter 2) due to the formation of very large band gaps, before describing the self-coupled SPPs excited with TE polarised light when the grating is oriented at a 90° azimuthal angle, and the dispersion of one of these self-coupled SPP bands for all grating orientations.

6.2 Previous Work on Deep Metal Gratings

About 20 years ago a grating with a depth of up to 200nm with a pitch of the order of 500nm was considered to be very deep. A theoretical study was undertaken by Andrewartha, Fox and Wilson [1979a,b] on perfectly conducting lamellar structures with approximately these parameters in the classical mount which described the dispersion curve of resonance anomalies obtained from identifying poles in the complex

wavelength plane. They showed the formation of flat bands, but only investigated small regions of ω-k space which did not include near normal incidence, or the zero order region of the spectrum. It is also questionable whether the modes they were identifying would be the same as SPPs on finitely conducting metal gratings, since no dispersion of the dielectric function of the metal was included which is an integral part of the dispersion relation of SPPs. Similar work was performed by Laks, Mills and Maradudin [1981] but, although they considered sinusoidal gratings with a frequency dependent dielectric function for the metal described by a free electron model, they only considered the non-radiative region of ω-k space, and for lower aspect ratio structures than those of Andrewartha *et al* [1997a,b].

Recent advances in computational power, and in the development of manufacturing techniques capable of producing high aspect ratio sub-micron structures, have stimulated interest in gratings with higher aspect ratios (where the depth of the grating is of the order of, or greater than, the grating pitch - this is our definition of the term deep). Experimentally it has been found that deep lamellar gratings can support highly localised resonances within the grooves with grating pitches of 1.75μm and depths of up to 1μm (Lopez-Rios, Mendoza, Garcia-Vidal, Sanchez-Dehesa and Pannetier [1998], Garcia-Vidal, Sanchez-Dehesa, Dechelette, Bustarret, Lopez-Rios, Fournier and Pannetier [1999]). These flat-banded resonances are very different to SPPs on shallow gratings, and have been explained as being due to hybrid waveguide-SPP resonances. However, the evolution of these bands as a function of groove depth was not investigated and has not been explained in any detail. An extension of the studies of deep gratings conducted by Porto, Garcia-Vidal and Pendry [1999] evaluated the band structures for lamellar transmission gratings with a pitch of 3.5μm and a depth of 4μm. This also showed flat-banded resonances in the infra-red region of the spectrum and predicted almost total resonant transmittance of the incident light in the zero-order region of the spectrum. Other recent developments have shown that resonant enhanced optical transmission can take place through hole arrays in classically opaque metal films (Ebbesen, Lezec, Ghaemi, Thio and Wolff [1998], Ghaemi, Thio, Grupp, Ebbesen and Lezec [1998]) which is also associated with SPP excitation (Saloman, Grillot, Zayato and de Fornel [2001], Martin-Moreno, Garcia-Vidal, Lezec, Pellerin, Thio, Pendry and Ebbesen [2001]).

All of the work described so far has related to diffractive structures. Until recently it had been assumed that, if the corrugation has a pitch less than half the wavelength of the incident radiation (in other words it is zero-order, or non-diffractive), the interface would act as a good mirror. The reason for this is that, as well as having no available diffracted orders, it was thought that SPPs would no longer be excited since, even with the added momentum available to the incident radiation due to scattering from the grating, the momentum of SPPs on such a structure is greater than the photon momentum available to the radiation in the zero-order region of the spectrum. However, Sobnack, Tan, Wanstall, Preist and Sambles [1998], and Tan, Preist, Sambles and Wanstall [1999] have shown that, for very deep zero-order monogratings, the SPP dispersion curve may be so severely modified from the shallow grating case that resonant absorption of light due to SPP excitation may occur within the zero-order region of the spectrum. A family of these resonances have been demonstrated to exist, and these have been termed self-coupled SPP resonances (for reasons which will be made clear later in this chapter). They are flat banded, possessing near zero group velocity over a large range of incident wavevectors (figure 6.2.1).

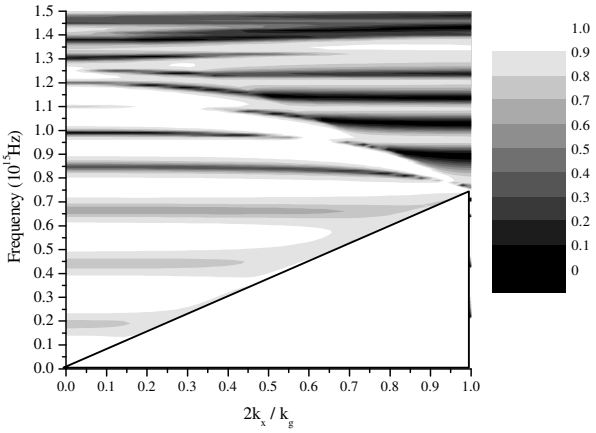

Figure 6.2.1. Reflectivity as a function of frequency and in-plane wavevector for TM polarised light incident on a 300nm deep, 50nm wide Gaussian-grooved, 200nm pitch silver grating held at a zero degree azimuthal angle. The bands are very flat for a large range of incident wavevectors.

In this chapter we shall investigate these self-coupled SPPs, and specifically we shall attempt to bridge the gap between the well understood band structure of SPPs on shallow gratings and those of the high aspect ratio structures on which they are excited. It will also extend the previous work of Tan *et al* [1999] to explore the SPP band structure for zero-order gratings at orientations other than that of the grooves perpendicular to the plane of incidence of the exciting EM radiation.

6.3 Band gaps and the Formation of Self-Coupled SPPs

In Chapter 2 the formation of band gaps in the SPP dispersion curves for shallow gratings was described, and it was mentioned that it was not possible to analytically obtain the size and central frequency of the band gap for deep gratings. Therefore, in this chapter we shall numerically obtain the dispersion curves for zero-order gratings for the visible region of the spectrum (λ_g = 200nm) for increasing depths, which will show the way band gaps develop for deeper gratings, and how they are responsible for the formation of a family of self-coupled SPPs in the zero-order region of the spectrum. In order to do this we have used computer codes based upon the method of Chandezon (Chapter 3). We have described the frequency dependent dielectric function of the metal as that of silver using the Drude model described in Chapter 2, with a plasma frequency of $\omega_p = 1.32 \times 10^{16}$ s^{-1}, and a relaxation time of $\tau = 1.45 \times 10^{-14}$ s.

If a grating which has a large first harmonic with a phase of +90° is investigated then a band structure which shows a large band gap with strong coupling to the lower energy branch is produced (described in chapter 2). An example of such a structure is a Gaussian groove profile, and it is this structure which is investigated in the following work in this chapter (figure 6.2.1). It is also used because previous work has shown the SPP resonances on deep Gaussian grooved gratings to be highly localised within the grooves. The use of Gaussian grooves on a grating of 200nm pitch reduces interactions between neighbouring grooves, allowing the band structure of more highly localised modes to be obtained.

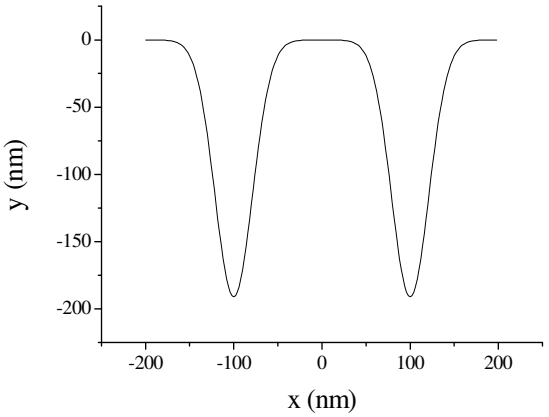

Figure 6.3.1 An example of the type of grating structure investigated. The profile is made up of a series of Gaussian grooves so that the depth, width and pitch of the grating may be altered independently, and so that in deep structures the SPP modes in neighbouring grooves will be largely isolated from each other.

Figure 6.3.2 shows the band structure, in the classical mount, for the first Brillouin zone of a 200nm pitch silver grating with 50nm wide Gaussian grooves for 10nm, 25nm, 50nm, and 75nm depths. In figure 6.3.2(a) the band structure looks very similar to that of a shallow sinusoidal grating. However, due to the large 1st harmonic needed to create the Gaussian profile, relatively large band gaps are opened (for example at between 1.1 and 1.2 x 10^{15}Hz at $k_x = 0$).

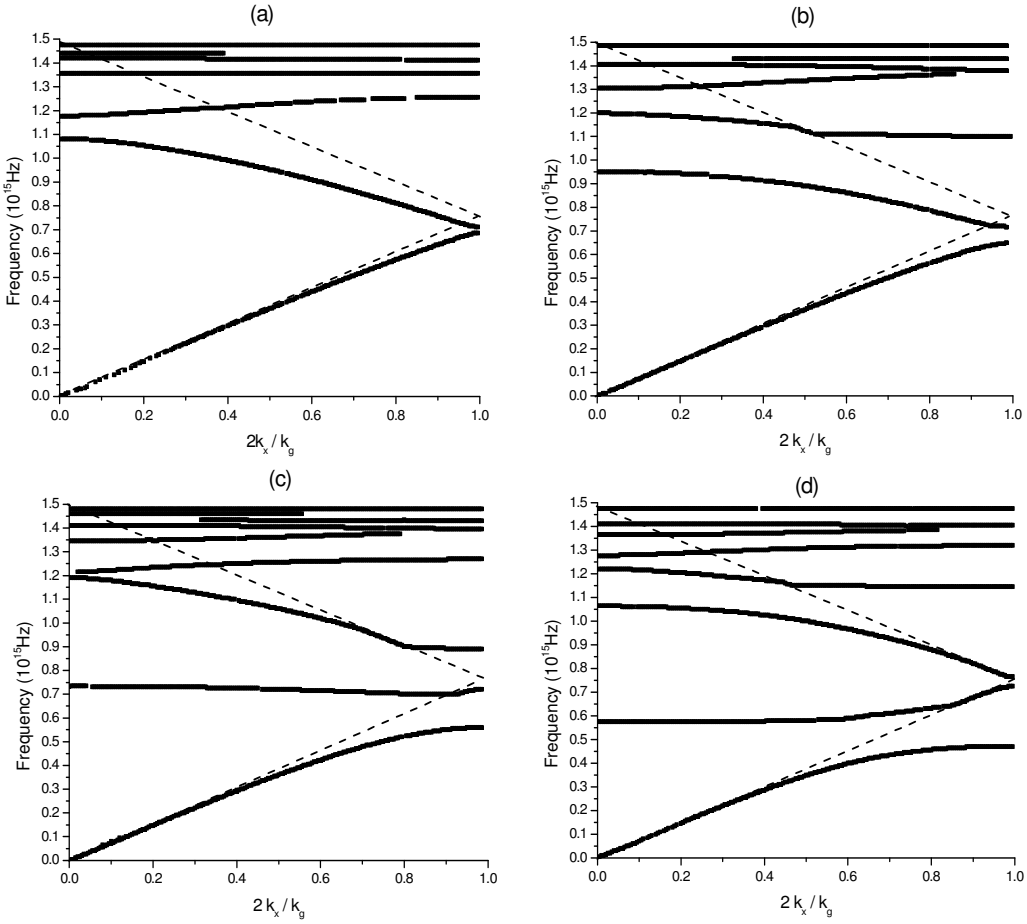

Figure 6.3.2 Dispersion curves for 200nm pitch, 50nm wide, Gaussian-grooved silver gratings for $k_z = 0$, with a) d = 10nm, b) d = 25nm, c) d = 50nm, and d) d = 75nm.

When the depth is increased these band gaps widen with the low energy solutions reducing in energy until, when the depth is 50nm (figure 6.3.2(c)), a very flat band is forming. On further increase of depth to 75nm (figure 6.3.2(d)) this branch, while remaining quite flat, has a positive gradient close to the light line, while the second order low energy branch has passed through the first order high energy branch and is in the process of forming a second flat band. The formation of these flat bands will be discussed in more detail later.

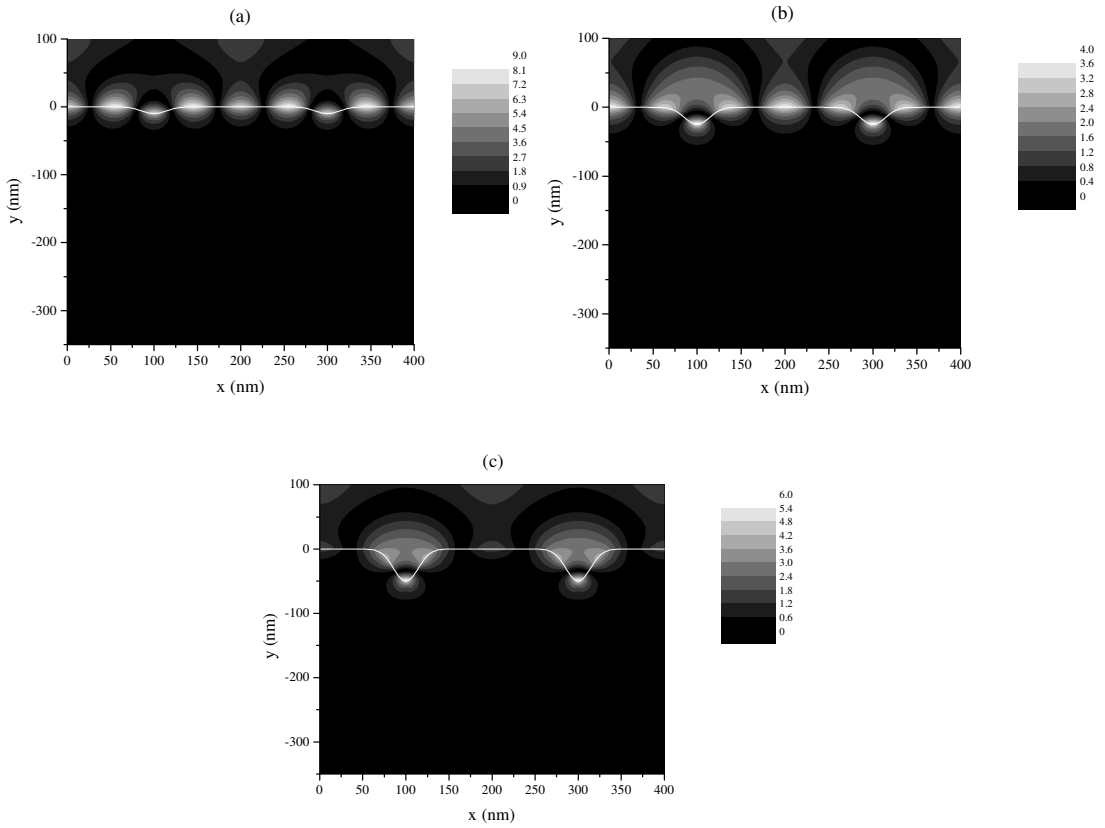

Figure 6.3.3 $|\mathbf{H}_z|$ for the second order SPP resonance for a) d = 10nm, f = 1.36 x 10^{15}Hz, b) d = 25nm, f = 1.3 x 10^{15}Hz, and c) d = 50nm, f = 1.19 x 10^{15}Hz.

To understand the nature of these resonances it is instructive to investigate their field profiles. Figure 6.3.3 shows the $|\mathbf{H}_z|$ (z is along the groove direction) component of the fields at $k_x = 0$ for the low energy branch of the second order TM resonance for 10nm (f = 1.36 x 10^{15}Hz), 25nm (f = 1.3 x 10^{15}Hz) and 50nm (f = 1.19 x 10^{15}Hz), deep gratings. For the resonance on a relatively shallow 10nm deep grating field maxima are observed on the tops of the grating between the grooves and also at the bottoms of the grating grooves. In addition to these there are two extra field maxima per grating period found on the grating groove 'shoulders'. (This arises because the second order SPP at $k_x = 0$ corresponds to a standing wave created by $\pm 2k_g$ scattering, thus the wave has 4 field amplitude maxima per grating period). When the depth is increased to 25nm the fields on the opposing shoulders of a groove appear to couple together, and by a depth of 50nm these coupled fields are almost entirely localised within the grating grooves. Since the fields of the SPPs on either side of the groove have coupled together these resonances can be described as self coupled SPPs (SCSPPs).

The dispersion curves for 100nm, 150nm and 300nm deep gratings are shown in figure 6.3.4. These show how the opening of even larger band gaps have pushed the low energy branches of the first, second and third order SPPs into the zero-order region of the spectrum. Indeed, the first order branch has been moved beyond the visible to the infra-red region of the spectrum. It is also noticeable that the character of the bands very much resembles that expected from anti-crossing of flat bands with the lightline. In fact, in previous work (Tan *et al* [1999]) it has been predicted that, for deep lamellar gratings, a series of almost flat bands with a small negative gradient would be expected. However, these flat bands interact with the free radiation and, since the density of states of the free radiation becomes divergent at grazing incidence, this leads to strong anti-crossing between the flat bands and the lightline as shown.

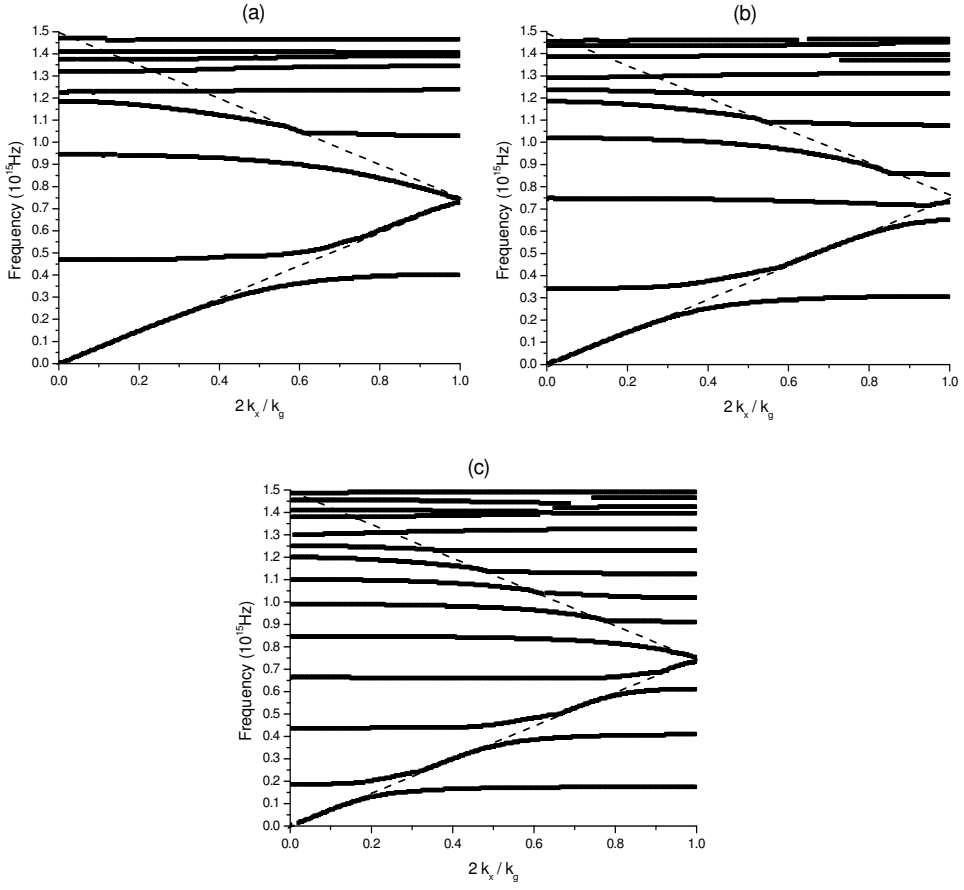

Figure 6.3.4 Dispersion curves for 200nm pitch, 50nm wide, Gaussian-grooved silver gratings for $k_z = 0$, with a) d = 100nm, b) d = 150nm, and c) d = 300nm.

The dispersion curve for the 300nm deep grating (figure 6.3.4(c)) shows several quite flat bands within the radiative regime. The $|\mathbf{H}_z|$ component of the fields for

successive flat bands at $k_x = 0$ are shown in figure 6.3.5. These show that there are a family of localised SCSPP resonances whose number of field maxima contained within the grating grooves is equal to the order of the mode (which corresponds to the multiple of k_g from which the branch originates).

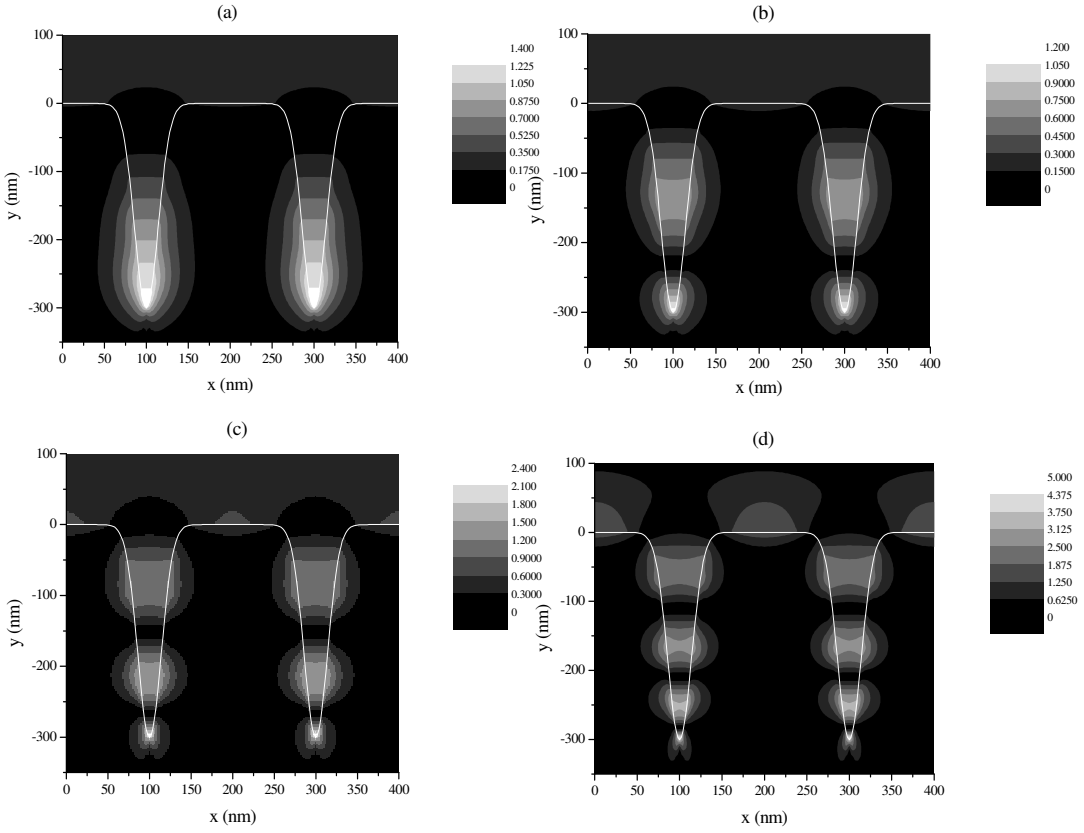

Figure 6.3.5. $|\mathbf{H}_z|$ for the first four SCSPPs on a 300nm deep, 50nm wide, Gaussian grooved grating with $k_z = 0$. a) f = 0.19 x 10^{15}Hz, b) f = 0.44 x 10^{15}Hz, c) f = 0.66 x 10^{15}Hz, and d) f = 0.85 x 10^{15}Hz.

The fields on the flat regions of the dispersion curves for k_x values away from the symmetry points of $k_x = 0$ and the Brillouin zone boundaries show similar field distributions to those observed at normal incidence, the principle difference being the intensity distribution of the fields. These intensity differences vary periodically along the grating, with the periodicity being determined by the ratio of $2k_x / k_g$ (e.g. at $2k_x / k_g = 0.5$ the fields in alternate grooves are the same (figure 6.3.6)).

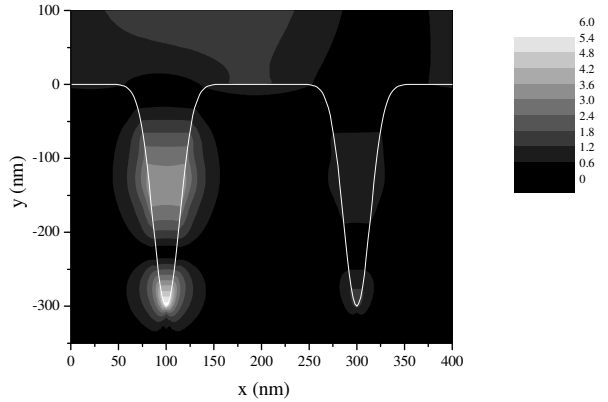

Figure 6.3.6 $|\mathbf{H}_z|$ for the second order SCSPP resonance on a 300nm deep, 50nm wide, Gaussian grooved grating with $2k_x / k_g = 0.5$ ($f = 0.46 \times 10^{15}$ Hz).

The nature of these resonances, and the mechanism by which they arise on short pitch gratings has been discussed. However, there are other aspects to the formation of these bands, and in order to discuss these it is useful to investigate the position of the SPPs in frequency as a function of depth (figure 6.3.7).

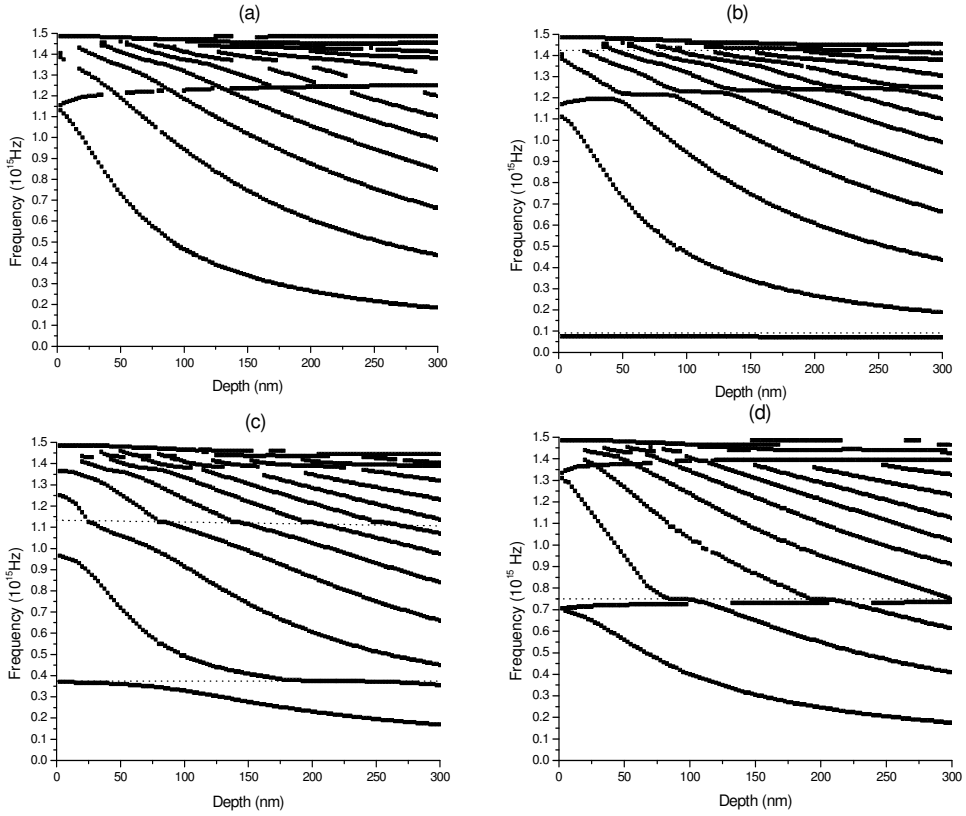

Figure 6.3.7 Mode frequency as a function of grating depth for 200nm pitch, 50nm wide, Gaussian-grooved gratings at $k_z = 0$. a) $2k_x / k_g = 0.0$, b) $2k_x / k_g = 0.1$, c) $2k_x / k_g = 0.5$, and d) $2k_x / k_g = 1.0$. Dotted lines are light-lines.

At $k_x = 0$ (figure 6.3.7(a)) the first order SPP (starting at $f \approx 1.15 \times 10^{15}$ Hz) shows a band gap opening as the depth is increased. The splitting is highly asymmetric with the low energy branch reducing in frequency rapidly with increasing depth, and the high energy mode changing comparatively little. In order to understand this asymmetry it is necessary to consider the coupling of the surface plasmon with the free radiation.

Earlier the SPP dispersion curve for a planar metal interface was described as a solution of Maxwell's equations. However, Maxwell's equations implicitly contain the coupling of the free radiation to the SP, and it is this coupling which creates the familiar flat-surface dispersion curve, rather than a straight line at ω_{sp}. When a grating structure is added and band gaps are opened at the Brillouin zone boundaries the high energy branches cannot increase in frequency above the light line (or its corresponding diffracted light line in the reduced zone scheme) since the coupling to the free radiation prevents it. For this reason the opening of the band gap is asymmetric; when the depth of the grating is increased the low energy branch may more readily reduce in energy relative to the high energy branch, which may only increase in energy up to that of the lightline, increasing asymptotically towards it with increasing depth.

When the band gaps increase at $k_x = 0$ the higher order low energy branches are free to pass through the lower order high energy branches with no interaction since this is at a symmetry point in ω - k space. At large depths the frequency at which these low energy SPP branches may be excited reduces as approximately 1/d. This can be simply predicted by considering the resonances as standing wave modes within the grooves.

As k_x is increased anti-crossing between higher order low energy branch SPPs and lower order high energy branch SPPs arises because the symmetry of the system has been broken. This can clearly be observed in figure 6.3.7(b) ($2k_x / k_g = 0.1$). At $2k_x / k_g = 0.5$ the anti-crossing is so strong that the branches are no longer well defined (figure 6.3.7(c)), and it is also noticeable that the coupling of the SPP bands with the free radiation causes them to tend towards the lightline (dotted lines at $f = 0.38 \times 10^{15}$Hz and 1.14×10^{15}Hz).

In figure 6.3.7(d) ($2k_x / k_g = 1.0$, at the Brillouin zone boundary) there is no longer any anti-crossing between the high and low energy SPP bands since it is another symmetry point in ω-k space. However, anti-crossing between the low energy branches and the lightline at $f = 0.75 \times 10^{15}$Hz is still evident.

Having described the SCSPP on metal gratings in the classical mount as a result of very large band gaps occurring in the dispersion curves of the SPPs we shall now extend this to the case of light incident at a 90° azimuthal angle.

6.4 Self Coupled Surface Plasmon Polaritons excited with TE polarised radiation

So far only the band structure in the ω-k_x plane with $k_z = 0$ (a 0° azimuthal angle) has been discussed. In this orientation only TM polarised radiation can be used to excite the SCSPP modes since a component of the radiation E-field must be perpendicular to the surface for excitation to occur. However, at normal incidence the case at a 0° azimuthal angle with TM polarised radiation is the same as for TE polarised radiation at an azimuthal angle of 90°. Therefore any SPP resonance excitable for one case should be excitable for the other. This is found to be the case, while the dispersion of these modes in the ω-k_z plane with $k_x = 0$ is significantly different from those in the other orientation (the ω-k_x plane with $k_z = 0$).

Figure 6.4.1 shows the dispersion curve along the k_z direction for the TE polarised resonances on a 300nm deep Gaussian grooved grating. There are three different features noticeable on this plot. Firstly, the diagonal line starting at the origin is the lightline. Moving away from this as k_z is increased is a curve which originates at the origin and is the flat surface SPP curve (slightly deformed by the grating in the x-direction). Thirdly, there are a series of bands which at normal incidence occur at the same frequencies as the TM excited resonances at normal incidence. These bands, though relatively flat, gently curve up in energy and, outside of the lightline, approach the grating-perturbed flat surface SPP curve. To understand the dispersion of these modes it is necessary to consider the SPP curve which has been scattered from the grating.

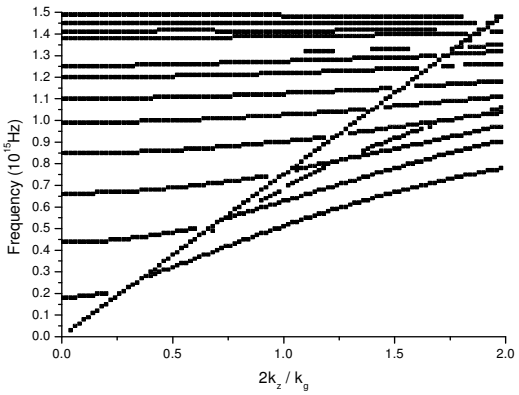

Figure 6.4.1 The dispersion curve for a 200nm pitch, 300nm deep, 50nm wide Gaussian-grooved grating with $k_x = 0$ and varying k_z.

In chapter 2 we discussed the fact that when a grating is added in the x-direction folding of the SPP bands at the Brillouin zone boundaries causes the horn shape of the SPP dispersion curve for a flat surface to fold into the region of ω-k space available to incident radiation. Another way of considering this band folding is to position the grating-deformed SPP curves at integer values of the grating vector in the x-direction to represent the scattering of the SPP from the grating surface. For a very shallow grating, where there is little deformation of the SPP dispersion curve, a slice in ω-k space through the plane at $k_x = 0$ will show the three types of features in figure 6.4.1. The lightline is present, as is the SPP horn centred at the origin. But there is also a curve which represents a slice through the scattered SPP curve centred at $k_x = k_g$. This is a band which rises in energy as k_z is increased and approaches the SPP curve centred at the origin at large values of k_z. Of course, at $k_x = 0$ two SPPs scattered from the grating are actually interacting; one scattered from $+k_g$, and one from $-k_g$. Therefore the mode is a standing wave in the x-direction, which propagates in the z-direction.

For a deep grating the band gaps and anti-crossing effects cause the SPP curve to deform in the x-direction, as discussed in the previous section. However, the effect of taking a slice through the scattered SPP curve in the ω-k_z plane at $k_x = 0$ produces bands which at normal incidence occur at a frequency defined by the deformation of the SPP dispersion curve caused by the grating. As k_z is increased the effect of the grating structure diminishes since the SPP is no longer propagating perpendicular to the grating grooves. This reduces the size of the band gaps causing the bands to curve up in frequency until at large k_z values they approach the dispersion curve centred at the origin.

It is noticeable in Figure 6.4.1 that, unlike the dispersion of these SCSPP bands in the k_x direction, there is no coupling with the free radiation at the light line. In the case of TM polarised radiation incident in the ω-k_x plane at $k_z = 0$ there is a periodic variation of the electric field in the direction of the grating vector when k_x is non-zero. Because of this the overlap integral of the wavefunctions of the incident light and the SCSPP mode will also be non-zero resulting in an interaction between them. By contrast, for TE polarised light incident in the ω-k_z plane at $k_x = 0$ there is no spatially periodic variation in the incident electric field in the direction of the grating vector for any values of k_z. Therefore, the overlap integral is zero and there can be no interaction. This is a special case and only occurs when $k_x = 0$. At all other orientations, where there is a finite k_x, coupling of the SCSPP mode with the free radiation can take place, and splitting occurs.

The form of the SPP dispersion curve for 0° and 90° azimuthal angles has now been described. In order to complete the discussion we shall now consider a single band of the SPP dispersion curve for any angle of incidence.

6.5 Band Structure for a SCSPP band for all grating orientations

When the grating is oriented such that both k_x and k_z are non zero there will be polarisation conversion due to the excitation of the SPP (as described in Chapter 2). The polarisation conserved and polarisation converted reflectivities as a function of frequency and in-plane wavevector for TM polarised light incident upon a 300nm deep, 40nm wide, Gaussian grooved grating oriented at a 45° azimuthal angle is shown in figure 6.5.1.

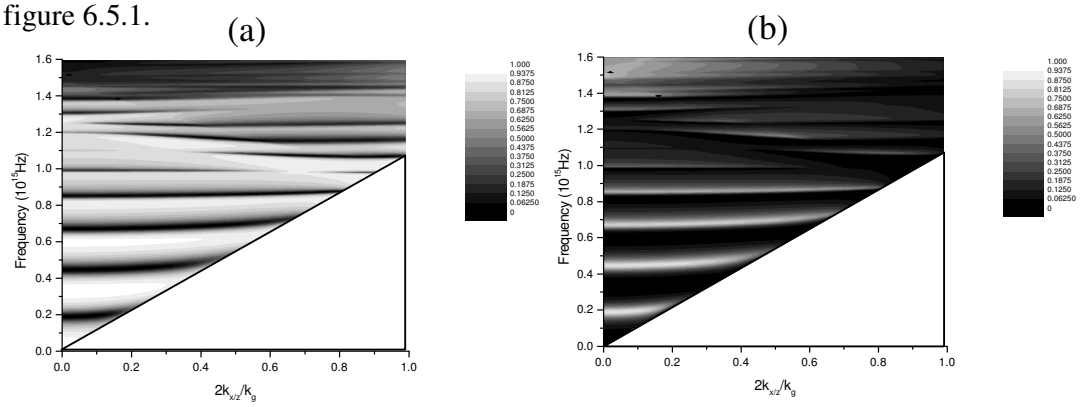

Figure 6.5.1 Polarisation conserved (a) and polarisation converted (b) reflectivities as a function of frequency and in-plane wavevector for TM polarised light incident upon a 300nm deep, 40nm wide, Gaussian grooved grating oriented at a 45° azimuthal angle.

It is clear that there is strong polarisation conversion occurring when the SCSPPs are excited. However, we are predominantly interested in obtaining the band structures for these resonances, and typically we have used the scattering matrices to obtain this in the previous sections. Unfortunately, due to the fact that there is anti-crossing between the SPP dispersion curve and the light line in the case of $k_z = 0$, and not for the case of $k_x = 0$, it is not possible to plot the band structure of a single SCSPP band using the scattering matrix technique. Therefore, we shall use a different method which will enable us to obtain the band structure in the radiative region only.

If we try to plot the band structure by identifying the minima in the reflectivities obtained from the structure we find that the results obtained are not in fact correct. This is because the polarisation conversion does not necessarily occur at the exact frequency of the SCSPP excitation, and this can produce a reflectivity minimum in the polarisation conserved reflectivity at a slightly different frequency to the true SCSPP excitation frequency. However, by calculating the absorption due to the excitation of a SCSPP resonance with both TM and TE polarised incident radiation it is possible to trace the SCSPP band for all possible grating orientations within the light line. The absorption (due to Joule heating in the metal) was calculated as a function of frequency for varying k_x and k_z within the lightline. By finding the peaks in the absorption the frequency at which resonant excitation of the SCSPP occurs may be found, and by plotting these the dispersion of a SCSPP band for all orientations of the grating with respect to the incident radiation may be obtained.

This has been performed for the second order SCSPP resonance on a 300nm deep, 40 nm wide Gaussian grooved grating, and the results are shown in figure 6.5.2. For the planes at $k_x = 0$, and $k_z = 0$, the curves have the forms described in the previous sections, and for values of k_x and k_z between these planes the band structure is seen to vary smoothly. In the ω-k_x plane at $k_z = 0$ this band can only be excited with TM polarised radiation, and in the ω-k_z plane at $k_x = 0$ only TE polarised radiation can excite it. For finite values of both k_x and k_z both polarisations can excite the mode, however the coupling strength will vary due to the magnitude of the E-field component perpendicular to the surface of the grating structure changing as the grating orientation with respect to the incident radiation is changed. Also, the intensity of the polarisation converted reflectivity will vary with azimuthal angle as described in chapter 2.

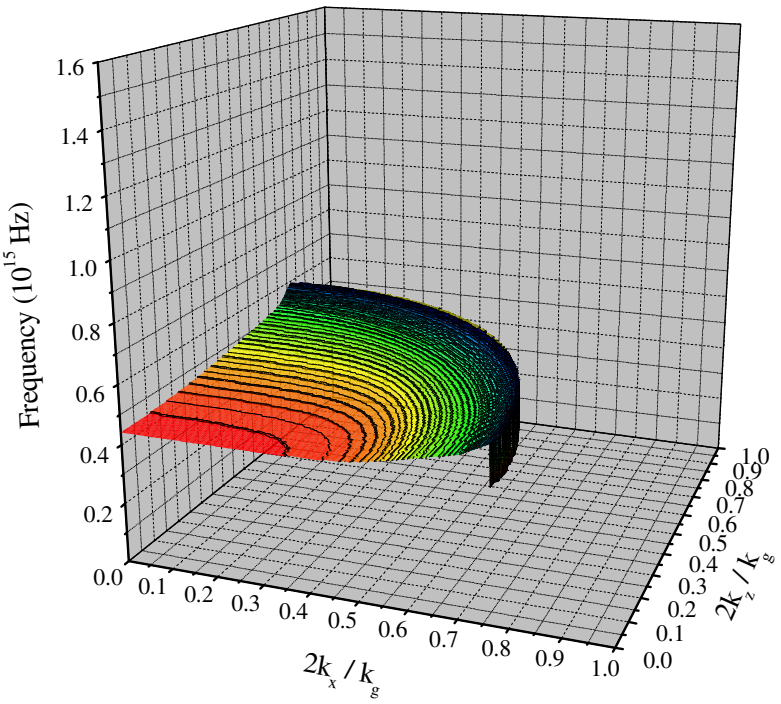

Figure 6.5.2 The frequency of maximum absorption due to the excitation of the second order SCSPP on a 200nm pitch, 300nm deep, 50nm wide Gaussian-grooved grating for varying k_x and k_z.

The fact that there is a relatively flat band producing absorption for all grating orientations could be useful for such application as selective absorbers. However, its use may be limited because the structure will only absorb particular polarisations at certain grating orientations. The use of a bi-grating structure could overcome this limitation and produce a structure which will absorb the same frequencies of the incident light for all polarisations at all azimuthal angles, and for a wide range of polar angles.

6.6 Summary

The form of the dispersion of a family of flat-banded self-coupled surface plasmon polaritons (SCSPPs) on very deep zero order gratings at $k_z = 0$ has been explored. Their band structure has been described by considering the deformation of the familiar shallow grating SPP dispersion curve caused by the large depth of the structure. Very large band gaps have been opened and, along with anti-crossing of the SCSPP

bands with the lightline, and anti-crossing between SPP bands away from the Brillouin zone boundaries, these have been shown to produce the dispersion curves obtained from the modelling. Also, the equivalent modes in the ω-k_z plane at $k_x = 0$, which are excitable with TE polarised radiation, have been shown to exist and are a product of the interaction of two SPPs scattered from $\pm k_g$ producing a SPP which is a standing wave in the x-direction (parallel to the grating vector) but which may propagate in the z-direction (parallel to the grating grooves). Finally, the dispersion of a single band of these modes has been calculated for all possible orientations of the grating with respect to the incident radiation, showing that a relatively flat banded mode can be excited for all azimuthal angles.

These flat bands may be of interest for use in selective absorbers or in resonant enhanced Raman scattering. They may also be of use for coupling radiation out of fluorescent species situated within the grating grooves.

Chapter 7

Surface Plasmon Polaritons on Narrow-Peaked Short-Pitch Metal Gratings

7.1 Introduction

In this chapter we shall extend our investigation into deep short-pitch gratings to consider structures consisting of a series of narrow Gaussian peaks. The band structures of the SPPs on these structures are found to be considerably different to those of the narrow Gaussian grooved structures considered in chapter 6.

In the first section we shall consider TM polarised light incident in the classical mount, where once again SPPs are found to be excited even in the zero-order region of the spectrum, and these may result in strong absorption of TM polarised radiation. For zero in-plane wavevector the SPP modes consist of a symmetric charge distribution on either side of the grating peaks, with a family of these modes existing corresponding to different numbers of field maxima per grating period. Because of the charge symmetry these modes may only be coupled to at polar angles away from normal incidence where strong resonant absorption may then occur. The dispersion of these SPP modes as a function of in-plane wavevector is found to be complex arising from the formation of very large band-gaps due to the harmonic content of the grating profile, the creation of pseudo high energy modes, and also through strong interactions between different SPP bands.

In the second section we will consider the case of light incident in the conical mount, and in particular when the azimuthal angle is 90° or 45°. When the azimuthal angle is 90° we find that the low energy bands produced by band gaps in the SPP dispersion curves may only be excited with TM polarised light, and describe how the coupling to the mode with this polarisation arises. Similarly the high energy bands may only be coupled to by TE polarised light. We also explain the dispersion of these modes in terms of anti-crossing between the SPP dispersion curves created by scattering from

the grating, and the SPP dispersion curve arising from the origin which is relatively unperturbed by the grating.

When the grating is oriented at a 45° azimuthal angle polarisation conversion can occur. In fact there is a mechanism which produces broad-band polarisation conversion on these structures (which is described in the third section of the chapter), and the consequence of exciting the SPPs is to either enhance or suppress the polarisation conversion depending upon whether the SPP occurs in a region of this broad-band polarisation conversion or not. This is described, as well as a brief description of the dispersion of the modes in this orientation.

For a review of previously published work on deep gratings see chapter 6, section 2.

7.2 SPPs on Narrow-Peaked Short-Pitch Gratings in the Classical Mount

In this section we shall consider the case of TM polarised light incident in the classical mount upon a grating consisting of a series of narrow Gaussian peaks of the form shown in figure 7.2.1

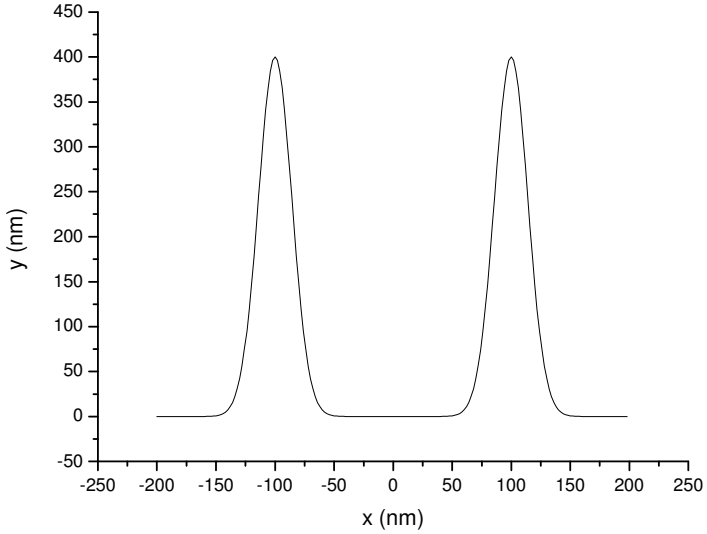

Figure 7.2.1 An example of the type of grating structure investigated. The profile is made up of a series of Gaussian peaks defined by the grating pitch, peak height, and peak width (FWHM).

In previous chapters we stated that the sign of any $2k_g$ component in the grating profile has no effect upon the band structure of the system, only upon the coupling strength to the modes. Therefore it would be expected that the band structure of Gaussian peaked gratings would be the same as those of the Gaussian grooved structures described in chapter 6. However, we also stated that this was only true for shallow gratings. In figure 7.2.2 we plot the TM reflectivity as a function of frequency and in-plane wavevector for a 200nm pitch grating consisting of a series of 400nm high, 40nm wide, Gaussian peaks in the classical mount.

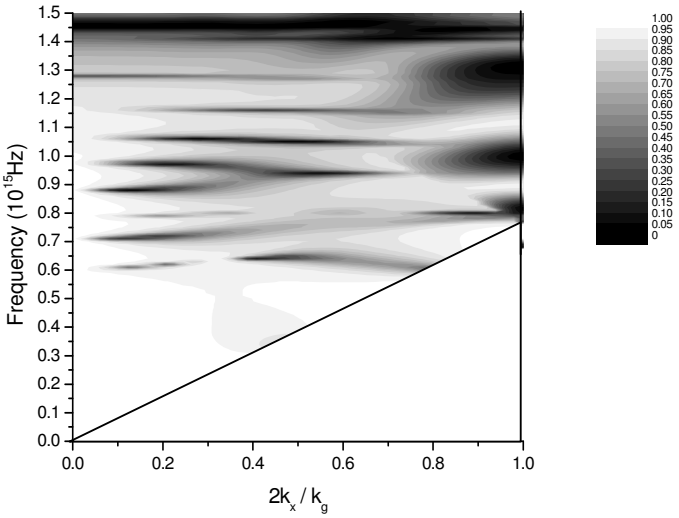

Figure 7.2.2 TM reflectivity as a function of frequency and in-plane wavevector for a 200nm pitch silver grating consisting of a series of 400nm high and 40nm wide (FWHM) Gaussian peaks with light incident at a 0° azimuthal angle.

It is clear from figure 7.2.2 that, when compared to the results obtained for gratings consisting of a series of Gaussian grooves in chapter 6, the band structure of the two systems are very different. Therefore we need to understand the way in which the band structures on the Gaussian peaked structures arise, and why this difference occurs. We shall also describe the nature of the SPPs on these structures by investigating the field profiles, an example of which is shown in figure 7.2.3. To our knowledge these resonances have not previously been observed or investigated.

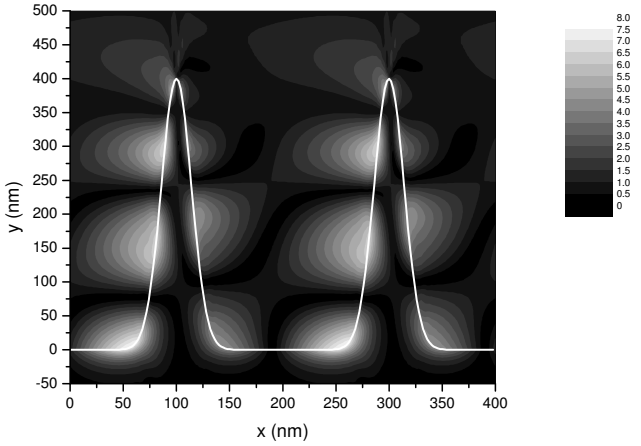

Figure 7.2.3 Time averaged $|\mathbf{H}_z|$ component of the fields of the SPP mode excited on a 200nm pitch silver grating consisting of a series of 400nm high and 40nm wide (FWHM) Gaussian peaks, with radiation of $f = 0.88 \times 10^{15}$ Hz incident at a 0° azimuthal angle and at $2k_x / k_g = 0.1$.

As described in chapter 2, when a sinusoidal grating is irradiated at normal incidence with EM radiation of the correct frequency and polarisation state to excite a SPP, what is actually excited are two SPPs, propagating in the +x and –x directions respectively. These interact to produce a standing wave which may have two possible field configurations relative to the grating profile, one with H-field maxima and minima on the sides of the grating peaks, and one with H-field maxima and minima at the peaks and troughs of the grating profile (for the 1st order SPP crossing point at normal incidence). Due to their different field distributions these two standing waves have different energies, and therefore a band gap is opened. The two solutions also correspond to symmetric and anti-symmetric charge distributions on either side of the grating peaks. At normal incidence it is impossible for incident radiation to couple to the symmetric charge distribution case since it would require the incident radiation E-fields to point in opposite directions on either side of the grating groove at the same instant in time. This, of course, is not possible for a normally incident plane wave. However, if the radiation is incident at some polar angle θ then the mode may be excited.

When the grating profile is more complicated than a pure sinusoid it is the details of the grating shape which determines which of these solutions is the high energy and which is the low energy. If we consider the 1st order branch at normal incidence

then it is the $2k_g$ component of the grating profile which determines the field distribution of the high and low energy branches on either side of the band gap.

If the phase between the k_g and $2k_g$ components is -90° the high energy branch consists of H-field maxima at the peaks and troughs of the grating since these are the positions at which the $2k_g$ component has its troughs (which correspond to an anti-symmetric charge distribution on either side of the grating grooves). This branch may then be coupled to at normal incidence, whereas the low energy branch may not since the H-field maxima of this mode occur at the shoulders of the grating grooves (which correspond to an anti-symmetric charge distribution). Therefore, the upper energy branch is coupled to at normal incidence, and not the lower energy branch. The Gaussian peaked profile investigated here is an example of this second case, although it also has higher harmonic components which will affect the higher order SPPs at normal incidence, and those at the Brillouin zone boundary produced by the periodicity of the grating.

For these other SPP branch crossing points, whether at the Brillouin zone boundary or at normal incidence, the formation of the band gaps is very similar. The component of the grating profile which couples together the two counter-propagating SPPs, and therefore the one that produces the band-gap, is always the sum of the two SPP scattering processes involved in creating the SPP at this point. For example, the 2nd order SPP branch crossing point at $2k_x = k_g$ arises through scattering from $-k_g$ and $+2k_g$, and therefore it is the $3k_g$ component of the grating profile which produces the band-gap.

Another point to note is that the number of field maxima (we shall consider the time averaged $|\mathbf{H}_z|$ component of the fields so that both maxima and minima in the instantaneous fields appear as maxima) per grating period for the SPP is determined by the scattering processes involved. For the 1st order SPP at the Brillouin zone boundary there is one field maximum per grating period. The 1st order SPP at normal incidence has two field maxima, since it arises from scattering by $\pm k_g$, and the second order SPP at the Brillouin zone boundary has three maxima since it arises from scattering by $-k_g$ and $+2k_g$ and so on. Therefore, SPPs at the Brillouin zone boundary have charge distributions which are not the same for each grating period, but which are instead the same for alternate peaks / grooves due to the fact that they have an odd number of field maxima per grating period.

For the narrow-peaked grating profiles investigated in this chapter the phases of the odd integer k_g components of the grating profile are always 0°. Therefore, the low energy branches of band-gaps arising at the Brillouin zone boundary have H-field maxima on the grating peak sides, and also at the midpoint between grating peaks.

For relatively shallow gratings ($d << \lambda_g$) the formation of band-gaps is well understood. However, when the grating height is increased so that it is of the order of the grating pitch the situation becomes more complex. The mode frequency, for $k_x = 0$, as a function of grating peak height for a 200nm pitch grating consisting of a series of 40nm wide Gaussian peaks at a 0° azimuthal angle is shown for normal incidence and the 1st Brillouin zone boundary in figure 7.2.4.

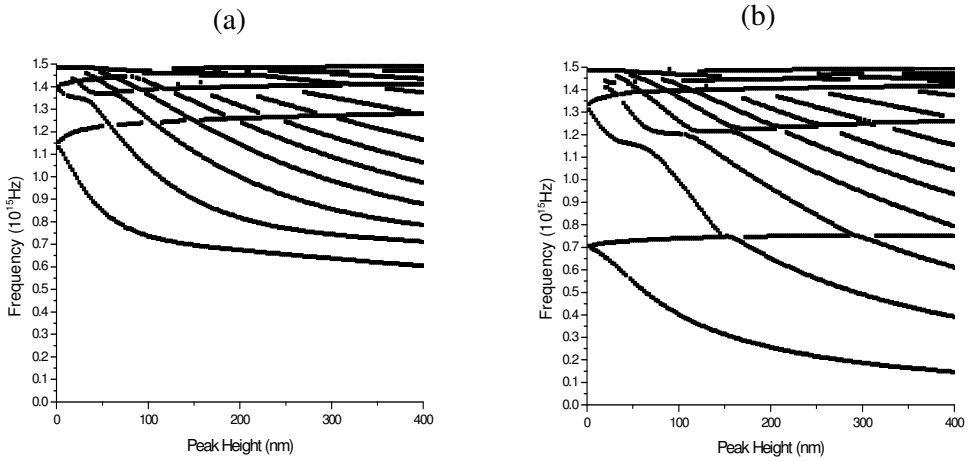

Figure 7.2.4 Mode frequency as a function of grating peak height for a 200nm pitch grating consisting of 40nm wide (FWHM) Gaussian peaks at a 0° azimuthal angle. a) $2k_x / k_g = 0$, and b) $2k_x / k_g = 1.0$.

In figure 7.2.4(a), for small peak heights, a band-gap corresponding to the first order SPP opens at $f \approx 1.14 \times 10^{15}$ Hz. The low energy branch rapidly reduces in energy with increasing peak height, whereas the high energy branch is observed to increase in energy relatively little. The reason for this asymmetric splitting is that the high energy branch may only rise in energy to asymptotically approach its associated diffraction edge, whereas the low energy branch has no such constraint.

As the peak heights increase, low energy branches of higher order SPPs are observed to reduce in energy through the high energy branch of the 1st order SPP. Also, the low energy branches appear to have a limit to the energy they may reduce to of $f \approx 0.6 \times 10^{15}$ Hz. However, this is not actually the case and this apparent limit arises at the frequency at which the lowest band is flat across the whole of the Brillouin zone. In

fact, this rapid reduction in the rate at which the low energy branches lower in frequency appears to occur at any point where a flat band has formed across the entire Brillouin zone.

This behaviour can best be observed in the plots of the mode frequency as a function of peak height at the BZ boundary (Figure 7.2.4(b)). In this plot there appears to be anti-crossing of two modes at a frequency of between 1.1 - 1.2 x 10^{15} Hz, although in fact there is no mode starting at f = 1.1 x 10^{15} Hz. However, this frequency is very close to that of the high energy mode at $k_x = 0$. The change in energy of this apparent anti-crossing as a function of peak height clearly follows that of the high energy mode at $k_x = 0$. This implies that when a low energy branch at either symmetry point occurs at the same frequency as a high energy branch at the other symmetry point (in other words, when the band is flat across the whole of the Brillouin zone) then the low energy branch follows the dispersion of the high energy mode as a function of peak height until the next higher-order low energy SPP branch anti-crosses with it.

Some understanding of this behaviour can be gained by investigating the field distributions for a high energy branch at $k_x = 0$, and for the low energy SPP at $2k_x = k_g$ which occurs at close to the same frequency. For a grating pitch of 200nm the first grating peak height at which this occurs is at d = 50nm, and the time averaged $|\mathbf{H}_z|$ component of the fields for the mode at near-normal incidence, together with that at the Brillouin zone boundary, are shown in figure 7.2.5.

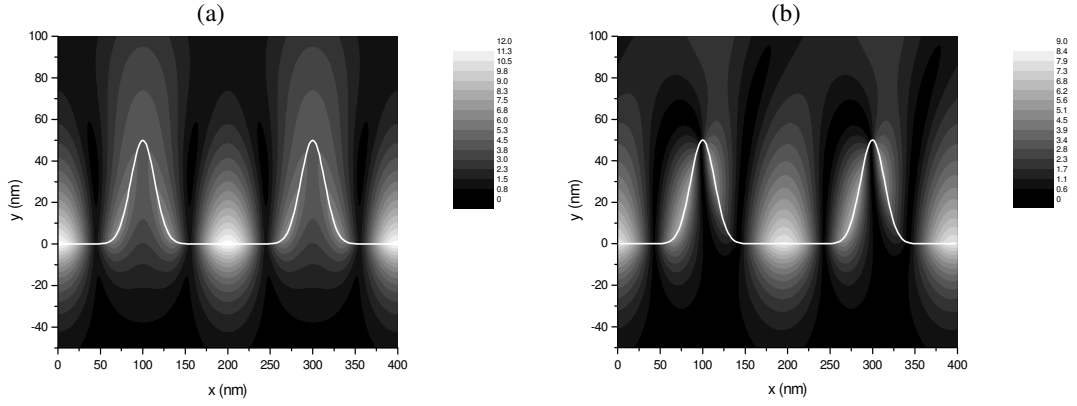

Figure 7.2.5 Time averaged $|\mathbf{H}_z|$ component of the fields of SPP modes excited on a 200nm pitch silver grating consisting of a series of 50nm high and 40nm wide (FWHM) Gaussian peaks at a 0° azimuthal angle. a) f = 1.22x10^{15} Hz and $2k_x / k_g = 0$, and b) f = 1.16x10^{15} Hz and $2k_x / k_g = 1.0$.

The obvious difference between these is that a maximum in the $|\mathbf{H}_z|$ distribution is evident on the grating peaks for the SPP at near-normal incidence, which is absent in the field distribution for the mode at the Brillouin zone boundary. However, in both cases it is clear that the strongest fields are those which occur at the grating troughs.

Since the dispersion in frequency as a function of grating peak height at d ≈ 50 nm is strongly affected by the 2nd order low energy mode this is far clearer for larger grating peak heights where the anti-crossing at the Brillouin zone boundary between this pseudo high energy mode and the higher order low energy modes is less pronounced.

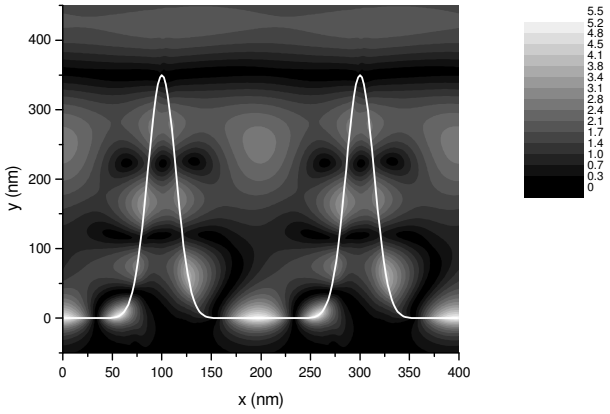

Figure 7.2.6 Time averaged $|\mathbf{H}_z|$ component of the fields of the SPP mode excited on a 200nm pitch silver grating consisting of a series of 350nm high and 40nm wide (FWHM) Gaussian peaks, with radiation of f = 1.26×10^{15} Hz incident at a 0° azimuthal angle and at $2k_x / k_g = 1.0$.

In figure 7.2.6 the time averaged $|\mathbf{H}_z|$ component of the fields are shown for d = 350nm for the pseudo high energy mode at the Brillouin zone boundary. Again, the strongest fields are found on the flat portion within the grating troughs, with the fields on the grating peaks being of a similar magnitude to the interference oscillation arising from the incident and reflected waves. In fact the spatial distribution of the strong fields in the grating with d = 350nm has the same periodicity as those with d = 50nm. Therefore, we believe that the pseudo high energy mode arises from the first order low energy mode, which then anti-crosses with the higher order low energy modes as they lower in frequency with increasing grating peak height. The additional field maxima on the peaks for the field distribution with d = 350nm arise due to the mixed character of the mode caused by the anti-crossing with the higher order low energy modes.

Similarly, the apparent low energy limit of the lowest energy SPP branch at normal incidence is also a frequency at which this occurs. In this case the 1st order low

energy SPP branch at $k_x = 0$ is almost flat with the 1st order high energy branch at the Brillouin zone boundary, and at this point the fields of the two modes are similar in the same way as described above. The apparent anti-crossing is far less pronounced in this case since the low energy modes are reducing in energy relatively slowly as they reach the frequency at which the anti-crossing takes place. Therefore it is not particularly clear for the grating height range investigated here, which is limited by the convergence limits of the modelling code. However, we believe that for larger grating peak heights these modes could be excited at lower frequencies.

From figure 7.2.4 the dispersion of these modes at normal incidence and at $2k_x / k_g = 1$ are well defined. However, these do not explain the complex dispersion behaviour shown in the reflectivity plot of figure 7.2.2. In order to gain some insight into how this complex dispersion with changing in-plane wavevector arises it is useful to investigate the TM reflectivities as a function of both frequency and in-plane wavevector for different heights of the grating peaks. These are shown in figure 7.2.7.

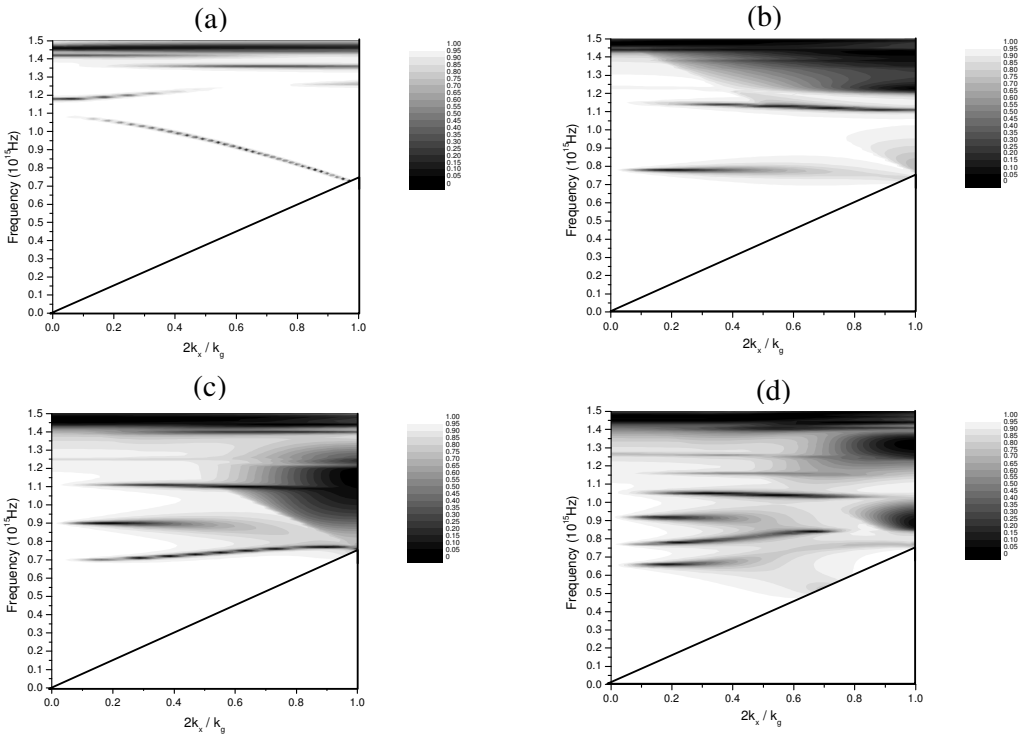

Figure 7.2.7 TM reflectivity as a function of frequency and in-plane wavevector for a 200nm pitch silver grating, consisting of a series of 40nm wide (FWHM) Gaussian peaks, and peak heights of a) 10nm, b) 75nm, c) 150nm, and d) 250nm.

Figure 7.2.7(a) shows the reflectivity from a 200nm pitch silver grating consisting of a series of 40nm wide, 10nm high, Gaussian peaks. This clearly shows the

band gap at $k_x = 0$ opened due to the large first harmonic component inherent in the Gaussian peaked profile.

In figure 7.2.7(b) the Gaussian peak heights have been increased to 75nm. The 1st order low energy branch at $k_x = 0$ has lowered in frequency sufficiently to produce a relatively flat band with the 1st order high energy branch at the Brillouin zone boundary. In addition, the 2nd order low energy branch at normal incidence has anti-crossed with the 1st order high energy branch and has produced a band at $f \approx 1.13 \times 10^{15}$ Hz.

It would be expected on such a grating that the 1st order SPP mode at normal incidence would produce a very broad and shallow reflectivity minimum since the radiative damping of a mode is known to increase with increasing depth / height of a grating. However, because these modes have fields which extend further into the metal than is the case with gratings consisting of narrow grooves (where the charge distribution is anti-symmetric on either side of the grating grooves), the resistive damping is also high. Since the resonance shape depends upon the ratio of the resistive and radiative damping components, with a reflectivity of zero when the two components are equal, this increase in the resistive damping compensates for the increase in the radiative damping. Therefore a deep reflectivity minimum is observed even on a high aspect ratio grating.

When the height of the peaks is increased to 150nm (Figure 7.2.7(c)) significant changes to the mode dispersions have occurred. The 1st order branch has lowered in energy very little since it is almost flat across the entire Brillouin zone. However, the end of the 2nd order SPP branch at the Brillouin zone boundary now coincides with that of the 1st order branch. Due to the fact that the modes may disperse in frequency as a function of grating height more rapidly at the Brillouin zone boundary than at normal incidence it is now closer in frequency to the 1st order SPP at normal incidence. Therefore the 1st order SPP has coupled more strongly to it leaving only weak coupling with the 2nd order SPP at $k_x = 0$.

In figure 7.2.7(d) the grating peak height has been increased to 250nm. The dispersion of the modes has developed from the 150nm case in that the lowest energy mode now weakly couples to a point on the lightline at $f \approx 0.52 \times 10^{15}$ Hz, as well as the 1st high energy mode at the Brillouin zone boundary. In figure 7.2.4(b) these non-radiative modes are evident below the light line, which occurs at $f = 0.75 \times 10^{15}$ Hz. The second order of these modes occurs at $f \approx 0.5 \times 10^{15}$ Hz and is an almost flat band across

to the lightline. This mode shows very weak coupling to radiation inside of the lightline since the branch of which it consisted before it become non-radiative remains coupled to the 1st order high energy mode at the Brillouin zone boundary.

The remaining band structure within the Brillouin zone can best be understood by considering which of the bands at normal incidence couple to which at the Brillouin zone boundary. This is simply achieved by comparing figure 7.2.7(d) with those of figure 7.2.4.

In this we define LE and HE to refer to the low and high energy modes respectively, followed by an integer which defines the order of the mode. From figure 7.2.4 and figure 7.2.7(d) it can be seen that LE1 at normal incidence links most strongly with HE1 at the Brillouin zone boundary, and LE2 links most strongly with LE3. Since LE4 at normal incidence is strongly coupled to LE4 at the Brillouin zone boundary after anti-crossing with the HE2 band, this leaves no branch at the Brillouin zone boundary for the LE3 branch at normal incidence to connect to. Therefore, it is observed to join only weakly with the LE3 band at the Brillouin zone boundary. For higher grating peak heights the number of SPP branches at normal incidence which have only weak coupling to the branches at the Brillouin zone boundary is increased, as is evident in figure 7.2.2. This is because the dispersion in frequency of the SPP modes at the Brillouin zone boundary as a function of grating peak height is more rapid, and therefore a higher number of SPP modes at the Brillouin zone boundary are at a lower frequency than the 1st order SPP at normal incidence.

(Note that, as the grating peak height is increased, one eventually reaches the point where the radius of curvature of the grating peaks is less than the electron mean free path or the skin depth, and therefore the use of a simple local dielectric function for the metal will become invalid. Therefore, the modelling for the gratings consisting of high peaks in this chapter may be flawed in these cases, however we believe that though this may alter the coupling to the modes it will not significantly change their dispersion).

Having described the SPP modes on these structures in the classical mount, in the next section we shall consider the case when the azimuthal angle is non-zero (the conical mount).

7.3 SPPs on Narrow-Peaked Short-Pitch Gratings in the Conical Mount

In the previous section we discussed the optical response of narrow-peaked short-pitch metal gratings in the classical mount. In this section we shall extend this to the case of non-zero azimuthal angles, and in particular where the azimuthal angle is either 90° or 45°. An SPP excited by TM polarised light on deep metal gratings oriented at a 90° azimuthal angle was first discovered by Watts, Preist and Sambles (1997), and this mode is evident in the results we shall present in this chapter. We shall discuss why this mode is excited on this structure and not on the Gaussian grooved structures discussed in chapter 6, and describe the dispersion of the mode with changing in-plane wavevector. We shall also discuss the effect of polarisation conversion on these structures when the azimuthal angle is 45°, although the case of polarisation conversion for normally incident light on these structures will be discussed in more detail in section 7.4.

The zeroth order reflectivity as a function of frequency and in-plane wavevector for a 200nm pitch silver grating consisting of 10nm deep and 40nm wide Gaussian peaks, and for an azimuthal angle of 90°, for both TM and TE polarised light is shown in figure 7.3.1.

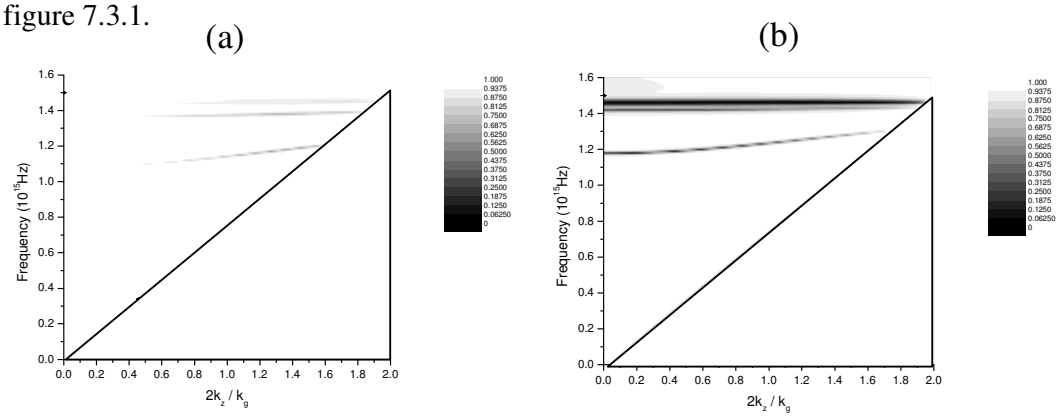

Figure 7.3.1 The zeroth order reflectivity from a 200nm pitch silver grating consisting of a series of 10nm deep 40nm wide Gaussian peaks, oriented at a 90° azimuthal angle, as a function of frequency and in-plane wavevector. a) TM polarised radiation, and b) TE polarised radiation.

The first point to note is that the frequencies at which the SPP bands occur at normal incidence (excepting the fact that there is no coupling in this orientation for the mode excited with TM polarised light) is the same as those which occur on the same structure for an azimuthal angle of 0°. Therefore, it appears that the modes are in fact

the same as those observed for that orientation. This is in fact the case, as will be shown later in the section.

In the classical mount there was no coupling to the low energy modes for normally incident TM polarised light due to the symmetry of the surface charge densities of the SPPs on the grating peaks. Since the surface charges are distributed symmetrically on either side of a grating peak for these low energy branches of the band gaps, and the fields of the incident light are always pointing in the same direction, the light may not couple to the mode. If the azimuthal angle is 90° then the surface charge densities always have this orientation (since $k_x = 0$), and for TE polarised light the fields will always point in the same direction. Therefore, these modes may not be coupled to for any polar angle of incidence. The high energy branches of the SPP band gaps have an anti-symmetric charge distribution on either side of a grating peak, and therefore these may be coupled to with TE polarised light. Since the high energy branch of the band gap disperses very little with increasing depth of the grating we will not discuss this mode any more in this chapter, except to mention that its dispersion with changing in-plane wavevector arises in the same way as that of the resonances observed on the narrow grooved structure at a 90° azimuthal angle described in chapter 6.

For the narrow grooved structures described in chapter 6 it was not possible to couple to the low energy SPP modes with TM polarised light when the azimuthal angle was 90°. This was due to the fact that, even though TM polarised light has a non-zero component of its electric field normal to the surface in this orientation, the charge distribution of the low energy modes required the incident light to have an E-field pointing in the same direction on either side of a grating peak *parallel to the grating vector* in order for coupling to occur. This could only be satisfied by TE polarised light. Conversely, for a grating consisting of a series of Gaussian peaks the charge distribution of the low energy modes is symmetric on either side of a grating peak, and in this case only TM polarised light may excite the modes since its E-field (for non-normal polar angles) has a normal component to the surface and points in the same direction *parallel to the grating grooves*, enabling it to couple to the symmetric charge distribution. This implies that whether TE or TM polarised light may couple to SPP modes when the azimuthal angle is 90° is defined by the phase of the $2k_g$ component of the grating profile with respect to the k_g component, since it is this which defines whether the modes consist of symmetric or anti-symmetric charge distributions.

An additional requirement for the coupling of TM polarised light to these modes is that the fields of the incident light must point in opposite directions at different points on the surface in the y-direction (normal to the average plane of the surface). This is because the surface charge distribution changes sign in the y-direction, and this requirement suggests that maximum coupling to the mode will occur when the y-component of the wavevector of the incident light is equal to the y-component of the SPP wavevector. Of course, this maximum coupling is likely to mean that the mode is over coupled and will not result in the lowest reflectivity minimum possible, but it is clear from this argument that the coupling to the mode will increase for higher values of k_z, and be very weak for angles near normal incidence. This can clearly be seen from figure 7.3.1(a) where the coupling to the first order SPP increases as k_z is increased.

If the depth of the grating is increased then the y-component of the SPP wavevector also increases, and therefore for deeper gratings the coupling to the SPP will be stronger for lower values of k_z. This is clearly seen if we calculate the same plots as in figure 7.3.1(a) but for increasing grating depths (figure 7.3.2).

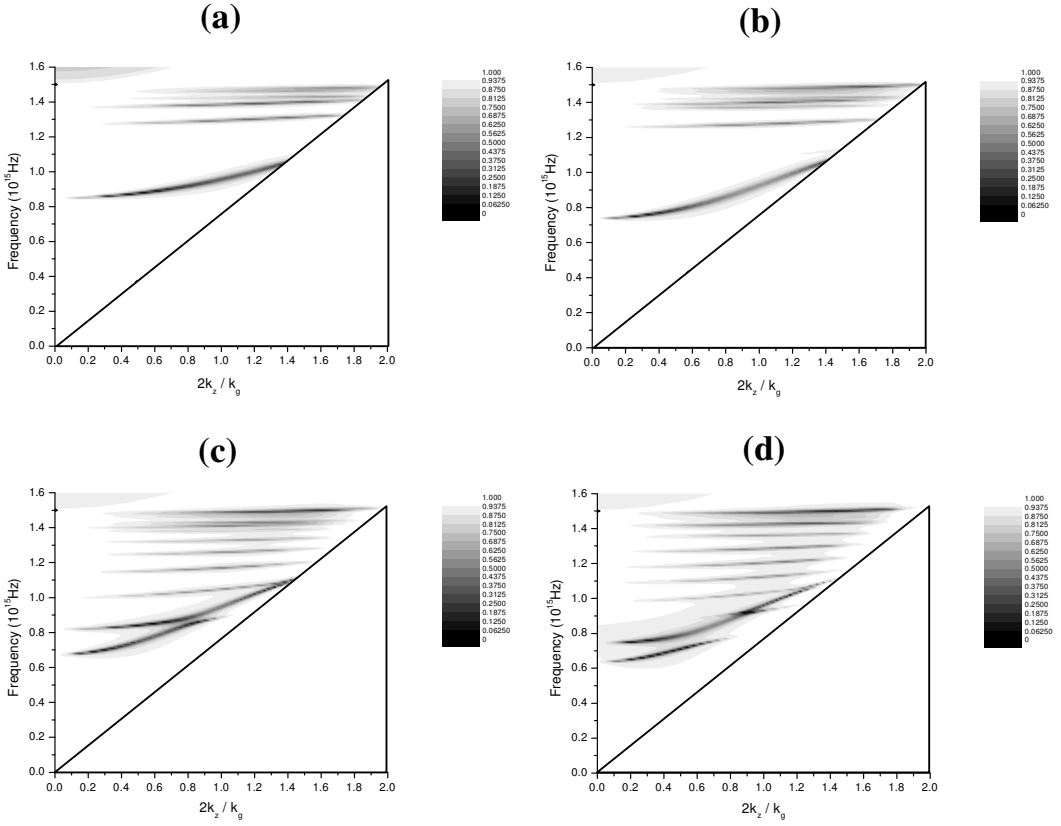

Figure 7.3.2 The TM reflectivity from 200nm pitch silver gratings consisting of a series 40nm wide Gaussian peaks, oriented at a 90° azimuthal angle, as a function of frequency and in-plane wavevector. a) d = 50nm, b) d = 100nm, c) d = 200nm, and d) d = 300nm

From figure 7.3.2 it is evident that the coupling to the lowest energy mode becomes stronger close to the $k_z = 0$ axis as the depth of the grating is increased. Also, the coupling to the higher order SPP modes (which have a larger y-component in their wavevector) couple less strongly close to the $k_z = 0$ axis. This is as expected from the arguments presented above.

If we investigate the field distributions of these modes we can show their nature. The $|\mathbf{H}_z|$ component of the fields (the component along the grooves) of the four lowest energy modes on the structure described for figure 7.3.2(c), and with $k_z = 0.4$ are shown in figure 7.3.3.

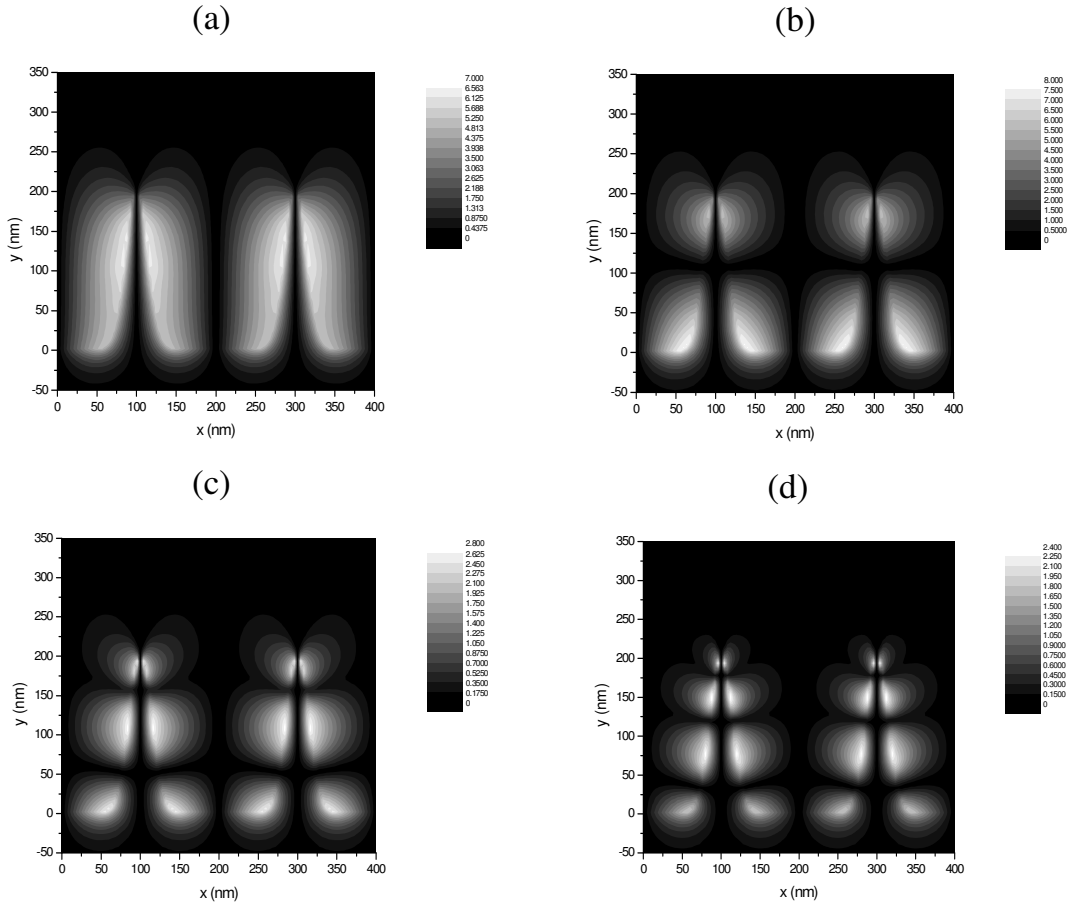

Figure 7.3.3 The $|\mathbf{H}_z|$ component of the fields (the component along the grooves) of the four lowest energy modes on the structure described for figure 7.3.2(c), and with $k_z = 0.4$. a) f = 0.721 x 10^{15}Hz, b) f = 0.833 x 10^{15} Hz, c) f = 1.008 x 10^{15}Hz, and d) f = 1.148 x 10^{15}Hz.

When the field distributions of figure 7.3.3 are compared to that shown in figure 7.2.3 for the SPP mode on a similar structure at an azimuthal angle of 0° it is clear that the modes are of the same origin. The main difference between the modes is that, even

though the SPPs at a 90° azimuthal angle are standing waves in the x-direction (similar to those at an azimuthal angle of 0°), they are also propagating in the z-direction (which the SPPs in the previous orientation are not). This can be shown by plotting the field distribution in the x-z plane through one of the SPP field maxima shown in figure 7.3.3. This has been performed for the SPP mode shown in figure 7.3.3(c) at y = 110nm (through the middle field maximum), and is shown in figure 7.3.4.

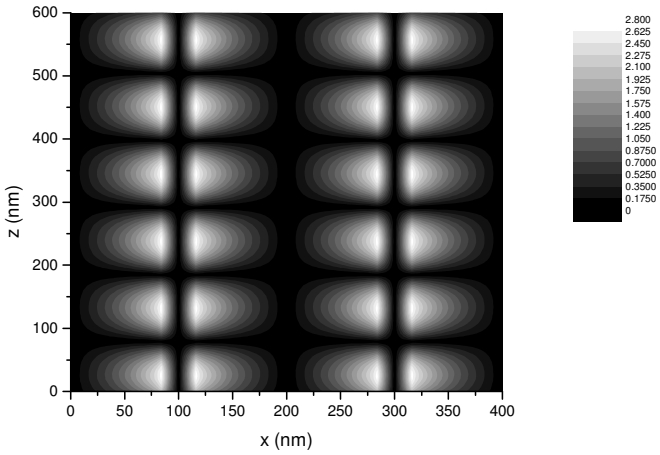

Figure 7.3.4 The $|\mathbf{H}_z|$ component of the fields in the x-z plane through the middle field maximum shown in figure 7.3.3(c) (y = 110nm).

It is clear, therefore, that the modes are produced by two SPPs propagating at angles to the grating grooves, with one propagating in the +x and +z directions, and one propagating in the –x and +z directions, so that the resulting mode is a standing wave in the x-direction, propagating in the z-direction. The periodicity of the fields in the z-direction shown in figure 7.3.4 is determined by the propagation angle of the SPPs.

The final features to note in the plots of figure 7.3.2 are the different dispersion with changing in-plane wavevector of the lowest energy SPP mode when compared to the higher order modes, and also the anti-crossing between this lowest energy SPP mode and the higher order SPP modes. The explanations for both of these are due to the same mechanism.

The charge distributions of the low energy branches of the SPP dispersion curves for these Gaussian peaked gratings have symmetrical charge distributions on either side of the peaks, and therefore have their maximum charge densities at the peaks and troughs of the grating profile. Since the flat regions between the peaks contain a maximum surface charge density, a maximum electric field intensity is also found in

these regions. Also the planar SPP dispersion curve (only slightly perturbed by the grating) which resides outside of the light line will have field maxima at these same regions. Therefore, anti-crossing may occur between the two bands, and this can be seen if, instead of plotting the reflectivities as we have so far in this chapter, we plot the band structure instead. We have done this for a 50nm deep grating, which resulted in the reflectivity plot of figure 7.3.2(a), and is shown in figure 7.3.5.

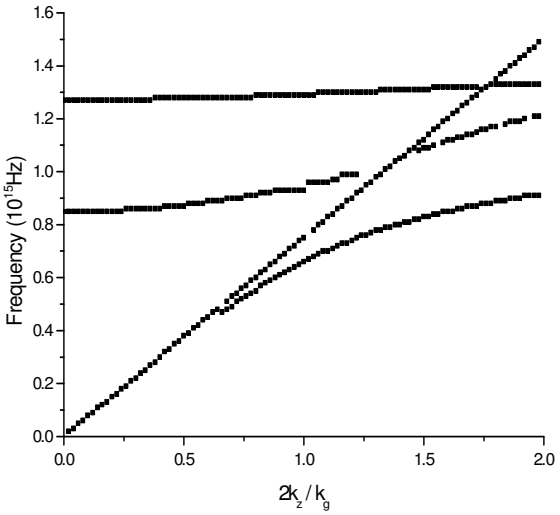

Figure 7.3.5 The band structure for a 200nm pitch silver grating consisting of a series of 40nm wide 50nm deep Gaussian peaks. The first and second order SPPs are shown as well as the light line and the SPP dispersion curve centred at the origin.

In figure 7.3.5 there are four lines which correspond to the light line (the diagonal line centred at the origin), the planar surface SPP dispersion curve (also centred at the origin), and two lines which are the first and second order SPPs (which originate at $f \approx 0.85 \times 10^{15}$Hz and $f \approx 1.28 \times 10^{15}$Hz). The peculiarity of this plot is that the planar surface SPP should approach $f \approx 1.5 \times 10^{15}$Hz (since this is the equivalent of ω_{sp} in this plot), yet it is clearly approaching a limit at a lower frequency than this. What is in fact happening is that anti-crossing is occurring between the first order SPP and the planar SPP curve, so that the planar SPP curve at high values of k_z is approaching the dispersion expected of the first order SPP, and the first order SPP is approaching that of the planar surface SPP. This is the cause of the increased gradient in the first order SPP when compared to the other SPP curves shown in figures 7.3.2(a) and (b). Similarly, when the higher order SPPs reduce in frequency such that they cross with the lowest frequency SPP anti-crossing may also take place, since all of these modes have strong

fields on the flat regions between the peaks. This is clearly seen in figures 7.3.2(c) and (d).

Having considered the dispersion and optical response of the Gaussian peaked gratings at a 90° azimuthal angle we shall now consider the case of the grating oriented at a 45° azimuthal angle. In this orientation polarisation conversion can occur, and we will primarily consider the effect of the SPPs on the polarisation conversion. Since we are considering the case of an azimuthal angle of 45° both TM and TE polarised light can couple to the SPP modes, and so the results of considering TM polarised incident light are much the same as considering that of TE polarised incident light. Therefore, it does not matter which we choose to investigate; we have chosen TM polarised light.

The polarisation conserved and converted TM reflectivities for 200nm pitch silver gratings consisting of a series of 40nm wide Gaussian grooves oriented at a 45° azimuthal angle as a function of frequency and in-plane wavevector for various depths are shown in figure 7.3.6.

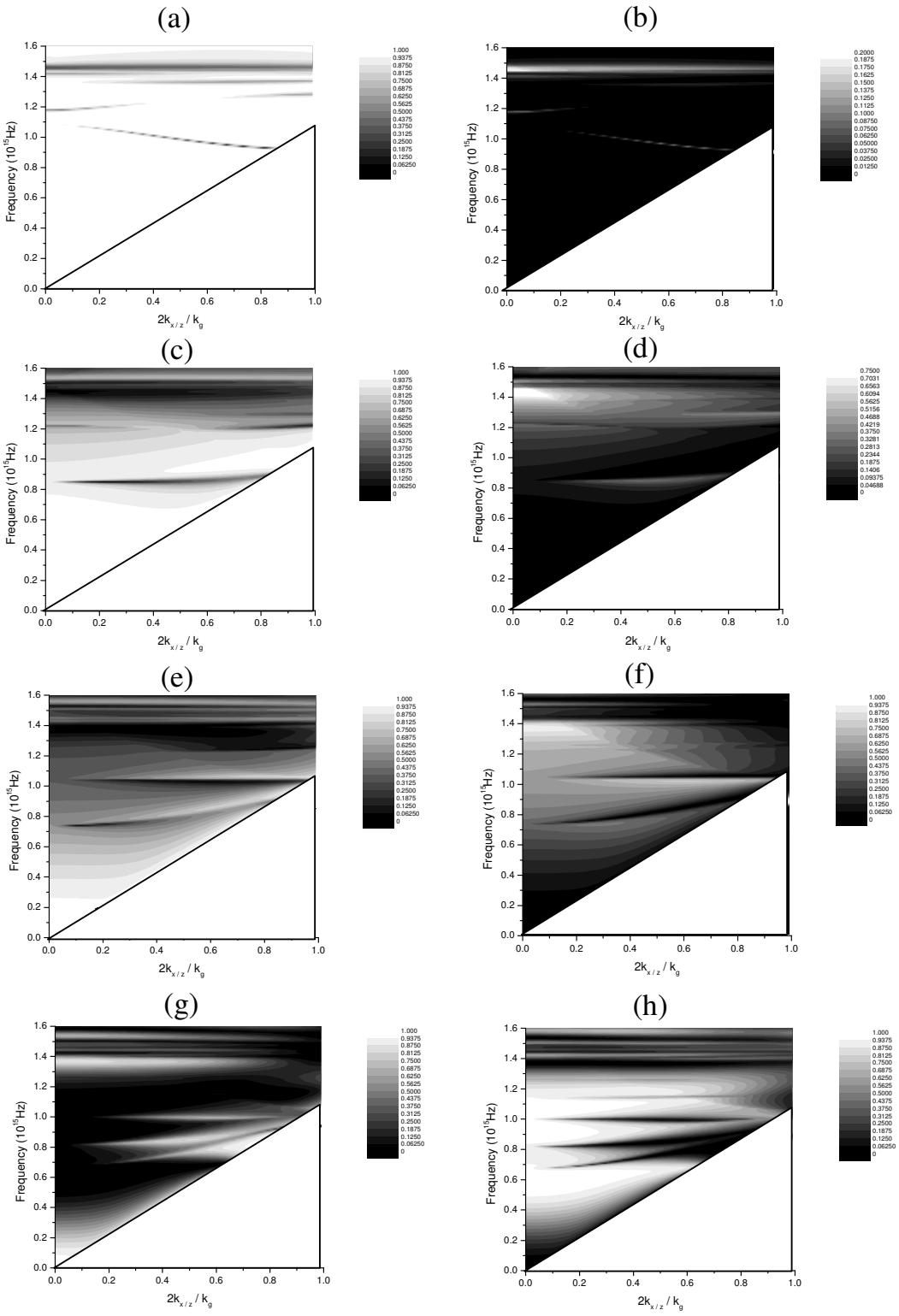

Figure 7.3.6 Zeroth order TM reflectivities for 200nm pitch silver gratings consisting of a series of 40nm wide Gaussian grooves oriented at a 45° azimuthal angle as a function of frequency and in-plane wavevector. a) 10nm deep polarisation conserved, b) 10nm deep polarisation converted, c) 50nm deep polarisation conserved, d) 50nm deep polarisation converted, e) 100nm deep polarisation conserved, f) 100nm deep polarisation converted, g) 200nm deep polarisation conserved, h) 200nm deep polarisation converted.

The most obvious point to note in figure 7.3.6 is the broad band polarisation conversion which occurs when the depth is large. This phenomenon will be considered in section 7.4 for normal incidence, and therefore will not be discussed here. We will concentrate on the SPP bands which are evident as relatively sharp features for non-normal angles of incidence.

The dispersion of the SPP bands in this orientation are a result of the evolution of the SPP dispersion curves for an azimuthal angle of 0° (described in section 7.2), to those of the SPP dispersion curves for an azimuthal angle of 90° (described earlier in this section). The bands exist for all possible grating orientations, and vary monotonically between these two limiting azimuthal angles, with the dispersion in the orientation considered here appearing slightly more like those of the dispersion when the azimuthal angle is 90°, rather than 0°.

The more interesting feature is the effect of these SPPs on the polarisation conversion from the structure. For shallow gratings the polarisation conversion is enhanced by the SPP, whereas for deeper gratings it is suppressed. In fact the minima in the polarisation converted reflectivities appear when the SPPs arise in a region of broad band polarisation conversion. To understand this we need to consider in more detail the process by which polarisation conversion occurs.

We have previously stated that polarisation conversion occurs due to the fact that both TM and TE polarised light can result from scattering of the SPP in to the specularly reflected order. Since, at a 45° azimuthal angle, the coupling into, and out of, the SPP mode is equal for both TM and TE polarised light it may be somewhat surprising that the polarisation converted reflectivity is often higher than 50%, and may in fact approach nearly 100%.

In order to understand this we will begin by considering the case of light incident upon a shallow grating oriented at a 0° azimuthal angle. When a SPP is excited a reflectivity minimum is produced, but we must consider where this lost energy has been dissipated. The excited SPP scatters out of the grating and into the specularly reflected order, whereupon it interferes destructively with the light which has been directly reflected from the surface. This produces the reflectivity minimum, but the energy has to have been dissipated somewhere. The only other energy loss channel available (if the structure is non-diffractive) is the absorption in the metal, and in this situation this is where the energy has been lost.

In the case of polarisation conversion when the azimuthal angle is 45° both TM and TE polarised light has been scattered into the specularly reflected order from the SPP. If we are considering TM polarised incident light then the TM polarised light scattered out of the SPP will destructively interfere with the directly reflected light as before. In this situation, however, there is an extra energy loss channel of the polarisation converted scattered light. Whether this energy is lost into absorption or the polarisation converted order depends upon the probability of scattering of the SPP from the surface. This is increased for deeper gratings, meaning that the majority of the energy will be distributed into the polarisation converted order, and therefore the polarisation conversion from the structure can approach 100%.

In order to explain the suppression of the polarisation conversion due to the SPP in the plots of figure 7.3.6 we must again consider the energy distributions between the orders. For a deep grating there is a broad band polarisation conversion region in which the SPPs are excited. If an SPP scatters into the specularly reflected order then destructive interference will occur in the polarisation converted order rather than the polarisation conserved order, due to the fact that the vast majority of the light in the specular reflected orders not due to the SPPs is already polarisation converted. The energy from the destructive interference has only two energy loss channels; absorption in the metal, and redistribution into the polarisation conserved order. The same arguments apply as before, meaning that maxima are observed in the polarisation conserved order which may approach 100%.

7.4 Broad Band Polarisation Conversion from Gratings at Normal Incidence

In Chapter 2 polarisation conversion from gratings at azimuthal angles which are not 0° or 90° was shown to originate via two different mechanisms (with a maximum in the conversion occurring at a 45° azimuthal angle in both cases). Firstly, it may be produced when a surface plasmon polariton (SPP) is excited, and secondly there is a dependence of the polarisation conversion on the depth of the grating caused by field loops being created within the grating grooves. The dispersion of the polarisation conversion produced via the excitation of SPPs follows that of the SPP dispersion

curves, with the frequency at which it occurs for normally incident light being predominantly determined by the grating pitch, and the dielectric function of the media. The dispersion of the polarisation conversion maximum which arises due to the creation of field loops as a function of the grating depth is given by,

$$f = \frac{c(2n-1)}{4d\cos\theta}$$

where n is an integer corresponding to the number of field loops contained within the grating grooves, and θ is the polar angle of incidence. Since we are interested only in normal incidence here $\theta = 0$, and $f \propto 1/d$.

These two mechanisms produce relatively narrow polarisation conversion bands, which have limited use for applications. In this section we explore a novel structure for extending the polarisation conversion band over a broader range, and in particular over the whole of the visible spectrum.

The only SPP modes which will be excited in the orientation of interest are the high energy branches of the SPP dispersion curve, since coupling to the low energy branches is prohibited due to the phase of the $2k_g$ component of the grating profile with respect to the k_g component. Due to the opening of the band gap the excitation frequency of these modes as a function of depth shows an increase from its original frequency at small d (determined by the pitch of the grating and the permittivities of the metal) to asymptotically approach its associated diffraction edge as the size of the band gap increases.

As previously stated both the polarisation conversion due to the excitation of the SPP, and that due to the field loops, give relatively narrow polarisation conversion resonances as a function of frequency. However, these two modes are able to interact since the overlap integral of their fields is non zero. It is this interaction, which produces anti-crossing and a resultant mode of mixed character, which enables the width of the polarisation conversion resonance as a function of frequency to be increased, producing a broad polarisation conversion band.

Figure 7.4.1 shows the polarisation conserved (a) and converted (b) reflectivities as a function of the frequency of the incident light and the height of the peaks for a 200nm pitch grating consisting of a series of 60nm wide Gaussian peaks (at full width half maximum), and oriented at an azimuthal angle of 45° for normally incident TM

radiation. The metal grating is modelled as silver using a Drude model with ω_p = 1.32 x $10^{16}s^{-1}$ and τ = 1.4x10^{-14}s. The very short pitch of the grating is used here to clearly demonstrate the origin of the broad polarisation conversion bands as there are fewer SPP bands and diffraction edges in the region of interest. Also, it is important that the grating be non-diffractive in the optical region of the spectrum at normal incidence requiring that the pitch be less than 400nm.

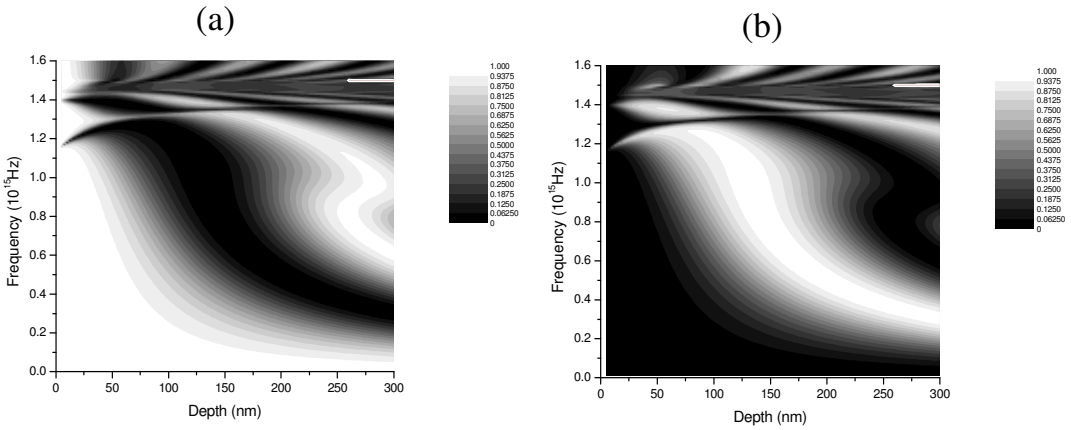

Figure 7.4.1 Reflectivity from a 200nm pitch silver grating consisting of a series of 60nm wide Gaussian peaks at a 45° azimuthal angle as a function of frequency and peak height. a) Polarisation conserved, and b) polarisation converted.

At small d a band originating at f = 1.18 x 10^{15}Hz is observed. This is the high energy branch of the first order SPP which, with increasing depth of the grating, would be expected to increase in frequency to asymptotically approach its corresponding diffraction edge at f = 1.5 x 10^{15}Hz. However, there is clear anti-crossing occurring with the mode dispersing with frequency as 1/d at a depth of 50-100nm.

If the width of the Gaussian peaks is narrowed to 40nm (figure 7.4.2) the interaction strength between the modes has altered, reducing the width of the gap produced by the anti-crossing. This change in the interaction strength has also produced back bending of the mixed character mode, and the polarisation conversion as a function of frequency for a depth of 180nm shows strong polarisation conversion over a broad frequency band from approximately 0.5 x 10^{15}Hz to 1.2 x 10^{15}Hz (a wavelength range of ~350nm).

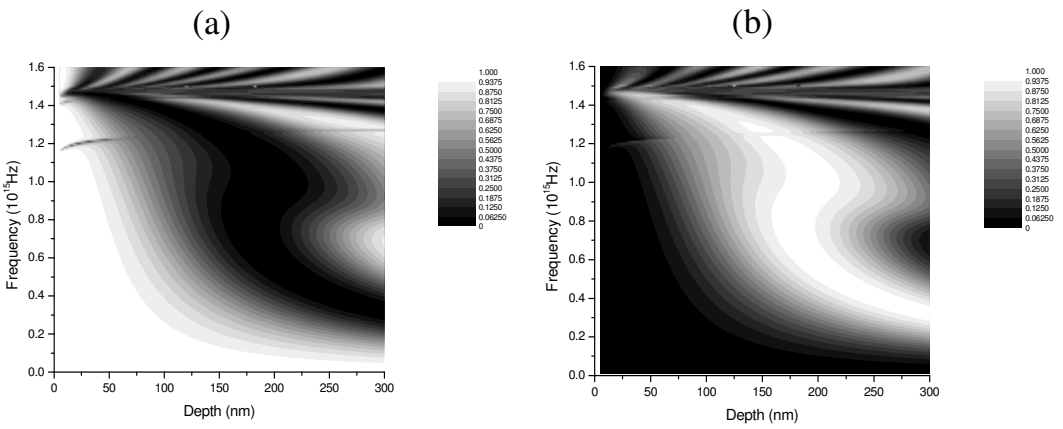

Figure 7.4.2 Reflectivity from a 200nm pitch silver grating consisting of a series of 40nm wide Gaussian peaks at a 45° azimuthal angle as a function of frequency and peak height. a) Polarisation conserved, and b) polarisation converted.

It is instructive at this point to investigate the fields of this mixed mode. Figure 7.4.3 shows the x and z components of the electric field for 3 different points on figure 7.4.2(a). Figure 7.4.3(a) shows the fields for a 100nm deep grating at a frequency of 1.37 x 10^{15}Hz. The x component shows strong SPP character in the enhanced fields localised at the interface, there is also a localised 'loop' of field in the z-component which is also due to the excitation of the SPP. It is noticeable that above the grating surface the two components are 180° out of phase with each other and it is this phase difference which is the origin of the polarisation conversion.

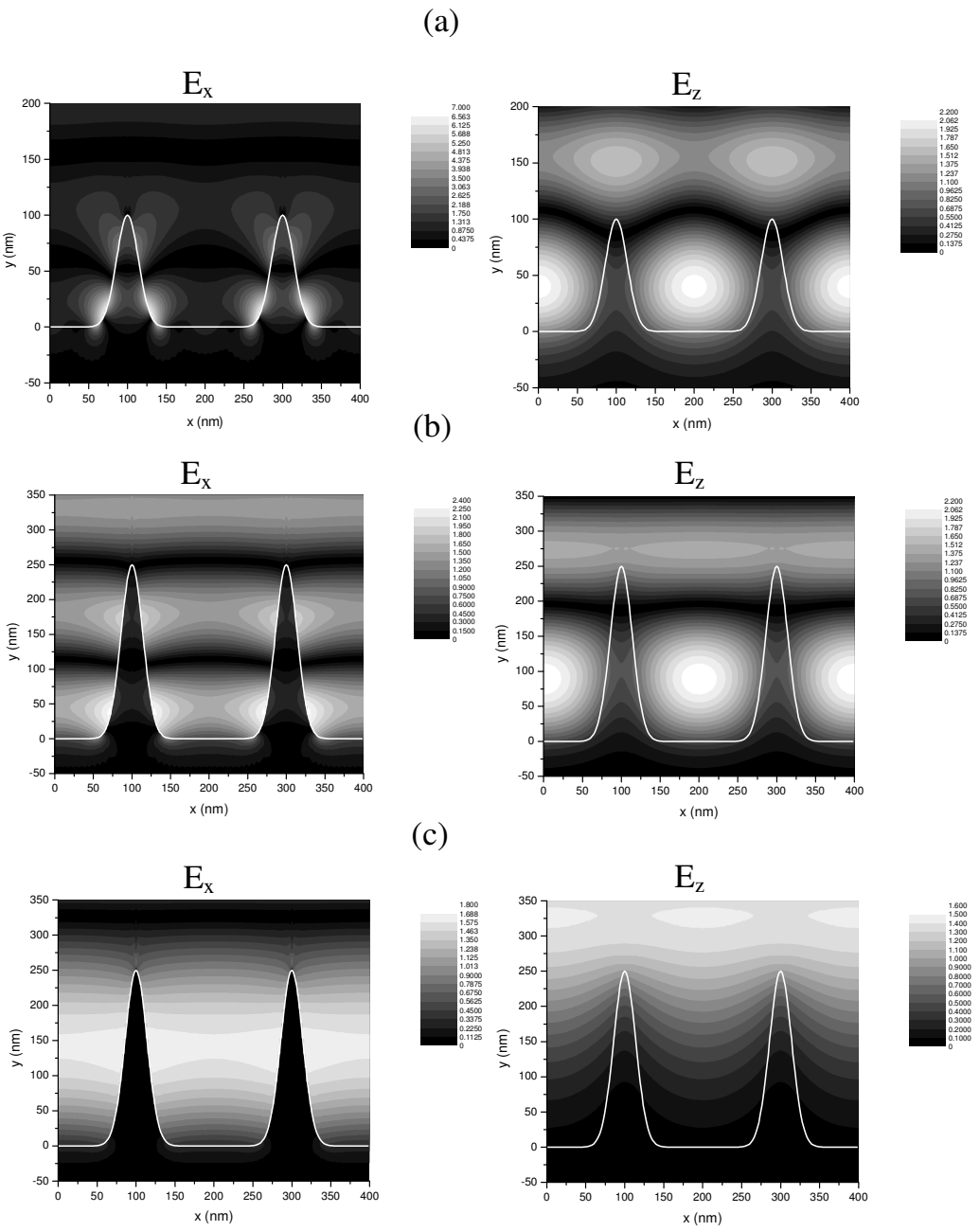

Figure 7.4.3 x and z components of the E fields at three points on the plot in figure 7.4.2(a). a) d = 100nm, f = 1.37 x 10^{15}Hz, b) d = 250nm, f = 0.97 x 10^{15}Hz, and c) d = 250nm, f = 0.41 x 10^{15}Hz.

For the point at d = 250nm and f = 0.97 x 10^{15}Hz (figure 7.4.3(b)) similar behaviour is observed, but the SPP character in the x component has reduced, as would be expected in anti-crossing behaviour away from the crossing point and close to the original dispersion of the second interacting mode. In figure 7.4.3(c) the behaviour has changed somewhat in that the SPP character in the x component has almost totally disappeared leaving just the interference between the incident light and that reflected

from the bottom of the grove. However, there is still a 180° phase difference between the two components producing the polarisation conversion.

The broad polarisation conversion band for this grating at 180nm deep only extends down to a frequency of 0.5 x 10^{15}Hz (600nm wavelength) which is still well into the visible region of the spectrum. Unfortunately, there is a limit to the low frequency extent of this band for a given depth given by the 1/d relation. However, by changing the pitch of the grating the frequency origin of the SPP mode is reduced and therefore the broad-band reduces in energy. Also, the interaction strength between the two modes is such that the broad-band is shifted to a higher depth where the 1/d relation means that the lower limit for the polarisation conversion band is lowered. For a 300nm pitch grating with 60nm wide Gaussian peaks the back bending of the resultant mixed mode is more pronounced than in the 200nm pitch case (Figure 7.4.4), but at a depth of 265nm the polarisation conversion is still >90% across a broad frequency band. This is shown in figure 7.4.4(c), which demonstrates the broad band extending across the whole of the visible region of the spectrum (from 400nm to 800nm wavelength).

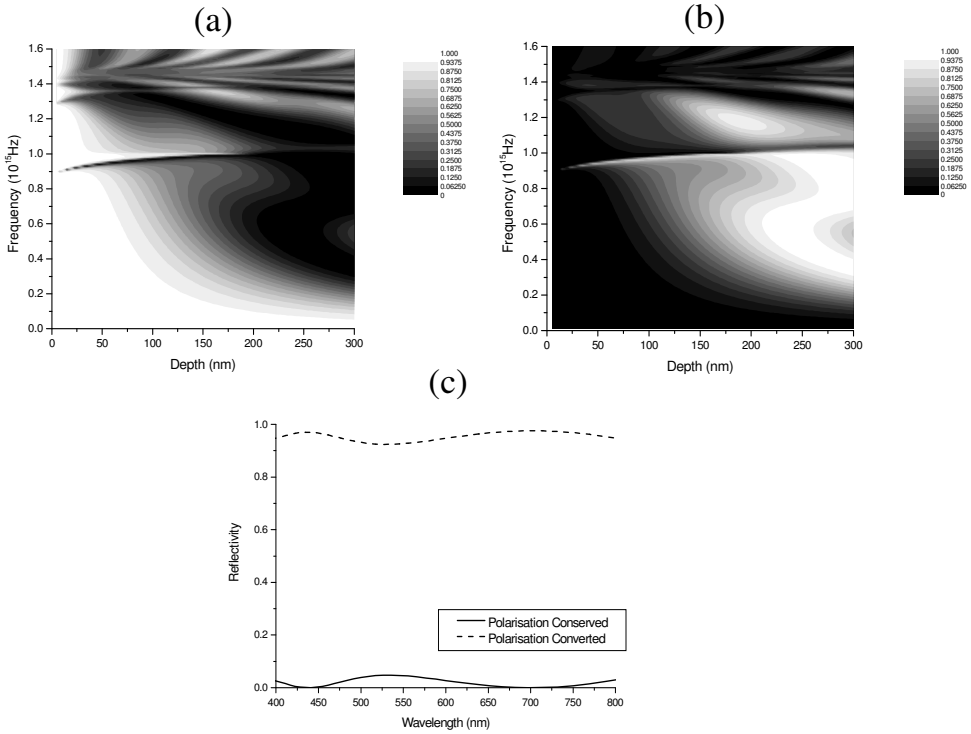

Figure 7.4.4 Reflectivity from a 300nm pitch silver grating consisting of a series of 60nm wide Gaussian peaks at a 45° azimuthal angle. a) Polarisation conserved reflectivity as a function of frequency and peak height, b) polarisation converted reflectivity as a function of frequency and peak height, and c) reflectivity as a function of wavelength for d=265nm.

In reality, the Drude model does not describe the permittivity of a metal particularly well, especially in regions of absorption. Therefore we have also modelled the polarisation conserved and converted reflectivities from similar structures to those described above, but with the permittivity of the silver described using a polynomial, which was fitted to experimentally derived figures. The results of these calculations are shown in Figure 7.4.5 for a 250nm pitch grating with 50nm wide Gaussian peaks, which were determined to be the optimum grating parameters in this case. The same broadband polarisation conversion is clearly visible, however less of the incident light is converted when compared to the grating described by the Drude model, especially at the high frequency region of the spectrum where there is increased absorption. This is clearly due to the more realistic values of the permittivity. However, the polarisation conversion from the grating remains >70% over the whole of the visible region of the spectrum.

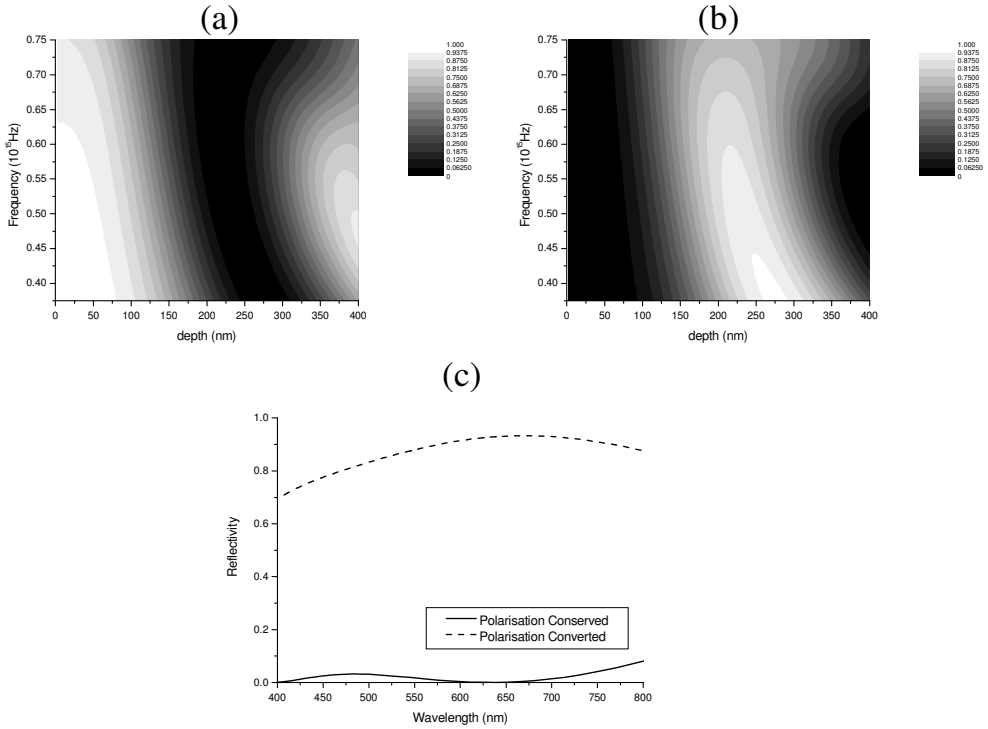

Figure 7.4.5 Reflectivity from a 250nm pitch silver grating consisting of a series of 50nm wide Gaussian peaks at a 45° azimuthal angle. Silver permittivity described by a polynomial fitted to experimentally determined values. a) Polarisation conserved reflectivity as a function of frequency and peak height, b) polarisation converted reflectivity as a function of frequency and peak height, and c) reflectivity as a function of wavelength for d=240nm.

Since there is absorption in the blue region of the spectrum for a silver grating we have also modelled the polarisation conversion from a similar aluminium structure

where the absorption at these wavelengths is less. The permittivity of the aluminium was described by a polynomial function fitted to experimentally derived values as in the silver case. The structure we obtained which produces the maximum polarisation conversion over a large wavelength range consists of a series of 232.5nm high, 60nm wide Gaussian grooves, with a grating pitch of 300nm. This is shown in Figure 7.4.6. For this structure there is a greater efficiency of polarisation conversion for lower wavelengths than in the silver case, however this increase in efficiency at the blue end of the spectrum is at the cost of a lowering of efficiency at the extreme red end of the spectrum.

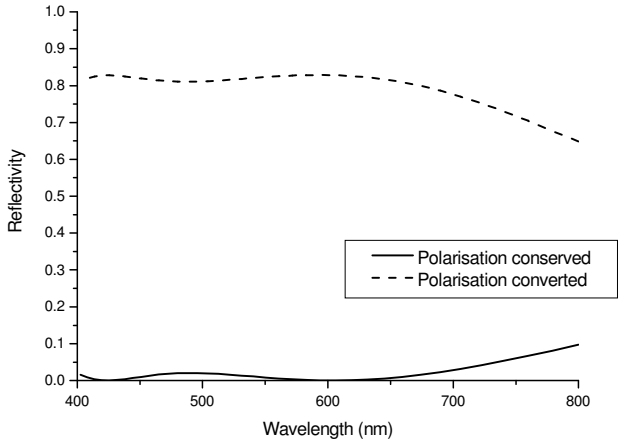

Figure 7.4.6 Reflectivity from a 300nm pitch aluminium grating consisting of a series of 232.5nm high, 60nm wide Gaussian peaks at a 45 degree azimuthal angle of 45°.

One particular application for this broad band polarisation conversion could be its use in increasing the amount of light extracted from back lit liquid crystal displays (LCDs), thereby increasing battery life times. In general, the amount of light extracted from a LCD is limited by the polariser on the outside of the structure, and this causes 50% of the light from the back-lighting to be lost. By placing a polarisation converting mirror behind the light source and using a metal grid polariser (to reflect the light of the incorrect polarisation required for transmission rather than absorb it) a far larger percentage of the light could be extracted. This mirror would need to polarisation convert over the entire visible region of the spectrum, and the broad band polarisation converting 'mirror' described in this section could be used for this.

7.5 Summary

In this chapter we have extended our investigation into deep short-pitch gratings to consider structures consisting of a series of narrow Gaussian peaks. In the first section we considered TM polarised light incident in the classical mount, where SPPs were found to be excited even in the zero-order region of the spectrum. For zero in-plane wavevector these SPP modes consisted of a symmetric charge distribution on either side of the grating peaks, with a family of these modes existing corresponding to different numbers of field maxima per grating period. The dispersion of these SPP modes as a function of in-plane wavevector was found to be relatively complex arising from the formation of very large band-gaps, and also through strong interactions between different SPP bands.

In the second section we considered the case of light incident in the conical mount, and in particular when the azimuthal angle was 90° or 45°. When the azimuthal angle was 90° we found that the low energy bands produced by band gaps in the SPP dispersion curves could only be excited with TM polarised light. Similarly the high energy bands could only be coupled to by TE polarised light. We also explained the dispersion of these modes in terms of anti-crossing between the SPP dispersion curves created by scattering from the grating, and the SPP dispersion curve arising from the origin which is relatively unperturbed by the grating.

When the grating was oriented at a 45° azimuthal angle polarisation conversion was shown to occur. We have shown (in the third section of this chapter) that there is a mechanism which produces broad band polarisation conversion on these structures, and the result of the SPPs is to either enhance or suppress the polarisation conversion depending upon whether the SPP occurs in a region of this broad band polarisation conversion or not.

Chapter 8

Experimental Observation of SPPs on Deep Zero-Order Silver Gratings for the Visible Region of the Spectrum

8.1 Introduction

In the previous two chapters the SPP modes which may be excited on deep zero-order metal gratings have been described. In this chapter a method for manufacturing such gratings is presented, and it is shown that resonant absorption of light may occur, which is believed to be due to the excitation of self-coupled SPPs. To our knowledge SPPs have never before been observed in the zero-order region of a grating.

Due to problems with the manufacturing process, which are inherent to the method, the surface of the grating structure is relatively rough, which causes diffuse scatter, and therefore a reduced specularly reflected order intensity. For this reason it has not been possible to fit the data obtained to theory, and therefore it is not possible to categorically attribute these reflectivity minima to the excitation of SPPs. However, the dispersion of these modes with changing in-plane wavevector is found to be very similar to those obtained for the SCSPPs in chapter 6, and therefore we believe that it is these modes which are being excited. The main difference in the optical response obtained from the gratings studied when compared to those investigated in chapter 6 is that broad band polarisation conversion occurs when the grating is oriented at a 45° azimuthal angle. This is more reminiscent of the structures investigated in chapter 7, and the reason that broad-band polarisation conversion occurs on these structures, even though it is the SCSPPs which appear to be excited, is explained.

8.2 Methodology

In chapters 6 and 7 the optical response of deep zero-order (for the visible region of the spectrum) grating structures has been described. In this section the method by which such gratings have been manufactured is described, and also the way in which the

optical response of these structures have been measured. There are three main stages in the fabrication process; grating fabrication in a photoresist layer, transfer of the grating into a SiO_2 substrate by reactive ion etching, and vacuum deposition of silver onto the SiO_2 grating. Each of these will be described in some detail, with the method used to obtain the optical response also described.

8.2.1 Sample Fabrication

The substrates on which the structures were manufactured were made of SiO_2, with the surface on which the structure was to be fabricated polished to $\lambda / 10$, and the reverse surface inspection-polished (meaning that there are no gross defects). The substrates were 6mm thick and 25mm square. They were cleaned thoroughly by immersing them in concentrated nitric acid for 1hr, before being cleaned with acetone and isopropanol using cotton buds. At this stage the inspection-polished side of the substrate was painted black, for reasons which will be explained later in this section. As a final cleaning stage the SiO_2 surface was 'drag cleaned' using lens tissue soaked in acetone in order to remove any final steaks or dust on the surface.

Immediately following this cleaning procedure the substrate was taken into a clean room for spin coating with photoresist. The photoresist used was Shipley SPR700, which is a broad-band photoresist sensitive to light of wavelengths between approximately 300 – 450nm. After exposure the photoresist becomes soluble in alkali solution, so that if it is exposed with a periodic intensity of light a grating structure may be produced. The method by which this is achieved will be described later in this section.

If the substrate is spin-coated at 4000rpm with non-diluted photoresist a film thickness of approximately 1.2μm is produced. For our purposes it was desirable to have a photoresist film thickness of approximately 300-400nm, since this was the limit of grating depth we could achieve, and by diluting the photoresist with a suitable solvent (obtained from Shipley), and by varying the spin speed, the thickness of the photoresist could be reduced to produce the thinner films required. Since any dust or imperfections on the SiO_2 surface will produce gross imperfections in the quality of the film the dust free environment of the clean room was necessary, and additional

precautions against dust were made by the use of an inert gas jet which was used on the surface immediately prior to the spin coating.

After the photoresist had been deposited the samples were placed in a light-proof box to avoid accidental exposure to light. For shallow gratings, and in its use in the electronics industry, the photoresist layer is typically 'soft-baked' for 1min. on a hotplate prior to exposure in order to remove any residual solvent, and to reduce its sensitivity to light. For the deep gratings we required we needed the photoresist to be as sensitive to light as possible, and therefore this soft-bake procedure was not used, but rather the samples were left overnight in order to ensure that the photoresist layer was solvent free.

The next step in the fabrication process was to produce the grating structure in the photoresist layer. In order to do this the photoresist layer was exposed to an interference pattern generated by a two-beam interferometer (Hutley [1982]). A schematic of the interferometer system used is shown in figure 8.2.1.1.

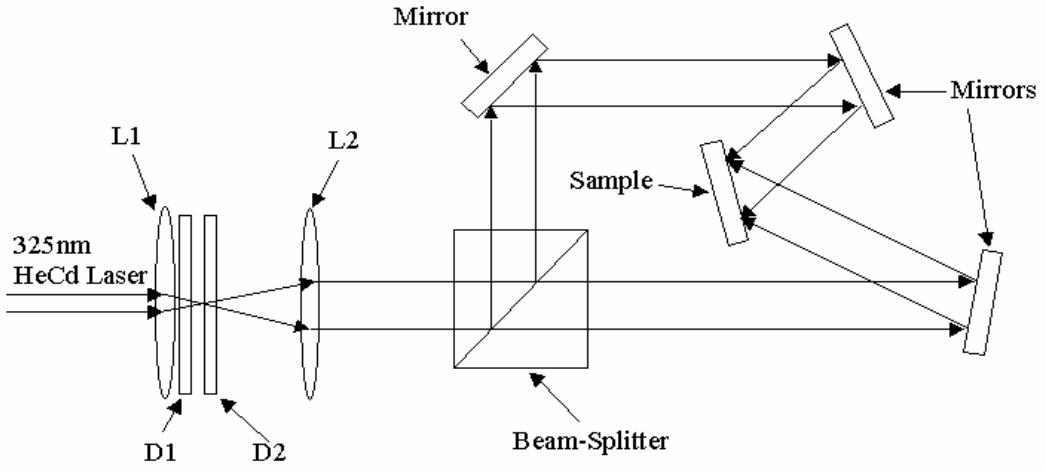

Figure 8.2.1.1 A schematic of the interferometer system used to manufacture the surface relief gratings.

The 325nm wavelength line of a HeCd laser was focused using a short focal length lens before being re-collimated using a longer focal length lens in such a way that the beam was expanded to a diameter of approximately $1cm^2$. This expanded beam passed through a 50:50 beam-splitter arrangement, manufactured by evaporating a thin layer of aluminium on the hypotenuse face of a 45° prism, which was then index matched to another 45° prism so that a cube was formed.

The two beams were reflected by a mirror arrangement shown in figure 8.2.1 so that they were incident upon a mirror which was placed in the position the sample would be for the exposure, such that the angle of incidence of the two beams (θ) on the sample were the same. The offset in position of the two mirrors is due to the need to have the path length of the two beams as similar as possible because of the limited coherence length of the HeCd laser. The incident angle θ defines the pitch of the grating to be produced, which is given by,

$$2\lambda_g \sin\theta = \lambda_0 \qquad 8.2.1$$

where λ_g is the desired grating pitch, and λ_0 is the wavelength of light used (in this case 325nm). The minimum pitch obtainable is, therefore, half of the wavelength of the light used, which in this case is 162.5nm, although due to practical limitations grating pitches of approximately 170nm were achieved. Great care is used in aligning the system so that the two beams overlay each other as precisely as possible.

Interferometer systems for creating diffraction gratings are typically used in clean rooms due to the fact that the diffraction rings created by dust particles can cause severe non-uniformity in the resultant grating. To overcome this problem, and enable the interferometer to be used in a standard laboratory, a two diffuser system was placed into the path of the laser beam between the two lenses, which also created a more uniform distribution of light over a larger area. This enabled the size of the grating area to be increased up to approximately $2cm^2$. The way in which the diffusers prevent the problems with dust in the environment of the system is as follows.

The first diffuser (D1) is stationary and is placed slightly before the focal point of the first lens (L1). This diffuser destroys the spatial coherence of the beam across its diameter, and gives a more uniform distribution of light which is essential if the grating is to be uniform over a large area (though with a stationary speckle pattern). The small speckles created by this diffuser prevent a constant interference pattern from being generated by dust particles moving through either of the two beams produced by the beam splitter. The second diffuser (D2) is then placed at the focal length of L1 and rotated. This rotation produces a time variation in the intensity of the speckle. This means that the speckle in the beam is averaged out throughout the time of exposure, giving a uniform distribution of light, and a uniform grating. Due to the fact that the speckles in the two beams need to overlap at the surface of the sample, and that the

speckle is very small, the alignment of the two beams needs to be very accurate. The two diffusers were made by lightly sand blasting silica plates and then etching in hydrofluoric acid for a couple of minutes. This creates a random pattern of 'microlenses' which are ideal for this application.

Once the interferometer was aligned the photoresist coated SiO_2 substrate was placed in position and exposed for the desired length of time. The black painted rear of the structure prevents any reflection from the rear of the substrate, which can cause standing waves normal to the surface that prevent the formation of the desired grating profile. The exposure time needed depends upon the source intensity and a couple of parameters; the grating pitch required, and the depth of grating to be fabricated. The dependence upon the grating pitch arises due to the reflection coefficient off the top surface of the photoresist altering with changing incident angle. This causes the exposure energy per unit time within the photoresist layer to be reduced for smaller grating pitches, and therefore shorter pitch gratings require longer exposure times. The depth dependence arises because the rate at which the photoresist dissolves in an alkali developer solution depends strongly upon the energy of exposure the photoresist has undergone. In fact, both the exposure and development processes are non-linear, and to create very deep gratings it is necessary to use higher exposure energies (longer exposure times) with short development times. These parameters were determined iteratively, with typical exposure times for deep short pitch gratings being of the order of 1hr.

After exposure the sample was immersed in Shipley Microposit 303 developer solution diluted to a concentration of 0.25M for approximately 2 seconds, before being immediately immersed in de-ionised water to wash off the developer solution. The water was then removed by an inert gas jet. Typical depths for a 170nm pitch grating which could be achieved by this process were of the order of 150 – 200nm, although due to the non-linear exposure and development processes the resulting profile is far from sinusoidal. The macroscopic uniformity of the gratings was poor due to the rapid development process, however they were uniform over the area of the beam spot used in the experimental determination of the sample's optical response, and were therefore deemed sufficient. An example of the grating profile obtained by taking a cross-section of one of these grating using a scanning electron microscope is shown in figure 8.2.1.2.

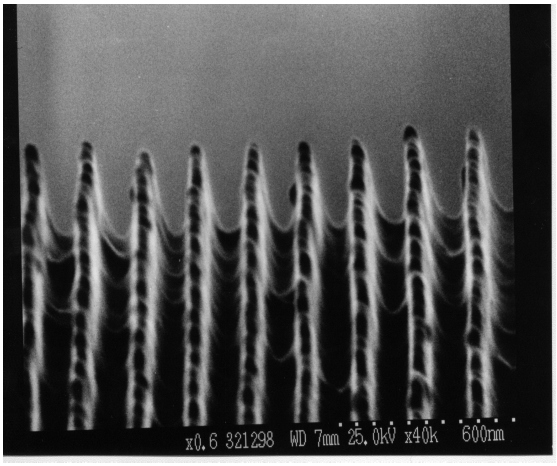

Figure 8.2.1.2 SEM image of a typical photoresist grating produced by the method described in this section.

This profile is very much like those investigated in chapter 7. Unfortunately, it is not possible to coat these gratings in metal by vacuum evaporation (to be described later in this section) since the nucleation kinetics are such that the metal preferentially deposits in the grating grooves, creating an array of metallic wires. If enough metal is evaporated so that the whole structure is immersed in metal, the top surface no longer resembles that of the photoresist grating, and no resonant features were evident in the reflectivity when they were investigated. Therefore, we decided to investigate the structure from the SiO_2 side since the evaporated metal film follows the bottom surface of the grating. This then enables gratings with a similar profile to those investigated in chapter 6 to be manufactured. However, the photoresist layer is then an extra optical layer in the equation. In order to overcome this the photoresist grating was etched so that the shape was transferred into the SiO_2 substrate using a Reactive Ion Etcher (RIE), before the metal deposition stage.

Reactive ion etching involves the use of a process gas which reacts with the substrate material and not the mask (in this case the photoresist grating) in order to transfer the grating into the substrate. A schematic of the system is shown in figure 8.2.1.3.

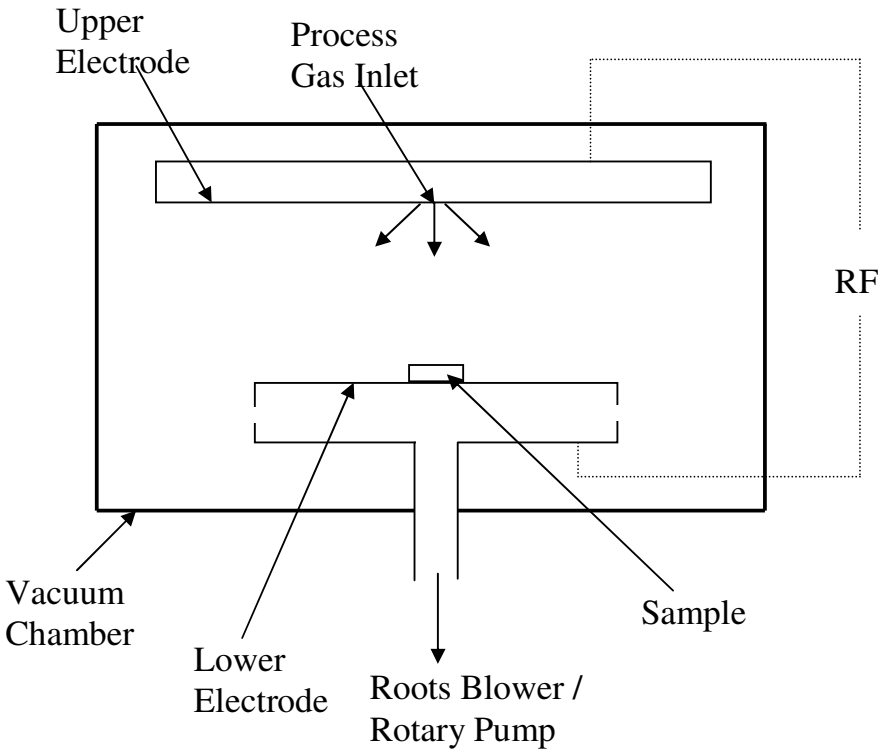

Figure 8.2.1.3 A schematic of the Reactive Ion Etcher (RIE) system.

The sample is placed in a vacuum chamber on top of one of the electrodes of an RF system. The RF creates ions in the process gas, and also causes these ions to bombard the sample so that a chemical reaction with the substrate material may occur. In order to etch the SiO_2 the process gas used is CHF_3 with the reaction mechanism being:

$$4CHF_3 + 3SiO_2 \rightarrow 3SiF_4 + 2H_2O + 4CO$$

All of the products of this reaction are gaseous at the system pressure used, and are removed by the vacuum system.

There are two main problems with this method. Firstly, the mask used is not binary, with a finite layer of photoresist above the SiO_2 substrate even at the troughs of the grating profile. And secondly, the CHF_3 ions are a precursor to a polymer which may be formed on the surface of the grating. This polymer is non-continuous and may act as many small 'micromasks', which can cause severe roughness in the resultant etched profile. In order to overcome these problems a small amount of O_2 was added to the process gas, which etches the photoresist slowly, and also reacts with the carbon in the system reducing the formation of the polymer. If the O_2 proportion is such that the

photoresist is etched by the O_2 at the same rate as the SiO_2 is etched by the CHF_3, then the resultant etched profile would be expected to be similar to that of the photoresist grating.

The parameters used in the etching process were: a pressure of 50mT, RF power of 180W, DC bias on the RF of –500V, gas flow rates of 20 standard cubic centimetres per minute for the CHF_3, and 4 standard cubic centimetres per minute for the O_2, and an etch time of approximately 15mins. With these parameters the photoresist grating was etched into the SiO_2 with an approximately 1:1 ratio, however, the resultant profile was found to be somewhat different to the initial photoresist grating profile (figure 8.2.1.4)

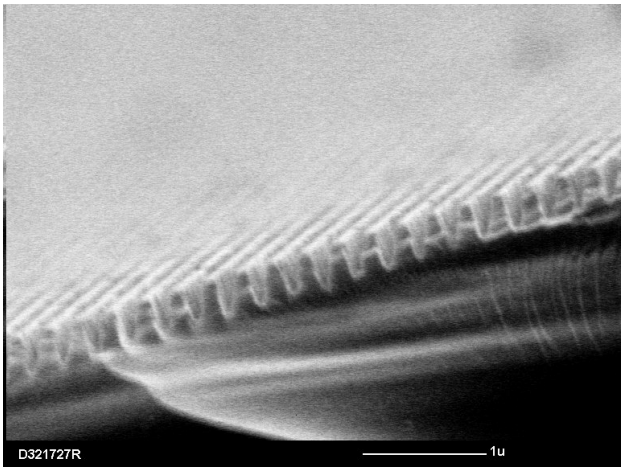

Figure 8.2.1.4 SEM image of the etched grating produced by reactive ion etching of a photoresist grating into the SiO_2 substrate.

This difference in the grating profiles is caused by polymer forming on the sides of the grating grooves. Reactive ion etching is typically performed using binary etch masks, and in order to achieve more anisotropic etching (in other words, the sidewalls of the grating are not etched whereas the bottoms of the grooves are) the fact that polymer may form on the sidewalls of the grating, but not on the tops and bottoms of the grooves, is utilised. The polymer then forms an etch resistant barrier on the sidewalls, so that only the bottoms of the grooves are etched. The reason that the polymer may form on the grating sidewalls is that there is no reactive gas incident upon them since the bombarding of the surface is in the vertical direction only.

For non-binary gratings there are no vertical sidewalls, although from the profile shown in figure 8.2.1.2 the gradient of the sides of the grating are very steep. Therefore,

it is likely that more polymer will form on the sides of the gratings than on the tops and bottoms of the grooves and peaks. This causes a flattening of the top of the grating which is evident in the etched profile shown in figure 8.2.1.4. Another consequence of the increased polymer formation on the sides of the grating is an increase in the surface roughness of the etched profile and, as will be seen later in this chapter, this is a considerable problem in producing good quality samples.

After the grating was etched in to the SiO_2 the sample was coated in silver by vacuum evaporation. A schematic of the system used to do this is shown in figure 8.2.1.5.

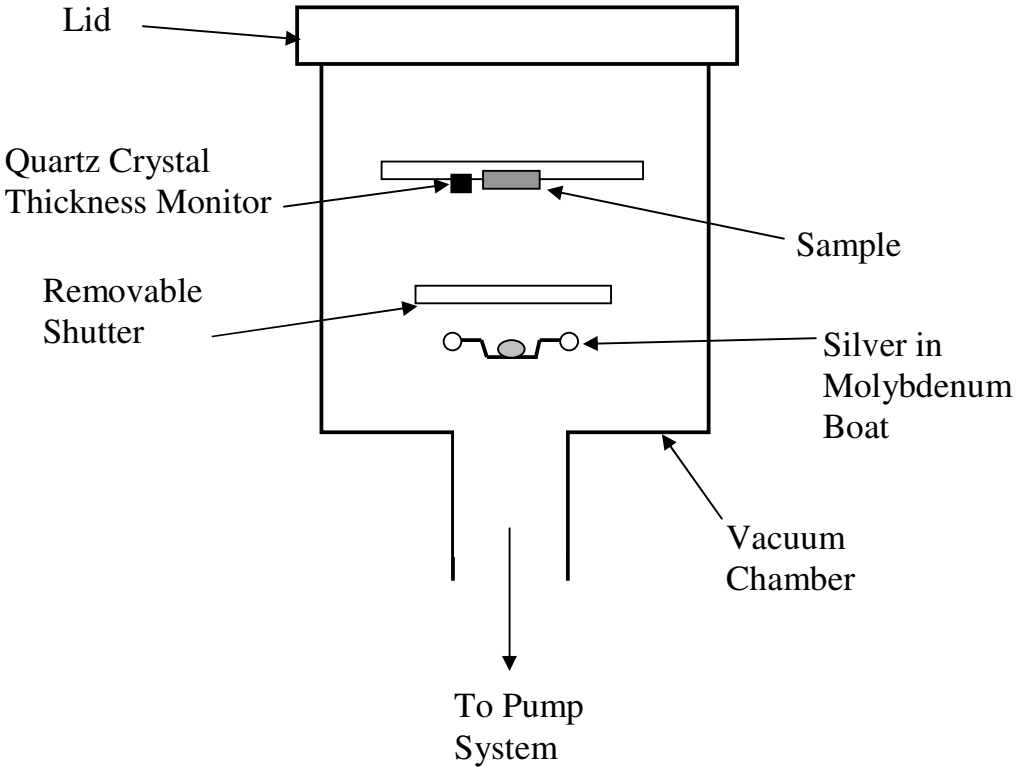

Figure 8.2.1.5 A schematic of the vacuum deposition equipment used to coat the grating in silver.

The sample is placed in a vacuum chamber with the face of the structure to be coated facing a molybdenum boat in which 99.999% pure silver is contained. The vacuum chamber is pumped down to a pressure of 1 x 10^{-6}mT by use of both a rotary pump (down to 1x10^{-2}mT) and a diffusion pump. When the base pressure has been obtained a current is passed through the molybdenum boat so that it heats sufficiently

for the silver to evaporate, with the silver vapour then condensing on to the structure. The rate at which the silver evaporates is controlled by the current being passed through the molybdenum boat, and is measured using a quartz crystal thickness monitor.

As described previously the metal preferentially deposits in the grating grooves, and therefore a film of sufficient thickness was needed to completely submerge the structure with optically thick silver. In order to do this a film thickness of approximately 500nm was evaporated onto the sample. After cooling the sample was removed from the chamber and the optical response measured.

8.2.2 Measurement of the Optical Response of the Structures

In order to measure the optical response of these structures a typical monochromator / rotating table system is used. A white light source enters a monochromator which outputs a beam of almost monochromatic light (the spectral width is approximately 0.25nm), with a wavelength range of 350nm to 850nm. Due to a low signal level blow 400nm only 400nm to 850nm is used in the experiment. The light is passed through an aperture to spatially filter the light, removing any scatter, before being passed through an optical chopper operating at approximately 1.3kHz. This allows the resultant modulated signal to be recovered from the two detectors using a pair of phase sensitive detectors. After the chopper the light passes through a polariser so that the optical response with either TM or TE polarised incident light can be investigated. A second aperture is then placed in the beam, before a beam splitter is used to remove a small percentage of the light, which then enters a reference detector. The sample is placed accurately at the centre of rotation of the rotating table at some polar angle θ to the incident beam, so that as θ is varied the area of sample being investigated always remains the same. The specularly reflected light then enters a second detector enabling the optical response of the structure to be determined. The detectors used are highly-sensitive photomultiplier tubes needed because of the low intensity signal obtained from the monomchromator source. A schematic of the system is shown in figure 8.2.2.1.

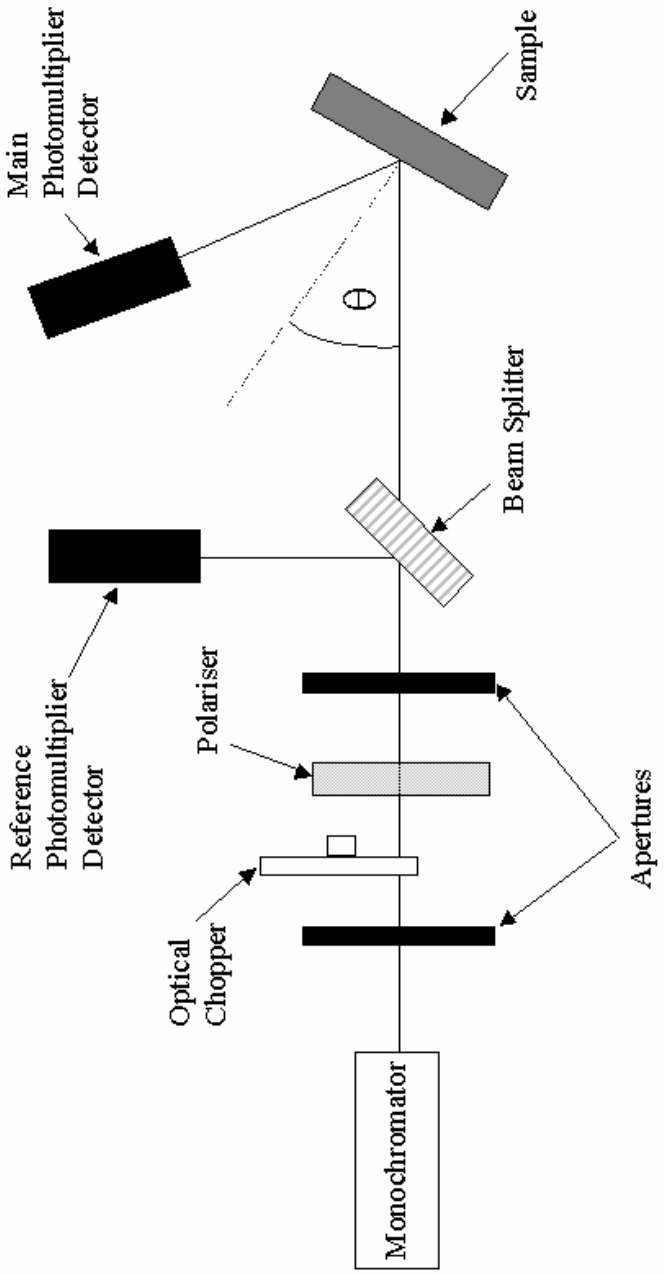

Figure 8.2.2.1 A schematic of the system used to measure the wavelength dependent reflectivities of the samples.

The samples were set-up in the system in two different ways. The aim was to record the wavelength dependent reflectivity of the sample for various values of sinθ in order to map out the dispersion of the modes, and compare them to the results shown in chapters 6 and 7. Due to the fact that the SiO_2 substrate was relatively thick there was an upper limit to the angle range we could investigate, with a maximum value of sinθ =

0.8. Since the light was incident through the SiO_2 substrate there was also a problem with the very low transmitted intensity through the sample for high incident angles due to the high specular reflection at the air / SiO_2 interface. Therefore, for higher values of sinθ (greater than 0.4), a 45° prism was index-matched to the sample so that high incident angles on the SiO_2 / silver interface could be achieved without the high incident angles on the air / SiO_2 interface.

Data were recorded as wavelength dependent reflectivity from the sample. By using the signal obtained from the reference detector, any fluctuations in intensity from the monochromator source could be averaged out since they would occur in both the reference and reflectivity signals. The signal was also normalised by dividing the results by a 'straight through run' in which the wavelength dependent signal with no sample in the system was measured. This enabled the absolute reflectivity to be determined.

By setting up the monochromator so that the maximum possible intensity output could be achieved, and by careful alignment of the system, a relatively high signal was obtained. Therefore, the signal to noise ratio was large, producing good quality data.

8.3 Results and Discussion

The wavelength dependent reflectivity for values of sinθ between 0.1 and 0.8 in steps of 0.1 were obtained for TM and TE polarised incident light over a wavelength range of 400-850nm, and for azimuthal angles of 0°, 45°, and 90°. The data sets which showed reflectivity features were those for TM polarised light at an azimuthal angle of 0°, TE polarised light at an azimuthal angle of 90°, and both polarisation for an azimuthal angle of 45°. The results for an azimuthal angle of 45° were very similar for both TM and TE polarised light, and therefore only the TM results will be shown.

8.3.1 TM polarised light incident at a 0° azimuthal angle

The wavelength dependent reflectivity of a 150nm deep 170nm pitch silver grating oriented at a 0° azimuthal angle for TM polarised incident light, and for various polar angles of incidence are shown in figure 8.3.1.1 (without prism – see section 8.2.2), and figure 8.3.1.2 (with prism).

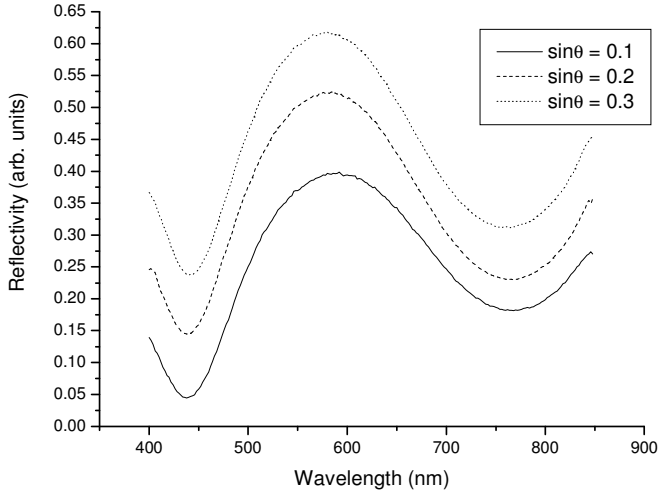

Figure 8.3.1.1 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep, 170nm pitch, silver grating oriented at an azimuthal angle of 0° with TM polarised incident light. The light is incident through the SiO_2 substrate.

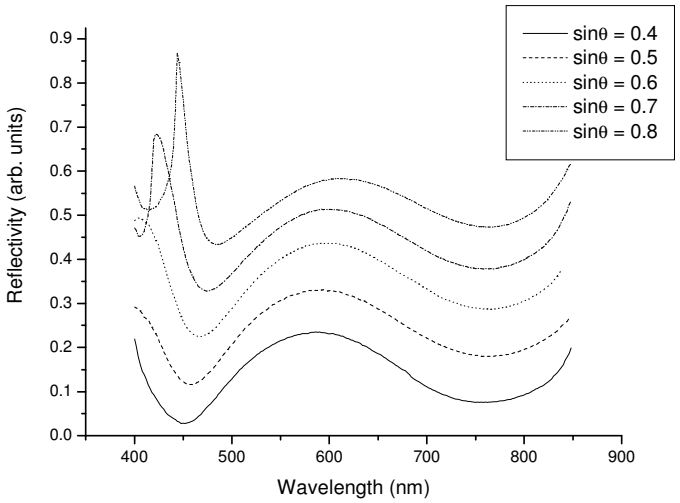

Figure 8.3.1.2 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep, 170nm pitch, silver grating oriented at an azimuthal angle of 0° with TM polarised incident light. The light is incident through a 45° SiO_2 prism which is index matched to the SiO_2 substrate of the sample.

The first point to note in these plots are the reflectivity values. In order to separate the various lines there is an offset of 0.1 on the reflectivity scale between the different data sets in both figure 8.3.1.1 and figure 8.3.1.2. Therefore, the reflectivity plotted is the absolute reflectivity *for only the lowest line in each plot*. The other point to note regarding the level of reflectivity is that it is very low when compared to a planar silver surface, and when compared to the reflectivities off resonance on the structures studied in chapters 6 and 7. This large reduction in the reflectivity is due to

the roughness caused by the reactive ion etching process which causes scatter of the incident light, and thereby a loss of energy in the specularly reflected order.

It is clear in the plots for low values of sinθ that there are two reflectivity minima at approximately 440 and 760nm wavelength. For higher values of sinθ an extra feature occurs at the low wavelength end of the spectrum which corresponds to the diffracted order pseudo critical edge. Even though the grating is 170nm in pitch it is only truly zero-order for wavelengths above approximately 500nm due to the refractive index of the SiO_2 substrate through which the reflectivity is measured.

If the positions in wavelength of the reflectivity minima are measured and converted to ω then they can be plotted in the same way as the dispersion curves shown in chapter 6. This has been performed and the results are shown in figure 8.3.1.3.

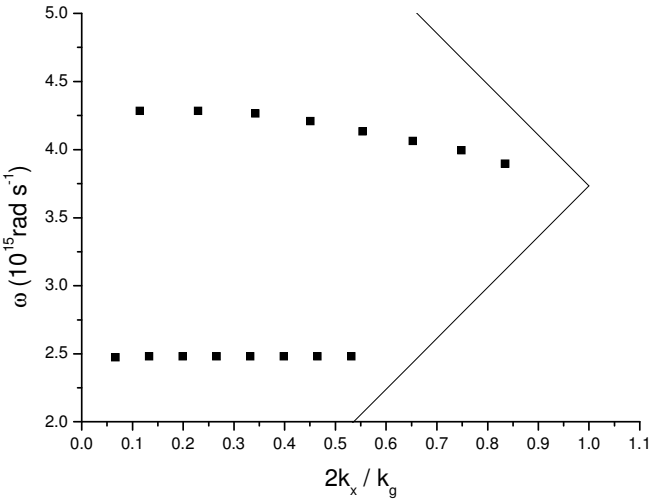

Figure 8.3.1.3 The dispersion of the reflectivity minima obtained from figure 8.3.1.1 and 8.3.1.2. The light line and diffracted order light line are also shown.

If figure 8.3.1.3 is compared with the dispersion curves of the self-coupled SPPs shown in figure 6.3.4(a) then it is noticeable that the dispersion of the two modes are very similar. The lower frequency band is very flat, changing little in frequency with increasing in-plane wavevector, with the higher frequency mode decreasing in frequency to approach the conjunction of the light line and diffracted order light line. This strongly suggests that the reflectivity minima are due to the excitation of self-coupled SPPs. Unfortunately, due to the increased scatter from the surface (and, therefore, lowering in the intensity of the specularly reflected order) it is not possible to

fit the wavelength dependent reflectivity data to theory, and therefore we cannot categorically attribute the reflectivity minima to these SPP resonances. Other problems in comparing this data to theory arise due to the large width of the modes, which is presumably due to broadening caused by a non-constant width and depth of the grating grooves. If an approximate comparison is attempted using the grating depth measured from the SEM of the cross section of the grating profile, it is found that the two modes occur at different frequencies than predicted, and are closer in frequency than would be expected if they were the first two SCSPP bands, but further apart than expected if they were the second and third SCSPP bands. This is possibly due to the fact that a single value for $\varepsilon(\omega)$ for the metal structure is not valid due to the surface roughness, although use of an intermediate cer-met layer to describe the rough surface is not able to correct for this. Therefore, it is not possible to say which order SCSPP bands have been excited.

8.3.2 TE polarised light incident at a 90° azimuthal angle

The wavelength dependent reflectivity of a 150nm deep, 170nm pitch, silver grating oriented at a 90° azimuthal angle for TE polarised incident light and for various polar angles of incidence are shown in figure 8.3.2.1 (without prism), and figure 8.3.2.2 (with prism).

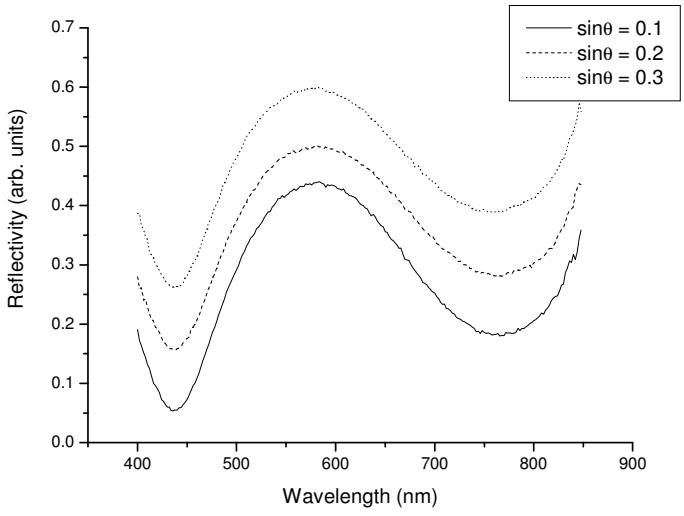

Figure 8.3.2.1 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep, 170nm pitch, silver grating oriented at an azimuthal angle of 90° with TE polarised incident light. The light is incident through the SiO_2 substrate.

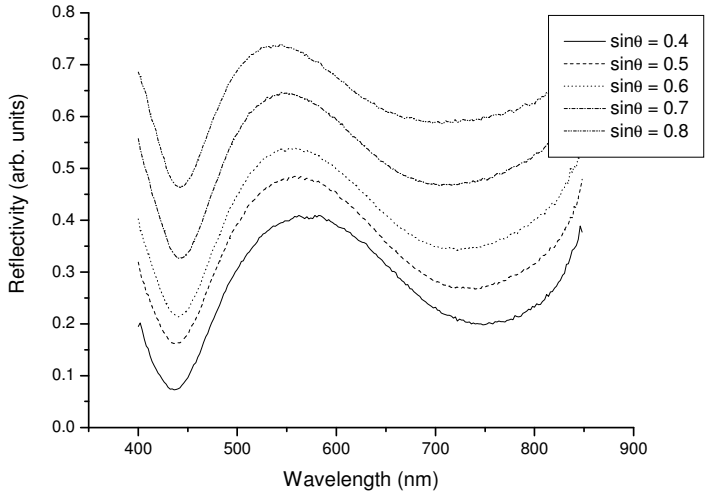

Figure 8.3.2.2 The wavelength dependent reflectivity for various angles of incidence of a 150nm deep, 170nm pitch, silver grating oriented at an azimuthal angle of 90° with TE polarised incident light. The light is incident through a 45° SiO_2 prism which is index matched to the SiO_2 substrate of the sample.

The results shown in figures 8.3.2.1 and 8.3.2.2 look very similar to those of figures 8.3.1.1 and 8.3.1.2, which for very low incident angles is not surprising since the case of TM polarised normally incident light on a grating oriented at a 0° azimuthal angle is the same as that of TE polarised normally incident light on a grating oriented at a 90° azimuthal angle. The dispersion of these modes, obtained in the same way as in section 8.3.1, is shown in figure 8.3.2.3.

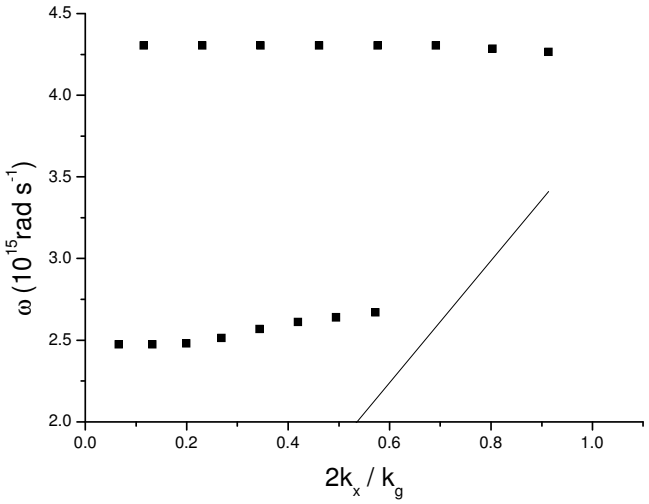

Figure 8.3.2.3 The dispersion of the reflectivity minima obtained from figure 8.3.2.1 and 8.3.2.2

If figure 8.3.2.3 is compared to figure 6.4.1 it is noticeable that the form of the dispersion of the modes are very similar, in that the modes are relatively flat, but curve up in energy for larger values of the in-plane wavevector. This agrees with the belief that these modes are SCSPPs, as described in section 8.3.1. The same discussion about the problems with comparing the data to theory also apply.

8.3.3 TM polarised light incident at a 45° azimuthal angle

When the azimuthal angle is neither 0° or 90° the polarisation of the incident light may be converted by the grating (as described in chapter 2), with the maximum in polarisation conversion occurring when the azimuthal angle is 45° Therefore, when the grating is oriented relative to the incident beam such that polarisation conversion occurs, it is necessary to measure both the polarisation conserved and converted reflectivities in order to determine the full optical response of the structure. This is achieved by placing a polariser in the reflected beam between the sample and the reflectivity detector, which can be set to allow light through of either the same polarisation state as the polariser in the beam before the sample, or the orthogonal polarisation state. Also, the results for TM polarised incident light are very similar to those for TE polarised light, and therefore only the results for TM polarised light are shown here.

In figure 8.3.3.1 the total wavelength dependent reflectivity is shown, and the polarisation conserved and converted wavelength dependent reflectivities obtained from the same structure investigated in sections 8.3.1 and 8.3.2, for the same polar angles, but with the grating oriented at a 45° azimuthal angle. The results obtained for $\sin\theta < 0.4$ were obtained without a prism matched to the SiO_2 substrate, whereas for those with $\sin\theta \geq 0.4$ the light was incident upon the structure via a prism.

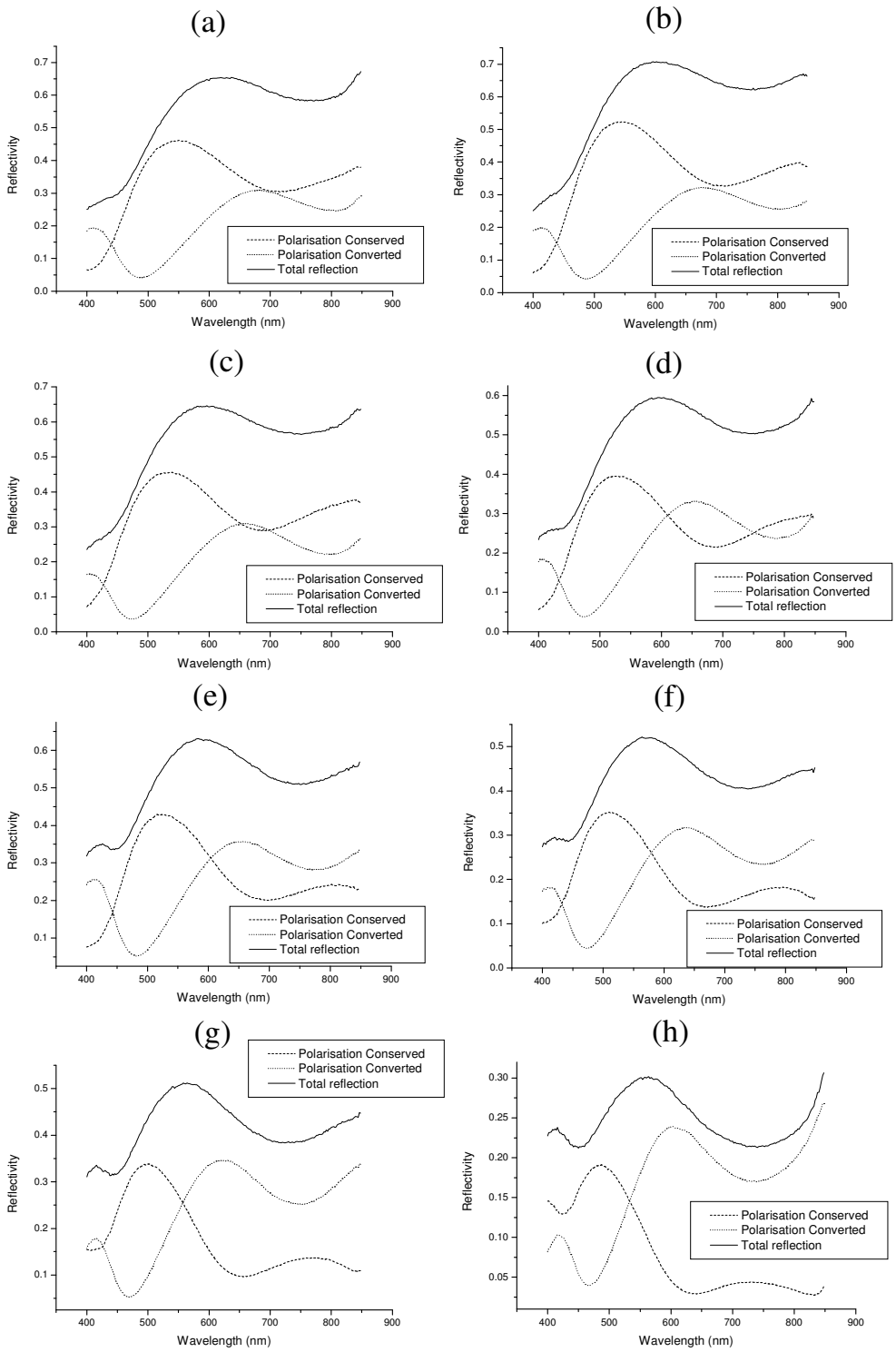

Figure 8.3.3.1 The polarisation conserved and converted wavelength dependent reflectivities for various angles of incidence of a 150 deep, 170nm pitch, silver grating oriented at an azimuthal angle of 45° with TM polarised incident light. For a) to c) the light is incident through the SiO_2 substrate, whereas for d) to h) it is incident through a 45° SiO_2 prism which is index matched to the SiO_2 substrate of the sample. a) $\sin\theta = 0.1$, b) $\sin\theta = 0.2$, c) $\sin\theta = 0.3$, d) $\sin\theta = 0.4$, e) $\sin\theta = 0.5$, f) $\sin\theta = 0.6$, g) $\sin\theta = 0.7$, h) $\sin\theta = 0.8$.

There are several features to note from the plots in figure 8.3.3.1. Firstly, it should be noted that the reflectivities measured are the absolute reflectivities, although the result obtained for $\sin\theta = 0.8$, exhibits a surprisingly low reflectivity when compared to the rest of the plots. This is possibly due to the increased effective beam area on the sample, causing a greater region of the grating to be sampled. This would cause any non-uniformity over the grating region to have a more pronounced effect upon the reflectivity obtained. Another point to note is that it is only the total reflectivity plots for the various polar angles (produced by combining the polarisation conserved and converted reflectivities), which show the position of the reflectivity minimum due to SPP excitation correctly, and therefore it is these which must be considered when the dispersion of the mode with changing in-plane wavevector is determined. It is also clear that the maximum in the polarisation conversion does not occur at the same frequency as the reflectivity minimum in the total reflectivity plots, and this contradicts the results obtained in chapter 6 regarding the polarisation conversion from structures exhibiting SCSPP excitation. In fact, it is more reminiscent of the broad-band polarisation conversion obtained from the narrow-peaked Gaussian structures investigated in chapter 7, since in some of the plots the polarisation converted reflectivity is greater than the polarisation conserved reflectivity over much of the wavelength range investigated. The reason for this is relatively simple.

In the case of a grating consisting of a series of Gaussian grooves the vast majority of the light is reflected from the top surface of the grating, and therefore there is very little polarisation conversion caused by the depth of the grating grooves (see chapter 2 for an explanation of the polarisation conversion mechanisms). For the Gaussian peaked gratings the case is very different, with considerable polarisation conversion occurring due to the depth of the grooves and, since the broad-band polarisation conversion is a result of an interaction between the SPP mediated polarisation conversion and the polarisation conversion caused by the depth of the grating grooves, the broad-band polarisation conversion does not occur on the narrow-grooved gratings, but does occur on the wider groove gratings. The structure investigated here is not as narrow grooved as those investigated in chapter 6 (as is evident from the SEM of the grating profile in figure 8.2.1.4), and therefore there will be increased polarisation conversion occurring due to the depth of the structure. This

will then lead to broad-band polarisation conversion like that obtained from the peaked structures investigated in chapter 7. The necessity of a $2k_g$ component in the grating profile which is +90° out of phase with the k_g component for SCSPPs to be excited is still satisfied even though the grooves are broader in the structure investigated.

By determining the positions in wavelength of the reflectivity minima in the total reflectivity curves in figure 8.3.3.1, and plotting them on a ω-k plot in the same way as in the previous sections the dispersion of the SPP mode can be obtained and is plotted in figure 8.3.3.2.

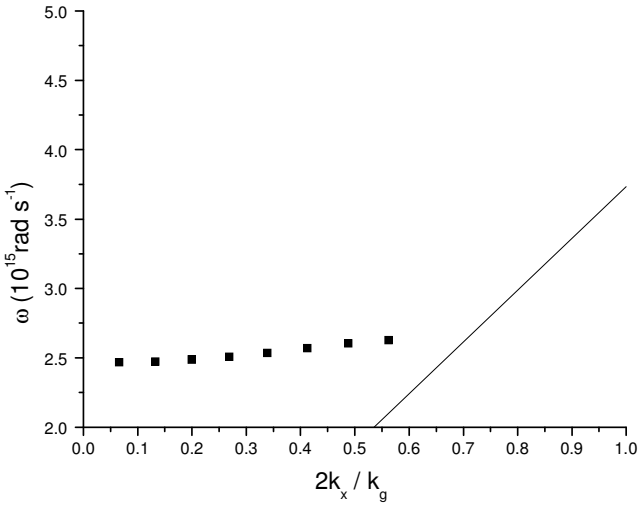

Figure 8.3.3.2 The dispersion of the reflectivity minima obtained from the total reflectivity plots of figure 8.3.3.1

The form of the dispersion of the mode with changing in-plane wavevector is very similar to that shown in figure 6.5.1, and this again suggests that the reflectivity minima are due to the excitation of SCSPPs.

8.4 Summary

In this chapter a method for manufacturing deep zero-order gratings for the visible region of the spectrum has been presented. Data has also been presented which show the optical response obtained from one of these gratings for TM polarised incident light when the grating was oriented at 0° and 45° azimuthal angles, and TE polarised incident light when the grating was oriented at a 90° azimuthal angle. These results

show that resonant absorption of light may occur, which is believed to be due to the excitation of self-coupled SPPs. To our knowledge this is the first time that SPPs have been observed on zero-order gratings.

Due to problems with the manufacturing process, which are inherent in the method, the surface of the grating structure was found to be relatively rough. This caused diffuse scatter, and therefore reduced the specularly reflected order intensity. For this reason it was not possible to fit the data obtained to theory, and therefore it was not possible to categorically attribute these reflectivity minima to the excitation of SCSPPs. However, the dispersion of these modes with changing in-plane wavevector was found to be very similar to those obtained for the SCSPPs in chapter 6, and therefore we believe that it is these modes which have been excited. The main difference in the optical response obtained from the gratings produced when compared to those investigated in chapter 6 was that broad band polarisation conversion occurred when the grating was oriented at a 45° azimuthal angle. The reason for this was explained in terms of the grating grooves being broader than those investigated in chapter 6.

Chapter 9

Conclusions and Future Work

9.1 Summary

The work in this thesis presents an original investigation of the optical response of various grating structures, with a view to advancing the understanding of the effects of short-pitch corrugations on the various orders produced. Both dielectric and metal gratings have been studied with either one or two interfaces, with several interesting results.

In the work on dielectric structures the main new results which are of interest occur on the two interface systems, where, if the slab thickness is very small and the two identical corrugations on each surface are in phase with each other, a total lack of transmitted diffracted orders is found to occur. This is shown to arise due to cancellation of the diffracted orders created at the two interfaces in the transmitted half-space. When the phase between the two gratings on each surface is changed there are two main results. Firstly, the structure acts as if it is effectively blazed for the first reflected diffracted order (the +1 and –1 diffracted orders are found to have different intensities). Secondly, when the two corrugations are in anti-phase with each other, it is found that almost all of the energy of normally incident light can be equally distributed between the +1 and –1 transmitted diffracted orders, with almost no intensity in either the reflected orders or zeroth transmitted order. Though this is found to be a special case, the range of grating parameters over which it occurs is shown to be relatively large, and the origin of the effect is discussed.

Similar original investigations have been undertaken with the slab material being a good metal. In this case SPPs are excited on both interfaces, but when the metal slab is bounded by dielectrics of different refractive indices (and therefore the frequency of excitation of the SPPs on the two metal / dielectric interfaces is different) the SPP on the transmitted side of the structure is found to be only very weakly excited. This effect is shown to arise due to the same transmitted diffracted order cancelling mechanism as is found on the dielectric structures. When the phase between the two surface corrugations is non-zero, coupling to this SPP is found to arise once more since this cancellation no

longer occurs. The way in which excitation of the SPPs can give maxima, minima, or Fano shaped resonances in the transmitted zeroth order is also discussed. This is ascribed to the phase difference between the light coupled out of the SPP, and that which passes straight through the structure.

If the bounding dielectrics have the same refractive index then coupled SPPs are found to occur on the thin metal slabs. These coupled SPPs are described and the effect of the thickness of the slab, the phase between the two corrugations, and the addition of higher harmonics to the description of the grating profile on the optical response of the structure are investigated. The explanation of the results obtained for coupled SPPs is more complicated than for the SPPs bound to only one interface, but it is shown how both transmission maxima and minima can occur in the transmitted zeroth order depending upon the phase between the two corrugations, and also the way in which band gaps are produced in the dispersion curves of the coupled SPPs. Typically there are two coupled SPP modes for a thin metal slab, the long range SPP and the short range SPP, however it is found that when there is a $2k_g$ component in the grating profile, and the two corrugations are anti-symmetric, a total of four coupled modes exist. These arise due to the different possible charge distributions of the modes with respect to the grating profile.

The next section of the thesis concerned the influence of very deep corrugations on the nature of the SPPs on a single interface dielectric / metal structure. Initially the case of a grating consisting of a series of narrow Gaussian grooves is explored, showing that there are a family of self-coupled SPPs, which are extremely flat banded with fields highly localised within the grooves, and which occur even in the zero-order region of the spectrum. These newly discovered modes are examined for all possible grating orientations with respect to the incident light, and it is shown that they arise due to the formation of very large band gaps in the SPP dispersion curves, and through interactions between the various SPPs which can be excited on the structures.

Following this, the inverse structure of a grating consisting of a series of narrow Gaussian peaks is studied for the first time. The dispersion and nature of the SPP modes which can be excited on the structure are investigated, and found to be radically different to those of the Gaussian grooved structures previously studied. The complex dispersion of these modes is described for gratings oriented at azimuthal angles of 0°, 45°, and 90°, and the polarisation of light which can excite the modes for each of these

orientations is also described. One surprising phenomenon which occurs is that of broad-band polarisation conversion when the gratings are oriented at a 45° azimuthal angle. The gratings show polarisation conversion of >80% over the whole of the visible region of the spectrum for normally incident light, and the way in which this occurs is described.

In the final chapter, results of an original experimental study on these deep zero-order metal gratings are presented. The fabrication method of the structures is described, along with the method used to obtain the optical response of the structure. The wavelength dependent reflectivity results obtained for various polar and azimuthal angles show reflectivity minima in the specularly reflected order. The dispersion of the modes with changing in-plane wavevector has been obtained from these results, showing very flat bands which are similar to those of the self-coupled SPPs described in chapter 6. Unfortunately, it was not possible to fit the data to theory due to a large amount of surface roughness caused by the manufacturing process. Therefore, it was not possible to categorically ascribe these modes to the self-coupled SPPs, although the evidence of the dispersion curves strongly suggested that this was the case.

9.2 Possible Applications

SPPs have, so far, formed the basis for relatively few applications. One field in which they have been somewhat successful is as chemical and biological sensors, where their sensitivity to the refractive index of the dielectric is used to determine changes in the local environment at the interface. The structures investigated in this thesis do not appear to lend themselves to this field, but there are several other possible applications for which they may be of use.

The SPPs on the thin metal slabs (chapter 5) could be used to enable more light to be coupled out of organic LEDs or, in fact, any light emitting structure where a proportion of the power is typically lost to surface EM modes or absorption in metal layers. This possibility is the subject of a patent application.

For the high-aspect ratio zero-order structures there are a number of possible applications. The narrow-grooved structures (chapter 6) produce selective absorption of light over a band of frequencies, which is relatively independent of the angle of incidence. For this reason these structures have received interest as stealth materials.

They have also received interest regarding a possible use as active surfaces for surface enhanced Raman spectroscopy. In this case they may be of particular interest since the molecules could fit entirely in to the grooves in which the enhanced fields of the SCSPP are contained.

In the case of the narrow-ridged structures (chapter 7) it is the broad-band polarisation conversion effect which lends itself to applications. Their possible use for increasing the amount of light emitted from back-lit LCDs is described in chapter 7. In addition to this they may be used as anti-counterfeiting materials. A system whereby a silver film which appears black in reflection through parallel polarisers, but silver through crossed polarisers, would be a simple, though very difficult to forge, anti-counterfeiting system.

9.3 Future Work

The majority of future work resulting from this thesis lies in the experimental observation of the numerous phenomena which have been predicted. The work on the thin slab structures presents many problems in regard to grating fabrication. The conformal structures may be manufactured using similar techniques to those described in chapter 8, and in fact some thin conformal metal slabs have been investigated in the past (for example Schröter and Heitmann [1999] investigated thin metal slabs bounded by dielectrics with different refractive indices). However, it is by no means obvious what methods could be used to fabricate the structures when the corrugations on the two interfaces are out of phase with each other, although any method by which this could be controllably accomplished would open up a large field of research. Similarly, it would be of great interest if the fabrication problems involved in manufacturing the deep zero-order structures described in chapters 6,7 and 8 could be overcome to give precise control over the form of the grating profiles. This could lead to several applications, and may be possible using advance fabrication techniques such as e-beam lithography.

In terms of further modelling work to more fully describe the structures investigated in this thesis there are several future avenues which could be explored. Firstly, the depth of the corrugations on the thin dielectric and metal slabs could be increased since, especially in the case of the metal slabs, the effect on the electromagnetic modes of the system of the deeper corrugations would significantly

alter the optical response obtained. Also, it would be interesting to investigate the optical response of deep lamellar single interface metal gratings, since similar results to those obtained in chapter 6 and 7 would be expected, and lamellar structures could certainly be manufactured using e-beam lithography equipment.

In addition there are numerous other structures which would be expected to show similar effects to those described in this thesis. The most obvious of these would be deep, short-pitch, bigrating structures (either square or hexagonal), since it would be expected that SCSPPs could be excited on these structures if the grooves were narrow, and that the coupling to these modes would be the same for both linear polarisations for all azimuthal angles of incidence. Also, bigratings of the narrow peaked structures would develop into arrays of narrow peaks, which could also have very interesting optical properties.

Other, more novel, structures could combine the use of dielectric and metal gratings to take advantage of the coupled modes of the systems, and it would also be interesting to explore the optical properties of non-single valued surfaces. In this case the optical properties of structures consisting of resonant grooves / cavities with a controlled opening on the top surface could be explored, with the coupling strength to the modes of the systems being controlled by the size of the openings.

Publications

I. R. Hooper and J. R. Sambles, *Dispersion of surface plasmon polaritons on short pitch metal gratings*, Phys. Rev. B **65** 165432 (2002)

I. R. Hooper and J. R. Sambles, *Surface Plasmon Polaritons on Narrow Peaked Short Pitch Metal Gratings*, (accepted by Phys. Rev. B)

I. R. Hooper and J. R. Sambles, *A Broad-Band Polarization Converting Mirror for the Visible Region of the Spectrum*, (accepted by Optics Letters)

I. R. Hooper and J. R. Sambles, *Surface Plasmon Polaritons on Narrow-Ridged Short-Pitch Metal Gratings in the Conical Mount*, (Submitted to J. Opt. Soc. Am. A)

A patent application on controlling emission from fluorescent species which involves the work on the thin metal slab structures has been filed in conjunction with QinetiQ (Malvern).

Other publications intended for submission in the future include:

1. A paper on the thin slab dielectric structures described in chapter 3.
2. A paper on the thin slab metal gratings bounded by dielectrics with different refractive indices described in chapter 5 (after resolution of issues pertaining to the patent application described above)
3. A paper on the coupled SPP modes found on thin slab metal gratings bounded by identical dielectrics described in chapter 5.
4. A paper on the narrow-peaked short-pitch metal gratings in the conical mount described in chapter 7.

References

Andrewartha, Fox and Wilson	1979(a)	Optica Acta **26** 69
Andrewartha, Fox and Wilson	1979(b)	Optica Acta **26** 197
Avrutsky IA, Svakhin AS, Sychugov VA	1989	J. Mod. Opt. **36** 1303
Barnes WL, Kitson SC, Preist TW, Sambles JR	1997	J. Opt. Soc. Am A **14** 1654
Barnes WL, Preist TW, Kitson SC, Sambles JR	1996	Phys. Rev. B **54** 6227
Botten C, Craig MS, McPhedran RC, Adams RC, Andrewartha JR	1981(b)	Optica Acta **28** 1087
Botten C, Craig MS, McPhedran RC, Adams RC, Andrewartha JR	1981(a)	Optica Acta **28** 413
Bryan-Brown GP, Sambles JR, Hutley MC	1990	J. Mod. Opt. **37** 1227
Chandezon J, Dupuis MT, Comet G, Maystre D	1982	J. Opt. Soc. Am. **72** 839
Chandezon J, Maystre D, Raoult D	1980	J. Optics (Paris) **11** 235
Chesler RB, Karr MA, Geusic JE	1970	Proc. IEEE **58** 1899
Depine RA, Lester M	2001	J. Mod. Opt. **48** 1405
Dakss ML, Kuhn L, Heidrich PF, Scott BA	1970	Appl. Phys. Lett. **16** 523
Dummery G, Filippi P	1970	CR Acad. Sci. Paris **270** 137
Dupta Gupta S, Varada GV, Agarwal GS	1987	Phys. Rev. B **36** 6331
Ebbesen TW, Lezec HJ, Ghaemi HF, Thio T, Wolff PA	1998	Nature (London) **391** 667

Elston SJ, Bryan-Brown GP, Sambles JR	1991	Phys. Rev. B **44** 6393
Fano U	1941	J. Opt. Soc. Am. **31** 213
Ferrell R	1958	Phys. Rev. **111** 1214
Flodstrom SA, Bachrach RZ	1976	Rev. Sci. Inst. **47** 1464
Garcia-Vidal FJ, Sanchez-Dehesa J, Dechelette A, Bustarret E, Lopez-Rios T, Fournier T, Pannetier B	1999	J. Lightwave Tech. **17** 2191
Ghaemi HF, Thio T, Grupp DE, Ebbesen TW, Lezec HJ	1998	Phys. Rev. B **58** 6779
Greffet JJ	1988	Phys Rev. B **37** 6436
Greffet JJ, Maasarani Z	1990	J. Opt. Soc. Am A **7** 1483
Harris JB, Preist TW, Sambles JR, Thorpe RN, Watts RA	1996	J. Opt. Soc. Am. A **13** 2041
Hibbins AP, Sambles JR, Lawrence CR	1998	J. Mod. Opt **45** 2051
Hickernell RK, Sarid D	1986	J. Opt. Soc. Am. B **3** 1059
Hutley MC, Vernill JP, McPhedran RC, Neviére M, Vincent P	1975	Nouv. Rev. Optique **6** 87
Hutley MC	1982	*Diffraction Gratings* (Academic Press)
Inagaki T, Goudonnet JP, Arakawa ET	1986	J. Opt. Soc. Am **7** 992
Inagaki T, Motosuga M, Arakawa ET, Goudonnet JP	1985	Phys. Rev. B **32** 6238
Johnson RH	1973	IEEE J. Quant. Elec. **QE-19** 255
Jory MJ, Bradberry GW, Cann PS, Sambles JR	1995	Meas. Sci. Technol. **6** 1193

Kim SH, Fonstad CG	1979	IEEE J. Quantum Elec. **QE-15** 1405
Knop K	1978(a)	J. Opt. Soc. Am. **68** 1206
Knop K	1978(b)	Appl. Opt. **17** 3598
Kretchmann E, Raether H	1968	Z. Naturf. A **23** 2135
Laks, Mills, and Maradudin	1981	Phys. Rev. B **23** 4965
Li L	1993	J. Opt. Soc. Am. A **10** 2581
Li L	1995	Opt. Comm. **114** 406
Li LF	1994	J. Opt. Soc. Am. A **11** 2816
Lopez C, Yndurain FJ, Garcia N	1978	Pys. Rev. B **18** 970
Lopez-Rios T, Mendoza F, Garcia-Vidal FJ, Sanchez-Dehesa J, Pannetier B	1998	Phys. Rev. Lett. **81** 665
Lorrain P, Corson DR, Lorrain F	1988	*Electromagnetic Fields and Waves*, 3rd ed. (W.H. Freeman and company) 646
Maradudin AA	1983	J. Opt. Soc. Am **73** 759
Martin-Moreno L, Garcia-Vidal FJ, Lezec HJ, Pellerin KM, Thio T, Pendry JB, Ebbesen TW	2001	Phys. Rev. Lett. **86** 1114
Maystre D	1978	J. Opt. Soc. Am. **68** 490
Mills DL	1977	Phys. Rev. B **15** 3097
Moaveni MK	1988	IEEE Proceedings **135** 318

Moaveni MK, Kalhor HA, Afrashteh A	1975	Comput. Elect. Eng **2** 265
Moharam MG, Gaylord TK	1977	J. Opt. Soc. Am. A **71** 811
Moharam MG, Gaylord TK	1982	J. Opt. Soc. Am. A **72** 1385
Gaylord TK, Moharam MG	1985	Proc. IEEE **73** 894
Moharam MG, Gaylord TK	1986	J. Opt. Soc. Am. A **3** 1780
Nash DJ, Cotter NPK, Wood EL, Bradberry GW, Sambles JR	1995	J. Mod. Opt. **42** 423
Neureuther A, Zaki K	1969	Alta. Freq. **38** 282
Neviére M, Maystre D, Vincent P	1977	J. Opt. Soc. Am. **8** 231
Neviére M, Carutti-Maari G, Cadilhac M	1971	Optics Communications **3** 48
Numata H	1982	Journal of the Physical Society of Japan **51** 2575
Nylander C, Liedberg B, Lind T	1982	Sensors and Actuators **3** 79
Otto A	1968	Z. Phys. **216** 398
Pai DM, Awada KA	1991	J. Opt. Soc. Am. A **8** 755
Palik ED (*editor*)	1985	*Handbook of Optical Constants of Solids*, Academic Press
Peng ST Tamir T	1974	Opt. Comm. **11** 405
Petit R	1966	Rev. Opt. **8** 353
Pipino A, Schatz GC	1994	J. Opt. Soc. Am. A **11** 2036

Plumey JP, Granet G, Chandezon J	1995	IEEE T Antenn. Propag. **43** 835
Pockrand I	1974	Phys. Lett. **49a** 259
Popov E, Mashev L	1986	Optica Acta **33** 593
Porto, Garcia-Vidal and Pendry	1999	Phys. Rev. Lett. **83** 2845
Preist TW, Cotter NPK, Sambles JR	1995	J. Opt. Soc. Am. A **12** 1740
Raether H	1988	*Surface Plasmons on Smooth and Rough Gratings* (Springer-Verlag, Berlin)
Rayleigh JWS	1907	Proc. Roy. Soc. A **79** 399
Rayleigh Lord	1907	Phil. Mag. **14** 213
Ritchie RH, Eldridge HB	1961	Phys. Rev. **126** 1935
Saloman, Grillot, Zayato and de Fornel	2001	Phys. Rev. Lett **86** 1110
Sarid D	1982	Phys. Rev. Lett. **47** 1927
Schröter U, Heitmann D	1999	Phys. Rev. B **60** 4992
Schwartz L	1965	Methodes Mathematique pour les Sciences Physics (Hermann, Paris)
Schwartz L	1966	Théorie des Distributions (Hermann, Paris)
Seshadri SR	1985	J. Appl. Phys. **58** 1733
Sobnack MB, Tan WC, Wanstall NP, Preist TW, Sambles JR	1998	Phys. Rev. Lett. **80** 5667

Stegeman GI, Burke JJ, Hall DG	1982	Appl. Phys. Lett. **41** 906
Steinmann W	1960	Phys. Rev. Lett. **15** 470
Strong J	1935	Phys. Rev. **48** 480
Tamir T (*editor*)	1979	*Integrated Optics*, (Springer-Verlag, Berlin)
Tan WC, Preist TW, Sambles JR, Wanstall NP	1999	Phys. Rev. B **59** 12661
Tibuleac S, Magnusson R	1997	J. Opt. Soc. Am. A **14** 1617
Tien PK	1977	J. Mod. Phys. **49** 361
Tien PK, Ulrich R, Martin RJ	1969	Appl. Phys. Lett. **14** 291
Tillin MD, Sambles JR	1988	Thin Solid Films **167** 73
Vincent P	1978	Optics Communications **26** 293
Vincent P, Neviére M, Maystre D	1978	Nucl. Instrum. Methods **152** 123
Watson TJ, Sambles JR	1991	Phil. Mag. B **65** 1141
Watts RA, Hibbins AP and Sambles JR	1999	J. Mod. Opt. **46** 2157
Watts RA, Preist TW and Sambles JR	1997	Phys. Rev. Lett. **79** 3978
Watts RA, Sambles JR	1997	Opt. Comm. **140** 179
Watts RA, Sambles JR, Harris JB	1997	Optics Com. **135** 189
Watts RA, Sambles JR, Hutley MC	1998	J. Mod. Opt **45** 639

Watts RA, Sambles JR, Hutley MC, Preist TW, Lawrence CR	1997	Nanotechnology **8** 35
Weber MG, Mills DL	1985	Phys. Rev. B **31** 2510
Wirgin A	1964	Rev. Opt. **9** 449
Wood El, Sambles JR, Cotter NPK, Kitson SC	1995	J. Mod. Opt. **42** 1343
Wood RW	1902	Phil. Mag. **4** 396
Yamasaki T	1995	IEICE Trans. Electron. **E78-C** 322