THEORIES OF SINGLY AND DOUBLY PERIODIC DIFFRACTION GRATINGS

by

L. C. Botten, B.Sc. (Hons.), University of Tasmania

A thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

in the

UNIVERSITY OF TASMANIA

HOBART

February, 1978

Except as stated herein, this thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief, this thesis contains no copy or paraphrase of material published by another person, except when due reference is made in the text of the thesis.

Lindsay Botten,

January 1978.

Any part of this thesis may be photocopied.

CONTENTS

SUMMARY

In this thesis, theories for the solution of diffraction problems involving both singly periodic and doubly periodic structures are presented. The studies have been motivated by two fields of application. The first of these is the use of singly periodic gratings in spectrographic instruments, with particular attention being devoted to diffraction anomalies and possible means for their reduction. The second pertains to the use of diffracting structures in systems having solar absorbing behaviour with the aim of optimizing their selective properties. The unity of the thesis rests not in the applications of the theories but in the rigorous electromagnetic methods used to establish the variety of formalisms presented.

The first three chapters are essentially introductory in their nature and contain an extensive review of integral formalisms for singly periodic gratings. Both infinite conductivity and finite conductivity theories are discussed and are applied to the study of higher order blaze effects, groove depth determination and the characterization of anomalies of the 'plasmon' type.

An original formalism for multi-layer transmission gratings is presented in chapter 4. The theory is then used in chapter 5 for the optimization of the selective properties of thin film solar absorbers. Surface roughening, modelled using a singly periodic profile modulation, is shown to improve the absorptance by up to ten percent.

Chapter 6 contains a new "integral-modal" treatment of the old problem involving the diffraction of a plane wave by a perfectly

conducting grating having a triangular profile with a right-angled apex. This is included since many of the concepts discussed therein are of relevance to the following chapter.

The next section, chapter 7, is concerned with the diffraction properties of bi-metallic gratings (which are structures composed of two species of metals). The evolution of this study is discussed and in doing so a number of theories culminating in a new and completely general formalism are presented. Numerical results obtained reveal that anomaly suppression can be achieved by overcoating the "off-blaze" facet of a triangular grating with a poorly conducting metal.

The remaining three chapters are devoted to the study of doubly periodic structures using modal formalisms. Chapter 8 is concerned with the diffraction properties of a crossed lamellar transmission grating, which is an inductive grid whose two mutually orthogonal axes of periodicity lie in spatially separated planes. The theory and a new amplitude constraint appropriate to a general Littrow mount are given together with some numerical results indicating promising solar selective behaviour.

The following chapter considers a singly periodic double grating composed of a pair of spatially separated lamellar transmission gratings. The theoretical formalism is presented together with a detailed discussion of the application of this structure as a long wavelength Fabry-Perot interferometer. Also contained in chapter 9, are the results of a comprehensive search for conservation relations (phase constraints) pertaining to singly periodic symmetric gratings.

Finally in chapter 10, a theory for inductive grids having circular apertures is discussed. It is shown that by inserting dielectric plugs

in the aperture and surrounding the grid with a symmetric pair of lossless thin films, the degraded transmission properties of such grids (caused by low hole to area fractions) can be substantially overcome.

ACKNOWLEDGEMENTS

It is with sincere gratitude that I wish to express my thanks to the following people for their assistance and also their support throughout the course of these studies.

Dr. Michael Waterworth (University of Tasmania), my supervisor, who introduced me to the fascinating subject of physical optics in my undergraduate career. During the past four years, his willingness to discuss my progress and problems and offer valuable suggestions as to further worthwhile avenues of research has been of great stimulus to me. More importantly, however, one came to regard him as being more a friend than a supervisor and it is with deep appreciation that his incomparable support is acknowledged.

Throughout the course of my project, it has been my priviledge to have worked closely with Dr. Ross McPhedran (University of Sydney) and Miss Jenny Adams (University of Tasmania). Ross has been an immense source of inspiration showing unbounded enthusiasm and interest in all of my investigations. His inimitable scientific method, his deep physical insight and his profoundly valued guidance, particularly during the preparation of this thesis, have made our collaboration a thoroughly exciting experience for the author. Jenny, who has been my co-worker in the investigations reported in chapters 8 and 9, has also been a tremendous source of motivation during the past two years. It has been her zeal, her thoroughness and her unfailing perseverance that have lent such a fresh approach to the subject and it is these qualities which have made our work together

such a rich and rewarding time for me. To both Ross and Jenny, I can offer only my most heart-felt thanks for your friendship and loyalty which have made these studies an experience that shall never be forgotten.

Dr. Ian Wilson (University of Tasmania) whose unqualified appreciation of the practical aspects of the subject has motivated many of my theoretical investigations and has made our collaboration a most enthralling experience. Also, my other fellow members of the Diffraction Grating Research Group of the University of Tasmania, Miss Barbara Brown and Mr. John Andrewartha, with whom I have worked from time to time and whose studies have almost invariably influenced, if not stimulated, my own efforts. To these people who have proffered so many useful suggestions and so graciously suffered my usual overly enthusiastic participation, I gratefully acknowledge your contributions and hold your "brotherhood" in most friendly regard.

During the latter half of 1977, I was priviledged to pursue my sutdies within the School of Physics at the University of Sydney. Thus, it is with great pleasure that I wish to thank the members of the Department of Theoretical Physics and the Solar Energy Group for the provision of research facilities and their overwhelming kindness during my stay. In particular, I should like to thank Dr. Graham Derrick for showing such an interest in my work and for giving me the benefit of his vast experience and wisdom in the concepts of modern physics. Sincere thanks must also be extended to Dr. David McKenzie for his sustained interest in my studies and for his many illuminating discussions, particularly with reference to the solar energy aspects of the work.

I should also like to mention the contributions of Mr. Denis Porteus and Mr. Marc Duldig of the University of Tasmania together with Dr. "Wibs" Smith, Dr. Robert Hewitt, Mr. Dennis Sams and Mr. Colin Mathers of the Univeristy of Sydney for their support and for providing me with such good cheer during my studies.

I reserve my very special thanks for Miss Pat Moroney who so willingly undertook the abhorrent task of typing this thesis in such a competent and painstaking fashion. More importantly, however, it has been her kindness and thoughtfulness that has made my stay at the University of Sydney such a memorable experience for me.

Finally, I turn to the immense contribution to my work that has been made by my parents. It has been their loyalty and limitless faith that has helped to carry me through the inevitable periods of frustration.

I should also like to acknowledge the financial assistance throughout my course afforded by a Commonwealth Post-Graduate Research Award.

CHAPTER 1

INTRODUCTION

This thesis is concerned with the theoretical analysis of diffraction problems involving a wide range of periodic structures. From a cursory examination of its contents, the rapid evolution in the complexity of problems solved becomes apparent. This not only emphasizes the progress made by the entire field but also reflects upon the changing priorities and demands imposed on the theory by exciting new applications.

When the work towards this thesis was commenced in 1974, the theory had only just reached the stage where the effects of finitely conducting materials could be successfully accommodated. Up until that time, comparisons between theory and experiment were made primarily with the aim of obtaining confirmation of the theory. Only four years later, the theoretical understanding has advanced to such an extent that the earlier situation is reversed with limitations in measurements often being revealed by confrontation with the theory. Furthermore, in this year of 1978, any aspiring grating theoretician must be prepared for the demands of doubly periodic structures, should he wish his work to be of relevance to the burgeoning field of solar energy research.

It is this rapid evolution which explains the wide variety of diffraction problems encountered in this thesis, which are solved by an equally wide variety of methods. Hence the unity of the thesis rests not so much in the form of the theories presented but in their

origin in rigorous applications of Maxwell's equations to the modelling of electromagnetic scattering processes. With these prefatory remarks, it is now appropriate to briefly examine the theoretical development of the subject.

The study of diffraction gratings dates back to the work of Fraunhofer (in 1821), who summarized their dispersive properties in the well known grating equation

$$\sin \theta_n = \sin \theta_i + \frac{n\lambda}{d} \quad ,$$

where θ_n denotes the angle of diffraction of the n^{th} order component of the diffracted field excited by a plane wave of wavelength λ incident upon a structure of period d at some angle θ_i. However, the more difficult problem concerned with the actual scattering mechanism is of a more recent origin, with major advances towards its solution being prevented until the advent of powerful digital computers during the mid-1960's.

Over the past few years, there have been a number of excellent reviews [1.1-5] of the literature in this subject. Since it would be inappropriate to duplicate these comprehensive works, it is the aim of this introduction to outline the important milestones in the development of the theory to its present state, and in doing so, to place the material contained in this thesis in its proper context.

The literature of grating theory abounds with diverse formulations proposed by authors who are frequently unaware of the contributions of others. Many of these formalisms represent no significant

advance in the theory in that they have not extended the range of groove depths, surface conductivities or types of structure capable of solution. As a rule, the most important advances have been made by the Marseille group headed by Professeur R. Petit, although isolated workers such as Bolomey [1.6], Wirgin [1.6-7] and Van den Berg [1.8] have made outstanding contributions.

Early attempts to characterize the energy properties of diffraction gratings involved the use of Kirchhoff's scalar theory. With such a technique it was possible to obtain a solution of the diffraction problem in closed form. However, any formalism based on the assumption of scalar optics cannot, in general, be an adequate representation of the scattering problem since one of the first experimental observations made concerning gratings was that they polarize light. Furthermore, as shown by McPhedran [1.3], such theories are also found wanting with respect to the fundamental constraint of energy conservation.

The first of the vector diffraction theories for plane reflection gratings was proposed by Lord Rayleigh [1.9]. He characterized the diffracted field at all points in space by a discrete expansion of plane waves (whose directions were given by the grating equation). The Rayleigh theory was successful in many ways and principally in that it could explain some of the properties of the then newly discovered Wood anomalies [1.10]. Since its original exposition, it has been utilized by numerous authors, but its results for gratings having deep grooves have tended to be unsatisfactory. In consequence, there has been much debate [1.11-12] in the literature

concerning the validity of plane wave expansions within the grating grooves (often referred to as the Rayleigh approximation or assumption).

The first author to avoid its use was Petit [1.13] who derived an integral equation treatment for the problem of the diffraction of a plane wave by a perfectly conducting grating. In his original work, Petit proposed that the integral equation be solved using a Fourier series technique which essentially limited his algorithm to profiles composed of a number of linear segments. Nevertheless, his theory was completely rigorous and the results obtained were of a far higher accuracy than those derived using either scalar methods or the Rayleigh assumption. As such, Petit must be regarded as being the "father" of modern grating theory.

In 1970, Pavageau and Bousquet [1.14] reported a further integral equation approach to the same problem. These authors suggested that their integral equation (involving the same fundamental unknown as Petit's theory, namely the surface current density) be solved using an iterative technique. Using such a method, they were able to obtain results of high accuracy for arbitrary profile shapes. Since then however, other workers in the field have shown the points matching technique to be more widely applicable than iteration.

It was not until the early 1970's that the problem of diffraction by a finitely conducting grating was solved successfully. Van den Berg [1.8], in 1971, derived a formulation of this problem involving the numerical solution of a pair of coupled integral equations. Although his theory was entirely rigorous, its solution did present certain

numerical difficulties arising from the need to characterize the field problem by a pair of unknowns. It was not until the following year that Maystre [1.4] was able to express these two unknowns in terms of a single entity analogous to the surface current density of perfectly conducting structures. In doing so, he was able to derive a single integral equation, thereby circumventing many of the previous numerical difficulties. In the opinion of the author, this has been the most important advance in integral formalisms made during this decade in that it has revolutionized the approach to diffraction grating problems. As will become evident from a perusal of the material presented in this thesis, it is Maystre's work which has most heavily influenced the author's investigations concerning singly periodic gratings.

A second approach to these same problems has also proved to be of immense value in recent years. This technique, referred to as the differential formalism, involves the direct integration of the Helmholtz equation using numerical means. The work of Nevière et al [1.15], conducted in parallel with the investigations of Maystre, realized a theory capable of accommodating a vast range of diffracting geometries with relative ease. However, Nevière's approach encountered some serious numerical difficulties when dealing with highly conducting metals and S polarized radiation.

There also exists a third and entirely separate class of formulations commonly referred to as the "modal formalisms", whose roots lie deep inside waveguide theory. However, to some extent they are allied to the integral formalisms, as evidenced by the work of Wirgin and Deleuil [1.16]. These require a knowledge of the exact nature of the

field within the grooves, which is matched to the Rayleigh expansion for free space by the method of moments. The successful use of these techniques requires the derivation of analytic expressions for the field modes and as such is restricted to relatively simple geometries such as lamellar gratings. Furthermore, their application is even further restricted since it has not as yet been possible to formulate modes for finitely conducting structures. Nevertheless, they do permit an enhanced understanding [1.17], of the diffraction problem and have also proved to be of great value in recent studies concerning doubly periodic structures, for which the integral formalisms would exceed the capabilities of present computers.

During 1976, an investigation [1.18] undertaken by the author, concerning the diffraction properties of a multi-layer transmission grating, extended the earlier integral formalisms of Maystre and Van den Berg to consider structures only slightly less general than those tackled by Nevière's differential theory. It was not until 1977 that it was realized that Maystre [1.19] had conducted a similar but more general study resulting in the formulation of a theory for multiple profile gratings. His latest treatment can accommodate diffracting geometries as diverse as those handled by the differential formalism. Since his algorithm is not afflicted with the same numerical problems as those referred to above, it must be taken that this is now the ultimate in conventional grating theories.

With the rapid advances made during the past few years, there now remain very few unsolved problems of significance to the grating theorist. One of the few outstanding questions concerns the properties

of "bi-metallic gratings", which are structures composed of two species of metals. The results of an investigation of this topic conducted by the author are presented in chapter 7 of this dissertation. The theories presented therein (which have been largely motivated by the works of Maystre) have revealed that anomaly suppression may be achieved by selectively overcoating the "off-blaze" facet of a triangular profile grating with a poorly conducting metal.

It has indeed been fortunate for grating theorists that the rapidly diminishing number of problems involving singly periodic structures available for solution has been accompanied by the opening up of a fresh avenue of research to which they may usefully apply their talents. Much of their expertise in classical electromagnetism is of great relevance to solar energy. A major portion of the work towards this thesis has been motivated by the problems posed in this new field and in this context the author has been particularly fortunate to have been given the benefit of a close interaction with members of the Solar Energy Group of the University of Sydney.

In general, the marked polarization effects exhibited by singly periodic gratings effectively eliminate their use in solar selective systems. Because of the desire to collect as much of the incident flux as possible, it is essential to use structures which exhibit no polarizance of the incident light, particularly for the important case of normal incidence. Hence, only those structures possessing square symmetry are acceptable and it is this that has initiated the theorist's interest in the diffraction properties of doubly periodic structures. Nevertheless, as evidenced by chapter 5 of this thesis, it is still

possible to usefully apply some of the conventional grating theories to these problems. In that chapter are contained the results of a study concerned with improvements to the design of thin film solar selective absorbers. It is shown that surface roughening of the appropriate scale (modelled using a singly periodic grating theory) can enhance the absorption of light by minimizing the interference action inside the relatively transparent film, if a grating absorption feature is located at those wavelengths for which the reflectance is gratest.

However, the great bulk of the material in this thesis pertaining to solar energy is concerned with a class of doubly periodic structures known as inductive grids. Until recent years, these were traditionally studied by electrical engineers interested in antenna systems. It was not until a suggestion by Horwitz [1.20] concerning the use of such structures as an effective solar-selective surface that investigations into their optical properties were instigated.

The relevant theoretical literature is of a comparatively recent origin, with the outstanding contributions being made by Chen [1.21-23] during the early years of this decade. In 1975, McPhedran and Maystre [1.24] extended Chen's analysis to study the solar selective properties of a type of inductive grid constructed by perforating a thick planar sheet of metal with periodically spaced rectangular apertures. Their work revealed that such structures having deep grooves and large hole-area fractions possessed excellent selective properties provided that they were operated in direct illumination and tracked the sun.

During 1976, the author became involved in these investigations, largely at the suggestion of his friend and colleague, Dr. Ross McPhedran of the University of Sydney. This facet of the work towards the thesis then proceeded on two major fronts. The first of these, concerned with an extension of the above theory to the case of inductive grids with circular apertures, was conducted jointly by author and Dr. Ross McPhedran. The results of this study, reported in chapter 10, involved a further expansion of Chen's original work and showed that the relatively poor packing achievable with circular apertures severely restricted the transmittance of the grid. In a subsequent extension of the analysis, made so as to accommodate the presence of a dielectric plug within the aperture and a pair of lossless symmetric uniform films surrounding the grid, it was discovered that by the appropriate choice of these additional agents, the degraded transmission properties could be restored to the level obtained by square-holed grids.

The second aspect of these grid studies was conducted concurrently with the above work in collaboration with another friend and co-worker, Miss Jenny Adams of the University of Tasmania, and also with Dr. Ross McPhedran. This study was concerned with the diffraction properties of a structure known as the crossed lamellar transmission grating, a perfectly conducting "grid" whose two mutually orthogonal axes of periodicity lie in spatially separated planes. This work, which has been accepted for publication [1.25], is detailed in chapter 8, where it is shown that such structures also possess excellent solar selective properties.

A continuation of this study by Miss Jenny Adams [1.26] has

generalized the above formalism to include non-orthogonal axes of periodicity. The material contained in chapter 9 has been largely motivated by the interesting case arising when the angle of inclination between the two arrays is reduced to zero. The structure so obtained is referred to as the double grating, for which an integral formalism was proposed by Blok and Mur [1.27] in 1972. In chapter 9 is presented a modal formalism for this problem, developed in collaboration with my co-worker in the subject, together with a discussion of the use of this structure as a Fabry-Perot interferometer. During the course of our investigation, a number of particularly intriguing analytic properties, characteristic of the symmetry and the lossless nature of the grating, came to light. Consequently, a thorough search for the origin of these phenomena was commenced, the results of which are comprehensively discussed in that same chapter.

Having placed the content of this thesis in context with the available literature and expertise in the field, it becomes immediately apparent from the wide variety of structures studied (shown in table (1.1)), that the only common link in these investigations is contained in them being rigorous applications of electromagnetic theory. Throughout the following chapters, considerable effort has been made to apply the formalisms to physically relevant problems, rather than to present the theories as bald exercises in physical and numerical analysis. Furthermore, within the theoretical sections, emphasis is placed not only upon the successes but also on the difficulties and failures encountered. This has been done so as to provide a true record of the investigations; but more importantly it is included in the hope that those who read this dissertation may profit from the author's own experiences.

Table 1.1

NATURE OF THE PROBLEM	METHOD OF SOLUTION	TYPE OF APPLICATION	CHAPTER
The perfectly conducting reflection grating of arbitrary profile.	Integral formalism	Spectroscopic	2
The finitely conducting reflection grating of arbitrary profile.	Integral formalism	Spectroscopic	3
The multi-layer transmission grating composed of one arbitrary periodic surface, surrounded by a set of plane interference layers.	Integral formalism	Solar	4,5
A triangular profile grating having a right-angled apex ruled in perfectly conducting metal and buried beneath a finitely conducting layer.	Integral-Modal formalism	-	6
Semi-finitely conducting bi-metallic reflection and transmission gratings.	Integral-Modal formalism	Spectroscopic	7
The buried cylinder grating.	Integral formalism	-	7
The general bi-metallic grating.	Integral formalism	-	7
The crossed lamellar transmission grating.	Modal methods	Solar	8
The double grating.	Modal methods	Interferometric	9
The round-holed inductive grid.	Modal methods	Solar	10

A brief outline of the theories for the variety of structures which are both reviewed and evolved within this thesis. The table also briefly mentions the nature of the physical problems to which these formalisms have been applied.

REFERENCES

[1.1] G. W. Stroke, in "Encyclopedia of Physics", 29 (1967) 644. (Berlin: Springer Verlag)

[1.2] R. Petit and D. Maystre, Rev. de Phys. appliquée, 7 (1972) 427.

[1.3] R. C. McPhedran, unpublished Ph.D. Thesis, (1973), University of Tasmania.

[1.4] D. Maystre, Thèse No. AO 9545, (1974), L'Université d'Aix-Marseille III, France.

[1.5] R. Petit, Nouv. Rev. Optique, 6 (1975) 129.

[1.6] J. C. Bolomey and A. Wirgin, Proc. I.E.E.E., 121 (1974) 794.

[1.7] A. Wirgin, Thèse No. AO 1429, (1967), Paris, France.

[1.8] P. M. Van den Berg, Ph.D. Thesis, (1971), Delft University of Technology, Delft, The Netherlands.

[1.9] Lord Rayleigh, Proc. Roy. Soc. A, 79 (1907) 399.

[1.10] R. W. Wood, Phil. Mag., 4 (1902) 396.

[1.11] R. Petit and M. Cadilhac, C. R. Acad. Sci. Paris, 262 (1966) 468.

[1.12] M. Nevière and M. Cadilhac, Optics Commun., 2 (1970) 235.

[1.13] R. Petit, Rev. d'Opt., 45 (1966) 249.

[1.14] J. Pavageau and J. Bousquet, Optica Acta, 17 (1970) 469.

[1.15] M. Nevière, P. Vincent and R. Petit, Nouv. Rev. Optique, 5 (1974) 65.

[1.16] A. Wirgin and R. Deleuil, J. Opt. Soc. Amer., 59 (1969) 1348.

[1.17] J. R. Andrewartha, J. R. Fox and I. J. Wilson, submitted for publication to Optica Acta.

[1.18] L. C. Botten, accepted for publication by Optica Acta.

[1.19] D. Maystre, accepted for publication by J. Opt. Soc. Amer.

[1.20] C. M. Horwitz, Opt. Commun. 11 (1974) 210.

[1.21] C-C. Chen, I.E.E.E., MTT-18, 18 (1970) 627.

[1.22] C-C. Chen, I.E.E.E., MTT-19, 19 (1971) 475.

[1.23] C-C. Chen, I.E.E.E., MTT-21, 21 (1973) 1.

[1.24] R. C. McPhedran and D. Maystre, Appl. Phys., 14 (1977) 1.

[1.25] J. L. Adams, L. C. Botten and R. C. McPhedran, accepted for publication by J. Optics (Paris).

[1.26] J. L. Adams, (1977), University of Tasmania, Report DGRG 77/7.

[1.27] H. Blok and G. Mur, Appl. Sci. Res., 26 (1972) 389.

CHAPTER 2

INFINITE CONDUCTIVITY INTEGRAL FORMALISMS

2.1 INTRODUCTION

In the first part of this chapter, a comprehensive discussion of the integral formalisms covering the years 1966 to 1970 is presented. Although no longer having the same physical relevance as their more sophisticated successors, they have, nevertheless, shaped the entire evolution of the subject. With the rapidly diminishing number of significant problems awaiting the attention of the grating theoretician, it is now perhaps appropriate to review the development of the subject. This is done not only to place the remainder of the material in this thesis in its proper perspective but also to do justice to the efforts of others such as Petit [2.1] and Pavageau and Bousquet [2.2], whose works have so greatly influenced, indeed stimulated, the contributions of the author.

The review material contained in this and in the following chapter pertains only to the integral equation treatments. However, this must not be taken as a denigration of other excellent works relying upon different methods of solution. Instead, it should be taken more as a comment upon the author, whose experience is restricted to this particular aspect of the subject.

The second part of this chapter is devoted to an application of these theories. In that section is presented the results of an investigation, concerned with the blaze properties of spectrographic diffraction

gratings, conducted jointly with Dr. Ian Wilson (University of Tasmania) and Dr. Ross McPhedran (University of Sydney). It represents a continuation of previous studies conducted by such authors as Maystre, Petit, McPhedran and Wilson (references [2.3-2.9]) whose efforts have been almost exclusively restricted to the case of a first order Littrow configuration. These latest investigations have revealed that gratings having triangular grooves are far less sensitive to the deleterious effects of angular deviation than are holographic gratings. It has been found that in the case of ruled gratings, angular deviation can often provide a useful reduction in the undesirable effects of anomalies. On the other hand, the use of gratings having sinusoidal or quasi-sinusoidal profiles with angular deviations much in excess of 10^{o} should be avoided. Also given within that section are the results of our studies concerning higher order blaze properties. These have shown that holographic gratings are far less suited to operation with a plurality of orders than are ruled gratings.

2.2 THEORETICAL PRELIMINARIES

It is the purpose of this section to outline a number of theoretical properties which are essential prerequisites to any study of grating problems.

2.2.1 Uniqueness and the radiation conditions

Firstly, the solution of such problems is uniquely determined by the outgoing wave condition and by the boundary conditions applying at the grating surface. The outgoing wave condition stipulates that the

field scattered by the grating should, at infinity, not contain any incoming wave contributions. When written in a mathematical form, it is referred to as Sommerfield's radiation condition (see for example [2.10], pp 56-57 and 562-566). Physically speaking, the condition is required because a unique solution to a scattering problem cannot be determined unless the incident field is specified uniquely. A corollary of the uniqueness condition states that if the far field radiation patterns excited by two separate beams incident upon the same structure are identical, then the field distributions for both problems are identical at every point in space.

2.2.2 The grating equation

Let us consider a plane wave, E^i, incident at some angle θ_i (measured from the grating normal and taken as being positive in an anti-clockwise sense) on a grating of period, d. It is our aim to calculate the angles of diffraction (θ_n) of the various orders (n) excited by the interaction of the incident wave with the grating surface. These angles are also measured from the grating normal and this time are taken as positive in a clockwise sense. Now the incident wave may be written in the form

$$E^i = \exp[i(\alpha_o x - \chi_o y)] \qquad (2.2.1)$$

where

$$\left.\begin{aligned} \alpha_0 &= k \sin\theta_i, \\ \chi_0 &= k \cos\theta_i, \\ \text{and} \quad k &= \frac{2\pi}{\lambda} \quad . \end{aligned}\right\} \tag{2.2.2}$$

A suppressed temporal dependence of $\exp(-i\omega t)$ is assumed throughout this thesis. Outside the grooves of the grating the scattered field may be expressed as a superposition of plane wave terms of the form

$$P_n(x,y) = \exp[i(\alpha_n x + \chi_n y)] \tag{2.2.3}$$

where $\alpha_n = k \sin\theta_n$,

and $\chi_n = \sqrt{(k^2 - \alpha_n^2)}$ if $k \geq |\alpha_n|$

$= i\sqrt{(\alpha_n^2 - k^2)}$ if $k < |\alpha_n|$.

The above choice of the sign of χ_n when it is purely imaginary has been made to ensure field convergence at $y = +\infty$.

Now, the incident field possesses the property which is commonly referred to as "pseudo-periodicity"

i.e. $E^i(x+d,y) = E^i(x,y) \exp(i\alpha_0 d)$.

This means that to an observer stationed at a distance of one grating period from the reference point, the incident field is phase-shifted by an an amount $\alpha_0 d$. Furthermore, this observer's view of the grating (taken to be infinite in extent) is identical to that of the observer stationed

at the reference origin. Both agree on the form of the radiation conditions at infinity and thus the field problem is essentially unaltered except for the phase shift of the incident field. From the uniqueness condition, the second observer sees the scattered field as being composed of terms of the form $\{P_n\}$, each shifted in phase by an amount $\alpha_0 d$

i.e. $P_n(x+d,y) = P_n(x,y) \exp(i\alpha_0 d)$

from which is obtained

$$\alpha_n d = \alpha_0 d + 2n\pi$$

$$\text{i.e.} \sin\theta_n = \sin\theta_i + \frac{n\lambda}{d} \qquad (2.2.4)$$

- the well known grating equation.

The above derivation, attributed to Petit [2.1], has relied totally upon the grating being assumed to be infinite in extent, since otherwise the "pseudo-periodicity" constraint would not hold for the entire field.

2.2.3 Boundary conditions

Let us now consider the form of the boundary conditions which apply at the grating surface. In two-dimensional diffraction problems, in which the incident wave-vector lies in the x-y plane (i.e. perpendicular to the grating grooves) two fundamental polarizations may be distinguished. In E|| or P polarization, the electric field vector of the incident wave is aligned with the grooves of the grating and the diffraction problem is

solved in terms of the electric field components parallel to the Oz axis. In H|| or S polarization, the magnetic field vector of the incident wave lies along the Oz axis and the diffraction problem is solved in terms of magnetic field quantities.

Given the grating behaviour for each of the two fundamental polarizations, it is a trivial matter to deduce its behaviour for an arbitrary polarization angle in a two dimensional problem. Furthermore for a three dimensional or conical diffraction problem, D. Maystre [2.11] has proved an elegant theorem which relates the behaviour to those in P and S polarized radiation for a two dimensional mount. However, Maystre's theorem does not hold for gratings of finite conductivity.

In the case of a perfectly conducting grating used in P polarized radiation, if E denotes the magnitude of the total (incident and scattered) electric field, then the fundamental boundary condition is

$$E = 0 \tag{2.2.5}$$

at the grating surface $y = f(x)$. For S polarization, the same boundary condition ($\underline{n} \times \underline{E} = 0$) is expressed by the equation

$$\frac{dH}{dn} = 0, \tag{2.2.6}$$

where the surface normal derivative is taken of the z-component of the total magnetic field. Since the grating has infinite conductivity, a surface current of non-zero density $\underline{j}_s$ can exist on its surface, where

$$\underline{n} \times \underline{H} = \underline{j}_s . \tag{2.2.7}$$

In these three boundary conditions, the left hand sides of the expressions must be interpreted as limits for a field point approaching the grating surface from free space.

For a grating of finite conductivity, there exist no free charges or surface currents at $y = f(x)$. Instead bulk currents given by Ohm's law

$$\underline{J} = \sigma \underline{E}$$

flow through the medium, assumed to be isotropic. Furthermore, in all subsequent analyses it is presumed that the grating is constructed of a non-magnetic material (i.e. a material whose relative permeability does not differ significantly from unity).

The finite conductivity boundary conditions are the continuity of the tangential components of $\underline{E}$ and $\underline{H}$. In the case of P polarized light, the exterior limit (in air) and the interior limit (in the grating) are related in the following manner:

$$E\Big|_{ext} = E\Big|_{int} , \tag{2.2.8}$$

$$\frac{\partial E}{\partial n}\Big|_{ext} = \frac{\partial E}{\partial n}\Big|_{int} . \tag{2.2.9}$$

(In the derivation of the latter constraint from the continuity of the tangential component of $\underline{H}$, it has been necessary to apply the assumption

that the grating proper is non-magnetic.) For S polarization, the corresponding boundary conditions are:

$$H\Big|_{ext} = H\Big|_{int} \tag{2.2.10}$$

and

$$\frac{\partial H}{\partial n}\Big|_{ext} = \frac{1}{r^2}\frac{\partial H}{\partial n}\Big|_{int} . \tag{2.2.11}$$

(Where r denotes the relative refractive index of the internal and external media.) Equation (2.2.4) implies that $\partial H/\partial n$ varies discontinuously at $y = f(x)$, with the relative magnitude of the discontinuity increasing with the refractive index jump at the grating surface. It is this fact which makes it far more difficult to obtain well-convergent differential formulations for highly conducting gratings in S polarized rather than P polarized light.

2.2.4 Edge conditions

In the solution of any diffraction problem it is imperative to consider the behaviour of the fields near profile edges. This behaviour must be understood and indeed compensated for if numerical algorithms are to be elaborated for profiles composed of linear segments (the most important case of which is the triangular profile). An excellent discussion of this topic for perfectly conducting surfaces has been given by Jones [2.10, pp. 562-569]. He shows that the uniqueness of the solution for a diffracting obstacle having an edge is preserved provided that the edge does not radiate any energy. He further demonstrates that field components parallel to the edge are bounded, and that if r denotes a radial distance measured from the edge and β denotes the included angle

at the edge then components perpendicular to the edge behave as

$$r^{(\beta-\pi)\ /\ (2\pi-\beta)}$$

for small values of r. Such components are singular if $\beta < \pi$, with the worst singularity ($r^{-\frac{1}{2}}$) occurring for $\beta = 0$ (i.e. at the edge of an infinitesimally thin sheet). In the important use of $\beta = \pi/2$, these field components behave as $r^{-\frac{1}{3}}$ for small r.

Meixner [2.12] has considered field behaviour near the edge of a finitely conducting surface. He showed that the effect of finite conductivity was to "soften" the singularity. (This would be expected on physical grounds since lowering the conductivity must decrease the impedance mismatch between air and the metal to thus decrease the rate of spatial variation of the fields.) Maystre [2.11], in his finite conductivity formulation, has "removed" these singularities by dividing the current density by a function having a singularity as bad or worse than that of the physical function. Thus, a function suitable for removing the singularity in an infinite conductivity treatment will also be effective for the same grating composed of finitely conducting metal.

2.3 A REVIEW OF THE INFINITE CONDUCTIVITY INTEGRAL FORMALISMS

In this section, it is proposed to review the works of Petit [2.1], Bolomey [2.13], Pavageau and Bousquet [2.2] and in doing so to trace the evolution of the infinite conductivity integral formalisms between the years 1966 and 1970. For the purposes of clarity, the discussion will be split into separate sections relating to P and S polarized radiation.

2.3.1 Formulations for P polarized radiation

The first of the rigorous diffraction theories was expounded by Petit [2.1]. In his original article, he considered an incident plane wave field of free space wavelength λ incident upon a structure (shown in figure 2.3.1) of period d at some angle θ_i. The incident field $\underline{E}^i$ of unit amplitude is written in the form

$$\underline{E}^i = \exp[i(\alpha_0 x - \chi_0 y)]\, \hat{\underline{k}}$$

$$\left.\begin{aligned} \text{where } \alpha_0 &= k \sin\theta_i, \\ \chi_0 &= k \cos\theta_i, \\ \text{and } k &= 2\pi/\lambda . \end{aligned}\right\} \qquad (2.3.2)$$

This field excites a surface current j_s (directed along the Oz axis) which in turn re-radiates a scattered field of the same polarization. This is given by

$$E^d(P) = E(P) - E^i(P) \qquad (2.3.3)$$

(after removing the vector dependence). Here, E(P) represents the total electric field at some point P in the space D_0.

The field E^d must obey the Helmholtz equation

$$(\nabla^2 + k^2)E^d(P) = 0 \qquad (2.3.4)$$

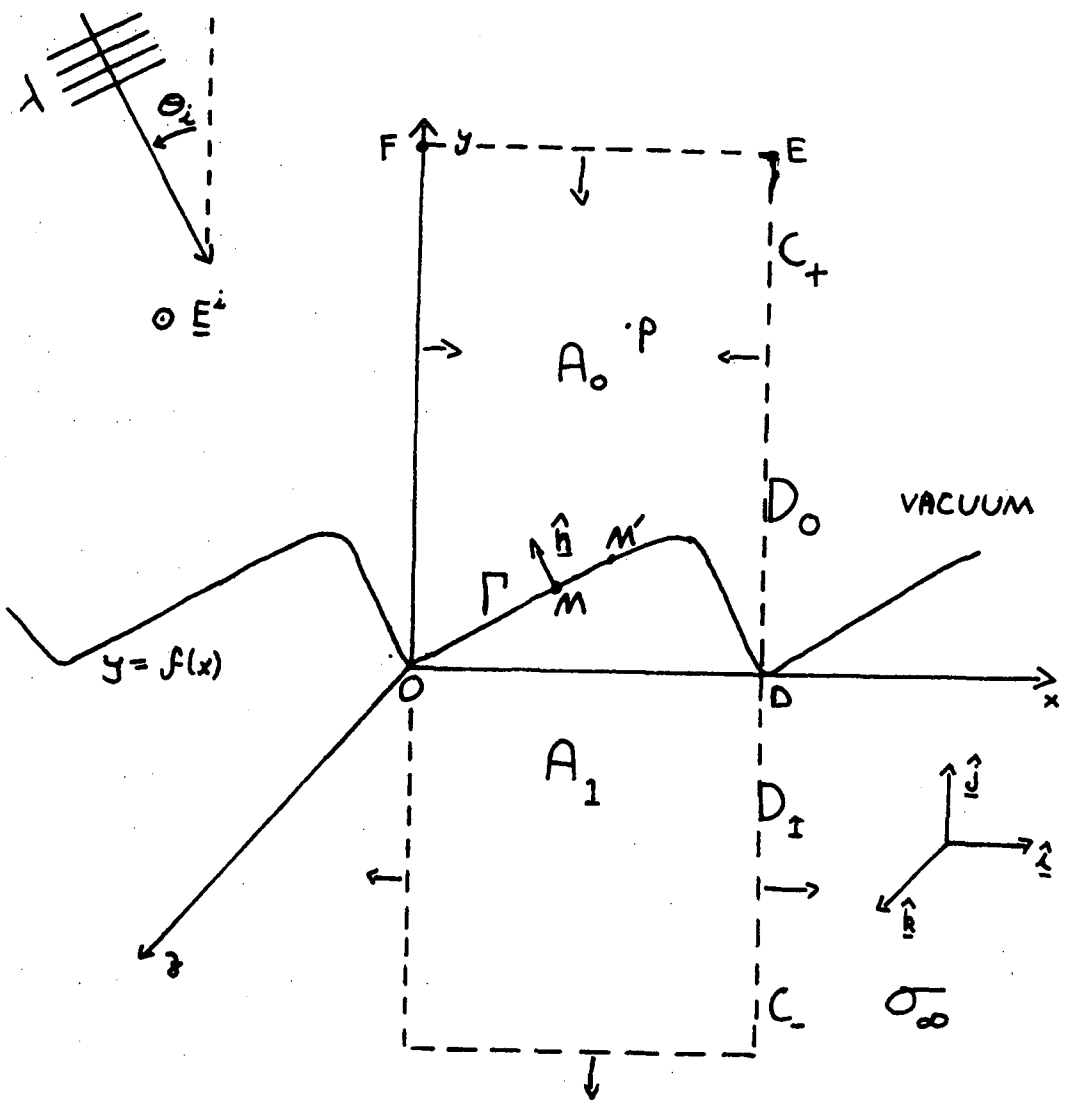


Figure 2.3.1 The geometry of the diffracting arrangement. The region D_0 represents free space while the domain D_1 represents the perfectly conducting grating.

and also the condition of outward going waves. Now for some point M on the profile Γ, the boundary condition appropriate to this particular polarization is

$$E^d(M) = - E^i(M).$$

In order to solve the diffraction problem, Petit then introduced a scalar function u defined by

$$\left.\begin{array}{lll} \text{(i)} \quad u(P) = E^d(P) & \text{if} & P \in D_0, \\ \text{(ii)} \quad u(P) = - E^i(P) & \text{if} & P \in D_1. \end{array}\right\} \qquad (2.3.5)$$

From this definition, it can immediately be seen that u is a function not only obeying the outgoing wave condition but also continuity across the boundary Γ. Although the continuity of u has been defined this does not constrain its normal derivative at the surface Γ to have the same property. (Of course this must also be expected from physical arguments, which are discussed later, since the discontinuity of $\partial u/\partial n$ is a measure of the surface current density. It is this discontinuity which enables the whole diffraction problem to be characterized uniquely.)

Having defined the problem, it only remains to choose the mechanism of the solution, Petit chose to convert the differential definition to the solution of an integral equation. He had to invoke Green's Theorem which involved the selection of an appropriate Green's function. The chosen function obeys the inhomogeneous Helmholtz equation:

$$(\nabla^2+k^2)\, G_0(P;M') = \delta(y-y') \sum_{n=-\infty}^{\infty} \delta(x-x'-nd)\, \exp(+i\alpha_0 nd)$$

where $P \equiv (x,y)$ and $M' \equiv (x',y')$. (2.3.6)

The right side of this expression embodies the pseudo-periodicity of the total field, a feature whose significance will be noted later. The solution of this differential equation is given by

$$G_0(P;M') = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\chi_n} \exp[i\alpha_n(x-x') + i\chi_n|y-y'|] \tag{2.3.7}$$

$$\left.\begin{aligned} \text{where } \alpha_n &= nK + \alpha_0, \\ \chi_n &= \sqrt{(k^2 - \alpha_n^2)}\ , \\ \text{and } K &= 2\pi/d. \end{aligned}\right\} \tag{2.3.8}$$

(The sign of the imaginary part of χ_n is chosen to be positive to ensure field convergence as $|y-y'| \to \infty$.) A second solution of equation (2.3.6) having the form

$$\frac{1}{2id} \sum_n \frac{1}{\chi_n} \exp[i\alpha_n(x-x') - i\chi_n|y-y'|]$$

also exists, but is rejected on physical grounds since it is a generator of inward going waves.

Now applying the two dimensional form of Green's Theorem to the region A_0 bounded by the contour C_+, it follows that

$$\iint_{A_0} [\nabla^2 G_0(P;M')\ u(M') - G_0(P;M')\ \nabla^2\ u(M')]dA'$$

$$= -\int_{C_+} [\frac{\partial G_0}{\partial n'}(P;M')\ u(M') - G_0(P;M')\ \frac{\partial u(M')}{\partial n'}]\ ds' .$$

Since there exists only one singularity within A_0, this reduces to the form

$$u(P) = \int_{C_+} [G_0 \frac{\partial u}{\partial n'} - \frac{\partial G_0}{\partial n'} u]\ ds' .$$

The pseudo-periodicity of both G_0 and u means that contributions from the line integrals along the linear segments OF and DE cancel with one another. Furthermore, the line integral along the segment FE tends to zero as the ordinate tends to infinity, a consequence of the radiation conditions.

Thus,

$$u(P) = \int_{\Gamma_+} [G_0 \frac{\partial u}{\partial n'} - \frac{\partial G_0}{\partial n'} u]\ ds'. \qquad (2.3.9)$$

A further application of Green's Theorem, this time to the region A_1 leads to

$$0 = \int_{\Gamma_-} [G_0 \frac{\partial u}{\partial n'} - \frac{\partial G_0}{\partial n'} u]\ ds'. \qquad (2.3.10)$$

(The left hand side of this reduces to zero because of the absence of a singularity within the region of integration. Once again, pseudo-periodicity and the radiation conditions simplify the contour integral

to a term involving only the grating profile.)

By now subtracting equations (2.3.9-10) from one another, noting that $\partial G_0/\partial n'$ is almost everywhere continuous on Γ and utilizing the defined continuity of u there, it may be shown that

$$u(P) = \int_{\Gamma} G_0(P;M')\ \xi(M')\ ds' \qquad (2.3.11)$$

$$\text{where } \xi(M') = \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_+} - \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_-} . \qquad (2.3.12)$$

(At this stage, it is appropriate to attach some physical significance to the rather abstract unknown, ξ. From its definition it follows that

$$\xi = \left.\frac{\partial E}{\partial n}\right|_{\Gamma_+} .$$

Now, since

$$\underline{n} \times \underline{H}\Big|_{\Gamma_+} = \underline{j}_s = j_s\, \hat{\underline{k}}$$

where $\underline{j}_s$ is the current flowing across the surface of the perfectly conducting metal, it follows that

$$\xi = -i\,\omega\,\mu_0\, j_s$$

i.e. ξ is analogous to the surface current.)

Now, to completely solve the diffraction problem, it is necessary to construct an integral equation constraining the behaviour of ξ.

This is done by applying the boundary condition

$$u(M) = - E^i(M)$$

for any point M lying on the profile Γ.

Thus

$$- E^i(M) = \int_\Gamma G_0(M;M')\ \xi(M')\ ds', \tag{2.3.13}$$

a Fredholm integral equation of the first kind. Its solution may be achieved using the "points-matching" technique, whereby the unknown is sampled at a discrete set of points. The above integral equation is then converted to a set of linear equations which may be solved using standard methods in numerical analysis (e.g. Gauss-Jordan elimination).

The solution outlined above relies upon the application of the boundary condition $\underline{n} \times \underline{E} = 0$. However, there exists a further method of solution using the boundary condition $\underline{n} \times \underline{H} = \underline{j}_s$ as its fundamental constraint. In the discussion below, attributed to Pavageau and Bousquet [2.2], the surface normal derivative of u is considered. Here,

$$\underline{n}.\nabla u(M) = \int_\Gamma \frac{\partial G_0(M;M')}{\partial n}\ \xi(M')\ ds'. \tag{2.3.14}$$

However, the normal derivative of u has a discontinuity of magnitude ξ at the profile surface and thus some care is needed in the evaluation of $\underline{n}.\nabla u$. From a theorem in Fourier analysis, it is known that the Fourier series of a discontinuous function converges to the arithmetic

mean of limits of the function on either side of the discontinuity -

$$\text{i.e. } \underline{n}.\nabla u = \tfrac{1}{2}\left[\left.\frac{\partial u}{\partial n}\right|_{\Gamma_+} + \left.\frac{\partial u}{\partial n}\right|_{\Gamma_-}\right]$$

$$= \left.\frac{\partial u}{\partial n}\right|_{\Gamma_+} - \tfrac{1}{2}\xi \ .$$

$$\text{Thus, } \left.\frac{\partial u}{\partial n}\right|_{\Gamma_+} = \tfrac{1}{2}\xi + \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n}\, \xi(M')\, ds'$$

$$\text{Now, } \underline{n} \times \underline{H} = \underline{j}_s$$

implies that

$$\left.\frac{\partial u}{\partial n}\right|_{\Gamma_+} = \xi - \frac{\partial E^i}{\partial n} \qquad (2.3.15)$$

and so one arrives at

$$\tfrac{1}{2}\xi(M) = \frac{\partial E^i(M)}{\partial n} + \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n}\, \xi(M')\, ds' \qquad (2.3.16)$$

- a Fredholm integral equation of the second kind. Pavageau and Bousquet proposed that this equation be solved iteratively. However, later workers found this method to be unreliable and resorted to the use of the points matching technique.

In our implementation of the theories, it is the usual practice to distribute the sampling points according to a multi-segment Gauss-Legendre rule. This method is advantageous since it enables a reasonable characterization of the unknown in regions of singular behaviour (e.g. at convex

edges) - a consequence of the clustering of the point distribution at either end of the segment.

Following the solution for ξ, it is a simple task to reconstruct the diffracted field amplitudes according to the expression:

$$u_n = \frac{1}{2id\chi_n} \int_\Gamma \phi(M') \exp[-i(\alpha_n x' + \chi_n f(x'))] \, ds'.$$

Here, the diffracted field is given by the Rayleigh expansion

$$u(P) = \sum_{n=-\infty}^{\infty} u_n \exp[i(\alpha_n x + \chi_n y)],$$

a representation appropriate to all points located above $y = \max(f(x))$.

2.3.2 Formulations for S polarized radiation

Let us now turn to a discussion of three formalisms describing the diffraction of an S polarized plane wave by a perfectly conducting reflection grating. Herein, it is proposed to review the works of Petit [2.1], Pavageau and Bousquet [2.2] and Bolomey [2.13].

For this polarization, the incident field is specified by

$$\underline{H}^i = \exp[i(\alpha_0 x - \chi_0 y)]\hat{\underline{k}}. \qquad (2.3.17)$$

This field excites a surface current traversing the grooves of the grating (with no component along the Oz axis) which in turn radiates an S polarized scattered field. This scattered field, H^d, is given by the difference between the total field and the incident field for all points

in free space

$$\text{i.e.} \quad H^d(P) = H(P) - H^i(P). \tag{2.3.18}$$

Once again this field obeys the Helmholtz equation and the radiation conditions. For this polarization the appropriate boundary conditions are

$$H\Big|_{\Gamma_+} = j_s \tag{2.3.19a}$$

and $\underline{n} \times \underline{E} = 0.$

This latter constraint may be reduced to the form

$$\frac{\partial H}{\partial n} = 0. \tag{2.3.19b}$$

by using Maxwell's equations.

In the original solution of this diffraction problem Petit defined a function u given by

$$\begin{aligned} &\text{(i)} \quad u(P) = H^d(P) \qquad \text{if} \qquad P \in D_0 \\ &\text{(ii)} \quad u(P) = -H^i(P) \qquad \text{if} \qquad P \in D_1. \end{aligned} \tag{2.3.20}$$

It is immediately evident that u has a discontinuity of magnitude j_s at the profile surface, whereas its normal derivative is continuous there. By pursuing an analysis similar to that described in the section dealing with P polarization formalisms it may be shown that

$$u(P) = \int_{\Gamma_+} \left[G_0 \frac{\partial u}{\partial n'} - \frac{\partial G_0}{\partial n'} u\right] ds'$$

and $$0 = \int_{\Gamma_-} \left[G_0 \frac{\partial u}{\partial n'} - \frac{\partial G_0}{\partial n'} u\right] ds'$$

for any point P in D_0. Subtracting these two expressions reveals that

$$u(P) = -\int_{\Gamma} \frac{\partial G_0}{\partial n'} (P;M') \, \xi(M') \, ds' \tag{2.3.21}$$

upon noting the continuity of $\partial u/\partial n$ at Γ.

Here, $$\xi(M) = u(M)\Big|_{\Gamma_+} - u(M)\Big|_{\Gamma_-} \tag{2.3.22}$$

$$= H(M)\Big|_{\Gamma_+}$$

$$= j_s(M).$$

In his original solution, Petit attempted to solve the diffraction problem by applying the boundary condition

$$\frac{\partial H}{\partial n} = 0$$

which resulted in an integral equation of the form

$$\frac{\partial H^i}{\partial n} = \int_{\Gamma_+} \frac{\partial^2 G_0}{\partial n \partial n'} (M;M') \, \xi(M') \, ds' \, . \tag{2.3.23}$$

However, the associated numerical implementation was plagued with difficulties caused by the divergent nature of integral kernel. Petit stated that he was forced to introduce a "renormalization" which he could not entirely justify.

It was not until the efforts of Pavageau and Bousquet [2.2] that a satisfactory solution to this diffraction problem was found. These authors applied the boundary condition

$$H\Big|_{\Gamma_+} = \xi$$

and were able to derive a stable Fredholm integral equation of the second kind whose kernel was everywhere convergent. In their derivation, they considered a field point M placed on Γ and so

$$u(M) = -\int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n} \xi(M')\, ds'.$$

However u is a discontinuous function and so its Fourier series tends to the arithmetic mean of the values on either side of the surface of discontinuity -

$$\text{i.e.} \quad u(M) = \tfrac{1}{2}[u(M)\Big|_{\Gamma_+} + u(M)\Big|_{\Gamma_-}]$$

$$= u(M)\Big|_{\Gamma_+} - \tfrac{1}{2}\,\xi\,.$$

$$\text{Thus,} H(M)\Big|_{\Gamma_+} = H^i(M) + \tfrac{1}{2}\,\xi(M) - \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n'} \xi(M')\, ds'$$

which gives rise to the final result

$$\tfrac{1}{2}\,\xi(M) \;=\; H^{i}(M) \;-\; \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n'}\,\xi(M')\,ds'. \qquad (2.3.24)$$

Let us turn now to a third solution of this problem. A discussion of this method, attributed to Bolomey [2.13], is included, not so much because of the use made of its numerical implementation, but because of many of the concepts involved in its derivation are of relevance to the finite conductivity theory reviewed in the following chapter.

Bolomey defined an unknown u by the following criteria:

(i) $u(P) = H^{d}(P)$ if $P \in D_0$;

(ii) $(\nabla^2 + k^2)\,u = 0$ elsewhere;

(iii) u is everywhere continuous and obeys the radiation conditions.

From these definitions it may be seen that u possesses many of the characteristics of that same function defined by Petit in his treatment of the P polarization problem. However, in the case of S polarized radiation, u represents a physical field quantity only in the space D_0.

By following an analysis similar to that discussed in the treatment of the P polarization problem, one arrives at an expansion for u (at any point in the region D_0) as a function of its derivative discontinuity on Γ.

$$\text{i.e.} \qquad u(P) = \int_{\Gamma} G_0(P;M')\, \xi(M')\, ds',$$

$$\text{where} \quad \xi(M') = \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_+} - \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_-} .$$

$$\text{Thus} \quad \left.\frac{\partial u(M)}{\partial n}\right|_{\Gamma_+} = \tfrac{1}{2}\, \xi(M) + \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n}\, \xi(M')\, ds'.$$

$$\text{Now} \quad \left.\frac{\partial u}{\partial n}\right|_{\Gamma_+} = \left.\frac{\partial H}{\partial n}\right|_{\Gamma_+} - \frac{\partial H_i}{\partial n} ,$$

and since $\partial H/\partial n$ is continuous across the interface and vanishes in the metal, it follows that

$$0 = \frac{\partial H^i(M)}{\partial n} + \tfrac{1}{2}\, \xi(M) + \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n}\, \xi(M')\, ds'. \qquad (2.3.25)$$

Equation (2.3.25) is a Fredholm integral equation of the second kind, solved by the points matching technique.

2.4 IMPLEMENTATIONS OF THE THEORY

The theories reviewed in the previous section have been implemented for use on large computer systems such as the Burroughs B6700 (University of Tasmania) and Control Data Cyber 76 (D.C.R., C.S.I.R.O.), and have produced results which are in excellent agreement with the constraints imposed by conservation of energy, the Reciprocity Theorem, and the boundary conditions. In most cases, only twenty sampling points per period are needed to give convergence to better than one percent, although

in some isolated examples thirty points are necessary to produce the required accuracy.

During previous years, members of our group have performed extensive studies (concerned with blaze optimization and anomaly properties) with the aid of these programs. In the course of this work, the generality of the algorithms was displayed. It was found that the programs could cope with wide ranges of groove depths (up to 1.3 × grating period) and wavelengths ($0.2 \leq \lambda/d \leq 2.0$) with relative ease.

2.5 FIRST, SECOND AND THIRD ORDER BLAZES OF DIFFRACTION GRATINGS

This section, which is an application of the infinite conductivity theories, is the result of a study conducted by the author in collaboration with Dr. Ian Wilson (University of Tasmania) and Dr. Ross McPhedran (University of Sydney). The material presented herein has been reported previously in the University of Tasmania Research Report DGRG76/1 [2.14] and in a paper [2.15] published recently by the Journal of Optics.

2.5.1 Introductory Remarks

Only recently has it become possible for designers of grating instruments to choose diffraction gratings of optimal performance for their particular application. In their choice, they can be guided by the results of blaze optimization studies [2.3-9] made using rigorous electromagnetic theories.

Maystre and Petit [2.3] have given a grid of efficiency curves for gratings of triangular groove profile having blaze angles ranging from

5^{o} to 45^{o} and used with first-order angular deviations of 0^{o}, $+ 15^{o}$ and $- 15^{o}$. Their curves display clearly the change in polarization by gratings with blaze angle, but unfortunately the concentration of points on some curves is locally insufficient, with the result that important anomalous effects are hidden.

McPhedran and Waterworth [2.4], in considering the optimal choice of apex angle for a triangular groove profile, advocated the use of a 90^{o} apex angle to give best blaze properties.

McPhedran, Wilson and Waterworth [2.5] have described an optimization of the groove depth of sinusoidal profile holographic gratings. They found the optimal depth to be 0.37 times the period. Similarly, Maystre and Petit [2.6-7] investigated and optimized the profile of holographic gratings of the type formed in logarithmic negative photoresists. The blaze of this class of gratings is clearly inferior to that of the optimized sinusoidal gratings.

Maystre et al [2.8] have given theoretical results concerning the efficiency of 3600 line/mm aluminium gratings used in the Littrow mounting in the near ultraviolet. Both echelette and sinusoidal profiles were treated. While their curves display well the form of the diffraction anomalies corresponding to the Littrow mounting, they give no information concerning the effect of angular deviation upon them.

McPhedran and Wilson [2.9] have reported on the application of the finite conductivity theory of D. Maystre [2.11] to grating optimization studies for the visible spectrum. They showed that the blaze optima predicted using the infinite conductivity theory [2.4-5] could be extended to the optical region. This investigation referred to gratings

having triangular, sinusoidal and distorted sinusoidal groove profiles.

Of these studies, all but one [2.3] have been confined to the Littrow mounting, and without exception only first order blazes have been discussed. This section reports the results of an extensive study into the effects of angular deviation on first-order grating performances. Second and third order blaze effects for the Littrow mount are considered for three important profile forms. *

The theoretical results quoted below have all been obtained using a formulation [2.2] in which the grating surface is taken to have infinite conductivity. Thus, the efficiency curves shown are of great generality in that they can be rescaled to cover a wide range of grating periods. The conclusions drawn here will be applicable from the visible to the millimetric-wave regions. (While it is true that the infinite conductivity theory may break down in the visible region if one is interested in calculating the precise form of anomalies, all previous investigations have shown that it is entirely adequate in studies where interest

---

* After the preparation of this chapter an important additional reference [2.29] came to the notice of the author. Loewen et al have extended some aspects of this investigation using Maystre's finite conductivity theory, and thus have indeed confirmed that our results obtained using the theory of Pavageau and Bousquet are valid for metallic gratings in the visible. It is also interesting to note that in reference [2.29] a comparison is made between experimental measurements and our curves for the first, second and third order blazes of the $26^{o}45'$ profile. This comparison bears out the accuracy of our predictions for this interesting grating.

is centred on the blaze properties of gratings in the visible.) However, instrument designers are cautioned against using the results in the ultra-violet region, where finite conductivity effects become very important [2.8, 2.16].

2.5.2 Angular deviation and blaze

Let us now consider the effect of angular deviation on first order efficiencies for triangular profile gratings. This effect is highly dependent on the polarization of the incident radiation. We will refer to P (or $E_{||}$) polarization when the electric field vector of the incident wave is aligned parallel to the direction of the grating ruling. The orthogonal polarization will be referred to as S (or $H_{||}$) polarization. Unpolarized radiation efficiencies (U) are the average of S and P efficiencies.

The study of the effect of angular deviation is simplified by the application of the Reciprocity Theorem. This theorem states that the efficiency curve as a function of wavelength for a grating used in order m with an angular deviation of D, is identical to the curve for the grating used in the same order with deviation - D [2.17]. The identity is unaffected by finite conductivity effects [2.18]. (Hence it is only $|D|$ which is important and throughout this work increasing or decreasing angular deviations will refer to $|D|$.)

For P polarization, angular deviation generally causes a decrease in peak efficiency, and a shift of the efficiency curve towards shorter wavelengths. If the P blaze peak is located near a Wood anomaly, then the position of the peak may well be an irregular function of wavelength. Such an anomaly is always associated with a small local increase of

efficiency; thus its movement and fluctuation in strength with angular deviation can control the behaviour of the blaze peak.

The situation for S polarized radiation is more complicated. The blaze peak is little affected in strength and position by variations of angular deviation. However, the instrument designer may be concerned with the changes in grating anomalies with deviation. S polarization anomalies are far stronger than P polarization anomalies. When they occur within the blaze region, they can give rise to rapid changes in the instrumental profile which are troublesome to users. The anomalies move with angular deviation according to a well known equation due to Lord Rayleigh [2.19]. In addition, their strength can be a function of angular deviation. Thus the control of this variable can provide a means of reducing (and even eliminating) the undesirable effects of these anomalies.

The complexity of the deviation effects necessitates their discussion on the basis of the behaviour for particular blaze angles. Discussed below are results obtained for four commonly used blaze angles ($10^{o}22'$, $21^{o}06'$, $26^{o}45'$ and $36^{o}52'$). The behaviour of gratings having intermediate blaze angles can be inferred by using the results presented here. All the gratings were chosen to have a 90^{o} apex angle, in order to provide best performance in the Littrow mounting [2.4, 2.20].

2.5.2.1 *The blaze angle $10^{o}22'$*

Figure (2.5.1) shows a grid of efficiency curves for the $10^{o}22'$ triangular grating profile, giving both S and P polarization behaviour for angular deviations between 0^{o} and -45^{o}. The radiation is incident

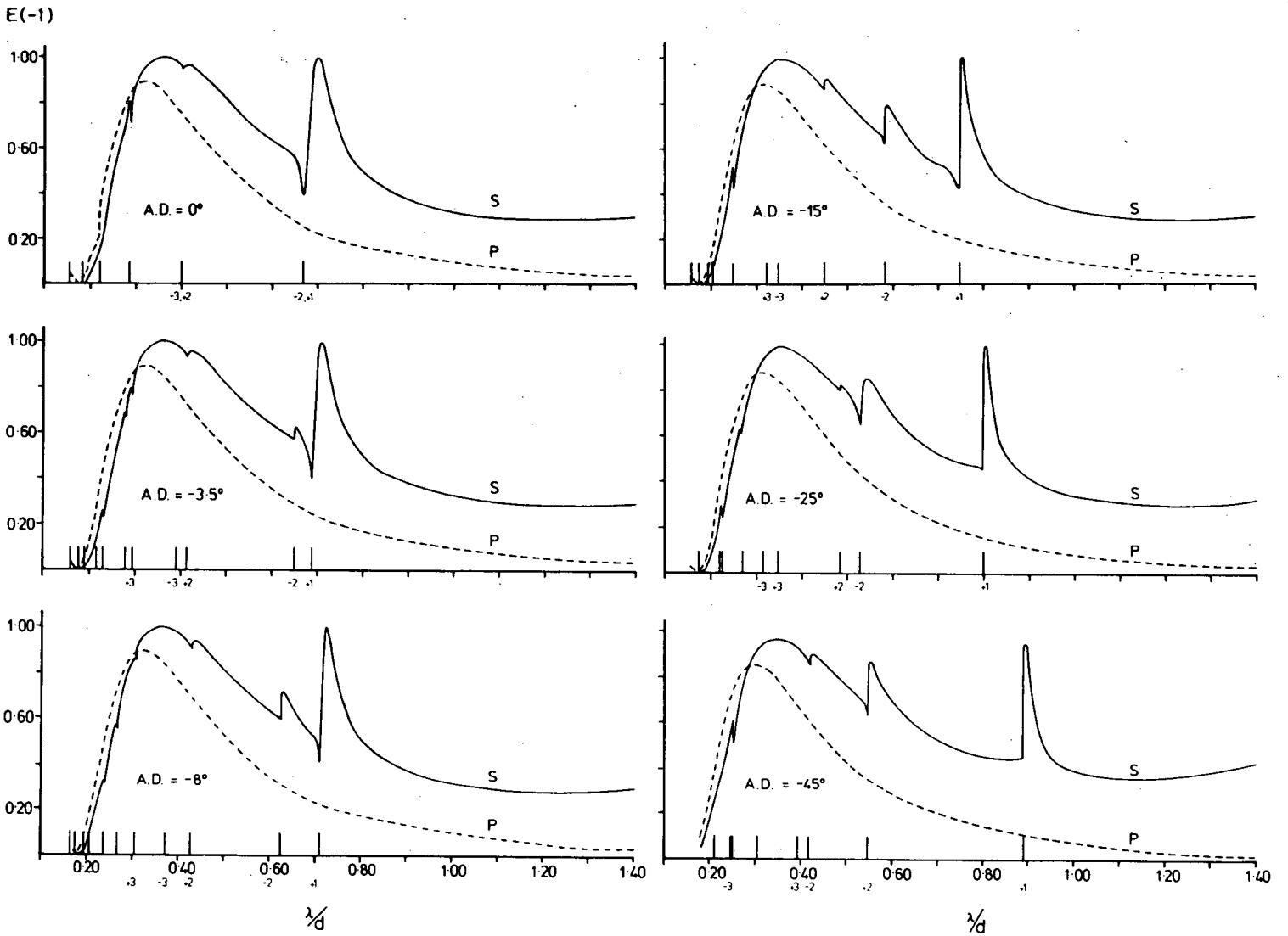


Figure 2.5.1 Blaze angle $10^{o}22'$ in constant angular deviation mounts.

on the profile on the side of the longer groove facet. The performance of the grating in unpolarized radiation and an estimate of the degree of polarization of the radiation diffracted by the grating can be readily obtained from these curves.

For this shallow blaze angle, P polarization efficiencies are only slightly affected by a variation in angular deviation. The decrease of D from 0^{o} to -45^{o} causes a drop in the maximum P efficiency of only 4% (from 89% to 85%). The wavelength shift of the position of maximum efficiency is correspondingly slight (from $\lambda/d = 0.320$ to 0.306).

The S blaze peak is also little affected by increasing the angular deviation. The increase of D from 0^{o} to -45^{o} causes a drop in blaze efficiency from 100% to 97% and a shift in the blaze wavelength from $\lambda/d = 0.36$ to 0.35.

In the Littrow configuration, the diffraction anomaly associated with the passing off of orders - 2 and + 1 gives rise to a strong resonance peak situated at the upper wavelength end of the blaze region. As D departs from zero the - 2 and + 1 anomalies separate and move to shorter and longer wavelengths respectively. Whether the two weaker anomalies are more or less of an annoyance than the single strong Littrow anomaly will depend on the particular application to which the grating is being put.

For an angular deviation of -25^{o}, a series of calculations were performed for a grating having the same blaze angle ($10^{o}22'$) but groove apex angles varying from 85^{o} to 110^{o}. No anomaly reduction was achieved by this profile variation and it was seen that the Littrow optimum angle of 90^{o} had not been altered by the introduction of angular deviation.

2.5.2.2 *The blaze angle of 21°06'*

In figure (2.5.2) is presented a set of efficiency curves for the 21°06' triangular profile grating, showing its behaviour for both S and P polarizations and angular deviations ranging from 0° to - 45°.

As for the previous blaze angle, P polarization behaviour is rather insensitive to variation in angular deviation. The change in D from 0° to - 45° causes a drop in the maximum P efficiency of 6% (from 82% to 76%) and a λ/d shift in its location from 0.58 to 0.56.

The same insensitivity to the variation of deviation is shown by the S polarization blaze peak. By changing D from 0° to - 45°, a 3% drop in blaze efficiency (from 100% to 97%) and a slight λ/d shift to the red (from 0.72 to 0.76) occur.

This blaze angle occurs towards the upper end of a range of angles in which diffraction anomalies are weak and hence have little effect on instrumental profiles. The region is centred on the blaze (19°30') for which the (- 2, + 1) Wood anomaly coincides with the blaze wavelength [2.16], and extends from 17° to 22°.

Anomalies associated with this groove profile are unlikely to trouble instrument users, whatever their choice of angular deviation in the range from 0° to - 45°.

2.5.2.3 *The blaze angle of 26°45'*

The grid of efficiency curves for this profile is given in figure (2.5.3). It is evident from the figure that angular deviation has quite a significant effect on the P polarization performance of this grating. In the Littrow mount, the peak P efficiency of 95% occurs at $\lambda/d = 0.67$, just after the (- 2, + 1) Wood anomaly. (The presence of

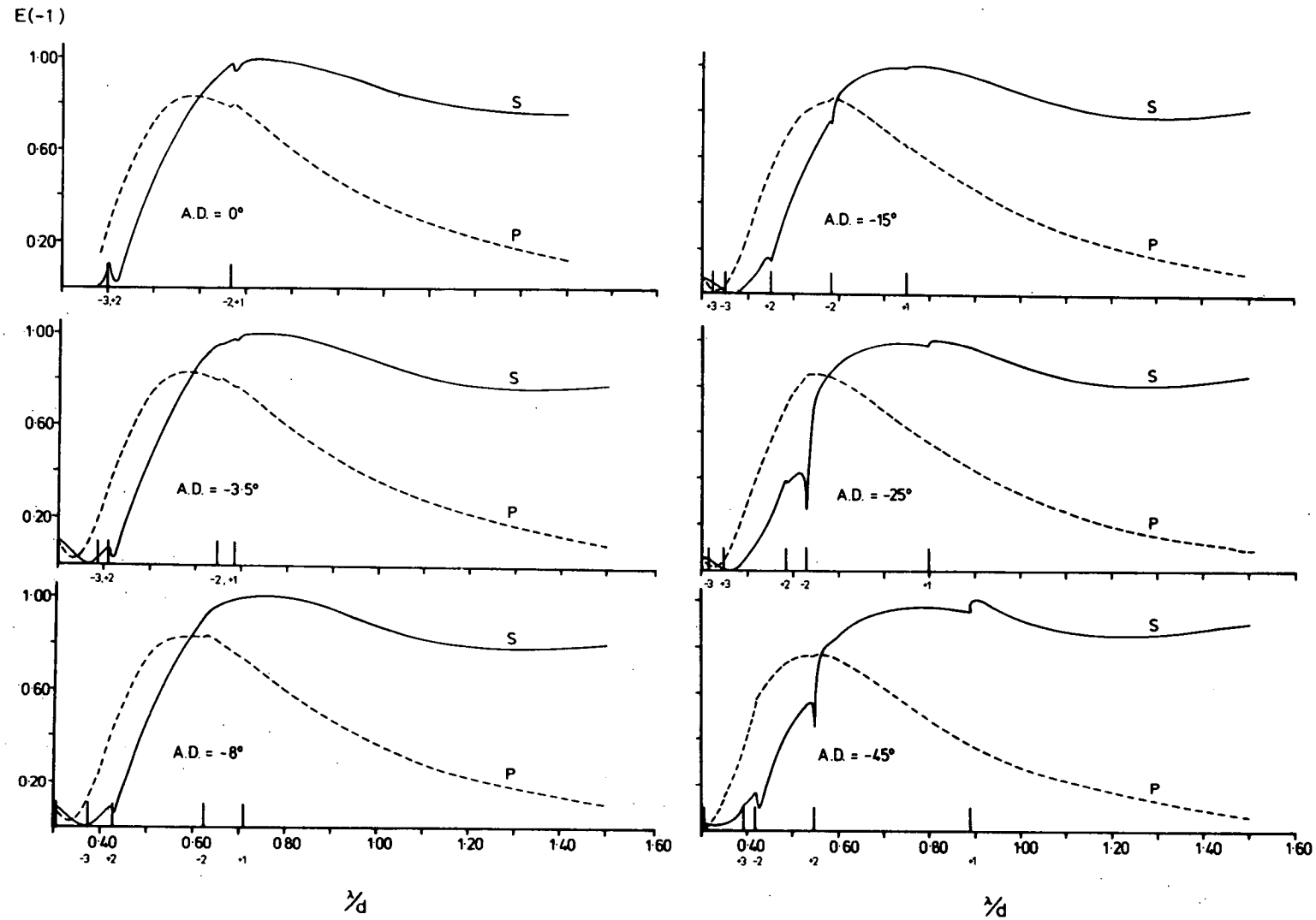


Figure 2.5.2 Blaze angle $21^{o}06'$ in constant angular deviation mounts.

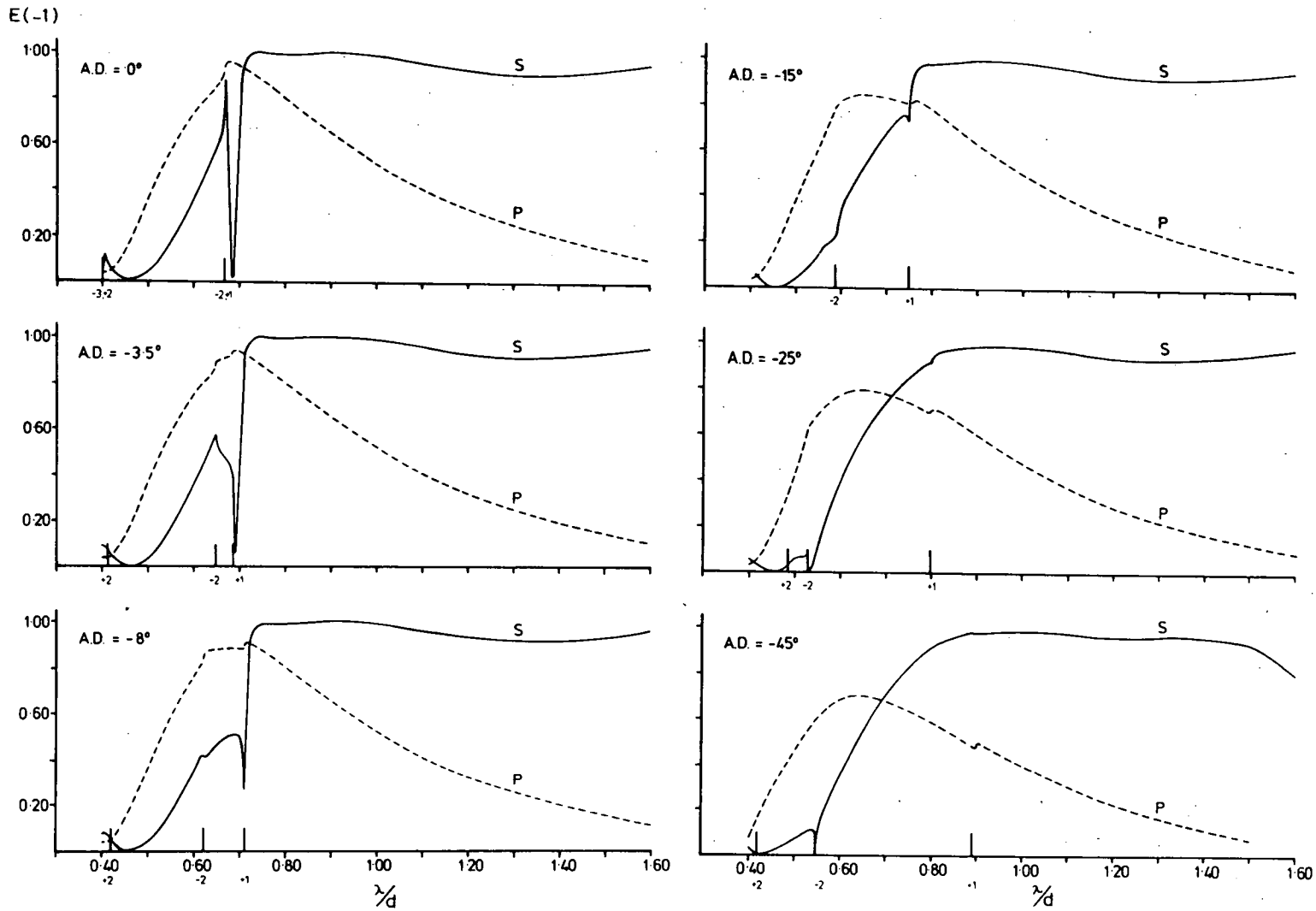


Figure 2.5.3 Blaze angle $26^{o}45'$ in constant angular deviation mounts.

this anomaly is in fact causing a significant increase in the maximum efficiency.) For an angular deviation of - 45^{o}, the peak efficiency of just 70% occurs at $\lambda/d = 0.63$.

The S polarization blaze is strong and wide, and is little affected by angular deviation. The Littrow peak of 100% occurs at $\lambda/d = 0.90$. For D = - 45^{o}, the maximum value of 99% occurs when λ/d is close to 1.00.

The most prominent feature of the S polarization Littrow efficiency curve is the very strong resonance minimum associated with the (- 2, + 1) Wood anomaly. This lies at the same wavelength as the P blaze peak, and hence would give rise to strong and rapid variations of U efficiency and polarizance if the Littrow mount was chosen for this grating. By increasing the deviation it is possible to separate and weaken the - 2 and + 1 anomalies. If an angular deviation of - 15^{o} is selected, then an S polarization blaze which is little disturbed by anomalies and a P blaze which is only 10% weaker than in the Littrow mount, are obtained. From the point of view of blaze performance, this value of angular deviation is clearly optimal for the $26^{o}45'$ grating.

Also, the effect of angular deviation on the behaviour of this grating, when used with radiation incident on the off-blaze facet of the groove, has been investigated. When only two real diffracted orders propagate, it has been established [2.21] using the Reciprocity Theorem that grating efficiencies are unchanged by moving from the on-blaze to the off-blaze side of the groove. This property holds only for the zero-order efficiencies when more than two orders propagate.

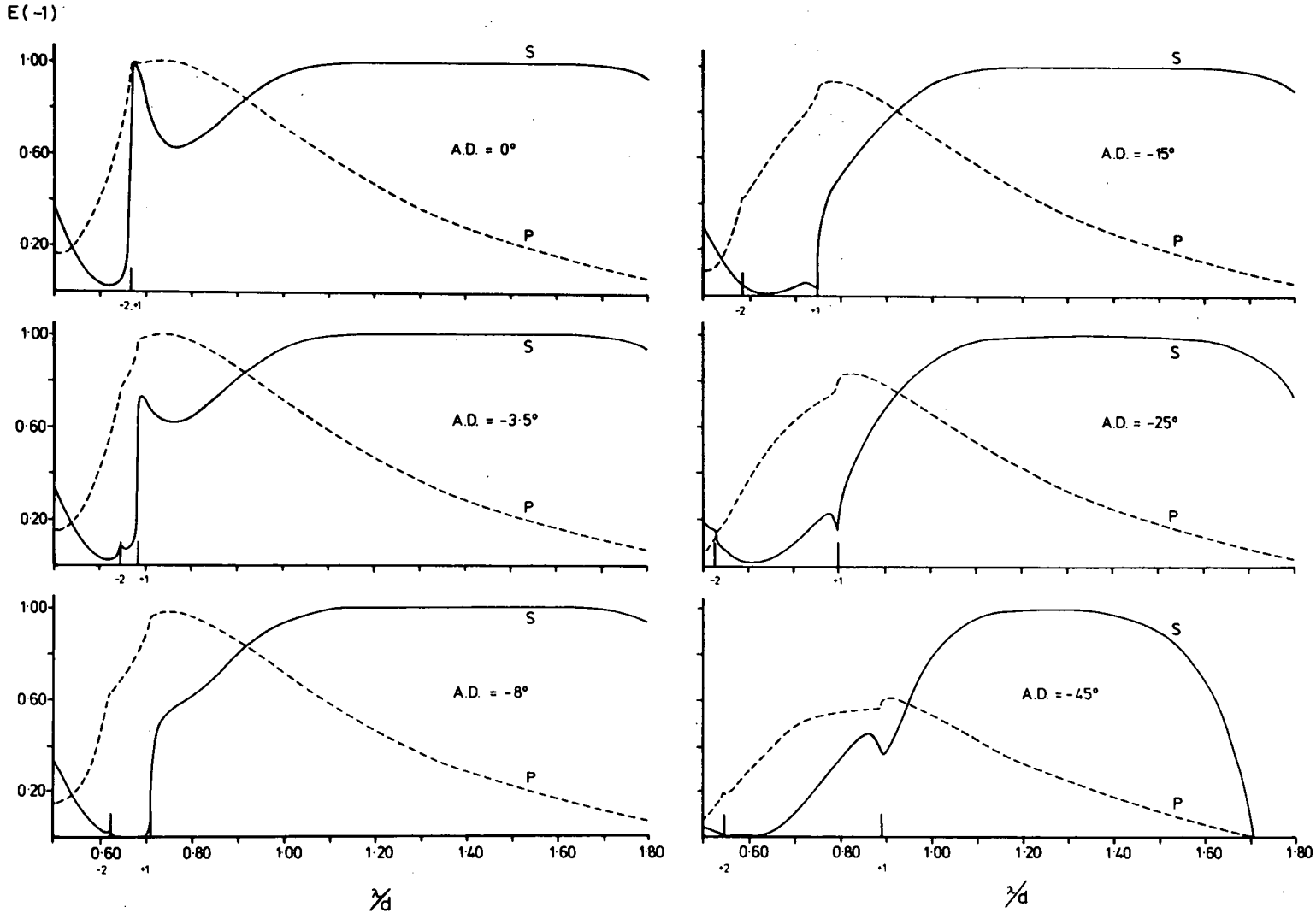


Figure 2.5.4 Blaze angle $36^{\circ}52'$ in constant angular deviation mounts.

Our results for this profile show that its use on the off-blaze side ensures the presence of a very strong S polarization Wood anomaly, irrespective of the value of angular deviation. This strong anomaly marks the transition from the off-blaze region of the curve to the region in which it must follow the on-blaze behaviour. In P polarized radiation, the weakening of the blaze with increasing deviation is rather more severe on the off-blaze than on the on-blaze side. (In fact for an angular deviation of -45^{o}, the peak P efficiency is only 49% for the off-blaze mounting, which is to be compared with 70% for the on-blaze side.)

It is decidedly advantageous to use the grating with incidence on its blazed side.

2.5.2.4 *The blaze angle of $36^{o}52'$*

The set of efficiency curves for this profile is shown in figure (2.5.4). The P polarization blaze of the grating is quite sensitive to the change of angular deviation. The Littrow peak efficiency of 100% occurs at $\lambda/d = 0.735$ whereas for $D = -45^{o}$ the maximum P efficiency is 60% when $\lambda/d = 0.85$. For this profile the efficiency maximum is always located on the long wavelength side of the + 1 Wood anomaly. This anomaly moves towards the red as D becomes more negative, taking the efficiency maximum with it.

Angular deviation also has an adverse effect on the S polarization blaze of the grating. While the efficiency maximum is maintained at 100%, the width of the maximum diminishes as D increases from 0^{o}.

If it is desired to have an anomaly-free instrument profile when using this grating, the angular deviation should have its magnitude in

the range 8^{o} to 15^{o}. The diffracted radiation is highly polarized throughout much of the blaze region. The S blaze which is formed for this blaze angle (and for all angles up to 45^{o}) is of a form well adapted for exploitation in laser cavity and mirror applications using the Littrow mount, as has been pointed out by Maystre and Petit [2.3]. Its great width would permit tuning over a wide range of lasing transitions.

2.5.2.5 *Blaze wavelength and angular deviation*

In textbooks and handbooks on gratings, one often finds an expression relating the blaze wavelength ($\lambda_{B,m}$) for order m and the angular deviation D to the first order Littrow blaze wavelength (λ_{B}). The expression is

$$m\lambda_{B,m} = \lambda_{B} \cos (D/2). \qquad (2.5.1)$$

It is based on geometrical optics considerations, and consequently has doubtful validity in situations where polarization effects become important.

Our rigorous studies of gratings have shown that this formula is in general not correct. In the examples cited above, the P polarization blaze peak tended to move towards shorter wavelengths as D increased in magnitude. However, these movements were generally smaller than those predicted by the use of equation (2.5.1). The tendency is for S polarization blaze peaks to move towards longer (not shorter) wavelengths with increasing deviation. Since the S blaze is so much wider than the P blaze, the movement of the latter controls the location of the U

blaze peak, which consequently moves towards the blue with increasing |D|. These movements are smaller than those predicted by equation (2.5.1).

2.5.3 Angular deviation effects on the blaze of holographic gratings

Let us now turn to the study of angular deviation effects in connection with holographic gratings. If these gratings are formed in a linear photoresist using an appropriate exposure technique, the resultant groove profile is close to a sinusoid. Non-linearity of the photoresist can produce distortions of the profile away from the sinusoidal form, and it has been shown [2.22] that these profile distortions can be exploited to widen the Littrow blaze associated with sinusoidal profiles. We have thus chosen to study both sinusoidal gratings and those with groove shapes incorporating photoresist distortions.

2.5.3.1 *Sinusoidal gratings*

It has been shown [2.5] that the optimal groove depth for sinusoidal gratings used in the Littrow mount is 0.37 times the period. This depth gives the widest useful first-order blaze.

In figure (2.5.5) a grid of efficiency curves for this optimal profile for both P and S polarized radiation are given, showing the effects of varying the angular deviation between 0^{o} and -45^{o}. (Similar grids have been computed for a wide range of groove depths about 0.37 and the trends evident in figure (2.5.5) apply generally.)

It is apparent from figure (2.5.5) that as D increases, the P polarization blaze weakens and becomes narrower. The blaze peak is carried to long wavelengths by the movement of the + 1 Wood anomaly. The S polarization blaze is always associated with efficiencies rising

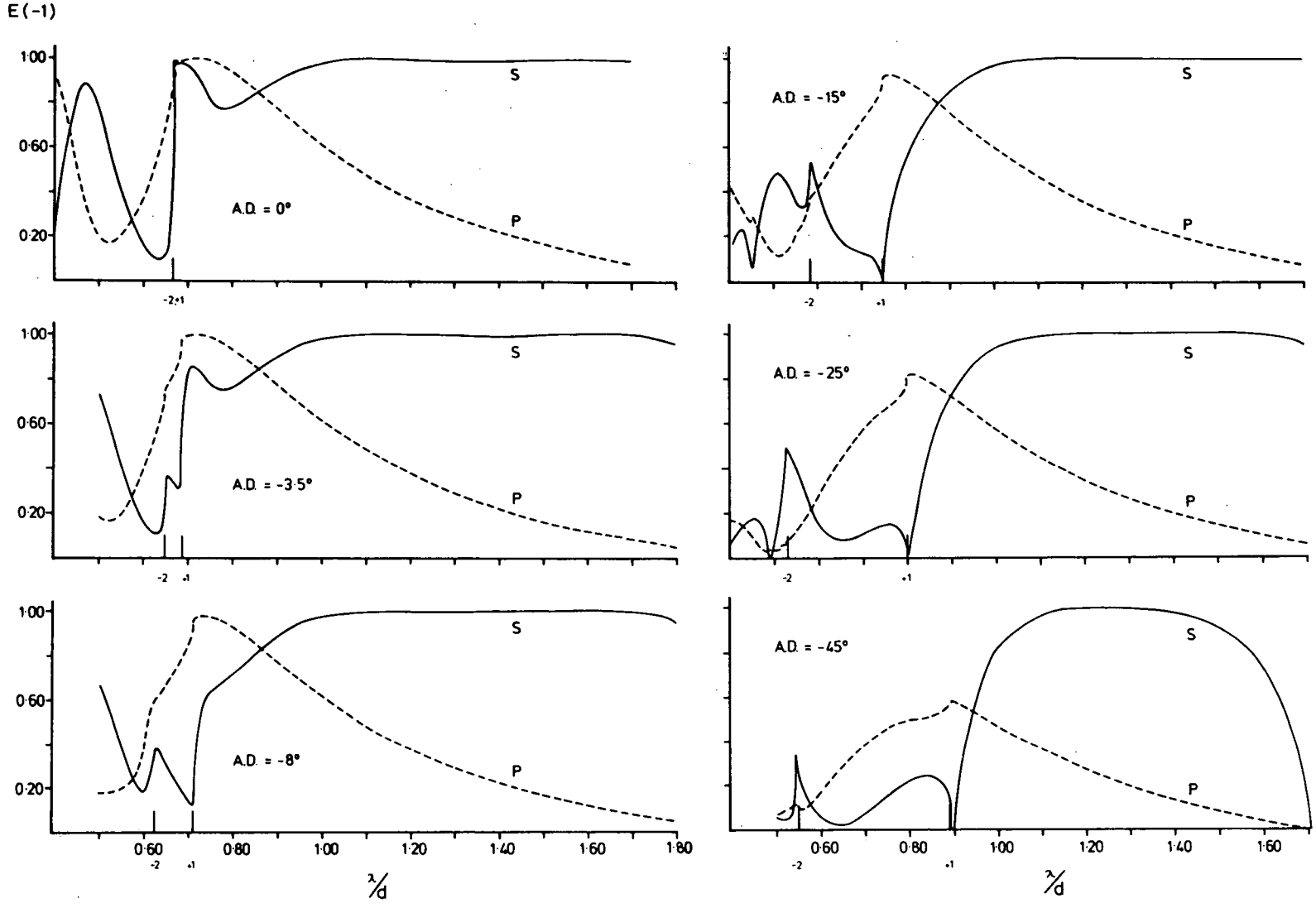


Figure 2.5.5 - Sinusoidal groove grating, h = 0.37 d, in constant angular deviation mounts.

to 100%, but with increasing deviation it diminishes in width.

Table (2.5.1) characterizes the deterioration of grating performance in P and U polarized radiation with increasing deviation. It is evident from the table that this grating is ill-suited to use with deviations much in excess of 10^{o}. The employment of larger deviations would result in a high polarization of the diffracted radiation and a need to operate the grating at relatively high angles of incidence.

The equivalent blaze angle (i.e. the Littrow incidence angle corresponding to the wavelength of peak U efficiency) of this grating in the Littrow mount is 20.5^{o} (for unpolarized radiation). Thus it is fair to compare its behaviour with that of the triangular profile grating of section (2.5.2.2). It can be seen that the triangular profile is much less sensitive to deleterious angular deviation effects than is the sinusoidal profile. While the former grating could be used in a wide variety of mountings, the latter should only be used near Littrow.

2.5.3.2 *Distorted sinusoidal profiles*

A set of efficiency curves is given in figure (2.5.6) referring to the P and S polarization behaviour associated with the optimum inverted distorted profile described by Wilson et al [2.22]. The effects of angular deviation on blaze for this profile are similar to those commented upon in section (2.5.3.1). It is to be noted that the P polarization blaze is more stable, both in strength and location, with increasing deviation than was the case with the purely sinusoidal profile. (This may be inferred from table (2.5.2), the analogue of table (2.5.1) for the optimum distorted profile.) The P blaze peak is always located above the + 1 Wood anomaly and is little affected by it.

TABLE 2.5.1

A.D.	P Polarization		Unpolarized		
	$(\lambda/d)_{max}$	E_{max}	$(\lambda/d)_{max}$	θ_{max}	E_{max}
0^{o}	0.72	1.00	0.70	20.5^{o}	0.996
-3.5^{o}	0.72	0.99	0.71	19.05^{o}	0.923
$-\ 8^{o}$	0.735	0.975	0.85	21.2^{o}	0.828
$-\ 15^{o}$	0.76	0.92	0.90	19.5^{o}	0.80
$-\ 25^{o}$	0.81	0.82	0.97	17.4^{o}	0.75
$-\ 45^{o}$	0.896	0.582	1.10	14.04^{o}	0.661

The effect of angular deviation on the order -1 performance of the optimum sinusoidal groove grating (h/d = 0.37).

TABLE 2.5.2

A.D.	P Polarization		Unpolarized		
	$(\lambda/d)_{max}$	E_{max}	$(\lambda/d)_{max}$	θ_{max}	E_{max}
0^{o}	0.94	1.00	0.92	27.39^{o}	0.996
-3.5^{o}	0.94	0.995	0.92	25.5^{o}	0.99
$-\ 8^{o}$	0.94	0.985	0.93	23.8^{o}	0.97
$-\ 15^{o}$	0.94	0.965	1.00	22.79^{o}	0.92
$-\ 25^{o}$	0.95	0.90	1.07	21.2^{o}	0.85
$-\ 45^{o}$	0.98	0.71	1.16	16.4^{o}	0.74

The effect of angular deviation on the order -1 performance of the optimum distorted groove grating.

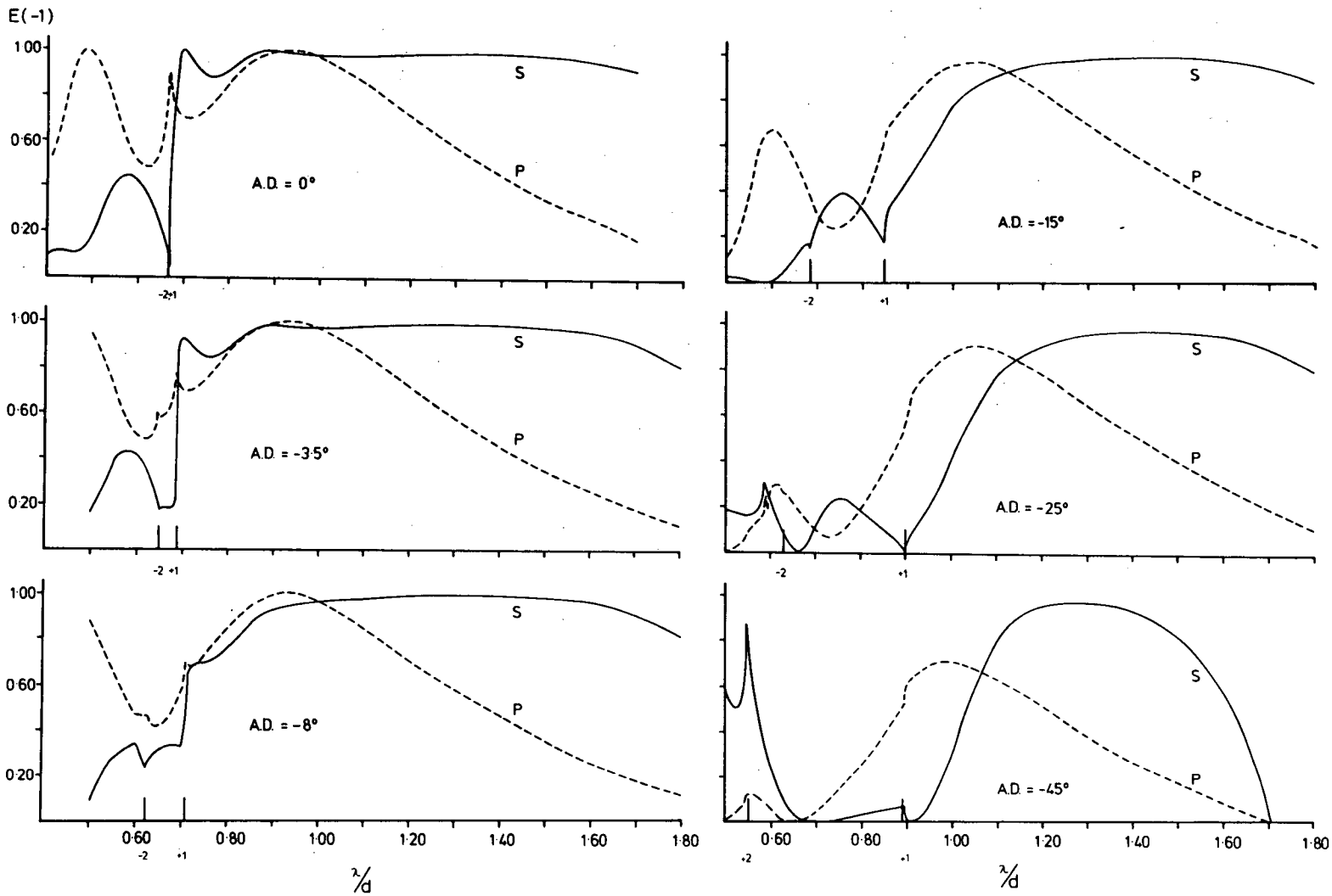


Figure 2.5.6 Distorted sinusoidal groove grating, h = 0.45 d, pre-exposure = 50 mJ/cm^2.

The S polarization blaze is slightly weakened by increasing deviation. However, the most important effect is the considerable curtailment of the blaze for deviations in excess of 15^{o}.

On the basis of the evidence presented here, this grating should not be used with deviations greater than 15^{o} if it is desired to achieve a strong blaze with only moderate polarization effects. On the other hand, the unpolarized radiation efficiency curve is free from the abrupt jump in the Littrow curve associated with the (- 2, + 1) Wood anomaly for $D = -15^{o}$.

The two quasi-sinusoidal groove profiles discussed in this section demonstrate less stability towards change of configuration than do the ruled gratings of sections (2.5.2.1-3). Thus designers have much less freedom in the design of grating instruments should they decide to use holographic rather than conventional gratings.

For all profile forms, it appears that increasing the groove depth relative to the period results in a blaze that is more sensitive to deviation changes. However, with ruled gratings one can always obtain a good Littrow blaze regardless of the value of the normalized groove depth. For holographic gratings, one is forced to use relatively deep grooves in order to obtain adequate spectral performance.

2.5.4 Higher order blaze studies for triangular profiles

Although the use of gratings in high orders is not uncommon, no theoretical and little experimental data is available in the literature on higher order efficiencies. Presented here are the results of a comprehensive study of second and third order efficiencies of gratings with the order of interest used in the Littrow mount. Although the relevant

experimental work [2.23] is still in progress, results so far obtained are in good agreement with the theoretical findings presented below.

2.5.4.1 *2nd order blaze studies*

A grid of efficiency curves for six blaze angles ranging between $13^{o}54'$ and $48^{o}55'$ is presented in figure (2.5.7). With increasing blaze angle, the location of the P efficiency peak moves further away from the S efficiency peak resulting in higher polarization of the diffracted radiation. For the two highest blaze angles ($36^{o}52'$ and $48^{o}55'$) the P blaze peak is associated with the (- 3, + 1) Wood anomaly resulting in higher maximum efficiencies. (Unfortunately in S polarized radiation, the anomaly is a strong minimum and so the ellipticity of the diffracted radiation approaches unity.)

In S polarized radiation, the diffracted efficiency always maximizes at 100%. This property has been demonstrated theoretically by Maréchal and Stroke [2.24]. They showed that

$$\frac{\lambda_{B,m}}{d} = \frac{2 \sin \alpha}{|m|} \qquad (2.5.2)$$

where $\lambda_{B,m}/d$ is the normalized blaze wavelength for order m in S polarized radiation and α is the groove facet angle (the grating being operated in a Littrow configuration).

In table (2.5.3), values typifying the behaviour of these gratings in P and U polarized radiation are presented. Here, it has been chosen to characterize the blaze width by quoting the wavelength interval $\Delta\lambda$ between successive points at which the efficiency is 60%, normalized with respect to the wavelength at which the efficiency is maximal. The

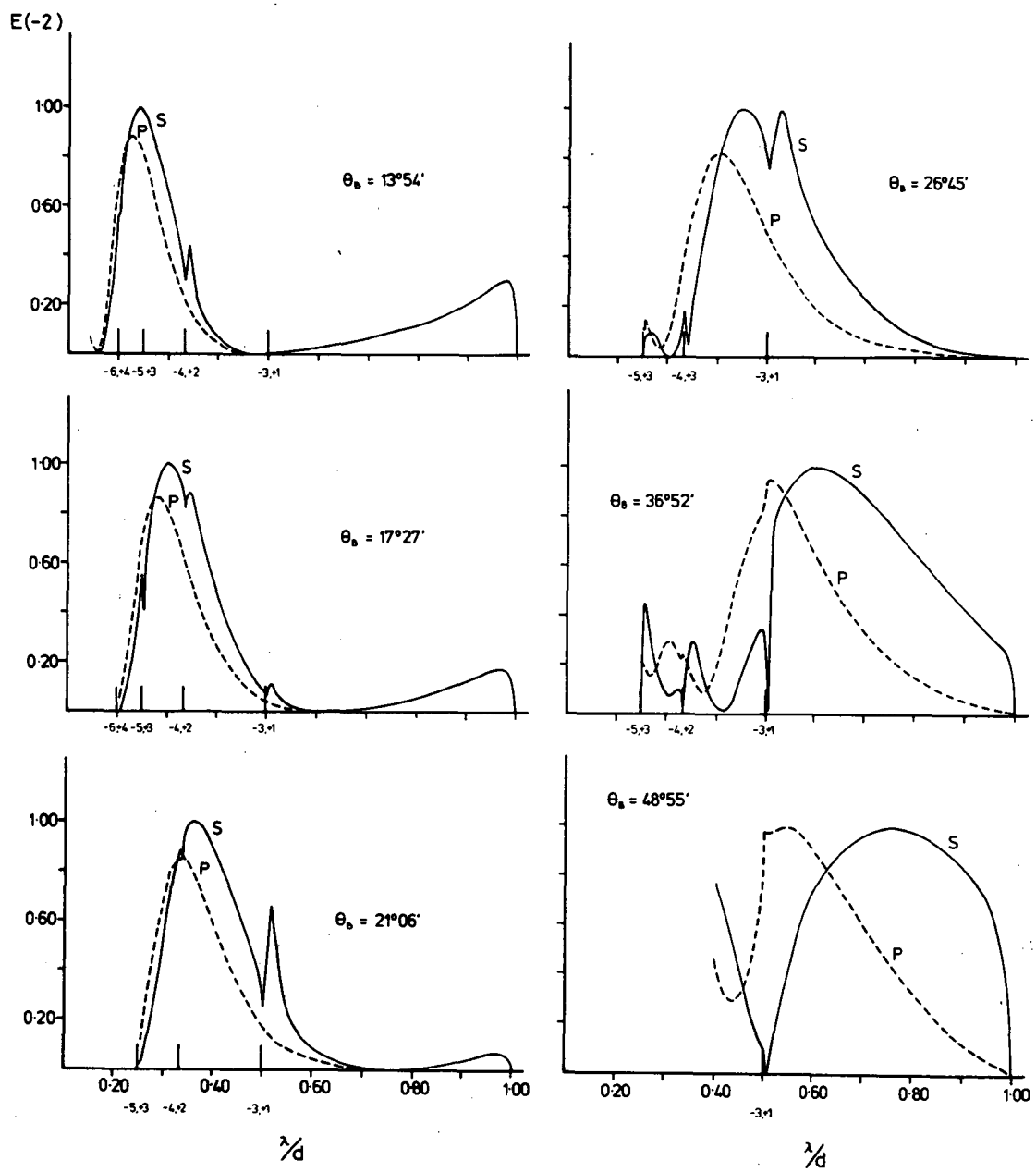


Figure 2.5.7 Littrow mount efficiency curves for order - 2 and a range of blaze angles (θ_B).

TABLE 2.5.3

θ_B	P Polarization			U Polarization		
	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$
$13^{o}54'$	0.23	0.88	0.31	0.23	0.93	0.39
$17^{o}27'$	0.28	0.87	0.29	0.29	0.92	0.39
$21^{o}06'$	0.34	0.86	0.32	0.35	0.91	0.38
$26^{o}45'$	0.40	0.83	0.33	0.43	0.87	0.50
$36^{o}52'$	0.51	0.95	0.34	0.54	0.88	0.37
$48^{o}55'$	0.54	0.99	0.41	0.63	0.83	0.49

A comparison of 2nd order spectral performance (in P & U polarized radiation) for a range of blaze angles.

TABLE 2.5.4

θ_B	P Polarization			U Polarization		
	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$
$17^{o}27'$	0.19	0.88	0.26	0.20	0.93	0.25
$21^{o}06'$	0.23	0.87	0.22	0.24	0.92	0.24
$26^{o}45'$	0.29	0.84	0.21	0.29	0.90	0.25
$36^{o}52'$	0.37	0.77	0.22	0.39	0.84	0.26
$48^{o}55'$	0.47	0.99	0.26	0.46	0.84	0.30

A comparison of 3rd order spectral performance (in P & U polarized radiation) for a range of blaze angles.

normalization was introduced since the blaze peak narrows as it moves towards shorter wavelengths.

As can be seen from the table, there are essentially two values of the normalized U-blazewidth, viz. 0.38 and 0.50. This quantity is dependent on the relative strengths of the (- 3, + 1) P and S polarization anomalies and also on the wavelength shift between the S and P blaze peaks. The position of the U blaze peak can be calculated to reasonable accuracy from equation (2.5.2) for facet angles not exceeding 30°. For deeper profiles, polarization effects are sufficiently strong to move the U peak significantly away from the S blaze position towards the blue.

2.5.4.2 *3rd order blaze studies*

A set of third order efficiency curves is shown in figure (2.5.8). Table (2.5.4) gives third order values corresponding to the second order values of table (2.5.3). The P polarization blaze efficiency decreases with increasing groove depth except in the case of the deepest profile studied. It is only for this profile that the peak lies on the long wavelength side of the (- 4, + 1) Wood anomaly, and that polarization effects are substantial.

The normalized U polarization blaze width remains constant at about 0.25 except for the $48^{\circ}55'$ profile, where the departure is again attributable to the (- 4, + 1) anomaly. (It is of interest that the passing off of the + 1 order has always been associated with the most important anomalous effects for all profiles so far considered.)

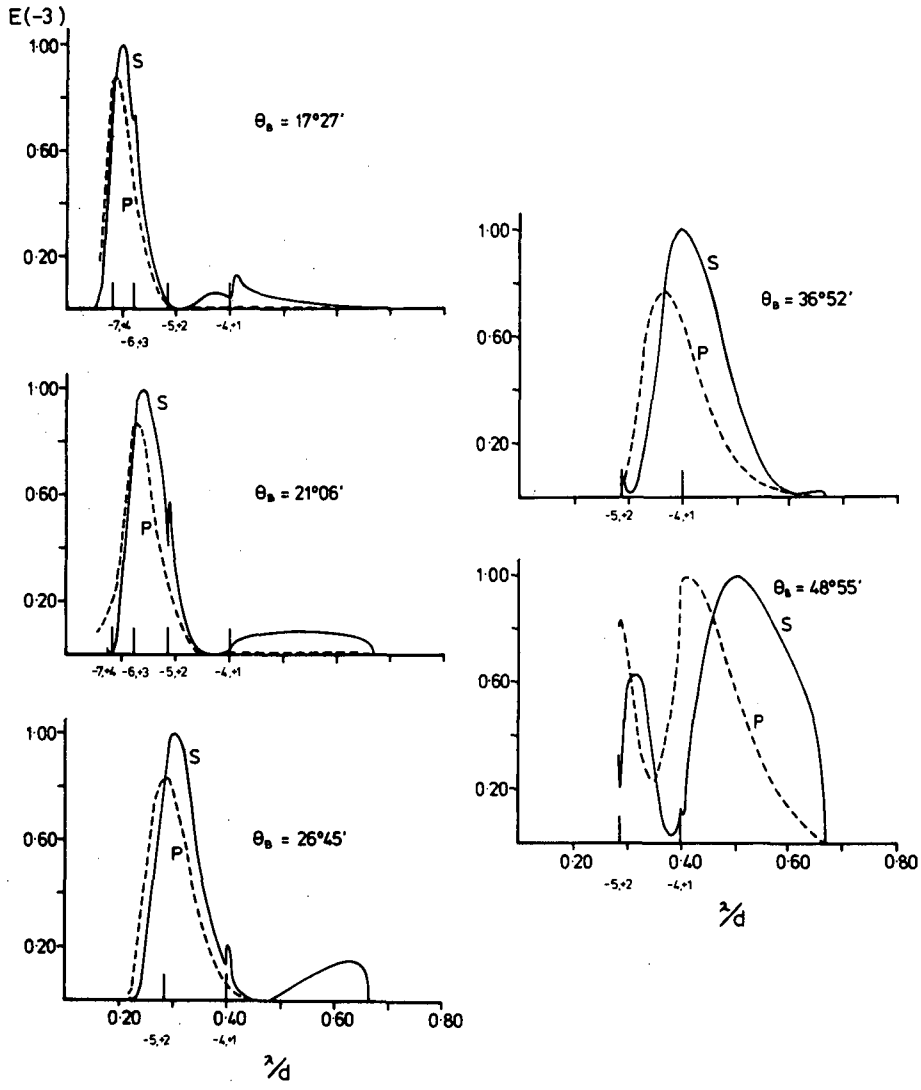


Figure 2.5.8 Littrow mount efficiency curves for order -3 and a range of blaze angles (θ_B).

2.5.4.3 *Overlapped blazes*

Let us now consider the exploitation of the blazes of orders - 2 and - 3 so as to obtain adequate efficiencies over an extended range of wavelengths. When the efficiency in one order falls below a specified minimum level, the efficiency in the next order must have already risen above it. Thus an overlap of first, second and third order blazes is necessary, the extent of the overlap depending upon the minimum efficiency level required and the blaze angle.

In the following, the performance obtained by specifying an efficiency minimum of 40% will be discussed briefly. For the two shallowest profiles, the overlap between the blazes is insufficient to enable an unbroken band of efficiencies exceeding 40%. For the deepest profiles, this criterion can be met. (The substantial (- 2, + 1) Wood anomaly in the first order of the $26^{\circ}45'$ profile necessitates use of a configuration other than the Littrow mount if the blaze is to be unbroken.)

As the groove depth increases the overlap between one blaze and the next increases, resulting in a reduction in the effective wavelength range over which the grating can successfully be operated. The ratio of the upper and lower wavelengths defining the useful range varies between 5 and 7.5 for the four deepest grooves.

2.5.5 Higher order blaze studies for holographic gratings

Shown in figure (2.5.9) are two sets of curves giving P and S blazes of a near-optimal sinusoidal grating. Tables (2.5.5-6) present performance details of gratings having a range of groove depths about the optimum. From table (2.5.5), it can be seen that the best performance in the second order will be achieved for groove depths close to the

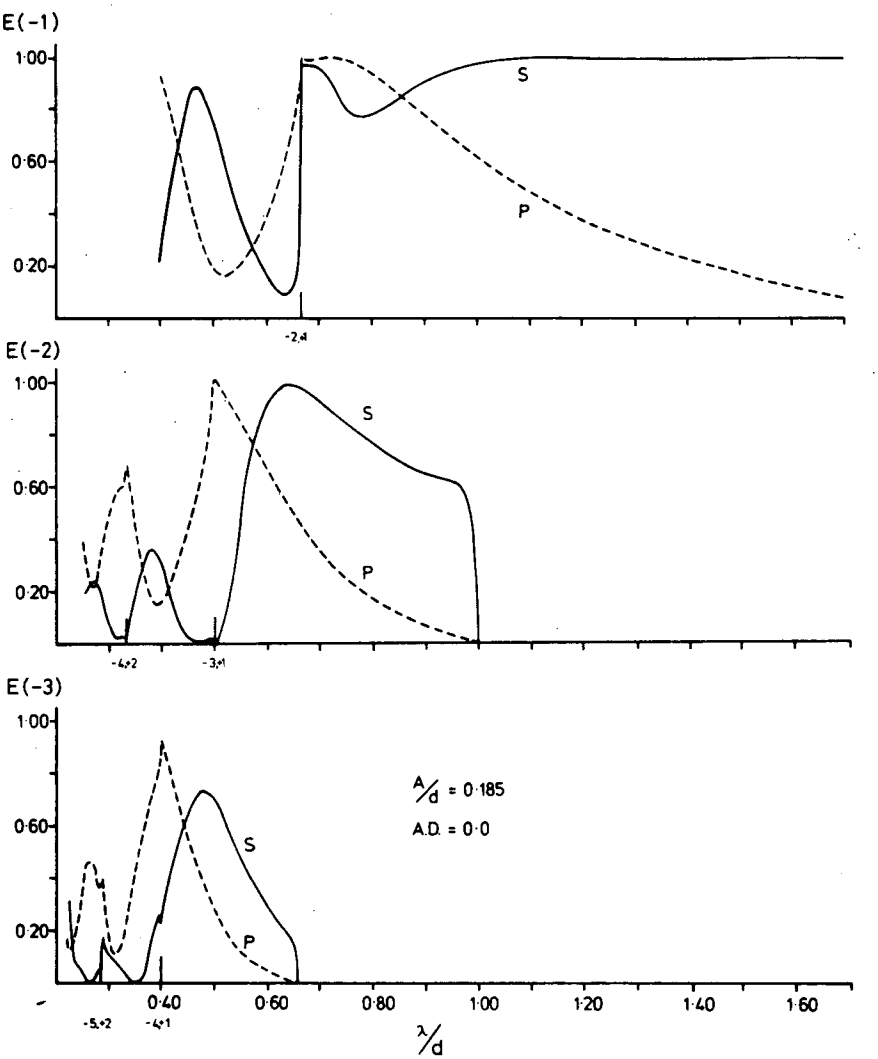


Figure 2.5.9 Littrow mount efficiency curves for orders -1, -2 and -3 for a sinusoidal grating, h = 0.37 d.

TABLE 2.5.5

A/d	P Polarization			U Polarization		
	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$
0.15	0.50	0.81	0.18	0.50	0.80	0.15
0.185	0.50	1.00	0.30	0.60	0.79	0.29
0.20	0.51	0.99	0.33	0.62	0.79	0.34
0.25	0.58	0.81	0.33	0.70	0.69	0.31

Variation of 2nd order blaze of sinusoidal groove gratings for a range of groove depths about the optimum.(As for the following tables, $\Delta\lambda/\lambda_{max}$ is obtained at the 60% efficiency level.)

TABLE 2.5.6

A/d	P Polarization			U Polarization		
	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$	$(\lambda/d)_{max}$	E_{max}	$\Delta\lambda/\lambda_{max}$
0.15	0.40	0.58	-	0.43	0.43	-
0.185	0.40	0.92	0.18	0.43	0.61	0.09
0.20	0.40	0.97	0.20	0.42	0.65	0.20
0.25	0.44	0.78	0.20	0.46	0.59	-

Variation of 3rd order blaze of sinusoidal groove gratings for a range of groove depths about the optimum.

previously determined first order optimum. (The precise choice of groove depth would depend on the choice of the minimum efficiency level associated with the bandwidth to be maximized.)

It is to be noted that the effects of polarization are quite marked for both second and third order blazes and increase rapidly with groove depth. The third order blazewidth is very narrow when compared with those associated with triangular profiles, despite the fact that the second order sinusoidal grating performance is only slightly reduced.

The overlapping of the first, second and third order blazes to achieve an extended bandwidth is also possible with these gratings, although the wavelength range over which an instrument may be used is considerably smaller than those discussed in section (2.5.4.3). The designer is also forced to use the grating with significantly higher angles of incidence.

Higher order blaze studies have also been undertaken for the optimal groove profile described by Wilson and McPhedran [2.22]. A summary of these studies is presented in figure (2.5.10) and table (2.5.7). The P blaze associated with this profile in the second and third orders is significantly broader than the same blaze for the optimum sinusoidal profile grating. However in both the second and third orders, the S blaze is very weak and narrow. Hence the incident radiation becomes severely polarized. Due to the weakness of the S blaze, it may not be feasible to produce an extended bandwidth by overlapping the diffracted orders.

2.5.6 Gratings having a plurality of blaze angles

In earlier sections, a method of extending the usable bandwidth

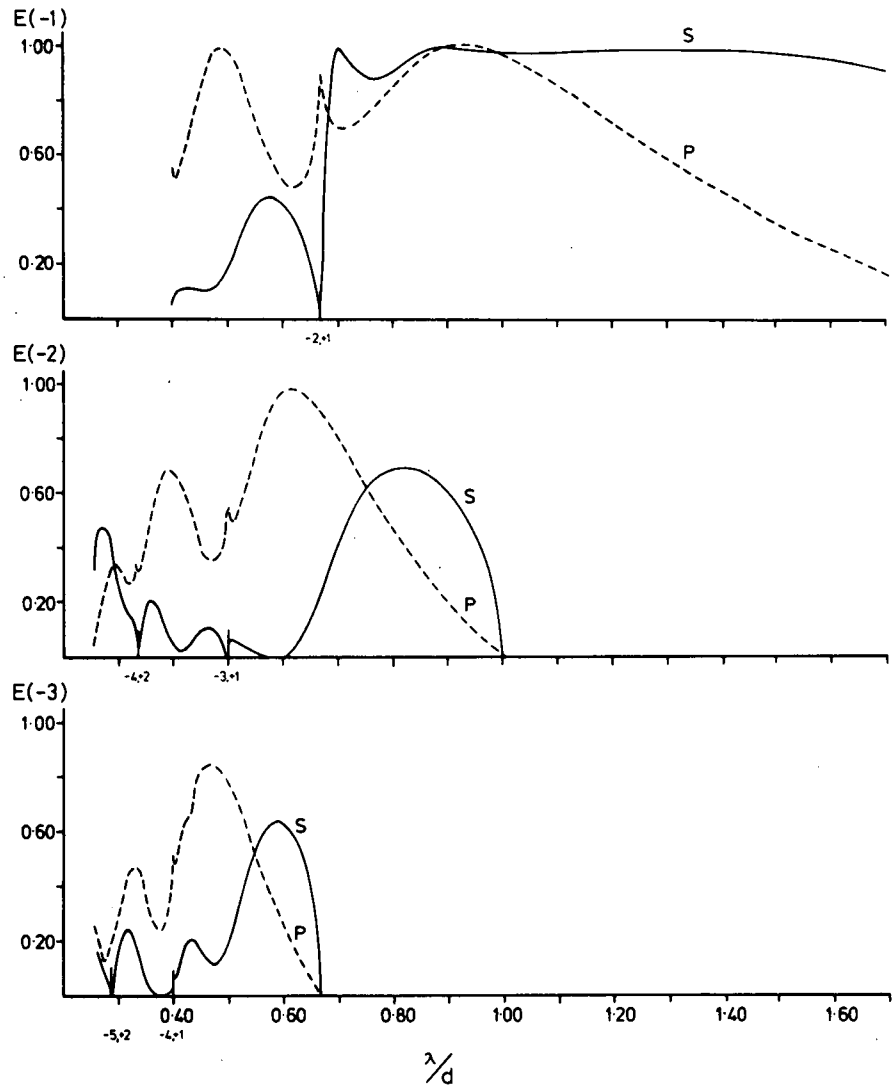


Figure 2.5.10 Littrow mount efficiency curves for orders -1, -2 and -3 for a distorted sinusoidal grating, h = 0.45 d, pre-exposure = 50 mJ/cm^2.

TABLE 2.5.7

Order	P Polarization			U Polarization		
	$(\lambda/d)_{max}$	E_{max}	$(\Delta\lambda/\lambda)_{max}$	$(\lambda/d)_{max}$	E_{max}	$(\Delta\lambda/\lambda)_{max}$
-1	0.94	1.00	0.67	0.92	0.996	1.01
-2	0.61	0.98	0.38	0.76	0.62	0.11
-3	0.47	0.84	0.26	0.55	0.52	

Blazes of the optimum distorted groove grating in orders -1, -2, and -3 for the Littrow mount.

of spectrographs by exploiting higher order blaze effects was discussed. An alternative technique has been advanced by J.J.J. Staunton [2.25]. It was suggested that one can incorporate a range of blaze angles into a groove, in order to give an extended spectral range in a single diffracted order. Staunton gives specific examples of groove profiles and sketched efficiency curves showing improved performance.

Efficiency curves for the profiles specified by Staunton, and also for other profiles of a similar form have been calculated. Staunton has suggested that with each blaze angle on the groove is associated a separate blaze peak on the efficiency curve. However, our studies have revealed that this in fact is not the case. The efficiency curves for P and S polarization correspond rather more closely to an average of efficiencies for purely triangular grooves than to an entirely new form of spectral performance. The performance of a grating having a plurality of blaze angles is in fact worse than that associated with the best blaze of any of its facets.

This example is quite interesting in that it again stresses that the blazing of gratings can only be understood and predicted accurately by using rigorous electromagnetic techniques. While geometric pictures of blazing can sometimes be useful, misleading results are frequently obtained.

2.5.7 Conclusions

Investigations into the effects of angular deviation on the spectral performance of ruled gratings have shown that they are highly dependent upon the polarization of the incident radiation. For P polarization, angular deviation generally results in a reduction of peak efficiencies

and a shift of the efficiency curves towards the blue. However, for the orthogonal polarization, the situation is more complex. The S polarization blaze is little affected in both strength and position by angular deviation. However, by controlling this variable, one can reduce and in some cases even eliminate the undesirable effects of anomalies.

P and S polarization blazes for shallow profiles are little affected by variation of angular deviation. As has been observed in earlier studies, gratings with facet angles in the range 17^{o} to 22^{o} have a blaze region relatively free of anomalies. This has been shown here to hold regardless of the choice of angular deviation.

For deeper triangular profiles (e.g. the $26^{o}45'$ blaze angle), angular deviation can be used to separate and weaken the troublesome - 2 and + 1 anomalies without affecting the strength of the S polarization blaze. An optimal choice of deviation for this grating appears to be 15^{o}.

On the other hand, quasi-sinusoidal gratings should not be used with deviations much in excess of 10^{o}. It has been shown that as deviation increases, the P polarization blaze becomes weaker and moves towards longer wavelengths with the + 1 Wood anomaly. The S polarization blaze, although always associated with an efficiency maximum of 100%, is curtailed by increasing deviations.

The blazes of all types of gratings become more sensitive to the effects of angular deviation as the groove depth is increased. Whereas good blazes can be obtained for triangular profile gratings of any depth, the instrument designer is forced to use relatively deep holographic

profiles so as to obtain adequate spectral performances. This means that the former type of grating lends itself to use in a far wider range of configurations than does the latter type.

Our studies have shown the inaccuracy of the blaze equation (2.5.1), which is often used to predict the effect of angular deviation on the blaze wavelength of ruled gratings. This expression predicts a movement of the blaze towards shorter wavelengths with increasing deviation. In fact, for S polarization the movement is towards the red. For unpolarized light, the movement is towards the blue, but is considerably smaller than would be expected from equation (2.5.1).

The higher order blazes of both triangular and quasi-sinusoidal profiles have also been investigated. For the former groove type, it has been shown that by overlapping the first, second and third order blazes one can obtain unpolarized light efficiencies in excess of 40% over a spectral region in which the wavelength changes by a factor of up to 7.5. For holographic profiles, this technique of using overlapping higher-order blazes is much less successful.

Also reviewed were the consequences of having a range of blaze angles in a single groove. Although geometrical optics arguments might suggest a consequent widening of blaze, our calculations indicate that "multiple-blazing" cannot be achieved in this way, and that having a plurality of blaze angles in fact worsens grating performance.

Finally, it should be pointed out that the grooves of holographic gratings are not inevitably near-sinusoidal in form. Sheridon [2.26] has suggested a method for producing quasi-triangular grooves holographically, and recent work [2.27, 2.28] has confirmed the promise of

this technique. The results presented here have illustrated the point that while gratings having near sinusoidal profiles may be just as useful as gratings having triangular profiles in specific circumstances, the latter are more versatile. Thus, it is advocated that more work be done concerning the production of quasi-triangular profiles by holographic means.

REFERENCES

[2.1] R. Petit, Rev. d'Opt., 45 (1966) 249.

[2.2] J. Pavageau and J. Bousquet, Optica Acta, 17 (1970) 469.

[2.3] D. Maystre and R. Petit, Nouv. Rev. Opt., 2 (1971) 115.

[2.4] R. C. McPhedran and M. D. Waterworth, Optica Acta, 20 (1973) 177.

[2.5] R. C. McPhedran, I. J. Wilson and M. D. Waterworth, Opt. Commun., 7 (1973) 331.

[2.6] D. Maystre and R. Petit, Opt. Commun., 2 (1970) 309.

[2.7] D. Maystre and R. Petit, Opt. Commun., 4 (1971) 25.

[2.8] D. Maystre, R. Petit, M. Duban and J. Gilewicz, Nouv. Rev. Opt., 5 (1974) 79.

[2.9] R. C. McPhedran and I. J. Wilson, Japan. J. Appl. Phys., 14 (1975) Suppl. 14-1, 159.

[2.10] D. S. Jones, "Theory of Electromagnetism", Pergamon Press (1964), New York.

[2.11] D. Maystre, Thèse No. AO 9545, (1974), L'Université d'Aix-Marseille III, France.

[2.12] J. Meixner, I.E.E.E., Trans. Ant. Prop., AP 20 (1972) 442.

[2.13] J. C. Bolomey, Thèse No. AO 5604, (1971), France.

[2.14] I. J. Wilson, L. C. Botten and R. C. McPhedran, (1976), University of Tasmania Research Report DGRG 76/1.

[2.15] I. J. Wilson, L. C. Botten and R. C. McPhedran, J. Optics (Paris), 8 (1977) 217.

[2.16] E. G. Loewen, D. Maystre, R. C. McPhedran and I. J. Wilson, Japan. J. Appl. Phys., 14 (1975) Suppl. 14-1, 143.

[2.17] R. C. McPhedran and M. D. Waterworth, Optica Acta, 19 (1972) 877.

[2.18] D. Maystre and R. C. McPhedran, Opt. Commun., 12 (1974) 164.

[2.19] J. E. Stewart and W. S. Gallaway, Appl. Optics, 1 (1962) 421.

[2.20] J. Verrill, Optica Acta, 23 (1976) 425.

[2.21] R. Petit, Optica Acta, 14 (1967) 301.

[2.22] I. J. Wilson, R. C. McPhedran and M. D. Waterworth, Opt. Commun., 9 (1973) 263.

[2.23] E. G. Loewen, personal communication.

[2.24] A. Maréchal and G. W. Stroke, C. R. Acad. Sci. Paris, 249 (1959) 2042.

[2.25] J. J. J. Staunton, (1962), U. S. Patent No. 3045532.

[2.26] N. K. Sheridon, Appl. Phys. Letters, 12 (1968) 316.

[2.27] M. C. Hutley, Optica Acta, 22 (1975) 1.

[2.28] H. A. Obermayer, Opt. Commun., 13 (1975) 426.

[2.29] E. G. Loewen, M. Nevière and D. Maystre, Appl. Optics, 16 (1977) 2711.

CHAPTER 3

THE EFFECTS OF FINITE CONDUCTIVITY

3.1 INTRODUCTION

Although the infinite conductivity theories, discussed in the previous chapter, are in excellent agreement with experiment in the far infrared and millimetric wave regions, severe discrepancies between their predictions and the actual behaviour of gratings in the visible and ultraviolet regimes often manifest themselves. In particular, with S polarized light, certain shallow gratings can be made to exhibit total absorption [3.1] of the incident radiation. Thus, there was a need for more sophisticated theories which could take into account the nature of the finitely conducting structures.

The first of these theories was proposed by Neureuther and Zaki [3.2] and Van den Berg [3.3]. However, the numerical implementation of their integral formalisms was complicated by the fact that two unknowns (the field and its normal derivative at the surface of the grating) were needed to completely characterize the diffraction problem. This difficulty was removed in 1972 when Maystre [3.4] first published his theory of finitely conducting gratings. In his integral formulation, he was able to express the previous two field quantities in terms of a single unknown, akin to the radiating surface currents of the infinite conductivity theories, and in doing so greatly simplified the numerical treatment of the problem.

The same problem has been tackled on another front with the aid of the differential formalism first developed by Cerutti-Maori [3.5]. During this decade, work by Nevière, Vincent and Petit [3.6] has increased the sophistication of this technique and today it may be regarded as being (in some ways) as powerful a tool as the integral formalisms.

Nevertheless, it is Maystre's method which has been thoroughly tested against experimental data. In all comparisons, it has been demonstrably accurate and with such successful confrontations the understanding of the nature of finite conductivity effects has been significantly enhanced. Furthermore, it is his method which has so greatly influenced the efforts of the author and thus it is felt appropriate to provide a review of this theory (in section (3.2)).

In the following section, a discussion of the effects of finite conductivity is presented in an attempt to gain some further understanding of the nature of the "plasmon" anomaly. Throughout this investigation, the author enjoyed many hours of illuminating discussion with Dr. Ross McPhedran and these are gratefully acknowledged. Within section (3.3), particular reference is made to the recent works of Maystre and Nevière [3.1, 3.7, 3.8], who so ably characterized the the anomalous behaviour in terms of poles and zeros of various field quantities. The rigorous diffraction theory has been confronted with the experimental data of Wheeler [3.9]. The results obtained provide further confirmation of the validity of Maystre's theory and reveal serious inadequacies in the first order perturbation theory proposed by Elson [3.10].

Finally, in section (3.4), a new experimental application of the theory is described. This study, conducted in collaboration with Dr. Ian Wilson, showed that the groove depth of a sinusoidal profile grating could be uniquely determined by correlating the predictions of the theory with the experimental results obtained with the aid of a gas laser, used in a particularly simple optical configuration.

3.2 MAYSTRE'S FINITE CONDUCTIVITY THEORY

Consider a P polarized plane wave field $\underline{E}_i$ incident upon the structure shown in figure (3.2.1). Here,

$$\underline{E}^i = \exp[i(\alpha_0 x - \chi_0 y)]\ \hat{\underline{k}} \qquad (3.2.1)$$

(where the relevant nomenclature is identical to that defined in chapter 2.) In region D_0, a P polarized diffracted field (having an outward going nature exterior to the grooves) is established, being given by

$$\underline{E}^d = \underline{E} - \underline{E}^i. \qquad (3.2.2)$$

Similarly in the region D_1, there also exists an outward going field, which in this case is the total field. These two fields obey the Helmholtz equations appropriate to the region in which they are propagating and also the radiation conditions as $|y| \to \infty$. Thus for a point P_0 within D_0

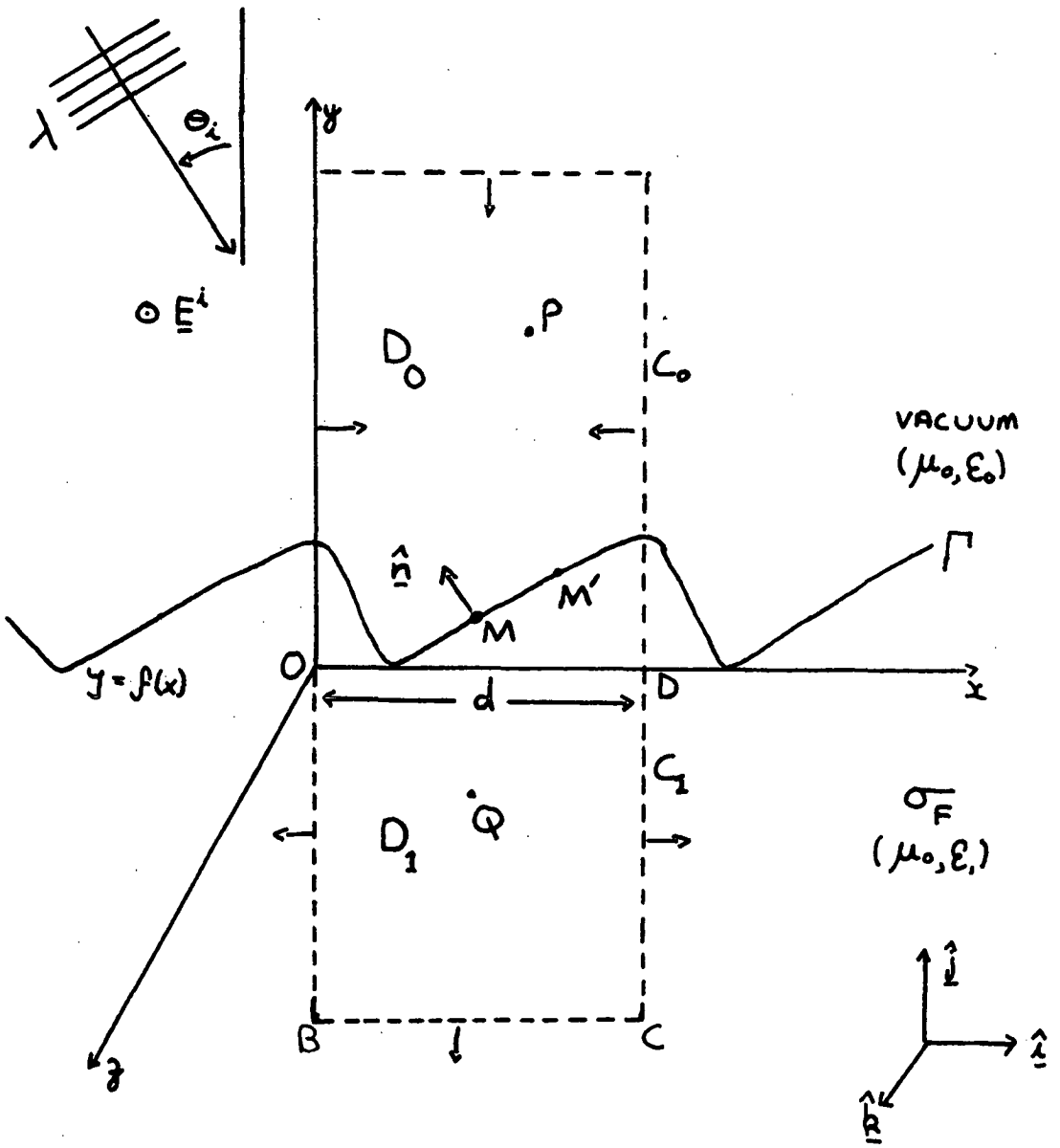


Figure 3.2.1 The geometry of the diffracting arrangement. The region D_0 represents free space while the domain D_1 represents the finitely conducting grating.

$$(\nabla^2 + k_0^2)\, E^d(P_0) = 0 \tag{3.2.3}$$

and for a point P_1 in D_1

$$(\nabla^2 + k_1^2)\, E(P_1) = 0. \tag{3.2.4}$$

(The wave-number in the domain D_1 is given by

$$k_1 = k_0 r$$

where r is the refractive index of that medium.)

The diffraction problem is constrained by the boundary conditions at Γ. These require the tangential components of both the total electric and total magnetic fields to be continuous across the interface separating D_0 and D_1. For this particular polarization these reduce to the form

$$E(M)\Big|_{\Gamma_+} = E(M)\Big|_{\Gamma_-} \tag{3.2.5}$$

and $$\frac{1}{\mu_0} \frac{\partial E(M)}{\partial n}\Big|_{\Gamma_+} = \frac{1}{\mu_1} \frac{\partial E(M)}{\partial n}\Big|_{\Gamma_-} \tag{3.2.6}$$

where μ_i is the magnetic permeability of the medium D_i. Throughout, the following analysis it will be assumed that both regions are non-magnetic, (an entirely reasonable assumption for optical diffraction gratings) and thus equation (3.2.6) simply states that the normal derivative of the electric field is continuous across Γ.

Having now completely defined the problem, it only remains to select an appropriate function which can completely and uniquely characterize it. Maystre [3.4] achieved this objective by defining a scalar function u by the following criteria:

(i) $u(P) = E^d(P)$ if $P \in D_0$,

(ii) $(\nabla^2 + k_0^2)\, u(P) = 0$ if $P \in D_1$,

(iii) u obeys the radiation conditions as $|y| \to \infty$,

(iv) u is continuous across the interface.

Now by pursuing an analysis is similar to that described in section (2.3) one arrives at an expression of the form

$$u(P) = \int_{\Gamma} G_0(P;M')\, \xi(M')\, ds' \qquad (3.2.7)$$

for some point P within D_0. The function ξ represents the discontinuity across Γ of the normal derivative of u

$$\text{i.e. } \xi(M') = \frac{\partial u(M')}{\partial n'}\bigg|_{\Gamma_+} - \frac{\partial u(M')}{\partial n'}\bigg|_{\Gamma_-} . \qquad (3.2.8)$$

Thus ξ may be thought of as being that current flowing in the surface of a perfectly conducting grating which would excite exactly the same diffracted wave field above Γ as would the finitely conducting grating. (Hence, the use of the term "pseudo current".)

It then follows that for some point M on the profile

$$E(M) = E^{i}(M) + \int_{\Gamma} G_0(M;M')\, \xi(M')\, ds' \qquad (3.2.9)$$

and

$$\frac{\partial E(M)}{\partial n} = \frac{\partial E^{i}(M)}{\partial n} + \tfrac{1}{2}\, \xi(M) + \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n}\, \xi(M')\, ds' \quad . \qquad (3.2.10)$$

Thus, E and $\partial E/\partial n$ can both be expressed in terms of the fundamental unknown ξ, thereby eliminating the need for the two separate unknowns used in earlier theories.

Up until this stage, the analysis has been identical to that discussed in the review of the infinite conductivity formalisms. However, from this point onwards, the treatments "diverge". This arises from the indirect nature of the boundary conditions which specify the continuity of certain field components. In order to couple, and hence constrain, the quantities E and $\partial E/\partial n$, it is necessary to derive an integral equation. This is accomplished by defining a Green's function appropriate to D_1 by

$$G_1(Q;M') = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\eta_n} \exp[i\alpha_n(x-x') + i\eta_n|y-y'|] \qquad (3.2.11$$

where $\eta_n = \sqrt{(k_1^2 - \alpha_n^2)}$

and applying Green's theorem to the region D_1. Thus,

$$\iint_{D_1} [\nabla^2 G_1(Q;M')\, E(M') - \nabla^2 E(M')\, G_1(Q;M')]\, dA'$$

$$= \int_{C_1} [\frac{\partial G_1(Q;M')}{\partial n'} E(M') - \frac{\partial E(M')}{\partial n'} G_1(Q;M')]\, ds' \qquad (3.2.12)$$

where the contour C_1 consists of the segments Γ_-, AB, CD and BC. The line integrals along AB and CD may be shown to cancel as a consequence of the field "pseudo-periodicity" whereas the line integral along BC vanishes due to the radiation conditions. Thus, expression (3.2.12) reduces to the form

$$E(Q) = \int_{\Gamma_-} [\frac{\partial G_1}{\partial n'} E - G_1 \frac{\partial E}{\partial n'}] \, ds'. \qquad (3.2.13)$$

Now, should Q tend to some point M on Γ, then effectively only "half of a singularity" is included within the region of integration and so

$$\tfrac{1}{2}E(M) = \int_{\Gamma_-} [\frac{\partial G_1}{\partial n'} E - G_1 \frac{\partial E}{\partial n'}] \, ds' \, . \qquad (3.2.14)$$

The above expression is an integral equation constraining the behaviour of E and $\partial E/\partial n$. By applying field continuity and substituting in the representations for E and $\partial E/\partial n$ given in equations (3.2.9-10) this reduces to the form

$$\xi_0(M) = \int_{\Gamma} \tau^E(M;M') \, \xi(M') \, ds' \qquad (3.2.15)$$

$$\text{where } \xi_0(M) = \tfrac{1}{2}E^i(M) - \int_{\Gamma} [\frac{\partial G_1(M;M')}{\partial n'} E^i(M') - G_1(M;M') \frac{\partial E^i(M')}{\partial n'}] \, ds' \qquad (3.2.16)$$

$$\text{and } \tau^E(M;M') = -\tfrac{1}{2} G_0(M;M') - \tfrac{1}{2} G_1(M;M') + \int_{\Gamma} [\frac{\partial G_1(M;M'')}{\partial n''} G_0(M'';M') - G_1(M;M'') \frac{\partial G_0(M'';M')}{\partial n''}] ds'' \qquad (3.2.17)$$

Equation (3.2.15) is a Fredholm integral equation of the first kind, which is once again solved using the points-matching technique. In general, its solution presents little, if any, additional difficulties not already circumvented in the the solution of earlier problems. However, some care is needed in the cases of

(i) profiles having convex edges, and

(ii) very highly conducting gratings.

For the former, Maystre [3.4] has described a treatment for the dividing out of the field singularity at the edge. With reference to the latter, the Green's function assumes a "spike-like" nature as the conductivity increases and thus the problem becomes particularly ill-conditioned since only very few sampling points actually characterise the behaviour of the unknown. Maystre [3.4] has also discussed a technique for removing this obstacle.

In the limit as the conductivity approaches infinity, the integral equation (3.2.15) may be shown to degenerate to the form originally derived by Petit (which was discussed in section (2.3.1)).

Let us now turn to the case of S polarized radiation and briefly discuss the salient differences between the treatments for the two fundamental polarizations. This time, the boundary conditions may be shown to be:

$$\left.\begin{aligned} H(M)\Big|_{\Gamma_+} &= H(M)\Big|_{\Gamma_-} \\ \text{and} \quad \frac{1}{\varepsilon_0}\frac{\partial H}{\partial n}\Big|_{\Gamma_+} &= \frac{1}{\varepsilon_1}\frac{\partial H}{\partial n}\Big|_{\Gamma_-} , \end{aligned}\right\} \qquad (3.2.18)$$

where ε_i represents the permittivity of the medium D_i. In the case of a non-magnetic medium the latter equation becomes

$$\left.\frac{\partial H}{\partial n}\right|_{\Gamma_-} = r^2 \left.\frac{\partial H}{\partial n}\right|_{\Gamma_+} .$$

For this polarization, the function u is given by

(i) $u(P) = H^d(P)$ if $P \in D_0$,

(ii) $(\nabla^2 + k_0^2)\, u(P) = 0$ if $P \in D_1$,

(iii) u obeys the radiation conditions as $|y| \to \infty$ and is continuous across Γ.

It then follows that for some point P lying within D_0,

$$u(P) = \int_\Gamma G_0(M;M')\, \psi(M')\, ds' \qquad (3.2.19)$$

where $\psi(M') = \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_+} - \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_-}$. (3.2.20)

(Unlike the previous analysis where it was shown that ξ could be regarded as a "pseudo-current", it is more difficult to attach a physical meaning to the quantity ψ, and so it has to be regarded as being a mathematical abstraction.) Thus, from equation (3.2.19) it follows that

$$H(M) = H^{i}(M) + \int G_0(M;M')\ \psi(M')\ ds' \tag{3.2.21}$$

$$\text{and } \frac{\partial H(M)}{\partial n}\bigg|_{\Gamma_+} = \frac{\partial H^{i}(M)}{\partial n} + \tfrac{1}{2}\ \psi(M) + \int \frac{\partial G_0(M;M')}{\partial n}\ \psi(M')\ ds' \ . \tag{3.2.22}$$

The derivation of an integral equation linking these two field quantities is performed in an analogous fashion to the previous derivation and an expression identical in form (except for the replacement of E by H) to equation (3.2.14) results. Finally, by implementing the continuity equations (3.2.18) and substituting the free space representations for H and $\partial H/\partial n$ into the integral equation, one arrives at the final result -

$$\psi_0(M) = \int \tau^{H}(M;M')\ \psi(M')\ ds' \tag{3.2.23}$$

$$\text{where } \psi_0(M) = \tfrac{1}{2}\ H^{i}(M) - \int_{\Gamma}[\frac{\partial G_1(M;M')}{\partial n'}\ H^{i}(M') - r^2\ G_1(M;M')\ \frac{\partial H^{i}(M)}{\partial n'}\ ]\ ds' \tag{3.2.24}$$

$$\text{and } \tau^{H}(M;M') = -\tfrac{1}{2}\ G_0(M;M') - \tfrac{1}{2}\ r^2\ G_1(M;M') + \int_{\Gamma}[\frac{\partial G_1(M;M'')}{\partial n''}\ G_0(M'';M') - r^2\ G_1(M;M'')\ \frac{\partial G_0(M'';M')}{\partial n''}\ ]\ ds'' \ .$$

3.3 THE PHYSICAL EFFECTS OF FINITE CONDUCTIVITY

Prior to 1972, it was believed that the effects of finite conductivity could be accounted for simply by rescaling the infinite conductivity results with the plane reflectance of the metallic surface. However, in that year, the first efficiency curves calculated for finitely conducting gratings were presented [3.11] and the inadequacy of this assumption was revealed. These curves showed that, although the rescaling concept was entirely adequate for P polarization, vast differences between the rescaled results and those predicted by Maystre's theory [3.4] for S polarization manifested themselves. In fact the disaccord between the two theories was of sufficient magnitude to give rise to doubts concerning the validity of the finite conductivity formalism.

However, after confrontation [3.12] with the experimental results of Hutley [3.13] and Hutley and Bird [3.14], confirmation of the finite conductivity theory was obtained. Hutley and Bird, in attempting to explain the failure of the infinite conductivity formalism when dealing with gratings composed of highly conducting metals, attributed this to the ability of the metal surface to support surface plasma oscillations in S polarized light (but not in P polarized light). They supported this suggestion with evidence of deep energy absorption features and also organized patterns in the stray light from holographic gratings.

A surface plasmon is a collective oscillatory motion of electrons in a charge density wave propagating along and mainly confined to a metallic surface. In the case of a finitely conducting medium, such plasmons represent an additional loss mechanism and it is for this reason that studies concerning their nature are warranted, with particular

reference being made to diffraction gratings. However, before considering the resonant behaviour of a corrugated surface, it is essential to discuss the free resonances of a plane surface.

From Fresnel's laws the amplitude reflection and transmission coefficients associated with a plane interface bounding two semi-infinite media, of refractive indices 1 (free space) and $\tilde{n}$ ($= n_r + in_i$) are given by:

(i) for P polarization

$$R_P = \frac{\cos\theta - \sqrt{\tilde{n}^2 - \sin^2\theta}}{\cos\theta + \sqrt{\tilde{n}^2 - \sin^2\theta}}$$

$$T_P = \frac{2\cos\theta}{\cos\theta + \sqrt{\tilde{n}^2 - \sin^2\theta}}$$

and

(ii) for S polarization

$$R_S = \frac{\tilde{n}^2\cos\theta - \sqrt{\tilde{n}^2 - \sin^2\theta}}{\tilde{n}^2\cos\theta + \sqrt{\tilde{n}^2 - \sin^2\theta}}$$

$$T_S = \frac{2\tilde{n}^2\cos\theta}{\tilde{n}^2\cos\theta + \sqrt{\tilde{n}^2 - \sin^2\theta}}$$

(In these expressions θ represents the angle of incidence.) Now for a resonance condition to occur, there must exist both reflected and transmitted fields without the presence of any incident field,

i.e. $R,T \rightarrow \infty$.

The above expressions reveal that no such resonance condition is possible for P polarized light. However, for the orthogonal polarization, the constraint is

$$\tilde{n}^2 \cos\theta = -\sqrt{\tilde{n}^2 - \sin^2\theta},$$

whose solution is given by

$$\sin\theta = \sqrt{\frac{\tilde{\varepsilon}}{1+\tilde{\varepsilon}}}$$

where $\tilde{\varepsilon} = \tilde{n}^2$.

(Of course, no such resonance can be achieved for a dielectric medium.) Thus, there exists a pole in the complex θ plane for the field amplitudes, and this is referred to as a plasmon.

There also exists a zero in the reflectivity resulting from the solution of the equation

$$\tilde{n}^2 \cos\theta = +\sqrt{\tilde{n}^2 - \sin^2\theta}$$

whose solution is also given by

$$\sin\theta = \sqrt{\frac{\tilde{\varepsilon}}{1+\tilde{\varepsilon}}} \quad .$$

Thus in the case of a plane interface, the pole and the zero coincide with one another in the sin θ plane. However, their relation is of an indirect nature in that the zero is associated with an incident field

and no reflected field, whereas the pole is associated with a free resonance characterized by reflected and transmitted fields existing without the presence of an incident field. For metals, these fields are inhomogeneous plane waves which must be regarded as being unphysical since their energy densities diverge as x tends to one of $-\infty$ or $+\infty$. However, the plasmon waves are bound to the interface (i.e. their energy densities converge to zero as $|y|$ tends to infinity).

If one writes the dispersion relation in the following form:

$$\sin\theta = \sqrt{\frac{\tilde{\varepsilon}}{1+\tilde{\varepsilon}}} = \gamma_r + i\,\gamma_i,$$

then for highly conducting metals, $\gamma_r > 1$ and γ_i is small and positive. Thus, to create a plasmon, that is to witness a physical resonance effect consequent upon proximity to the pole in the complex sin θ plane, one must have an evanescent wave whose value of sin θ is approximately equal to γ_r. This was observed in the classic experiment of Otto [3.15] concerning frustrated total internal reflection, where a deep absorption resonance was exhibited in the reflected field.

It is possible to excite such resonances with homogeneous plane waves provided that there exists some momentum transfer mechanism by which the pole in the sin θ plane can be moved into the range

$$|\mathrm{Re}(\sin\theta)| \leq 1.$$

Such a transfer of momentum can be achieved by surface roughening, the most regular example of which is the diffraction grating. In the case of the grating this is summarized by the expression

$$\pm \gamma_r = \sin \theta_i + \frac{n\lambda}{d} ,$$

where a homogeneous plane wave incident at some real angle θ excites a diffracted order n whose direction sine is given $\pm \gamma_r$. Since γ_r has a value slightly greater than 1, it follows that the plasmon anomaly is located just on the long wavelength side of the Wood anomaly given by

$$\pm 1 = \sin \theta_i + \frac{n\lambda}{d} .$$

The above analysis is somewhat crude in that the dispersion relation used therein applies only to plasmons associated with a plane surface. Nevertheless, it may be regarded as a useful first approximation for gratings.

A rigorous derivation of the dispersion relation appropriate to a periodically roughened surface necessitates a full solution of the diffraction problem. During the past decade, there have been numerous attempts by the plasmon theorists to formulate a viable theory describing the behaviour of the grating. However, these have relied upon the assumption that the grating can be regarded as a first order perturbation of a plane surface (e.g. the work of Elson [3.10]). Maystre and Petit [3.1] have demonstrated the questionable nature of this assumption by finding a very shallow sinusoidal grating whose reflectance for a particular angle of incidence is reduced from approximately 90% (for the plane interface) to exactly zero.

In the first order perturbation theories, a probability of plasmon excitation is calculated. However a feature of the method is that these probabilities $\{P_n\}$ may exceed unity. In an attempt to re-scale the probabilities to lie in the range [0,1], Elson is forced to replace P_n by the quantity

$$1 - \exp(- P_n).$$

Nevertheless, even with such ad hoc assumptions, these theories are in qualitative agreement with the experimental results. It is possible for them to give reasonable quantitative agreement over a restricted spectral range, but only at the expense of making gross adjustments to both the optical constants and the profile shape of the grating. These discrepancies are revealed in the thesis of Wheeler [3.9], which contains a vast range of experimental data and corresponding curves generated by the Elson theory. A confrontation of the experimental results with the curves shown in figures (3.3.1-2) calculated using Maystre's finite conductivity formalism confirmed the validity of the rigorous diffraction theory. For most cases, it was found that this theory could emulate the experimental data to a high degree of accuracy, using only the published optical constants and profile specification. In those examples for which agreement between the rigorous theory and the experiment was somewhat less than perfect, the nature of the discrepancy suggested the presence of base line drift in Wheeler's apparatus.

On the other hand, the inadequacies of the perturbation theories are exemplified by the fact that a reasonable theoretical fit to the

Figure 3.3.1 S polarization zeroth order efficiency spectra for a 1200 g/mm. triangular profile grating having a blaze angle of $5^{o}10'$ and a right-angled apex ruled in aluminium. Graphs for wavelengths of

(a)	5200 Å	(d)	5800 Å
(b)	5400 Å	(e)	6000 Å
(c)	5600 Å	(f)	6200 Å

showing both the zeroth order efficiency (———) and the total diffracted energy (-----) are presented.

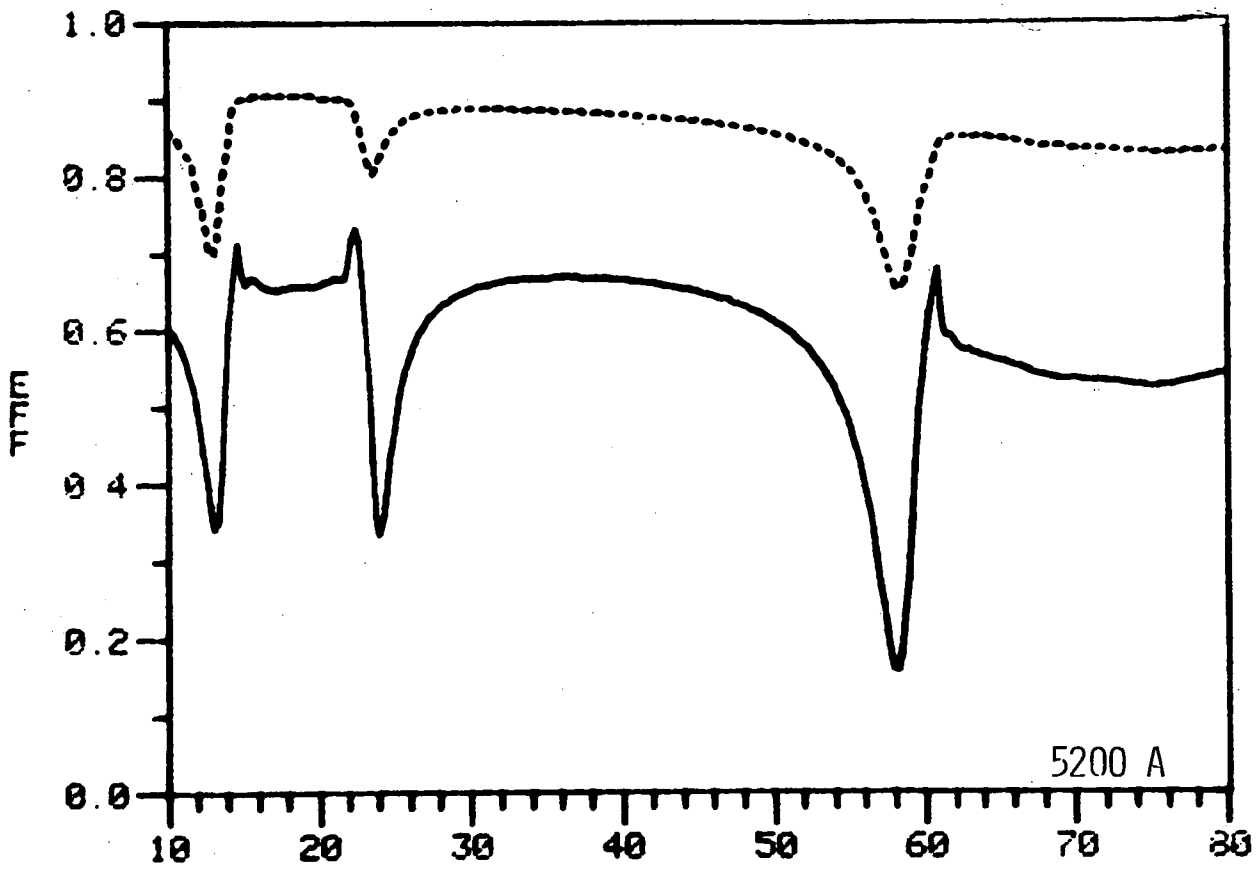


Figure 3.3.1(a)

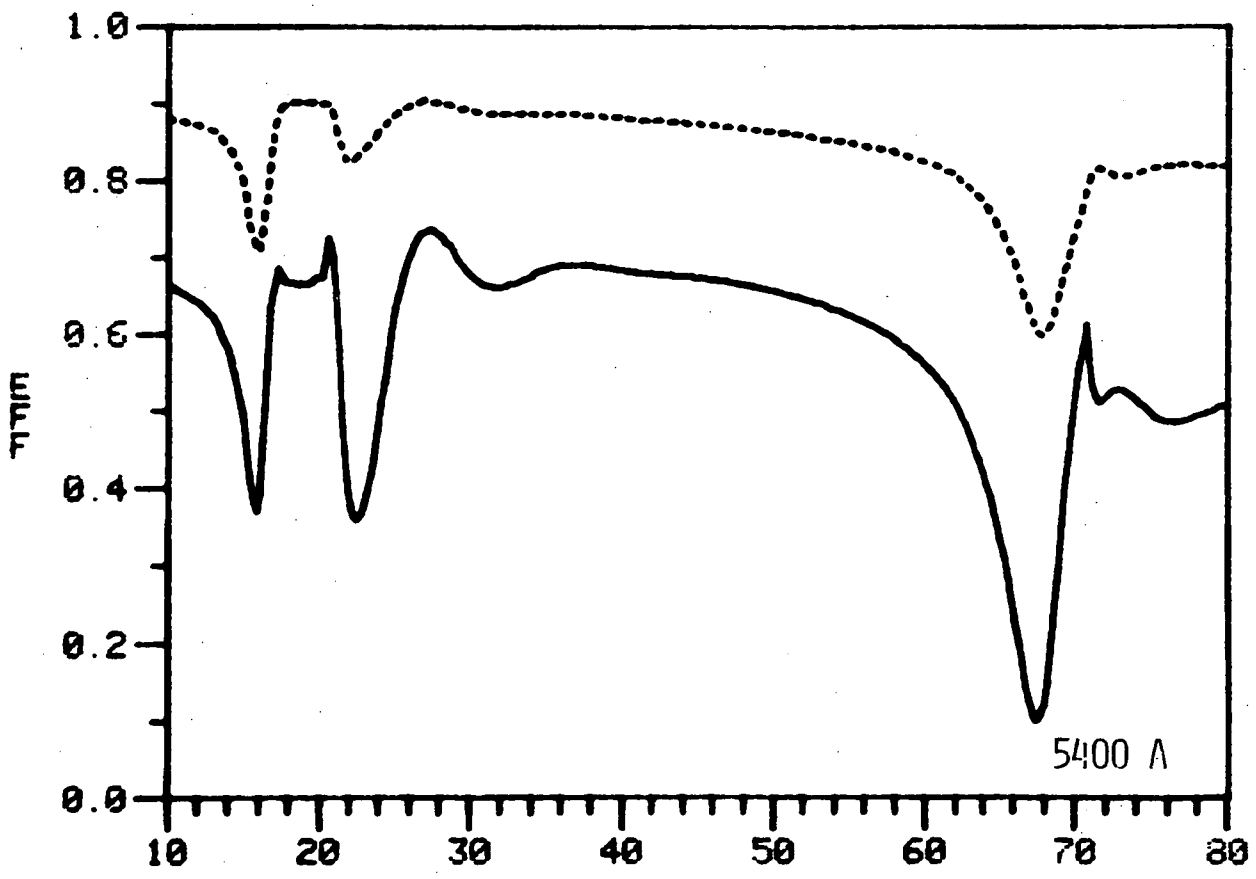


Figure 3.3.1(b)

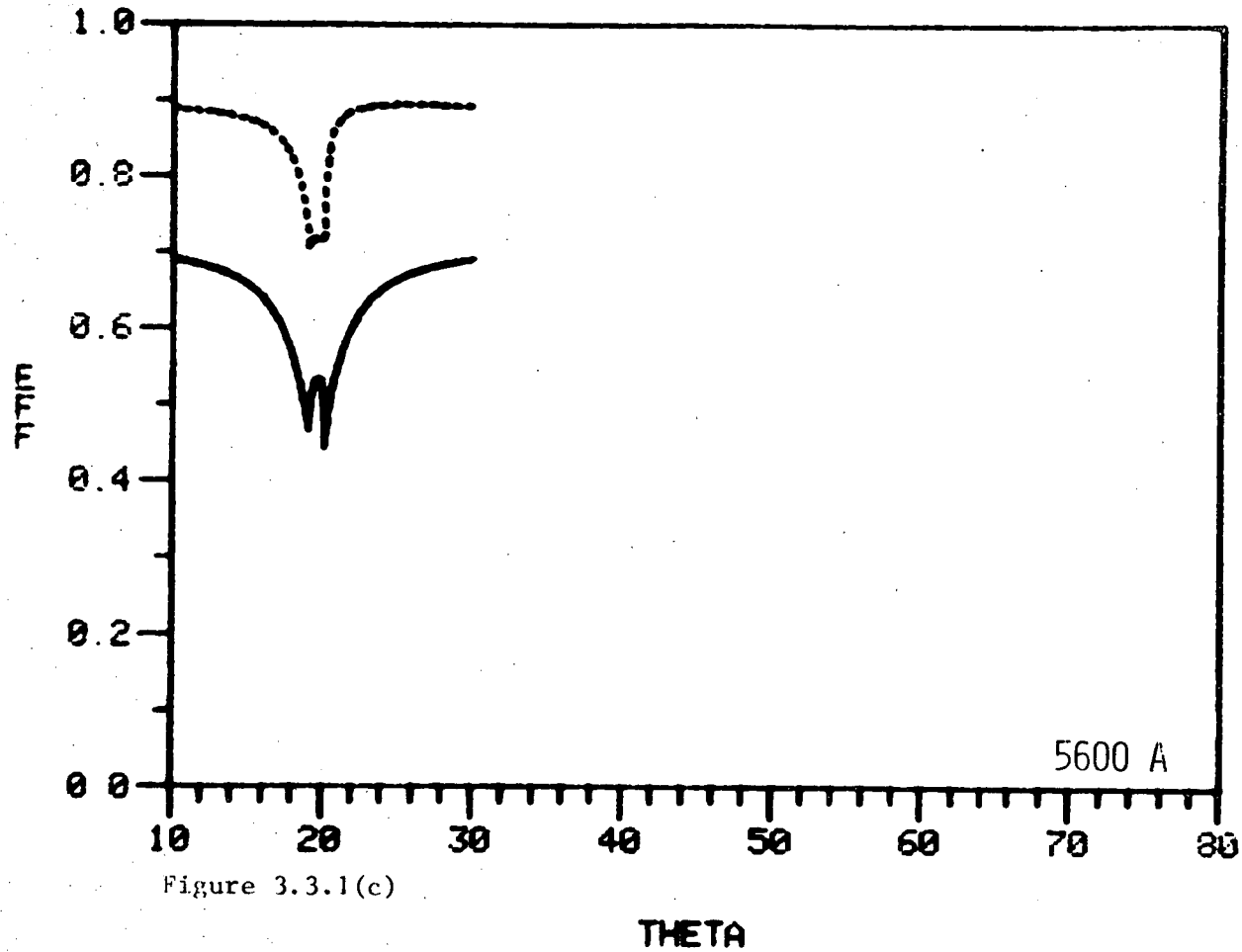


Figure 3.3.1(c)

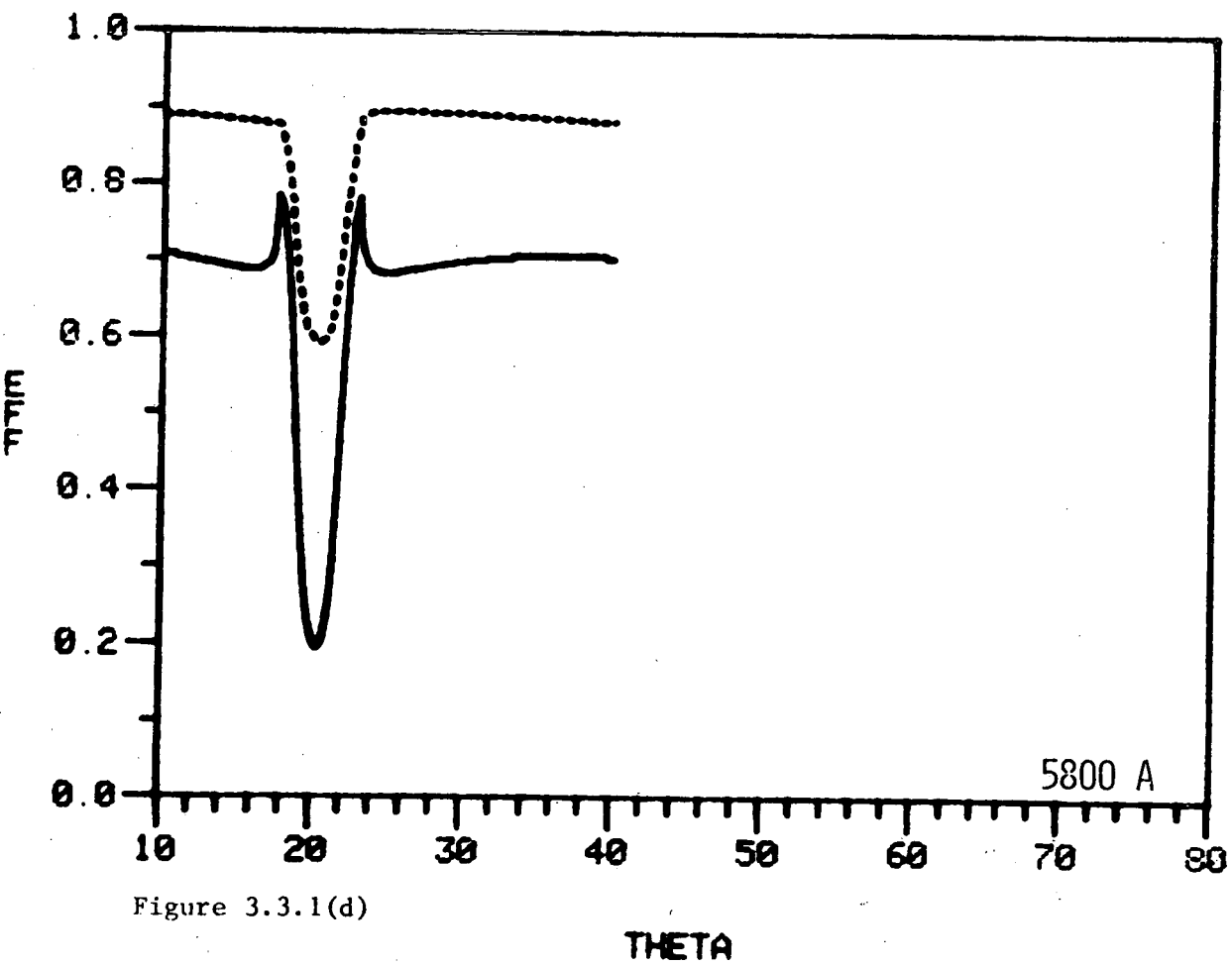


Figure 3.3.1(d)

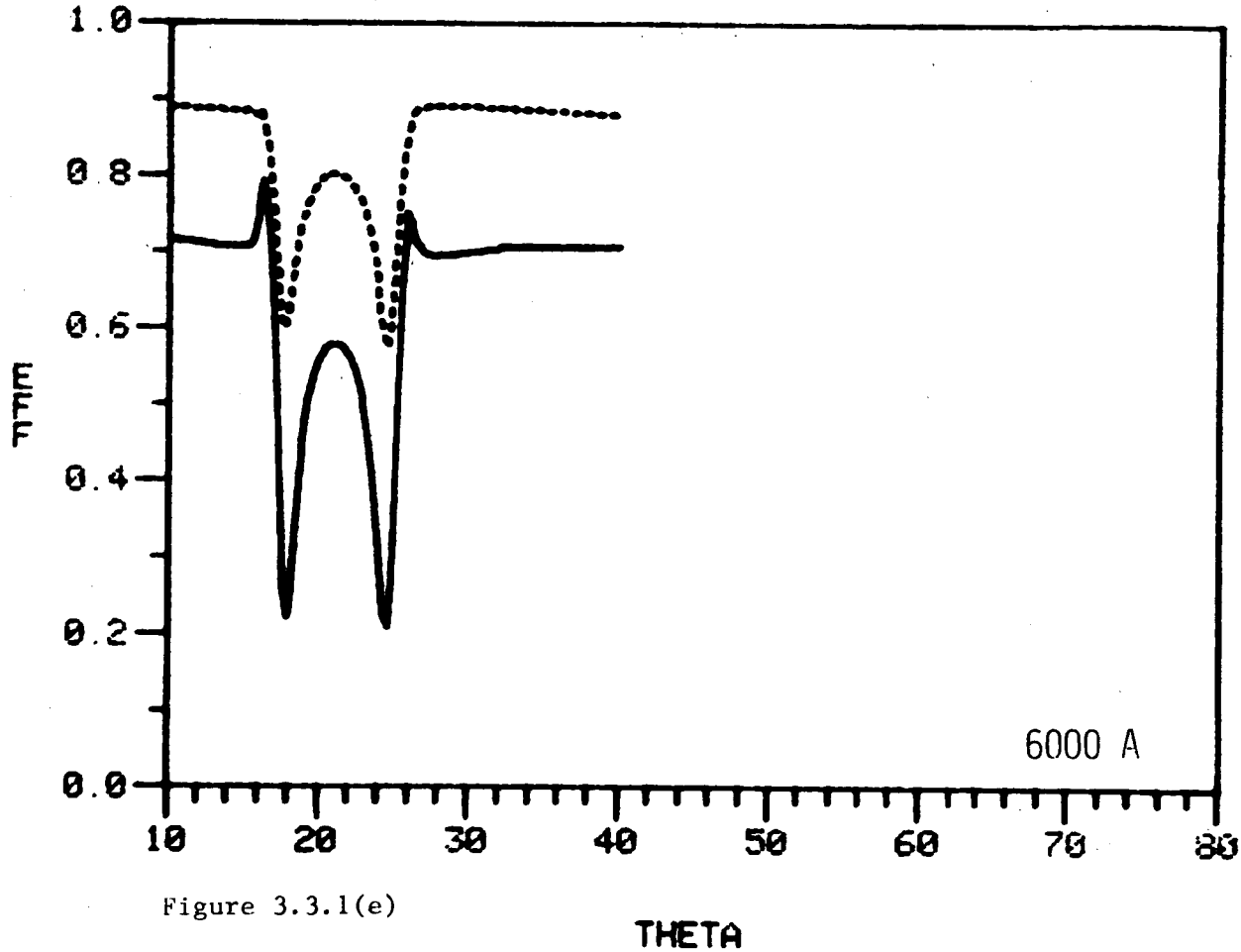


Figure 3.3.1(e)

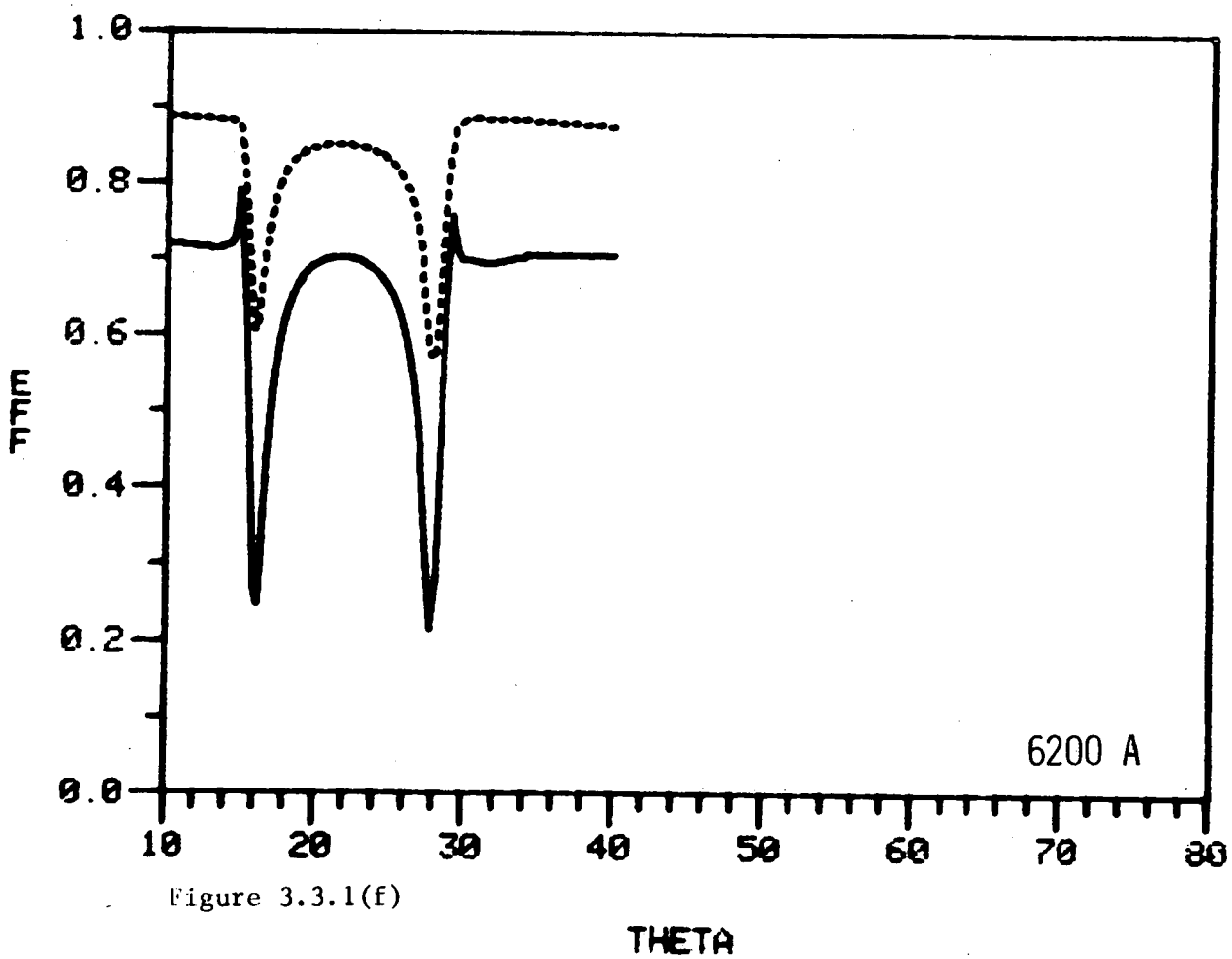


Figure 3.3.1(f)

Figure 3.3.2 S polarization zeroth order efficiency spectra for a 600 g/mm. triangular profile grating having a blaze angle of $5^{\circ}10'$ and a right-angled apex ruled in aluminium. Graphs for wavelengths of

(a) 4700 Å

(b) 5500 Å

(c) 6200 Å

showing both the zeroth order efficiency (———) and the total diffracted energy (------) are presented.

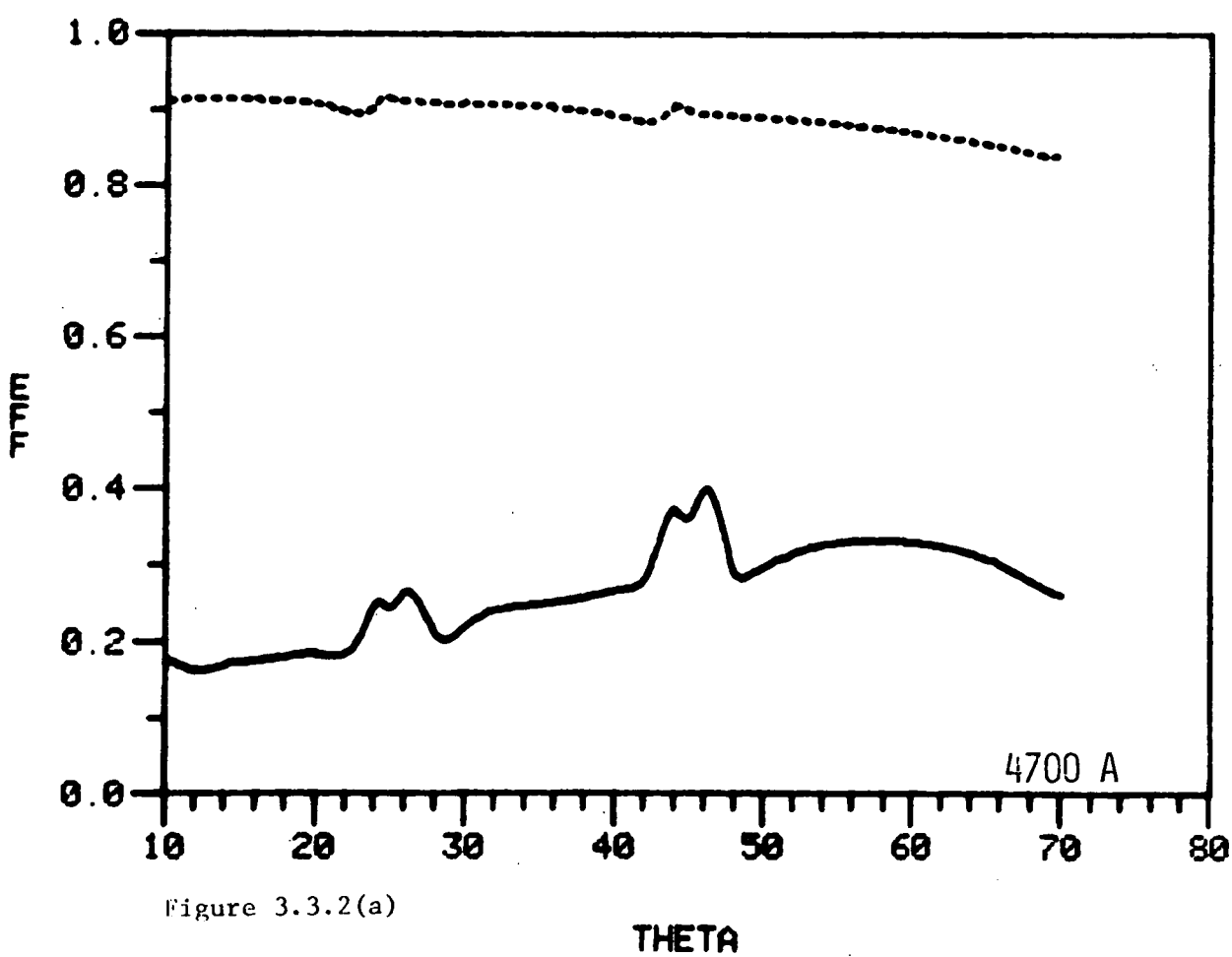


Figure 3.3.2(a)

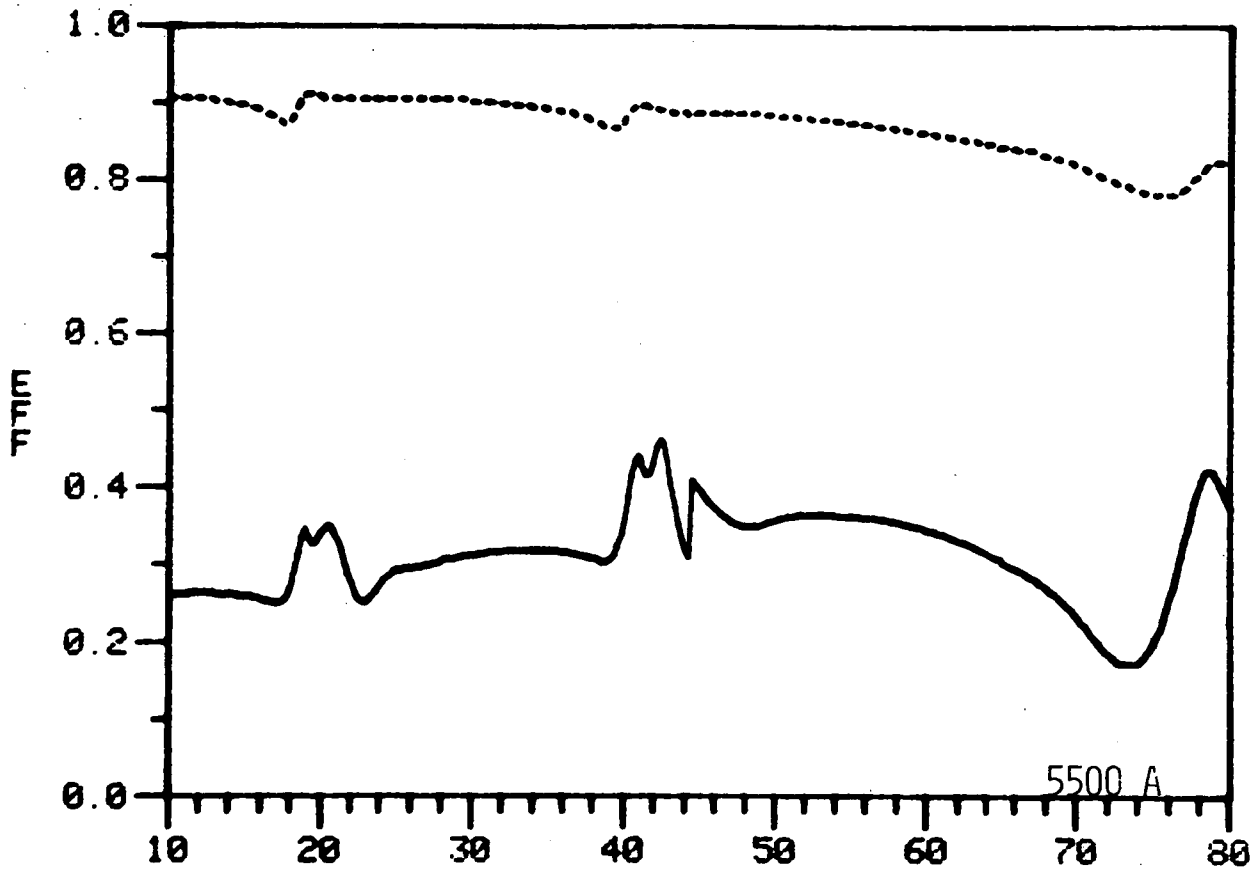


Figure 3.3.2(b)

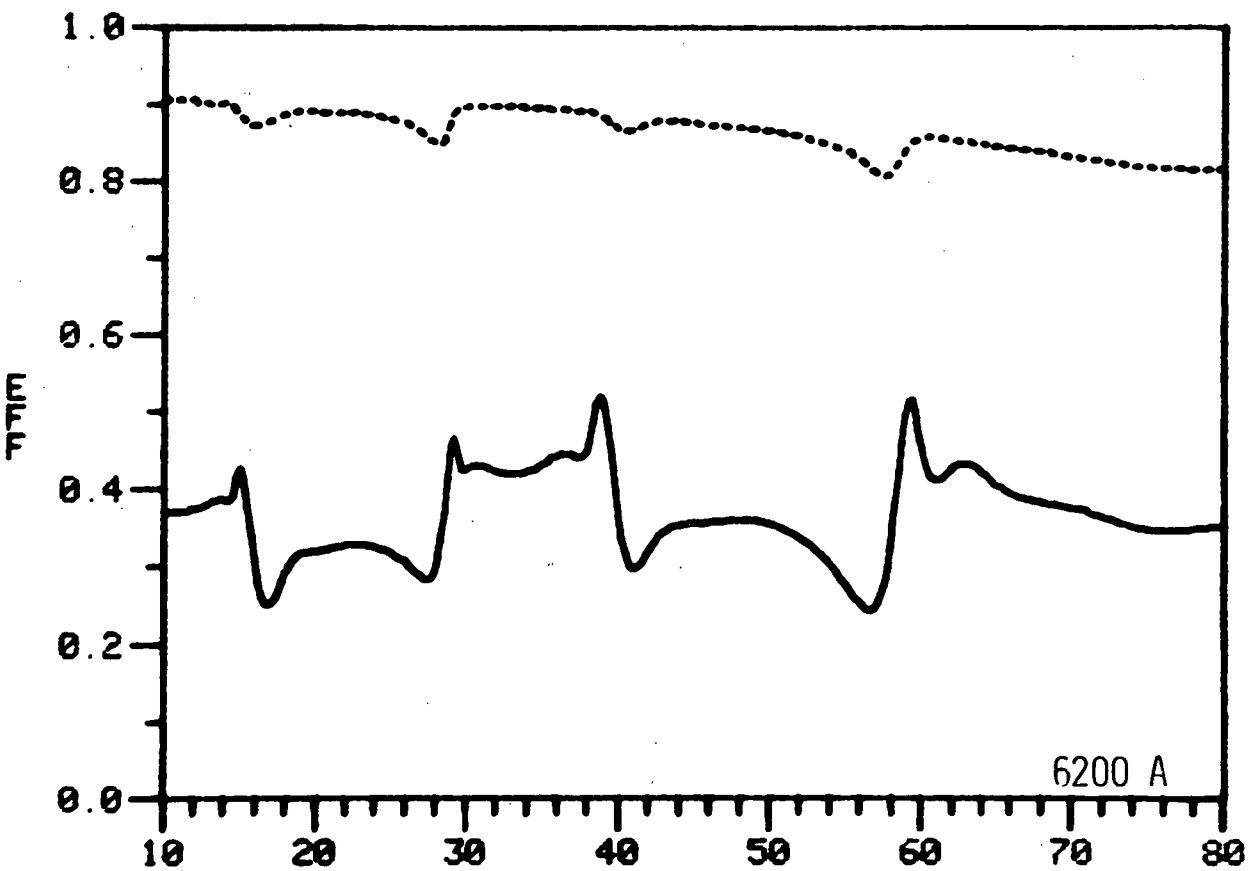


Figure 3.3.2(c)

experimental data can only be achieved if the fundamental parameters, (the optical constants and the profile shape) are regarded as variables. The source of this major shortcoming becomes evident upon a close inspection of this theory. As a consequence of the use of first-order theory, the dispersion relations of the pole for both the grating and the plane interface are identical. Consequently, the optical constants need be adjusted if both the position and the shape of the theoretically predicted resonance is to match the experimental results (since for real gratings, the resonance position is a function of groove depth).

Although the rigorous diffraction theory is not afflicted with such fundamental shortcomings, its solution is not obtained in closed form and hence the interpretation of the plasmon anomaly is essentially camouflaged. In an attempt to remove this hindrance to the complete understanding of the problem, Maystre and Nevière [3.7] modified their algorithm to accommodate inhomogeneous plane waves and thus were able to calculate the trajectories of both poles and zeros (in the complex sin θ plane) for various field quantities. In doing so, they achieved a particularly elegant characterization of the physical nature of the plasmon. Their study, restricted to sinusoidal gratings, revealed that the pole of the zeroth order field coefficient (which was originally stationed just above the Re(sin θ) axis for a plane interface) moved away from the axis in such a way that Im(sin θ_p) increased monotonically with groove depth (θ_p being the complex angle at which the pole occurred). On the other hand, the zero (which coincided with the pole for zero groove depth) moved monotonically downwards (in the sin θ plane).

This behaviour must be expected on physical grounds, since if the pole did cross the real axis then a catastrophic situation would occur. For one particular groove depth, the zero can lie precisely on the real axis and hence, for the case of a single propagating order, the total absorption (i.e. the Brewster phenomenon) of a monochromatic beam is achieved. This has been substantiated experimentally by Hutley and Maystre [3.16].

Having demonstrated the existence of both poles and zeros, Maystre and Nevière were then able to show that the zeroth order field amplitude was proportional to

$$\frac{\sin \theta_i - (\sin \theta_Z - \frac{n\lambda}{d})}{\sin \theta_i - (\sin \theta_P - \frac{n\lambda}{d})}$$

in the vicinity of the passing off of the order n. (Here θ_i denotes the real angle of incidence and θ_P and θ_Z represent the complex angles of the pole and zero respectively. A continuation [3.8] of their analysis has also achieved a satisfactory explanation of the anomalous behaviour [3.17] exhibited in P polarized light by gratings conformally coated with a dielectric film.

From the above expression, it can be readily seen that the sharpness of any resonant behaviour manifesting itself in the real orders is dependent upon the proximity of the pole to the Re(sin θ) axis. The quantities θ_Z and θ_P are dependent not only upon the profile shape but also upon the optical constants of the metal. In an attempt to elucidate the refractive index dependence of the resonances, a number of calculations were performed, modelling the performance of a 5^o blaze angle

triangular profile ruled in fictitious metals having wavelength-independent optical constants. The results of this study are shown in figures (3.3.3) and (3.3.4) which are in reasonable qualitative agreement with curves calculated by Wheeler [3.9] using the theory of Elson. From these graphs, it is evident that the imaginary part of the refractive index dominates the sharpness of the resonances. It is evident that in the case of metals whose ratio n_i/n_r is large, the pole must lie close to the real axis. Although no attempt was made to characterize this behaviour by actually calculating the pole position, one can still gain some understanding of these phenomena by using, as a first approximation, the dispersion relation appropriate to plane surfaces (and hence the first order perturbation theory).

The absorption in the zeroth order, shown in figures (3.3.3-4) corresponds to the passing off of the $(+1)^{th}$ order of diffraction. Thus, according to the plane interface dispersion relation, the resonance minimum should be centred on

$$\sin \theta_i = \mathrm{Re}\left(\sqrt{\frac{\tilde{\varepsilon}}{1+\tilde{\varepsilon}}}\right) - \frac{\lambda}{d} .$$

In the case of $n_i >> n_r$, it may be shown that

$$\sqrt{\frac{\tilde{\varepsilon}}{1+\tilde{\varepsilon}}} \simeq \left[1 + \frac{n_r^2 + \tfrac{1}{2}}{n_i^2}\right] + i \left[\frac{n_r^2 + \tfrac{1}{2}}{n_i^2}\right] \frac{n_r}{n_i} .$$

Although this representation is incapable of exactly predicting the position of the resonance, it does, nevertheless, produce reasonable qualitative agreement with the data. For a constant value of n_r, the

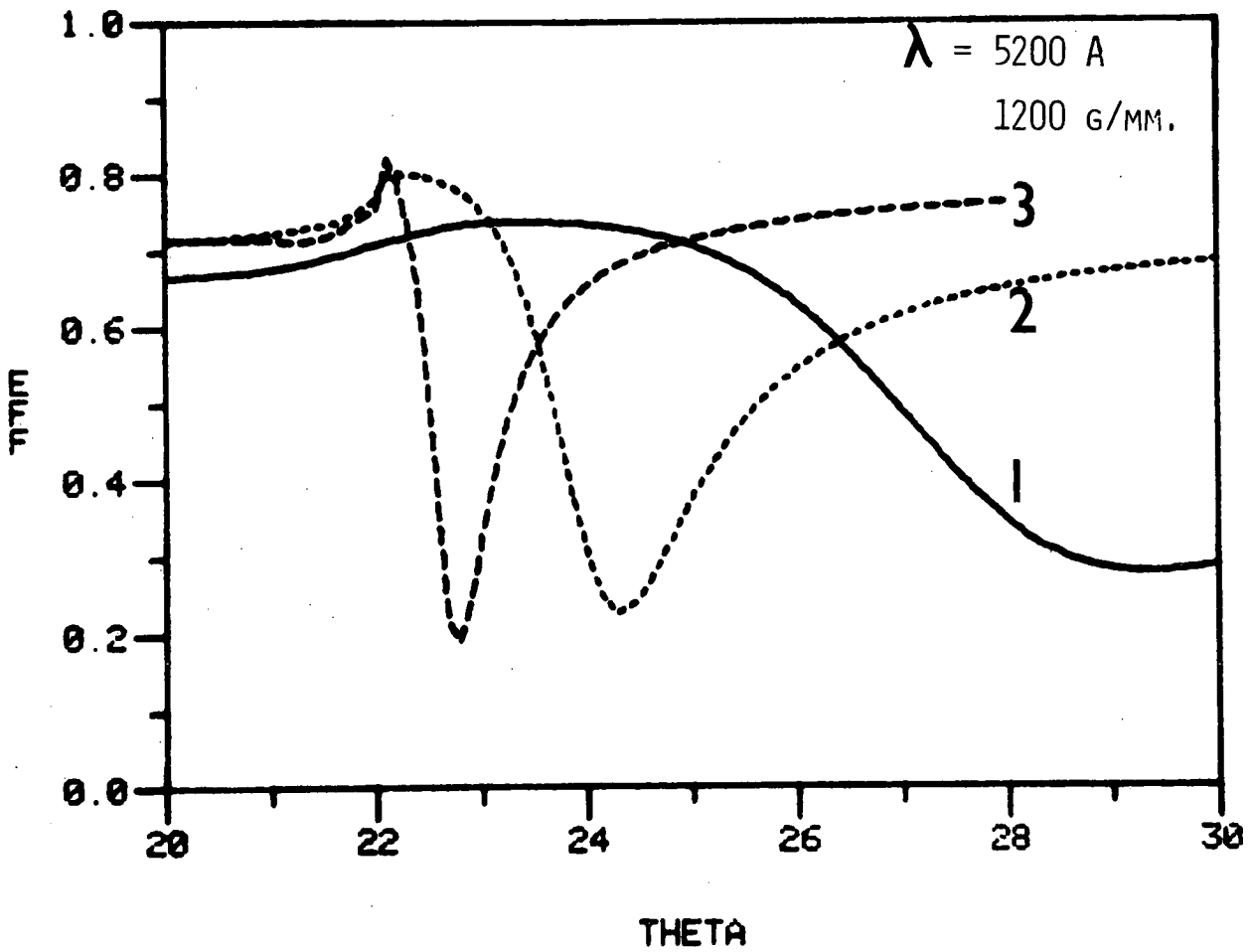


Figure 3.3.3 The variation of the 0^{th} order efficiency spectra with the imaginary part of the refractive index. The grating, characterized by a 5^{o} blaze angle and a right-angled apex is ruled in three fictitious metals of wavelength independent refractive index:

1. $\tilde{n} = 0.2 + i\ 3.0$
2. $\tilde{n} = 0.2 + i\ 5.0$
3. $\tilde{n} = 0.2 + i10.0$

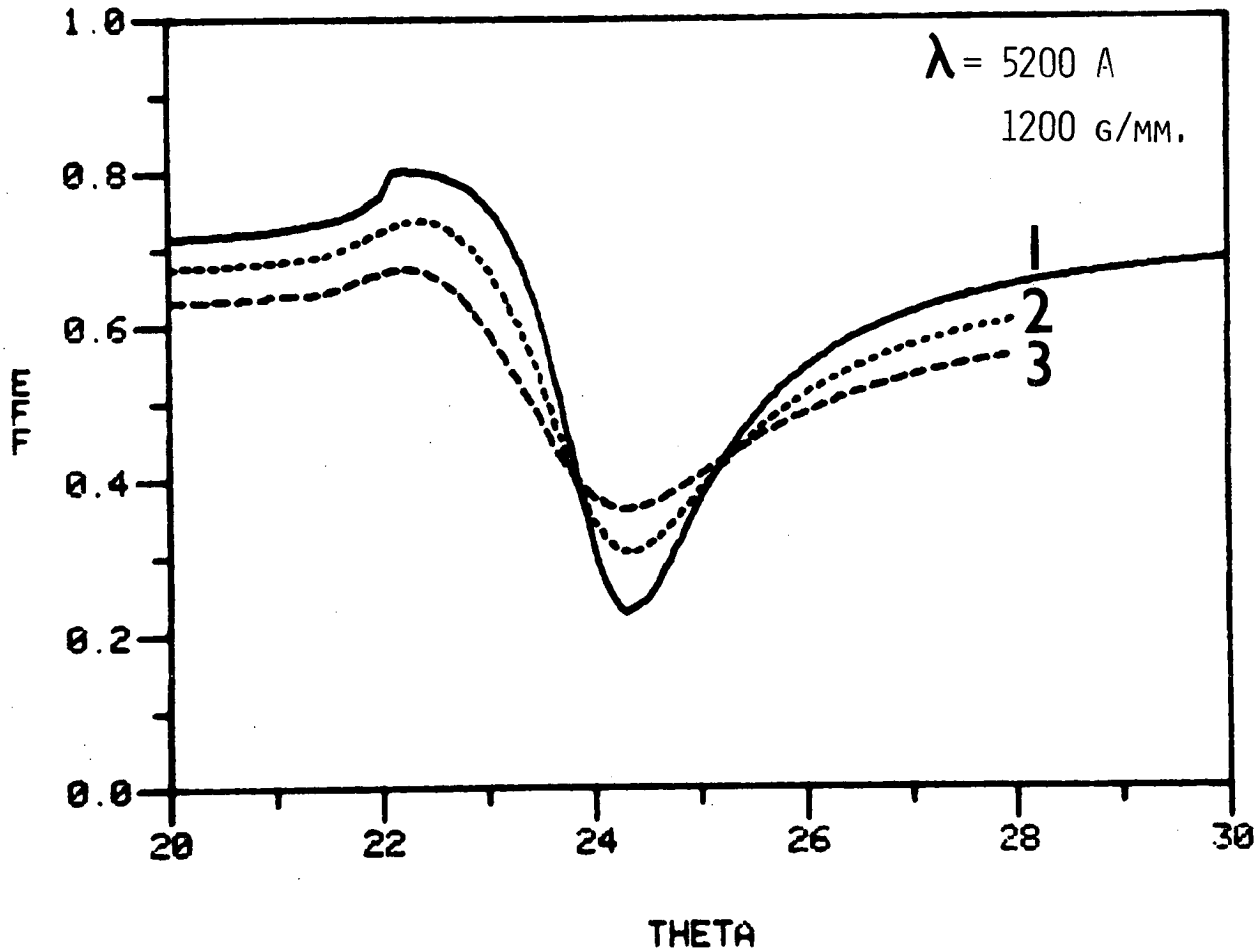


Figure 3.3.4 The variation of the 0^{th} order efficiency spectra with the real part of the refractive index. The grating, characterized by a 5^{o} blaze angle and a right-angled apex is ruled in three fictitious metals of wavelength independent refractive index:

1. $\tilde{n} = 0.2 + i\ 5.0$
2. $\tilde{n} = 0.5 + i\ 5.0$
3. $\tilde{n} = 0.9 + i\ 5.0$

pole approaches the real axis as n_i^{-3}, thereby explaining the marked decrease in bandwidth as n_i becomes large. Furthermore, the pole moves back towards the Wood anomaly as $\frac{1}{n_i^2}$. For the curves corresponding to a constant value of n_i, a slower decrease in plasmon bandwidth as n_r tends to zero is observed. This is evidenced by the term $n_r(n_r^2 + \frac{1}{2})$ which also explains the slowness of the resonance in its move back to the Wood anomaly.

In summary, it must be clear that the rigorous diffraction theories have far surpassed the perturbation treatments, as far as predicting and understanding the nature of grating anomalies is concerned. However, it must be stated that our understanding is still very incomplete in that the original problem has only been transformed into another problem whose solution is more easily interpreted. Not until such time as the pole and zero trajectories can be predicted analytically (rather than numerically) will it be possible to achieve full insight into the nature of grating absorption features.

3.4 GROOVE DEPTH DETERMINATION OF SINUSOIDAL PROFILE GRATINGS USING A LASER

This section discusses a new experimental application of the finite conductivity theories. These investigations were conducted in collaboration with Dr. Ian Wilson of the University of Tasmania and have been reported in a paper published by "Applied Optics".

3.4.1 Introductory comments

In 1967, Labeyrie [3.18] in France and Rudolph and Schmahl [3.19] in Germany demonstrated the feasibility of making useful diffracting

structures in photoresist. Of particular interest are the plane diffraction gratings and the various types of focusing gratings that have evolved following the work of these early researchers. Notable advances in fabrication techniques have been made in the past few years, especially by M. C. Hutley, National Physical Laboratory, England, Rudolph and Schmahl, Göttingen Solar Observatory, F.D.R and workers at Jobin-Yvon, France. It is now possible to make photoresist or holographic gratings with spectroscopic properties comparable with and in some cases superior to those of the best ruled gratings.

Such advances in the development work on holographic gratings have been accompanied by an equally rapid evolution in the theoretical research. Thus, there has been considerable incentive and scope for using the recently developed grating theories to guide experimental physicists in their choice of exposure parameters that will yield gratings of high spectral performance. In 1970, Maystre and Petit [3.20] used the rigorous infinite conductivity theory of Pavageau and Bousquet [3.21] to investigate numerically the efficiency characteristics of plane reflection gratings made in a logarithmic negative photoresist. Their work was followed by a contribution from McPhedran et al. [3.22] concerning the blaze optimization of holographic gratings having sinusoidal groove shapes. From this study, it was established that high spectral performance is critically dependent on the depth of the sinusoidal profiles. Further investigations [3.23, 3.24] using the finite conductivity theory of Maystre [3.4] confirmed the validity of the optimum groove depths predicted in earlier infinite conductivity theory investigations.

Many applications in which diffracting structures are used do not necessarily require the high spectral performance sought after by designers of fast grating instruments. One such application, which is rapidly increasing in importance, is the use of the grating coupler in the field of integrated optics. Already, experimental and theoretical studies have indicated the importance of having a precise knowledge of the groove depth of the diffracting structure forming the grating coupler.

In the past, the groove profile of diffraction gratings has been established using delicate and time-consuming electron-microscope techniques. The electron-microscope has the resolution and the depth of field required to evaluate groove profiles, but its usefulness depends entirely on the proper presentation of the specimens. Pröger [3.25] appears to have made the first extensive use of the electron microscope to measure grating groove shapes. Further notable advances were made in 1965 by Anderson et al [3.26] who used a technique employing the evaporation of platinum past an asbestos whisker lying across the grooves. The resulting shadow cast by the whisker could then be used to determine the groove shape with some success. The most accurate groove determinations have been made using techniques similar to those described by Bennett [3.27]. This method employs a thin replica section at the grating surface bent over a pin so that the grooves may be viewed in silhouette. Care must be taken when using this method to select a groove that is representative of the grating and free from distortions introduced by preparation of the sample.

A technique that circumvents many drawbacks of electron-microscope groove profile determination employs a Rank Precision Industries Talystep. The Talystep has been used to advantage by Verrill [3.28, 3.29] and Hutley [3.30] to determine the groove shape of both ruled and holographic gratings. The principle of operation of this device is easily understood. A diamond stylus is drawn across the surface of the groove, and its vertical movement is converted to an electrical signal using a sensitive electromechanical transducer. The output may then be amplified and displayed on a chart recorder. A groove profile of a 1200 mm^{-1} grating may be measured using the Talystep in a matter of minutes.

The electron-microscope and Talystep techniques can be usefully employed for groove profile determination. However, another technique suggested by Beesley and Castledine [3.31] may be more attractive. They suggested that by measuring the diffracted efficiency of a sinusoidal groove grating and comparing this value with theory, it should be possible to estimate the groove depth. The concept is soundly based except that Beesley and Castledine attempted to implement the idea using the scalar optics theory of Kogelnik [3.32].

3.4.2 Laser determination of groove depth - Sinusoidal grooves

In our laboratory, Beesley and Castledine's approach has been explored using a gas laser, in a very simple experimental arrangement, coupled with a facility for computing grating efficiencies using a rigorous finite conductivity theory. Rather than an electron beam or mechanical probe, this technique involves the use of the electromagnetic

theory and a laser to determine the groove depth of a reflection grating. In this section, the potential use of the method to the important group of reflection gratings having sinusoidal grooves will be discussed. The gratings considered are coated with evaporated aluminium of complex refractive index as measured by Hunter [3.33] and have groove frequencies of $1200\,mm^{-1}$ and $1800mm^{-1}$. The electromagnetic theory used is the rigorous formalism of Maystre and the results presented here, obtained from our implementation of his theory, have been estimated to be accurate to ± 2%.

The essence of the gas laser technique for measuring groove depths is the use of the laser and a suitable detector to measure the diffracted efficiency at certain angles of incidence in the two fundamental planes of polarization, and compare these measured efficiencies with those obtained numerically. With the judicious choice of both the angle of incidence and the wavelength of the laser, it is possible to measure the diffracted efficiency under circumstances that are very sensitive to changes in groove depth. Natural choices for the optical arrangements in which to execute the efficiency measurements are the Littrow and normal incidence configurations. Both of these arrangements are ideally suited to the purpose since they can be achieved with ease, experimentally. In the Littrow arrangement, the first order beam is diffracted back along the path of the incident laser beam. In the normal incidence case, the zero order beam is similarly diffracted back along the path of the incident beam. If the distance between the laser and the grating under test is sufficiently large, both the Littrow and normal incidence configurations may be realized to a high degree of precision.

Having decided upon the optical configuration to be employed for the efficiency testing, it is now necessary to select a test wavelength for which the Littrow and normal incidence efficiencies are very sensitive to changes in groove depth. Initially, it was hoped that the 0.6328μm He-Ne laser line would be suitable for testing optical gratings. However, the numerical studies indicated that for sinusoidal groove gratings the most sensitive normalized wavelength (λ/d) range is $0.70 \leq \lambda/d \leq 0.85$. This means that the He-Ne laser line is only useful for testing gratings with groove frequencies near $1200mm^{-1}$. For groove frequencies in the vicinity of $1800mm^{-1}$ it was necessary to choose the 0.4579μm argon ion laser line.

With the optical configuration and efficiency testing wavelength established, all that is necessary is to now measure the zeroth order efficiency, E(0), in P and S polarized light for the case when one of the first orders is in a Littrow configuration. As a double check, similar efficiency measurements are made on one of the first order diffracted beams, E(-1), while normal incidence of the laser beam prevails. These efficiency measurements are then compared with the numerical results of the next section to yield the grating groove depth, h.

The whole experimental procedure takes only a matter of minutes to perform. In our laboratory, a Spectra Physics Model 404 laser power meter with digital readout is used to measure the absolute efficiencies of the zero and first order beams. It is usually necessary to have some control over the incident laser power, a problem which is easily solved using a variable beam splitter. In order to rotate the plane of polarization of the laser beam, a $\frac{1}{2}\lambda$ plate followed by a polaroid filter is

used. Commercially available polarization rotators or Kerr cells could serve equally well. Using such apparatus, it is estimated that the experimental efficiencies can be measured to within ± 1%.

3.4.3 Numerical results for 1200mm^{-1} and 1800mm^{-1} sinusoidal groove gratings

Examples of the numerical results obtained using our computer program are shown in figures (3.4.1-2). In figure (3.4.1) is shown the absolute efficiency of the order zero as a function of the angle of incidence (θ_i) for the two fundamental planes of polarization and for various groove depths (h). The important angle of incidence occurs when $\theta_i = \theta_\ell \simeq 24.3^\circ$. This is the Littrow angle for the $(-1)^{th}$ order, occurring where efficiencies are varying rapidly with groove depth. Similarly in figure (3.4.2) the efficiency changes with groove depth for the order - 1 are depicted. In this figure, the important angle of incidence is $\theta_i = 0^\circ$, corresponding to rapid changes in the efficiency of the zeroth order.

The behaviour of the absolute efficiency with groove depth at these two important angles of incidence are conveniently summarized in figure (3.4.3). For the 1200mm^{-1} sinusoidal groove grating, the numerical results are similar to those just presented for the 1800mm^{-1} grating. For this reason, only the results pertinent to the final step in the groove depth determination are shown in figure (3.4.4). (It should be remembered that these results are for an efficiency testing wavelength of 0.6328μm.)

When the grating grooves are sufficiently distorted away from the sinusoidal profile, one would expect the form of the diffracted field to

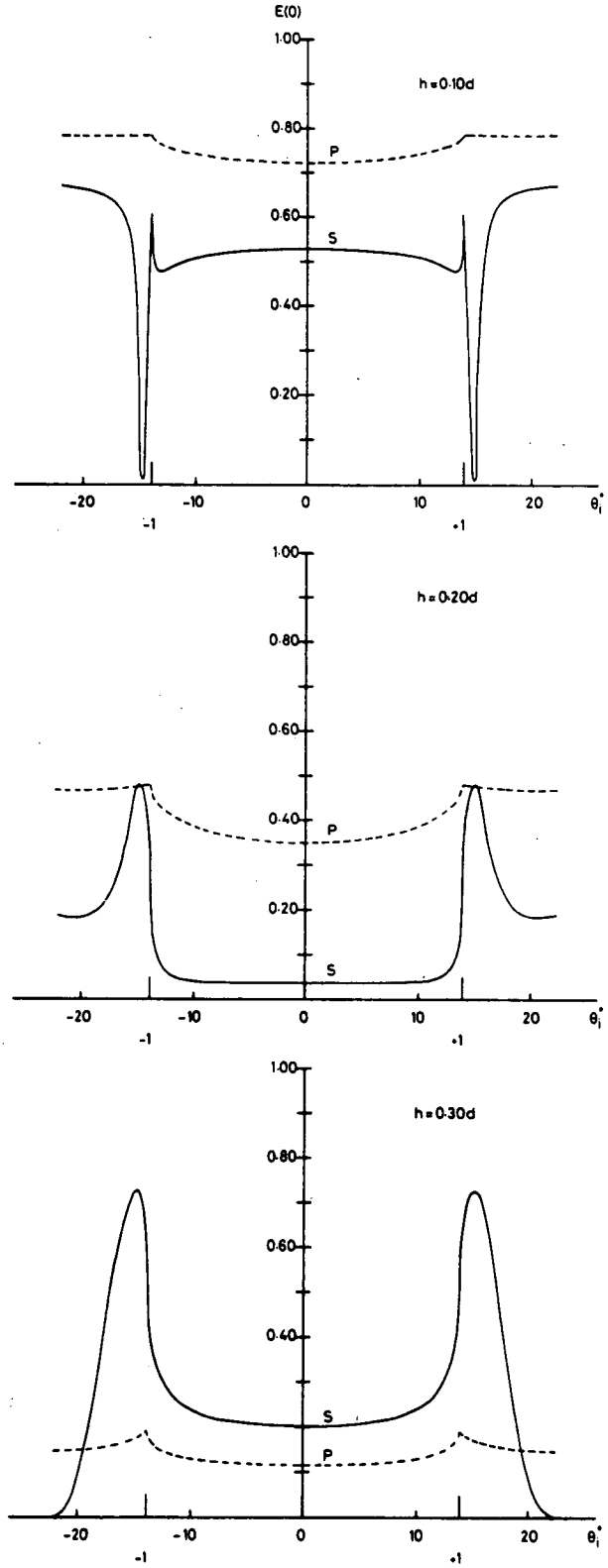


Figure 3.4.1 Zero order efficiencies as a function of the angle of incidence θ_i^o for various groove depths h. The sinusoidal groove grating has an evaporated aluminium surface, 1800 grooves/mm, and λ = 0.4579 μm.

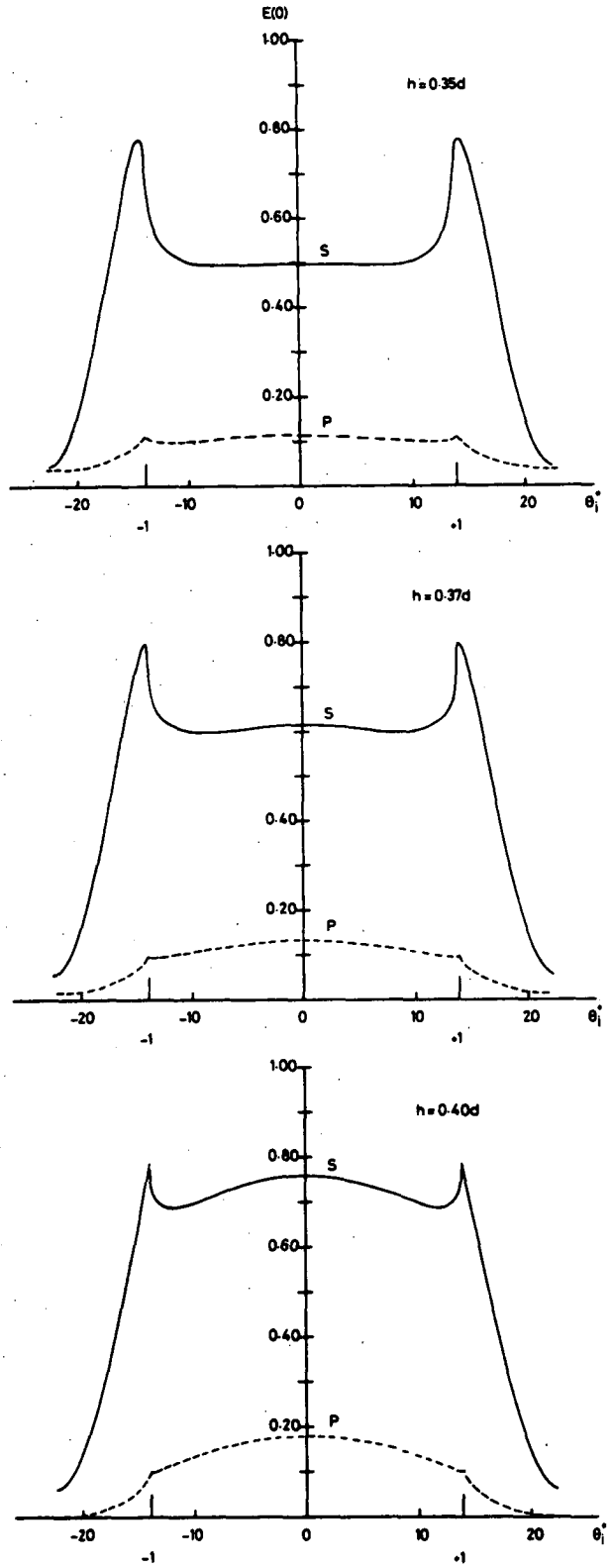


Figure 3.4.1 ...continued.

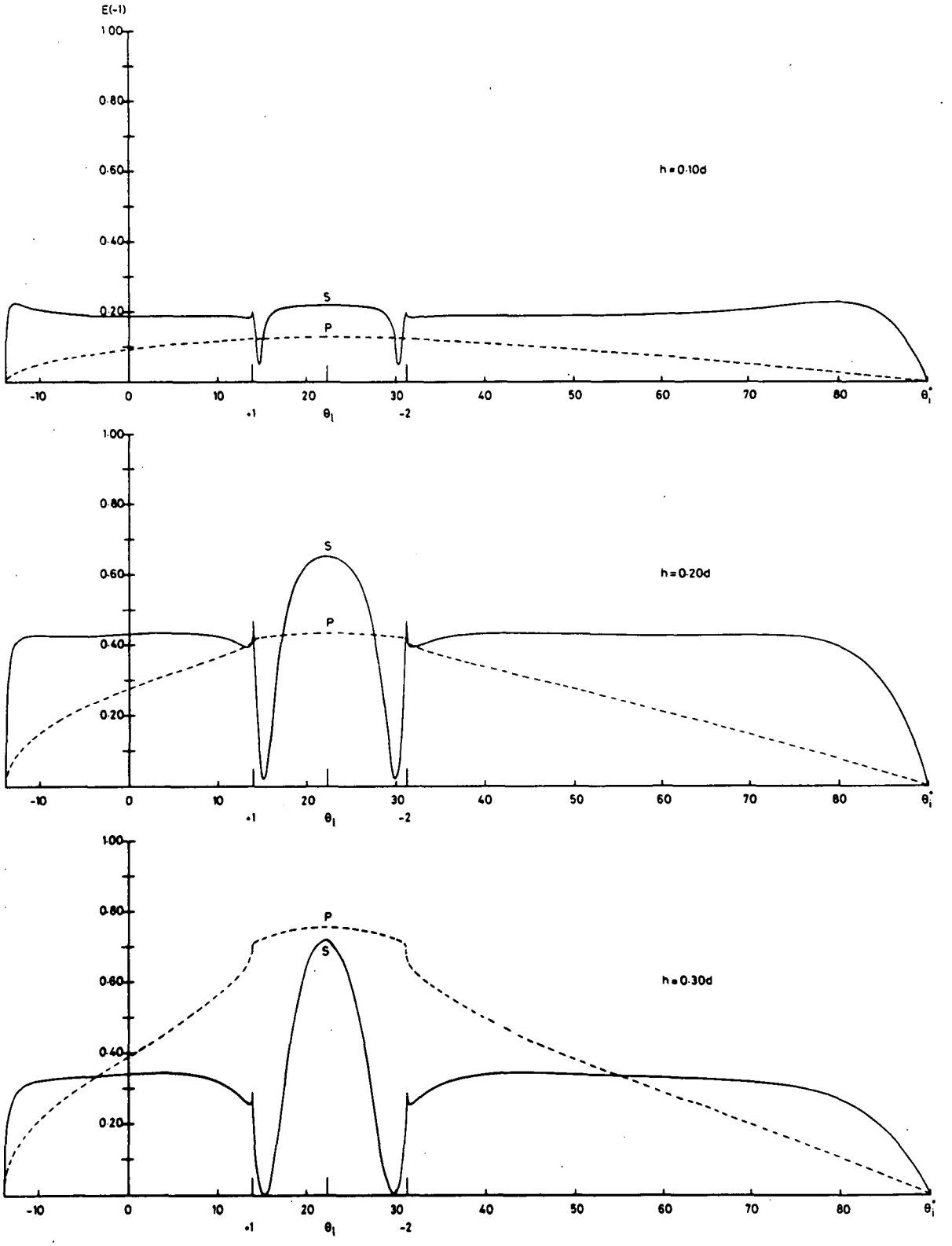


Figure 3.4.2 Order - 1 efficiencies as a function of the angle of incidence θ_i^o for various groove depths h. The sinusoidal groove grating has an evaporated aluminium surface, 1800 grooves/mm, and λ = 0.4579 μm.

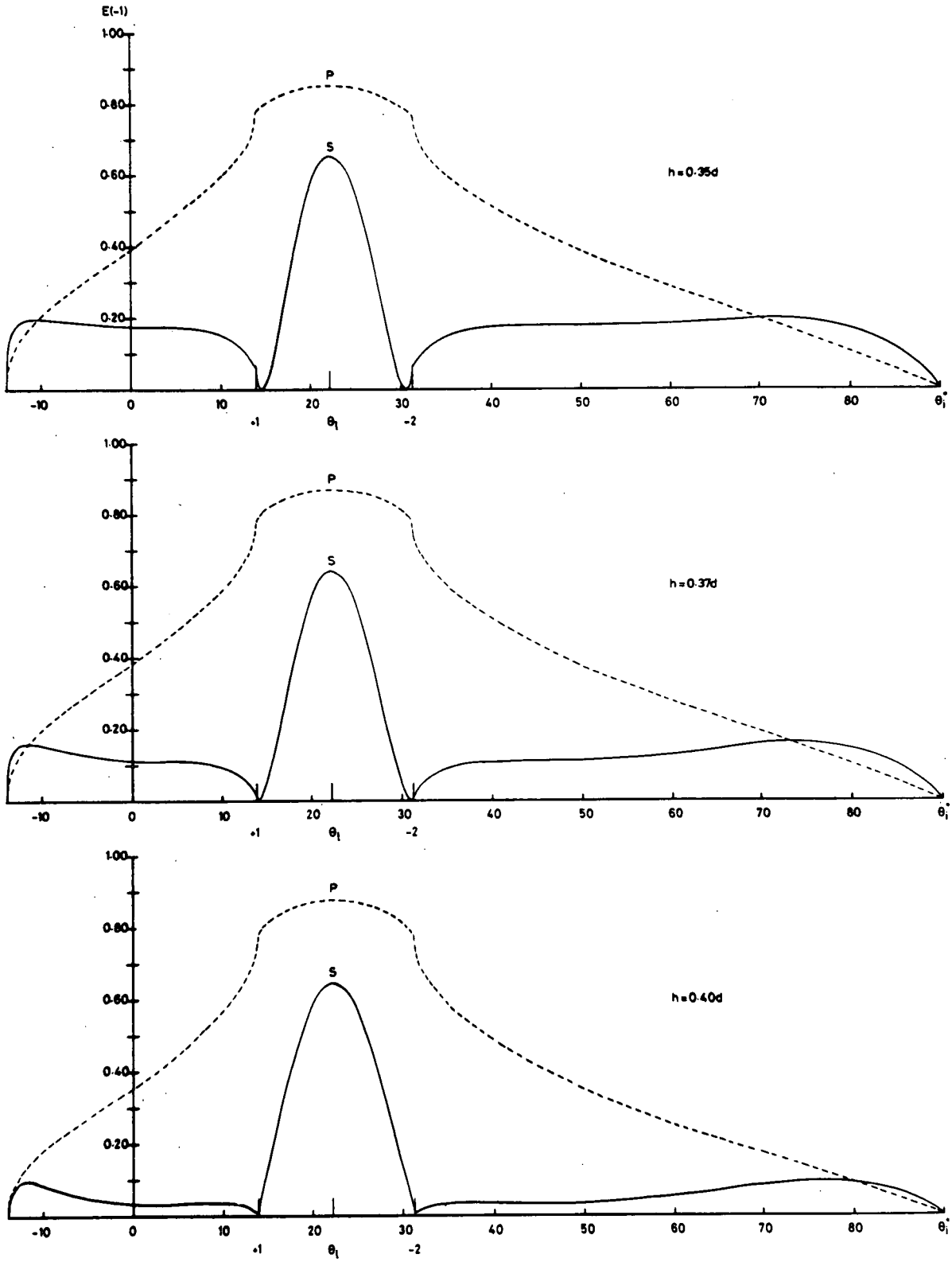


Figure 3.4.2 ... continued.

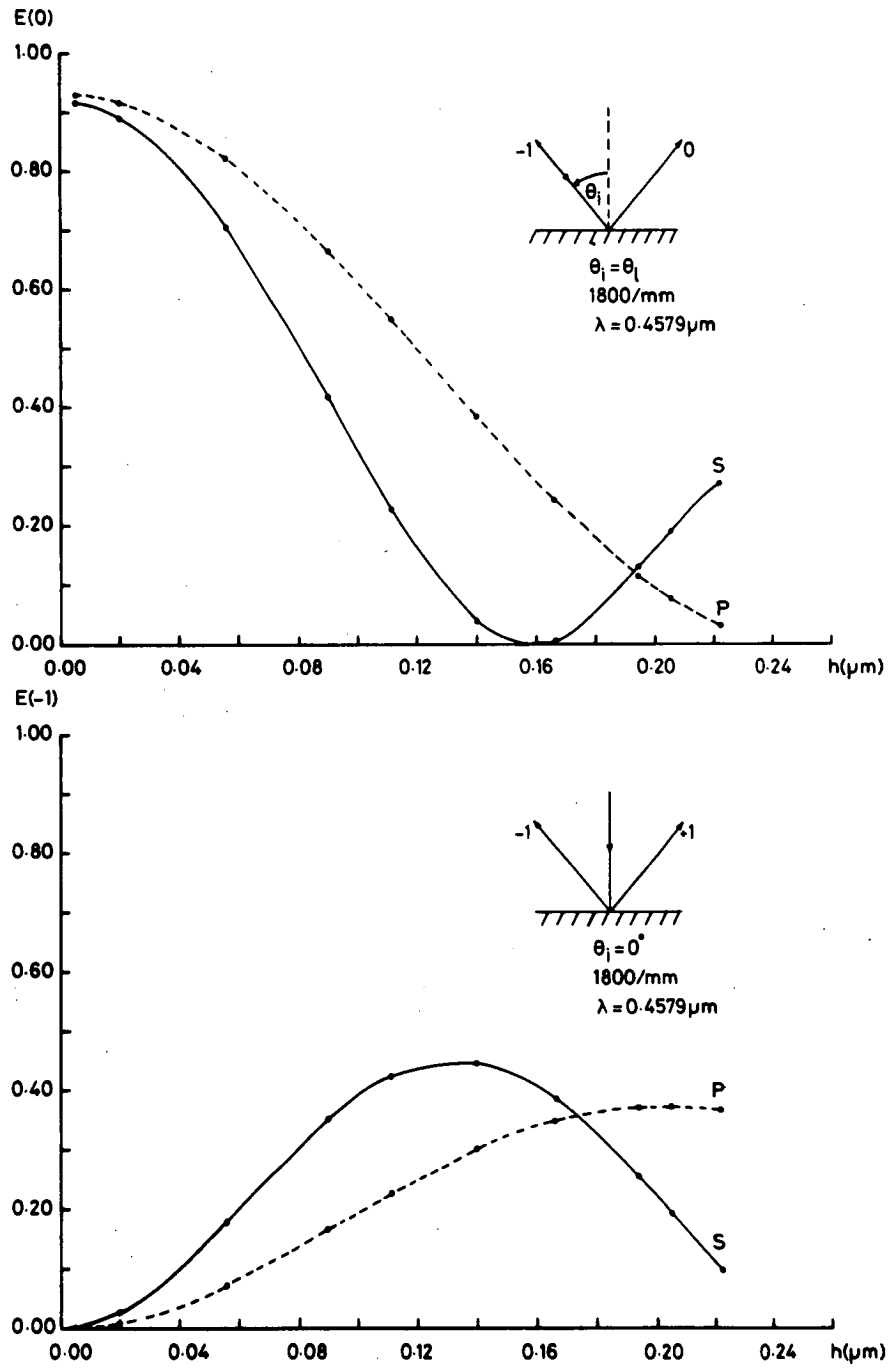


Figure 3.4.3 Littrow and normal incidence efficiencies as a function of groove depth h for an 1800 groove/mm sinusoidal groove grating with λ = 0.4579 μm.

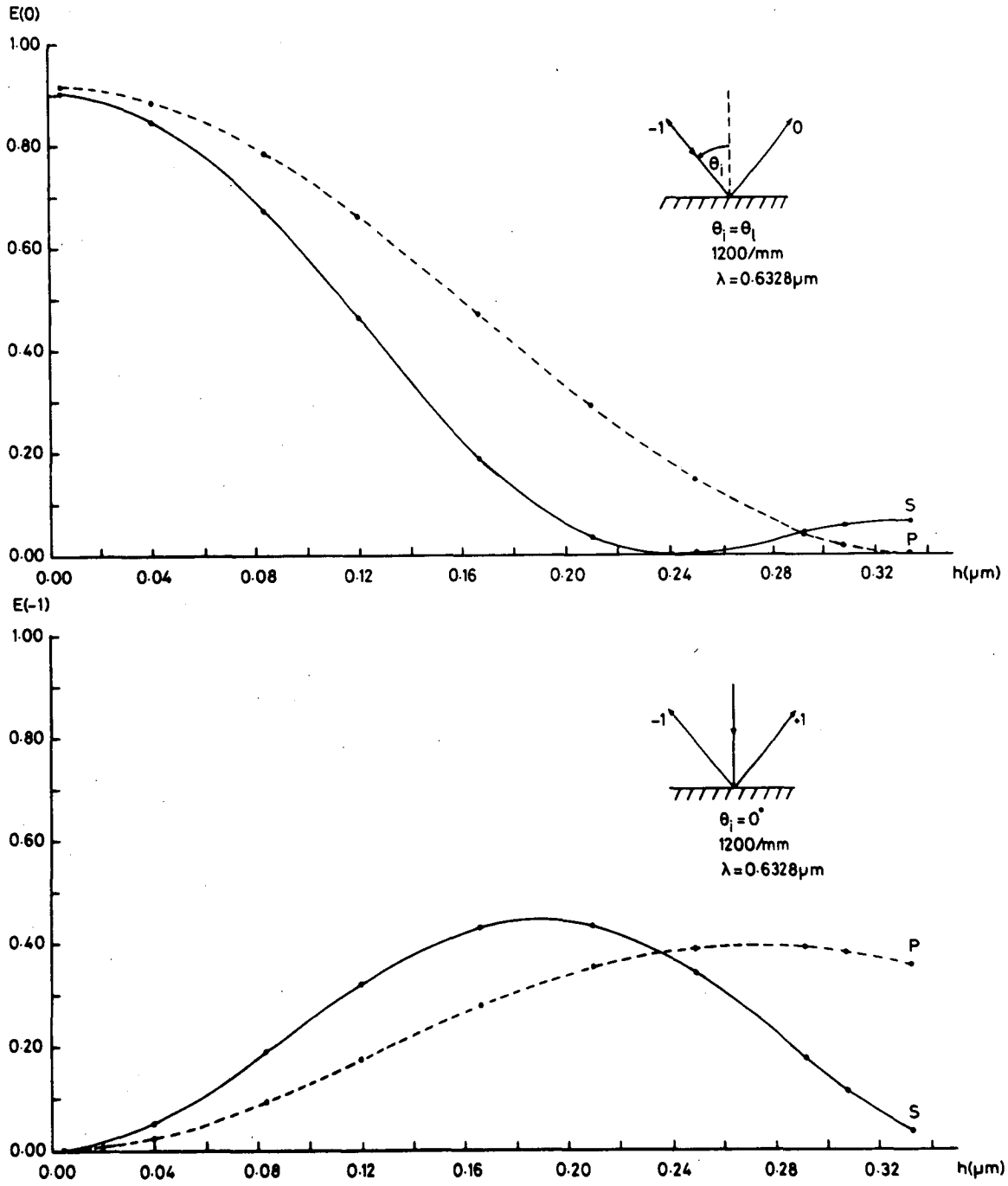


Figure 3.4.4 Littrow and normal incidence efficiencies as a function of groove depth h for a 1200 groove/mm sinusoidal groove grating with λ = 0.6328 μm.

reflect this change. This has indeed been found to be the case as shown by Wilson et al [3.34]. It would be useful to make some assessment of the error incurred in assuming a perfect sinusoidal groove geometry when some degree of distortion is present. Unfortunately it is difficult to quantify this effect because of the variety of groove distortions that may be possible. However, for the type of photoresist distortion described in reference [3.34] it is only for low pre-exposures (0-100 mJ/cm^2) and hence moderate to high groove distortions that appreciable departures from the results for perfect sinusoidal grooves occur. It is just these departures which may alert the experimenter to the presence of some degree of groove distortion. When significant groove distortion is present, it is necessary to examine in detail the extended features of the efficiency curve in order to determine the groove depth. Numerical results for distorted sinusoidal profile can be found in reference [3.35].

3.4.4 Concluding Remarks

In the previous discussion, the usefulness of the rigorous electromagnetic theories in guiding experimenters towards the attainment of a desired groove depth has been demonstrated. Although this section has dealt only with the case of sinusoidal groove profiles, the entire method suggests itself as a viable alternative to the Talystep and electron-microscope for the solution of a wide range of practical problems.

REFERENCES

[3.1] D. Maystre and R. Petit, Opt. Commun., 17 (1976) 196.

[3.2] A. Neureuther and K. Zaki, Alta Frequenza, 38 (1969) 287.

[3.3] P. M. Van den Berg, Ph.D. Thesis, (1971), Delft University of Technology, Delft, The Netherlands.

[3.4] D. Maystre, Thèse No., AO 9545, (1974), L'Université d'Aix-Marseille III, France.

[3.5] G. Cerutti-Maori, R. Petit and M. Cadilhac, C. R. Acad. Sci. Paris, 8B (1969) 1060.

[3.6] M. Nevière, P. Vincent and R. Petit, Nouv. Rev. Opt., 5 (1974) 65.

[3.7] D. Maystre and M. Nevière, J. Optics (Paris), to appear.

[3.8] M. Nevière, D. Maystre and P. Vincent, J. Optics (Paris), 8 (1977) 231.

[3.9] C. E. Wheeler, E. T. Arakawa and R. H. Ritchie, (1976), Oak Ridge National Laboratory, Report ORNL/TM-5185.

[3.10] J. M. Elson, J. Opt. Soc. Amer., 66 (1976) 687.

[3.11] R. Petit, D. Maystre and M. Nevière, in Space Optics, ed. B. J. Thompson and R. R. Shannon (National Academy of Sciences, Washington D.C.) (1974) p. 667.

[3.12] R. C. McPhedran and D. Maystre, Proc. Sym. on Diffraction Gratings and Grating Instruments, Spec. Soc. Japan, Tokyo (1974), 13.

[3.13] M. C. Hutley, Optica Acta, 20 (1973) 607.

[3.14] M. C. Hutley and V. M. Bird, Optica Acta, 20 (1973) 771.

[3.15] A. Otto, Z. Physik, 216 (1968) 398.

[3.16] M. C. Hutley and D. Maystre, Opt. Commun., 19 (1976) 431.

[3.17] M. C. Hutley, J. F. Verrill, R. C. McPhedran, M. Nevière and P. Vincent, Nouv. Rev. Opt., 5 (1974) 65.

[3.18] A. Labeyrie, in "Proceedings of the Conference on Optics", Marseilles (Centre National d'Etudes Spatiales) (1967) Report 00015/PR/ED.

[3.19] D. Rudolph and G. Schmahl, Umsch. Wiss. Tech., 67 (1967) 225.

[3.20] D. Maystre and R. Petit, Opt. Commun., 2 (1970) 309.

[3.21] J. Pavageau and J. Bousquet, Optica Acta, 17 (1970) 469.

[3.22] R. C. McPhedran, I. J. Wilson and M. D. Waterworth, Opt. Commun., 7 (1973) 331.

[3.23] D. Maystre, R. Petit, M. Duban and J. Gilewicz, Nouv. Rev. Opt., 5 (1974) 79.

[3.24] R. C. McPhedran and I. J. Wilson, Japan. J. Appl. Phys., 14, Suppl. 14-1 (1975) 159.

[3.25] H. J. Pröger, Exp. Tech. Phys., 9 (1961) 236.

[3.26] W. A. Anderson, G. L. Griffin, C. F. Mooney and R. S. Wiley, Appl. Opt., 4 (1965) 999.

[3.27] J. M. Bennett, J. Phys. E, 2 (1969) 816.

[3.28] J. F. Verrill, J. Phys. E, 6 (1973) 1199.

[3.29] J. F. Verrill, Optica Acta, 23 (1976) 425.

[3.30] M. C. Hutley, J. Phys. E, 9 (1976) 513.

[3.31] M. J. Beesley and J. G. Castledine, Appl. Optics, 9 (1970) 2720.

[3.32] H. Kogelnik, in "Modern Optics", J. Fox, Ed. Polytechnic Press, Brooklyn, (1967).

[3.33] W. R. Hunter, U. S. Naval Research Laboratory, personal communication.

[3.34] I. J. Wilson, R. C. McPhedran and M. D. Waterworth, Opt. Commun., 9 (1973) 263.

[3.35] I. J. Wilson, Ph.D. Thesis, (1977), University of Tasmania.

CHAPTER 4

A NEW FORMALISM FOR TRANSMISSION GRATINGS

This chapter is based on a research report [4.1] and a paper [4.2] of the same title which has been accepted for publication in the journal, "Optica Acta".

4.1 INTRODUCTION

Prior to 1971, all rigorous diffraction grating theories relied upon the assumption that the surface of the structure was perfectly reflecting. The inappropriate nature of this assumption for dielectric transmission gratings, together with the inadequacy of agreement between experiment and infinite conductivity predictions for metallic reflection gratings in the visible, necessitated the formulation of new theories that take into account the conductivity of the diffracting media. Several such theories have been proposed and these fall into two broad categories, namely the integral formalisms of Van den Berg [4.3] and Maystre [4.4] and the differential formalism of Nevière et al [4.5].

The integral method developed by Maystre is intrinsically similar to that of Van den Berg yet involves the solution of only one quantity which links the two unknowns of the latter formulation. Van den Berg also analysed the problems of diffraction by a coated reflection grating and by a transmission grating. In each case, however, his method involved the solution of two field quantities.

Numerical implementation of the above theories has concentrated

largely upon two problems, the reflection grating and the grating coupler. The differential approach of Nevière has attempted to handle both but is plagued by numerical difficulties when dealing with highly conducting metals and S polarized radiation.

It is the aim of this chapter to develop an integral formalism capable of treating a wide range of structures consisting of a single grating surface and a series of plane interference films placed above and below it. The formalism may be regarded as applying a 'bridge' between the efforts of Van den Berg and Maystre. Its solution involves the computation of only one unknown, analogous to the infinite conductivity surface current density.

The method has been designed to supply the ability to solve diffraction problems in a number of fields with its generality being such that it satisfies many of the requirements of grating spectroscopists. It has also been used in the field of solar energy research with an application of the theory pertinent to this end being presented in the following chapter.

The development of the method forms part of efforts, currently underway in Australia and France, designed at providing formulations which can deal with the most general grating structures capable of manufacture. In this context, D. Maystre [4.6] in Marseille has recently developed a theory of even greater generality than that discussed in this chapter. His formalism is capable of solving the diffraction problem involving a multiple-profile grating.

Also discussed in this chapter are four theoretical tests which have been used to verify the formalism and its associated computer

program. A review of the concepts of energy conservation, and reciprocity is presented together with a discussion of the a priori agreement of the grating formalism and Fresnel's Laws in the case of zero groove depth. Particular reference is made to the derivation of a new constraint relating the phases and efficiencies of the propagating orders excited when a lossless symmetric transmission grating is operated in a first order Littrow mount. Some sample numerical results, from all of the above tests, supporting the theory are also given.

4.2 CONCEPTS OF THE METHOD

Consider the diffraction of a plane wave by a grating depicted in figure (4.1a). The problem which has been solved differs from Maystre's original work due to the presence of the two plane surfaces S_u and S_ℓ. These, together with the profile Γ, give rise to an upward and downward going series of plane waves in media D_1 and D_2. In D_1, Γ represents the source of upward going waves and S_u the source of downward going waves. Naturally in D_2, Γ and S_ℓ have the reverse roles.

It should be noted that the upward going field, ϕ_2^+, in D_2 is incident upon the grating surface Γ and hence its plane wave nature is preserved within the grooves of Γ. This essentially constitutes a definition of the incident field on the surface Γ_-. The total field within the groove is defined to be the sum of ϕ_2^+ and a second field ϕ_2^- where ϕ_2^- can be represented as a sum of outward-going plane waves beneath the lowest point of Γ. However, it cannot be represented as a plane wave expansion with constant coefficients within the grooves of Γ. This is the Rayleigh approximation which has been the subject of

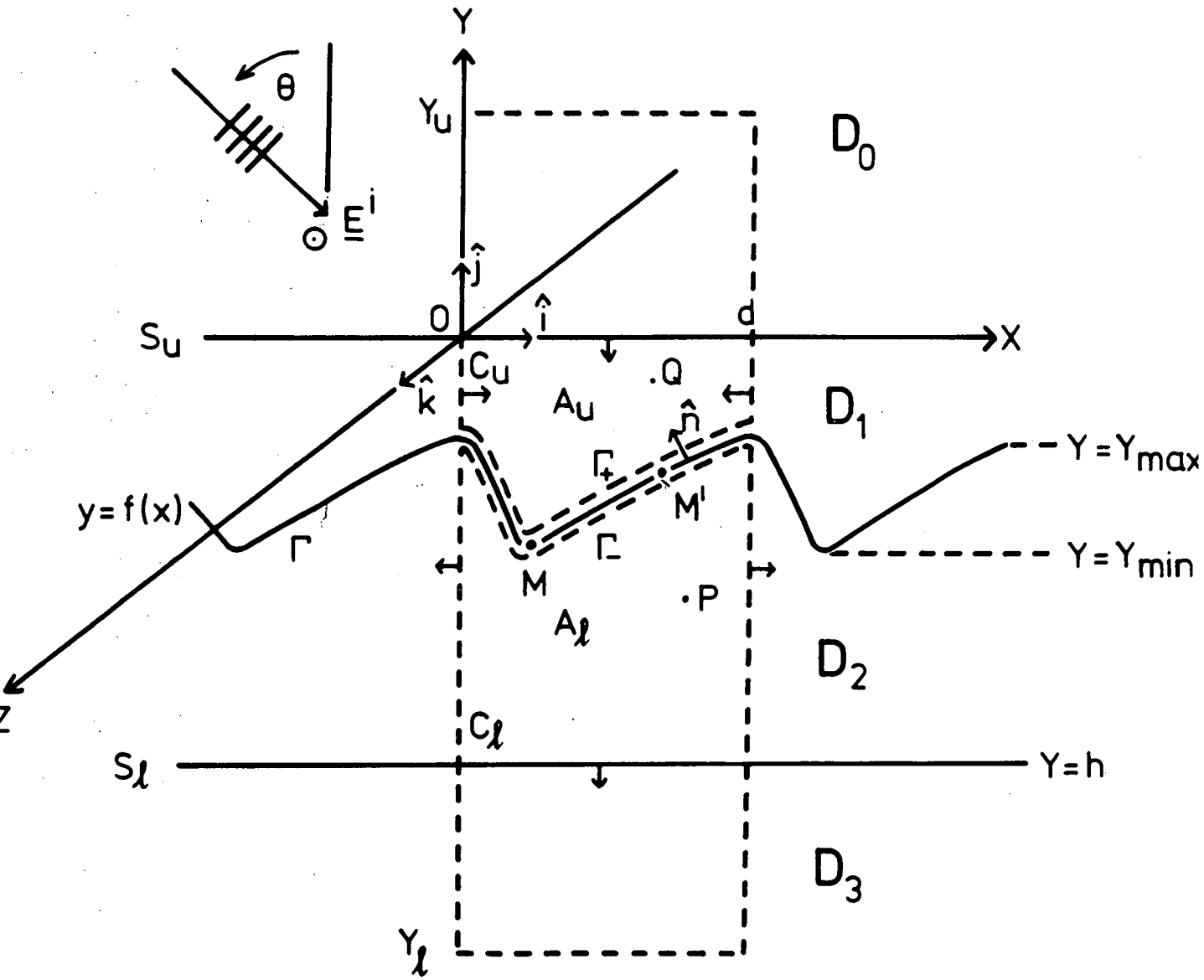


Figure 4.1a The geometry of the diffracting arrangement. Also shown is the notation used in the theoretical analysis.

controversy [4.7, 4.8] for many years.

The application of boundary conditions at the plane interface S_ℓ enables the derivation of a simple reflection relation linking the two sets of Rayleigh coefficients $\{\phi^+_{2,n}\}$ and $\{\phi^-_{2,n}\}$. The downward (or outward) going plane wave coefficients can be expressed in terms of field quantities associated with Γ and thus the field within D_2 can be given as a function of the field and its normal derivative immediately beneath Γ. With the aid of the Kirchhoff-Helmholtz formula, this field can be thought of as originating from some fictitious surface current, ξ. Using this quantity, both the field and its normal derivative just beneath Γ can be derived.

Similarly in the region D_1, the total field can be expressed as an expansion of upward and downward going waves. By considering the reflection and transmission of plane waves at the boundary S_u, it is possible to write an expression for the field in D_1 as a function of the field quantities, again expressible in terms of ξ, associated with the upper side of the interface Γ. Finally by utilizing field continuity at Γ, a single Fredholm integral equation of the first kind in the variable ξ is obtained.

Generalization of the structure to allow for a series of plane interference layers, placed both above and below the basic structure, is easily accomplished by applying a transmission line analysis to relate the plane wave coefficients in D_1 and D_2 to the reflected and transmitted amplitudes in the semi-infinite outer media.

4.3 THE THEORETICAL FORMALISM

4.3.1 Notation

The problem to be solved concerns the diffraction of a P polarized monochromatic plane wave of (free space) wavelength $\lambda = 2\pi/k_0$ incident at some angle θ on the structure of period $d = 2\pi/K$ shown in figure (4.1a). The interfaces of S_u, Γ and S_ℓ define four regions having refractive indices r_0, r_1, r_2 and r_3 in the domains D_0, D_1, D_2 and D_3 respectively. (The refractive index of region D_0 is chosen to be unity i.e. that of free space.) The surface Γ is defined by the periodic function $y = f(x)$, where $f(x+d) = f(x)$.

The incident field is defined by

$$\underline{E}^i = \phi^i \hat{\underline{z}} = \exp[i(\alpha_0 x - \chi_0 y)]\, \hat{\underline{z}} \qquad (4.3.1)$$

where $\alpha_0 = k_0 \sin\theta$

and $\chi_0 = k_0 \cos\theta$.

The total electric field $\underline{E}$ has the same polarization as the incident field and must obey the following Helmholtz equations:

$$(\nabla^2 + k_i^2)\, E(P_i) = 0 \qquad \text{for} \quad P_i \in D_i \qquad (4.3.2)$$

where $k_i = k_0 r_i$ for $i = 0,1,2$ or 3.

Both the tangential components of $\underline{E}$ and $\underline{H}$ are continuous across all interfaces and since the media are assumed to be non-magnetic

(i.e. $\mu_i = \mu_0 \ \forall i$), it follows that the field quantity, E, and its normal derivative, $\partial E/\partial n$, are continuous across all surfaces. Furthermore, the diffracted field, E^d, must obey the condition of outgoing waves as $|y| \to \infty$.

The choice of the particular polarization enables any vector dependence in the equations to be removed. In D_0 the scalar field representing the entire electric field is given by the following Rayleigh expansion:

$$\phi(x,y) = \exp[i(\alpha_0 x - \chi_0 y)] + \sum_{n=-\infty}^{\infty} \phi_n^r \exp[i(\alpha_n x + \chi_n y)] \qquad (4.3.3)$$

$$\text{for} \quad 0 \le y \le \infty$$

where $\alpha_n = \alpha_0 + nK$

and $\chi_n = \sqrt{k_0^2 - \alpha_n^2}$.

Only those waves for which $\mathrm{Im}(\chi_n) = 0$ are propagating (i.e. non-evanescent) with the angle of diffraction θ_n given by

$$\sin \theta_n = \alpha_n / k_0 . \qquad (4.3.4)$$

In a similar manner, the grating field ϕ_1^g, in medium D_1 can be written as an expansion of waves travelling in both the positive and negative y-directions. For the region $y_{max} \le y \le 0$, ϕ_1^g can be expressed as a Rayleigh expansion:

$$\phi_1^g(P_1) = \phi_1^+(P_1) + \phi_1^-(P_1) \tag{4.3.5}$$

where $\phi_1^+(x,y) = \sum_{n=-\infty}^{\infty} \phi_{1,n}^+ \exp[i(\alpha_n x + \Gamma_n y)]$

$$\phi_1^-(x,y) = \sum_{n=-\infty}^{\infty} \phi_{1,n}^- \exp[i(\alpha_n x - \Gamma_n y)]$$

with $\Gamma_n = \sqrt{k_1^2 - \alpha_n^2}$.

In the region D_2, for which $h \le y \le y_{min}$, the grating field ϕ_2^g is given by the following Rayleigh expansion:

$$\phi_2^g(P_2) = \phi_2^+(P_2) + \phi_2^-(P_2) \tag{4.3.6}$$

where $\phi_2^+(x,y) = \sum_{n=-\infty}^{\infty} \phi_{2,n}^+ \exp[i(\alpha_n x + \eta_n y)]$

$$\phi_2^-(x,y) = \sum_{n=-\infty}^{\infty} \phi_{2,n}^- \exp[i(\alpha_n x - \eta_n y)]$$

and $\eta_n = \sqrt{k_2^2 - \alpha_n^2}$.

Finally, in region D_3 ($y \le h$) the scalar transmitted field is specified by

$$\phi^t(P_3) = \sum_{n=-\infty}^{\infty} \phi_n^t \exp[i(\alpha_n x - \delta_n y)] \tag{4.3.7}$$

where $\delta_n = \sqrt{k_3^2 - \alpha_n^2}$.

4.3.2 Reflection and transmission at plane boundaries

At the interface at y = 0, field continuity states that

$$\left.\begin{aligned} \phi(x,0^+) &= \phi_1^g(x,0^-) \\ \text{and} \quad \left.\frac{\partial\phi}{\partial y}\right|_{y=0^+} &= \left.\frac{\partial\phi_1^g}{\partial y}\right|_{y=0^-} \end{aligned}\right\} \qquad (4.3.8)$$

From the above Rayleigh expansions it follows that

$$\phi_{1,n}^+ + \phi_{1,n}^- = \delta_{n,0} + \phi_n^r$$

$$\text{and} \quad \frac{\Gamma_n}{\chi_n}(\phi_{1,n}^+ - \phi_{1,n}^-) = -\delta_{n,0} + \phi_n^r$$

$$\left.\begin{aligned} \text{whence} \quad \phi_{1,n}^- &= T_{1,0}^E \delta_{n,0} + R_{1,n}^E \phi_{1,n} \\ \phi_n^r &= (T_{1,0}^E - 1)\,\delta_{n,0} + (R_{1,n}^E + 1)\,\phi_{1,n}^+ \\ \text{where} \quad R_{1,n}^E &= \frac{\Gamma_n - \chi_n}{\Gamma_n + \chi_n} \\ \text{and} \quad T_{1,n}^E &= \frac{2\chi_n}{\Gamma_n + \chi_n} \end{aligned}\right\} \qquad (4.3.9)$$

In a similar fashion, continuity of E and $\partial E/\partial n$ at the interface S_ℓ yields

$$\left.\begin{aligned} \phi^{+}_{2,n} \exp(i\eta_n h) &= R^{E}_{2,n}\, \phi^{-}_{2,n} \exp(-i\eta_n h) \\ \text{and} \quad \phi^{t}_{n} \exp(-i\delta_n h) &= (R^{E}_{2,n}+1)\, \phi^{-}_{2,n} \exp(-i\eta_n h) \\ \text{where} \quad R^{E}_{2,n} &= \frac{\eta_n - \delta_n}{\eta_n + \delta_n} \ . \end{aligned}\right\} \qquad (4.3.10)$$

4.3.3 Choice of an unknown function

In general, the Rayleigh expansions introduced in section (4.3.1) are incapable of representing the diffracted field within the grooves of Γ. Thus it is necessary to choose some function which can accurately characterize the form of the diffracted fields everywhere in space.

Consider a function u(P) defined by

$$\left.\begin{aligned} &\text{(i)} && u(P) = \phi^{g}_{2}(P) && \text{if} \quad P \in D_2 \\ &\text{(ii)} && (\nabla^2 + k_2^2)\, u(P) = 0 && \text{if} \quad P \notin D_2 \\ &\text{(iii)} && u(P) \text{ is everywhere continuous and obeys the radiation conditions as } |y| \to \infty \ . \end{aligned}\right\} \qquad (4.3.11)$$

Although the continuity of u has been defined, this does not constrain its normal derivative to have the same property. In fact, it is the absence of such continuity which enables the selection of a unique unknown whose infinite conductivity analogue is the surface current density.

By choosing a Green's function $G_2(R;R')$ obeying the inhomogeneous Helmholtz equation

$$\left.\begin{aligned}(\nabla^2+k_2^2)\, G_2(R;R') &= \delta_c(R;R') \\ &= \delta(y-y') \sum_{n=-\infty}^{\infty} \delta(x-x'-nd)\, \exp(i\alpha_0 nd)\end{aligned}\right\} \qquad (4.3.12)$$

whose solution is given by

$$G_2(R;R') = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\eta_n} \exp[i\alpha_n(x-x')+i\eta_n|y-y'|]$$

where $R \equiv (x,y)$ and $R' \equiv (x',y')$ represent arbitrary points in space, it follows from Green's theorem that

$$\iint_{A_\ell} [G_2(P;M')\nabla^2 u(M') - u(M')\nabla^2 G_2(P;M')]\, dA'$$

$$= \int_{C_\ell} [G_2(P;M') \frac{\partial u(M')}{\partial n_{M'}} - u(M') \frac{\partial G_2}{\partial n_{M'}}(P;M')]\, ds'$$

where M' is a point on the contour C_ℓ and $P \in A_\ell$. That is,

$$u(P) = -\int_{\Gamma-} [G_2 \frac{\partial u}{\partial n'} - \frac{\partial G_2}{\partial n'} u]\, ds' + \int_{S_\ell} [G_2 \frac{\partial u}{\partial y'} - \frac{\partial G_2}{\partial y'} u]\, ds' \qquad (4.3.13)$$

The expression for u can be expressed as a sum of a purely upward travelling wave field (u^+) together with a field of more general form (u^-) which is defined by:

$$u(P) = u^+(P) + u^-(P).$$

In the region $y \in [h, y_{min}]$ these two fields are given by the following Rayleigh expansions:

$$\left.\begin{aligned} u^+(P) &= \sum_n u_n^+ \exp[i(\alpha_n y + \eta_n y)] \\ \text{and} \quad u^-(P) &= \sum_n u_n^- \exp[i(\alpha_n x - \eta_n y)] \end{aligned}\right\} \tag{4.3.14}$$

With these definitions it is easily shown that

$$\int_{S_\ell} [G_2 \frac{\partial u^+}{\partial y'} - u^+ \frac{\partial G_2}{\partial y'}]\Big|_{y=h^+} ds'$$

$$= \int_0^d \sum_{m,n} \{\frac{i}{2id} \frac{\eta_m}{\eta_n} u_m^+ \exp[i\alpha_n(x-x') + \eta_n(y-h)] \exp[i(\alpha_m x' + \eta_m h)]$$

$$+ \frac{i}{2id} \frac{\eta_n}{\eta_n} u_m^+ \exp[i\alpha_n(x-x') + \eta_n(y-h)] \exp[i(\alpha_m x' + \eta_m h)]\} dx'$$

$$= \sum_n u_n^+ \exp[i(\alpha_n x + \eta_n y)] .$$

$$\text{Thus,} \quad u^+(P) = \int_{S_\ell} [G_2 \frac{\partial u^+}{\partial y'} - u^+ \frac{\partial G_2}{\partial y'}] dx' . \tag{4.3.15}$$

In a similar manner,

$$0 = \int_{S_\ell} [G_2 \frac{\partial u^-}{\partial y'} - u^- \frac{\partial G_2}{\partial y'}] dx' . \tag{4.3.16}$$

The results of equations (4.3.15) and (4.3.16) are interpreted as meaning that S_ℓ is a source of upward-going waves only.

Green's theorem may also be used to show that

$$u^-(P) = -\int_{\Gamma_-} [G_2 \frac{\partial u^-}{\partial n'} - u^- \frac{\partial G_2}{\partial n'}] \, ds' \qquad (4.3.17)$$

$$\text{and} \quad 0 = \int_{\Gamma_-} [G_2 \frac{\partial u^+}{\partial n'} - u^+ \frac{\partial G_2}{\partial n'}] \, ds' \qquad (4.3.18)$$

Thus Γ_- can be regarded as a source of downward-going waves only. Having now identified the sources of the wave-trains u^+ and u^-, the addition of (4.3.17) and (4.3.18) yields

$$u^-(P) = -\int_{\Gamma_-} [G_2 \frac{\partial u}{\partial n'} - u \frac{\partial G_2}{\partial n'}] \, ds' . \qquad (4.3.19)$$

Now the domain of convergence ($h \le y \le y_{min}$) of the Rayleigh expansion for $u^+(P)$ can be extended into the grooves of Γ by noting that it is a wave field incident upon Γ. (That is, the integrity of its plane wave expansion is preserved within the grooves.)

Hence everywhere in D_2,

$$u^+(P) = \sum_{n=-\infty}^{\infty} u_n^+ \exp[i(\alpha_n x + \eta_n y)]$$

where $u_n^+ = R_{2,n}^E \, u_n^- \exp(-2i\eta_n h)$.

$$\text{Now} \quad u_n^- = -\int_{\Gamma_-} [\frac{\partial u}{\partial n'} g_{2,n} - \frac{\partial g_{2,n}}{\partial n'} u] \, ds'$$

$$\text{where} \quad g_{2,n} = \frac{1}{2id\eta_n} \exp[-i\alpha_n x' + i\eta_n f(x')] .$$

That is,

$$u^{+}(P) = -\sum_{n} R^{E}_{2,n} \int_{\Gamma_{-}} [g_{2,n} \frac{\partial u}{\partial n'} - \frac{\partial g_{2,n}}{\partial n'} u] \, ds' \times \exp[i\alpha_n x + i\eta_n (y-2h)]. \tag{4.3.20}$$

By now adding (4.3.17) and (4.3.20) it follows that

$$u(P) = u^{+}(P) + u^{-}(P)$$

$$= -\int_{\Gamma_{-}} [G^{E}_{2}(P;M') \frac{\partial u}{\partial n'} - \frac{\partial G^{E}_{2}}{\partial n'} (P;M) \, u] \, ds' \tag{4.3.21}$$

where

$$G^{E}_{2}(P;M') = G_2(P;M') + \frac{1}{2id} \sum_{n} \frac{R^{E}_{2,n}}{\eta_n} \exp[i\alpha_n (x-x') + i\eta_n (y+f(x') - 2h)]. \tag{4.3.22}$$

The second element of the kernel G^{E}_{2} has no singularities and thus

$$(\nabla^2 + k_2^2) \, G^{E}_{2}(P;M') = (\nabla^2 + k_2^2) \, G_2(P;M') = \delta_c .$$

By now applying Green's theorem around a contour C_u within the domain D_1 it may be shown that

$$\int_{C_u} [G^{E}_{2} \frac{\partial u}{\partial n'} - \frac{\partial G^{E}_{2}}{\partial n'} u] \, ds' = -\iint_{A_u} [G^{E}_{2} \nabla^2 u - \nabla^2 G^{E}_{2} u] \, dA'$$

$$= 0 .$$

The pseudo-periodicity of the electromagnetic field causes cancellation of contributions to the line integral (around C_u) from the vertical

segments of C_u and so

$$\int_{\Gamma_+} [G_2^E \frac{\partial u}{\partial n'} - \frac{\partial G_2^E}{\partial n'} u] \, ds' = \int_{S_u} [G_2^E \frac{\partial u}{\partial y'} - \frac{\partial G_2^E}{\partial y'} u] \, dx' . \qquad (4.3.23)$$

Since the function u(Q) obeys $(\nabla^2+k_2^2)\, u(Q) = 0$, $\forall\ Q \notin D_2$, it follows that there are no field boundaries or discontinuities outside D_2. Thus in the region $y \in [y_{max},0]$, u may be expanded in terms of an outward-going plane wave field. Evaluation of G_2^E along the contour S_u shows that this kernel is a generator of upward-going plane wave terms. Some easy manipulation shows that the right hand side of (4.3.23) reduces to zero, and thus

$$0 = \int_{\Gamma_+} [G_2^E \frac{\partial u}{\partial n'} - \frac{\partial G_2^E}{\partial n'} u] \, ds' . \qquad (4.3.24)$$

Now adding equation (4.3.24) to equation (4.3.21) gives

$$u(P) = \int_{\Gamma} [G_2^E \, (\frac{\partial u}{\partial n'}\Big|_{\Gamma_+} - \frac{\partial u}{\partial n'}\Big|_{\Gamma_-}) - \frac{\partial G_2^E}{\partial n'} \, (u\Big|_{\Gamma_+} - u\Big|_{\Gamma_-})] \, ds' .$$

As u, by definition, is everywhere continuous, then

$$\left.\begin{aligned} u(P) &= \int_{\Gamma} G_2^E(P;M')\, \xi(M') \, ds' \\ \text{where } \xi(M') &= \frac{\partial u}{\partial n'}\Big|_{\Gamma_+} - \frac{\partial u}{\partial n'}\Big|_{\Gamma_-} . \end{aligned}\right\} \qquad (4.3.25)$$

This quantity, ξ, is to be used as the fundamental unknown.

In particular, for a point $M \equiv (x,f(x))$ on Γ

$$u(M) = \phi_2^g(M) = \int_\Gamma G_2^E(M;M')\ \xi(M')\ ds' \ . \qquad (4.3.26)$$

In calculating the normal derivative of ϕ_2^g on Γ, it must be noted that

$$\left.\frac{\partial \phi_2^g}{\partial n}\right|_{\Gamma_-} = \left.\frac{\partial u}{\partial n}\right|_{\Gamma_-} \ .$$

However, as $\underline{n}.\nabla u$ is a discontinuous function,

$$n(M)\ .\ \nabla u(M) = \tfrac{1}{2}\left[\left.\frac{\partial u}{\partial n}\right|_{\Gamma_-} + \left.\frac{\partial u}{\partial n}\right|_{\Gamma_+}\right] = \left.\frac{\partial u}{\partial n}\right|_{\Gamma_-} + \tfrac{1}{2}\xi(M) \ .$$

Thus $$\left.\frac{\partial \phi_2^g}{\partial n}\right|_{\Gamma_-} = -\tfrac{1}{2}\xi(M) + \int_\Gamma \underline{n}_M\ .\ \nabla G_2^E(M;M')\ \xi(M')\ ds'. \qquad (4.3.27)$$

4.3.4 The redundancy of one of Van den Berg's two coupled integral equations

In Van den Berg's original formulation [4.3] of the diffraction problem, he sought to solve for two fundamental unknowns E and $\partial E/\partial n$ at the grating surface. His choice of two unknowns necessitated the solution of a pair of coupled integral equations, thereby complicating the numerical implementation of the formalism. The analysis presented in the previous section has, however, coupled these two field quantities in terms of a more fundamental unknown ξ. Thus, it is the aim of this section to demonstrate that such coupling makes redundant one of Van den Berg's two integral equations. By returning to equation (4.3.21) of the previous section, it follows that for some point M on Γ;

$$\tfrac{1}{2}\phi_2^g(M) = -\int_{\Gamma_-} [G_2^E(M;M') \frac{\partial \phi_2^g}{\partial n'} - \frac{\partial G_2^E}{\partial n'}(M;M')\, \phi_2^g]\, ds' \,. \qquad (4.3.28)$$

Now substituting expressions (4.3.26) and (4.3.27) in the above equation, the right hand side is evaluated as follows:

$$\text{R.H.S.} = -\int_\Gamma \{G_2^E(M;M') [-\tfrac{1}{2}\xi(M') + \int_\Gamma \underline{n}_{M'} . \nabla G_2^E(M';M'')\, \xi(M'')\, ds'']$$

$$- \underline{n}_{M'} . \nabla G_2^E(M;M') \int_\Gamma G_2^E(M';M'')\, \xi(M'')\, ds''\}\, ds'$$

$$= \tfrac{1}{2}\int_\Gamma G_2^E(M;M')\, \xi(M')\, ds' - \int_\Gamma \xi(M'') \int_\Gamma [G_2^E(M;M') \frac{\partial G_2^E}{\partial n_{M'}}(M';M'')$$

$$- \frac{\partial G_2^E}{\partial n_{M'}}(M;M')\, G_2^E(M';M'')\,]\, ds'\, ds'' \,.$$

The innermost integral of the above expression may be evaluated to be zero (using Green's theorem) and thus,

$$\text{R.H.S.} = \tfrac{1}{2}\int_\Gamma G_2^E(M;M')\, \xi(M')\, ds' = \text{L.H.S.}$$

This demonstrates that all the information contained in the solution of (4.3.28) is represented in the expansions of ϕ_2^g and $\frac{\partial \phi_2^g}{\partial n}$ expressed as functions of ξ.

4.3.5 Integral representation for ϕ_1^- and ϕ_1^+ in D_1

Our mathematical analysis of the problem so far has concentrated on expressing the field quantities in the region D_2 in terms of a

single unknown. It is now necessary to show that medium D_1 imposes a further relation between the field and its normal derivative on Γ.

Consider the upward travelling field ϕ_1^+ in the region D_1. From Green's theorem,

$$\iint_{A_u} [\phi_1^+ \nabla^2 G_1 - G_1 \nabla^2 \phi_1^+] \, dA' = - \int_{C_u} [\phi_1^+ \frac{\partial G_1}{\partial n'} - G_1 \frac{\partial \phi_1^+}{\partial n'}] \, ds'$$

where $G_1 = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\Gamma_n} \exp[i\alpha_n(x-x') + i\Gamma_n|y-y'|]$.

Since ϕ_1^+ can be represented by an upward going plane wave expansion in the region $y_{max} < y < 0$, then

$$\int_{S_u} [\phi_1^+ \frac{\partial G_1}{\partial n'} - \frac{\partial \phi_1^+}{\partial n'} G_1] \, dx' = 0$$

and thus $$\phi_1^+(Q) = \int_{\Gamma_+} [\frac{\partial \phi_1^+}{\partial n'} G_1 - \frac{\partial G_1}{\partial n'} \phi_1^+] \, ds' \, . \qquad (4.3.29)$$

In a similar manner,

$$0 = \int_{\Gamma_+} [\frac{\partial \phi_1^-}{\partial n'} G_1 - \frac{\partial G_1}{\partial n'} \phi_1^-] \, ds' \, . \qquad (4.3.30)$$

The addition of equations (4.3.29) and (4.3.30) then yields

$$\phi_1^+(Q) = \int_{\Gamma_+} [\frac{\partial \phi_1^g}{\partial n'} G_1 - \frac{\partial G_1}{\partial n'} \phi_1^g] \, ds' \, . \qquad (4.3.31)$$

Equations (4.3.29) and (4.3.30) demonstrate that Γ is a source of upward-going waves only. A similar treatment shows S_u to be a source of

downward-going waves. Thus a plane wave expansion for ϕ_1^+ is valid only in the region $y_{max} \leq y \leq 0$, whereas ϕ_1^-, a field approaching Γ, can be represented adequately by a Rayleigh expansion at all points within A_u.

Expressing this in a more rigorous manner, it can be readily seen that $G_1(Q;M')$ is a generator of upward-going plane waves in the region $y \in [y_{max},0]$. That is,

$$G_1(Q;M') = \sum_n g_{1,n} \exp[i(\alpha_n x + \Gamma_n y)]$$

where $g_{1,n} = \frac{1}{2id\Gamma_n} \exp(-i\alpha_n x' - i\Gamma_n f(x'))$.

This enables ϕ_1^+ to be expressed in terms of the following Rayleigh expansion:

$$\left.\begin{aligned} \phi_1^+(Q) &= \sum_n \phi_{1,n}^+ \exp[i(\alpha_n x + \Gamma_n y)] \\ \text{where } \phi_{1,n}^+ &= \int_{\Gamma_+} [\frac{\partial \phi_1^g}{\partial n'} g_{1,n} - \frac{\partial g_{1,n}}{\partial n'} \phi_1^g] \, ds' \, . \end{aligned}\right\} \quad (4.3.32)$$

4.3.6 Derivation of an integral equation

In the previous section, it was shown that

$$\phi_1^+(Q) = \int_{\Gamma_+} [G_1 \frac{\partial \phi_1^+}{\partial n'} - \frac{\partial G_1}{\partial n'} \phi_1^+] \, ds'$$

for a point Q completely enclosed within D_1. Should Q tend to M, a

point on the surface Γ_+, then

$$\tfrac{1}{2}\,\phi_1^+(M) = \int_{\Gamma_+} [G_1(M;M')\,\frac{\partial\phi_1^+}{\partial n'}(M') - \frac{\partial G_1}{\partial n'}(M;M')\,\phi_1^+(M')]\,ds' \qquad (4.3.33)$$

where the multiplicative factor of "½" is a consequence of the singularity of the Green's function G_1 being half-enclosed within the domain of integration, A_u. Similarly, it may be shown that

$$-\tfrac{1}{2}\phi_1^-(M) = \int_{\Gamma_+} [\frac{\partial\phi_1^-}{\partial n'}\,G_1(M;M') - \phi_1^-\,\frac{\partial G_1}{\partial n'}(M;M')]\,ds' \qquad (4.3.34)$$

By now adding equations (4.3.33) and (4.3.34), it follows that

$$\phi_1^-(M) = \tfrac{1}{2}\phi_1^g(M) - \int_{\Gamma_+} [\frac{\partial\phi_1^-}{\partial n'}\,G_1(M;M') - \phi_1^-\,\frac{\partial G_1}{\partial n'}(M;M')]\,ds' \qquad (4.3.35)$$

The field ϕ_1^- is incident upon Γ and as stated earlier its Rayleigh expansion in the region $y \in [y_{max},0]$ is preserved within the grooves of Γ. A consequence of this is the close resemblance of equation (4.3.35) to the infinite conductivity formalism of Pavageau and Bousquet [4.9] with the quantity ϕ_1^- being a measure of the effective plane wave field incident upon the grating profile.

$$\text{Now, } \phi_1^-(M) = \sum_n \phi_{1,n}^-\,\exp[i(\alpha_n x - \Gamma_n f(x))]$$

$$= \sum_n (T_{1,0}^E\delta_{n,0} + R_{1,n}^E\phi_{1,n}^+)\,\exp[i(\alpha_n x - \Gamma_n f(x))] \qquad (4.3.36)$$

As before, it must be stressed that the Rayleigh coefficients $\{\phi_{1,n}^+\}$

are meaningful only in the space $y \in [y_{max}, 0]$.

By now substituting equations (4.3.32) and (4.3.36) into expression (4.3.35), the following integral equation is derived:

$$\tfrac{1}{2}\phi_1^g(M) - \int_{\Gamma_+} \left[\frac{\partial \phi_1^g}{\partial n'} G_1^E(M;M') - \phi_1^g \frac{\partial G_1^E}{\partial n'}(M;M')\right] ds' =$$

$$= T_{1,0}^E \exp[i(\alpha_0 x - \Gamma_0 f(x))] \qquad (4.3.37a)$$

where

$$G_1^E(M;M') = G_1(M;M') + \frac{1}{2id} \sum_n \frac{R_{1,n}^E}{\Gamma_n} \exp[i\alpha_n(x-x') - i\Gamma_n(f(x)) + f(x'))]. \qquad (4.3.37b)$$

Clearly, equation (4.3.37a) links the two field unknowns $\phi_1^g(M')$ and $\partial\phi_1^g(M')/\partial n'$. Application of the boundary conditions appropriate to this state of polarization

$$\phi_1^g(M)\Big|_{\Gamma_+} = \phi_2^g(M)\Big|_{\Gamma_-}$$

and $$\frac{\partial \phi_1^g}{\partial n}(M)\Big|_{\Gamma_+} = \frac{\partial \phi_2^g}{\partial n}(M)\Big|_{\Gamma_-}$$

and substitution of expressions (4.3.26) and (4.3.27) into equation (4.3.37a) yields the final result, a Fredholm integral equation of the first kind.

$$\int_\Gamma \tau^E(M;M')\, \xi(M') ds' = T_{1,0}^E \exp[i(\alpha_0 x - \Gamma_0 f(x))] \qquad (4.3.38)$$

where $\tau^E(M;M') = \frac{1}{2} G_1^E(M;M') + G_2^E(M;M') -$

$$- \int_\Gamma [G_1^E(M;M'') \frac{\partial G_2^E}{\partial n_{M''}} (M'';M') - \frac{\partial G_1^E}{\partial n_{M''}} (M;M'') \; G_2^E(M'';M')] \, ds'' \tag{4.3.39}$$

4.3.7 Reconstruction of the field amplitudes

Equation (4.3.38) can be solved numerically using a points-matching technique in which the unknown function ξ is sampled over a single period by a set of Dirac delta functions (see for example Harrington [4.10]). The above integral equation can be solved in closed form only for the particular case of zero groove depth. Further discussion of this is presented in section (4.4.4), dealing with Fresnel's Laws.

Following the numerical solution of equation (4.3.38) for the unknown ξ, the field and its normal derivative at the surface Γ_- are reconstructed using the expressions:

$$\phi_2^g(M) = \int_\Gamma G_2^E(M;M') \, \xi(M') \, ds'$$

and $$\frac{\partial \phi_2^g}{\partial n} (M) \Big|_{\Gamma_-} = -\frac{1}{2}\xi(M) + \int_\Gamma \frac{\partial G_2^E}{\partial n} (M;M') \, \xi(M') \, ds'.$$

Use of the continuity equations

$$\phi_1^g(M) \Big|_{\Gamma_+} = \phi_2^g(M) \Big|_{\Gamma_-}$$

and $$\frac{\partial \phi_1^g}{\partial n} (M) \Big|_{\Gamma_+} = \frac{\partial \phi_2^g}{\partial n} (M) \Big|_{\Gamma_-}$$

then enables reconstruction of the upward-going Rayleigh coefficients $\{\phi^{+}_{1,n}\}$ using the following expression:

$$\phi^{+}_{1,n} = \int_{\Gamma_+} [\frac{\partial \phi^{g}_{1}}{\partial n'} g_{1,n} - \frac{\partial g_{1,n}}{\partial n'} \phi^{g}_{1}] \, ds' .$$

The reflected field amplitudes can then be calculated using

$$\phi^{r}_{n} = (T^{E}_{1,0}-1) \, \delta_{n,0} + (R^{E}_{1,n}+1) \, \phi^{+}_{1,n} .$$

The computation of the transmitted field coefficients also follows using a similar procedure.

4.3.8 Modification of the theory for S polarization

The above analysis has discussed a method for studying the diffraction properties of a structure consisting of a single grating surface surrounded by a pair of thin films and used with P polarized radiation. In this section, the theory for S polarized incident radiation is outlined and the salient differences between the two treatments noted. As will easily be seen, such differences only arise due to the different boundary conditions imposed.

In an analogous manner, scalar wave functions ψ (representing magnetic field quantities) are used in place of the functions ϕ discussed earlier. For this problem, the new unknown ξ is given by

$$\xi(M) = \frac{\partial u}{\partial n}\Big|_{\Gamma_+} - \frac{\partial u}{\partial n}\Big|_{\Gamma_-} \qquad (4.3.40)$$

where u(P) is now defined by

(i) $u(P) = \psi_2^g(P)$ for $P \in D_2$

(ii) $(\nabla^2+k_2^2)\, u(P) = 0$ otherwise

(iii) $u(P)$ is everywhere continuous.

From these definitions

$$\psi_2^g(M) = \int_\Gamma G_2^H(M;M')\, \xi(M')\, ds' \tag{4.3.41}$$

and

$$\frac{\partial \psi_2^g}{\partial n}(M)\Big|_{\Gamma_-} = -\tfrac{1}{2}\xi(M) + \int_\Gamma \frac{\partial G_2^H}{\partial n}(M;M')\, \xi(M')\, ds' \tag{4.3.42}$$

where

$$G_2^H(M;M') = G_2(M;M') + \frac{1}{2id}\sum_n \frac{R_{2,n}^H}{\eta_n}\exp[i\alpha_n(x-x') + i\eta_n(f(x)+f(x')-2h)] \tag{4.3.43a}$$

and

$$R_{2,n}^H = \frac{\eta_n - (\frac{r_2}{r_3})^2\,\delta_n}{\eta_n + (\frac{r_2}{r_3})^2\,\delta_n} \quad . \tag{4.3.43b}$$

The S polarization analogue of equation (4.3.37a) is

$$\tfrac{1}{2}\,\psi_1^g(M) - \int_{\Gamma_+} [\frac{\partial \psi_1^g}{\partial n'}(M')\, G_1^H(M;M') - \psi_1^g(M')\, \frac{\partial G_1^H}{\partial n'}(M;M')]\, ds'$$

$$= T_{1,0}^H \exp[i(\alpha_0 x - \Gamma_0 f(x))] \tag{4.3.44a}$$

where

$$G_1^H(M;M') = G_1(M;M') + \sum_n \frac{R_{1,n}^H}{2id\Gamma_n} \exp[i\alpha_n(x-x') - i\Gamma_n(f(x) + f(x'))] \quad (4.3.44b)$$

and $$R_{1,n}^H = \frac{\Gamma_n - r_1^2 \chi_n}{\Gamma_n + r_1^2 \chi_n}$$

with $$T_{1,n}^H = \frac{2\chi_n}{\chi_n + \Gamma_n/r_1^2} \quad .$$

Application of the S polarization boundary conditions at Γ,

$$\psi_1^g(M)\Big|_{\Gamma_+} = \psi_2^g(M)\Big|_{\Gamma_-} \text{ and } \frac{1}{r_1^2}\frac{\partial \psi_1^g}{\partial n}(M)\Big|_{\Gamma_+} = \frac{1}{r_2^2}\frac{\partial \psi_2^g}{\partial n}(M)\Big|_{\Gamma_-}$$

enables the analogue of equation (4.3.39) to be written in the form

$$\int_\Gamma \tau^H(M;M')\, \xi(M')\, ds' = T_{1,0}^H \exp[i(\alpha_0 x - \Gamma_0 f(x))] \qquad (4.3.45)$$

where

$$\tau^H(M;M') = \tfrac{1}{2}\frac{r_1^2}{r_2^2} G_1^H(M;M') + \tfrac{1}{2} G_2^H(M;M') - \int_\Gamma \left[\frac{r_1^2}{r_2^2} G_1^H(M;M'') \frac{\partial G_2^H}{\partial n_{M''}}(M'';M') - \frac{\partial G_1^H}{\partial n_{M''}}(M;M'')\, G_2^H(M'';M')\right] ds'' \quad . \qquad (4.3.46)$$

Equation (4.3.46), a Fredholm integral equation of the first kind, is also solved using the points-matching technique. It should be stressed that this integral formulation converges equally well for S and P polarizations, unlike the differential formalism of Nevière et al [4.5].

4.3.9 Generalization of the theory to accommodate multiple interference films both above and below the grating surface

It is the purpose of this section to generalize the above formalism to accommodate the presence of a series of plane interference films placed behind the structure illustrated in figure (4.1a). Of course, a similar treatment can be applied to a series of films placed in front of the same structure.

Consider the structure depicted in figure (4.1b). The nomenclature used is defined as follows:

(i) a set of (L + 1) media separated by a set of L contours $\{S_i\}$;

(ii) surface m separates media m and (m + 1);

(iii) the direction cosine for the n^{th} order plane wave within medium m is $\gamma_{m,n}$.

The treatment, based on a transmission line analysis, is restricted to the case of a P polarized field. (However, it is a relatively trivial exercise to convert this treatment to cover the case of an S polarized field.)

Due to the chosen polarization, the appropriate boundary conditions are

(i) continuity of the electric field at any interface, and

(ii) continuity of its normal derivative at any interface.

The wave field in medium D_m may be written as a Rayleigh expansion.

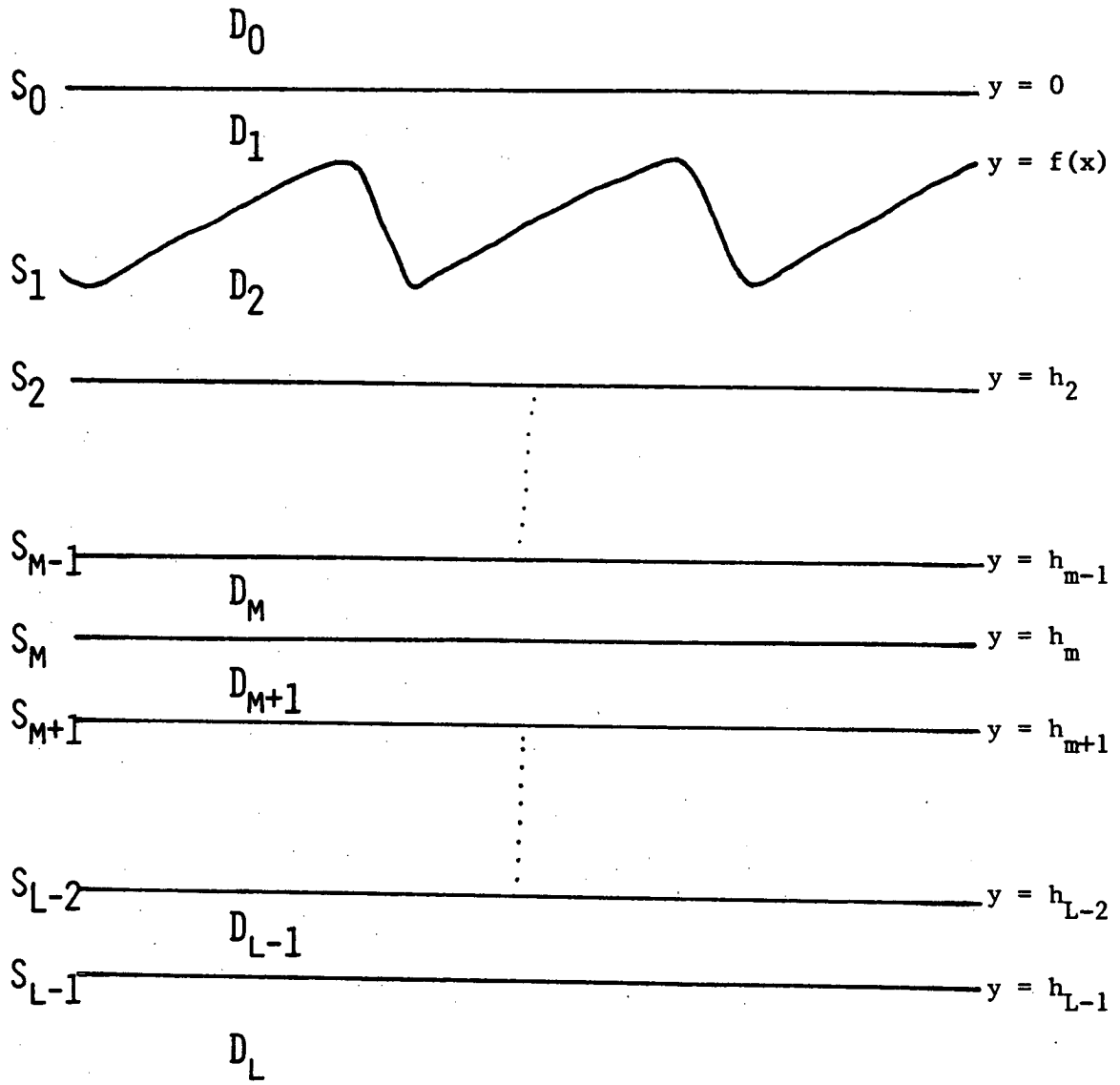


Figure 4.1b The geometry of the multi-layer structure.

$$\text{i.e.}\quad \phi_m^g = \sum_n [\phi^+_{m,n} \exp(i\gamma_{m,n} y) + \phi^-_{m,n} \exp(-i\gamma_{m,n} y)] \exp(i\alpha_n x) \qquad (4.3.47)$$

The continuity of the field and its normal derivative of the plane surface S_m at height $y = h_m$ imply that

$$\left.\begin{aligned} \overline{\phi^+_{m,n}} + \overline{\phi^-_{m,n}} &= \overline{\overline{\phi^+_{m+1,n}}} + \overline{\overline{\phi^-_{m+1,n}}} \\ \text{and}\quad \overline{\phi^+_{m,n}} - \overline{\phi^-_{m,n}} &= \frac{\gamma_{m+1,n}}{\gamma_{m,n}} \left(\overline{\overline{\phi^+_{m+1,n}}} - \overline{\overline{\phi^-_{m+1,n}}}\right) \end{aligned}\right\} \qquad (4.3.48)$$

where $\overline{\phi^+_{m,n}} = \phi^+_{m,n} \exp(i\gamma_{m,n} h_m)$

$$\overline{\phi^-_{m,n}} = \phi^-_{m,n} \exp(-i\gamma_{m,n} h_m)$$

$$\overline{\overline{\phi^+_{m+1,n}}} = \phi^+_{m+1,n} \exp(i\gamma_{m+1,n} h_m)$$

and $\overline{\overline{\phi^-_{m+1,n}}} = \phi^-_{m+1,n} \exp(-i\gamma_{m+1,n} h_m)$.

Hence $\overline{\overline{\phi^+_{m+1,n}}} = \overline{\phi^+_{m+1,n}} \exp[i\gamma_{m+1,n} (h_m - h_{m+1})]$

and $\overline{\overline{\phi^-_{m+1,n}}} = \overline{\phi^-_{m+1,n}} \exp[-i\gamma_{m+1,n} (h_m - h_{m+1})]$.

With the nomenclature defined above, equation (4.3.48) can be written as a matrix equation

$$\overline{\Phi}_m = \sigma_{m+1} \overline{\Phi}_{m+1} \qquad (4.3.49a)$$

$$\text{where } \bar{\Phi}_m = \begin{bmatrix} \overline{\phi^+_{m,n}} \\ \overline{\phi^-_{m,n}} \end{bmatrix} \tag{4.3.49b}$$

$$\text{and } \sigma_m = \tfrac{1}{2}\left(1 + \frac{\gamma_{m,n}}{\gamma_{m-1,n}}\right) \exp[+ i\gamma_{m,n}(h_m - h_{m-1})] \times$$

$$\times \begin{bmatrix} P_{m,n} & R^E_{m,n} \\ P_{m,n} R^E_{m,n} & 1 \end{bmatrix} \tag{4.3.49c}$$

$$\left.\begin{aligned} \text{with } R^E_{m,n} &= \frac{\gamma_{m-1,n} - \gamma_{m,n}}{\gamma_{m-1,n} - \gamma_{m,n}} \\ \text{and } P_{m,n} &= \exp[2i\gamma_{m,n}(h_{m-1} - h_m)] \quad . \end{aligned}\right\} \tag{4.3.49d}$$

By using this matrix representation, it is then possible to handle the cascade of a series of thin films in a particularly elegant manner. The upward and downward going plane wave amplitudes at the lower boundary of the medium D_2 and the penultimate layer, D_{L-1} are related by the matrix equation

$$\bar{\Phi}_2 = \sigma_T \bar{\Phi}_{L-1} \tag{4.3.50a}$$

$$\text{where } \sigma_T = \prod_{i=3}^{L-1} \sigma_i \quad . \tag{4.3.50b}$$

Thus the reflection coefficient for plane waves striking the boundary S_2 is given by

$$R^{E}_{2,n} = \frac{\overline{\phi^{+}_{2,n}}}{\overline{\phi^{-}_{2,n}}} = \frac{[\ {}_{11}\sigma_T\ R^{E}_{L,n} + {}_{12}\sigma_T\ ]}{[\ {}_{21}\sigma_T\ R^{E}_{L,n} + {}_{22}\sigma_T\ ]} \qquad (4.3.51)$$

where the numerical subscripts represent matrix co-ordinates. With knowledge of this effective reflection coefficient, the solution of the diffraction problem then proceeds along the usual lines relying upon the solution of the integral equation given by expression (4.3.38). The reconstruction of the field quantities in all the plane layers can be achieved by performing a calculation which is essentially the "inverse" of that described within this section.

4.4 THEORETICAL PROPERTIES OF TRANSMISSION GRATINGS

4.4.1 Introduction

The prime motivation for any investigation into the theoretical properties of diffraction gratings must surely be the necessity to have physically meaningful tests of the validity of any numerical model. Such properties, although having no obvious relevance to the experimenter, have proved invaluable to the theoretician in isolating discrepancies in their physical models and errors in their associated numerical implementation.

Of course, the most fundamental of these checks must be conservation of energy. Early models such as scalar and Rayleigh theories were soon found to be wanting with respect to this fundamental constraint. Over the past decade authors such as Petit [4.11] and Van den Berg [4.3] have presented clear discussions of this concept and demonstrated that

their integral formalisms obey this criterion to high precision.

Reciprocity, or the concept of inverse return as it is sometimes known, must surely follow energy conservation in fundamental physical importance. The form of the reciprocity theorem discussed later in this section applies only to the specific cases of P and S polarization. The concept of reciprocity, pertaining to diffraction gratings, can be stated as follows: Consider a plane wave incident upon the grating surface. Let the efficiency of the p^{th} order diffracted wave, propagating at angle θ_p, be ρ_p. Should a wave be returned along the same path as the p^{th} order wave emerged, then its p^{th} order diffracted wave will emerge along the original incident path, also with efficiency ρ_p.

Contributions by such authors as Uretsky [4.12], Petit [4.13, 4.14] and Maystre [4.14] have demonstrated the significance of the reciprocity theorem for perfectly conducting gratings. However, it was not until the work of Maystre and McPhedran [4.15] that the theorem was shown to be valid for finitely conducting reflection gratings. These same authors [4.16, 4.17] further generalized the theorem to include the case of two-dimensional diffracting structures and derived forms applicable to the return of both reflected and transmitted orders. Subsequent work by Nevière and Vincent [4.18] yielded a form of the reciprocity theorem appropriate to singly periodic transmission gratings. The same result was derived independently by the author, but during the preparation of the manuscript of reference [4.1] the published work of Nevière and Vincent appeared in the literature.

The above two criteria are by far the most important of all the physical tests. Thus, in an attempt to meet one of the aims of this

thesis, the presentation of an overview of the current theories, these concepts are briefly discussed in sections (4.4.2) and (4.4.3).

However, there are a further two properties which have highlighted some early inadequacies of the formalism and its associated numerical implementation. The first of these is the a priori agreement of the classical diffraction theory with Fresnel's laws when the groove modulation is reduced to zero. The second of these properties relates the phases and efficiencies of the reflected and transmitted orders excited by a lossless symmetric grating when used in a first order Littrow mount between the (-1) and (-2,+1) Wood anomalies. The search for this constraint was motivated originally by the need to explain the $\pi/2$ phase difference between the $(-1)^{th}$ and 0^{th} orders for a symmetric, perfectly reflecting grating operated in a $(-1)^{th}$ order Littrow configuration. These are discussed in sections (4.4.4) and (4.4.5) respectively.

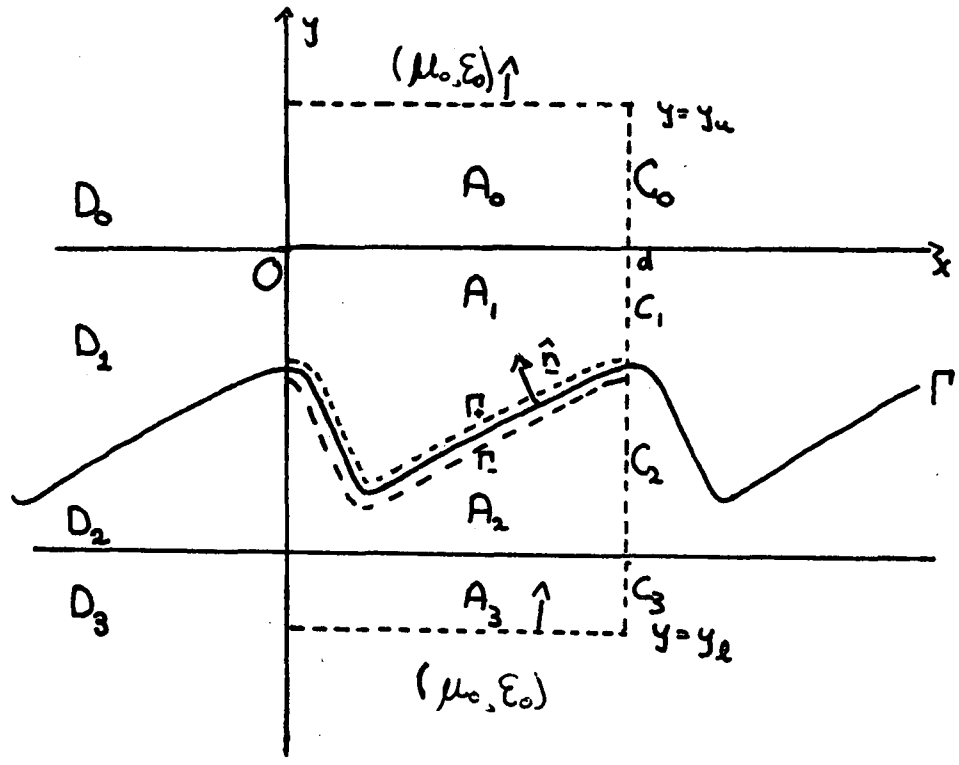
4.4.2 Energy conservation

The work presented herein is essentially a review of the method used by Van den Berg [4.3]. Consider a lossless grating, as illustrated in figure (4.1c), illuminated with S polarized radiation. The nomenclature used in this analysis is identical to that prescribed in the previous section, with the total magnetic fields in the domains D_0, D_1, D_2 and D_3 represented by $\psi = (\psi^i + \psi^r)$, ψ_1^g, ψ_2^g and ψ^t respectively.

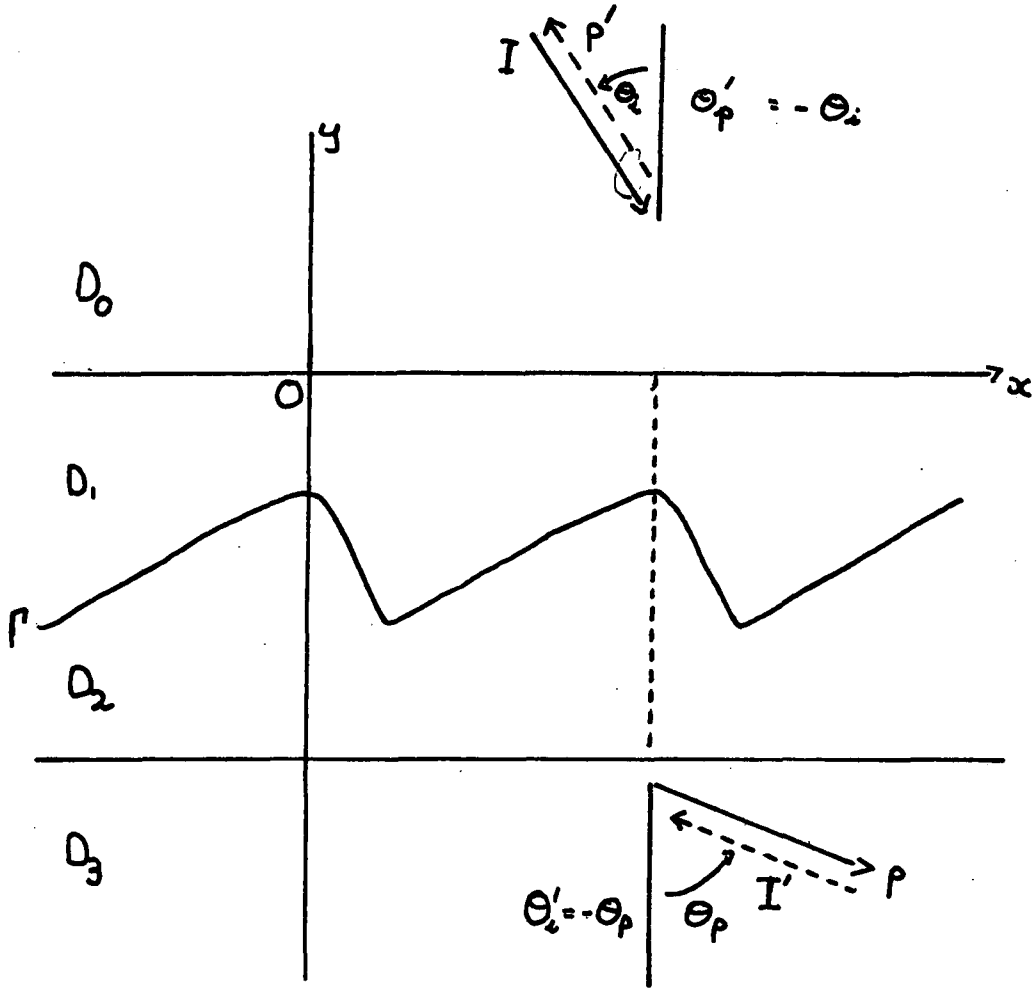
From Green's theorem,

$$\iint_{A_0} [\psi \nabla^2 \psi^* - \psi^* \nabla^2 \psi] \, dA = \int_{C_0} [\psi \frac{\partial \psi^*}{\partial n} - \psi^* \frac{\partial \psi}{\partial n}] \, ds$$

where C_0 is the contour bounding the region A_0. (Here the asterisk



Nomenclature for sections (4.4.2) and (4.4.3).



Diffracting arrangement for the Reciprocity Theorem (transmission).

Figure 4.1c.

represents complex conjugation.) The integrand of the left hand side vanishes as a consequence of the lossless nature of the medium D_0. {It is interesting to note that this quantity is analogous to the term $\underline{E}.\underline{J}$ (where $\underline{J} = \sigma\underline{E}$) in the Poynting theorem and thus represents the ohmic loss of electromagnetic energy.}

Now, since A_0 is lossless and ψ is a pseudo-periodic field quantity, it follows that

$$\int_o^d \left[\psi\frac{\partial\psi^*}{\partial y} - \psi^* \frac{\partial\psi}{\partial y}\right] \Big|_{y=y_u} dx = \int_o^d \left[\psi \frac{\partial\psi^*}{\partial y} - \psi^* \frac{\partial\psi}{\partial y}\right] \Big|_{y=0^+} dx$$

$$= \frac{1}{r_1^2} \int_o^d \left[\psi_1^g \frac{\partial\psi_1^{g*}}{\partial y} - \psi_1^{g*} \frac{\partial\psi_1^g}{\partial y}\right] \Big|_{y=0^-} dx$$

(after application of the S polarization boundary conditions at the surface y = 0).

A further application of Green's theorem to the region A_1, (which is also assumed to be lossless) bounded by the contour C_1 yields the results

$$\int_o^d \left[\psi \frac{\partial\psi^*}{\partial y} - \psi^* \frac{\partial\psi}{\partial y}\right] \Big|_{y=y_u} dx = \frac{1}{r_1^2} \int_{\Gamma_+} \left[\psi_1^g \frac{\partial\psi_1^{g*}}{\partial n} - \psi_1^{g*} \frac{\partial\psi_1^g}{\partial n}\right] ds$$

$$= \frac{1}{r_2^2} \int_{\Gamma_-} \left[\psi_2^g \frac{\partial\psi_2^{g*}}{\partial n} - \psi_2^{g*} \frac{\partial\psi_2^g}{\partial n}\right] ds \ .$$

The continuous usage of Green's theorem within each layer of the structure and application of the S polarization boundary conditions at the interfaces separating these layers leads to the final result:

$$\int_0^d [\psi \frac{\partial \psi^*}{\partial y} - \psi^* \frac{\partial \psi}{\partial y}] \Big|_{y=y_u} dx = \int_0^d [\psi^t \frac{\partial \psi^{t*}}{\partial y} - \psi^{t*} \frac{\partial \psi^t}{\partial y}] \Big|_{y=y_\ell} dx. \quad (4.4.1)$$

Now in the regions D_0 to D_3 the fields ψ and ψ^t may be expressed in terms of the following Rayleigh expansions:

$$\psi(x,y) = \sum_{n=-\infty}^{\infty} [\exp(-i\chi_0 y)\, \delta_{n,0} + \psi_n^r \exp(i\chi_n y)] \exp(i\alpha_n x) \forall (x,y) \in D_0 \quad (4.4.2a)$$

$$\text{and} \quad \psi^t(x,y) = \sum_{n=-\infty}^{\infty} \psi_n^t \exp[i(\alpha_n x - \chi_n y)] \forall (x,y) \in D_3. \quad (4.4.2b)$$

(Note that both D_0 and D_3 are assumed to be free space.)
These definitions, together with the orthogonality of the plane wave basis $\{\exp(i\alpha_n x)\}$ enable equation (4.4.1) to be reduced to the form

$$\sum_{n\in\Omega} \frac{\chi_n}{\chi_0} [|\psi_n^r|^2 + |\psi_n^t|^2] = 1 \quad (4.3)$$

where $\Omega = \{ n | \mathrm{Im}(\chi_n) = 0 \}$.

Tables (4.1a) and (4.1b) demonstrate the conservation of energy propagating normally to the surface. In these tables the following nomenclature has been used:

$$\rho_n^R = \frac{\cos\theta_n}{\cos\theta_i} |\psi_n^r|^2 \quad - \quad \text{efficiency of the } n^{th} \text{ order reflected wave}$$

$$\text{and} \quad \rho_n^T = \frac{\cos\theta_n}{\cos\theta_i} |\psi_n^t|^2 \quad - \quad \text{efficiency of the } n^{th} \text{ order transmitted wave.}$$

4.4.3 Reciprocity - its application to the Transmission grating

Reciprocity relations are probably the most rigorous test of any grating formalism. Let us consider an incident plane wave, I, giving rise to a set of reflected and transmitted diffracted waves. Let the efficiency of the p^{th} order wave (in either reflection or transmission) propagating at an angle θ_p be ρ_p. If a wave I' is now returned along the same path as this p^{th} order wave, then its p^{th} order diffracted wave will emerge along the original incident path, again with efficiency ρ_p.

The basic approach presented here was strongly motivated by the work of Maystre and McPhedran [4.15]. Although Nevière and Vincent [4.18] have published a reciprocity theorem applicable to transmission gratings, the method detailed below is entirely the work of the author and the investigation was conducted independently of the above study.

Consider the problem illustrated in figure (4.1c). The incident P polarized wave I establishes a field distribution given by the following Rayleigh expansions

$$\text{(i)} \quad \phi_0(P) = \exp[i(\alpha_0 x - \chi_0 y)] + \sum_n \phi_n^r \exp[i(\alpha_n x + \chi_n y)] \quad \text{for } P \in D_0 \tag{4.4.4a}$$

$$\text{and (ii)} \quad \phi_3(P) = \sum_n \phi_n^t \exp[i(\alpha_n x - \chi_n y)] \quad \text{for} \quad P \in D_3 \, . \tag{4.4.4b}$$

If a wave is then returned to the lower plane face of the grating at an angle $\theta_i' = -\theta_p$, then its p^{th} order transmitted wave propagates along the inverse path of the initial incident wave at an angle θ_i. This sets up a diffracted field distribution given by

$$\phi_0'(x,y) = \sum_n \phi_n^{t'} \exp[i(\alpha_n' x + \chi_n' y)] \quad \text{in} \quad D_0 \tag{4.4.5a}$$

$$\text{and} \quad \phi_3'(x,y) = \exp[i(\alpha_0' x + \chi_0' y)] + \sum_{n=-\infty}^{\infty} \phi_n^{r'} \exp[i(\alpha_n' x - \chi_n' y)] \text{ in } D_3. \tag{4.4.5b}$$

Now, since $\theta_i' = -\theta_p$ it follows that $\alpha_0' = -\alpha_p$.

Hence $\alpha_n' = -\alpha_{p-n}$ and $\chi_n' = \chi_{p-n}$.

By now applying Green's theorem (in a layer by layer fashion) to the entire structure as detailed in section (4.4.2), it may be shown that

$$\int_o^d [\phi_0' \frac{\partial \phi_0}{\partial y} - \phi_0 \frac{\partial \phi_0'}{\partial y}]\Big|_{y=y_u} dx = \int_o^d [\phi_3' \frac{\partial \phi_3}{\partial y} - \phi_3 \frac{\partial \phi_3'}{\partial y}]\Big|_{y=y_\ell} dx. \tag{4.4.6}$$

On explicitly evaluating this expression using the field quantities defined in equations (4.4.4) and (4.4.5), it follows that

$$\chi_0 \phi_p^{t'} = \chi_p \phi_p^t$$

$$\text{and thus } \rho_p^{T'} = \rho_p^T, \tag{4.4.7}$$

TABLE 4.1a

n	$\theta_i = -\theta'_n$	$\theta'_i = -\theta_n$	ρ_n^R	$(\rho_n^R)'$	ρ_n^T	$(\rho_n^T)'$
-2	19.75^o	$+73.86^o$	0.0012	0.0003	0.0007	0.0007
-1	19.75^o	$+18.07^o$	0.0198	0.0191	0.0482	0.0482
0	19.75^o	-19.75^o	0.2431	0.2461	0.6623	0.6625
+1	19.75^o	-80.38^o	0.0028	0.0041	0.0212	0.0213

Demonstration of the Reciprocity Theorem for the reflected and transmitted orders excited by the grating depicted in figure 4.2a. Terms denoted by a superscripted ' correspond to the orders of diffraction returned as incident waves, and are to be compared with the corresponding unprimed quantities. ρ denotes the efficiency of the reflected and transmitted orders (n) shown by superscripts (R) and (T) respectively. The reflected and transmitted efficiency sums were 0.2669 and 0.7324 respectively giving agreement with energy conservation to within 0.07%. The above calculations were performed with 20 sampling points per period.

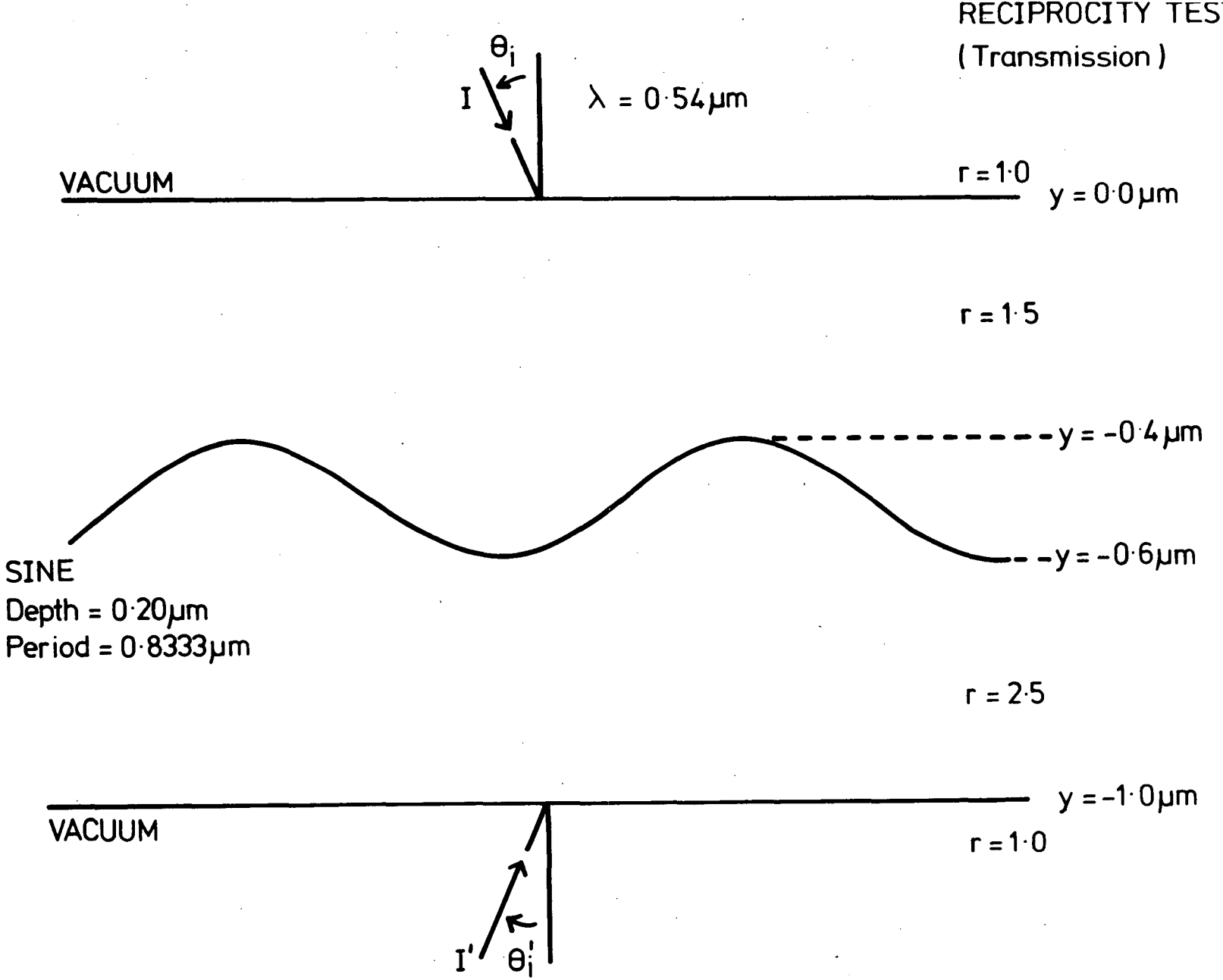


Figure (4.2a) The geometrical arrangement for the reciprocity test detailed in table (4.1a).

TABLE 4.1b

n	$\theta_i = -\theta'_n$	$\theta'_i = -\theta_n$	ρ_n^R	$(\rho_n^R)'$	ρ_n^T
-2	19.75^o	$+73.86^o$	0.0108	0.0108	0.0030
-1	19.75^o	$+18.07^o$	0.0249	0.0256	0.2904
0	19.75^o	-19.75^o	0.3281	0.3281	0.1187
+1	19.75^o	-80.38^o	0.0282	0.0276	0.1994

Demonstration of energy conservation of the Reciprocity Theorem for the various orders excited by the grating depicted in figure (4.2b). The reflected and transmitted efficiency sums were 0.3919 and 0.6115 respectively thereby giving agreement with energy conservation to within 0.34%. These calculations were performed using 30 sampling points per period.

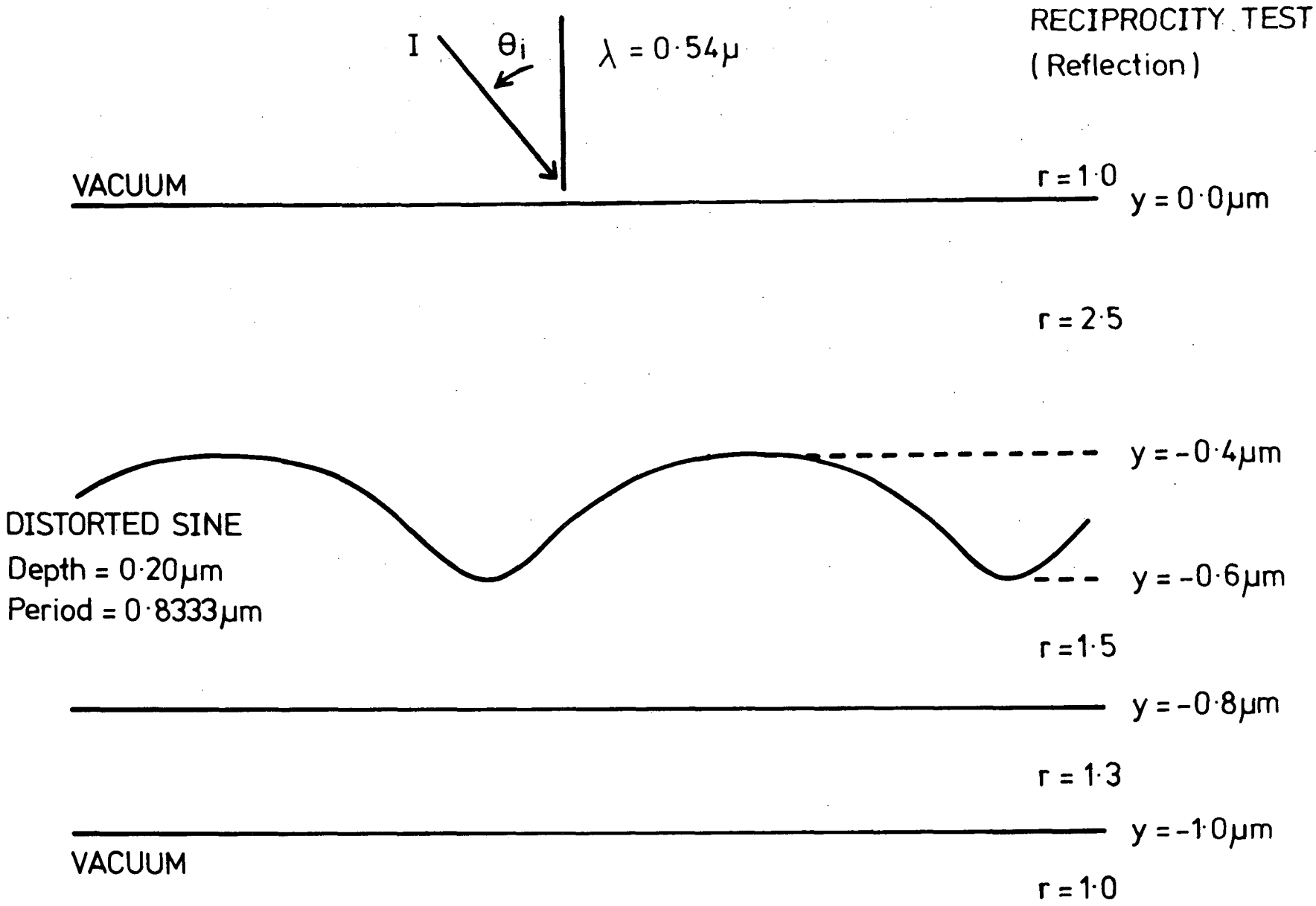


Figure (4.2b) The geometrical arrangement for the reciprocity test detailed in table (4.1b).

i.e. the efficiency of the transmitted order p is invariant under the conditions of inverse return. The entire formalism has been thoroughly tested against reciprocity and excellent agreement has been obtained. A summary of the results substantiating these remarks is presented in tables (4.1a) and (4.1b).

4.4.4 Agreement with Fresnel's laws

Any rigorous formalism for diffraction gratings must be capable of reproducing the results of Fresnel's laws when the groove depth is reduced to zero.

Consider a P polarized plane wave incident upon the optical flat illustrated in figure (4.1d). The structure consists of three separate domains D_0, D_1 and D_2 specified by their optical constants (μ_0,ε_0) (μ_0,ε_1) and (μ_0,ε_2). In an attempt to show that Fresnel's laws are a "subset" of the general diffraction formalism, the analytic solution of the integral equation

$$\tfrac{1}{2}\,\phi_1^g(M) - \int_{\Gamma_+} \left[\frac{\partial \phi_1^g}{\partial n'}\, G_1^E(M;M') - \phi_1^g\, \frac{\partial G_1^E}{\partial n'}\,(M;M')\right] ds'$$

$$= T_{1,0}^E \exp[i(\alpha_0 x - \Gamma_0 f(x))] \qquad (4.4.8)$$

is considered. Since all interfaces are planar, the field can be expanded in terms of Rayleigh expansions in all domains. It is now the aim of this section to demonstrate that such a structure can excite only field expansions containing zeroth order contributions and no higher order components. This is a reasonable assertion in view of the fact there exists no momentum transfer mechanism for the case of planar

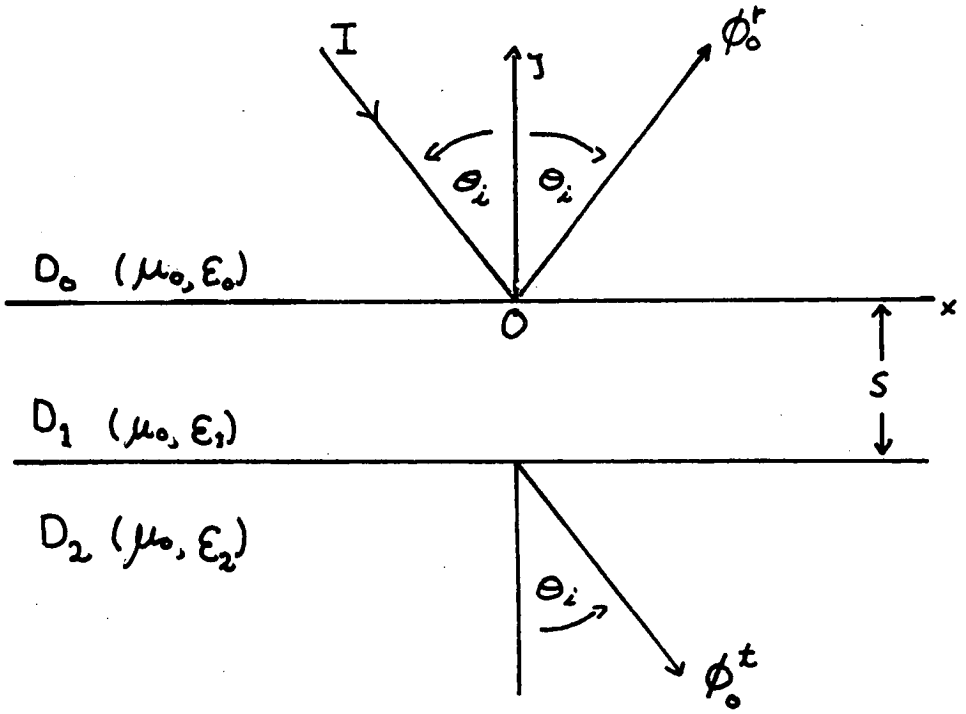


Figure 4.1d Fresnel's Laws.

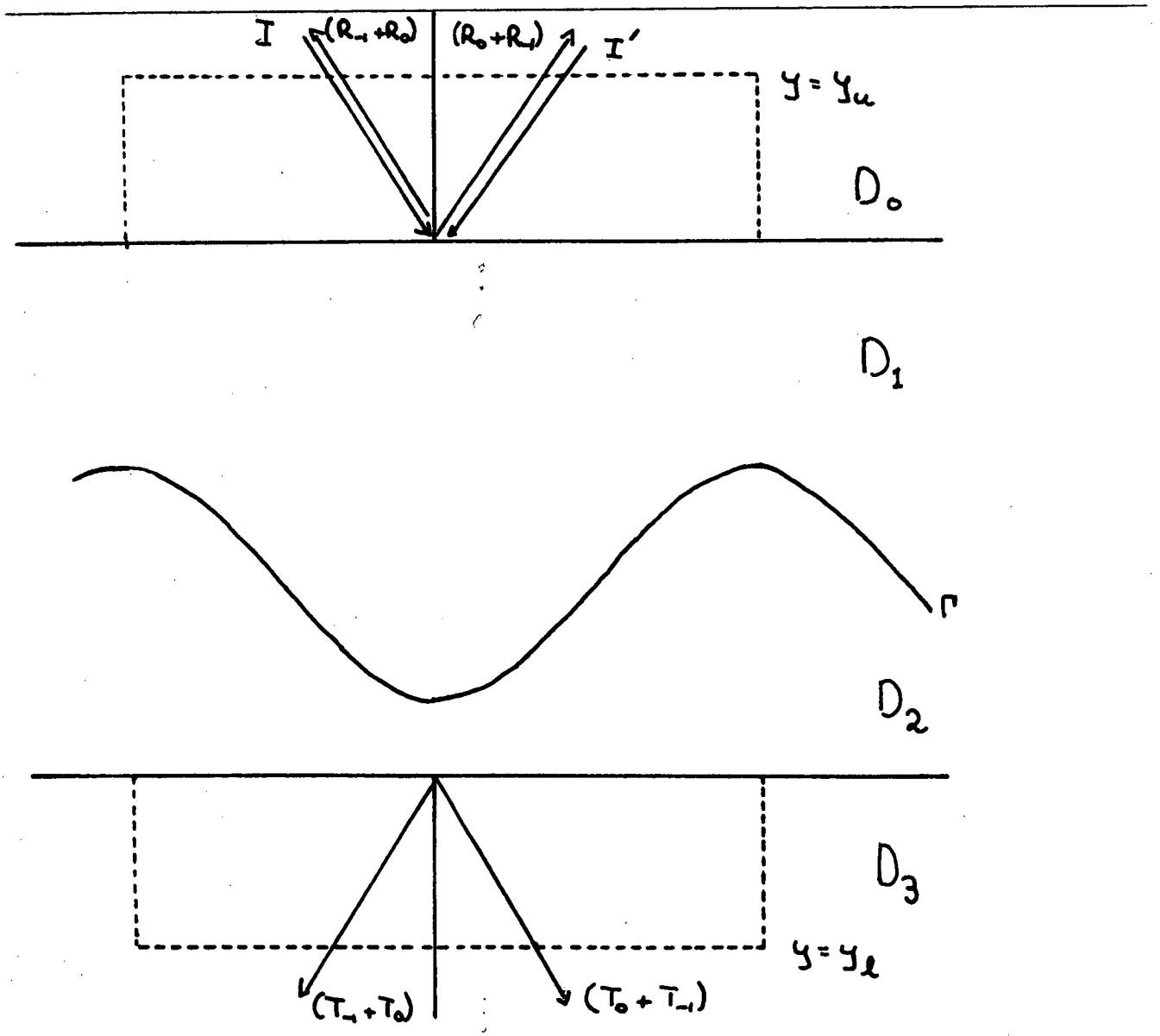


Figure 4.1e A first order Littrow mount for a left-right symmetric grating.

interfaces and thus no coupling between the incident field and the dispersive components can occur.

Now, at the interface y = - s the field can be written in the form:

$$\phi_1^g = \phi^t = \sum_n \overline{\phi_n^t} \exp(i\alpha_n x)$$

and its normal derivative expressed by

$$\frac{\partial \phi_1^g}{\partial y} = \frac{\partial \phi^t}{\partial y} = - \sum_n i\eta_n \overline{\phi_n^t} \exp(i\alpha_n x).$$

The Green's function G_1^E and its normal derivative are considerably simplified by the assumption of a plane interface at y = - s. In this case

$$G_1^E = \frac{1}{2id} \sum_n \frac{1}{\Gamma_n} (1+\hat{R}_{1,n}^E) \exp[i\alpha_n(x-x')] \qquad (4.4.9)$$

and

$$\frac{\partial G_1^E}{\partial y'} = - \frac{1}{2d} \sum_n \hat{R}_{1,n}^E \exp[i\alpha_n(x-x')] \qquad (4.4.10)$$

where $\hat{R}_{1,n}^E = R_{1,n}^E \exp(2i\Gamma_n s)$.

Thus, equation (4.4.8) reduces to the form

$$\tfrac{1}{2} \sum_n \left[1 - \hat{R}^E_{1,n} + \frac{\eta_n}{\Gamma_n}\left(1 + \hat{R}^E_{1,n}\right)\right] \overline{\phi^t_n} \exp[i\alpha_n(x-x')] = T^E_{1,0} \exp[i(\alpha_0 x + \Gamma_0 s)]$$

$$\left.\begin{aligned} \text{i.e.} \quad \phi^t_0 &= \frac{t_{01}t_{12}\exp(i\beta)}{1 + r_{01}r_{12}\exp(2i\beta)} \\ \text{and} \quad \overline{\phi^t_n} &= 0 \qquad \forall\ n \neq 0. \end{aligned}\right\} \qquad (4.4.11)$$

where $t_{01} = 2\chi_0/(\chi_0 + \Gamma_0)$

$t_{12} = 2\Gamma_0/(\Gamma_0 + \eta_0)$

$r_{01} = (\chi_0 - \Gamma_0)/(\chi_0 + \Gamma_0)$

$r_{12} = (\Gamma_0 - \eta_0)/(\Gamma_0 + \eta_0)$

and $\beta = \Gamma_0 s.$

Furthermore,

$$\left.\begin{aligned} \phi^r_0 &= \frac{r_{01} + r_{12}\exp(2i\beta)}{1 + r_{01}r_{12}\exp(2i\beta)} \\ \text{and} \quad \phi^r_n &= 0 \qquad \forall\ n \neq 0. \end{aligned}\right\} \qquad (4.4.12)$$

Expressions (4.4.11) and (4.4.12) are the classical forms of Fresnel's laws for a slab of thickness s. The program has been checked using this criterion and a sample result is presented below.

The physical data used for the test are summarized as follows:

$\lambda = 0.5\mu m$ $\quad \theta_i = 19.75^o$ $\quad s = 0.4\mu m$

$r_0 = 1.0$ $\quad r_1 = 2.5$ $\quad r_2 = 1.0$

	ρ_0^R	ρ_0^T
Program results	.9835	.0162
Fresnel's laws (P Pol.)	.9835	.0165

Although the comparison of results obtained from the grating theory and the predictions of Fresnel's laws does not have the same physical significance as reciprocity, the use of such a test has, nevertheless, proved to be a valuable tool in the elimination of minor algebraic and programming errors. As such, this property should be regarded as an initial verification of any grating formalism.

4.4.5 A phase constraint associated with the Littrow Mount

The initial investigation into the phase properties of a lossless symmetric grating used in a first order Littrow configuration was motivated by the need to explain the $\pi/2$ phase difference between the amplitudes of the $(-1)^{th}$ and 0^{th} orders for a perfectly conducting reflection grating. Since then, the analysis has been extended to include all lossless transmission gratings. In the case of a first order Littrow configuration with only two orders propagating in both reflection and transmission, the result may be stated as follows:

$$\frac{\cos(\delta_0^R - \delta_{-1}^R)}{\cos(\delta_0^T - \delta_{-1}^T)} = -\sqrt{\frac{\rho_0^T \rho_{-1}^T}{\rho_0^R \rho_{-1}^R}} \qquad (4.4.13)$$

where δ and ρ denote respectively the phase and efficiency of the various orders (denoted by the numerical subscripts) in either reflection or transmission (denoted by the superscripts R and T respectively).

The simplifying assumption, allowing only two orders in propagation, is not as restricting as it initially appears. Since the blaze region of sinusoidal and quasi-sinusoidal groove profile gratings operated in a $(-1)^{th}$ order Littrow mount lies between the (-2,+1) and (-1) Wood anomalies, it is felt that this phase constraint is of some relevance. The analysis has been generalized to cope with the propagation of more than two orders. However, the result is not as analytically elegant as that expressed in equation (4.4.13), and for this reason is not included in the current study.

Consider a plane wave incident upon the structure illustrated in figure (4.1e). The field distribution established by the P polarized incident wave is given by

$$\text{(i)}\quad E(x,y) = \sum_n [\exp(-i\chi_0 y)\ \delta_{n,0} + R_n \exp(i\chi_n y)] \exp(i\alpha_n x) \quad \text{in } D_0$$

and

$$\text{(ii)}\quad E(x,y) = \sum_n T_n \exp[i(\alpha_n x - \chi_n y)] \quad \text{in } D_3$$

where D_3 is now assumed to represent free space. This assumption is made in order to constrain the number of propagating reflected and transmitted orders to be always identical. Since the grating is being operated in a Littrow mount in the $(-1)^{th}$ order, the discrete values of α_n are specified by

$$\alpha_n = -\alpha_{-1-n} = (n + \tfrac{1}{2})\ K.$$

By now operating the grating in a Littrow mount in the $(+1)^{th}$ order, a corresponding wave field given by

$$\text{(i)}\quad E'(x,y) = \sum_n [\exp(-i\chi_0' y)\,\delta_{n,0} + R_n' \exp(i\chi_n' y)]\exp(i\alpha_n' x) \quad \text{in } D_0$$

and

$$\text{(ii)}\quad E'(x,y) = \sum_n T_n \exp[i(\alpha_n' x - \chi_n' y)] \quad \text{in } D_3$$

is established. In this case, the values of α_n' are specified by

$$\alpha_n' = (n - \tfrac{1}{2})K = -\alpha_{-n}.$$

After invoking the left-right symmetry of the optical arrangement, it follows that $R_n' = R_{-n}$ and $T_n' = T_{-n}$, and thus the above expressions for E' may be rewritten in the form

$$\text{(i)}\quad E'(x,y) = \sum_n [\exp(-i\chi_0 y)\,\delta_{n,-1} + R_{-1-n} \exp(i\chi_{-1-n} y)]\exp(i\alpha_n x) \text{ in } D_0$$

and

$$\text{(ii)}\quad E'(x,y) = \sum_n T_{-1-n} \exp[i(\alpha_n x - \chi_{-1-n} y)] \quad \text{in } D_3.$$

Now by defining the superposition of these two fields in regions D_0 and D_3 to be $E^T = E + E'$, a simple application of Green's theorem yields

$$\int_0^d [E^T \frac{\partial E^{T*}}{\partial y} - E^{T*} \frac{\partial E^T}{\partial y}]\Big|_{y=y_u} dx = \int_0^d [E^T \frac{\partial E^{T*}}{\partial y} - E^{T*} \frac{\partial E^T}{\partial y}]\Big|_{y=y_\ell} dx$$

where the asterisk denotes complex conjugation. (It should be noted that

the above relation holds only if all media are lossless.)

Upon explicit integration, the above expression simplifies to the form

$$|R_0 + R_{-1}|^2 + |T_0 + T_{-1}|^2 = 1. \qquad (4.4.14)$$

Since conservation of energy supplies the constraint

$$|R_0|^2 + |R_{-1}|^2 + |T_0|^2 + |T_{-1}|^2 = 1$$

equation (4.4.14) reduces to

$$\mathrm{Re}(R_0 R^*_{-1} + T_0 T^*_{-1}) = 0$$

or $$\frac{\cos(\delta_0^R - \delta_{-1}^R)}{\cos(\delta_0^T - \delta_{-1}^T)} = -\sqrt{\frac{\rho_0^T \rho_{-1}^T}{\rho_0^R \rho_{-1}^R}} \; . \qquad (4.4.13)$$

The same result has been derived for the orthogonal polarization. An important corollary of this result occurs in the case of a grating having a perfectly reflecting surface. Under these circumstances the result $\psi_0^R - \psi_{-1}^R = (\ell + \tfrac{1}{2})\pi$ (for an integer ℓ) may be derived.

These phase properties have proved themselves invaluable in the isolation of errors in both the formalism and the computer program. Table (4.2) presents numerical evidence confirming this property. Further discussion of such properties is presented in chapters 8 and 9.

Pol.	m	λ/d	ψ_{-1}^{R}	ψ_{o}^{R}	ψ_{-1}^{T}	ψ_{o}^{T}	ρ_{-1}^{R}	ρ_{o}^{R}	ρ_{-1}^{T}	ρ_{o}^{T}	$\frac{\cos(\psi_{o}^{R}-\psi_{-1}^{R})}{\cos(\psi_{o}^{T}-\psi_{-1}^{T})}$	$-\sqrt{\frac{\rho_{o}^{T}\rho_{-1}^{T}}{\rho_{o}^{R}\rho_{-1}^{R}}}$
P	20	1.2	165.79	-135.50	89.42	-129.95	.1682	.3732	.0723	.3715	- .672	- .654
P	30	1.2	164.96	-135.61	88.92	-130.32	.1684	.3794	.0722	.3758	- .657	- .652
S	20	1.2	-128.35	- 18.43	108.59	-171.36	.1231	.1169	.0839	.6749	-1.972	-1.984
S	30	1.2	-128.50	- 18.50	108.48	-171.47	.1231	.1167	.0840	.6759	-1.979	-1.988
P	20	1.6	- 83.18	-175.64	57.37	143.39	.1076	.5594	.2237	.1098	- .618	- .637
P	30	1.6	- 83.18	-175.34	57.34	143.45	.1074	.5598	.2233	.1096	- .633	- .638
S	20	1.6	133.50	4.91	85.36	161.03	.0070	.1205	.0831	.7191	-2.520	-2.538
S	30	1.6	133.45	4.91	85.23	160.97	.0771	.1205	.0832	.7191	-2.529	-2.538

TABLE 4.2

Results for both fundamental polarizations of light confirming the Littrow phase property. m denotes the number of sampling points per period. The geometry of the optical arrangement is shown in figure (4.3).

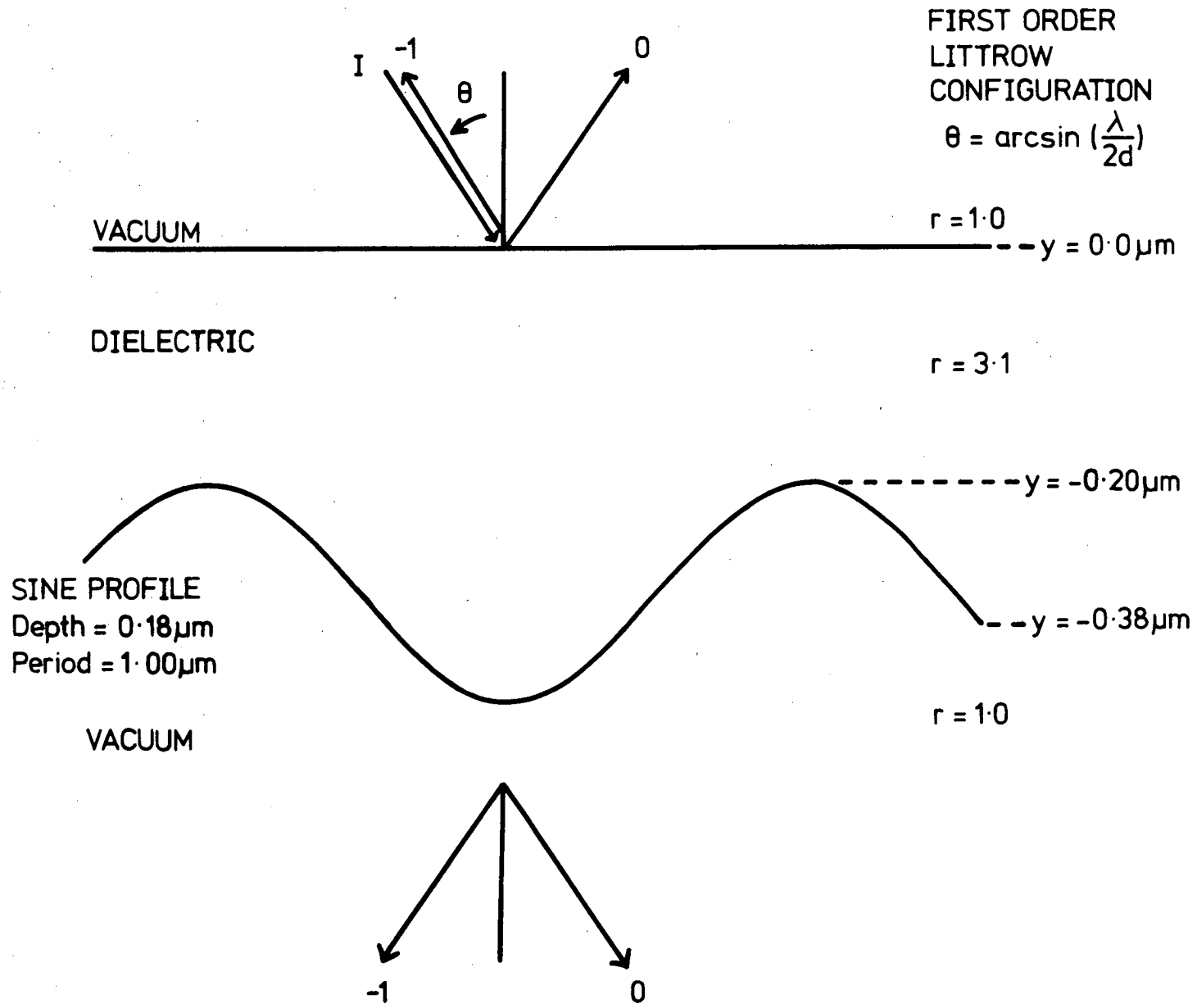


Figure 4.3 The optical arrangement for the Littrow phase property detailed in table (4.2). The high value of the refractive index (3.1) was chosen to provide a sufficiently large reflected energy component.

4.5 CONCLUSIONS

This chapter has outlined a theoretical formalism which reduces a wide range of diffraction grating problems to the solution of a single Fredholm integral equation of the first kind. A new phase constraint, appropriate to lossless symmetric gratings operated in a first order Littrow configuration, has been derived and has shown itself to be a powerful test of the theory.

The numerical implementation presents no difficulties not already circumvented in the solution of previous diffraction problems. A computer program written in FORTRAN IV and operated on Burroughs B6700 (Hobart) and Control Data Cyber 72 (Sydney) computers has been used to verify the accuracy of the method. Its generality is illustrated by figure (4.4), which shows the variety of problems which can be solved simultaneously by the algorithm. The program is almost as generally applicable as that devised by Nevière et al [4.5] but does not encounter the same serious numerical difficulties associated with highly conducting metals and S polarized radiation.

In the following chapter, an application of this theory pertinent to the field of solar energy research is discussed. The investigation, concerning the effects of surface roughening on the selective properties of thin film solar absorbers, was instigated at the suggestion of Dr. Ross McPhedran and Dr. David McKenzie of the Solar Energy Group of the University of Sydney.

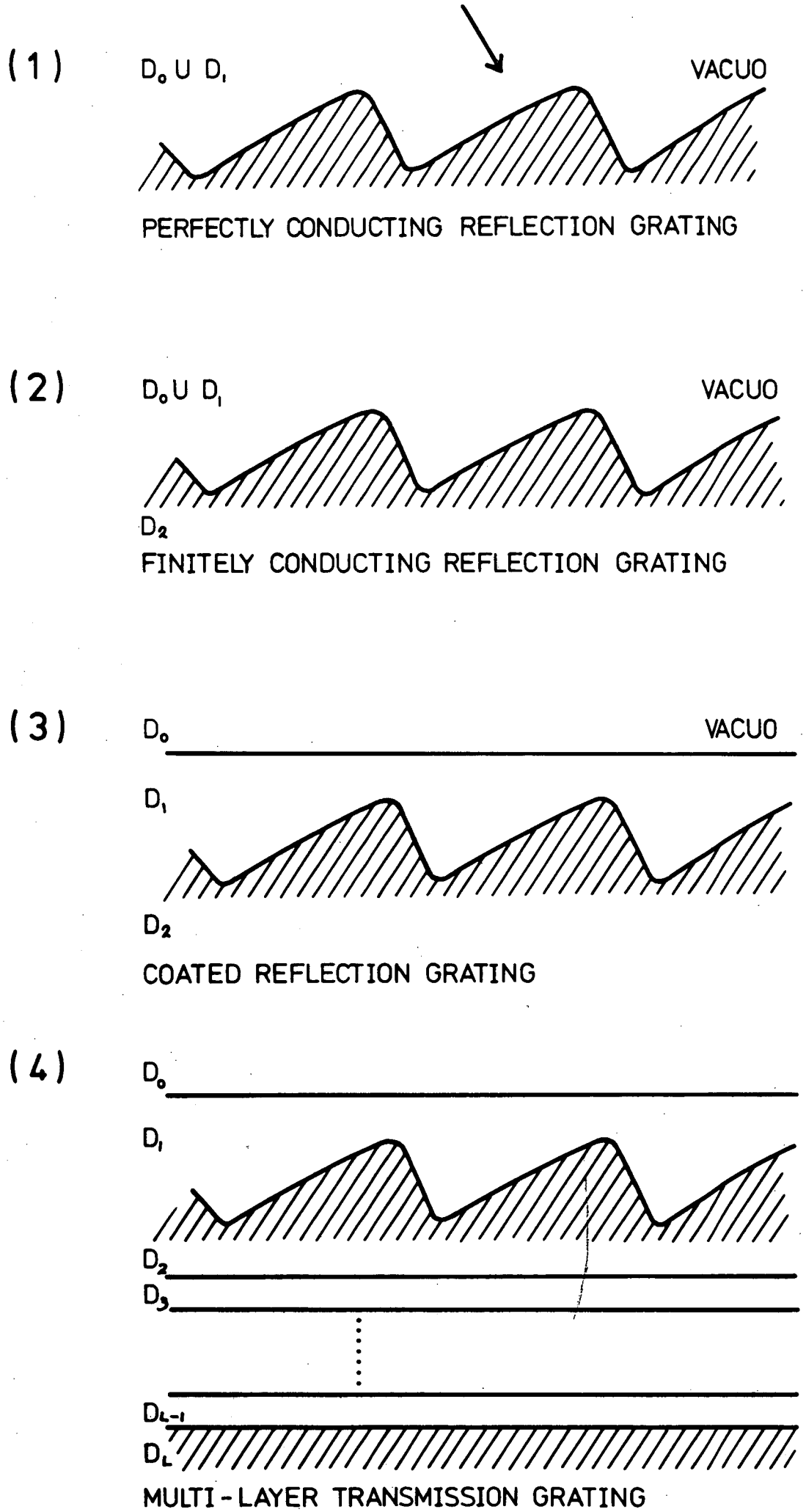


Figure 4.4 The four diffraction problems solved simultaneously by the program.

REFERENCES

[4.1] L. C. Botten, accepted for publication by Optica Acta.

[4.2] L. C. Botten, (1977), University of Tasmania, Report DGRG 77/2.

[4.3] P. M. Van den Berg, Ph.D. thesis, (1971), Delft University of Technology, Delft, The Netherlands.

[4.4] D. Maystre, Thèse No. AO 9545, (1974), L'Université d'Aix-Marseille III, France.

[4.5] M. Nevière, P. Vincent, and R. Petit, Nouv. Rev. Optique, 5 (1974) 65.

[4.6] D. Maystre, personal communication.

[4.7] R. Petit and M. Cadilhac, C. R. Acad. Sci. Paris, 262 (1966) 468.

[4.8] M. Nevière and M. Cadilhac, Optics Commun., 2 (1970) 235.

[4.9] J. Pavageau and J. Bousquet, Optica Acta, 17 (1970) 469.

[4.10] R. F. Harrington, (1968), "Field Computation by Moment Methods", (New York: McMillan) Pp 9-11.

[4.11] R. Petit, Rev. Optique, 45 (1966) 249.

[4.12] J. L. Uretsky, Ann. Phys., 33 (1965) 400.

[4.13] R. Petit, Optica Acta, 14 (1967) 301.

[4.14] D. Maystre and R. Petit, Optics Commun., 4 (1972) 380.

[4.15] D. Maystre and R. McPhedran, Optics Commun., 12 (1974) 164.

[4.16] R. McPhedran and D. Maystre, (1976), School of Physics, University of Sydney, internal report SP76-1.

[4.17] R. McPhedran and D. Maystre, Appl. Phys., 14 (1977) 1.

[4.18] M. Nevière and P. Vincent, Optica Acta, 23 (1976) 557.

CHAPTER 5

IMPROVEMENTS IN THE DESIGN OF SOLAR SELECTIVE THIN FILM ABSORBERS

This chapter is based on a paper which has been accepted for publication in the journal, "Optics Communications". The study, which was undertaken in collaboration with Mr. Ian Ritchie of the Solar Energy Group of the University of Sydney, represents an application of the theory presented in the previous chapter. It is a pleasure to thank Dr. Ross McPhedran and Dr. David McKenzie for suggesting the topic and for their sustained interest and many valuable discussions during the course of the investigation.

5.1 INTRODUCTION

It has long been recognised that surfaces which selectively absorb solar radiation can be made by depositing a stack of thin films upon a highly reflecting metal substrate. By an appropriate choice of the optical properties and thickness of the film stack, it is possible to design a collector absorbing in the wavelength range of 0.2μm to 2.0μm yet transparent to infrared radiation (i.e. having a low thermal emittance). In a previous study undertaken by Ritchie [5.1] it has been shown that a single layer interference film can produce solar absorptance values greater than 90% provided that $n < 1.5$ and $0.2 < k < 0.8$ (where n and k are the real and imaginary parts of the complex refractive index).

However as the performance of these absorbers can be degraded either by ion diffusion or by poor film adhesion during frequent thermal cycling over large temperature ranges, it is imperative that the above theoretical constraint be tempered by this more practical consideration of selecting materials having a suitable refractory nature. In view of this restriction, it may be necessary to use absorbing films having less than optimal characteristics.

Thus it was the aim of these studies to

(i) determine whether the integrated absorptivity could be improved by surface roughening and

(ii) compare the overall performance of a roughened surface with that obtained by grading the refractive index of the absorber, a discussion of which has already been presented by Ritchie and Window [5.1].

Although experimental work [5.2] has shown that surface roughening produces beneficial results, this study to our knowledge, is the first theoretical attempt to understand the effects of the roughening parameter. Clapham and Hutley [5.3] in discussing the reduction of lens reflection by the "moths-eye" principle postulated that the diminished reflection losses of a biperiodic grating structure could be explained by the refractive index of the lens material being graded with air (thereby more adequately matching the impedances of free space and the lens material). This same concept of refractive index grading has been adopted by Handa et al [5.4] and Peng and Tamir [5.5] to study the properties of dielectric lamellar gratings. Jacobsson and Martensson

[5.6] used a similar approach in their work on graded anti-reflection coatings.

Our treatment has utilized a rigorous electromagnetic theory developed recently (see chapter 4) to consider a single dimensional roughening modelled by a sinusoidal grating profile. It has been our experience that the effects of surface roughening and refractive index grading are markedly different with roughening producing rather localized improvements in absorptivity. By locating these grating absorption features at those wavelengths for which the absorptivity produced by interference action was lowest, it was possible to raise the overall absorptance without significantly affecting the integrated emittance.

On the other hand, refractive index grading, by minimising reflection losses (due to the more adequate matching of the impedances of free space and the substrate) tends to smooth the entire absorption spectrum thereby removing any deep absorption minima resulting in an overall improvement in the integrated absorptance. However, grading can produce an undesirable side effect in moving the cut-off wavelength further towards the infrared thus increasing the thermal emittance.

5.2 SURFACE ROUGHENING

The structure studied consisted of an idealized graphite layer deposited on a copper substrate and used in normally incident radiation. Due to the absence in the available literature of adequate refractive index data for graphite, it was necessary to use the wavelength independent value of $(1.95 + i\ 0.34)$ throughout. Refractive index values of copper were taken from the A.I.P. Handbook [5.8].

As has already been mentioned we modelled a single dimensional surface roughening using a sinusoidal profile and thus the parameters of our study became:

(i) the mean thickness (t) of the layer,

(ii) the period (d) of the surface modulation,

(iii) the choice of roughening either the air-graphite (A-G) interface or the graphite-copper (G-C) interface and

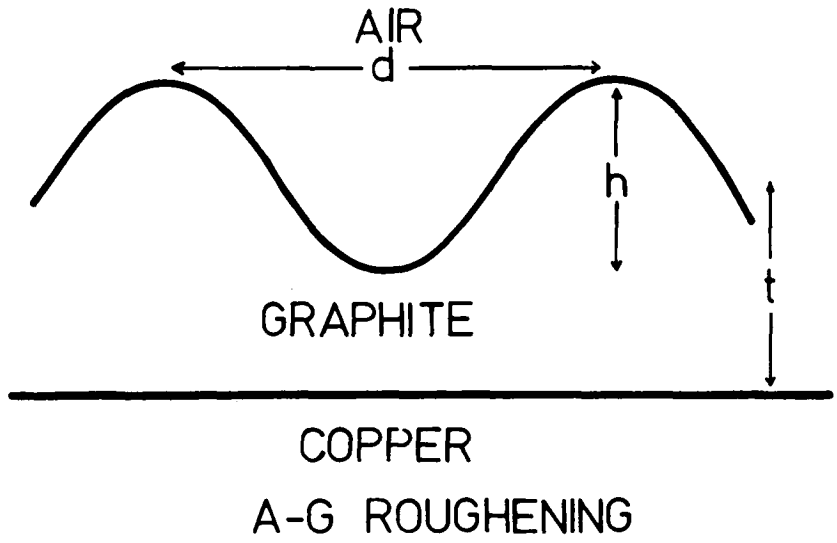
(iv) the depth (h) of the roughening.

The structures modelled are depicted in figure 5.1.

5.2.1 The choice of an appropriate graphite layer thickness

This parameter was chosen with the aim of achieving a good absorptance value (a) and a good absorptance to emittance ratio (a/e) (assuming an emitter temperature of 700 K) for the unroughened structure. Using the above refractive indices, we surmised that a suitable thickness would be 0.12μm. With this value of t the integrated absorptance and integrated emittance were calculated to be 76% and 3.5% respectively. As can be seen from table (5.1) and figure (5.2) plane structures having a layer thickness of less than 0.12μm have inadequate solar absorptance since the cut-off appears at too short a wavelength. On the other hand, a layer thickness much in excess of 0.12μm extends the cut-off wavelength well into the infrared thereby contributing to an increased thermal emittance.

The rather disappointing absorptance value results from the non-optimal nature of graphite (contributing significantly to rather large reflection losses) and also from the presence of a deep absorption minimum coinciding with the peak of the solar spectrum.



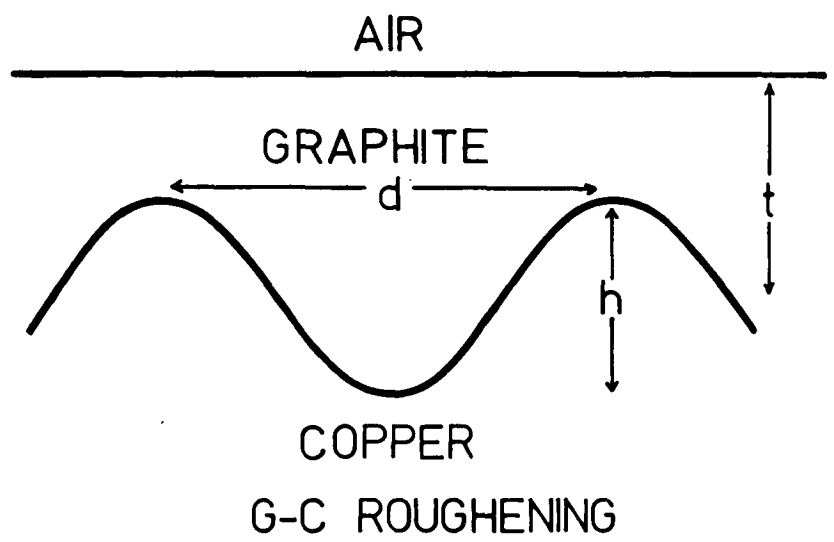


Figure 5.1 : The geometry of the structure used in modelling A-G and G-C roughening.

TABLE 5.1

t (μm)	a (%)	a/e
0.05	60.1	35
0.08	71.0	31
0.10	74.0	25
0.12	76.3	20
0.15	78.7	13
0.20	81.5	7.4
0.30	84.4	3.4
0.40	86.0	2.1

Variation of absorptance and emittance values for a uniform graphite film with thickness.

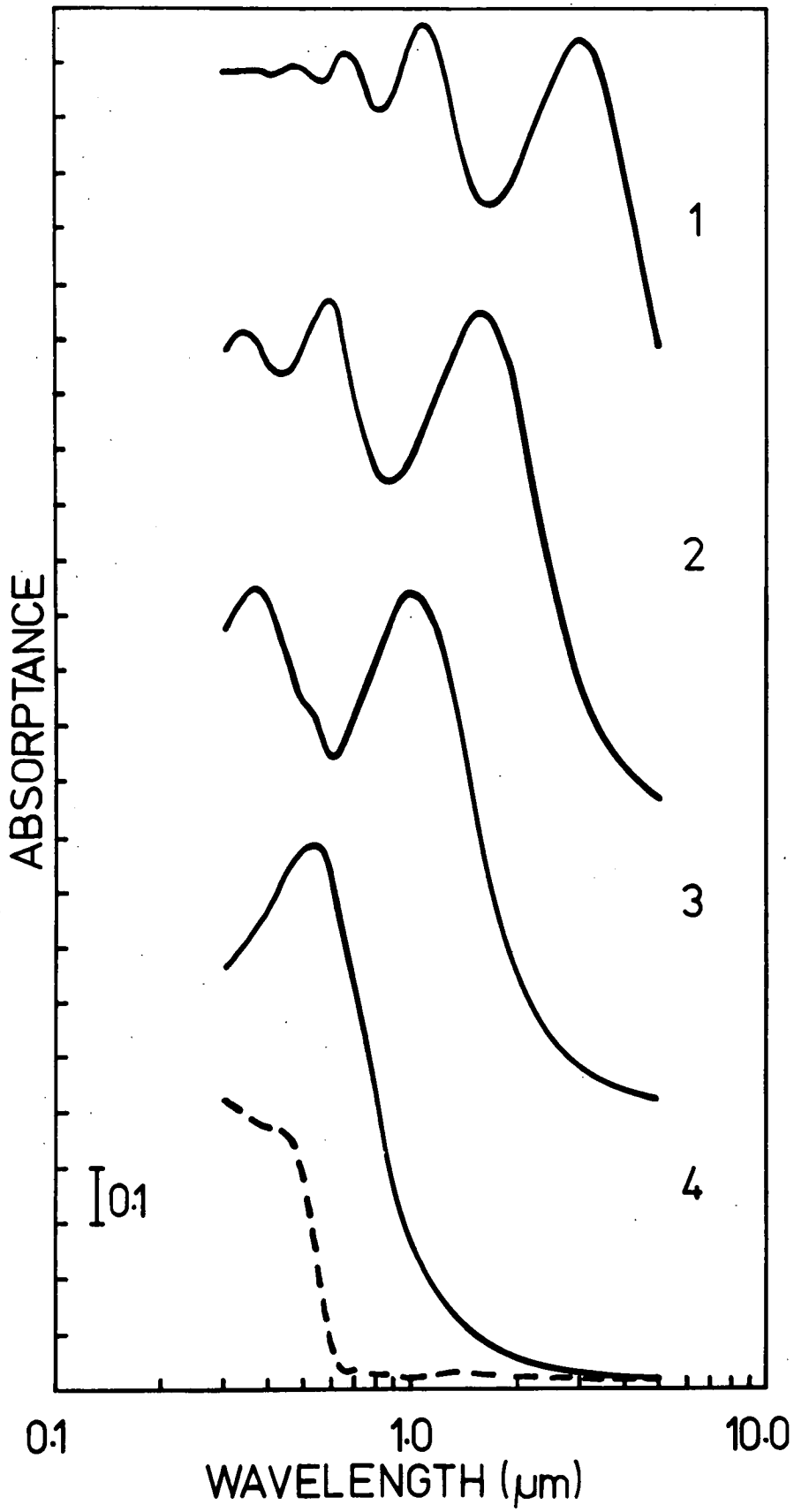


Figure 5.2 : Effect of layer thickness on the absorptance of an unroughened uniform film. Shown here are curves for thickness of

(1) 0.40μm (2) 0.20μm (3) 0.12μm

(4) 0.05μm (5) 0.00μm (corresponding to the absorptivity of a copper substrate).

5.2.2 The choice of the scale of roughness

It was then our objective to improve the general performance by eliminating the deleterious absorption minimum. This was achieved by using to advantage a grating absorption feature known as a surface plasmon, a phenomenon which has long plagued experimental physicists using diffraction gratings for spectrographic purposes. A substantial review of the study and the history of plasmons has been presented by Wheeler et al [5.9]. The very nature of these plasmons (or surface plasma oscillations) which are excited on the long wavelength side of Wood anomalies is determined by the shape of the roughened surface and also the optical properties of the various media surrounding the interface. The excitation of a plasmon always represents an additional loss or absorption mechanism and often they manifest themselves as deep and relatively sharp absorption bands (in the case of metallic reflection gratings).

Thus, to excite a surface plasmon centred on the absorption minimum at approximately 0.6μm, it was necessary to station a Wood anomaly at a wavelength of 0.5μm. In the case of normally incident radiation this resulted in the grating period being chosen to be 0.5μm.

5.2.3 The effects of surface roughening

The results of our studies for both roughened A-G and G-C interfaces are presented in figures (5.3) and (5.4) and table (5.2). Perhaps the most striking feature of these graphs is the polarization independent behaviour exhibited by shallow gratings ($h < .1\mu m$). This is taken to mean that the values of a and e obtained for this one dimensional roughening give a good indication as to the values which could be

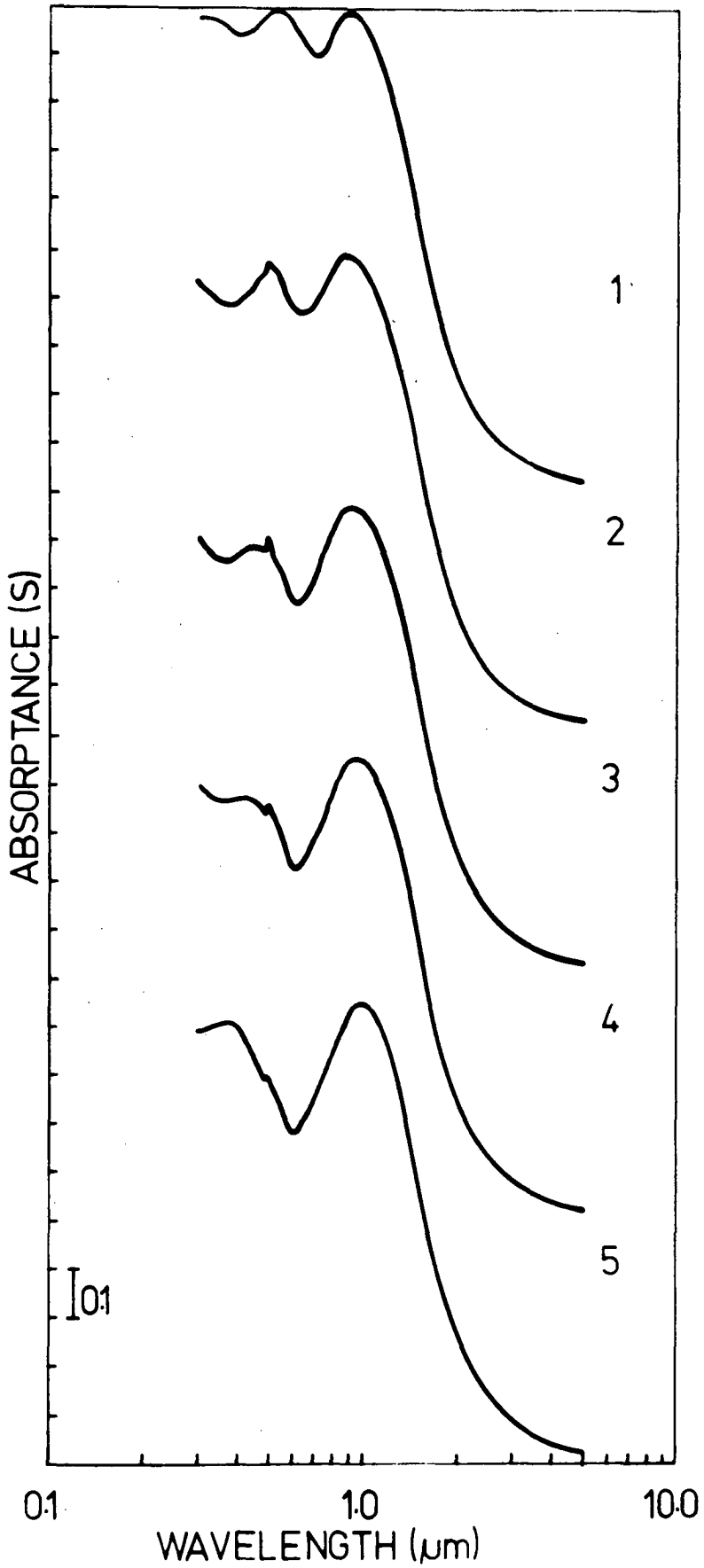


Figure 5.3 a : Variation of absorptance(in S polarized light) with groove depth for A-G roughening characterized by a groove period of 0.50μm and a mean thickness of 0.12μm. Depicted are curves for groove depths of

(1) 0.20μm (2) 0.15μm (3) 0.10μm

(4) 0.08μm (5) 0.05μm

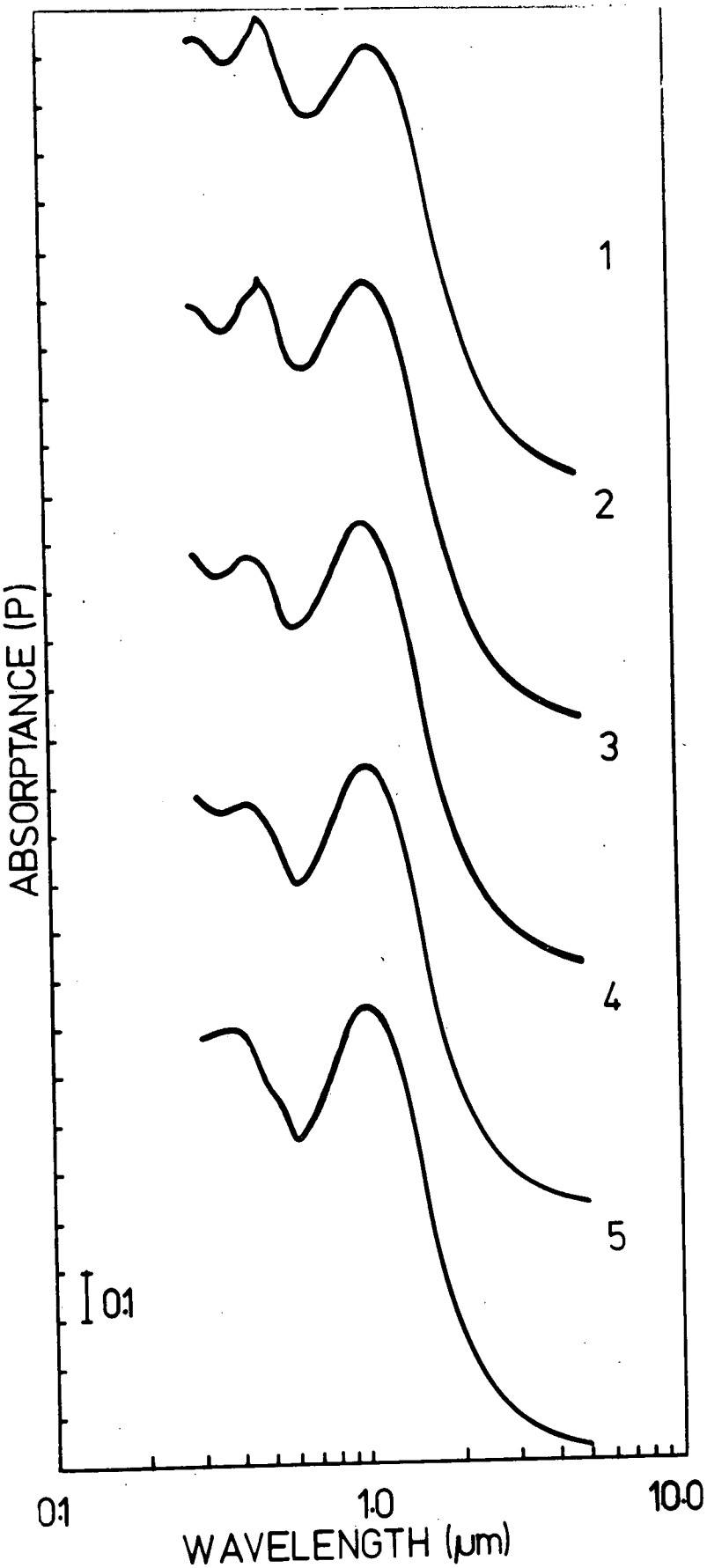


Figure 5.3b : Variation of absorptance (in P polarized light) with groove depth for A-G roughening characterized by a groove period of 0.50μm and a mean thickness of 0.12μm. Depicted are curves for groove depths of

(1) 0.20μm (2) 0.15μm (3) 0.10μm

(4) 0.08μm (5) 0.05μm

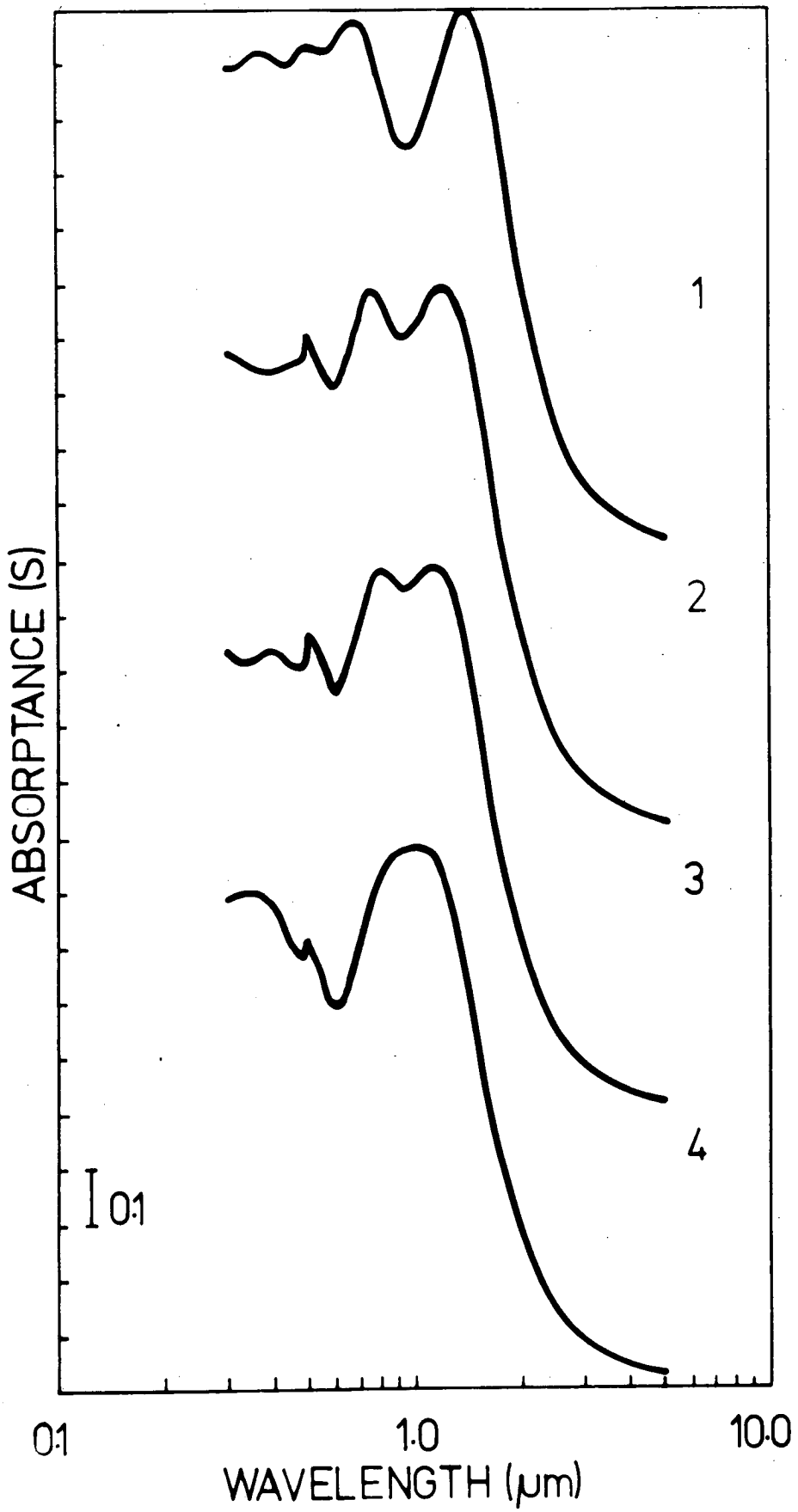


Figure 5.4a : Variation of absorptance (S polarization curves are given) with groove depth for G-C roughening specified by the same parameters as for A-G roughening. Presented are curves for depths of (1) 0.15μm (2) 0.10μm (3) 0.08μm (4) 0.05μm.

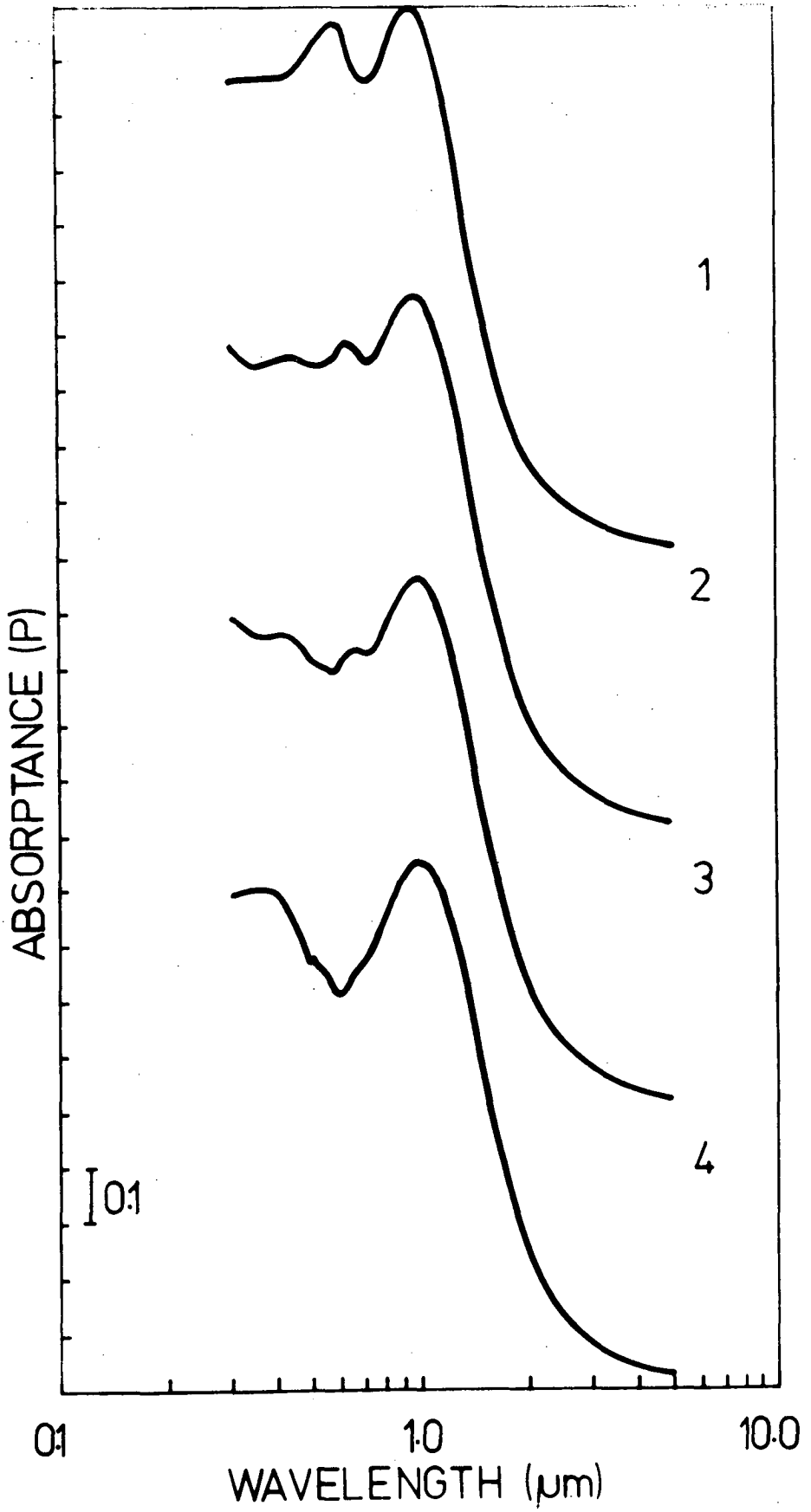


Figure 5.4b : Variation of absorptance (P polarization curves are given) with groove depth for G-C roughening specified by the same parameters as for A-G roughening. Presented are curves for depths of

(1) 0.15μm (2) 0.10μm (3) 0.08μm (4) 0.05μm

TABLE 5.2

Roughening Type	h (μm)	a (%)	e (%)	λ_c (μm)
A-G	0.00	76.3	3.5	1.58
"	0.05	75.1	3.8	1.58
"	0.08	77.2	3.8	1.57
"	0.10	81.8	3.9	1.56
"	0.15	83.0	4.1	1.54
"	0.20	88.4	6.2	1.54
G-C	0.00	76.3	3.5	1.58
"	0.05	81.2	3.8	1.64
"	0.08	85.4	3.9	1.72
"	0.10	87.2	4.0	1.74
"	0.15	87.8	4.4	1.98

The effect of groove depth on the spectral performance of a roughened surface characterized by a mean graphite thickness of 0.12μm and a roughness period of 0.5μm.

obtained using two-dimensional roughening. In all cases however, the S polarization results are a more accurate representation of the actual physical phenomena associated with roughening. With any type of roughening, the plasmons, which can be excited by any surface irregularity, propagate across the corrugations as is the case with S polarized radiation.

We note that the excitation of plasmons occurs in both P and S polarized light. This is to be contrasted with highly conducting metals, where only S polarized plasmons are observed. However, our results are in accordance with the theoretical studies of Hutley et al [5.10] and the experiments of Palmer [5.11] who showed that P polarization diffraction anomalies could be greatly enhanced by overcoating the grating with a dielectric material.

For both A-G and G-C roughening we see a continuous improvement in the spectral performance with increasing groove depth. This improvement is impaired in the case of G-C roughening only for relatively deep grooves ($h < 0.15\mu m$). In this case the presence of a deep resonance anomaly at a wavelength of approximately $1.0\mu m$, enhanced by increasing groove depth, eventually reverses the upward trend in absorptance values. This leads us to believe that to achieve good performance throughout the entire wavelength range, the field on the groove tops should not be considerably stronger than the field in the groove troughs. In contrast with the results obtained for A-G roughening, for which improvement is monotonic with increasing groove depth, G-C roughening using deep grooves does tend to promote an increased 'relative' field activity on the groove tops due to the attenuation of

the incident field in propagating through the graphite to the groove troughs.

Another significant feature of our results occurs for groove depths, $h < 0.10\mu m$. As can be seen from the accompanying graphs and tables, the improvement in integrated absorptance with increasing groove depth is more gradual for A-G roughening than for G-C roughening. Wheeler [5.9] has shown that the probability of plasmon excitation at the grating surface is considerably higher if one of the media is metal-like. For G-C roughening, the copper substrate on the lower side of the interface fulfils this criterion and so enhances the chance of establishing a plasmon at a given groove depth even though the incident field actually reaching the grating surface has been diminished by absorption within the graphite. For A-G roughening of the same groove depth, the graphite is not sufficiently 'metal-like' to give rise to a plasmon of the strength required to counter the presence of the interference absorption minimum. Thus to create a plasmon whose strength is comparable with one obtained using G-C roughening, it is necessary to use increased groove depths.

It is interesting to note that reflection losses for wavelengths less than $0.5\mu m$ also appear to have been reduced even though additional orders of diffraction are available to carry the reflected energy. Clapham and Hutley [5.3] argued that the period needed to be sufficiently small ($d < 0.3\mu m$) to disallow any additional modes of energy propagation. However, such a choice of period would shift the previously mentioned plasmon to shorter wavelengths thereby leaving the primary absorption minimum caused by interference action virtually unaffected.

The position of the cut-off wavelength, λ_c, (i.e. the wavelength for which a = 50%) is also affected by groove depth. With A-G roughening, λ_c moves only slightly towards shorter wavelengths. However with G-C roughening the cut-off wavelength moves towards the 'red' as groove depth increases thereby increasing the absorptance but also the thermal emittance.

We have also investigated the spectral performance of a structure with an absorbing layer having an optimal refractive index of 1.4 + i 0.5 and thickness of 0.179μm. We have achieved some improvement with the integrated absorptance being raised from 87% to 90% (using 'air-absorber' roughening characterized by a period of 0.58μm and a depth of 0.2μm). However as the interference action is only very weak, the action of the grating in reducing its effects cannot be as dramatic.

For this reason, we believe that surface roughening of this type can be usefully applied only to single layer structures whose absorbing medium is of a non-optimal nature. We further stress that although we have modelled the roughening with a uniperiodic structure, our results give a good indication of the performance obtainable with an arbitrary form of roughening. In particular, the S polarization results are appropriate since it is possible to establish surface plasmons which propagate across the corrugations for any type of roughening.

5.3 REFRACTIVE INDEX GRADING

In this section we discuss two different types of refractive index matching, involving grading the graphite layer with air and copper thereby effectively smoothing the A-G and G-C interfaces. This was done

in an attempt to correlate these results with those obtained for A-G and G-C roughening respectively. All calculations were performed by adjusting the layer thickness and the grading profile so that the resultant structure contained an amount of graphite equivalent to a uniform film of thickness 0.12μm.

The grading profile is characterized by a parameter (γ) originally defined by Jacobsson [5.12]. In describing the film grading of two materials A and B

$$\gamma = \frac{r_B(0)}{r_A(T)}$$

where $r_A(0) = 0$, $r_B(T) = 0$ and $r_i(t)$ is the rate of deposition of material i at some time t varying linearly with time from t = 0 to t = T. Using the same linear mixing theory [5.12], the dielectric constant of the composite layer (of thickness t) at some depth y is given by

$$\varepsilon(y) = \frac{\varepsilon_A - \gamma\,\varepsilon_B}{1 - \gamma} - \frac{\gamma(\varepsilon_A - \varepsilon_B)}{(1-\gamma)\,[1 - (1-\gamma^2)y/t]^{\frac{1}{2}}} .$$

A full description of the grading profiles so obtained is given in a recent paper by Ritchie [5.1].

Although linear mixing theory is entirely adequate for the case of A-G grading [5.1], we note that it may not be rigorously applicable to G-C grading. In this case the more rigorous approach adopted by McKenzie and McPhedran [5.13, 5.14, 5.15] may be needed. However for simplicity the linear mixing theory is used here.

5.3.1 A-G Grading

Here γ is given by

$$\gamma = \frac{r_{air}(0)}{r_{graphite}(T)} \quad .$$

Naturally for low values of γ the absorption spectrum closely resembles that obtained for a uniform film. Figure (5.5) and table (5.3) demonstrate that as γ increases, thereby effectively smoothing the refractive index variation across the A-G interface, the interference action becomes weaker and absorptance values rise. However, the overall improvement in absorptance, gained by a more adequate matching of optical impedances, must be tempered by noting that the film now has a decreased infrared reflectance. Thus, on the basis of achieving both good absorptance and low emittance values we suggest that $\gamma = 1$ is most appropriate.

5.3.2 G-C Grading

For this type of grading

$$\gamma = \frac{r_{graphite}(0)}{r_{copper}(T)}$$

and thus large values of γ most closely resemble a uniform layer. As γ decreases, the movement of the cut-off towards shorter wavelengths causes diminished values of the absorptance. Our results, figure (5.6) and table (5.3), show the mean refractive index of the layer to be dominated by copper, thereby contributing to a substantially curtailed absorptance spectrum (which in the limit as γ approaches zero tends to

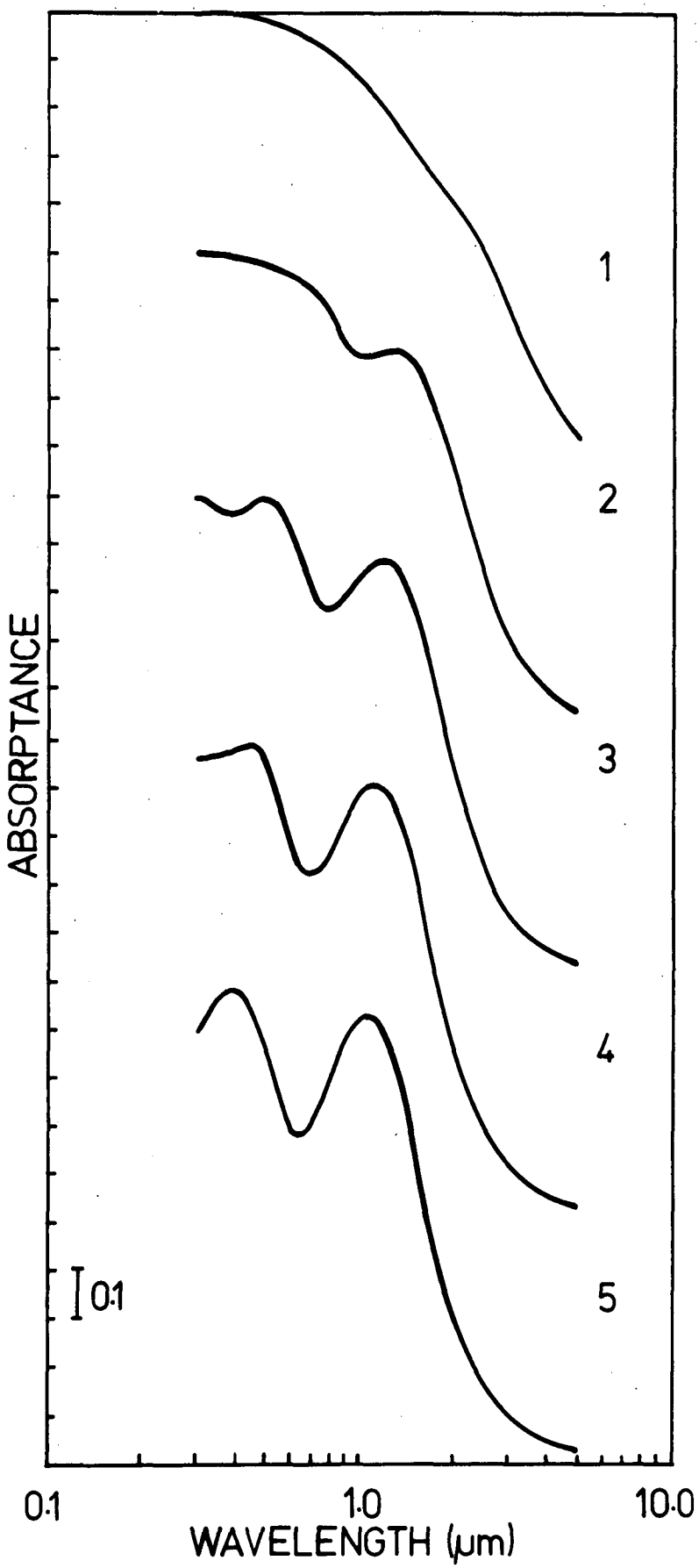


Figure 5.5 : Effect of the parameter γ on the absorptance of an A-G graded film. Shown here are curves for the following values of γ

(1) 5.0 (2) 2.0 (3) 1.0 (4) 0.5 (5) 0.2

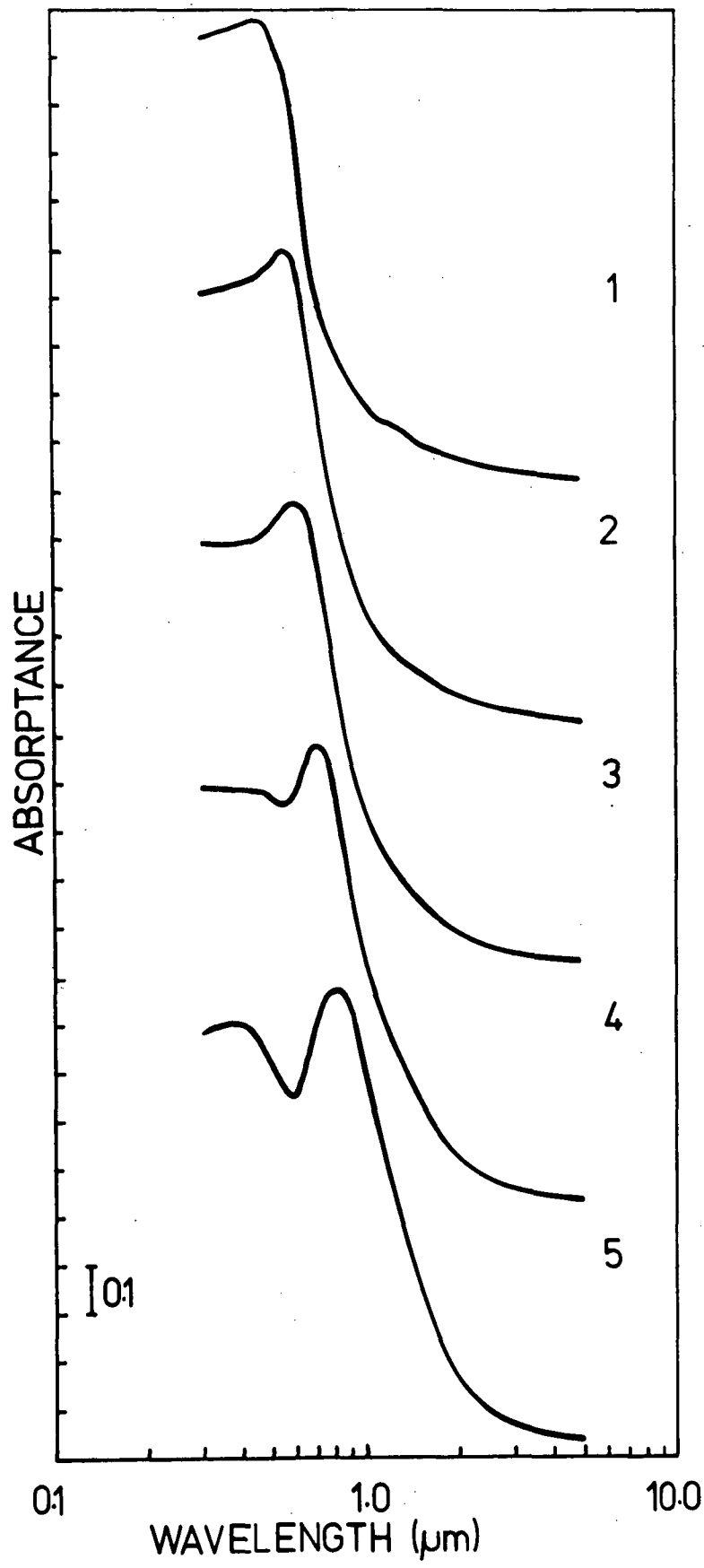


Figure 5.6 : Variation of absorptance with γ for a G-C graded film. Curves for the following γ are presented

(1) 5.0 (2) 2.0 (3) 1.0 (4) 0.5 (5) 0.2

TABLE 5.3

Grading Type	t (μm)	γ	a (%)	λ_c (μm)
A-G	.1477	0.2	78.7	1.65
"	.1888	0.5	81.9	1.75
"	.2586	1.0	85.3	1.87
"	.3994	2.0	87.8	2.10
"	.8247	5.0	89.7	2.45
G-C	.8247	0.2	49.8	0.66
"	.3994	0.5	58.5	0.76
"	.2586	1.0	66.3	0.90
"	.1888	2.0	72.2	1.02
"	.1477	5.0	76.5	1.28

The effect of the parameter, γ, on the spectral performance of a graded film containing a mass thickness of 0.12μm of graphite.

that of pure copper). However this performance could be improved by using an increased mass thickness of graphite to shift the cut-off towards longer wavelengths and still maintain the sharp filtering action exhibited in the curve for $\gamma = 0.2$.

5.4 CONCLUSIONS

We have presented evidence of the effectiveness of both roughening and refractive index grading in improving the performance of a selective surface. Although both techniques are capable of reducing or eliminating undesirable interference features, the mechanisms by which they accomplish this objective are entirely dissimilar.

In the case of surface roughening, the interference action is masked effectively by the excitation of surface plasmons, active only over restricted wavelength ranges. Results obtained using our relatively simple model lead us to believe that roughening in a random fashion to the scale of approximately 0.5μm would effectively remove the undesirable interference features. Those wavelengths longer than 1.5μm are unable to resolve corrugations of this magnitude. The filtering action is thus unaffected and the emittance values substantially unaltered. The polarization independent performance exhibited by a one-dimensional roughening leads us to believe that the results obtained from a two dimensional roughening would not be dissimilar. It is known [5.16] that only a very small number of grooves is required before the energy properties of a structure become equivalent to those of a grating and thus only local regularity of the roughened surface is needed to ensure applicability of the results presented here.

On the other hand, refractive index grading, by smoothing the optical discontinuities between the various media, thus achieves a more effective impedance matching than does a uniform film. It improves absorptance values by eliminating interference action. However, grading does effect the entire spectrum (unlike roughening) and tends to shift the cut-off towards longer wavelengths resulting in increased emittances. Better filtering could be obtained by grading the absorber material fairly sharply with the substrate material and using a layer thickness substantially greater than those referred to in the previous section. However too great a thickness would give rise to bulk absorption with a concomitant increase in emittance.

Although our current study has been restricted to normally incident radiation, work is proceeding, in the case of roughening, to examine the effects of hemispherically incident radiation on collector performance. Ritchie and Window [5.1] have already shown that results obtained for graded films used in off-axis configurations are similar to those obtained for on-axis operation except in that the optimum layer thickness is somewhat greater.

In summary, we stress that although both grading and roughening have merit in enhancing the spectral performance of a selective surface, there can be no universal panacea since the design of any solar collector must be governed by the task that it is expected to undertake. Experimental work on techniques of producing roughening of selective absorbers on the scale defined in this chapter would be valuable. Until such work is undertaken, it will not be possible to judge whether surface roughening can make the transition from a theoretically desirable feature to one of practical importance.

REFERENCES

[5.1] I. T. Ritchie and B. Window, Appl. Opt., 16 (1977) 438.

[5.2] B. O. Seraphin and A. B. Meinel, "Optical Properties of Solids - New Developments", B. O. Seraphin, Editor, North-Holland, Amsterdam (]976), Chap. 17.

[5.3] P. B. Clapham and M. C. Hutley, Nature, 244 (1973) 281.

[5.4] K. Handa, S. T. Peng and T. Tamir, Appl. Phys., 5 (1975) 325.

[5.5] S. T. Peng and T. Tamir, Appl. Phys., 7 (1975) 35.

[5.6] R. Jacobsson and J. O. Martensson, Appl. Opt., 5 (1966) 29.

[5.7] L. C. Botten, submitted for publication to Optica Acta.

[5.8] American Institute of Physics Handbook, Ed. 3, McGraw-Hill, New York (]972), 6-133.

[5.9] C. E. Wheeler Jr., E. T. Arakawa and R. M. Ritchie, Oak Ridge National Laboratory, Report ORNL/TM-5185 (1976).

[5.10] M. C. Hutley, J. F. Verrill, R. C. McPhedran, M. Nevière and P. Vincent, Japan J. Appl. Phys., 14 Suppl. 14-1 (1975) 167.

[5.11] C. H. Palmer, J. Opt. Soc. Amer., 42 (1952) 269.

[5.12] R. Jacobsson, Optica Acta, 10 (1963) 309.

[5.13] D. R. McKenzie and R. C. McPhedran, Nature, 265 (1977) 128.

[5.14] R. C. McPhedran and D. R. McKenzie, to appear in Proc. Roy. Soc.

[5.15] D. R. McKenzie, R. C. McPhedran and G. H. Derrick, submitted for publication to Proc. Roy. Soc.

[5.16] C. H. Palmer and F. W. Phelps Jr., J. Opt. Soc. Amer., 58 (1968) 1184.

CHAPTER 6

AN INTEGRAL MODAL THEORY FOR BURIED GRATINGS

6.1 INTRODUCTION

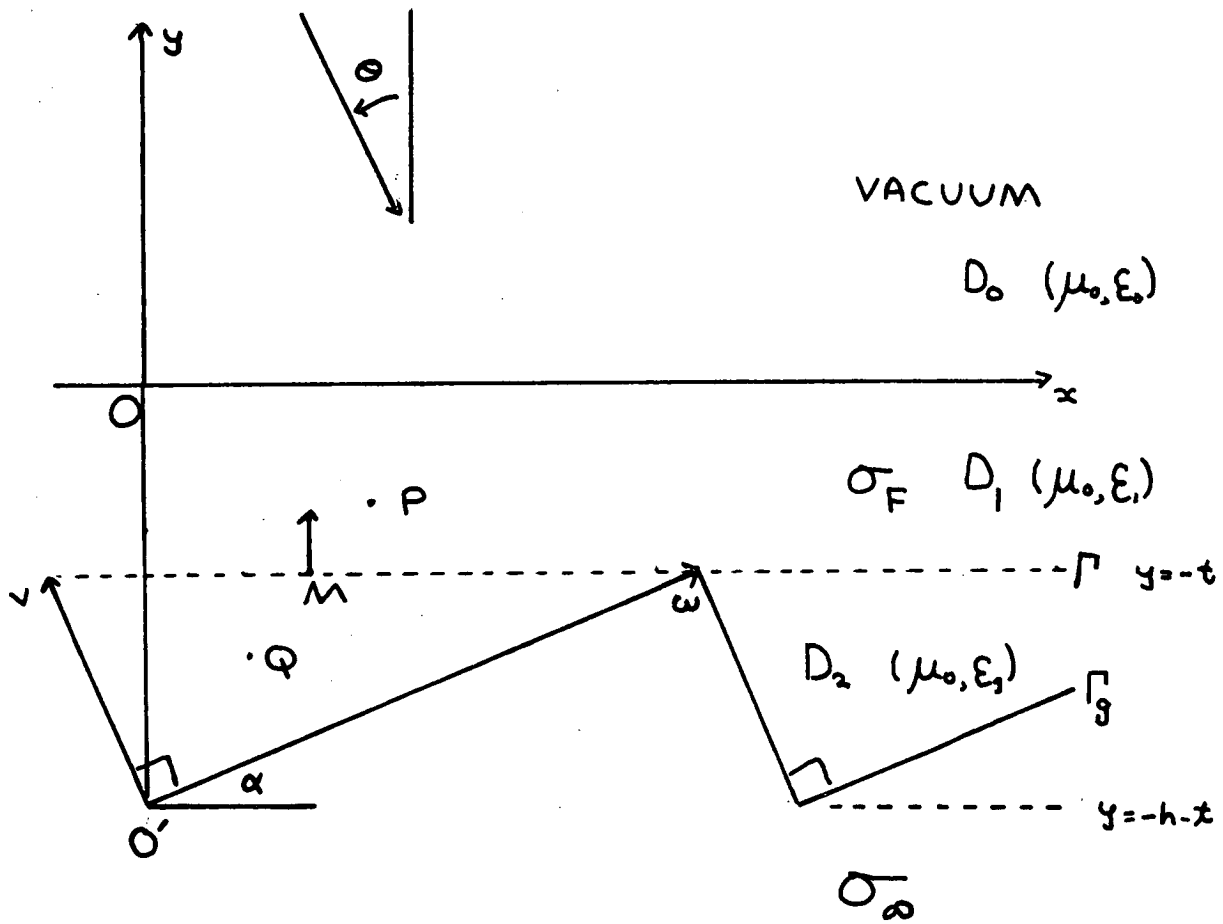
This chapter concerns itself with the diffraction of a plane wave by a triangular profile grating having a right angled apex, buried beneath a dielectric layer. The treatment evolved here can best be described as being an "integral-modal" formalism i.e. the technique relies upon the solution of an integral equation, yet encompasses features characteristic of waveguide modal formalisms.

This problem has already been discussed by both Van den Berg [6.1] and the author, using a simplified version of the theory presented in chapter 4. However, the treatment presented below differs fundamentally from the above in that the simple geometry and the infinite conductivity of the structure permit the successful use of a Green's function which implicitly incorporates both the field boundary conditions and the nature of the triangular groove. This converts the sampling surface from the actual profile to a planar surface.

6.2 THE THEORETICAL TREATMENT

6.2.1 Notation and plane wave expansions

Consider a P polarized plane wave of wavelength $\lambda = 2\pi/k_0$ (in free space) incident at some angle θ upon the structure of period $d = 2\pi/K$ shown in figure (6.2.1). The incident field is defined by



$$\frac{\varepsilon_1}{\varepsilon_0} = r^2$$

$$\bar{x} = x \qquad \bar{y} = y + h + t$$

$$w = \bar{x} \cos\alpha + \bar{y} \sin\alpha$$

$$v = -\bar{x} \sin\alpha + \bar{y} \cos\alpha$$

Figure 6.2.1 The geometry of the buried grating.

$$\underline{E}^i = \exp[i(\alpha_0 x - \chi_0 y)]\ \hat{\underline{z}} \qquad (6.2.1)$$

where $\alpha_0 = k_0 \sin\theta$

and $\chi_0 = k_0 \cos\theta$.

In the region D_0, the total electric field ϕ_0 obeys the Helmholtz equation

$$(\nabla^2 + k_0^2)\ \phi_0(P) = 0$$

and thus for all points P lying above the plane interface y = 0, the field may be expressed in terms of the Rayleigh expansion

$$\phi_0(x,y) = \exp[i(\alpha_0 x - \chi_0 y)] + \sum_{p=-\infty}^{\infty} \phi_p^r \exp[i(\alpha_p x + \chi_p y)] \qquad (6.2.2)$$

where $\alpha_p = pK + \alpha_0$

and $\chi_p = \sqrt{k_0^2 - \alpha_p^2}$ if $k_0 \geq |\alpha_n|$

$= i\sqrt{\alpha_p^2 - k_0^2}$ if $k_0 < |\alpha_n|$.

Now, in the region D_1, bounded by the planes y = 0 and y = -t, the total field ϕ_1 obeys

$$(\nabla^2 + k_1^2)\ \phi_1 = 0$$

where $k_1 = k_0 r$,

r being the refractive index of D_1. As D_1 is open-ended in the

x-direction, the field can again be expanded in terms of a Rayleigh series, this time involving both upward and downward going wave terms

$$\text{i.e. } \phi_1(x,y) = \sum_{p=-\infty}^{\infty} [\phi_{1,p}^{+} \exp(i\eta_p y) + \phi_{1,p}^{-} \exp(-i\eta_p y)] \exp(i\alpha_p x) \quad (6.2.3)$$

$$\text{where} \qquad \eta_p = \sqrt{k_1^2 - \alpha_p^2} \ .$$

Once again, the branch of the square root is chosen such that the evanescent terms are correctly attenuated.

By using the treatment evolved in chapter 4, it may be shown that the plane wave coefficients $\{\phi_{1,p}^{+}\}$ and $\{\phi_{1,p}^{-}\}$ are related by Fresnel's laws:

$$\left.\begin{aligned} \text{i.e.} \qquad \phi_{1,p}^{-} &= T_0^E \, \delta_{p,0} + R_p^E \, \phi_{1,p}^{+} \\ \text{and} \qquad \phi_p^r &= (T_0^E - 1) \, \delta_{p,0} + (R_p^E + 1) \, \phi_{1,p}^{+} \end{aligned}\right\} \quad (6.2.4)$$

$$\text{where} \qquad R_p^E = \frac{\eta_p - \chi_p}{\eta_p + \chi_p}$$

$$\text{and} \qquad T_p^E = \frac{2\chi_p}{\eta_p + \chi_p} \ .$$

6.2.2 Choice of a suitable unknown characterizing the diffraction problem

In accordance with the treatment of chapter 4, a function u is introduced where:

(i) $u(P) = \phi_1(P)$ if $P \in D_1$,

(ii) $(\nabla^2 + k_1^2)\, u = 0$ elsewhere,

and

(iii) u is everywhere continuous and obeys the radiation conditions as $|y| \to \infty$.

It must be noted that u represents a physical field quantity only within region D_1. For all other points in space, it must be taken as being a mathematical abstraction. Of course, this is an obvious statement for the case of points lying in D_0 since, in general, u and the physical field, E, do not obey the same Helmholtz equation. However, for points inside D_2, the abstract nature of the above definitions is less clear. This arises from the fact that there is no change in the optical properties upon traversing the plane interface Γ and hence no discontinuous behaviour exhibited by either the real field or its normal derivative at Γ. Hence, one may be led to suggest that both u and $\partial u/\partial n$ are continuous across Γ. However, this speculation is quickly quashed after performing an analysis similar to that described in chapter 4, where it was shown that u may be expressed as a function of the discontinuity in $\frac{\partial u}{\partial n}$ at Γ. If $\frac{\partial u}{\partial n}$ were continuous across Γ then u (and hence the physical field) would vanish everywhere within D_1, an entirely unreasonable suggestion. It thus follows that u must possess a derivative discontinuity on Γ.

This argument may be further pursued by considering the wave equations appropriate to both u and the physical field E beneath the plane y = -t. Part (ii) of the definition of u stipulates that

$$(\nabla^2 + k_1^2)\, u = 0 \qquad \forall \quad y < -t.$$

This together with the continuity of u means that $\partial u/\partial n$ is continuous across the actual grating profile Γ_g. However, this is not the case for the physical field where $\partial E/\partial n$ is discontinuous at Γ_g, a consequence of the existence of a surface current. This manifests itself as a distributive term (of magnitude $\partial E/\partial n\ \delta_{\Gamma_g}$) in the Helmholtz equation -

$$\text{i.e.} \quad (\nabla^2 + k_1^2)\, E = \frac{\partial E}{\partial n}\, \delta_{\Gamma_g}.$$

Thus, it is clear that u does not represent the physical field quantity beneath Γ. In short, the derivative discontinuity in the physical field at Γ_g may be transformed to a derivative discontinuity in the fictitious field, u, at Γ.

Now, by following the treatment outlined in chapter 4 it may be shown that

$$u(P) = \int_{\Gamma_+} \left[G_1 \frac{\partial u}{\partial n'} - \frac{\partial G_1}{\partial n'}\, u\right] ds'$$

$$+ \int_0^d \left[G_1 \frac{\partial u}{\partial n'} - \frac{\partial G_1}{\partial n'}\, u\right]\Big|_{y' = 0^-} dx' , \qquad (6.2.5)$$

where $G_1(P;M') = \frac{1}{2id} \sum_{p=-\infty}^{\infty} \frac{1}{\eta_p} \exp[i\alpha_p(x-x') + i\eta_p |\, y-y'\,|\,]$.

After splitting the field u into its upward and downward going wave trains

$$\text{viz.} \quad u(P) = u^{+}(P) + u^{-}(P) ,$$

the line integral (equation (6.2.5)) over the contour Γ_+ may be identified as being a source of upward going waves

$$\text{i.e.} \quad u^{+}(P) = \int_{\Gamma_+} [G_1 \frac{\partial u}{\partial n'} - \frac{\partial G_1}{\partial n'} u] \; ds'.$$

Because of the open ended nature of D_1, it is possible to expand u^+ everywhere within that region in terms of a Rayleigh series:

$$\left.\begin{aligned} \text{viz.} \quad u^{+}(P) &= \sum_{p=-\infty}^{\infty} u_p^{+} \exp[i(\alpha_p x + \eta_p y)] \\ \text{where} \quad u_p^{+} &= \int_{\Gamma_+} [G_{1,p} \frac{\partial u}{\partial n'} - \frac{\partial G_{1,p}}{\partial n'} u] \; ds' \\ \text{and} \quad G_{1,p} &= \frac{1}{2id\eta_p} \exp[i(-\alpha_p x' + \eta_p t)]. \end{aligned}\right\} \qquad (6.2.6)$$

Furthermore, the incoming field u^- can also be expressed in terms of a Rayleigh expansion whose coefficients $\{u_p^-\}$ are given by

$$u_p^{-} = T_0^E \delta_{p,0} + R_p^E u_p^{+}.$$

After some simple manipulation, one then arrives at the final result:

$$u(P) = u^{+}(P) + u^{-}(P)$$

$$= E^{i}_{eff}(P) + \int_{\Gamma_+} [G^{E}_{1} \frac{\partial u}{\partial n'} - \frac{\partial G^{E}_{1}}{\partial n'} u] \, ds' \qquad (6.2.7)$$

where $E^{i}_{eff}(P) = T^{E}_{0} \exp[i(\alpha_0 x + \eta_0 t)]$

and $G^{E}_{1}(P;M') = G_1(P;M') + \frac{1}{2id} \sum_{p=-\infty}^{\infty} \frac{R^{E}_{p}}{\eta_p} \exp[i \alpha_p (x-x') + i\eta_p (t-y)].$

It is to be noted that,

$$(\nabla^2 + k_1^2) G^{E}_{1} = (\nabla^2 + k_1^2) G_1$$

$$= \delta(y-y') \sum_{p=-\infty}^{\infty} \delta(x-x'-pd) \exp(i\alpha_0 pd)$$

(i.e. G^{E}_{1} exhibits the same singular behaviour as G_1).

Thus, by a second application of Green's theorem, this time to the region $x \in [0,d]$, $y \in (-\infty, -t^{-}]$, it follows that

$$0 = \int_{\Gamma_-} [G^{E}_{1} \frac{\partial u}{\partial n'} - \frac{\partial G^{E}_{1}}{\partial n'} u] \, ds', \qquad (6.2.8)$$

since there are no singularities lying within the region of integration. Subtraction of equations (6.2.7) and (6.2.8) then yields

$$u(P) = E^{i}_{eff}(P) + \int_{\Gamma} G^{E}_{1}(P;M') \, \xi(M') \, ds'$$

where $\xi = \frac{\partial u}{\partial n}\Big|_{\Gamma_+} - \frac{\partial u}{\partial n}\Big|_{\Gamma_-}$.

Hence, at some point M on Γ

$$E(M) = E^{i}_{eff}(M) + \int_{\Gamma} G_1^E (M;M')\, \xi(M')\, ds' \qquad (6.2.9)$$

and

$$\frac{\partial E(M)}{\partial n}\bigg|_{\Gamma_+} = \frac{\partial E^{i}_{eff}(M)}{\partial n} + \tfrac{1}{2}\xi(M) + \int_{\Gamma} \frac{\partial G_1^E}{\partial n} (M;M')\, \xi(M')\, ds'. \qquad (6.2.10)$$

These equations reveal that both E and $\partial E/\partial n$ can be expressed in terms of a single function ξ which, to some extent, can be regarded as being the analogue of a radiating surface current located in the plane $y = -t$. It is now necessary to establish the constraints imposed on the physical systems by the boundary conditions appropriate to the actual grating profile, Γ_g.

6.2.3 Derivation of an integral equation

As has been stated in the introduction to the chapter, this treatment is applicable only to a buried grating of infinite conductivity, having a triangular profile whose apex angle is 90°. However, such restrictions give rise to a relatively elegant solution of the problem and one is led to suggest that the formulation encompasses a number of the desirable features exhibited by modal formalisms. Previous approaches to the solution of this class of problems using integral equation techniques have relied upon sampling the unknown at a set of discrete points on the grating surface proper. Admittedly these methods (such as the solution of the diffraction problem of a multilayer transmission grating) are far more general in their application, in that they are restricted by neither a profile type nor the grating

conductivity. However, the problem solved herein does possess a certain intrinsic simplicity in that the field everywhere within the groove (i.e. region D_2) is expressible in terms of field quantities on the planar interface Γ.

The proposition that this formulation is "modal" in its nature is perhaps slightly misleading. No attempt is made to derive an explicit modal expansion characterizing the fields within the grooves. However, the restrictions placed on the structure do permit the derivation of a Green's function closely emulating the field boundary conditions and embracing the rectangular geometry of the groove.

Thus with these prefactory remarks, let us now seek an integral equation constraining the two field quantities E and $\partial E/\partial n$ on Γ. However, before proceeding with the analysis, it is necessary to define the following co-ordinate transformation:

$$\bar{x} = x, \qquad \bar{y} = y + h + t,$$
$$w = \bar{x}\cos\alpha + \bar{y}\sin\alpha,$$
$$v = -\bar{x}\sin\alpha + \bar{y}\cos\alpha.$$

With the aid of these rotated co-ordinates, it is now possible to introduce a Green's function G_2^E defined by the inhomogeneous Helmholtz equation:

$$(\nabla^2 + k_1^2)\, G_2^E(Q;M') = \sum_{m=-\infty}^{\infty} [\delta(w-w'+2mc) - \delta(w+w'+2mc)]\,[\delta(v-v') - \delta(v+v')] \tag{6.2.11}$$

whose solution is given by

$$G_2^E(Q;M') = -\frac{i}{c}\sum_{m=1}^{\infty}\frac{1}{\mu_m}\sin(\frac{m\pi w}{c})\sin(\frac{m\pi w'}{c})$$

$$\{\exp[i\mu_m|v-v'|] - \exp[i\mu_m|v+v'|]\}$$

where $c = \frac{d}{\cos\alpha}$ and $\mu_m = \{k_1^2 - \frac{m^2\pi^2}{c^2}\}^{\frac{1}{2}}$. (6.2.12)

The original derivation of this Green's function is attributed to Wirgin (see for example reference [6.2]). As can be seen from equation (6.2.12), the astute location of the singularities has yielded a function vanishing everywhere on the metal walls of the groove, thereby emulating the field boundary conditions.

Now, from an application of Green's Theorem to the groove D_2 bounded by the contour C_2, it follows that:

$$\iint_{D_2}[G_2^E(Q;M')\,\nabla^2E(M') - \nabla^2G_2^E(Q;M')\,E(M')]\,dA'$$

$$= \int_{C_2}[G_2^E(Q;M')\,\frac{\partial E}{\partial n'}(M') - \frac{\partial G_2^E}{\partial n'}(Q;M')\,E(M')]\,ds'.$$

This reduces to the following form after noting that only one singularity is located within D_2 and also that both G_2^E and E vanish on the metal walls:

$$E(Q) = \int_{\Gamma}[\frac{\partial G_2^E}{\partial n'}E - G_2^E\frac{\partial E}{\partial n'}]\,ds'. \quad (6.2.13)$$

Thus, by the appropriate choice of the Green's function, it has been possible to relocate the sampling surface at the plane Γ.

Should the point Q now tend to a point M lying on Γ, then effectively only "half" of the singularity lies within the region of integration, and so

$$\tfrac{1}{2}E(M) = \int_{\Gamma_-} \left[\frac{\partial G_2^E}{\partial n'}(M;M')\, E(M') - G_2^E(M;M') \frac{\partial E(M')}{\partial n}\right] ds'. \qquad (6.2.14)$$

This is the desired result, an integral equation constraining the behaviour of E and $\partial E/\partial n$ at Γ and hence characterizing the diffraction problem. By now applying the continuity equations at Γ and substituting the representations of E and $\partial E/\partial n$ (at Γ_+) in terms of the fundamental unknown ξ one arrives at the final result:

$$\xi_0(M) = \int_\Gamma \tau^E(M;M')\, \xi(M')\, ds', \qquad (6.2.15)$$

where $$\xi_0(M) = \tfrac{1}{2} E^i_{eff}(M) + \int_\Gamma \left[G_2^E(M;M') \frac{\partial E^i_{eff}(M')}{\partial n'} - \frac{\partial G_2^E}{\partial n'}(M;M')\, E^i_{eff}(M')\right] ds' \qquad (6.2.16a)$$

and

$$\tau^E(M;M') = -\tfrac{1}{2} G_1^E(M;M') - \tfrac{1}{2} G_2^E(M;M') + \int_\Gamma \left[\frac{\partial G_2^E}{\partial n''}(M;M'')\, G_1^E(M'';M') - G_2^E(M;M'') \frac{\partial G_1^E}{\partial n''}(M'';M')\right] ds'' \qquad (6.2.16b)$$

This is a Fredholm integral equation of the first kind, whose numerical solution presents no difficulties not already circumvented in other other problems. Field reconstruction takes place in a manner analogous to that outlined in chapter 4.

6.2.4 Modification of the theory for S polarization

The derivation of the theory for S polarized light is entirely similar to the discussion in the previous two sections, with the only differences arising from the application of boundary conditions appropriate to this polarization. Again, a function u is defined, this time in terms of the magnetic field quantities. Here, u is given by the following criteria:

(i) $u(P) = H(P)$ if $P \in D_1$,

(ii) $(\nabla^2 + k_1^2)\, u = 0$ elsewhere,

(iii) u is everywhere continuous and obeys the radiation conditions as $|y| \to \infty$.

A treatment similar to that outlined in section (6.2.2) reveals that for some point M on Γ

$$H(M) = H^i_{eff}(M) + \int_\Gamma G_1^H(M;M')\, \xi(M')\, ds' \qquad (6.2.17)$$

and

$$\frac{\partial H(M)}{\partial n}\bigg|_{\Gamma_+} = \frac{\partial H^i_{eff}(M)}{\partial n} + \tfrac{1}{2}\, \xi(M) + \int \hat{\underline{n}}.\nabla G_1^H(M;M')\, \xi(M)\, ds' \qquad (6.2.18)$$

where

$$\xi(M) = \frac{\partial u}{\partial n}\bigg|_{\Gamma_+} - \frac{\partial u}{\partial n}\bigg|_{\Gamma_-},$$

and

$$G_1^H(P;M') = G_1(P;M') + \frac{1}{2id} \sum_p \frac{R_p^H}{\eta_p} \exp[i\alpha_p(x-x') + i\eta_p(t-y)] \ .$$

For this polarization, the reflection and transmission coefficients at the interface y = 0 are given by

$$R_p^H = \frac{\eta_p - r^2\chi_p}{\eta_p + r^2\chi_p}$$

and $$T_p^H = \frac{2r^2\chi_p}{\eta_p + r^2\chi_p}$$

and so the effective incident field upon Γ becomes

$$H_{eff}^i(M) = T_0^H \exp[i(\alpha_0 x + \eta_0 t)] .$$

In deriving an integral equation linking the two field quantities H and $\partial H/\partial n$, one again selects a Green's function representative of the geometry and implicity including the nature of the boundary condition on the metal surfaces. For S polarization the normal derivative of H vanishes on the metal and thus one locates the singularities in accordance with the desire to also render $\partial G_2^H/\partial n$ void on the groove walls. In this case

$$(\nabla^2 + k_1^2)\, G_2^H(Q;M') = \sum_{m=-\infty}^{\infty} [\delta(w-w'+2mc) + \delta(w+w'+2mc)][\delta(v-v') + \delta(v+v')],$$

whose solution is given by

$$G_2^H(Q;M') = -\frac{i}{2c} \sum_{m=0}^{\infty} \frac{\varepsilon_m}{\mu_m} \cos\left(\frac{m\pi w}{c}\right) \cos\left(\frac{m\pi w'}{c}\right)$$

$$\{\exp[i\mu_m|v-v'|] + \exp[i\mu_m|v+v'|]\}$$

where $\varepsilon_m = 1$ if $m = 0$

$= 2$ otherwise.

The subsequent analysis pertaining to the derivation of the integral equation is identical to that presented in section (6.2.3) and so, for reasons of brevity, is omitted.

6.2.5 Verification of the formalism

The theories for both fundamental polarizations have been implemented on a Control Data Cyber 72 computer (University of Sydney) and have yielded results in accordance with

(i) conservation of energy,

(ii) the Reciprocity Theorem,

(iii) the Littrow phase constraint (chapters 4 and 9), and

(iv) the Maréchal-Stroke relationship.

6.3 CONCLUSIONS

In this chapter, a new formulation of the diffraction problem involving buried gratings has been outlined. Although restricted to a triangular profile grating having a right angled apex ruled in metal of infinite conductivity, it is of some theoretical interest in that the above assumptions make possible an integral formulation based heavily on modal methods, such as those used in studying lamellar gratings.

It would be both dishonest and incorrect to suggest that this treatment makes any significant advance to this field of research in view of the past works. However, one feels that this theory warrants

an exposition on the basis of its elegance. Moreover, much of the experience gained in this analysis proved to be of considerable value in the formulation of the bi-metallic grating theories discussed in the following chapter.

To the best of the author's knowledge, the method described above represents one of the first integral formalisms based upon the actual geometry of the structure. These "integral-modal" techniques are limited by a poor understanding of the nature of fields within the grooves and as such one feels that their application is currently restricted to groove profiles possessing some semblance of rectangular symmetry.

REFERENCES

[6.1] P. M. Van den Berg, Ph.D. Thesis (1971), Delft University of Technology, Delft, the Netherlands.

[6.2] A Wirgin, C. R. Acad. Sci. Paris, 262B (1966) 870.

CHAPTER 7

A STUDY OF BI-METALLIC GRATINGS

Before entering into the body of this chapter, I wish to thank Dr. Ross McPhedran for suggesting this fascinating topic to me. His sustained interest in all facets of the study, particularly his generous and highly valued discussions involving the results of the numerical experiments are very much appreciated. A paper, concerning various aspects of the material presented in this chapter, has been submitted for publication to the journal "Optics Communications".

7.1 INTRODUCTION

The studies presented in this chapter were largely motivated by the recent experimental work of Sugahara, Kita and Shimotakahara [7.1], who reported the successful suppression of spectrographic grating anomalies using a replication process involving the deposition of aluminium on the "on-blaze" facet of a silver-aluminium alloy mother grating. Such a structure immediately introduces the concept of the bi-metallic grating i.e. a grating composed of two metals arranged such that both the profile and conductivity of the structure are periodically modulated.

At the start of this study, there was little if any theoretical interpretation of the diffraction properties of such structures. Thus, armed only with a basic understanding of integral formalisms, it was decided that the only reasonable approach to the problem was to adopt

a systematic and evolutionary strategy. Hence it is appropriate for this chapter to reconstruct the basic steps of the study.

The theoretical discussion commences with a treatment of two structures involving very specific assumptions as to their geometry and conductivity and eventually evolves a completely general formulation involving no restrictions whatsoever. At the time of the submission of this thesis, it had not been possible to numerically implement this theory. However, the encouraging results produced by its more "primitive" predecessors encourage the belief that such an exercise is indeed worthwhile. Throughout the theoretical sections of the chapter, considerable emphasis is placed on both the difficulties encountered and the errors which have been detected and corrected, in an effort to re-create some of the atmosphere of the study.

The numerical results which have been obtained using a "roughened lamellar reflection grating theory" indeed confirm the experimental evidence of the Japanese authors [7.1] concerning anomaly suppression. Useful reductions in anomaly strength can be achieved by the creation of "lossy plasmons" having a bandwidth sufficient to mask any sharp spectral features. It has been discovered that both upper and lower limits need to be placed on the conductivity of the lossy medium which resides on the off-blaze facet of the grating. The upper bound arises from arguments concerned with inadequate plasmon bandwidths causing poor anomaly suppression. On the other hand, the lower bound needs to be enforced because of two factors:

(i) the increasing absorption curtails the blazewidth, and/or

(ii) the increasing transparency of the overcoated film permits

resonant behaviour inside the off-blaze facet which manifests itself as additional anomalous spectral features.

The first two sections involve the derivation of the theory (from which these results were calculated) and a thorough scrutiny of the numerical experiments. It is with these introductory remarks that the discussion is now turned to the analysis of the "roughened lamellar reflection grating".

7.2 THE THEORETICAL TREATMENT FOR A SEMI-FINITELY CONDUCTING BI-METALLIC REFLECTION GRATING

7.2.1 Notation and the geometry of the grating

The structure which is considered within this section is illustrated in figure (7.2.1). As can be seen from the diagram, the grating is composed of an array of cylinders, alternating in conductivity, deposited on a perfectly conducting substrate (medium D_3). The formalism described below relies upon the assumption that D_1, a medium of finite conductivity, and D_2, a medium of infinite conductivity, are separated by vertical linear segments. This choice has been made in order to facilitate a "semi-modal" treatment which implicitly incorporates the boundary conditions of the field at all points on the metal walls of D_2. Although it has been necessary to define the metal walls to be linear segments, the shape of the interface Γ, separating D_1 and D_2 from free-space (region D_0) can be arbitrary.

Now, consider a plane P polarized wave of free-space wavelength $\lambda = 2\pi/k_0$ incident at some angle upon the structure shown in figure (7.2.1) having period $d = 2\pi/K$. The incident field is given by

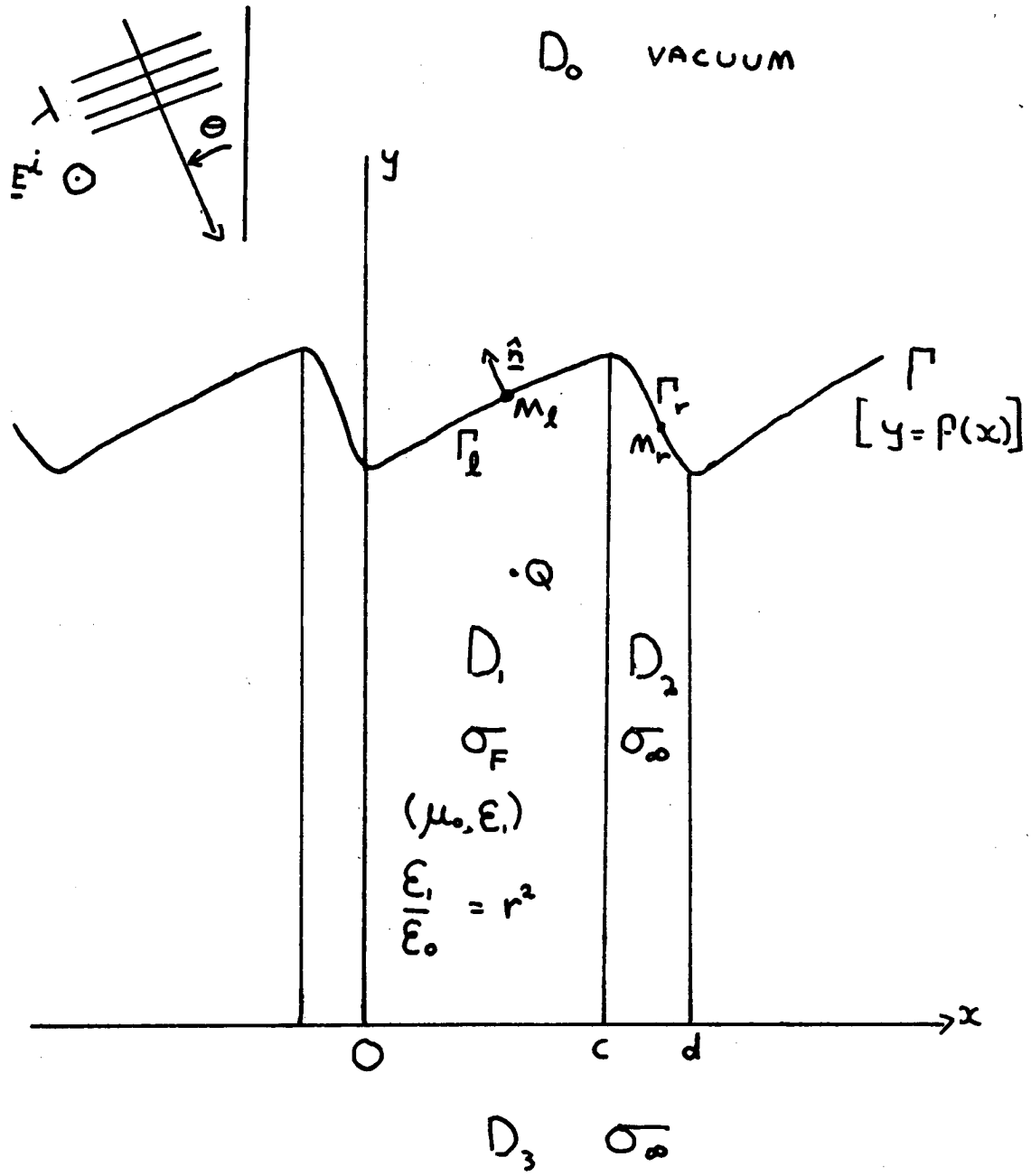


Figure 7.2.1 The diffracting arrangement of the roughened lamellar reflection grating. Medium D_0 represents free space and D_1 is a finitely conducting (σ_F) insertion (of refractive index r) in an otherwise perfectly conducting (σ_∞) structure.

$$\underline{E}^i = \exp[i(\alpha_0 x - \chi_0 y)]\, \hat{\underline{k}}$$

where the direction sine and cosine are given by

$$\alpha_0/k_0 = \sin\theta$$

and $$\chi_0/k_0 = \cos\theta.$$

7.2.2 Characterization of the electric field in free space

In the region D_0, the diffracted field is given by

$$E^d = E - E^i \tag{7.2.2}$$

and obeys

(i) the Helmholtz equation $(\nabla^2 + k_0^2)\, E^d(x,y) = 0$

and

(ii) the radiation conditions as y tends to infinity.

In order to uniquely characterize the diffracted field distribution it is necessary to consider a function u defined by:

$$\left.\begin{aligned} &\text{(i)} \quad u(P) = E^d(P) \qquad \forall\ P \in D_0, \\ &\text{(ii)} \quad u \text{ is everywhere continuous,} \\ &\text{and} \\ &\text{(iii)} \quad (\nabla^2 + k_0^2)\, u(P) = 0 \quad \forall\ P \notin D_0. \end{aligned}\right\} \tag{7.2.3}$$

By now following the method originally devised by Maystre [7.2], it may be shown that

$$u(P) = \int_{\Gamma} G_0(P;M')\, \xi(M')\, ds' \qquad (7.2.4a)$$

$$\text{where } \xi(M') = \left.\frac{\partial u}{\partial n}\right|_{\Gamma_+} - \left.\frac{\partial u}{\partial n}\right|_{\Gamma_-} \qquad (7.2.4b)$$

$$\text{and } G_0(P;M') = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\chi_n} \exp[i\alpha_n(x-x') + i\,\chi_n|y - y'|] \qquad (7.2.4c)$$

$$\text{with } \chi_n = \sqrt{k_0^2 - \alpha_n^2} \text{ and } y' = f(x').$$

By allowing the point P to tend to some point M on the profile, it may be shown that

$$E(M) = E^i(M) + \int_{\Gamma} G_0(M;M')\, \xi(M')\, ds' \qquad (7.2.5a)$$

$$\text{and } \left.\frac{\partial E(M)}{\partial n}\right|_{\Gamma_+} = \frac{\partial E^i(M)}{\partial n} + \tfrac{1}{2}\xi(M) + \int_{\Gamma} \frac{\partial G_0(M;M')}{\partial n}\, \xi(M')\, ds' \,. \qquad (7.2.5b)$$

Equations (7.2.5a) and (7.2.5b) demonstrate that both E and $\frac{\partial E}{\partial n}$ at the surface Γ can be expanded in terms of a single unknown, thereby facilitating a relatively easy numerical treatment. It is impossible to emphasize the crucial nature of Maystre's approach sufficiently strongly. Surely, it must be the opinion of all those who have been actively involved in theoretical grating research during the past five years that his method has completely revolutionized the solution of classical diffraction problems using integral formalisms. Subsequent work by the author has been largely stimulated by the efforts of Maystre, with variations of his treatment being used to solve such problems as the diffraction of a plane wave by a multi-layer transmission grating

(chapter 4) and a buried grating (chapter 6). It is unfortunate that the above description cannot do justice to the subtleties of his technique, simply because of the need for brevity. However, a more adequate review of the concepts involved is presented elsewhere (chapters 2 and 3).

7.2.3 Derivation of an integral equation

Having defined the existence of an unknown ξ which can characterize the field and its normal derivative at the surface Γ, it is now necessary to impose the physical boundary conditions in order to derive a pair of integral equations from which the quantity ξ can be deduced. However before launching ourselves into the detailed mathematical analysis of the problem, it is appropriate to define some additional notation. Let Γ be divided into two segments, Γ_ℓ and Γ_r, such that $x \in [0,c]$ and $x \in [c,d]$ respectively. On Γ_ℓ the relevant boundary condition is the continuity of the tangential component of E and its normal derivative

$$\text{i.e.} \quad E(M_\ell)\Big|_{\Gamma_{\ell+}} = E(M_\ell)\Big|_{\Gamma_{\ell-}} \tag{7.2.6a}$$

$$\frac{\partial E}{\partial n}\Big|_{\Gamma_{\ell+}} = \frac{\partial E}{\partial n}\Big|_{\Gamma_{\ell-}} \tag{7.2.6b}$$

whereas on Γ_r it is given by

$$E(M_r) = 0. \tag{7.2.6c}$$

Thus for those points M_r lying on Γ_r, it follows from equation (7.2.5a) that

$$- E^i(M_r) = \int_\Gamma G_0(M_r;M')\, \xi(M')\, ds' \tag{7.2.7}$$

- a Fredholm integral equation of the first kind. This constraint, equation (7.2.7), is purely a consequence of what may be termed the "direct" boundary condition given by expression (7.2.6c). However, the continuity of E across Γ_ℓ, as expressed by equation (7.2.6a), is somewhat less "direct". This necessitates the derivation of a separate representation for the fields within D_1, which can be matched to field prescriptions derived earlier (equations 7.2.5a and 7.2.5b) by use of the boundary condition for the surface Γ_ℓ.

The chosen "semi-rectangular" geometry of the groove D_1 enables one to implicitly enforce the boundary condition E = 0 everywhere on the metal walls of D_2 by defining a Green's function G_1 obeying

$$(\nabla^2 + k_1^2)\, G_1(Q;M') = \sum_{m=-\infty}^{\infty} [\delta(x-x'+2mc) - \delta(x+x'+2mc)] \times [\delta(y-y') - \delta(y+y')] \qquad (7.2.8)$$

where $Q \equiv (x,y)$ and $M' \equiv (x',y')$.

Here, k_1 is the wave number appropriate to the region D_1 and is given by

$$k_1 = k_0 r$$

where r is the refractive index of D_1. The definition of G_1, attributed to Wirgin [7.3], can only be described as being remarkably astute. The solution of the inhomogeneous Helmholtz equation (expression (7.2.8)) is given by

$$G_1 = -\frac{i}{c}\sum_{m=1}^{\infty}\frac{1}{\mu_m}\sin\left(\frac{m\pi x}{c}\right)\sin\left(\frac{m\pi x'}{c}\right)\exp[i\mu_m|y-y'|]$$

$$+\frac{i}{c}\sum_{m=1}^{\infty}\frac{1}{\mu_m}\sin\left(\frac{m\pi x}{c}\right)\sin\left(\frac{m\pi x'}{c}\right)\exp[i\mu_m|y+y'|] \qquad (7.2.9)$$

where

$$\mu_m = \sqrt{k_1^2 - \left(\frac{m\pi}{c}\right)^2},$$

and implicitly incorporates boundary conditions on the metal walls of D_1 emulating those of the field quantity E. It is indeed unfortunate that such analytically elegant formulations do not appear to exist for either the full finite conductivity treatment or for variations in the geometry of the structure.

Having now defined G_1, it is possible to apply Green's theorem to the region D_1 as follows:

$$\iint_{D_1}[\nabla^2 G_1(Q;M_\ell') E(M_\ell') - G_1(Q;M_\ell')\nabla^2 E(M_\ell')]\,dA'$$

$$= \int_{C_\ell}\left[\frac{\partial G_1(Q;M_\ell')}{\partial n'} E(M_\ell') - G_1(Q;M_\ell')\frac{\partial E(M_\ell')}{\partial n'}\right] ds' \qquad (7.2.10)$$

Since

$$(\nabla^2 + k_1^2) E(M_\ell') = 0 \qquad \forall\ M_\ell' \in D_1$$

and both $E(M_\ell') = 0$

and $G(Q;M_\ell') = 0$

for all points M_ℓ' located on the metal walls of D_1, equation (7.2.10)

reduces to

$$E(Q) = \int_{\Gamma_{\ell}^-} [\frac{\partial G_1(Q;M'_{\ell})}{\partial n'} E(M'_{\ell}) - G_1(Q;M'_{\ell}) \frac{\partial E(M'_{\ell})}{\partial n'}] ds' . \quad (7.2.11)$$

Should Q lie on Γ_{ℓ} then effectively "half a singularity" lies within D_1 and "half" lies without and so

$$\tfrac{1}{2} E(M_{\ell}) = \int_{\Gamma_{\ell}^-} [\frac{\partial G_1}{\partial n'} E - G_1 \frac{\partial E}{\partial n'}] ds' \quad (7.2.12)$$

for some point M_{ℓ} on Γ_{ℓ}^- .

By invoking the boundary conditions expressed in equations (7.2.6a) and (7.2.6b) and subsequently substituting the external representations of E and $\frac{\partial E}{\partial n}$ (given by equations (7.2.5a) and (7.2.5b)) into equation (7.2.12) one arrives at the final result:

$$\xi_0^E(M_{\ell}) = \int_{\Gamma} \tau^E(M_{\ell};M') \xi(M') ds' \quad (7.2.13a)$$

where

$$\xi_0^E(M_{\ell}) = \tfrac{1}{2} E^i(M_{\ell}) - \int_{\Gamma} [\frac{\partial G_1(M_{\ell};M'_{\ell})}{\partial n'} E^i(M'_{\ell}) - G_1(M_{\ell};M'_{\ell}) \frac{\partial E^i(M'_{\ell})}{\partial n'}] ds' \quad (7.2.13b)$$

and

$$\tau^E(M_{\ell};M') = -\tfrac{1}{2} G_0(M_{\ell};M') - \tfrac{1}{2} A(M') G_1(M_{\ell};M') + \int_{\Gamma_{\ell}} [\frac{\partial G_1(M_{\ell};M''_{\ell})}{\partial n''} G_0(M''_{\ell};M') - G_1(M_{\ell};M''_{\ell}) \frac{\partial G_0(M''_{\ell};M')}{\partial n''}] ds'' . \quad (7.2.13c)$$

Here A(M') is a Heaviside function which samples the function only on Γ_ℓ,

$$\left.\begin{aligned} \text{i.e. } A(M') &= 1 \quad \forall\ 0 \le x' \le c \\ &= 0 \quad \forall\ c < x' < d. \end{aligned}\right\} \tag{7.2.14}$$

Expressions (7.2.7) and (7.2.13a) are two Fredholm integral equations of the first kind which are solved numerically using the points-matching technique. Reconstruction of the diffracted field, given by the Rayleigh representation

$$E^d(x,y) = \sum_{n=-\infty}^{\infty} R_n \exp[i(\alpha_n x + \chi_n y)]$$

appropriate for all points such that $y \ge \max(f(x))$ ($f(x)$ being the profile Γ), takes place using the expression

$$R_n = \frac{1}{2id\chi_n} \int_\Gamma \xi(M') \exp[-i\alpha_n \chi' - i\chi_n f(x')]\, ds'.$$

7.2.4 Necessary modifications to the theory to cover the case of S polarized radiation

The treatment of the diffraction problem for the case of S polarized light follows along lines similar to those elaborated in the previous sections with the only differences arising as a result of the application of boundary conditions appropriate to this polarization. Thus, it is the purpose of this section to briefly outline the formalism and to expand upon the salient differences between the two treatments.

Once again, the method of Maystre is followed in order to derive free space representations for the magnetic field quantity H. This time, the variable u is defined by

(i) $u(P) = H^d(P)$ if $P \in D_0$

(ii) $(\nabla^2 + k_0^2)\, u(P) = 0$ elsewhere

and

(iii) u is everywhere continuous.

From this definition, one can show that

$$H(P) = H^i(P) + \int_\Gamma G_0(P;M')\, \xi(M')\, ds'$$

and so for a point M lying on Γ

$$H(M) = H^i(M) + \int_\Gamma G_0(M;M')\, \xi(M')\, ds' \qquad (7.2.15a)$$

and $$\left.\frac{\partial H}{\partial n}\right|_{\Gamma_+} = \frac{\partial H^i}{\partial n} + \tfrac{1}{2}\,\xi(M) + \int_\Gamma \frac{\partial G_0(M;M')}{\partial n}\, \xi(M')\, ds; \qquad (7.2.15b)$$

where $$\xi(M') = \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_{ext}} - \left.\frac{\partial u}{\partial n'}\right|_{\Gamma_-} .$$

Now, in deriving integral equations to be solved numerically using the points matching technique it is necessary to apply the following boundary conditions:

$$\left.\begin{aligned} &\text{(i)} \quad H(M_\ell)\big|_{\Gamma_\ell -} = H(M_\ell)\big|_{\Gamma_\ell +} \\ &\text{and} \quad \frac{\partial H(M_\ell)}{\partial n}\bigg|_{\Gamma_\ell -} = r^2\, \frac{\partial H(M_\ell)}{\partial n}\bigg|_{\Gamma_\ell +} \end{aligned}\right\} \qquad (7.2.16)$$

for a point M_ℓ on the interface Γ_ℓ

and

$$\text{(ii)}\quad \frac{\partial H(M_r)}{\partial n} = 0 \quad \text{(a consequence of } \underline{n} \times \underline{E} = 0) \tag{7.2.17}$$

on the vacuum-metal interface Γ_r.

Thus for those points M_r on Γ_r it follows from equation (7.2.17) and (7.2.15b) that

$$-\frac{\partial H^i(M_r)}{\partial n} = \tfrac{1}{2}\,\xi(M_r) + \int_\Gamma \frac{\partial G_0(M_r;M')}{\partial n}\,\xi(M')\,ds' \tag{7.2.18}$$

- a Fredholm integral equation of the second kind.

For those points M_ℓ on Γ_ℓ, the indirect nature of the field continuity expressed in equations (7.2.16) necessitates the formulation of a distinct representation for the magnetic field in the groove region D_1 which can be linked to the free space expansions using expressions (7.2.16). Once again it is advantageous to select a Green's function G_1 emulating the boundary condition $\frac{\partial H}{\partial n} = 0$ everywhere on the metal walls of D_1. Such a function is defined by the inhomogeneous Helmholtz equation

$$(\nabla^2 + k_1^2)\,G_1(Q;M') = \sum_{m=-\infty}^{\infty} [\delta(x-x'+2mc) + \delta(x+x'+2mc)] \times [\delta(y-y') + \delta(y+y')]$$

whose solution is

$$G_1(Q;M') = -\frac{i}{2c} \sum_{m=0}^{\infty} \frac{\varepsilon_m}{\mu_m} \cos\left(\frac{m\pi x}{c}\right) \cos\left(\frac{m\pi x'}{c}\right) \{ \exp[i\mu_m|y-y'|] + \exp[i\mu_m|y+y'|] \}.$$

Using this Green's function and applying a treatment similar to that given in section (7.2.3), one arrives at the integral expression

$$\tfrac{1}{2}H(M_\ell)\Big|_{\Gamma_\ell^-} = \int_{\Gamma_\ell^-} \left[\frac{\partial G_1(M_\ell;M'_\ell)}{\partial n'} H(M'_\ell) - G_1(M_\ell;M'_\ell) \frac{\partial H(M'_\ell)}{\partial n'}\right] ds' \qquad (7.2.19)$$

which yields the final result

$$\xi_0^H(M_\ell) = \int_\Gamma \tau^H(M_\ell;M') \, \xi(M') \, ds' \qquad (7.2.20a)$$

where

$$\xi_0^H(M_\ell) = \tfrac{1}{2} H^i(M_\ell) - \int_{\Gamma_\ell} \left[\frac{\partial G_1(M_\ell;M'_\ell)}{\partial n'} H^i(M'_\ell) - r^2 \, G_1(M_\ell;M'_\ell) \frac{\partial H^i(M'_\ell)}{\partial n'}\right] ds' \qquad (7.2.20b)$$

and

$$\tau^H(M_\ell;M') = -\tfrac{1}{2} G_0(M_\ell;M') - \tfrac{1}{2} r^2 \, A(M') \, G_1(M_\ell;M') + \int_\Gamma \left[\frac{\partial G_1(M_\ell;M''_\ell)}{\partial n''} G_0(M''_\ell;M') - r^2 \, G_1(M_\ell;M''_\ell) \frac{\partial G_0(M''_\ell;M')}{\partial n''}\right] ds'' , \qquad (7.2.20c)$$

a Fredholm integral equation of the first kind.

7.2.5 Verification of the formalism

Computer programs (for both fundamental polarizations) implementing the theories described above have been written in FORTRAN IV for operation on the Control Data Cyber 72 of the University of Sydney. Results generated from a number of numerical experiments have been in excellent agreement with the criteria of energy conservation and reciprocity. Accuracy to within half a percent has been achieved.

7.3 NUMERICAL EXPERIMENTS

7.3.1 Introductory Comments

In the final remarks of the previous section, mention was made of the successful numerical implementation of the theory. However, glib comments such as these do much to mask the immense effort and frustration involved in eliminating errors from the computer program. In the experience of the author, it is the normal practice for the formalism to be basically sound with the "debugging" stages concerned only with the removal of minor algebraic errors in the implementation. In this case, however, these roles were essentially reversed and for this reason it is appropriate to briefly describe some of the difficulties encountered.

Two computer programs, one for each of the fundamental polarizations of light, were written. The internal structure of these programs is very similar with the necessary sequence of calculations being performed outlined as follows:

(i) the computation of the kernel matrices;
(ii) the setting up of linear equations characteristic of the integral equation being solved;
(iii) the solution of these linear equations; and
(iv) the reconstruction of the field quantities using the pseudo-current density [from (iii)].

In the evolution of these programs, it was decided to first implement the P polarization formalism and then modify it for the case of S polarized light. However, in haste, certain changes regarding S polarization boundary conditions on the face Γ_r were omitted and

consequently the program was modelling an unphysical structure having a "magnetic wall" on one facet of the profile. This program yielded results which were in excellent agreement with the constraints imposed by conservation of energy and reciprocity and so it was presumed that the numerical implementation represented a correct interpretation of the theory.

It was not until the commencement of a detailed series of numerical experiments that one was led to suspect the validity of the computer program. The strong similarities in the performance of the structure exhibited in both P and S polarized light led one to either suspect the numerical implementation or alternatively to search for some hitherto unknown physical mechanism that could cause such a phenomenon. In an attempt to resolve this dilemma, a number of field reconstructions were performed both inside the groove and on the surface, Γ. With the aid of these calculations, the nature of the problem (i.e. the imposition of an incorrect boundary condition) was immediately exposed. Following the correction of this error, results which were readily understood in the light of past experience were forthcoming.

These comments have been included for two major reasons:

(i) the desirability of presenting a true record of the evolution of this theory, and

(ii) the need to dispel any absolute reliance on the results of the reciprocity theorem.

Of these two reasons, the latter is by far the most crucial. Although it had long been appreciated that the reciprocity theorem did not constitute a test of the physical validity of a grating formalism, one

was, nevertheless, inclined to regard its results with almost blind faith. Thus, it has been the aim of these few remarks to place on record a particular case for which the reciprocity theorem gave rise to a totally unjustified belief in the validity of the formalism. In view of these comments, it is felt that the consequences of the reciprocity theorem need to be tempered by a field reconstruction (with the results of both being thoroughly scrutinized) before the physical validity of any grating theory can be assured.

7.3.2 Results and discussion

As stated earlier, this study was largely motivated by the recent experimental work of Sugahara, Kita and Shimotakahara [7.1] who reported that grating anomalies could be significantly reduced by coating the respective facets of a blazed grating with metals of different conductivities. Thus, in an attempt to theoretically understand this phenomenon it was decided to implement the theory of the previous section as the first stage of an extensive examination of the diffraction properties of bi-metallic gratings.

It is pleasing to be able to report the success of this investigation, with partial suppression of anomalies having been achieved. However, this suppression is accompanied by a small but noticeable reduction in the blaze performance. The anomaly reduction is a consequence of the excitation of lossy (or inelastically scattered) plasmons in the finitely conducting medium. As such, the choice of the refractive index of the plug is critical to the successful elimination of anomalies.

The current study has been restricted to a structure of the type shown in figure (7.3.1). It is to be noted that the perfectly conducting

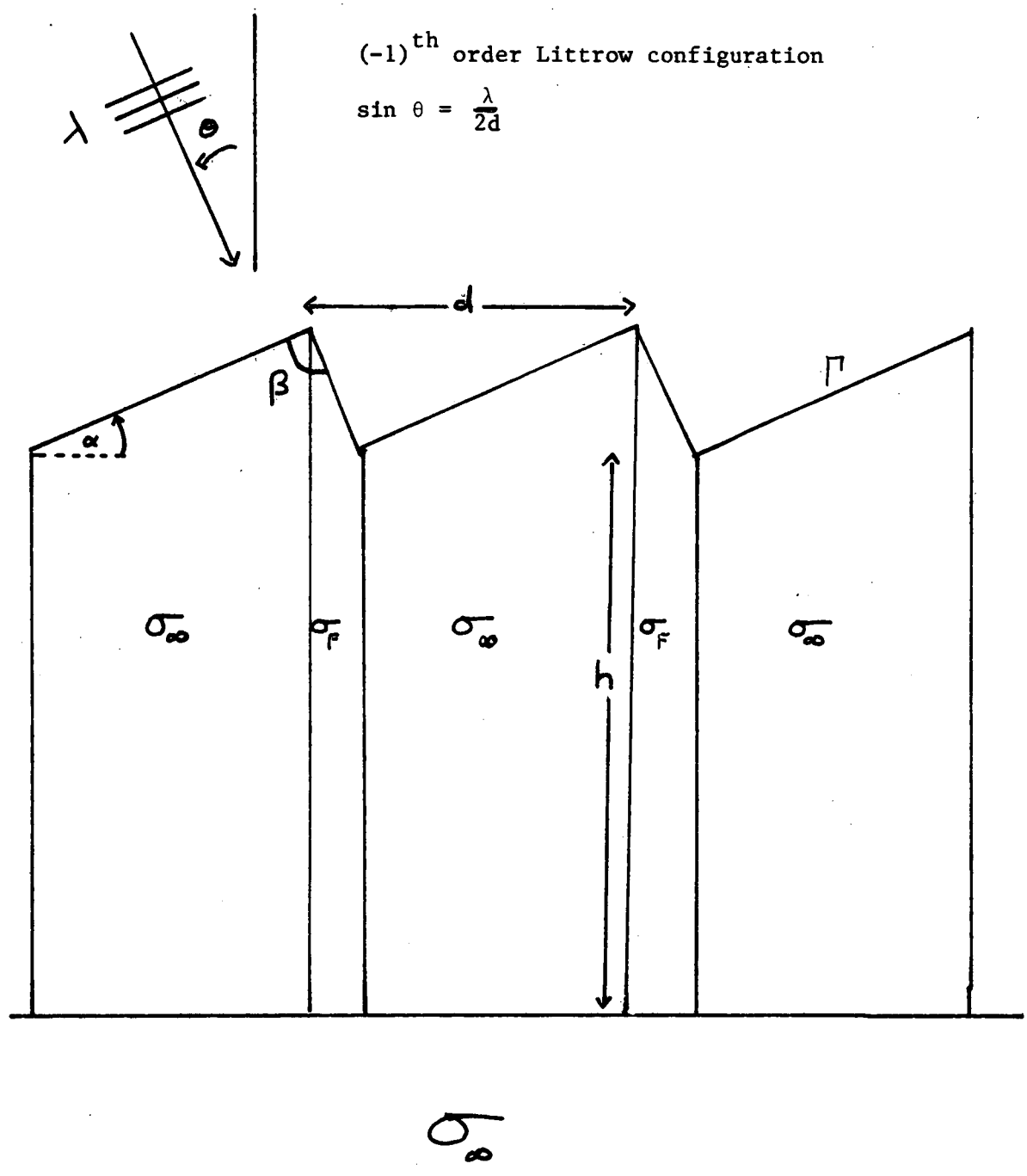


Figure 7.3.1

The geometry of the blazed bi-metallic grating characterized by the two angles α and β. The structure is composed of two species of metals:

(i) a perfectly conducting segment (denoted by σ_∞) and

(ii) a finitely conducting plug (denoted by σ_F) of refractive index r_1 residing on a perfectly conducting substrate displaced a distance h from the lower edge of the actual grating profile Γ.

material is located on the "on-blaze" facet of the grating in order to maintain the overall blaze at as high a level as possible.

The performance of two gratings, both exhibiting anomalies of an insidious nature, is now discussed in some detail.

7.3.2.1 *The 11.5° blaze angle grating*

The first of the gratings to be considered was characterized by a blaze angle of 11.5° and an apex angle of 90°. The grating structure was displaced from the ground plane by an amount h/d = 0.4 such that the perfectly conducting substrate was effectively "hidden" (from the incident field) by the opaque metal plug. The grating was operated in a $(-1)^{th}$ order Littrow mount and figures (7.3.2) show the efficiency of the blazed order and the total diffracted energy spectrum for several values of the plug refractive index.

The reason for evaluating the performance of this grating becomes immediately evident upon inspection of figure (7.3.2a). Clearly, the aim of this numerical study was to weaken the very strong S polarization resonance anomaly occurring on the long wavelength side of the (-2,+1) Wood anomaly. As can be seen from figures (7.3.2), this objective has been largely achieved. However, it is thought that further reductions in the strength of this anomaly could be expected had even more weakly conducting plugs been used. These results, particularly the total energy spectra, are in accord with the work of Strong [7.4], who noted that the strongest anomalies were exhibited by highly reflecting metals.

It is interesting to note that the Wood anomaly (at $\lambda/d = 2/3$), manifesting itself as a very sharp dip in the efficiency spectrum immediately before the resonance, is virtually unaffected by the presence

Figure 7.3.2

Diffracted energy spectra for a triangular profile grating having a blaze angle of 11.5^{o} and a right-angled apex. The grating is displaced from the plane of the substrate by an amount h/d=0.4 and is operated in a $(-1)^{th}$ order Littrow configuration. On each of the graphs are curves displaying

(i) $(-1)^{th}$ order efficiency in P polarized radiation (- - - - -),

(ii) $(-1)^{th}$ order efficiency in S polarized radiation (————),

(iii) total diffracted energy in P polarized radiation (- - - -),

and

(iv) total diffracted energy in S polarized radiation (—•—•—).

Graphs for plugs having wavelength independent refractive indices of

(a) infinite conductivity,

(b) 1.5 + i 5.0 and

(c) 1.5 + i 2.0

are presented.

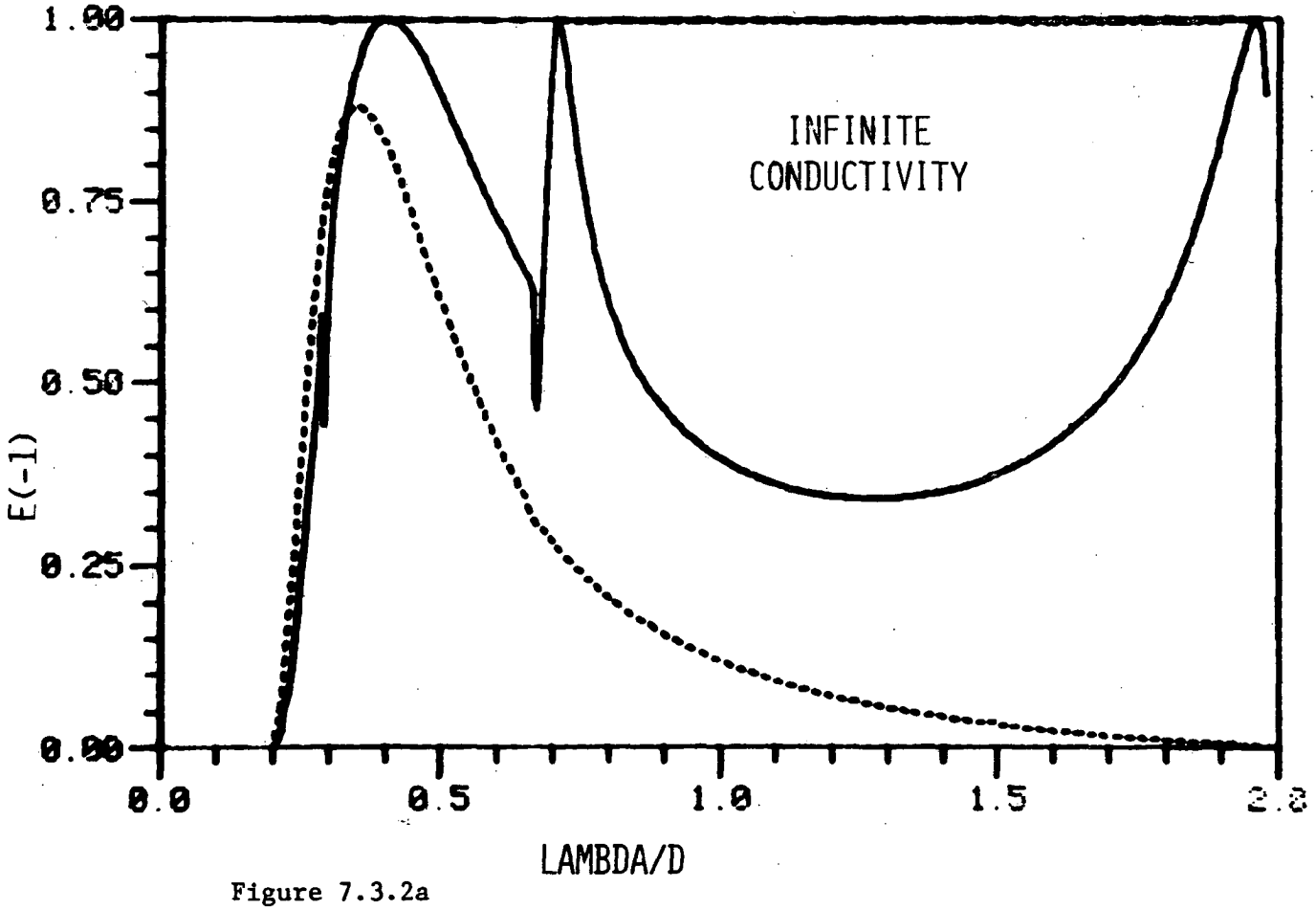


Figure 7.3.2a

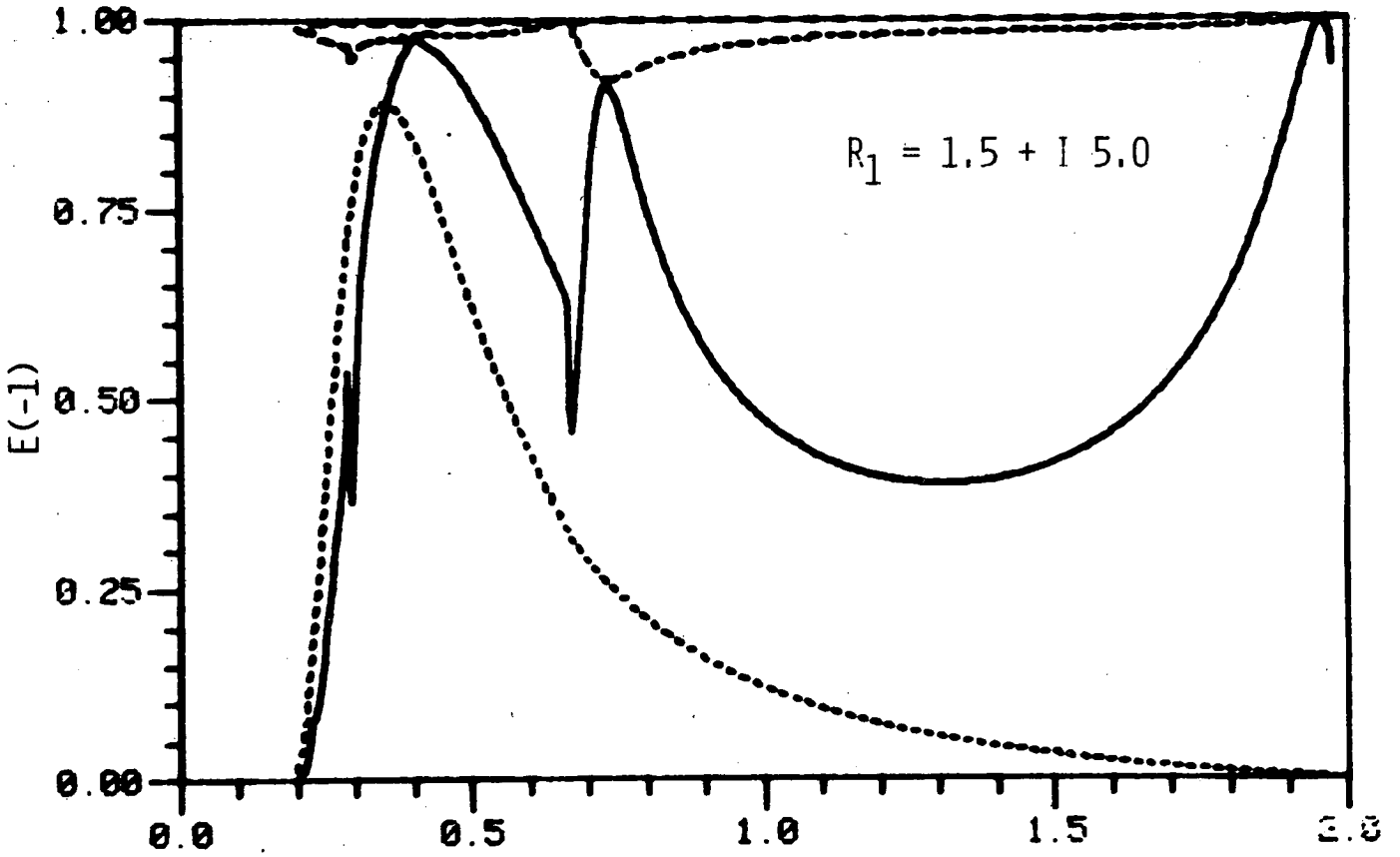


Figure 7.3.2b

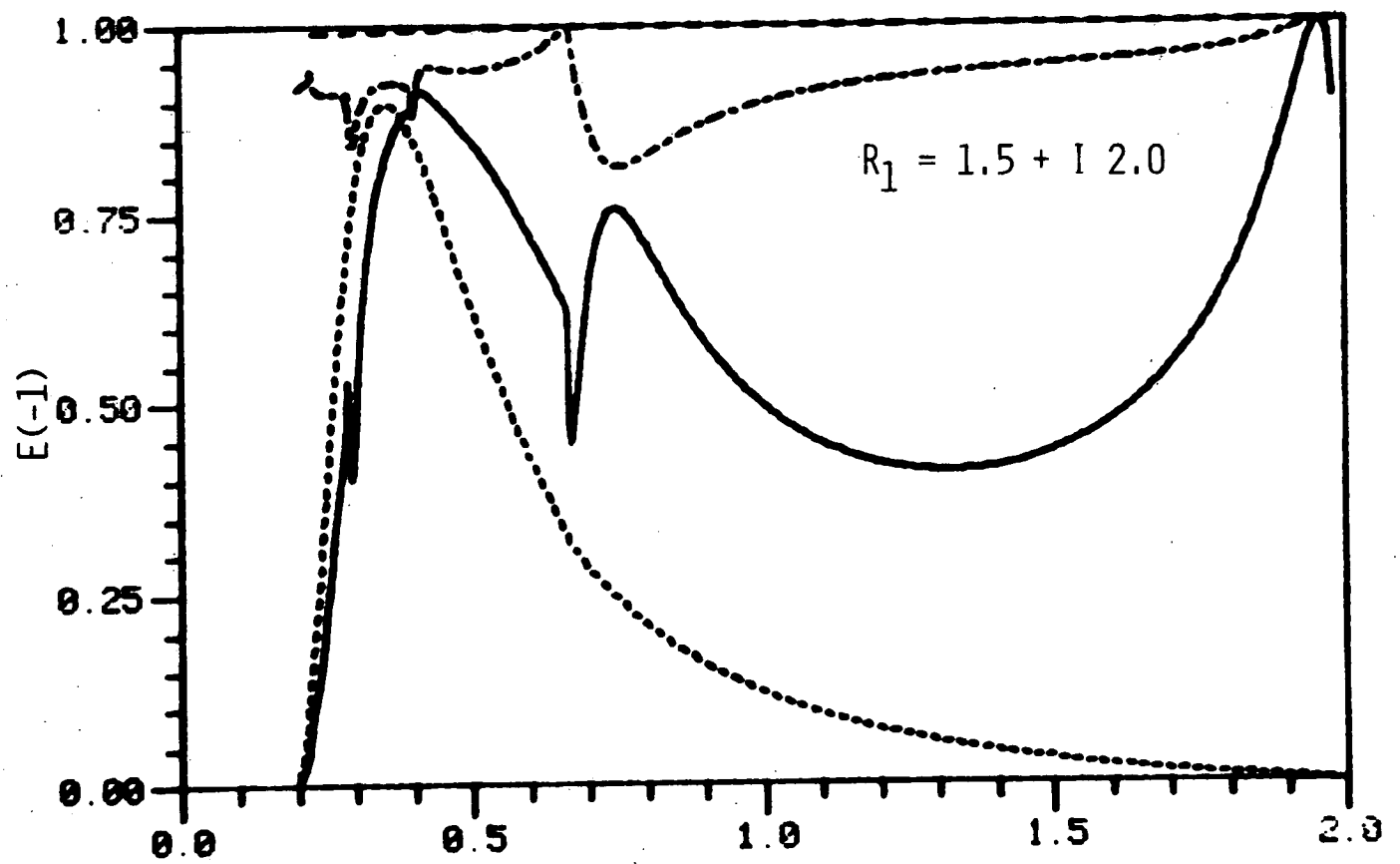


Figure 7.3.2c

of the plug. This phenomenon, together with the evidence shown in the total diffracted energy spectra, confirms that the effect of anomaly suppression is achieved using plasmon excitation. Note also the strengthening and broadening of the plasmons with decreasing conductivity. This increase in strength is a direct consequence of the diminished reflectance of the plug. The broadening effect is in agreement with the observations of Wheeler [7.5] (and also section (3.3) of this thesis) who noticed a strong correlation between the half width of these resonances and the ratio $Im(r_1)/Re(r_1)$.

Figures (7.3.2) also show that the S polarized blazewidth is substantially unaffected even though the blaze peak has been weakened. Note that the P polarization efficiency spectra are almost independent of the plug refractive index. Of course, this result is to be expected in view of the direction of current flow.

7.3.2.2 *The 19^o blaze angle grating*

In this study, a grating having a blaze angle of 19^o and an apex angle of 120^o, operated in a $(-1)^{th}$ order Littrow configuration was considered. Once again, the basic structure was elevated by an amount $h/d = 0.4$ above the substrate plane. As can be seen from figure (7.3.3a), this geometry, when composed entirely of perfectly conducting metal, exhibits a spike-like resonance directly following the Rayleigh wavelength at $\lambda/d = 2/3$. For an apex angle of 90^o, this spike is totally suppressed by the geometrical concept of blazing known as the Maréchal-Stroke relationship [7.6], which constrains the $(-1)^{th}$ order S polarization efficiency to be unity at this wavelength. However, as demonstrated by McPhedran and Waterworth [7.7] in their paper dealing

Figure 7.3.3

Diffracted energy spectra for a triangular profile grating having a blaze angle of 19^{o} and an apex angle of 120^{o}. The grating is displaced from the "ground" plane by an amount of h/d=0.4 and is operated in a $(-1)^{th}$ order Littrow configuration. On the various graphs are curves displaying:

(i) $(-1)^{th}$ order efficiency in P polarized light (- - - - -),

(ii) $(-1)^{th}$ order efficiency in S polarized light (————),

(iii) total diffracted energy in P polarized light (– – – →)

and

(iv) total diffracted energy in S polarized light (— — — —).

Graphs for plugs having wavelength refractive indices of:

(a) infinite conductivity,

(b) 1.5 + i 5.0,

(c) 1.5 + i 2.0,

(d) 1.5 + i 1.0,

(e) 1.5 + i 0.5 and

(f) 1.5 + i 0.2

are shown.

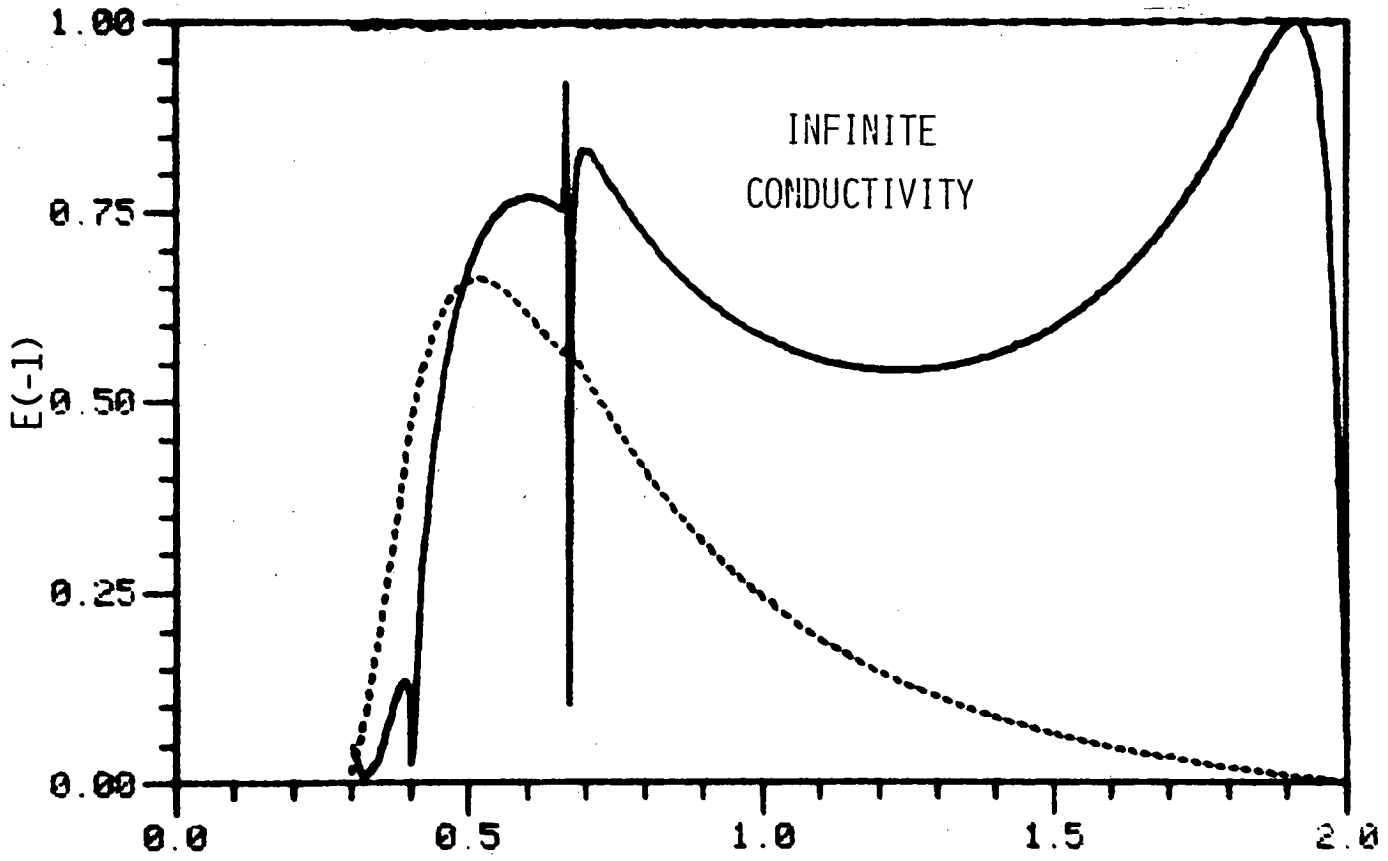


Figure 7.3.3a

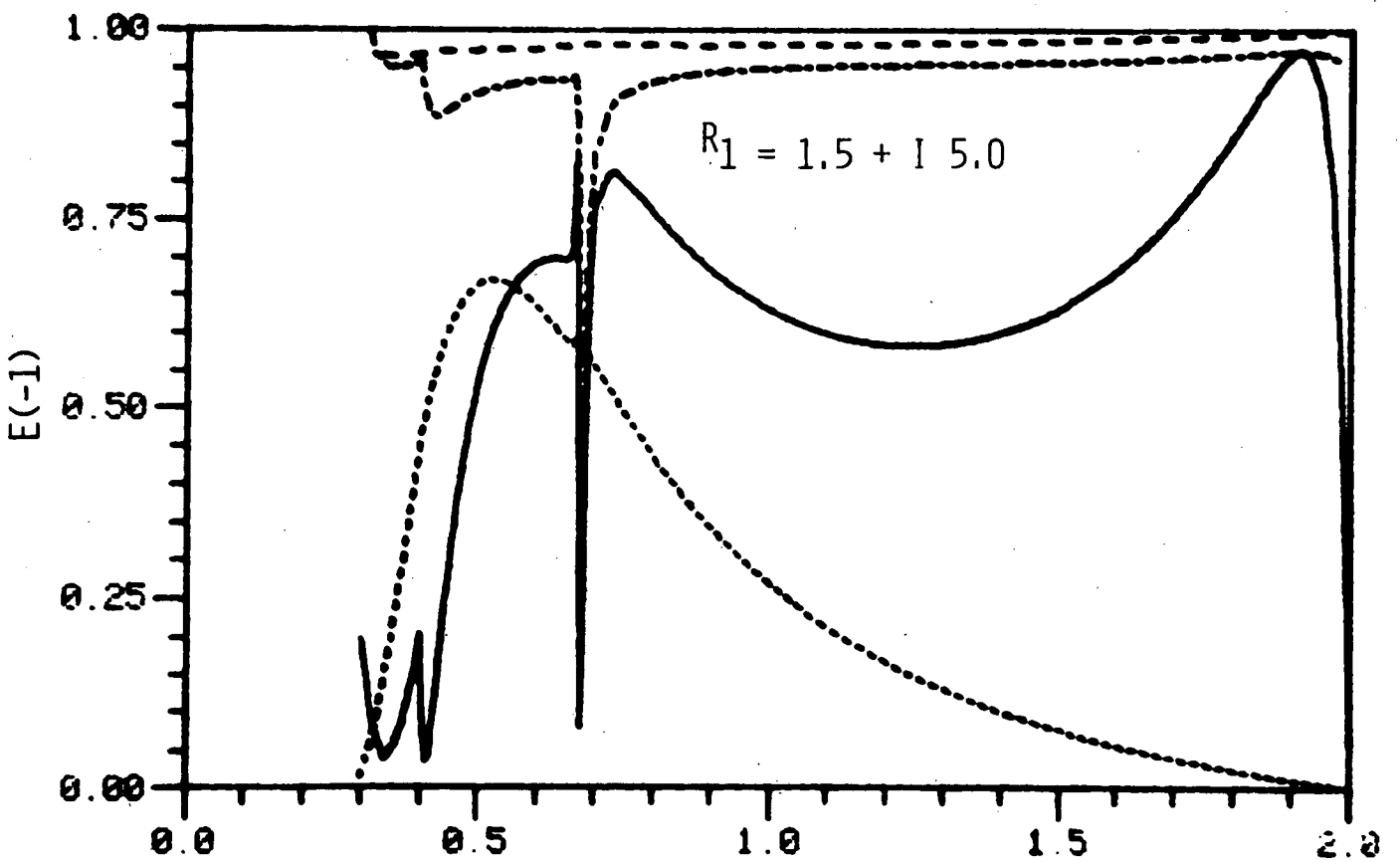


Figure 7.3.3b

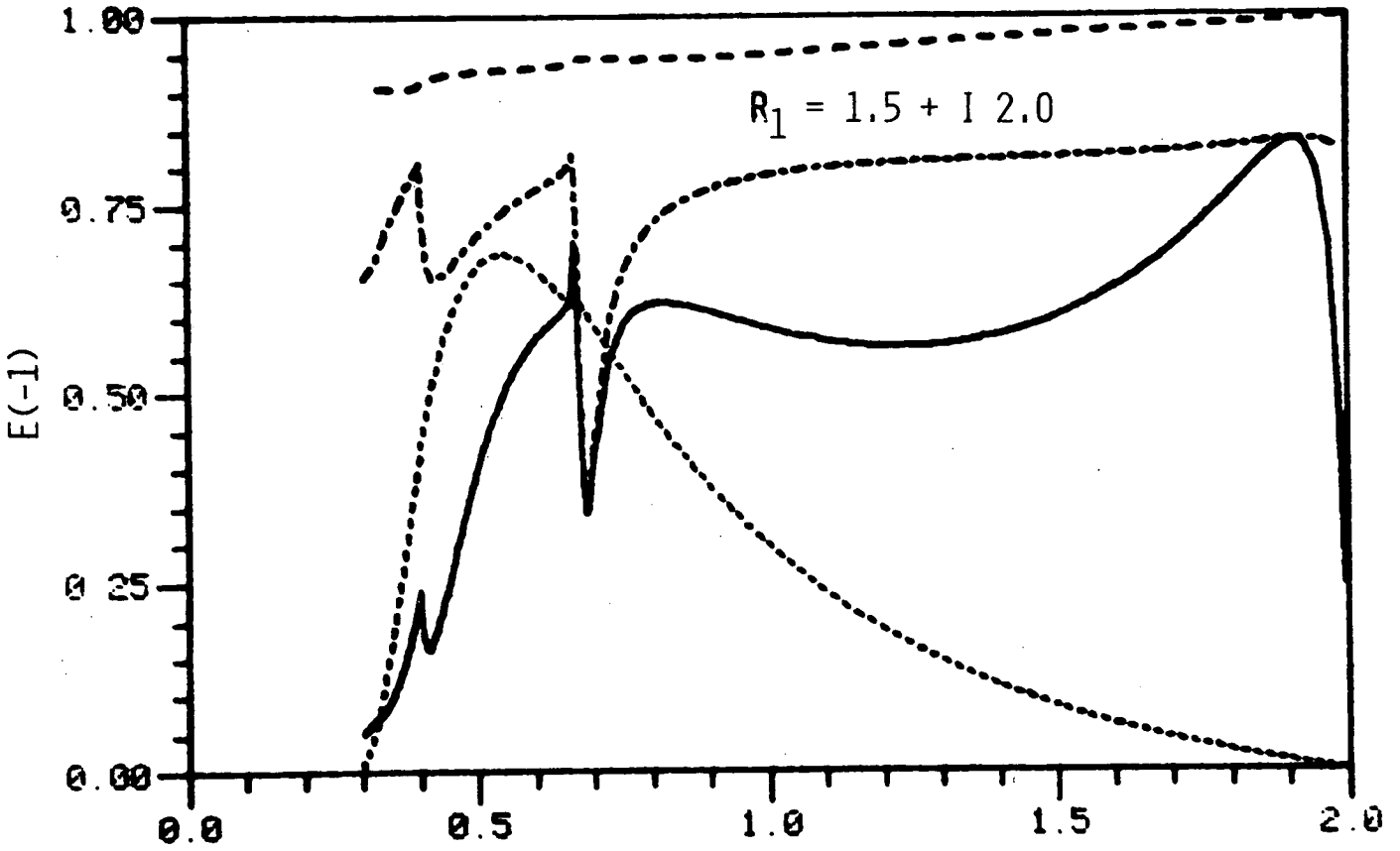


Figure 7.3.3c

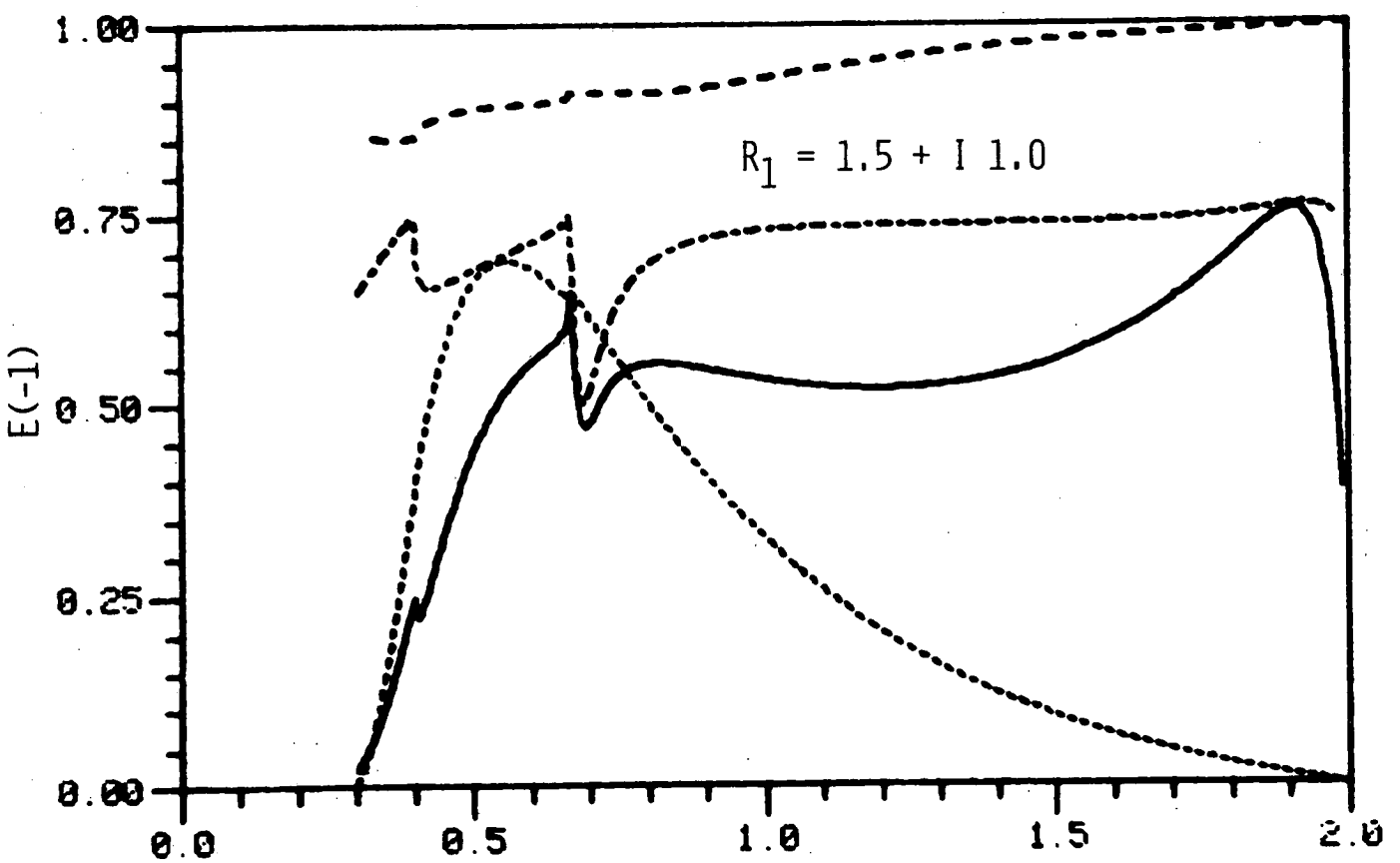


Figure 7.3.3d

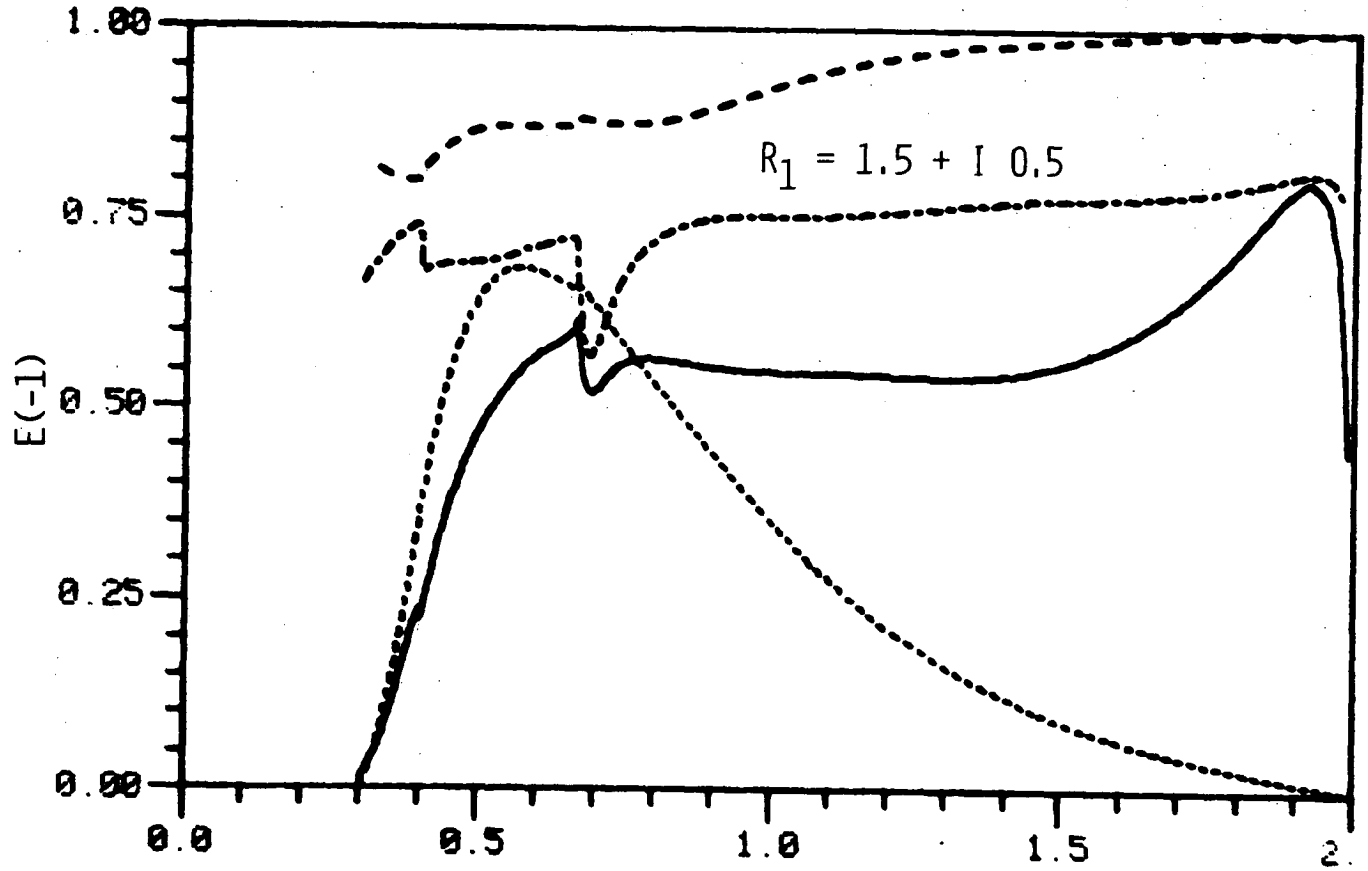


Figure 7.3.3e

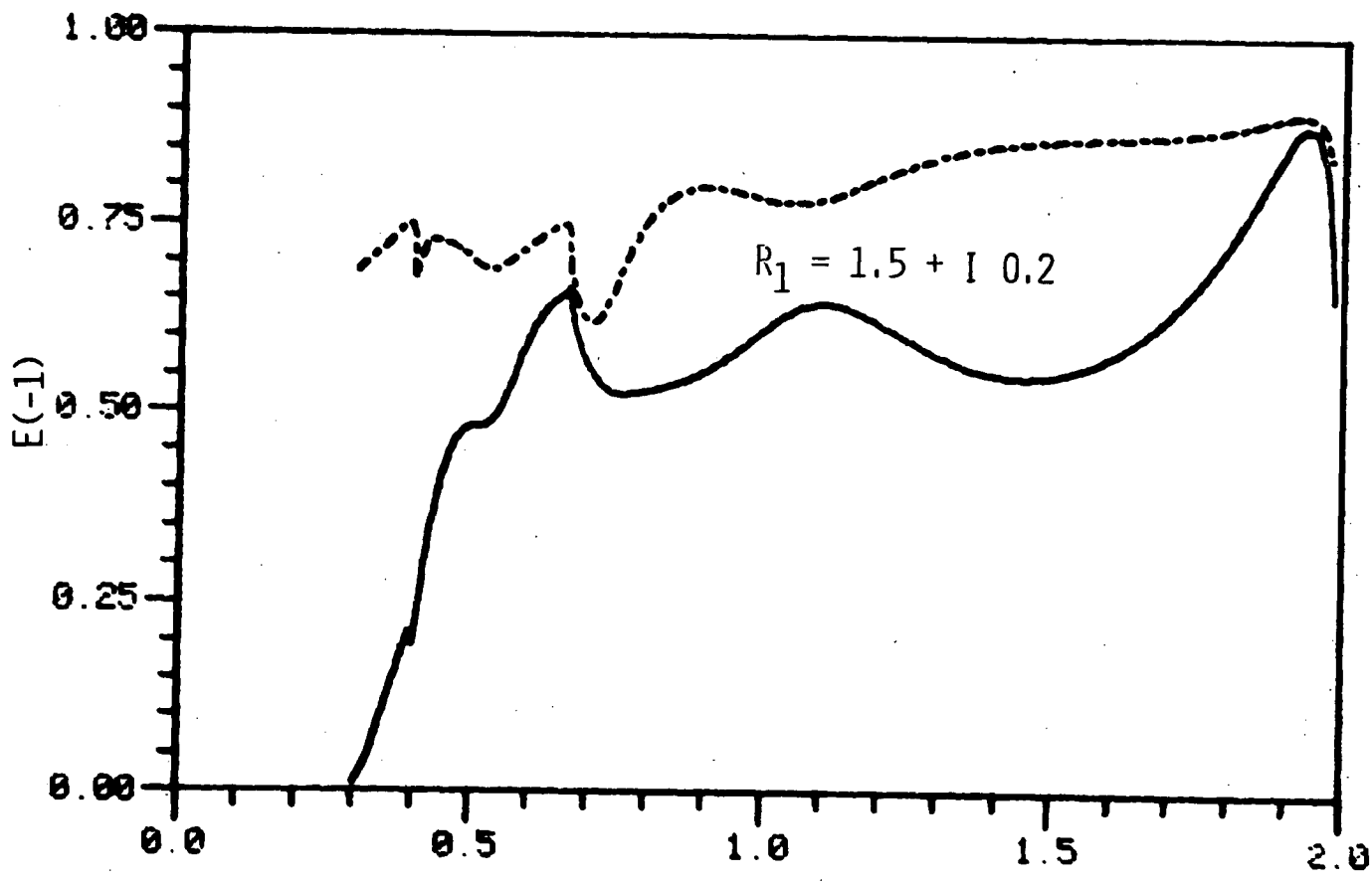


Figure 7.3.3f

with the blaze optimization of triangular profile gratings, any perturbation of the apex angle from 90^{o} (for this blaze angle) introduces anomalous behaviour at the passing off of the orders -2 and +1.

Thus, the insidious nature of this anomaly means that it is appropriate for this grating to be used as a significant test of the anomaly suppression hypothesis. Clearly, an inspection of figures (7.3.3) reveals that this hypothesis has been substantially verified. Once again the performance in P polarized light is observed to be independent of the value of the refractive index of the plug. This, together with the anomaly reduction existing for S polarization, tends to diminish the polarizance of the grating.

The S polarization efficiency spectra show a weakening of both the Wood anomaly (the leading edge) and the accompanying resonance (the deep spike) with decreasing conductivity. This is to be contrasted with the results of the previous study (for the 11.5^{o} apex angle grating) for which the Wood anomaly was largely unaltered. In that example, the Wood anomaly and its resonance were substantially displaced from one another.

These two phenomena are entirely dissimilar in nature with the Wood anomaly being a reaction to the passing on or off of various orders, whereas the resonance anomaly is a consequence of either elastic or inelastic plasmon scattering. Although the position of the former is a function of the periodicity of the structure and also of the incidence parameters, its strength is dependent on the geometry and the conductivity of the structure. On the other hand, the position and the strength of the resonance are governed by both of the above parameters.

Thus, one is led to surmise that the Wood anomaly can only be significantly affected if the resonance lies sufficiently close to the Rayleigh wavelength for inelastic plasmon scattering to smear any rapid variations in the spectral response. This is indeed the case for the grating currently under consideration, but not for the structure discussed earlier.

From figures (7.3.3) it can be seen that the resonance bandwidth increases monotonically with decreasing conductivity until $\mathrm{Im}(r_1) \lesssim 0.2$. This is also supported by the total energy curves which show a monotonic increase in the absorption baseline until the above conductivity constraint is reached. Moreover, should the plug conductivity drop below that specified by $\mathrm{Im}(r_1) = 0.2$, any field resonances characteristic of the cavity now manifest themselves as features in the plane wave spectra.

Figures (7.3.4) show the spectral response of the $(-1)^{th}$ order for this grating with a variety of lossless plugs. The most startling feature of these graphs is the prevalence of anomalous behaviour for plugs of high refractive index. To some extent, this can be understood in terms of an effective wavelength reduction permitting resonant behaviour in an increased number of the cavity modes. However, a more adequate explanation may be found using the recent theory of Andrewartha, Fox and Wilson [7.8] concerning poles (in the complex wavelength plane) of the modal amplitudes. Although their formalism is currently restricted to perfectly conducting lamellar gratings, it is hoped that their work may be extended so as to shed some light on these fascinating phenomena.

7.3.3 Concluding remarks

In summary, anomaly suppression can indeed be achieved by overcoating the "off-blaze" segment with a weakly conducting material. On

Figure 7.3.4

S polarization diffracted energy spectra for a bi-metallic grating having a triangular profile characterized by a blaze angle of 19^0 and an apex angle of 120^0. The bi-metallic insertion consists, in each case, of a lossless dielectric material. On each of the graphs are shown curves for:

(i) $(-1)^{th}$ order efficiency in S polarized light (————) and

(ii) total diffracted energy in S polarized light (- - - - - -),

the grating being operated in a $(-1)^{th}$ order Littrow mount. Graphs for plugs having the following wavelength independent refractive indices are presented:

(a) infinite conductivity (for comparison purposes),

(b) 1.0 + i 0.0,

(c) 1.5 + i 0.0 and

(d) 2.5 + i 0.0.

In figure (7.3.4d) it is to be noted that isolated regions of poor convergence, exemplified by the total energy curve, are exhibited. Furthermore, a total of 216 points have been calculated for this graph and even so the spectral performance has not been completely characterized.

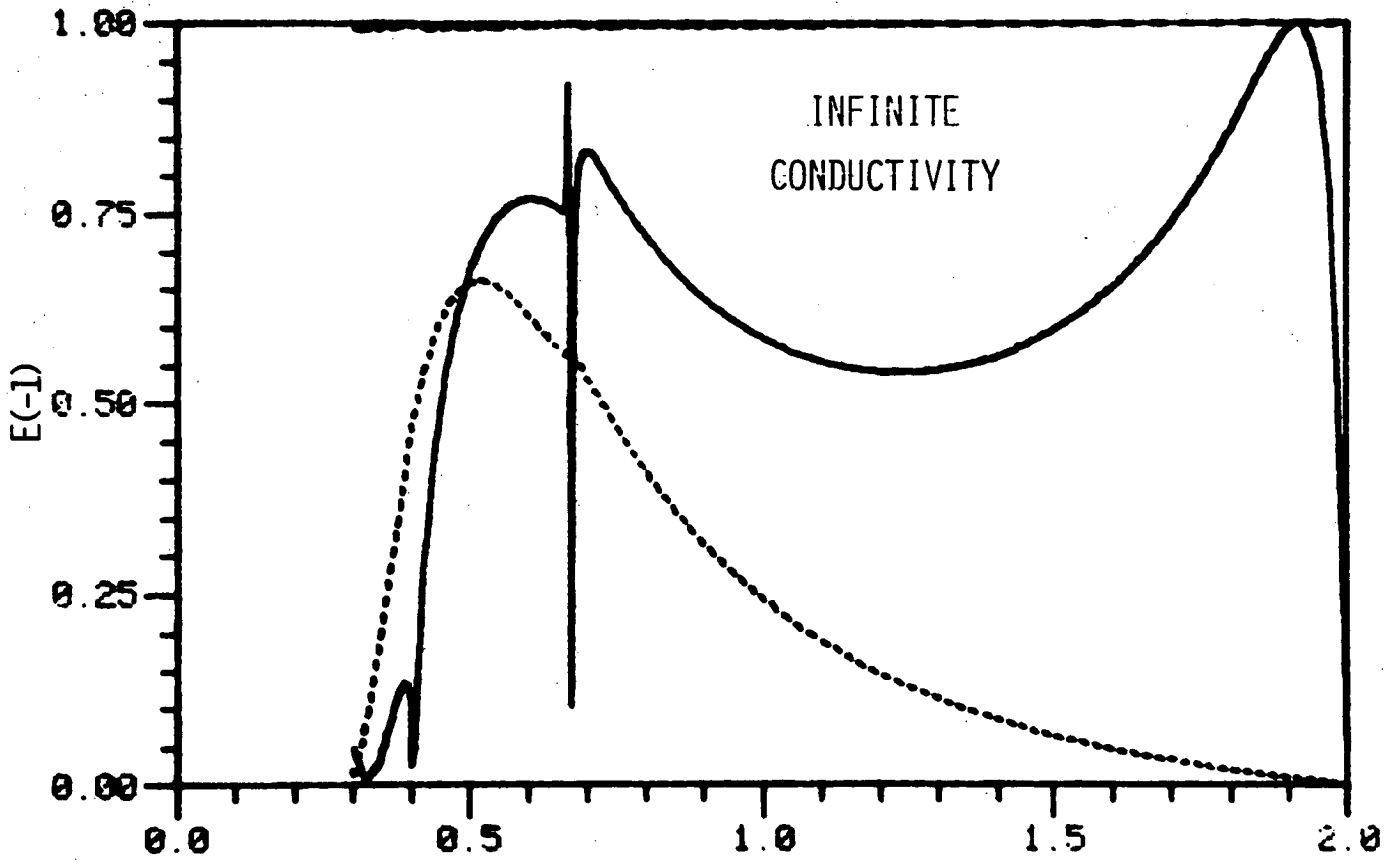


Figure 7.3.4a

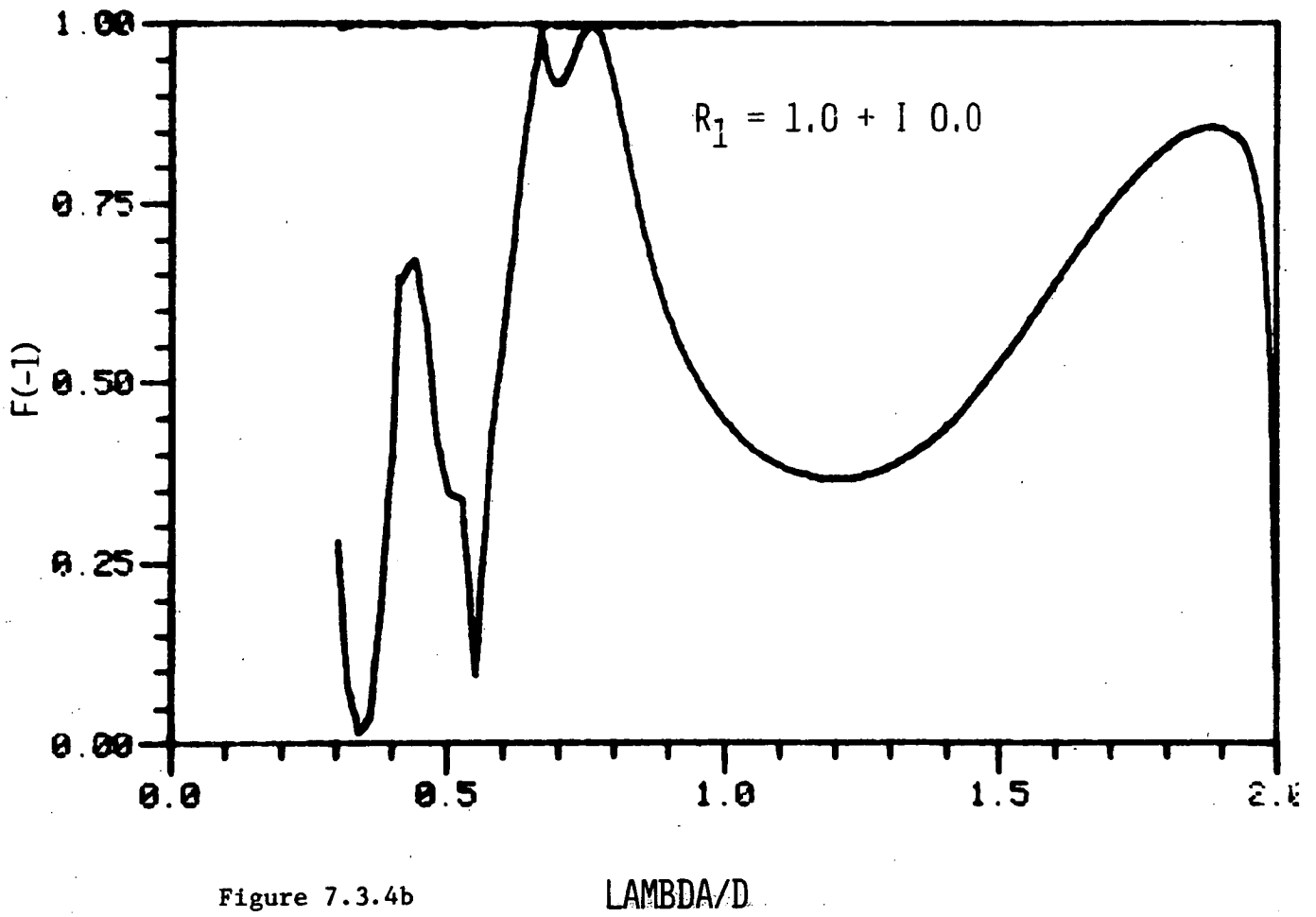


Figure 7.3.4b

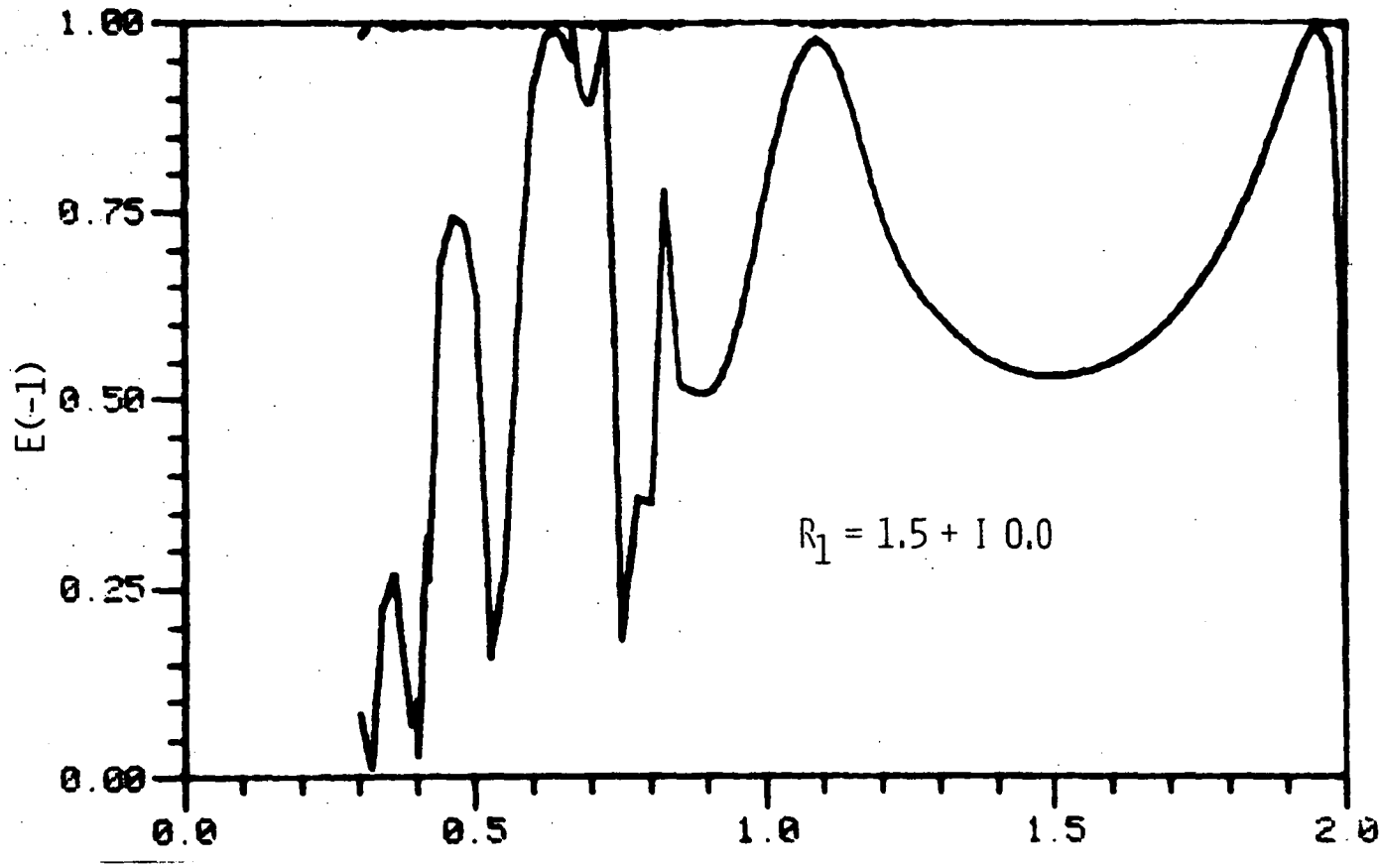


Figure 7.3.4c

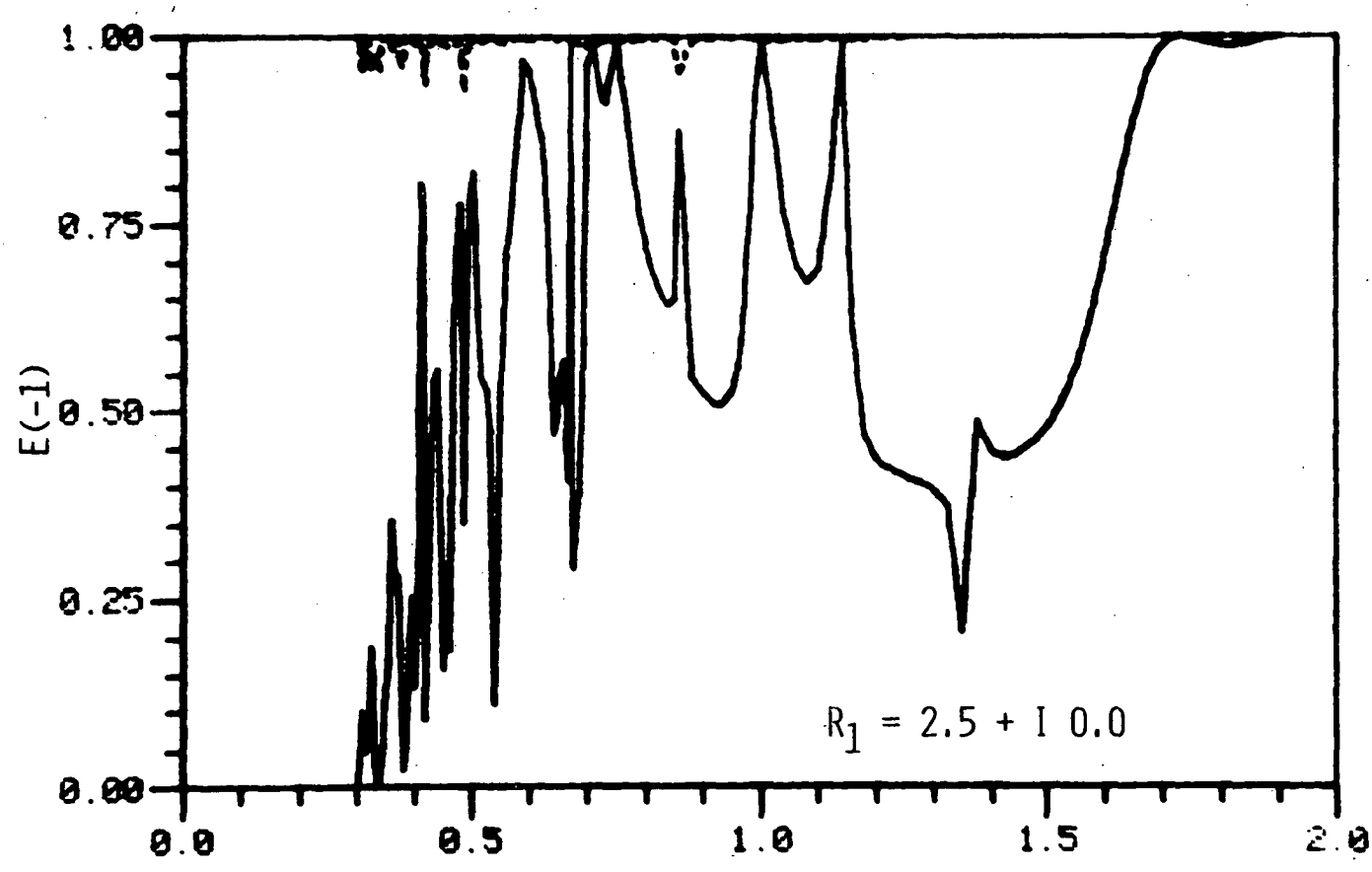


Figure 7.3.4d

the basis of the evidence presented above, it appears that the mechanism of anomaly reduction can be explained in terms of the excitation of lossy plasmons. The metal overcoating needs to have a relatively low conductivity in order to create plasmon resonances of a sufficiently broad nature so as to mask any sharp spectral features. This introduces an upper bound on the value of the plug conductivity. However it is also necessary to enforce a lower bound as dictated by the two criteria mentioned in section (7.1). Since both are dependent on the grating and film geometries and also on the conductivity of the overcoated material, it is impossible to make any general or categorical statements concerning them.

This clearly indicates the need for a more powerful theory capable of dealing with arbitrary geometries and conductivities. Such a theory is proposed in a later section of this chapter. Nevertheless, this study, undertaken using a rather restricted formalism, has yielded some very useful results in the sphere of anomaly suppression and in this sense it can be classed as a worthwhile exercise. Further promising results, of considerable importance to spectrographic grating designers, are anticipated using more general theories.

7.4 THE FORMALISM FOR A SEMI-FINITELY CONDUCTING BI-METALLIC LAMELLAR TRANSMISSION GRATING

7.4.1 Specification of the problem

This treatment of the diffraction problem concerning a bi-metallic transmission grating follows closely upon the analysis presented in section (7.2). The model presented herein is restricted to the structure

illustrated in figure (7.4.1) - an array of periodically spaced perfectly conducting prisms (having vertical walls) isolated from one another by a material of finite conductivity. In this treatment it is appropriate to subdivide space into four regions:

(i) D_0 - a semi-infinite medium representing free space characterized by a wave vector of magnitude $k_0 = 2\pi/\lambda$,

(ii) D_1 - a finitely conducting medium of refractive index r_1 whose spatial range is defined by $x \in [0,c]$, $y \in [f_\ell(x), f_u(x)]$,

(iii) D_2 - a perfectly conducting prism

and

(iv) D_3 - another semi-infinite medium whose refractive index is r_3, i.e. $k_3 = k_0 r_3$.

Once again, the vertical walls of the prisms have been chosen to be linear in order to facilitate a "semi-modal" treatment (i.e. a formulation implicitly exploiting the boundary conditions upon the metal walls). The profile faces $\Gamma_u (\equiv \Gamma_u^\ell \cup \Gamma_u^r)$ and $\Gamma_\ell (\equiv \Gamma_\ell^\ell \cup \Gamma_\ell^r)$ can be of arbitrary shape.

The problem to be treated in this section is restricted to the case of P polarized incident radiation. Naturally, the conversion of the theory to cover the case of S polarization follows in a manner analogous to that presented in section (7.2).

7.4.2 Characterization of the field problem

By pursuing an analysis similar to that undertaken in section (7.2.2), the fields in the regions D_0 and D_3 can be uniquely represented

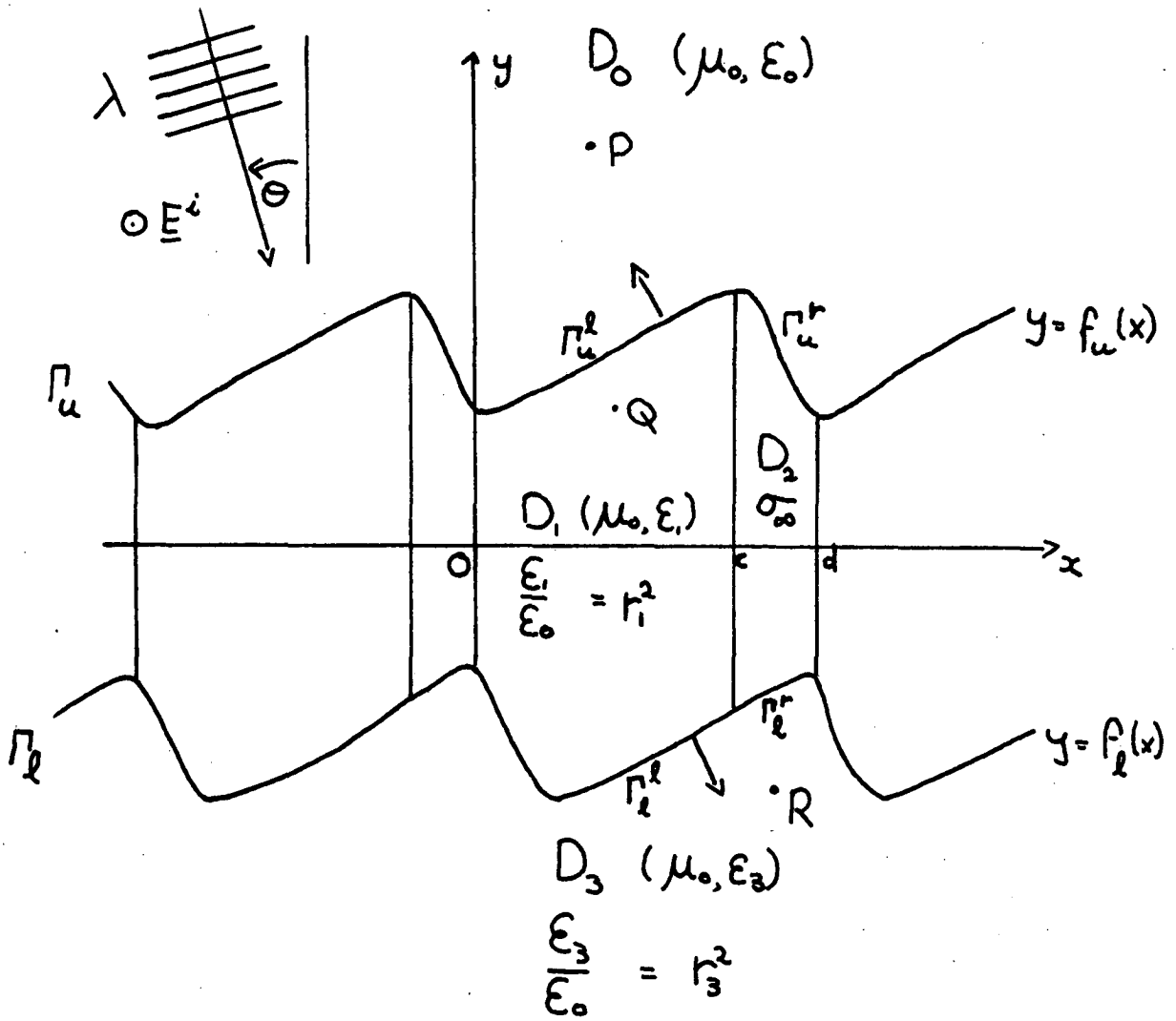


Figure 7.4.1 The diffracting geometry of the roughened lamellar transmission grating. The structure is composed of an array of perfectly conducting prisms (D_2) having vertical walls and separated from one another by a series of finitely conducting spacers (D_1) of refractive index r_1. The entire grating resides upon a substrate (D_3) of refractive index, r_3.

as functions of the "pseudo" surface currents on the interfaces Γ_u and Γ_ℓ respectively. This is achieved by defining functions u and v as follows:

$$\left.\begin{aligned} u(P) &= E^d(P) \quad \text{if} \quad P \in D_0 \\ (\nabla^2 + k_0^2)\, u(P) &= 0 \quad \text{if} \quad P \notin D_0 \\ u &\text{ is everywhere continuous} \end{aligned}\right\} \tag{7.4.1}$$

and

$$\left.\begin{aligned} v(R) &= E(R) \quad \text{if} \quad R \in D_3 \\ (\nabla^2 + k_3^2)\, v(R) &= 0 \quad \text{if} \quad R \notin D_3 \\ v &\text{ is everywhere continuous.} \end{aligned}\right\} \tag{7.4.2}$$

From these definitions, it is possible to extract the "pseudo" currents ξ and ψ, existing on the surfaces Γ_u and Γ_ℓ respectively, which uniquely specify the entire diffraction problem.

Here,

$$\xi(M_u) = \left.\frac{\partial u}{\partial n}\right|_{\Gamma_u +} - \left.\frac{\partial u}{\partial n}\right|_{\Gamma_u -} \tag{7.4.3}$$

and

$$\psi(M_\ell) = \left.\frac{\partial v}{\partial n}\right|_{\Gamma_\ell -} - \left.\frac{\partial v}{\partial n}\right|_{\Gamma_\ell +} \tag{7.4.4}$$

where M_u and M_ℓ are points lying on Γ_u and Γ_ℓ respectively.

After application of Green's Theorem to the appropriate domains, the electric fields in the regions D_0 and D_3 are specified by the following expansions

$$E(P) = E^i(P) + \int_{\Gamma_u} G_0(P;M_u') \, \xi(M_u') \, ds' \qquad \text{for} \quad P \in D_0 \tag{7.4.5}$$

and $$E(R) = \int_{\Gamma_\ell} G_3(R;M_\ell') \, \psi(M_\ell') \, ds' \qquad \text{for} \quad R \in D_3 \tag{7.4.6}$$

where the Green's function G_3 obeys the inhomogeneous Helmholtz equation

$$(\nabla^2 + k_3^2) \, G_3(R;R') = \sum_{n=-\infty}^{\infty} \delta(x-x'-nd) \exp(i\alpha_0 nd) \, \delta(y-y')$$

whose solution is given by

$$G_3(R;R') = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\eta_n} \exp[i\alpha_n(x-x') + i\eta_n|y-y'|] \tag{7.4.7}$$

with $$\eta_n = \sqrt{k_3^2 - \alpha_n^2} \, .$$

Thus it follows that

$$E(M_u) = E^i(M_u) + \int_{\Gamma_u} G_0(M_u;M_u') \, \xi(M_u') \, ds' \tag{7.4.8a}$$

$$\frac{\partial E(M_u)}{\partial n}\bigg|_{\Gamma_u +} = \frac{\partial E^i(M_u)}{\partial n} + \tfrac{1}{2} \, \xi(M_u) + \int_{\Gamma_u} \frac{\partial G_0(M_u;M_u')}{\partial n} \, \xi(M_u') \, ds' \tag{7.4.8b}$$

and

$$E(M_\ell) = \int_{\Gamma_\ell} G_3(M_\ell; M'_\ell)\, \psi(M'_\ell)\, ds' \qquad (7.4.9a)$$

$$\frac{\partial E(M_\ell)}{\partial n} = \tfrac{1}{2}\, \psi(M_\ell) + \int_{\Gamma_\ell} \frac{\partial G_3(M_\ell; M'_\ell)}{\partial n}\, \psi(M'_\ell)\, ds'. \qquad (7.4.9b)$$

7.4.3 Derivation of an integral equation

In the previous section, the existence of "pseudo" currents on the upper and lower interfaces of the structure uniquely characterizing the fields in D_0 and D_3, was demonstrated. In order to solve the diffraction problem, one must now impose field boundary conditions in an attempt to link these two entities. This is achieved by application of

(i) the direct boundary condition E = 0 on the metal surfaces Γ_u^r and Γ_ℓ^r,

and

(ii) the indirect boundary conditions involving the continuity of E and $\frac{\partial E}{\partial n}$ at the spatially separated interfaces Γ_u^ℓ and Γ_ℓ^ℓ.

The direct boundary conditions lead to

$$-E^i(M_u^r) = \int_{\Gamma_u} G_0(M_u^r; M'_u)\, \xi(M'_u)\, ds' \qquad (7.4.10a)$$

$$0 = \int_{\Gamma_\ell} G_3(M_\ell^r; M'_\ell)\, \psi(M'_\ell)\, ds' \qquad (7.4.10b)$$

for points M_u^r and M_ℓ^r on Γ_u^r and Γ_ℓ^r respectively. However, the indirect nature of the boundary conditions on the spatially separated interfaces Γ_u^ℓ and Γ_ℓ^ℓ necessitates the derivation of a separate field expansion in D_1 (characteristic of that region) before attempting to apply the continuity equations. Since the electric field vanishes on the vertical

linear walls of the domain D_1, it is appropriate to utilize a Green's function whose boundary conditions emulate those of the field quantity. For this problem, the chosen Green's function is

$$G_1(Q;M') = -\frac{i}{c}\sum_{m=1}^{\infty}\frac{1}{\mu_m}\sin(\frac{m\pi x}{c})\sin(\frac{m\pi x'}{c})\exp[i\,\mu_m|y-y'|] \qquad (7.4.11a)$$

obeying the differential equation

$$(\nabla^2 + k_1^2)\,G_1(Q;M') = \sum_{m=-\infty}^{\infty}[\delta(x-x'+2mc) - \delta(x+x'+2mc)]\,\delta(y-y'). \qquad (7.4.11b)$$

Applying Green's Theorem to the domain D_1 leads to the following field representation for the electric field at some point Q (within D_1).

$$E(Q) = -\int_{\Gamma_u^\ell}[G_1(Q;M_u^{\ell\prime})\frac{\partial E(M_u^{\ell\prime})}{\partial n'} - \frac{\partial G_1(Q;M_u^{\ell\prime})}{\partial n'}E(M_u^{\ell\prime})]\,ds'$$
$$-\int_{\Gamma_\ell^\ell}[G_1(Q;M_\ell^{\ell\prime})\frac{\partial E(M_\ell^{\ell\prime})}{\partial n'} - \frac{\partial G_1(Q;M_\ell^{\ell\prime})}{\partial n'}E(M_\ell^{\ell\prime})]\,ds'\,. \qquad (7.4.12)$$

This expression involves no contribution from line integrals along the vertical segments, a consequence of the "simple" groove geometry which has enabled a "semi-modal" treatment of the problem implicitly incorporating the direct boundary conditions within D_1.

Should the field be sampled at points M_u^ℓ and M_ℓ^ℓ on Γ_u^ℓ and Γ_ℓ^ℓ respectively, the following pair of coupled integral equations results:

$$\tfrac{1}{2}E(M_u) = -\int_{\Gamma_u^\ell} [G_1(M_u^\ell;M_u^{\ell\prime}) \frac{\partial E}{\partial n'}(M_u^{\ell\prime}) - \frac{\partial G_1}{\partial n'}(M_u^\ell;M_u^{\ell\prime}) E(M_u^{\ell\prime})]\, ds'$$

$$- \int_{\Gamma_\ell^\ell} [G_1(M_u^\ell;M_\ell^{\ell\prime}) \frac{\partial E}{\partial n'}(M_\ell^{\ell\prime}) - \frac{\partial G_1}{\partial n'}(M_u^\ell;M_\ell^{\ell\prime}) E(M_\ell^{\ell\prime})]\, ds' \quad (7.4.13a)$$

and

$$\tfrac{1}{2}E(M_\ell) = -\int_{\Gamma_u^\ell} [G_1(M_\ell^\ell;M_u^{\ell\prime}) \frac{\partial E}{\partial n'}(M_u^{\ell\prime}) - \frac{\partial G_1}{\partial n'}(M_\ell^\ell;M_u^{\ell\prime}) E(M_u^{\ell\prime})]\, ds'$$

$$- \int_{\Gamma_\ell^\ell} [G_1(M_\ell^\ell;M_\ell^{\ell\prime}) \frac{\partial E}{\partial n'}(M_\ell^{\ell\prime}) - \frac{\partial G_1}{\partial n'}(M_\ell^\ell;M_\ell^{\ell\prime}) E(M_\ell^{\ell\prime})]\, ds' \,. \quad (7.4.13b)$$

Field continuity conditions across the boundaries Γ_u^ℓ and Γ_ℓ^ℓ reduce equations (7.4.13) to the following forms, upon substitution of representations for E and $\frac{\partial E}{\partial n}$ just exterior to the region D_1 (equations (7.4.8-9)):

$${}_u\xi_0(M_u^\ell) = \int_{\Gamma_u} {}_u\tau_u(M_u^\ell;M_u')\, \xi(M_u')\, ds' + \int_{\Gamma_\ell} {}_u\tau_\ell(M_u^\ell;M_\ell')\, \psi(M_\ell')\, ds' \quad (7.4.14a)$$

where

$${}_u\xi_0(M_u^\ell) = \tfrac{1}{2}E^i(M_u^\ell) + \int_{\Gamma_u^\ell} [G_1(M_u^\ell;M_u^{\ell\prime}) \frac{\partial E^i}{\partial n'}(M_u^{\ell\prime}) - \frac{\partial G_1}{\partial n'}(M_u^\ell;M_u^{\ell\prime}) E^i(M_u^{\ell\prime})]ds' \quad (7.4.14b)$$

$${}_u\tau_u(M_u^\ell;M_u') = -\tfrac{1}{2}G_0(M_u^\ell;M_u') - \tfrac{1}{2}A(M_u')\, G_1(M_u^\ell;M_u^{\ell\prime})$$

$$+ \int_{\Gamma_u^\ell} [\frac{\partial G_1}{\partial n''}(M_u^\ell;M_u^{\ell\prime\prime})\, G_0(M_u^{\ell\prime\prime};M_u') - G_1(M_u^\ell;M_u^{\ell\prime\prime}) \frac{\partial G_0}{\partial n''}(M_u^{\ell\prime\prime};M_u')]ds'' \quad (7.4.14c)$$

$$ {}_{u}\tau_{\ell}(M_u^{\ell};M_{\ell}') = -\tfrac{1}{2}A(M_{\ell}')\, G_1(M_u^{\ell};M_{\ell}^{\ell}') + \int_{\Gamma_{\ell}} [\frac{\partial G_1(M_u^{\ell};M_{\ell}^{\ell}")}{\partial n"} G_3(M_{\ell}^{\ell}";M_{\ell}') - G_1(M_u^{\ell};M_{\ell}^{\ell}") \frac{\partial G_3(M_{\ell}^{\ell}";M_{\ell}')}{\partial n"}]\, ds" \quad (7.4.14d) $$

and

$$ {}_{\ell}\xi_0(M_{\ell}^{\ell}) = \int_{\Gamma_u} {}_{\ell}\tau_u(M_{\ell}^{\ell};M_u')\, \xi(M_u')\, ds' + \int_{\Gamma_{\ell}} {}_{\ell}\tau_{\ell}(M_{\ell}^{\ell};M_{\ell}')\, \psi(M_{\ell}')\, ds' \quad (7.4.15a) $$

where

$$ {}_{\ell}\xi_0(M_{\ell}^{\ell}) = \int_{\Gamma_u^{\ell}} [G_1(M_{\ell}^{\ell};M_u^{\ell}') \frac{\partial E^i(M_u^{\ell}')}{\partial n'} - \frac{\partial G_1(M_{\ell}^{\ell};M_u^{\ell}')}{\partial n'} E^i(M_u^{\ell}')]ds' \quad (7.4.15b) $$

$$ {}_{\ell}\tau_u(M_{\ell}^{\ell};M_u') = -\tfrac{1}{2}A(M_u')\, G_1(M_{\ell}^{\ell};M_u^{\ell}') + \int_{\Gamma_u^{\ell}} [\frac{\partial G_1(M_{\ell}^{\ell};M_u^{\ell}")}{\partial n"} G_0(M_u^{\ell}";M_u') - G_1(M_{\ell}^{\ell};M_u^{\ell}") \frac{\partial G_0(M_u^{\ell}";M_u')}{\partial n"}]\, ds" \quad (7.4.15c) $$

$$ {}_{\ell}\tau_{\ell}(M_{\ell}^{\ell};M_{\ell}') = -\tfrac{1}{2}G_3(M_{\ell}^{\ell};M_{\ell}') - \tfrac{1}{2}A(M_{\ell}')\, G_1(M_{\ell}^{\ell};M_{\ell}^{\ell}') + \int_{\Gamma_{\ell}^{\ell}} [\frac{\partial G_1(M_{\ell}^{\ell};M_{\ell}^{\ell}")}{\partial n"} G_3(M_{\ell}^{\ell}";M_{\ell}') - G_1(M_{\ell}^{\ell};M_{\ell}^{\ell}") \frac{\partial G_3(M_{\ell}^{\ell}";M_{\ell}')}{\partial n"}]\, ds" \quad (7.4.15d) $$

Expressions (7.4.10a), (7.4.10b), (7.4.14a), and (7.4.15a) are the desired coupled integral equations completely characterizing the diffraction problem. The calculation of the unknowns ξ and ψ enables the total reconstruction of the electromagnetic field in all domains.

7.4.4 Concluding remarks

At the commencement of this study concerning the diffraction properties of "bi-metallic" gratings, it was immediately evident that the

problem was of a "non-trivial" nature. Since this class of structures was entirely unfamiliar to the author, it was decided that the only reasonable method of attack on the problem at hand was to follow a slow, painstaking and systematic program of study in an attempt to gain some experience, indeed confidence, in dealing with such problems. Furthermore, before attempting to evolve a fully general formulation, it was thought necessary to establish whether or not any physically significant result could arise from such an effort.

Thus, in this context, the formalisms and results discussed in this and the previous sections represent the first attempt in characterizing such a problem. Of course, it must be admitted that these treatments involve massive restrictions concerning groove geometry and surface conductivity. However, as has been demonstrated above, these assumptions have led to a relatively elegant formulation of the diffraction problem and the results so obtained have indicated the need for further studies culminating in the development of a fully general theory.

7.5 DIFFRACTION BY A BURIED CYLINDER GRATING

7.5.1 Introductory comments

As can be seen from figure (7.5.1), the structure to be considered within this section consists of a periodically spaced array of imperfectly conducting cylinders buried in a planar slab of finitely conducting material. The manner in which this section fits into the evolutionary framework leading to a completely general theory may not be immediately evident and thus one feels duty bound to justify the inclusion of this formalism.

The previous sections have dealt with structures involving specific restrictions with regard to their groove geometry and surface conductivity. These restrictions have enabled the formulation of semi-modal theories implicitly incorporating the field boundary conditions on the groove walls. However, such elegant techniques are not transferable to the more general problem. This remark must be qualified by saying that such formalisms are achievable only in those cases for which it is possible to derive analytic expressions for the field modes. Currently, the derivation of analytic modes is restricted to structures having highly specific geometries and perfectly conducting surfaces. The techniques by which analytic modes may be constructed for generalized geometries are not at all evident; let alone involving the more difficult transition from infinite to finite conductivity (for which the direct boundary conditions are replaced by constraints of a completely indirect nature).

Thus, until such time as modal formulations have made a number of significant advances, these methods will have to be abandoned in favour of the tried and proven integral formalisms, limited by neither structural geometry nor surface conductivity. Hence, in making the transition from infinite to finite conductivity, with the concomitant reversion to full integral formalisms, it was felt advisable to follow the usual systematic approach and consider a structure having a relatively simple geometry. As can be seen from the diagram, the restrictions concerning conductivity and groove (or cylinder) geometry have been lifted but at the expense of constraining the slab geometry to be rectangular. However it was felt that such a theory warranted some closer attention since:

(i) it represented a continuation in the study of a class of problems (e.g. the multi-layer transmission grating) in which some experience had been already gained;

and

(ii) furthermore, it provided a useful "bench-mark" against which the fully general treatment could be tested.

7.5.2 Specification of the problem

Consider an array of finitely conducting cylinders of arbitrary cross-section arranged periodically inside a planar slab of finitely conducting material. The entire space is divided into a set of four distinct regions $\{D_i\}$ (where $i \in [0,3]$) having refractive indices $\{r_i\}$ such that $r_0 = r_3 = 1$ (i.e. free space).

The problem which will now be discussed concerns the diffraction of a P polarized wave of free space wavelength λ striking the structure of period d, illustrated in figure (7.5.1), at some angle θ_i. Here the electric component of the incident field is given by

$$E^i = \exp [i(\alpha_0 x - \chi_0 y)]. \qquad (7.5.1)$$

7.5.3 Field representation within the dielectric slab

By applying Green's Theorem to a subregion of D_1 (such that $x \in [0,d]$) bounded by the contour C it follows that

$$E(P) = \int_C [G_1(P;M') \frac{\partial E(M')}{\partial n'} - \frac{\partial G_1(P;M')}{\partial n'} E(M')]ds' \qquad (7.5.2)$$

for some point P lying within the contour C. Here the Green's function G_1 is given by

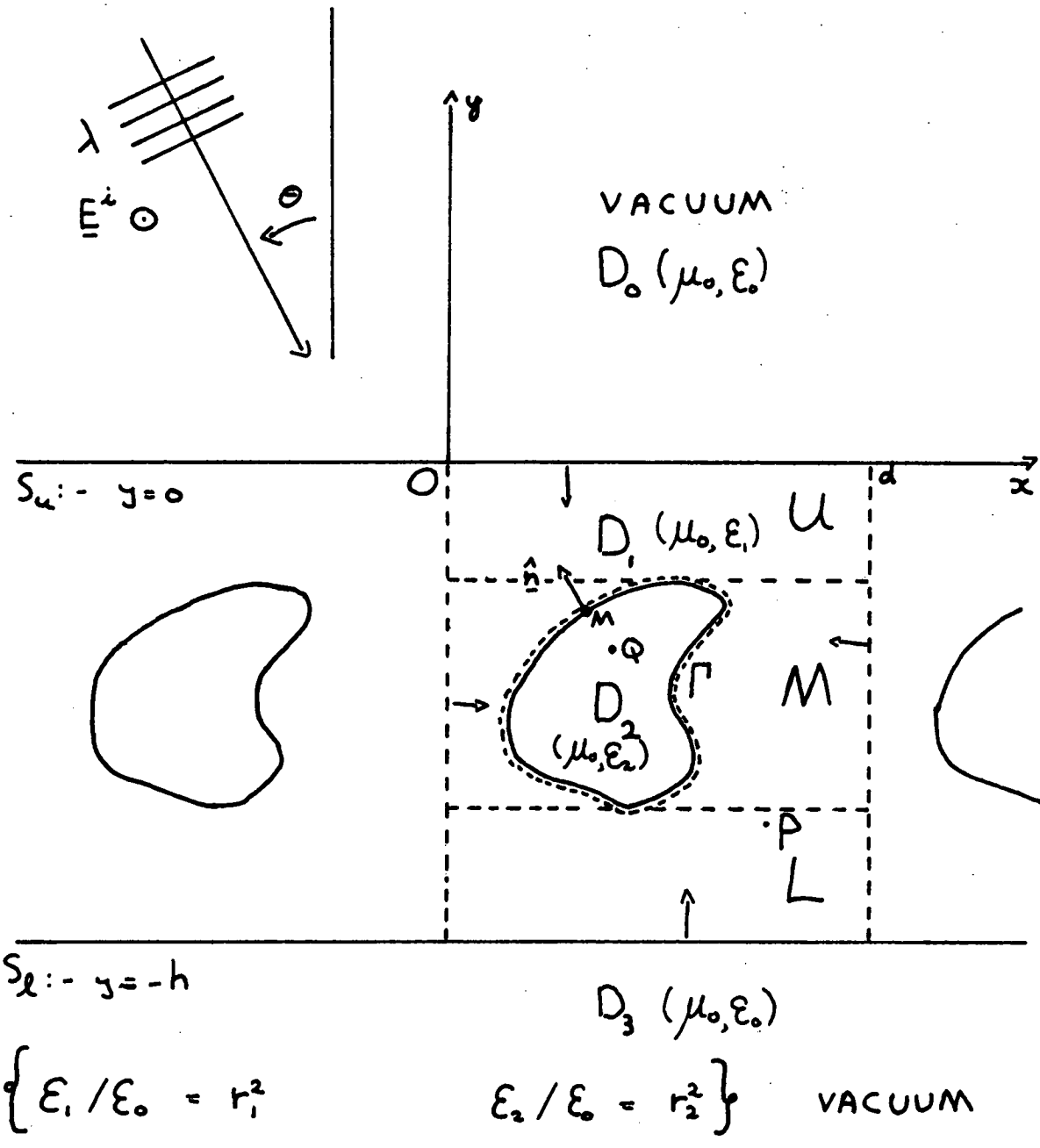


Figure 7.5.1

The geometry of the buried cylinder grating. The structure is composed of an array of infinitely long cylinders (denoted by D_2) of refractive index r_2 buried within a planar dielectric slab (denoted by D_1) of refractive index r_1 and surrounded on either side by free space.

$$G_1(P;M') = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\Gamma_n} \exp[i\alpha_n(x-x') + i\Gamma_n|y-y'|] \qquad (7.5.3)$$

where $\Gamma_n = \sqrt{k_1^2 - \alpha_n^2}$ with $k_1 = k_0 r_1$.

Since the region lying within C is a multiply-connected domain, the contour C is composed of

(i) the horizontal linear segments S_u and S_ℓ,

(ii) the vertical linear segments at $x' = 0,d$

and

(iii) the cylinder profile Γ.

However, from pseudo-periodicity arguments, the contributions to the line integral from the vertical linear segments may be shown to cancel with one another and so

$$\left.\begin{aligned} E(P) = &\int_{\Gamma_{ext}} \left[G_1 \frac{\partial E}{\partial n'} - \frac{\partial G_1}{\partial n'} E\right] ds' \\ &- \int_{S_u} \left[G_1 \frac{\partial E}{\partial y'} - \frac{\partial G_1}{\partial y'} E\right] dx' \\ &+ \int_{S_\ell} \left[G_1 \frac{\partial E}{\partial y'} - \frac{\partial G_1}{\partial y'} E\right] dx' . \end{aligned}\right\} \qquad (7.5.4)$$

Now, it may be shown that the line integrals along the segments S_u and S_ℓ contribute only downward and upward going plane waves respectively and since both wave trains are approaching the grating proper (i.e. the array of cylinders) their plane wave nature within the "grooves" (i.e. the region bounded in the x-direction by the cylinders)

must be preserved. If the region D_1 is further subdivided into domains U, M and L it follows that

$$E(P) = \sum_n [E_n^{U+} \exp(i\Gamma_n y) + E_n^{U-} \exp(-i\Gamma_n y)] \exp(i\alpha_n x) \qquad (7.5.5a)$$

for some point $P \equiv (x,y)$ within U

and $$E(P) = \sum_n [E_n^{L+} \exp(i\Gamma_n y) + E_n^{L-} \exp(-i\Gamma_n y)] \exp(i\alpha_n x) \qquad (7.5.5b)$$

for $P \in L$.

Within the region U, contributions to the upward going field are supplied by wave reflections at the interface S_ℓ and the outward going radiation in the form of plane waves from the cylinder surface Γ. Naturally, within the region L, downward going terms are generated by outward going radiation from the cylinders and reflections from the interface S_u. However, inside the domain M it is not possible to express the field in terms of a plane wave expansion having constant coefficients. To do so would be to invoke the Rayleigh approximation which has been shown previously to be a dubious assumption.

Now, consider the field representation within U. The contribution from the line integral around Γ yields a term of the form

$$\sum_n a_n^+ \exp(i\alpha_n x + i\Gamma_n y) \qquad (7.5.6a)$$

where $$a_n^+ = \int_{\Gamma_{ext}} [G_{1,n}^+ \frac{\partial E}{\partial n'} - \frac{\partial G_1^+}{\partial n'}{}_{,n} E] \, ds' \qquad (7.5.6b)$$

with $G_{1,n} = \frac{1}{2id\Gamma_n} \exp(-i\alpha_n x' - i\Gamma_n y')$. (7.5.6c)

(Here, the cylinder is specified by the relation $y' = f(x')$.)

By now utilizing these plane wave expansions by invoking the orthogonality of the basis $\{\exp(i\alpha_n x)\}$ over a period length it follows from equation (7.5.4) that

$$E_n^{U+} \exp(i\Gamma_n y) + E_n^{U-} \exp(-i\Gamma_n y) = (E_n^{L+} + a_n^+) \exp(i\Gamma_n y) + E_n^{U-} \exp(-i\Gamma_n y)$$

$$\text{i.e.} \quad E_n^{U+} = E_n^{L+} + a_n^+ . \tag{7.5.7}$$

Within the region L, the contribution from the line integral around the cylinder yields a downward going plane wave expansion of the form

$$\sum_n a_n^- \exp[i(\alpha_n x - \Gamma_n y)] \tag{7.5.8a}$$

$$\text{where } a_n^- = \int_{\Gamma_{ext}} [G_{1,n}^- \frac{\partial E}{\partial n'} - \frac{\partial G_{1,n}^-}{\partial n'} E] \, ds' \tag{7.5.8b}$$

$$\text{and } G_{1,n}^- = \frac{1}{2id\Gamma_n} \exp(-i\alpha_n x' + i\Gamma_n y') . \tag{7.5.8c}$$

By now performing some manipulations similar to those described above, it may be shown that

$$E_n^{L-} = E_n^{U-} + a_n^- . \tag{7.5.9}$$

Expressions (7.5.7) and (7.5.9) represent two equations of constraint on the physical system, currently involving the four sets of

unknowns $\{E_n^{U+}\}$, $\{E_n^{U-}\}$, $\{E_n^{L+}\}$ and $\{E_n^{L-}\}$. By now applying Fresnel's Laws at the interfaces S_u and S_ℓ it is possible to introduce a further two constraints upon the physical system. These are :

$$E_n^{U-} = T_0 \, \delta_{n,0} + R_n E_n^{U+} \tag{7.5.10}$$

and $$E_n^{L+} = R_n \exp(2i\Gamma_n h) \, E_n^{L-} \tag{7.5.11}$$

where the reflection and transmission coefficients are given by the respective expressions :

$$R_n = \frac{\Gamma_n - \chi_n}{\Gamma_n + \chi_n} \tag{7.5.12a}$$

and $$T_n = \frac{2\chi_n}{\Gamma_n + \chi_n} \quad . \tag{7.5.12b}$$

The symbol χ_n denotes the direction cosine of the n^{th} order plane wave in free space and is given by

$$\chi_n = \sqrt{k_0^2 - \alpha_n^2} \quad . \tag{7.5.12c}$$

The solution of equations (7.5.7), (7.5.9), (7.5.10) and (7.5.11) then enables the inward going plane wave coefficients $\{E_n^{U-}\}$ and $\{E_n^{L+}\}$ to be found in terms of the coefficients $\{a_n^+\}$ and $\{a_n^-\}$ and thus as functions of field quantities on the surface Γ. These representations are given by the following expressions:

$$E_n^{U-} = \frac{T_0}{1 - R_0^2 \exp(2i\Gamma_0 h)} \delta_{n0} + \frac{R_n}{1 - R_n^2 \exp(2i\Gamma_n h)} a_n^+ + \frac{R_n^2 \exp(2i\Gamma_n h)}{1 - R_n^2 \exp(2i\Gamma_n h)} a_n^- \qquad (7.5.13a)$$

and

$$E_n^{L+} = \frac{R_0 T_0 \exp(2i\Gamma_0 h)}{1 - R_0^2 \exp(2i\Gamma_0 h)} \delta_{n0} + \frac{R_n \exp(2i\Gamma_n h)}{1 - R_n^2 \exp(2i\Gamma_n h)} a_n^- + \frac{R_n^2 \exp(2i\Gamma_n h)}{1 - R_n^2 \exp(2i\Gamma_n h)} a_n^+ \qquad (7.5.13b)$$

It is interesting to consider the physical interpretation of the denominator $[1 - R_n^2 \exp(2i\Gamma_n h)]$. This term can be written as the sum of the following infinite geometric series,

$$\text{i.e.} \quad \frac{1}{1 - R_n^2 \exp(2i\Gamma_n h)} = \sum_{p=0}^{\infty} [R_n \exp(i\Gamma_n h)]^{2p} .$$

Such a series would arise if one were to adopt a multiple-scattering approach to the problem.

Now, with the aid of expressions (7.5.13), (7.5.8), (7.5.6) and (7.5.4) it is possible to express the electric field at any point P completely enclosed within C as a function of field quantities directly associated with the cylinder Γ

$$\begin{aligned} \text{i.e.} \quad E(P) = {} & \frac{T_0}{1 - R_0^2 \exp(2i\Gamma_0 h)} \exp(i\alpha_0 x - i\Gamma_0 y) \\ & + \frac{R_0 T_0 \exp(2i\Gamma_0 h)}{1 - R_0^2 \exp(2i\Gamma_0 h)} \exp(i\alpha_0 x + i\Gamma_0 y) \\ & + \int_{\Gamma_{ext}} [G_1^T(P;M') \frac{\partial E(M')}{\partial n'} - \frac{\partial G_1^T(P;M')}{\partial n'} E(M')] \, ds' \end{aligned} \qquad (7.5.14a)$$

where the Green's function G_1^T is given by

$$G_1^T = G_1 + G_1^A + G_1^B + G_1^C + G_1^D \tag{7.5.14b}$$

where $$G_1^A = \frac{1}{2id} \sum_n \frac{1}{\Gamma_n} \frac{R_n}{1 - R_n^2 \exp(2i\Gamma_n h)} \exp[i\alpha_n(x-x') + i\Gamma_n(-y-y')] \tag{7.5.14c}$$

$$G_1^B = \frac{1}{2id} \sum_n \frac{1}{\Gamma_n} \frac{R_n^2}{1 - R_n^2 \exp(2i\Gamma_n h)} \exp[i\alpha_n(x-x') + i\Gamma_n(-y+y'+2h)] \tag{7.5.14d}$$

$$G_1^C = \frac{1}{2id} \sum_n \frac{1}{\Gamma_n} \frac{R_n}{1 - R_n^2 \exp(2i\Gamma_n h)} \exp[i\alpha_n(x-x') + i\Gamma_n(y+y'+2h)] \tag{7.5.14e}$$

and $$G_1^D = \frac{1}{2id} \sum_n \frac{1}{\Gamma_n} \frac{R_n^2}{1 - R_n^2 \exp(2i\Gamma_n h)} \exp[i\alpha_n(x-x') + i\Gamma_n(y-y'+2h)]. \tag{7.5.14f}$$

It is to be noted that the effective incident field in equation (7.5.14a) is composed of two terms, an upward and a downward going wave. The downward going field arises from a transmission of the original incident field in free space through the boundary S_u whereas the upward going field results from a reflection of this transmitted component at the lower interface S_ℓ. For clarity, these two terms will be grouped and referred to by the symbol E^i_{eff} - the effective incident field.

$$\text{i.e.} \quad E(P) = E^i_{eff}(P) + \int_{\Gamma_{ext}} [G_1^T \frac{\partial E}{\partial n'} - \frac{\partial G_1^T}{\partial n'} E] \, ds' \, .$$

Consider now the properties of the kernel G_1^T. By inspection, it may be seen that none of the functions G_1^A, G_1^B, G_1^C or G_1^D possess any singularities within the region D_1 and that all are absolutely convergent series. Thus, one infers that the singular behaviour of G_1^T is characterized purely by the singularities of G_1.

$$\text{i.e.} \quad (\nabla^2 + k_1^2)\, G_1^T = (\nabla^2 + k_1^2)\, G_1 = \sum_{n=-\infty}^{\infty} \delta(x-x'-nd)\, \exp(i\alpha_0 nd)\, \delta(y-y').$$

By now sampling the field at a point M on Γ it follows that

$$\tfrac{1}{2}E(M) = E^i_{eff}(M) + \int_{\Gamma_{ext}} [G_1^T(M;M') \frac{\partial E\,(M')}{\partial n'} - \frac{\partial G_1^T(M;M')}{\partial n'} E(M')]\, ds', \quad (7.5.15)$$

a consequence of the fact that only "half a singularity" lies within the domain of integration. This expression is the desired integral equation uniquely characterizing the diffraction problem.

The equation requires a knowledge of both the field and its normal derivative on the cylinder surface and in this sense the problem is underdetermined. This situation may be alleviated however, by once again turning to Maystre's technique (referenced in previous sections) and defining the existence of a "pseudo-current" on the inner surface of Γ. This quantity is capable of linking both E and $\frac{\partial E}{\partial n}$ and in doing so provides a further equation of constraint on the physical problem. By introducing a function v specified by the following criteria :

$$\left.\begin{array}{lll} \text{(i)} & v(Q) = E(Q) & \text{if } Q \in D_2, \\ \text{(ii)} & (\nabla^2 + k_2^2)\, v = 0 & \text{elsewhere,} \\ \text{(iii)} & v \text{ is everywhere continuous and obeys the radiation conditions as } |y| \to \infty, & \end{array}\right\} \quad (7.5.16)$$

the "pseudo-current", ξ, may be defined by

$$\xi(M) = \left.\frac{\partial v}{\partial n}\right|_{\Gamma_{ext}} - \left.\frac{\partial v}{\partial n}\right|_{\Gamma_{int}} . \qquad (7.5.17)$$

Following an analysis similar to that discussed in earlier sections of this chapter, it may be shown that

$$E(Q) = \int_{\Gamma} G_2(Q;M')\, \xi(M')\, ds' \quad \text{for} \quad Q \in D_2 \qquad (7.5.18)$$

where

$$G_2(Q;M') = \frac{1}{2id} \sum_{n=-\infty}^{\infty} \frac{1}{\eta_n} \exp[i\alpha_n(x-x') + i\eta_n|y-y'|]$$

where

$$\eta_n = \sqrt{k_2^2 - \alpha_n^2} . \qquad (7.5.19)$$

Furthermore, for a point M on Γ, it follows that

$$E(M) = \int_{\Gamma} G_2(M;M')\, \xi(M')\, ds' \qquad (7.5.20a)$$

and

$$\left.\frac{\partial E(M)}{\partial n}\right|_{\Gamma_{int}} = -\tfrac{1}{2}\xi(M) + \int_{\Gamma} \frac{\partial G_2(M;M')}{\partial n}\, \xi(M')\, ds' \qquad (7.5.20b)$$

and it is these representations together with the boundary conditions that yield a completely and uniquely determined solution of the diffraction problem.

It only remains to substitute these representations into expression (7.5.15) to derive the final equation:

$$\int_{\Gamma} \tau^{E}(M;M')\, \xi(M')\, ds' = E^{i}_{eff}(M) \qquad (7.5.21a)$$

where $\tau^{E}(M;M') = \tfrac{1}{2}\, G_2(M;M') + \tfrac{1}{2}\, G_1^{T}(M;M')$

$$- \int_{\Gamma} \left[G_1^{T}(M;M'') \frac{\partial G_2(M'';M')}{\partial n''} - \frac{\partial G_1^{T}(M;M'')}{\partial n''} G_2(M'';M') \right] ds'' \qquad (7.5.21b)$$

- a Fredholm integral equation of the first kind which is solved using standard techniques.

Reflected and transmitted field amplitudes in free space may be reconstructed by performing the following sequence of calculations:

(a) reconstruction of E and $\frac{\partial E}{\partial n}$ on Γ using equations (7.5.20),

(b) determination of the cylinder radiation coefficients $\{a_n^{+}\}$ and $\{a_n^{-}\}$ from equations (7.5.6) and (7.5.8)

and

(c) computation of the coefficient set $\{E_n^{L-}\}$ and $\{E_n^{U+}\}$ from which can be obtained the reflected and transmitted field amplitudes after a further application of Fresnel's Laws.

7.5.4 Application to the case of a generalized lamellar grating

The previous theory can be now specifically applied to the case of a generalized lamellar grating. As can be seen from figure (7.5.2) the structure possesses many of the characteristics of a lamellar grating but is far more general in the sense that

(i) the geometry of the groove is no longer restricted to having a rectangular cross-section

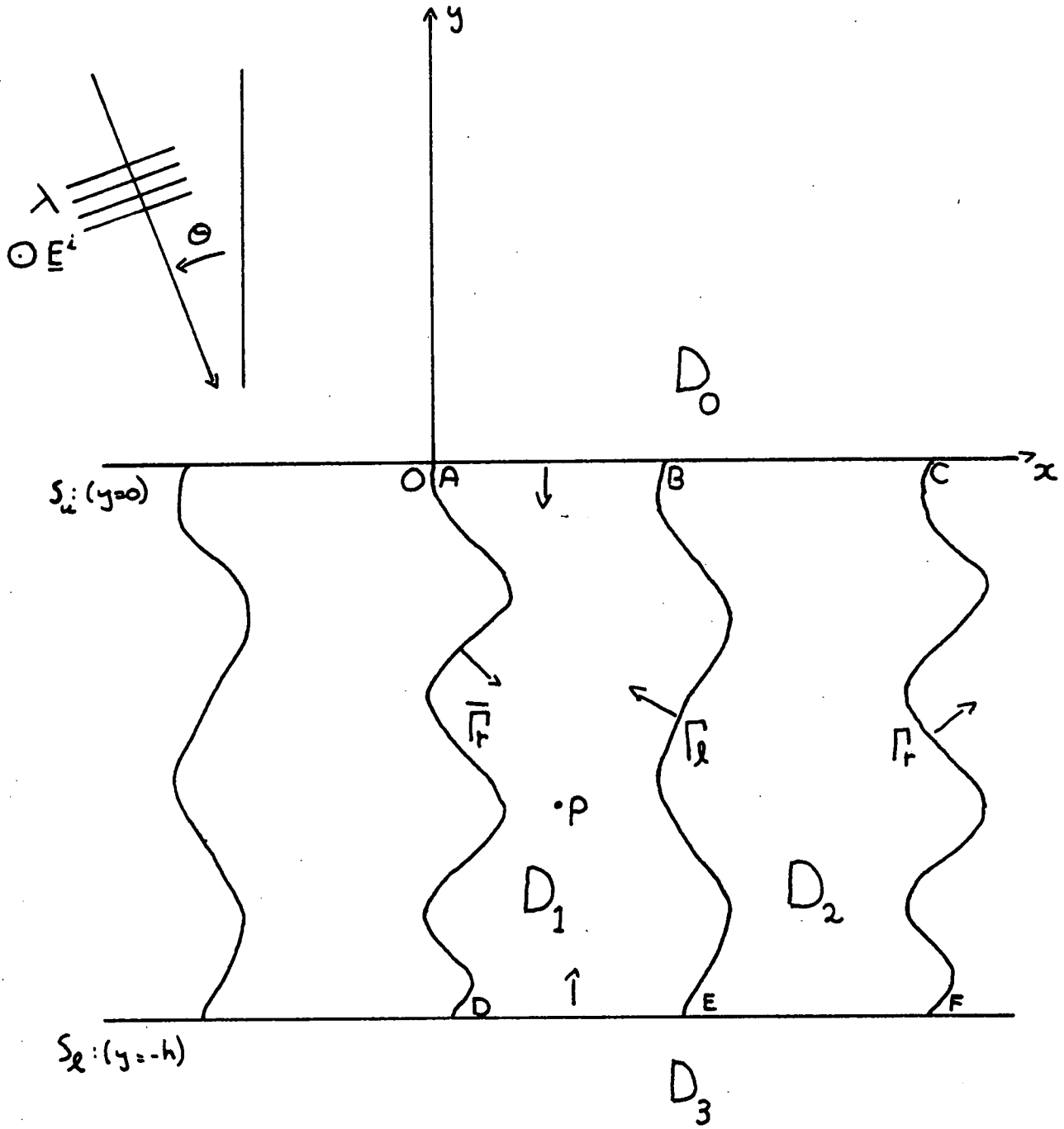


Figure 7.5.2

The geometry of the generalized lamellar grating.

and

(ii) the constraint of requiring a perfectly conducting segment in the groove profile has been removed.

Both of these assumptions were necessary to previous modal formulations of this problem.

However, the formalism outlined in the previous section cannot be directly applied to this case without some modification to its derivation. The difficulty arises in attempting to identify the existence of plane waves in the space D_1 when the cylinder, D_2, is such that it completely fills the space $y \in [-h,0]$.

In an attempt to alleviate this dilemma, consider the field representation for the electric field in D_1 given by the Kirchhoff-Helmholtz formula:

$$\begin{aligned} E(P) &= \int_{C_1} [G_1(P;M') \frac{\partial E(M')}{\partial n'} - \frac{\partial G_1(P;M')}{\partial n'} E(M')] \, ds' \\ &= (\int_{\overline{\Gamma}_r} + \int_{\Gamma_\ell}) [G_1 \frac{\partial E}{\partial n'} - \frac{\partial G_1}{\partial n'} E] \Big|_{ext} ds' \\ &\quad - \int_{AB} [G_1 \frac{\partial E}{\partial y'} - \frac{\partial G_1}{\partial y'} E] \, dx' \\ &\quad + \int_{DE} [G_1 \frac{\partial E}{\partial y'} - \frac{\partial G_1}{\partial y'} E] \, dx'. \end{aligned} \tag{7.5.22}$$

From field pseudo-periodicity it follows that

$$[G_1 \frac{\partial E}{\partial n'} - \frac{\partial G_1}{\partial n'} E]\Big|_{\Gamma_r} = [G_1 \frac{\partial E}{\partial n'} - \frac{\partial G_1}{\partial n'} E]\Big|_{\overline{\Gamma}_r}$$

and after some elementary manipulation equation (7.5.22) reduces to the form:

$$E(P) = \int_{\Gamma_{ext}} [G_1 \frac{\partial E}{\partial n'} - \frac{\partial G_1}{\partial n'} E] \, ds'$$
$$- \int_0^d [G_1 \frac{\partial E}{\partial y'} - \frac{\partial G_1}{\partial y'} E]\Big|_{y'=0^-} dx'$$
$$+ \int_0^d [G_1 \frac{\partial E}{\partial y'} - \frac{\partial G_1}{\partial y'} E]\Big|_{y'=-h^+} dx' \qquad (7.5.23)$$

where $\Gamma = \Gamma_\ell \cup BC \cup \Gamma_r \cup FE$.

This expression takes a form identical to equations (7.5.4) derived in the previous section.

Now, to avoid any explicit reference to plane wave expansions it is necessary to write the field and its normal derivative on the plane interfaces in terms of the following series:

$$E(x,0^-) = \sum_n f_n^U \exp(i\alpha_n x) \qquad (7.5.24a)$$

$$\frac{\partial E}{\partial y}(x,0^-) = \sum_n d_n^U \exp(i\alpha_n x) \qquad (7.5.24b)$$

$$E(x,-h^+) = \sum_n f_n^L \exp(i\alpha_n x) \qquad (7.5.24c)$$

and

$$\frac{\partial E}{\partial y}(x,-h^+) = \sum_n d_n^L \exp(i\alpha_n x) \qquad (7.5.24d)$$

where

$$\left.\begin{aligned} d_n^U &= -2i\,\chi_0\,\delta_{n,0} + i\,\chi_n\,f_n^U \\ \text{and}\quad d_n^L &= -i\,\chi_n\,f_n^L\ . \end{aligned}\right\} \qquad (7.5.25)$$

Now, by using these definitions the electric field at any point P in D_1 (such that P does not lie on the boundary of D_1) is given by:

$$\begin{aligned} E(P) = &\int_{\Gamma_{ext}} [G_1 \frac{\partial E}{\partial n'} - \frac{\partial G_1}{\partial n'} E] \quad ds' \\ &+ \sum_n E_n^{U-} \exp[i(\alpha_n x - \Gamma_n y)] \\ &+ \sum_n E_n^{L+} \exp[i(\alpha_n x + \Gamma_n y)] \end{aligned} \qquad (7.5.26a)$$

where

$$E_n^{U-} = \frac{\chi_0}{\Gamma_0}\,\delta_{n,0} + \tfrac{1}{2}(1 - \frac{\chi_n}{\Gamma_n})\,f_n^U \qquad (7.5.26b)$$

and

$$E_n^{L+} = \tfrac{1}{2}(1 - \frac{\chi_n}{\Gamma_n})\,f_n^L \exp(i\Gamma_n h)\ . \qquad (7.5.26c)$$

Should the point P now tend to a point M^U on AC, equation (7.5.23) can be re-written in the form:

$$\tfrac{1}{2}\,E(M^U) = (\int_{\Gamma_{ext}} + \int_{AC} + \int_{DF}) [G_1 \frac{\partial E}{\partial n'} - \frac{\partial G_1}{\partial n'} E] \quad ds'\ . \qquad (7.5.27)$$

As before, the contribution from the line integral around the contour Γ yields a series of upward going plane waves. Similar remarks apply to the contributions from the linear segment DF. After some elementary

but tedious manipulation of equation (7.5.27) the following constraint is derived

$$\tfrac{1}{2} f_n^U = a_n^+ + E_n^{L+} + \frac{i\, d_n^U}{2\,\Gamma_n}$$

or

$$f_n^U = T_0\, \delta_{n0} + \frac{2\, a_n^+}{1 + \chi_n/\Gamma_n} + R_n\, f_n^L \exp(i\Gamma_n h). \qquad (7.5.28)$$

A similar line of argument, derived by considering a point M^L on DF, yields

$$f_n^L = T_0 \exp(i\Gamma_0 h)\, \delta_{n0} + \frac{2\, a_n^-}{1 + \chi_n/\Gamma_n} \exp(i\Gamma_n h) + R_n \exp(i\Gamma_n h)\, f_n^U \qquad (7.5.29)$$

It then follows from these equations that

$$E_n^{U-} = \frac{T_0}{1 - R_0^2 \exp(2i\Gamma_0 h)}\, \delta_{n0} + \frac{R_n}{1 - R_n^2 \exp(2i\Gamma_n h)}\, a_n^+ + \frac{R_n^2 \exp(2i\Gamma_n h)}{1 - R_n^2 \exp(2i\Gamma_n h)}\, a_n^-$$

which is an identical expression to that derived in the previous section. Thus, it has been established that the earlier formulation derived for the pure cylinder problem is entirely valid when applied to the case of a generalized lamellar grating.

7.6 A FULLY GENERAL BI-METALLIC GRATING THEORY

7.6.1 Introductory remarks

The material presented in this section represents the final stage in the evolution of a fully generalized bi-metallic grating theory. Previous sections have outlined a number of formalisms which have been concerned with somewhat simplified diffracting structures (with underlying assumptions being made concerning both geometry and groove conductivity). However this formulation, relying upon the linear operator approach adopted by Maystre in his latest work [7.9], has eliminated all such restrictions and is in fact the realization of the ultimate goal.

Although at the time of preparation of this thesis it had not been possible to implement the theory numerically, and thus test the formulation against such constraints as reciprocity and energy conservation, it is, nevertheless, felt that the formalism rests on a firm basis. On the surface, this may appear to be a bold remark, but in view of the experience gained from the development of previous theories, together with the fact that it has been possible to analytically test certain aspects of the formalism, one feels that a measure of confidence in the theory is justified.

However, the study is still incomplete and it is not possible to comment further until such time as the associated computer program has been written and thoroughly tested. It is felt that such an effort is warranted and indeed stimulated by the promising results given in section (7.3). Clearly, there is much scope for future work and it is hoped to report some interesting results at a later stage.

7.6.2 Solution of the diffraction problem

Consider a P polarized wave field incident upon the structure shown in figure (7.6.1). This geometry is a generalization of that discussed in the previous section. Here the interfaces Γ_0 and Γ_2 bounding the semi-infinite media D_0 and D_2 can assume arbitrary profile shapes. Once again, the inserted metallic element of this structure has been chosen to be the cylinder D_1. However this represents no restriction to the formulation of a completely general bi-metallic grating theory as evidenced by the treatment of section (7.5.4).

In seeking to uniquely characterize the field distribution, it is once again necessary to apply the powerful technique developed by Maystre [7.2] and define the existence of "pseudo-currents" $\{\phi_j\}$ on each of the interfaces Γ_j for $j \in [0,2]$. This is accomplished by defining a set of functions $\{u_j\}$ by

$$\left.\begin{aligned} &\text{(i)} \quad u_j(P_j) = E(P_j) - E^i(P_j)\,\delta_{j,0} \quad \text{if} \quad P_j \in D_j \\ &\text{(ii)} \quad (\nabla^2 + k_j^2)\,u_j = 0 \quad \text{if} \quad P_j \notin D_j \\ &\text{(iii)} \quad u_j \text{ is everywhere continuous and obeys the radiation conditions.} \end{aligned}\right\} \qquad (7.6.1)$$

By performing analyses identical to those presented elsewhere in this thesis, it follows that the electric field at any point P_j within the medium D_j is given by

$$E(P_j) = E^i(P_j)\delta_{j0} + \int_{\Gamma_j} G_j(P_j;M_j')\,\phi_j(M_j')\,ds' \qquad (7.6.2a)$$

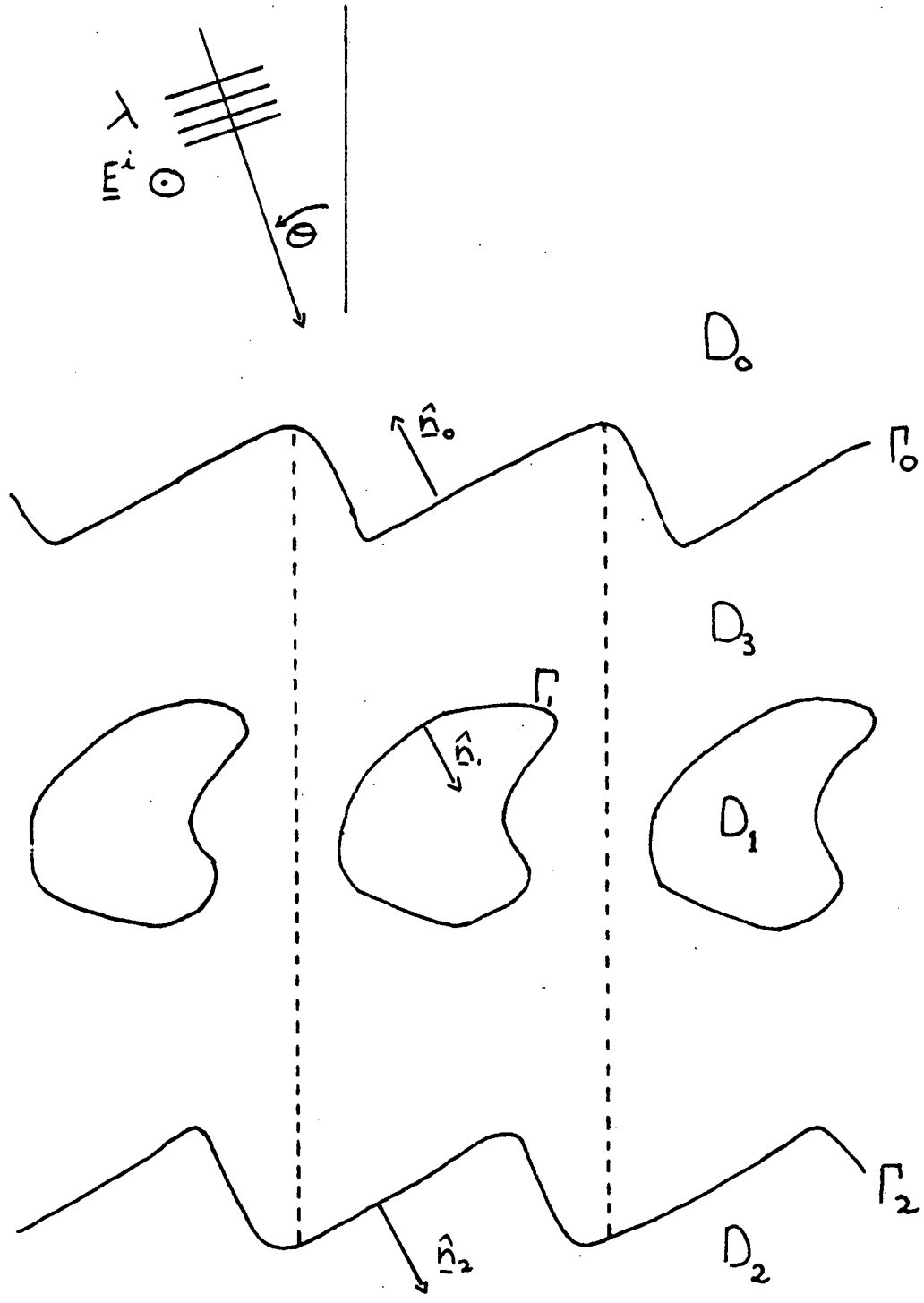


Figure 7.6.1

The geometry of the fully general bi-metallic grating.

where the "pseudo-current" ϕ_j appropriate to the interface Γ_j (bounding D_j) is defined by

$$\phi_j(M_j) = \left.\frac{\partial u_j}{\partial n}\right|_{\Gamma_j,\text{int}} - \left.\frac{\partial u_j}{\partial n}\right|_{\Gamma_j,\text{ext}} . \qquad (7.6.2b)$$

(The terms "int" and "ext" refer to the derivative being taken at points just interior and exterior to the medium D_j respectively. The unit normals $\hat{\underline{n}}_j$ to the surface Γ_j are inwardly directed towards the medium D_j.)

The field is sampled by a Green's function G_j (appropriate to the medium D_j) whose form is identical to those used in previous sections

$$\text{i.e.} \quad G_j(P_j;M_j') = \frac{1}{2id}\sum_n \frac{1}{\eta_{j,n}} \exp[i\alpha_n(x-x') + i\eta_{j,n}|y-y'|] \qquad (7.6.3)$$

$$\text{where} \quad \eta_{j,n} = \sqrt{k_j^2 - \alpha_n^2} \quad \text{and} \quad k_j = \frac{2\pi}{\lambda} r_j .$$

(Here r_j is the refractive index of the medium D_j.)

Now, for the particular case of a point M_j lying on the surface Γ_j, it may be shown that

$$E(M_j) = E^i(M_j)\,\delta_{j,0} + \int_{\Gamma_j} G_j(M_j;M_j')\,\phi_j(M_j')\,ds' \qquad (7.6.4a)$$

and

$$\left.\frac{\partial E(M_j)}{\partial n}\right|_{\Gamma_j,\text{int}} = \frac{\partial E^i(M_j)}{\partial n}\,\delta_{j,0} + \tfrac{1}{2}\,\phi_j(M_j) + \int_{\Gamma_j} \frac{\partial G_j(M_j;M_j')}{\partial n}\,\phi_j(M_j')\,ds' . \qquad (7.6.4b)$$

By using the Kirchhoff-Helmholtz formula, the electric field at some

point Q, completely enclosed within D_3, may be expressed in terms of these quantities.

$$\text{i.e.} \quad E(Q) = \sum_{j=0}^{2} \int_{\Gamma_{j,ext}} \left[\frac{\partial G_3(Q;M_j')}{\partial n'} E(M_j') - G_3(Q;M_j') \frac{\partial E(M_j')}{\partial n'}\right] ds' . \quad (7.6.5)$$

Now for a point M_j lying on the interface $\Gamma_{j,ext}$ (i.e. just interior to the domain D_3)

$$\tfrac{1}{2} E(M_i) = \sum_{j=0}^{2} \int_{\Gamma_{j,ext}} \left[\frac{\partial G_3(M_i;M_j')}{\partial n'} E(M_j') - G_3(M_i;M_j') \frac{\partial E(M_j')}{\partial n'}\right] ds'$$

$$\forall \; i \in [0,2]. \quad (7.6.6)$$

Equation (7.6.6) together with the field continuity conditions links the unknowns $E\big|_{\Gamma_j}$ and $\frac{\partial E}{\partial n}\big|_{\Gamma_j}$ and thus gives rise to a set of three coupled integral equations in the $\{\phi_j\}$

$$\text{i.e.} \quad A_{00}\,\phi_0 + A_{01}\,\phi_1 + A_{02}\,\phi_2 = R_0 \quad (7.6.7a)$$

$$A_{10}\,\phi_0 + A_{11}\,\phi_1 + A_{12}\,\phi_2 = R_1 \quad (7.6.7b)$$

$$A_{20}\,\phi_0 + A_{21}\,\phi_1 + A_{22}\,\phi_2 = R_2 \,. \quad (7.6.7c)$$

Here, the linear operator A_{ij} is defined by

$$A_{ij}\,\phi_j = \int_{\Gamma_j} A_{ij}(M_i;M_j')\,\phi_j(M_j')\,ds'$$

and the kernel is given by

$$A_{ij}(M_i;M'_j) = -\tfrac{1}{2} G_3(M_i;M'_j) - \tfrac{1}{2} G_i(M_i;M'_i)\,\delta_{ij}$$

$$+ \int_{\Gamma_j} \left[\frac{\partial G_3(M_i;M''_j)}{\partial n''_j} G_j(M''_j;M'_j) - G_3(M_i;M''_j) \frac{\partial G_j(M''_j;M'_j)}{\partial n''_j}\right] ds'' \,. \tag{7.6.8}$$

The right hand sides of equations (7.6.7) are given by

$$R_i(M_i) = \tfrac{1}{2} E^i(M_0)\,\delta_{i0}$$

$$- \int_{\Gamma_0} \left[\frac{\partial G_3(M_i;M'_0)}{\partial n'} E^i(M'_0) - G_3(M_i;M'_0) \frac{\partial E^i(M'_0)}{\partial n'}\right] ds' \,. \tag{7.6.9}$$

7.6.3 Reduction of the coupled integral equations

The above set of coupled integral equations may, of course, be solved using the points matching technique. However it is more elegant and perhaps computationally more efficient to reduce them analytically so as to involve the solution of only one unknown. In accordance with the analysis presented in section (7.5) it was decided to reduce these equations by eliminating unknowns at the surfaces Γ_0 and Γ_2. This decision was made so that the results of this formalism could be directly compared with the buried cylinder grating theory. As such, comments regarding the choice of reduction mechanism are interspersed throughout the following analysis.

Before commencing the solution of the operator equation, it is worthwhile to discuss the properties of these operators. It may be shown that kernels of the form A_{ii} possess logarithmic singularities which may be removed by integration. On the other hand, kernels of the

form A_{ij} such that $i \neq j$ can be shown to be absolutely convergent.

At this stage, a question arises concerning which of the operators can be satisfactorily inverted in order to explicitly isolate the functions ϕ_0 and ϕ_2. Any analysis involving the inversion of absolutely convergent operators will give rise to operators whose kernels exhibit divergent behaviour. Although this represents no hindrance to a successful analytic reduction (possible in the case of planar upper and lower interfaces), one must reject any numerical implementation of this treatment on the grounds of gross instability. Thus, the only reasonable approach involves the inversion of operators of the type A_{ii}

$$\text{e.g.} \quad \phi_0 = A_{00}^{-1} R_0 - A_{00}^{-1} A_{01} \phi_1 - A_{00}^{-1} A_{02} \phi_2. \tag{7.6.10}$$

On substituting this expression for ϕ_0 into equations (7.6.7b) and (7.6.7c), it follows that

$$\left.\begin{aligned} B_1 \phi_1 + B_2 \phi_2 &= S \\ \text{and} \quad C_1 \phi_1 + C_2 \phi_2 &= T \end{aligned}\right\} \tag{7.6.11}$$

$$\left.\begin{aligned} \text{where } B_1 &= A_{11} - A_{10} A_{00}^{-1} A_{01} & C_1 &= A_{21} - A_{20} A_{00}^{-1} A_{01} \\ B_2 &= A_{12} - A_{10} A_{00}^{-1} A_{02} & C_2 &= A_{22} - A_{20} A_{00}^{-1} A_{02} \\ S &= R_1 - A_{10} A_{00}^{-1} R_0 & T &= R_2 - A_{20} A_{00}^{-1} R_0 \end{aligned}\right\} \tag{7.6.12}$$

It only remains to eliminate the variable ϕ_2. Although this can be done by inverting either of the operators B_2 or C_2, the choice is not arbitrary. Because of the absolute convergence of the operator B_2,

it is preferable to invert the logarithmically-divergent operator C_2.

Thus,

$$\phi_2 = C_2^{-1} T - C_2^{-1} C_1 \phi_1$$

and so,

$$(B_1 - B_2 C_2^{-1} C_1) \phi_1 = S - B_2 C_2^{-1} T . \qquad (7.6.13)$$

7.6.4 Agreement with the buried cylinder grating formalism

Since it had not been possible to numerically implement the above formalism by the submission date of the thesis, it was of paramount importance to find some criterion against which the formalism could be meaningfully tested. The only available check having any significance was a comparison of this theory and the treatment of the buried cylinder grating outlined in section (7.5). Thus, it is the purpose of this section to analytically evaluate the operator equation (7.6.13) and derive an expression in agreement with equations (7.5.21).

As before, the interfaces Γ_0 and Γ_2 are represented by the planes $y = 0$ and $y = -h$ respectively and so

$$\left.\begin{aligned} A_{00} &= -\frac{1}{4id} \sum_n \left(\frac{1}{\eta_{3,n}} + \frac{1}{\eta_{0,n}}\right) \exp[i\alpha_n (x-x')] \\ A_{02} &= -\frac{1}{4id} \sum_n \left(\frac{1}{\eta_{3,n}} - \frac{1}{\eta_{2,n}}\right) \exp[i\alpha_n (x-x') + i\eta_{3,n} h] \\ A_{10} &= -\frac{1}{4id} \sum_n \left(\frac{1}{\eta_{3,n}} - \frac{1}{\eta_{0,n}}\right) \exp[i\alpha_n (x-x') - i\eta_{3,n} f(x)] \end{aligned}\right\} \qquad (7.6.14)$$

continued page

$$A_{12} = -\frac{1}{4id}\sum_n \left(\frac{1}{\eta_{3,n}} - \frac{1}{\eta_{2,n}}\right) \exp[i\alpha_n(x-x') + i\eta_{3,n}(h + f(x))]$$

$$A_{20} = -\frac{1}{4id}\sum_n \left(\frac{1}{\eta_{3,n}} - \frac{1}{\eta_{0,n}}\right) \exp[i\alpha_n(x-x') + i\eta_{3,n}\, h] \qquad (7.6.14)$$

$$A_{22} = -\frac{1}{4id}\sum_n \left(\frac{1}{\eta_{3,n}} + \frac{1}{\eta_{2,n}}\right) \exp[i\alpha_n(x-x')]$$

with

$$R_0 = \tfrac{1}{2}\left(1 - \frac{\eta_{0,n}}{\eta_{3,n}}\right) \exp(i\alpha_0 x)$$

$$R_1 = -\tfrac{1}{2}\left(1 + \frac{\eta_{0,n}}{\eta_{3,n}}\right) \exp[i\alpha_0 x - i\eta_{3,0}\, f(x)] \qquad (7.6.15)$$

and

$$R_2 = -\tfrac{1}{2}\left(1 + \frac{\eta_{0,n}}{\eta_{3,n}}\right) \exp[i\alpha_0 x + i\eta_{3,0}\, h].$$

Now, in the space [0,d], on which is defined an orthogonal basis $\{\exp(i\alpha_n x)\}$ it is possible to establish the existence of an identity operator I such that for an operator O,

$$OO^{-1} = O^{-1}O = I.$$

Here,

$$I = \frac{1}{d}\sum_n \exp[i\alpha_n(x-x')]. \qquad (7.6.16)$$

Thus,

$$A_{00}^{-1} = -\frac{4i}{d}\sum_n \left(\frac{\eta_{3,n} \times \eta_{0,n}}{\eta_{3,n} + \eta_{0,n}}\right) \exp[i\alpha_n(x-x')]\ . \qquad (7.6.17)$$

After performing a vast quantity of tedious manipulation it follows that

$$
\left.
\begin{aligned}
S &= -t_{0,0}\exp[i\alpha_0 x - i\eta_{3,0} f(x)] \\
T &= -t_{0,0}\exp[i\alpha_0 x + i\eta_{3,0} h] \\
B_2 &= -\frac{1}{4id}\sum_n \left(\frac{1}{\eta_{3,n}} - \frac{1}{\eta_{2,n}}\right)\exp[i\alpha_n(x-x')] \\
&\quad \{\exp[i\eta_{3,n}(h+f(x))] + r_{0,n}\exp[i\eta_{3,n}(h-f(x))]\} \\
C_2 &= -\frac{1}{4id}\, n \left(\frac{1}{\eta_{3,n}} + \frac{1}{\eta_{2,n}}\right)\exp[i\alpha_n(x-x')] \\
&\quad \{1 - r_{0,n}\, r_{2,n}\exp(2i\eta_{3,n} h)\}
\end{aligned}
\right\} \qquad (7.6.18)
$$

and thus,

$$
\begin{aligned}
S - B_2 C_2^{-1} T = &\frac{-t_{0,0}}{1 - r_{2,0}\, r_{0,0}\exp(2i\eta_{3,0} h)}\exp[i\alpha_0 x - i\eta_{3,0} f(x)] \\
&\frac{-r_{2,0}\, t_{0,0}\exp(2i\eta_{3,0} h)}{1 - r_{2,0}\, r_{0,0}\exp(2i\eta_{3,0} h)}\exp[i\alpha_0 x + i\eta_{3,0} f(x)] \qquad (7.6.19)
\end{aligned}
$$

where the reflection and transmission factors for the n^{th} order wave at the plane boundary Γ_1 are respectively given by

$$
r_{m,n} = \frac{\eta_{3,n} - \eta_{m,n}}{\eta_{3,n} + \eta_{m,n}} \qquad (7.6.20a)
$$

and

$$
t_{m,n} = \frac{2\eta_{m,n}}{\eta_{3,n} + \eta_{m,n}} \,. \qquad (7.6.20b)
$$

Expression (7.6.19) is a measure of the effective incident field on the cylinder Γ_1. Inspection of this equation and the first two terms of

equation (7.5.14a) from the previous section reveals that they are indeed of similar form. Some further manipulation also demonstrates that the kernel τ^E defined by equation (7.5.21b) and the operator $(B_1 - B_2 C_2^{-1} C_1)$ are analogous quantities.

7.6.5 Concluding remarks

Although this section has concentrated on the derivation of a theory describing the diffraction properties of a finitely conducting bi-metallic cylinder grating, the treatment can be extended to cover the case of a fully general bi-metallic grating, for which the cylindrical element partially coincides with the upper and lower interfaces Γ_0 and Γ_2.

Although no rigorous testing of this formalism has been attempted using a computer implementation, the theoretical verification discussed in the previous sub-section can be considered as having some significance. However, until such time as a computer program has been written and its results scrutinized thoroughly, it will be necessary to regard this formalism with due caution.

It is only with hindsight that the true value of the analysis presented in section (7.6.4) becomes apparent. To suggest that the above reduction only represents a successful test of the formalism is to do the entire method a grave injustice. Much valuable experience has been gained from this exercise, particularly in highlighting numerical difficulties before they manifest themselves as obstinate discrepancies in the results of the associated computer implementation. At this stage, it is felt that the current version of the theory should present no significant numerical difficulties, although confirmation of this statement can result only from the actual implementation.

7.7 SUMMARY

The aim of this chapter has been two-fold in that it has attempted to

(i) re-create the evolution of the bi-metallic grating theories and

(ii) outline the method of anomaly suppression.

In respect of the second of these aims, anomaly suppression can indeed be accomplished by overcoating the "off-blaze" facet with a weakly conducting material. The mechanism appears to involve the creation of "lossy plasmons" whose bandwidth is sufficiently broad to smear any rapid spectral response. The results obtained from the numerical experiments, using a relatively restricted formalism, indicate the need to enforce both upper and lower bounds on the conductivity of the overcoated film.

Clearly, on the basis of these results, there exists much scope for future theoretical work having great relevance to grating designers. In the near future, it is hoped to implement the theory of section (7.6) in an attempt to glean a more thorough understanding of the diffraction properties of this fascinating class of structures.

REFERENCES

[7.1] K. Sugahara, T. Kita and T. Shimotakahara, Proc. Sym. on Diffraction Gratings and Grating Instruments, Spec. Soc. Japan, Tokyo (1974) 21.

[7.2] D. Maystre, Thèse No. AO 9545, (1974), L'Université d'Aix-Marseille III, France.

[7.3] A. Wirgin, C. R. Acad. Sci. Paris, 262B (1966) 870.

[7.4] J. Strong, Phys. Rev., 49 (1936) 291.

[7.5] C. E. Wheeler, E. T. Arakawa and R. H. Ritchie, (1976), Oak Ridge National Laboratory, Report ORNL/TM-5185.

[7.6] A. Maréchal and G. W. Stoke, C. R. Acad. Sci. Paris, 249 (1959) 2047.

[7.7] R. C. McPhedran and M. D. Waterworth, Optica Acta, 20 (1973) 177.

[7.8] J. R. Andrewartha, J. R. Fox and I. J. Wilson, Optica Acta, to appear.

[7.9] D. Maystre, J. Opt. Soc. Amer., to appear.

CHAPTER 8

THE CROSSED LAMELLAR TRANSMISSION GRATING

The investigations which are reported in this chapter were conducted in collaboration with Miss Jenny Adams and Dr. Ross McPhedran, and it is with great pleasure that the author warmly acknowledges their contributions. The material which is discussed herein has been taken from a paper [8.1], accepted for publication by the "Journal of Optics", and a University of Tasmania internal research report [8.2].

8.1 INTRODUCTION

This chapter, and also the following chapter, are concerned with the diffraction properties of a class of structures referred to, within our group, as "bi-gratings". These consist of a pair of spatially separated singly periodic lamellar transmission gratings whose axes of periodicity are inclined with respect to one another at some arbitrary angle.

The literature of crossed and spatially separated gratings dates back to the year 1939 when the experimental work of Esau et al [8.3] and the theoretical work of Wessel [8.4] were published. Wessel chose to characterize the behaviour of the grating in terms of an impedance which he defined as being the ratio of the electric field (parallel to the wires) to the current flowing in one of the wires. His theory assumed that the wires were composed of a perfectly conducting metal and had a circular cross-section whose radius was much smaller than the wavelength of the incident radiation. This restriction was enforced so

that there would be no spatial variation of the current distribution across the wire, an assumption which greatly simplified the analysis and permitted a solution of the problem in closed form. (Furthermore, his theory was valid only for wavelengths in excess of twice the grating period, i.e. only a single propagating order.)

In 1949, Franz [8.5] confronted the experiments of Esau et al. (concerned with the diffraction properties of a structure made from several parallel grids) with a simplified version of Wessel's method and was able to obtain reasonable agreement with the experimental data.

Lewis and Casey [8.6], in 1951, discussed the application of a pair of parallel grids in a microwave interference filter. A further paper [8.7] by these same authors demonstrated that the use of double grids, (or a pair of aligned wire gratings used in P polarized radiation), in long wavelength Fabry-Perot interferometers was preferable to the traditional design employing highly conducting thin films, (for which the transmission properties were particularly poor). However, their theory was also afflicted with the same restrictions as those applying to Wessel's formalism. In particular, the need to assume the wavelength to be far greater than the dimensions of the individual wires greatly diminished the generality of their theory. Although this assumption permitted the derivation of an analytic solution, it restricted the theory to interferometers of low finesse - a consequence of the reduced reflectance of each array.

Following upon these early studies, there have been numerous other investigations on both the theoretical and experimental fronts. Almost invariably however, the theories have relied upon impedance techniques,

the application of which must be open to debate. Nevertheless, even though the theoretical development had stagnated, many valuable experimental investigations were published. Notable contributions were made by such authors as Mitsuishi et al, Ulrich, Renk and Genzel (references [8.8-13]).

The first departure from the use of impedance theories was the work of Adodina et al [8.14-15] who adopted a "multiple-scattering" approach to the problem. However, their theory was also plagued by a number of restrictions. In particular, it assumed that the only mechanism for interaction between the two arrays was a single propagating order. Even for very long wavelengths, this involved the restriction of the analysis to relatively large array separations in order to prevent any evanescent field coupling of the two arrays. However, within the terms of reference laid down by the design of an interferometer, it was a justifiable assumption.

Since these early investigations, the greatly increased understanding of diffraction problems together with the advent of powerful digital computers (thereby removing the need to derive solutions in closed form) has enabled the theoretician to arrive at fully rigorous formalisms characterizing such problems. In recent years, the theory of doubly periodic structures has been predominantly studied by electrical engineers concerned with antenna systems. However, the extensive use of inductive grids as interference filters and interferometers and a recent suggestion by Horwitz [8.16] that such structures could provide an effective solar selective surface have motivated many of the current optical investigations.

These optical studies follow closely upon the success of rigorous electromagnetic theories describing the diffraction of a plane wave by a singly periodic grating. During the past decade, the development of grating theories has largely followed two basic approaches, namely the integral and differential formalisms. Although these methods have the advantage that they can accommodate a diversity of diffracting geometries, they are, unfortunately, totally unsuited to the study of doubly periodic structures. It is the need for immense computer resources, which are currently unavailable, instead of any physical restriction that limits their application.

In consequence, it has been necessary to develop a further class of formalisms referred to as the "Modal Methods", whose roots are deeply entrenched in waveguide theory. Such techniques rely upon the choice of appropriate waveguide modes to specify the fields within the apertures and Rayleigh expansions to characterize the field in the semi-infinite spaces outside the apertures. To date, however, it has been possible to apply modal formalisms only to relatively simple aperture geometries, for which analytic expressions of the field modes can be derived.

In the early years of this decade, the great bulk of the literature pertaining to inductive grids (made by perforating a metallic sheet with a periodic array of apertures) was contributed by Chen [8.17-19]. In 1976, his theory was generalized by McPhedran and Maystre [8.20-21] to study the solar selective applications of such structures. Their analysis revealed that the earlier impedance theories were appropriate only to grids of infinitesimal thickness. It was also shown that such grids possessed excellent solar selective properties provided

they were used in direct sunlight and employed diurnal tracking.

The first rigorous study of spatially separated gratings was performed in 1972 by Blok and Mur [8.22]. These authors, who considered a pair of infinitesimally thin gratings whose axes of periodicity were aligned with one another, derived an integral formalism describing the diffraction problem. However, it was unfortunate that they failed to capitalize on the opportunity to discuss the application of their theory to the design of interferometers.

The material presented in this chapter, arose from the success of the study undertaken by McPhedran and Maystre, and it is believed that this is the first rigorous theoretical treatment of a bi-grating referred to as a "crossed lamellar transmission grating" - a structure composed of a pair of spatially separated singly periodic lamellar transmission gratings arranged such that their axes of periodicity are orthogonal to one another. This work was originally commenced in 1976 by my co-worker in this subject, Miss Jenny Adams, as her Honours project [8.23], but evolved rapidly to the stage where both the author and Dr. Ross McPhedran became involved in these studies. As such, the author should like to express his sincere gratitude for their stimulating collaboration during this work.

In the following section an outline of the derivation of the theoretical formalism is presented. Because of the need for brevity, it is only possible to present a précis of what would otherwise be a particularly lengthy discussion of a very complex theory. However, a complete version of the treatment is supplied in our research report [8.2]. Section (8.3), which is concerned with a discussion of the validity of the formalism, presents a new amplitude constraint appropriate

to a Littrow configuration. This property is appropriate to all symmetric grids but has proved to be a particularly valuable convergence test of this formalism, for which the "open-ended" nature of the modes tends to strain the available computing resources to the limits of their storage capacity. In that same section, a critical discussion of the use of boundary conditions applying to crossed gratings is also presented. This is drawn totally from our own experience in misusing the boundary conditions, and is included largely in the hope that those who may read this dissertation may profit from our errors. Finally, in section (8.4), a thorough discussion of the spectral characteristics of this structure is given with particular reference being made to their application in solar selective systems. This aspect of the studies has confirmed the design parameters of conventional grids (presented by McPhedran and Maystre) but has revealed that the additional array separation parameter may be successfully utilized to tune the transmission bandwidth.

8.2 THE THEORETICAL FORMALISM

8.2.1 Notation

Consider a plane wave, having wavelength λ, incident upon a pair of mutually perpendicular, perfectly conducting lamellar transmission gratings arranged in two parallel planes in free space.

The nomenclature defined for the problem is summarized in figures (8.2.1a) and (8.2.1b). For simplicity it is assumed that the temporal dependence of $\exp(-j\omega t)$ is suppressed throughout the analysis. The incident electric and magnetic fields are defined by

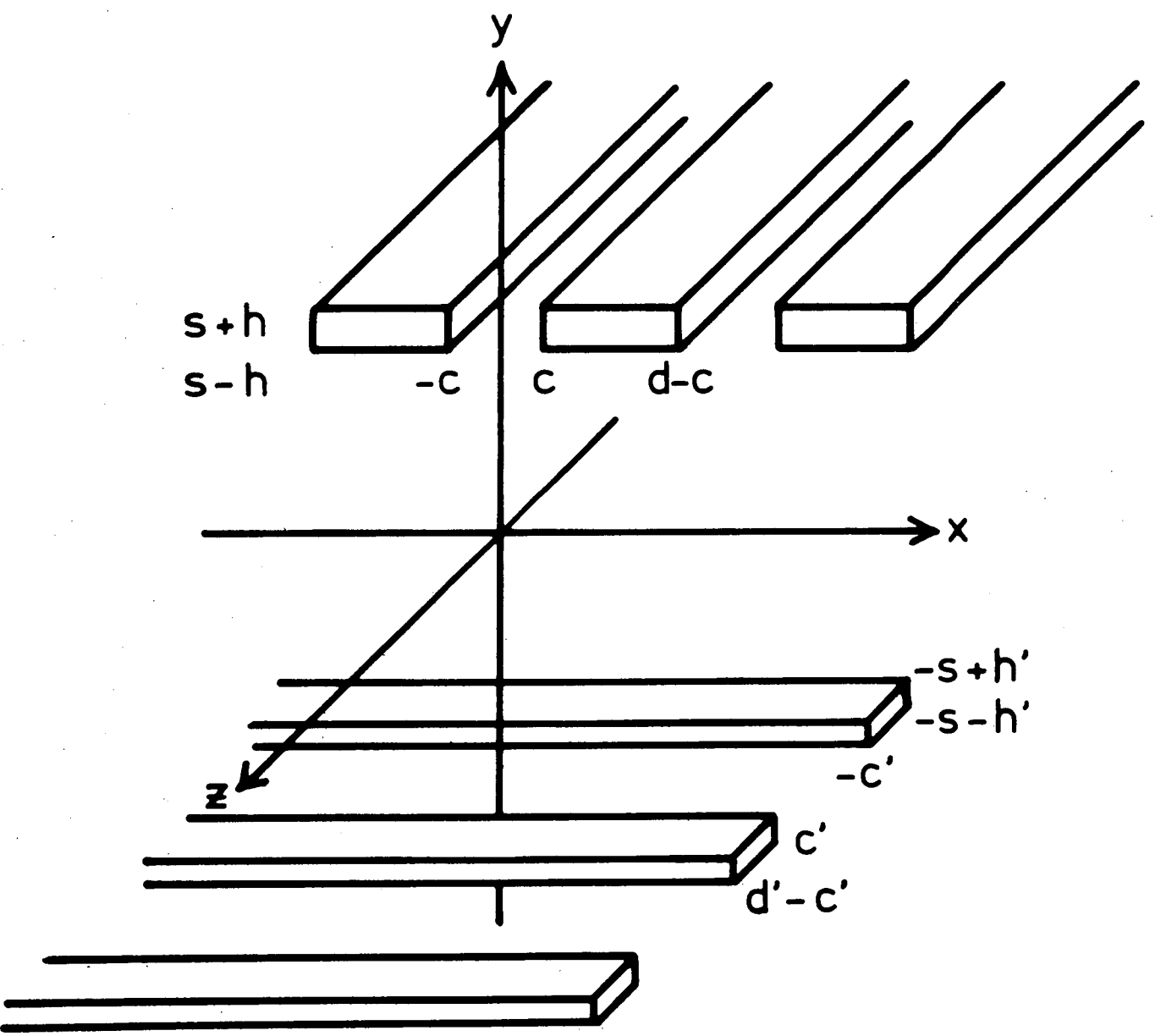


Figure 8.2.1a The general structure of the crossed lamellar transmission grating.

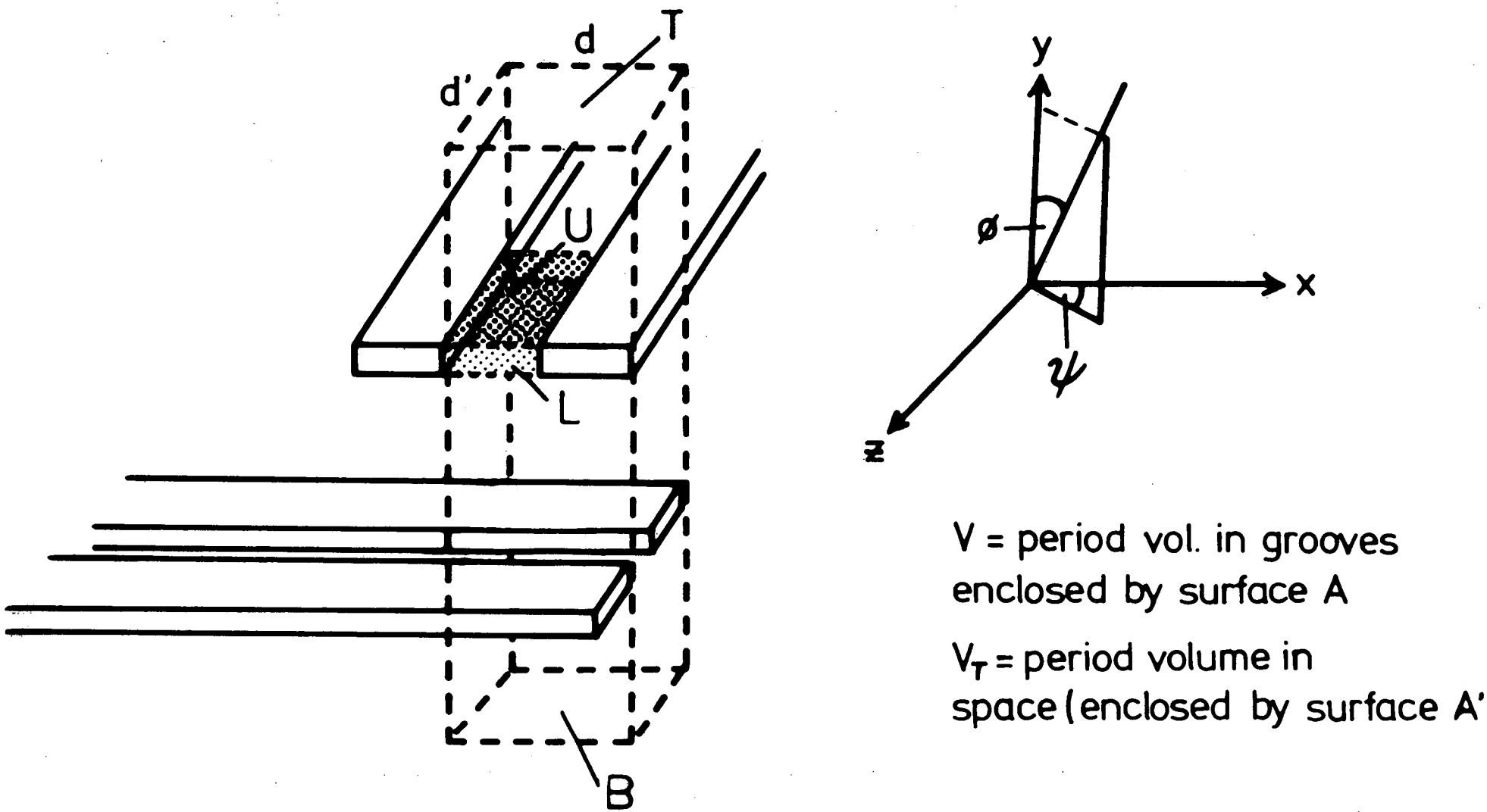


Figure 8.2.1b The volumes of integration used in the derivation of the modal expansions for the grating and also in the derivation of the amplitude constraint for the Littrow mount.

$$\underline{E}^i = \underline{A}\, R^i(x,y,z)$$

and $$\underline{H}^i = \underline{B}\, R^i(x,y,z)$$

where $R^i(x,y,z) = \exp[jk_0(\alpha_0 x - \beta_{00} y + \gamma_0 z)]$

and $k_0 = \frac{2\pi}{\lambda}$, $\alpha_0 = \sin\phi \cos\psi$

$\beta_{00} = \cos\phi$, and $\gamma_0 = \sin\phi \sin\psi$.

The incident field excites surface current distributions on both gratings and these re-radiate diffracted fields. The dual periodicity of the grating structure constrains the free-space fields to be composed of an infinite series of discrete plane waves specified by the integers p and q. Upward-going waves are associated with plane wave terms of the form

$$R_{pq}(x,y,z) = \exp[jk_0(\alpha_p x + \beta_{pq} y + \gamma_q z)].$$

Downward-going waves are made up from terms of the form

$$\hat{R}_{pq}(x,y,z) = \exp[jk_0(\alpha_p x - \beta_{pq} y + \gamma_q z)].$$

Here, $\alpha_p = \alpha_0 + p\,\frac{\lambda}{d}$

$$\gamma_q = \gamma_0 + q\,\frac{\lambda}{d},$$

and $$\beta_{pq} = \sqrt{1 - \alpha_p^2 - \gamma_q^2} \qquad \text{if} \qquad \alpha_p^2 + \gamma_q^2 \le 1$$

$$= j\sqrt{\alpha_p^2 + \gamma_q^2 - 1} \qquad \text{if} \qquad \alpha_p^2 + \gamma_q^2 > 1\ .$$

8.2.2 Edge and boundary conditions

The conditions of continuity appropriate to this problem are as follows:

$\underline{n} \times \underline{E} = 0$ on all metal surfaces

$\underline{n} \times \underline{E}$ is continuous across all air interfaces, which implies E_x and E_z are continuous across all air interfaces

$\underline{n} \times \underline{H}$ is continuous across all air interfaces, which implies H_x and H_z are continuous across these interfaces.

(Here $\underline{n}$ is a unit normal to the surface.)

It should be noted that $\underline{n} \times \underline{H}$ cannot be continuous across an air-metal interface since the incident field excites a surface current distribution on the metal.

However, a knowledge of the boundary conditions alone is not sufficient to adequately characterize the entire problem. In a problem such as has been solved, it is essential to analyse the nature of the fields in the vicinity of the metal edges and understand the singular behaviour exhibited by certain field components. Any misunderstanding of the boundary and edge conditions can certainly lead to nonphysical conclusions, a subject which is discussed at some length in section (8.3.4).

Consider a co-ordinate system, the z-axis of which is oriented along the edge and the x axis taken to be tangential to one sheet of the surface. Let the tangential plane to the other sheet be $\theta = 3\pi/2$ where r, θ and z are the components of a right handed cylindrical polar co-ordinate system. (Here r denotes a radial distance measured from the edge.) Now, by using a treatment as detailed by Jones [8.24] and Meixner [8.25], the following field behaviours may be expected –

$$\frac{\partial E_z}{\partial y}, \quad \frac{\partial E_z}{\partial x}, \quad E_x, \quad E_y \quad \propto \quad r^{-1/3} \qquad \text{(integrable singularities)}$$

$$\frac{\partial E_x}{\partial x}, \quad \frac{\partial E_y}{\partial y}, \quad \frac{\partial E_x}{\partial y}, \quad \frac{\partial E_y}{\partial x} \quad \propto \quad r^{-4/3} \qquad \text{(non-integrable singularities)}$$

and $$E_z \quad \propto \quad r^{1/3} \qquad \text{(no singularity).}$$

It is to be noted that the non-integrable singularities are partial derivatives taken normal to the edge of field components, themselves normal to the edge. Similar analyses reveal that H_x, H_y and H_z have at worst integrable singularities at the above-mentioned surfaces.

Hence, it is seen that in constructing any crossed grating formalism, one cannot impose non-physical boundary conditions such as the continuity of $\partial E_x/\partial y$ or $\partial E_z/\partial y$. (For the case of this structure, $\partial E_x/\partial y$ possess a non-integrable singularity at the surfaces $y = s \pm h$. On the other hand, $\partial E_z/\partial y$ has a non-integrable singularity at the edges of the lower array.)

Instead, one must impose boundary conditions on physical field quantities which involve no non-integrable singularities. To completely determine the diffraction problem, four field components must be constrained and these are chosen to be E_x, E_z, H_x and H_z. (At first sight, one might think that the Oy axis was the preferred axis and that E_y and H_y were the natural choices for the unknowns. However, these may be quickly rejected by considering the important case of normal incidence for which

$$\phi = 0, \qquad A_y = 0, \qquad B_y = 0.$$

For this example, any formalism based on E_y and H_y would be singular.)

8.2.3 Rayleigh expansions

The fields in the region $y \geq (s + h)$ are expressed as the sum of the incident wave and of all upward-going plane waves. Similarly, the fields in the region $y \leq -(s + h')$ are specified by a series of downward-going waves, while in the region $-(s - h') \leq y \leq (s - h)$, an expansion consisting of both upward and downward travelling waves is used. Thus, the four field components E_x, E_z, H_x and H_z can be expressed in terms of Rayleigh expansions, with the electric and magnetic field coefficients related through Maxwell's equations.

Typical expansions for the electric and magnetic fields are:

$$E_x(x,y,z) = A_x R^i(x,y,z) + \sum_{p,q} B_{pq} R_{pq}(x,y,z) \qquad \text{for } y \geq (s + h)$$

$$= \sum_{p,q} B^+_{pq} R_{pq}(x,y,z) + B^-_{pq} \hat{R}_{pq}(x,y,z) \qquad \text{for } -(s-h') \leq y \leq (s-h)$$

$$= \sum_{p,q} \hat{B}_{pq} \hat{R}_{pq}(x,y,z) \qquad \text{for } y \leq -(s+h')$$

and

$$H_x(x,y,z) = B_x R^i(x,y,z) + \sum_{p,q} X_{pq} R_{pq}(x,y,z) \qquad \text{for } y \geq (s + h)$$

$$= \sum_{p,q} X^+_{pq} R_{pq}(x,y,z) + X^-_{pq} \hat{R}_{pq}(x,y,z) \qquad \text{for } -(s-h') \leq y \leq (s-h)$$

$$= \sum_{p,q} \hat{X}_{pq} \hat{R}_{pq}(x,y,z) \qquad \text{for } y \leq -(s+h').$$

In the above expressions, the summation indices p and q lie in the range $(-\infty,\infty)$.

Maxwell's equations can be used to derive relations of the following form between the plane wave coefficients:

$$\text{e.g.} \quad X_{pq} = \frac{jk_0}{\beta_{pq}} [\alpha_p \gamma_q B_{pq} + (1 - \alpha_p^2) C_{pq}],$$

where $\{C_{pq}\}$ are the Rayleigh coefficients of E_z in the region $y \geq (s+h)$.

8.2.4 Modal structure

The field in each aperture is specified by waveguide modal expansions, each distinct mode of which satisfies the boundary conditions on the aperture walls. The grooves in the upper array are open-ended in the z-direction and as a consequence the modes in this region will have a z-dependence characterized by the pseudo-periodicity of the Rayleigh orders. Similar remarks apply to the lower array.

In this section, attention is focussed on the mode structure appropriate to the upper array. The boundary conditions relevant to the problem are :

$$E_y = E_z = 0 \qquad \text{and} \qquad H_x = 0$$

on the aperture walls. Using these boundary conditions, it is possible to rigorously derive modal expansions for these field quantities. The mode structure for the field components E_x, H_y and H_z can then be deduced from these expansions. In the case of these three components, there are no simple boundary conditions from which a similar rigorous derivation of the mode structure can be deduced.

The following treatment is based on the work of Wirgin and Deleuil [8.26]. However, the generalization from singly to doubly periodic gratings entails sufficient difficulties to warrant an exposition of the derivation.

Consider the field component E_z which obeys the Helmholtz equation

$$(\nabla^2 + k_0^2)\, E_z(P) = 0$$

for any point P lying within the unit period groove cell V defined by $x \in [-c,c]$, $z \in [0,d']$. From an application of Green's theorem to this cell it follows that

$$\iiint_V [G_0 \nabla^2 E_z - E_z \nabla^2 G_0]\, dV = - \iint_A [G_0 \frac{\partial E_z}{\partial n'} - E_z \frac{\partial G_0}{\partial n'}]\, dA'$$

where $\underline{n}$ is an inwardly directed unit normal vector to the surface A bounding the volume V. Here G_0 is a Green's function whose form is still to be determined. As in all modal formalisms, the choice of the Green's function is totally governed by the geometry of the aperture and the field boundary conditions on the metal walls. Since it has been possible to expand the fields in the regions $y \in [s+h,\infty)$ and $y \in [-s+h',s-h]$ in terms of Rayleigh expansions, the solution of the field problem (performed by matching the groove and free space field representations at the aperture boundaries U and L) would be greatly simplified should the Green's function implicitly incorporate the field properties within the groove. This has the desirable effect of eliminating any contributions to the area integral from the metallic walls of the aperture.

This is achieved by following the method of singularity imaging as discussed by Wirgin [8.27]. In this case, the chosen Green's function obeys the inhomogeneous Helmholtz equation:

$$(\nabla^2 + k_0^2)\, G_0(P;M') = \delta_A(P;M')$$

$$= \sum_{m=-\infty}^{\infty} [\delta(x-x' + 4mc) - \delta(x+x' + (4m + 2)c)]$$

$$\times\, \delta(y-y') \times \sum_{q=-\infty}^{\infty} \delta(z-z' + qd')\, \exp(-jqk_0\gamma_0 d')$$

where $P = (x,y,z)$ and $M' = (x',y',z')$.

In the above expression, the singularities have been located such that the Green's function vanishes on the metal walls $x = -c, c$ and possesses the pseudo-periodicity characteristic of the Rayleigh expansion for the z-dependence. (The latter is a consequence of the open-ended nature of the grooves in the z direction.)

However, before attempting to solve this differential equation, it is necessary to obtain a solution to the simpler expression

$$(\nabla^2 + k_0^2)\, G_p(\underline{r};\underline{r}') = \delta(x-x')\, \delta(y-y')\, \delta(z-z')$$

where $\underline{r} = \underline{OP}$ and $\underline{r}' = \underline{OM}'$.

This is achieved by expanding the Green's function in an orthonormal, continuum basis of plane waves:

$$\psi(\tilde{\underline{k}}_0;\underline{r}) = \frac{1}{(2\pi)^{3/2}} \exp(j\, \tilde{\underline{k}}_0 \cdot \underline{r})$$

where $\tilde{\underline{k}}_0 = k_0(\alpha,\beta,\gamma)$.

The Green's function is then written in the form

$$G_p(\underline{r};\underline{r}') = \iiint_{-\infty}^{\infty} A(\tilde{\underline{k}}_0;\underline{r}')\, \psi(\tilde{\underline{k}}_0;\underline{r})\, d\alpha\, d\beta\, d\gamma.$$

Now since G_p must obey the above inhomogeneous Helmholtz equation, it follows that

$$A(\tilde{\underline{k}}_0;\underline{r}') = \frac{k_0^3}{k_0^2 - \tilde{k}_0^2}\, \psi(\tilde{\underline{k}}_0;\underline{r}')$$

(upon noting the orthogonality of the basis $\{\psi\}$).

Hence,

$$G_p(\underline{r};\underline{r}') = \frac{k_0^3}{8\pi^3} \iiint_{-\infty}^{\infty} \frac{\exp[jk_0\alpha(x-x') + jk_0\beta|y-y'| + jk_0\gamma(z-z')]}{k_0^2(1 - \alpha^2 - \beta^2 - \gamma^2)}\, d\alpha\, d\beta\, d\gamma\ .$$

This expression may be even further reduced by explicitly integrating with respect to β. In this integration, the contour β_T is chosen so as to completely exclude a pole whose physical interpretation is a field sink [8.28]. Thus, by defining

$$\nu^2 = 1 - \alpha^2 - \gamma^2$$

it may be shown that

$$\int_{-\infty}^{\infty} \frac{\exp[jk_0\beta|y-y'|]}{k_0^2(\nu^2 - \beta^2)}\, d\beta$$

$$= \int_{\beta_T} \frac{\exp[jk_0\beta|y-y'|]}{k_0^2(\nu^2 - \beta^2)}\, d\beta$$

$$= -2\pi j \frac{\exp[jk_0\nu|y-y'|]}{2\, k_0^2\nu}\,.$$

Thus,

$$G_p(\underline{r};r') = -\frac{jk_0}{8\pi^2} \iint_{-\infty}^{\infty} \frac{\exp[jk_0\alpha(x-x') + jk_0\nu|y-y'| + jk_0\gamma(z-z')]}{\nu}\, d\alpha\, d\gamma.$$

Now, by appropriately superposing the "primitive" Green's functions according to the specification laid down by the definition of G_0 it follows that

$$G_0 = -\frac{k_0}{4\pi^2} \iint_{\infty} \frac{1}{\nu} \{ \sum_{m=-\infty}^{\infty} \exp(jk_0\alpha\, 4mc)\ \exp[jk_0\alpha(x+c)]\sin[k_0\alpha(x'+c)]$$

$$\cdot\ \exp[jk_0\nu|y-y'|]\ \cdot \sum_{q=-\infty}^{\infty} \exp[jk_0(\gamma-\gamma_0)\, qd']$$

$$\cdot\ \exp[jk_0\gamma(z-z')]\ \}\, d\alpha\, d\gamma.$$

From Poisson's summation formula, it may be shown that

$$\sum_{m=-\infty}^{\infty} \exp[jk_0\alpha\, 4mc] = \frac{\pi}{2ck_0} \sum_{m=-\infty}^{\infty} \delta(\alpha - \frac{m\pi}{2ck_0})$$

and

$$\sum_{q=-\infty}^{\infty} \exp[jk_0(\gamma-\gamma_0)qd'] = \frac{2\pi}{d'k_0} \sum_{q=-\infty}^{\infty} \delta(\gamma-\gamma_q)$$

and thus G_0 may be shown to reduce to the form:

$$G_0 = \frac{1}{2icd'} \sum_{m=1}^{\infty} \sum_{q=-\infty}^{\infty} \{\frac{1}{k_0 \nu_m^q} \sin[\frac{m\pi}{2c}(x+c)] \sin[\frac{m\pi}{2c}(x'+c)]$$

$$\exp[jk_0 \nu_m^q |y-y'|] \exp[jk_0 \gamma_q (z-z')] \}$$

where

$$k_0 \nu_m^q = \sqrt{k_0^2(1 - \gamma_q^2) - (\frac{m\pi}{2c})^2} .$$

With this definition of G_0, the previous application of Green's theorem yields the following integral representation for E_z within the aperture V.

$$E_z(P) = \iint_{U \cup L} [G_0(P;M') \frac{\partial E_z(M')}{\partial n'} - \frac{\partial G_0(P;M')}{\partial n'} E_z(M')] \, dA'.$$

Simple manipulation of this expression leads to

$$E_z(P) = \sum_{m=1}^{\infty} \sum_{q=-\infty}^{\infty} \{c_m^q \sin[k_0 \nu_m^q (y-s)] + d_m^q \cos[k_0 \nu_m^q (y-s)]\} \, MEZU_m^q(x,z)$$

where the modes of E_z (and also E_y and H_x) take the form

$$MEZU_m^q(x,z) = g_m^q \sin[\frac{m\pi}{2c}(x+c)] \exp[jk_0 \gamma_q z]$$

where

$m \in [1,\infty)$ and $q \in (-\infty,\infty)$.

Here, $g_m^q = \sqrt{\frac{\varepsilon_m}{2cd'}}$

and $\varepsilon_m = 1$ if $m = 0$

$= 2$ if $m \neq 0$.

It then follows from Maxwell's equations that the corresponding mode structure of the field components E_x, H_y and H_z takes the form:

$$MEXU_m^q(x,z) = g_m^q \cos[\frac{m\pi}{2c}(x+c)] \exp[jk_0\gamma_q z]$$

for all $m \in [0,\infty)$ and $q \in (-\infty,\infty)$.

Thus, by writing E_x as a superposition of these modes, it follows that

$$E_x(P) = \sum_{m=0}^{\infty} \sum_{q=-\infty}^{\infty} \{a_m^q \sin[k_0\nu_m^q(y-s)] + b_m^q \cos[k_0\nu_m^q(y-s)]\} MEXU_m^q(x,z).$$

Although each mode of E_x obeys

$$\frac{\partial}{\partial x} MEXU_m^q(\pm c,z) = 0$$

it is incorrect to say that

$$\frac{\partial}{\partial x} E_x(\pm c,y,z) = 0.$$

Such an assertion would involve the term-by-term differentiation of the above series, an action which is unjustified in view of the divergent nature of E_x at a groove edge.

8.2.5 Solution of the field equations

The initial specification of the problem involves 32 infinite sets of coefficients, composed of 16 sets of modal coefficients and 16 sets of Rayleigh coefficients, characterizing the field quantities E_x, E_z, H_x and H_z. Boundary conditions for the tangential components of $\underline{E}$ and $\underline{H}$ are then applied at the surfaces $y = (s \pm h)$ and $-(s \pm h')$. and the resulting equations are solved using the method of moments.

A typical electric field equation is derived from the continuity of E_x at $y = s + h$ and gives

$$A_x R^i(x,s+h,z) + \sum_{p,q} B_{pq} R_{pq}(x,s+h,z)$$

$$= 0 \qquad \forall \quad x \in (c,d-c)$$

$$= \sum_{\substack{m=0 \\ q=-\infty}}^{\infty} (a_m^q \sin k_0 \nu_m^q h + b_m^q \cos k_0 \nu_m^q h)\ MEXU_m^q(x,z)$$

$$\forall \quad x \in [-c,c]$$

This equation is multiplied throughout by $\bar{R}_{PQ}(x,0,z)$ (where the bar denotes complex conjugation) and the resulting expression is integrated over a rectangular area defined by $z \in [0,d']$, $x \in [-c,c]$. The orthogonality of the plane wave terms is expressed by :

$$\int_0^{d'} \int_0^{d} R_{pq}(x,0,z)\ \bar{R}_{PQ}(x,0,z)\ dxdz = dd'\ \delta_{pP}\ \delta_{qQ}$$

and this results in

$$A_x^*\ \delta_{P0}\ \delta_{Q0} + B_{PQ}^* = \sum_{m=0}^{\infty} (a_m^Q \sin k_0 \nu_m^Q h + b_m^Q \cos k_0 \nu_m^Q h)\ IXU_m^P$$

where $A_x^* = A_x \exp[-jk_0\beta_{00}(s+h)]$

$B_{PQ}^* = B_{PQ} \exp[jk_0\beta_{PQ}(s+h)]$

and $IXU_m^P = \left(\frac{\varepsilon_m}{2cd'}\right)^{\frac{1}{2}} \frac{1}{d} \int_{-c}^{c} \cos \frac{m\pi(x+c)}{2c} \exp[-jk_0\alpha_p x]\, dx$

and $\varepsilon_m = 1$ if $m = 0$, $\varepsilon_m = 2$ if $m \neq 0$.

A typical magnetic field equation is derived from the continuity of H_x at $y = s + h$ and leads to

$$B_x R^i(x,s+h,z) + \sum_{p,q} X_{pq} R_{pq}(x,s+h,z)$$

$$= \sum_{\substack{m=1 \\ q=-\infty}}^{\infty} (A_m^q \sin k_0\nu_m^q h + B_m^q \cos k_0\nu_m^q h)\, MEZU_m^q(x,z) \quad \forall \begin{cases} x \in [-c,c] \\ z \in [0,d']. \end{cases}$$

(Here the coefficients $\{A_m^q\}$ and $\{B_m^q\}$ represent the y-anti-symmetric and y-symmetric components of H_x.)

This equation is then multiplied by $\overline{MEZU}_M^Q$ and integrated over an area bounded by $x \in [-c,c]$, $z \in [0,d']$ to give

$$dd' \sum_p (X_{pQ}^* + B_x^* \delta_{p0}\delta_{Q0})\, \overline{IZU}_M^P = A_M^Q \sin k_0\nu_M^Q h + B_M^Q \cos k_0\nu_M^Q h, \quad M \in [1,\infty]$$

where $X_{pQ}^* = X_{pQ} \exp[jk_0\beta_{pQ}(s+h)]$

$B_x^* = B_x \exp[-jk_0\beta_{00}(s+h)]$

and $IZU_m^P = \left(\frac{\varepsilon_m}{2cd'}\right)^{\frac{1}{2}} \frac{1}{d} \int_{-c}^{c} \sin \frac{m\pi(x+c)}{2c} \exp[-jk_0\alpha_p x]\, dx.$

In deriving the above expression, it has been necessary to invoke the orthogonality of the modes, viz.

$$\int_{-c}^{c}\int_{0}^{d'} MEZU_m^q \, \overline{MEZU}_n^r \, dxdz = \delta_{qr}\delta_{mn}.$$

Sixteen matrix equations, relating the 32 unknown infinite sets of Rayleigh and modal coefficients, can be derived using similar analyses for all field components at the surfaces $y = s \pm h$ and $y = -(s \pm h')$. Maxwell's equations provide the further 16 constraints relating these unknowns and are used to eliminate the coefficients corresponding to the magnetic field components from the set of equations. Algebraic manipulation is then used to eliminate the remaining Rayleigh coefficients and eight independent matrix equations in the eight infinite sets of modal coefficients specifying the electric field are finally derived. (A full description of both the theory and the numerical implementation has been documented in our report [8.2].) Reconstruction of the Rayleigh coefficients enables the reflected and transmitted efficiencies associated with each real diffracted order to be evaluated.

It should be noted that this set of equations cannot be decoupled, as was possible for the inductive grid [8.20], and thus the numerical solution for three modal indices and five Rayleigh orders in both arrays involves the inversion of a matrix with 100 × 100 complex elements. Such a solution requires approximately three minutes of processor time on the Burroughs B6700 computer. Thus, it is machine computation time and array storage requirements which limit the number of modal quantities which can be determined. Should too great a number of Rayleigh orders be propagating, then the matrices to be generated would exceed the resources

of the computer. Consequently, it has been necessary to restrict the numerical solution to normalized wavelengths (λ/d and λ/d') in excess of 0.5.

8.3 THEORETICAL PROPERTIES OF THE FORMALISM

This section is devoted both to the discussion of the theoretical tests used in checking the validity of the theoretical formalism and the numerical accuracy of our solution, and also to a discussion of the correct use of boundary conditions when applied to grid structures.

8.3.1 Theoretical tests of the formalism

In order to obtain a solution to the diffraction problem it is necessary to truncate the infinite series defining the fields and thus it is desirable to have some indication as to the extent of these truncation errors.

The formalism outlined above has been thoroughly tested against the well known constraints of conservation of energy and reciprocity. However, as has been shown by McPhedran and Maystre [8.21], energy conservation is analytically satisfied by any grid formalism independently of truncation errors. This criterion cannot therefore be used to test the physical accuracy of the numerical solutions to the equations describing the scattering problem. Thus, to date, the only real test of convergence of such solutions has been reciprocity. (Evidence of the agreement of the theory with the two-dimensional version of the reciprocity theorem is presented in section (8.3.3).)

However, during the course of these investigations, a further test was devised to examine the validity of this formalism, and in general, any grid formalism. This test, which is discussed in the following sub-section involves a constraint upon the electric field amplitudes when the grid is operated in a Littrow configuration.

8.3.2 An amplitude constraint for the Littrow mount

The search for such a property pertaining to any grid having rectangular axes of symmetry was motivated by previous work (discussed in chapter 4) which showed that the phases of the reflected and transmitted orders for any lossless, symmetric, singly periodic transmission grating operated in a $(-1)^{th}$ order Littrow mount between the (-2,+1) and (-1) Wood anomalies, were related to the reflected and transmitted efficiencies as follows:

$$\frac{\cos(\delta_0^R - \delta_{-1}^R)}{\cos(\delta_0^T - \delta_{-1}^T)} = -\sqrt{\frac{\rho_0^T \rho_{-1}^T}{\rho_0^R \rho_{-1}^R}}$$

where δ and ρ represent the phase and efficiency respectively of the various reflected (R) and transmitted (T) orders.

Throughout the following analysis it will be assumed that field quantities are physical entities, in contrast to those used in the numerical solution of the problem (which are represented by truncated series). It should also be noted that the relation to be derived relies on the coordinate axes being located symmetrically within the apertures.

In the space $y \geq (s + h)$

$$E_x = \sum_{p,q} [A_x \exp(- jk_0\beta_{00}y)\ \delta_{p0}\delta_{q0} + B_{pq} \exp(jk_0\beta_{pq}y)] \exp[jk_0(\alpha_p x + \gamma_q z)].$$

Now, should the grid be operated in a (f,g) order Littrow configuration then the values of α_p and γ_q are given by

$$\left.\begin{aligned} \alpha_p &= -\alpha_{f-p} = (p - f/2)\ \lambda/d \\ \text{and}\quad \gamma_q &= -\gamma_{g-q} = (q - g/2)\ \lambda/d'. \end{aligned}\right\}$$

By now operating the grating in a (-f,-g) order Littrow mount, the corresponding wave field is specified by

$$E_x^{\dagger} = \sum_{p,q} [A_x \exp(- jk_0\beta_{00}y)\ \delta_{p0}\delta_{q0} + B_{pq}^{\dagger} \exp(jk_0\beta_{pq}^{\dagger}y)] \exp[jk_0(\alpha_p^{\dagger}x + \gamma_q^{\dagger}z)]$$

$$\left.\begin{aligned} \text{where}\ \alpha_p^{\dagger} &= (p + f/2)\ \lambda/d = -\alpha_{-p} \\ \text{and}\quad \gamma_q^{\dagger} &= (q + g/2)\ \lambda/d' = -\gamma_{-q}\ . \end{aligned}\right\}$$

The symmetry of the optical arrangement implies $B_{pq}^{\dagger} = B_{-p,-q}$ and thus the above expression for $E_x^{\dagger}$ becomes

$$E_x^{\dagger} = \sum_{p,q} [A_x \exp(-jk_0\beta_{00}y)\delta_{p,f}\delta_{q,g} + B_{f-p,g-q} \exp(jk_0\beta_{\substack{f-p,\\ g-q}}y)]\exp[jk_0(\alpha_p x+\gamma_q z)].$$

If these two fields are then superposed, the total electric and magnetic field quantities, given by $\underline{E}^T$ and $\underline{H}^T$ respectively, can be written in the form

$$\underline{E}^T = \underline{E} + \underline{E}^{\dagger}$$
$$\underline{H}^T = \underline{H} + \underline{H}^{\dagger}\ .$$

Thus, the total flux of energy through the surface T, defined in figure (8.2.1b) is given by

$$\tfrac{1}{2}\mathrm{Re}\Big[\iint_T \left[E_z^T\overline{H}_x^T - E_x^T\overline{H}_z^T\right] dxdz = \xi_{\delta\delta} + \xi_{\delta c} + \xi_{cc}$$

where $\xi_{\delta\delta} = \dfrac{dd'}{\omega\mu_0}\,\mathrm{Im}\{A_z\overline{B}_x - A_x\overline{B}_z\}$

$$\xi_{\delta c} = \frac{dd'}{\omega\mu_0}\,\mathrm{Im}\{A_z\exp(-jk_0\beta_{00}y)\,[\overline{X}_{00}\exp(-jk_0\beta_{00}y) + \overline{X}_{f,g}\exp(-jk_0\beta_{00}y)]$$

$$+\ \overline{B}_x\exp(jk_0\beta_{00}y)\,[C_{00}\exp(jk_0\beta_{00}y) + C_{f,g}\exp(jk_0\beta_{00}y)]$$

$$-\ A_x\exp(-jk_0\beta_{00}y)[\overline{Z}_{00}\exp(-jk_0\beta_{00}y) + \overline{Z}_{f,g}\exp(-jk_0\beta_{00}y)]$$

$$-\ \overline{B}_z\exp(jk_0\beta_{00}y)\,[B_{00}\exp(jk_0\beta_{00}y) + B_{f,g}\exp(jk_0\beta_{00}y)]\ \}$$

$$\xi_{cc} = \frac{dd'}{2\omega\mu_0}\,\mathrm{Im}\sum_{p,q}\{[C_{pq}\exp(jk_0\beta_{pq}y) + C_{FG}\exp(jk_0\beta_{FG}y)][\overline{X}_{pq}\exp(-jk_0\overline{\beta}_{pq}y) +$$

$$+\ \overline{X}_{FG}\exp(-jk_0\overline{\beta}_{FG}y)] - [B_{pq}\exp(jk_0\beta_{pq}y) + B_{FG}\exp(jk_0\beta_{FG}y)]\times$$

$$\times\,[\overline{Z}_{pq}\exp(-jk_0\overline{\beta}_{pq}y) + \overline{Z}_{FG}\exp(-jk_0\overline{\beta}_{FG}y)]\}.$$

In the above expressions, the Rayleigh coefficients $\{X_{pq}\}$ and $\{Z_{pq}\}$ refer to the x and z components of the magnetic field. (These are related to the electric field coefficients $\{B_{pq}\}$ and $\{C_{pq}\}$ through Maxwell's equations.) Also, the symbols F and G are defined as follows

$$F = f - p \qquad \text{and} \qquad G = g - q.$$

After application of Maxwell's equations, it may be shown that

$$\xi_{\delta c} = \frac{dd'}{\omega\mu_0} \mathrm{Im}\{2\mathrm{Re}[\bar{C}_{00}B_x - \bar{B}_{00}B_z + \bar{C}_{fg}B_x - \bar{B}_{fg}B_z) \exp(-2jk_0\beta_{00}y)]\} .$$

$$= 0,$$

after noting that

$$\beta_{fg} = \beta_{00}.$$

The expression for ξ_{cc} can be expanded in the form

$$\xi_{cc} = \xi_{cc1} + \xi_{cc2} + \xi_{cc3} + \xi_{cc4}$$

where

$$\xi_{cc1} = \frac{dd'}{2\omega\mu_0} \mathrm{Im}\{\sum_{pq} [C_{pq}\bar{X}_{pq} - B_{pq}\bar{Z}_{pq}] \exp[jk_0(\beta_{pq} - \bar{\beta}_{pq})y]\}$$

$$\xi_{cc2} = \frac{dd'}{2\omega\mu_0} \mathrm{Im}\{\sum_{pq} [C_{FG}\bar{X}_{FG} - B_{FG}\bar{Z}_{FG}] \exp[jk_0(\beta_{FG} - \bar{\beta}_{FG})y]\}$$

$$\xi_{cc3} = \frac{dd'}{2\omega\mu_0} \mathrm{Im}\{\sum_{pq} [C_{pq}\bar{X}_{FG} - B_{pq}\bar{Z}_{FG}] \exp[jk_0(\beta_{pq} - \bar{\beta}_{FG})y]\}$$

$$\xi_{cc4} = \frac{dd'}{2\omega\mu_0} \mathrm{Im}\{\sum_{pq} [C_{FG}\bar{X}_{pq} - B_{FG}\bar{Z}_{pq}] \exp[jk_0(\beta_{FG} - \bar{\beta}_{pq})y]\} .$$

Only terms representing real orders of diffraction are capable of producing non-trivial contributions to the expression ξ_{cc}. Thus, after some simple manipulation it may be shown that

$$\iint_T S_y dA = \frac{dd'}{2\omega\mu_0} \mathrm{Im}[2(A_z\bar{B}_x - A_x\bar{B}_z) + \sum_{p,q\in\Omega}(C_{pq}\bar{X}_{pq} - B_{pq}\bar{Z}_{pq}$$

$$+ C_{FG}\bar{X}_{FG} - B_{FG}\bar{Z}_{FG} + C_{pq}\bar{X}_{FG} - B_{pq}\bar{Z}_{FG} + C_{FG}\bar{X}_{pq} - B_{FG}\bar{Z}_{pq})]$$

where $\Omega = \{(p,q) \mid \mathrm{Im}(\beta_{pq}) = 0\}$.

A similar result can be derived for the region $y \leq -(s + h')$. Now since $\iint_T S_y dA = \iint_B S_y dA$ physically within our formalism, the following relation can be derived:

$$\sum_{(p,q \in \Omega)} \frac{1}{\beta_{pq}} \{\alpha_p \gamma_q (C_{pq}\bar{B}_{FG} + B_{pq}\bar{C}_{FG} + \hat{C}_{pq}\bar{\hat{B}}_{FG} + \hat{B}_{pq}\bar{\hat{C}}_{FG})$$

$$+ (1 - \alpha_p^2)(C_{pq}\bar{C}_{FG} + \hat{C}_{pq}\bar{\hat{C}}_{FG}) + (1 - \gamma_q^2)(B_{pq}\bar{B}_{FG} + \hat{B}_{pq}\bar{\hat{B}}_{FG})\} = 0 ,$$

since all other terms cancel because of the constraint imposed by the conservation of energy. This relation, termed the Littrow expression, for a general Littrow mount is the two-dimensional analogue of the expression derived for lossless singly periodic transmission gratings.

By way of an example, if a crossed grating is used in a (-1,-1) order Littrow mount with only four real propagating orders, the amplitude constraint becomes:

$$\alpha_0\gamma_0 \,\mathrm{Re}[C_{00}\bar{B}_{-1-1} + B_{00}\bar{C}_{-1-1} - C_{0-1}\bar{B}_{-10} - B_{0-1}\bar{C}_{-10}$$

$$+ \hat{C}_{00}\bar{\hat{B}}_{-1-1} + \hat{B}_{00}\bar{\hat{C}}_{-1-1} - \hat{C}_{0-1}\bar{\hat{B}}_{-10} - \hat{B}_{0-1}\bar{\hat{C}}_{-10}]$$

$$+ (1-\alpha_0^2)\,\mathrm{Re}[C_{00}\bar{C}_{-1-1} + C_{0-1}\bar{C}_{-10} + \hat{C}_{00}\bar{\hat{C}}_{-1-1} + \hat{C}_{0-1}\bar{\hat{C}}_{-10}]$$

$$+ (1-\gamma_0^2)\,\mathrm{Re}[B_{00}\bar{B}_{-1-1} + B_{0-1}\bar{B}_{-10} + \hat{B}_{00}\bar{\hat{B}}_{-1-1} + \hat{B}_{0-1}\bar{\hat{B}}_{-10}] = 0.$$

The Littrow expression is in general not precisely satisfied by the computed field amplitudes and thus it can be used as a convergence check

for the algorithm. However, if the truncated set of Rayleigh orders is chosen in the following field symmetric fashion:

$$\begin{aligned}\Omega_1 = \{(p,q)\,|\; & f - n \le p \le n && \text{if} && f \le 0;\\ & -n \le p \le f + n && \text{if} && f > 0;\\ & g - m \le q \le m && \text{if} && g \le 0;\\ & -m \le q \le g + m && \text{if} && g > 0\}\end{aligned}$$

where n and m are arbitrary positive integers, then the constraint is analytically satisfied. These properties can be demonstrated as follows.

The Littrow expression was derived from a consideration of integrals of the form:

$$\tfrac{1}{2}\mathrm{Re}\int_{T/B}\int (E_z^T\bar{H}_x^T - E_x^T\bar{H}_z^T)\,dx\,dz.$$

Now,

$$E_z^T\bar{H}_x^T - E_x^T\bar{H}_z^T = (E_z + E_z^\dagger)(\bar{H}_x + \bar{H}_x^\dagger) - (E_x + E_x^\dagger)(\bar{H}_z + \bar{H}_z^\dagger)$$

$$= (E_z\bar{H}_x - E_x\bar{H}_z) + (E_z^\dagger\bar{H}_x^\dagger - E_x^\dagger\bar{H}_z^\dagger) + (E_z^\dagger\bar{H}_x - E_x^\dagger\bar{H}_z) + (E_z\bar{H}_x^\dagger - E_x\bar{H}_z^\dagger).$$

The first two terms are those used in the derivation of energy conservation for the separate fields and hence can be shown to be analytically preserved across the free-space groove interfaces, regardless of truncation errors [8.20].

Now, in the case of a (f,g) order Littrow mount let us define a further set Ω_2 such that

$$\Omega_2 = \{(p,q) \mid \begin{array}{lll} f' - n \le p \le n & \text{if} & f' \le 0; \\ -n \le p \le f' + n & \text{if} & f' > 0; \\ g' - m \le q \le m & \text{if} & g' \le 0; \\ -m \le q \le g' + m & \text{if} & g' > 0\} \end{array}$$

where $f' = -f, \quad g' = -g.$

Consider the continuity of the term $E_x \overline{H}_z^{\dagger}$ across the interface U (see figure 8.2.1b) at $y = s + h$. In the following analysis, it is taken that $(p,q) \in \Omega_1$ and $(r,s) \in \Omega_2$.

$$\text{Then, } I(U) = \int_0^d \int_0^{d'} E_x \overline{H}_z^{\dagger} \, dA'$$

$$= \int_0^d \int_0^{d'} \sum_{p,q} [A_x \exp\{-2jk_0\beta_{00}(s+h)\} \delta_{p0}\delta_{q0} + B_{pq}] \exp\{jk_0(\alpha_p x + \gamma_q z)\} \times$$

$$\exp\{jk_0\beta_{pq}(s+h)\} \times \sum_{r,s} [\overline{B}_z^{\dagger} \exp\{2jk_0\beta_{00}(s+h)\} \delta_{r0}\delta_{s0} + \overline{Z}_{rs}^{\dagger}] \times$$

$$\exp\{-jk_0(\alpha_{f+r}x + \overline{\beta}_{f+r,g+s}(s+h) + \gamma_{g+s}z)\} \, dx \, dz$$

since $\alpha_r^{\dagger} = \alpha_{f+r}$ and $\gamma_s^{\dagger} = \gamma_{g+s}$.

Now f+r and g+s are elements of Ω_1, and thus

$$I(U) = \sum_{p,q} dd' [A_x \exp\{-2jk_0\beta_{00}(s+h)\}\delta_{p0}\delta_{q0} + B_{pq}] \times \exp[jk_0\beta_{pq}(s+h)] \times$$

$$\times \; [B_z^{\dagger} \exp\{2jk_0\overline{\beta}_{fg}(s+h)\}\delta_{pf}\delta_{qg} + \overline{Z}^{\dagger}_{\substack{p-f, \\ q-g}}] \times \exp[-jk_0\overline{\beta}_{pq}(s+h)]$$

Now the continuity of E_x over the aperture yields

$$dd'[A_x \exp\{-2jk_0\beta_{00}(s+h)\}\delta_{p0}\delta_{q0} + B_{pq}] = \sum_{m,\ell}(a_m^{\ell *} + b_m^{\ell *}) < MEXU_m^{\ell}, R_{pq} >_A \times \times \exp[-jk_0\beta_{pq}(s+h)]$$

where $\ell \in \Omega_1$,

$$a_m^{\ell *} = a_m^{\ell} \sin(k_0 \nu_m^{\ell} h),$$

$$b_m^{\ell *} = b_m^{\ell} \cos(k_0 \nu_m^{\ell} h)$$

and the inner product is defined by

$$<f,g>_A = \iint_A f\bar{g}\, dA.$$

In the above, A refers to the area U of figure (8.2.1b) and so

$$I(U) = \sum_{m,\ell}(a_m^{\ell *} + b_m^{\ell *})\sum_{p,q}[B_z^{\dagger}\exp\{2jk_0\beta_{fg}(s+h)\}\delta_{pf}\delta_{qg} + \bar{Z}^{\dagger}_{\substack{p-f,\\ q-g}}] \times \times \exp[-jk_0\bar{\beta}^{\dagger}_{pq}(s+h)] < MEXU_m^{\ell}, R_{pq} >_A.$$

Also continuity of $H_z^{\dagger}$ across the groove free-space interface, U, gives

$$\sum_{r,s}[B_z^{\dagger}\exp\{-2jk_0\beta_{00}^{\dagger}(s+h)\}\delta_{r0}\delta_{s0} + Z_{rs}^{\dagger}] < R_{rs}^{\dagger}, MEXU_m^{t\dagger} >_A \exp[jk_0\beta_{rs}^{\dagger}(s+h)] = C_m^{t*\dagger} + D_m^{t*\dagger}$$

where r, s, t $\in \Omega_2$. (In this expression the C_m^{t*} and D_m^{t*} refer to the modal coefficients for the z component of the magnetic field evaluated at the surface y = s + h. The attached "dagger" refers to the modes established by operating the grating in the (-f,-g) order Littrow configuration.)

Now by writing F = f + r, G = g + s and H = t + g, then F, G and H are elements of Ω_1. Since $MEXU_m^{t\dagger} = MEXU_m^{g+t}$, and also $R_{rs}^{\dagger} = R_{FG}$ consequent

upon $\alpha_r^{\dagger} = \alpha_{f+r}$ and also $\gamma_s^{\dagger} = \gamma_{g+s}$ it follows that

$$\sum_{F,G} [B_z^{\dagger} \exp -2jk\beta_{0fg}(s+h)\delta_{Ff}\delta_{Gg} + Z^{\dagger}_{\substack{F-f, \\ G-g}}] \exp[jk_0\beta_{FG}(s+h)] \langle R_{FG}, MEXU_m^H \rangle A$$
$$= C_m^{H-g*\dagger} + D_m^{H-g*\dagger}.$$

Hence, the expression for I(U) now becomes

$$I(U) = \sum_{m,\ell} (a_m^{\ell*} + b_m^{\ell*}) \overline{(C_m^{\ell-g*\dagger} + D_m^{\ell-g*\dagger})} \quad \text{where} \quad \ell \in \Omega_1.$$

Evaluating the flux of $E_z \bar{H}_x^{\dagger}$ across the surface A, defined to be unit cell at height $y = (s+h)^-$, leads to

$$I(A) = \int_0^c \int_0^{d'} E_z H_x^{\dagger} \, dxdz$$
$$\int_0^c \int_0^{d'} \sum_{m,\ell} (a_m^{\ell*} + b_m^{\ell*}) \, MEXU_m^{\ell} \sum_{n,t} \overline{(C_n^{t*\dagger} + D_n^{t*\dagger})} \, \overline{MEXU}_n^{t\dagger} \, dxdz$$

where $\ell \in \Omega_1$ and $t \in \Omega_2$.

Because,

$$MEXU_n^{t\dagger} = MEXU_n^{g+t}$$

and also $H = g + t$

(which is an element of Ω_1), then

$$I(A) = \int_0^c \int_0^{d'} \sum_{m,\ell} (a_m^{\ell *} + b_m^{\ell *})\, MEXU_m^{\ell} \sum_{n,H} (\overline{C_n^{H-g*\dagger} + D_n^{H-g*\dagger}})\, \overline{MEXU}_n^{H}\, dx\, dz$$

$$= \sum_{m,\ell} (a_m^{\ell *} + b_m^{\ell *})(\overline{C_m^{\ell-g*\dagger} + D_m^{\ell-g*\dagger}})$$

i.e. $I(A) = I(U)$.

Hence, the flux of $E_x \bar{H}_z^{\dagger}$ is analytically preserved. In a similar manner it may also be shown that the flux of $E_z \bar{H}_x^{\dagger}$ is preserved.

Thus, the Littrow expression is analytically satisfied if the field problem is solved using a truncated set of Rayleigh orders chosen to be elements of Ω_1. (That is, if the Rayleigh orders are chosen to be symmetric about the Littrow order.) However, if the set of Rayleigh orders chosen is at all "asymmetric" then there are terms in the expression of continuity which cannot be matched across this interface. The magnitude of this "overhang" gives us some feeling for the convergence of the field problem.

The following table presents some evidence of these assertions using results obtained from the computer program.

Table 8.3.2

Rayleigh Orders	0,-1,-1	0,-1,1,-2,2	0,-1,1,-2
Waveguide Modes	0,1,2	0,1,2	0,1,2
L.H.S. of Littrow Expression	3.83×10^{-3}	3.78×10^{-4}	5.96×10^{-11}

Demonstration of the properties of the Littrow Expression for a grating defined by $d/d' = 1.0$, $2c/d = 2c'/d = 0.9$, $S/d = 0.6$, $2h/d = 2h'/d = 0.3$, operated in a (-1,-1) Littrow configuration. The incident radiation was characterized by $\lambda/d = 1.2001$, $\phi = 58.0596^{o}$, $\psi = 45^{o}$ and $\delta = 0^{o}$. In each case, the same set of Rayleigh orders and waveguide modes were used to describe the diffracted fields in the two arrays.

8.3.3 Reciprocity

From [8.20], the form of the reciprocity theorem for a doubly periodic grating is $\beta_{00}(\underline{A} \cdot \underline{U}') = \beta'_{00}(\underline{A}' \cdot \underline{U})$. Here $\underline{U}$ is the vectorial amplitude of the electric field of the $(p,q)^{th}$ order of diffraction, excited by an incident field of vectorial amplitude $\underline{A}$. The symbols carrying a "prime" correspond to a second diffraction problem obtained by choosing a reflected order (p,q) of the first problem and returning an incident beam with the same wavelength λ along the direction of this propagating order ($\beta'_{00} = \beta_{pq}$). The polarization of the electromagnetic field may be chosen arbitrarily in both problems.

The following numerical example pertains to a grating characterized by $d = 1.0$, $d' = 1.2$, $2c = 0.25$, $2c' = 0.4$, $S = 0.4$, $2h = 2h' = 0.6$

and incident radiation at a wavelength of 1.1. (In the above, all lengths are expressed in arbitrary units.)

The first diffraction problem was defined by incident angles $\phi = 30^{o}$, $\psi = 35^{o}$, $\delta = 90^{o}$. The (-1,0) order was returned to give a second diffraction problem defined by the angles $\phi' = 48.46421^{o}$, $\psi' = -22.52765^{o}$. δ' was chosen to be 90^{o}. The (-1,-1) order was also returned to give a third diffraction problem defined by the angles $\phi' = 69.37018^{o}$, $\psi' = 42.37084^{o}$ and $\delta' = 90^{o}$.

Two tests were undertaken. The first used four waveguide modes and five Rayleigh orders to describe the diffracted fields in both arrays. Results for this test, obtained from our computer program, were

Problem 1:

$B_{-1,0} = (0.36188, 8.39001^{o})$

$C_{-1,0} = (0.06508, -123.06761^{o})$

$B_{-1,-1} = (0.02329, 78.68925^{o})$

$C_{-1,-1} = (0.00000, 108.52216^{o})$

E.T.* = 0.00281.

(* E.T. = Total Energy Transmission.)

Problem 2:

$B'_{-1,0} = (0.41715, 9.86503^{o})$

$C'_{-1,0} = (0.02551, 56.97048^{o})$

E.T. = 0.00454.

Problem 3:

$B'_{-1,-1} = (0.02116, 83.22531^{o})$

$C'_{-1,-1} = (0.00003, -124.52589^{o})$

E.T. = 0.53397.

The left and right hand sides of the reciprocity relation, corresponding to the return of the (-1,0) order were calculated to be (0.35183, 11.61073°) and (0.35127, 11.43950°) respectively. Results obtained from the return of the (-1,-1) order yielded values for the left and right hand sides of the reciprocity relation of (0.01732, 83.24818°) and (0.01721, 78.69184°) respectively.

The second test used five waveguide modes and a "field symmetric" selection of Rayleigh orders to describe the fields in both arrays. For example, the first problem used the Rayleigh orders (0, -1, 1, -2) for both arrays.

Results obtained for our computer program for this test were:

Problem 1:

$B_{-1,0}$ = (0.36502, 7.54186°)

$C_{-1,0}$ = (0.07047, -122.44265°)

$B_{-1,-1}$ = (0.02373, 81.81319°)

$C_{-1,-1}$ = (0.00000, 106.44642°)

E.T. = 0.00291.

Problem 2:

$B'_{-1,0}$ = (0.42049, 8.96581°)

$C'_{-1,0}$ = (0.02760, 57.55056°)

E.T. = 0.00465.

Problem 3:

$B'_{-1,-1}$ = (0.02143, 81.78939°)

$C'_{-1,-1}$ = (0.00003, -125.00659°)

E.T. = 0.54323.

In this case, the left and right hand sides of the reciprocity relation, corresponding to the return of the (-1,0) order were calculated to be (0.35512, 10.88171°) and (0.35512, 10.88171°) respectively, while those corresponding to the return of the (-1,-1) order were (0.01753, 81.81571°) and (0.01753, 81.81571°) respectively.

8.3.4 A critical discussion of the use of boundary conditions

It is the intention of this section to give some insight as to the consequences of misunderstanding the electromagnetic boundary conditions. This discussion is based on our experience in misusing the boundary conditions which culminated in meaningless physical conclusions. As such, it is felt worthwhile to provide some understanding of the logic which resulted in the derivation of erroneous answers in the hope that those who read this may learn from our experience and not waste valuable effort in duplicating our mistakes. It is our belief that we have profitted from these experiences in that we now have a more adequate comprehension of the physical significance of the boundary conditions. Thus, for this reason this section is devoted to one of the most commonly misunderstood aspects of classical electromagnetism.

Before commencing the detailed discussion, let us preface the analysis with some preliminary remarks concerning the edge conditions (outlined in section (8.2.2)). There, it was shown that for right angled metallic edges

(i) those components of the electric field which are normal to the edge have integrable singularities no worse than $O(r^{-1/3})$;

(ii) the tangential component of the electric field has no singularity and behaves as $O(r^{1/3})$;

(iii) the tangential component of the magnetic field is non-singular, behaving as $O(r^0)$;

(iv) tangential partial derivatives of field components have the same behaviour as their respective components;

(v) partial derivatives of tangential field components taken normal to the edge have integrable singularities of order $O(r^{-1/3})$; and

(vi) partial derivatives taken normal to the edge of field components themselves normal to the edge have non-integrable singularities and behave as $O(r^{-4/3})$.

From the boundary conditions, which are derived using Maxwell's equations, it is found that the field components E_x, E_z, H_x and H_z are continuous across the air-air interface separating the regions for which it is necessary to use different field expansions. Here it is to be noted that all of these field components have either no singularities or at worst singularities that can be removed upon integration.

Now, for any theory of singly periodic diffraction gratings concerned with a particular polarization, these boundary conditions reduce to specifying only

(i) the continuity of the appropriate field components across the interface, and

(ii) the continuity of its partial derivative with respect to the grating normal (i.e. $\partial/\partial y$).

Thus, it was with such previous knowledge that, in our first attempt at solving the diffraction problem of the crossed lamellar transmission grating, we invoked the continuity of E_x, E_z, $\partial E_x/\partial y$ and $\partial E_z/\partial y$. Such a

theory produced results not in accordance with conservation of energy which in turn led us to suspect the validity of the assumption involving the continuity of the two partial derivatives $\partial E_x/\partial y$ and $\partial E_z/\partial y$. After some close scrutiny of the edge conditions it became apparent that the construction of any crossed grating formalism involving the above assumption was doomed to failure since each of the partial derivatives $\partial E_x/\partial y$, $\partial E_z/\partial y$ possess a non-integrable singularity on half of the metal edges. In the case of the crossed lamellar transmission grating $\partial E_x/\partial y$ possesses non-integrable singularities on the edges of the upper grating (located at $y = s \pm h$), whereas $\partial E_z/\partial y$ has non-integrable singularities at the edges of the lower array (located at $y = -s \pm h'$).

Having identified the source of the discrepancy it became a simple matter to explain why this should produce results contrary to the conservation of energy. If H_z is expressed in the form

$$j\omega\mu_0 H_z = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}$$

then it may be seen that the assumed continuity of $\partial E_x/\partial y$ (at the interfaces located at $y = s \pm h$), together with the concomitant assumption that it should possess no non-integrable singularities, leads to the result that H_z has a non-integrable singularity. Clearly this is in disagreement with the boundary conditions derived from Maxwell's equations but, furthermore, results in the introduction of a non-integrable singularity into the expression for the normal component of the Poynting vector, S_y. Thus, any numerical treatment of the problem, although incapable of characterizing the nature of the singularity, is unstable and cannot physically ensure conservation of energy.

This incorrect interpretation of the boundary condition also led us to believe in the existence of a property which was termed "energy-decoupling". By this, it was meant that the energy associated with each Cartesian component of the field should be independently conserved. This is probably the most concise statement of the consequences of the misinterpretation of the boundary conditions referred to above. Stated in a mathematical form, it was believed that

$$\sum_{p,q\in\Omega} [\beta_{pq}(|B_{pq}|^2 + |\hat{B}_{pq}|^2)] = \beta_{00}|A_x|^2$$

for the x-component, and

$$\sum_{p,q\in\Omega} [\beta_{pq}(|C_{pq}|^2 + |\hat{C}_{pq}|^2)] = \beta_{00}|A_z|^2$$

for the z-component of the electric field where

$$\Omega = \{(p,q)|\beta_{pq} \varepsilon \mathbb{R}\} .$$

These conclusions were drawn largely on the basis of our experience with singly periodic lamellar gratings for which the spatial invariance of the energy flux $\int_0^d S_y \, dx$ is equivalent to the y-invariance of the integrals

(i) $\mathrm{Im}\{\int_0^d E_z \, \partial\overline{E}_z/\partial y \, dx\}$ for P polarized radiation, and

(ii) $\mathrm{Im}\{\int_0^d H_z \, \partial\overline{H}_z/\partial y \, dx\}$ for S polarized radiation.

These two expressions involve integrands which are either non-singular

or have at worst singularities which can be removed by integration. However, in the course of our analysis for S polarization, the fact that $\mathrm{Im}\{\int_0^d H_z\, \partial\overline{H}_z/\partial y\, dx\}$ and $\mathrm{Im}\{\int_0^d E_x\, \partial\overline{E}_x/\partial y\}$ were in no way equivalent quantities for singly periodic gratings (having the generating axis aligned with the field component H_z) was overlooked. For the case of P polarization, one can demonstrate that

$$\sum_{n\in\Omega_s} \beta_n[|C_n|^2 + |\hat{C}_n|^2] = \beta_0$$

where $\Omega_s = \{n|\mathrm{Im}(\beta_n) = 0\}$.

However, for S polarization it is incorrect to suggest that

$$\sum_{n\in\Omega_s} \beta_n[|B_n|^2 + |\hat{B}_n|^2] = \beta_0$$

since the integrand $E_x\, \partial\overline{E}_x/\partial y$ has a non-integrable singularity at the groove edges thereby destroying the y-invariance of $\int_0^d E_x\, \partial\overline{E}_x/\partial y\, dx$ in all space.

Now in making the transition from the singly periodic lamellar grating to the crossed lamellar transmission grating, one finds that

(i) $E_x\, \partial\overline{E}_x/\partial y$ has a non-integrable singularity on the edges of the upper array, and

(ii) $E_z\, \partial\overline{E}_z/\partial y$ has a non-integrable singularity on the edges of the lower array.

However, although it is not possible to ensure the physical invariance of

$$\text{Im} \iint_{\text{period}} E_x \frac{\partial \bar{E}_x}{\partial y}\, dA \quad \text{and} \quad \text{Im} \iint_{\text{period}} E_z \frac{\partial \bar{E}_z}{\partial y}\, dA$$

throughout the entire domain, it can be stated categorically that $\text{Im} \iint E_z\, \partial \bar{E}_z/\partial y\, dA$ is y-invariant everywhere in the space $y > -s + h'$, which leads to

$$\beta_{00}|A_z|^2 = \sum_{p,q\varepsilon\Omega} \beta_{pq}[|C_{pq}|^2 + |C^-_{pq}|^2 - |C^+_{pq}|^2]$$

$$+ \sum_{p,q\varepsilon\bar{\Omega}} 2|\beta_{pq}|\ \text{Im}[C^+_{pq}\ \bar{C}^-_{pq}]$$

and that $\text{Im} \iint E_x\, \partial E_x/\partial y\, dA$ has the same property everywhere in the space $y < s - h$ from which one derives

$$0 = \sum_{p,q\varepsilon\Omega} \beta_{pq}[-|\hat{B}_{pq}|^2 + |B^-_{pq}|^2 - |B^+_{pq}|^2]$$

$$+ \sum_{p,q\varepsilon\bar{\Omega}} 2|\beta_{pq}|\ \text{Im}[B^+_{pq}\ \bar{B}^-_{pq}].$$

Here, $\bar{\Omega}$ refers to the set of evanescent orders.
Furthermore, it is interesting to note that although neither $\text{Im} \iint E_x\, \partial \bar{E}_x/\partial y\, dA$, $\text{Im} \iint E_y\, \partial \bar{E}_y/\partial y\, dA$, nor $\text{Im} \iint E_z\, \partial \bar{E}_z/\partial y\, dA$ are invariant throughout all space, the quantity:

$$\text{Im}\ \{ \iint [E_x \frac{\partial \bar{E}_x}{\partial y} + E_y \frac{\partial \bar{E}_y}{\partial y} + E_z \frac{\partial \bar{E}_z}{\partial y}]\, dA \}$$

is independent of the variable y. Such a relation is assured both physically and analytically by the energy conservation criterion.

In summary, it is hoped that this section has provided some conception of the consequences of misunderstanding the electromagnetic

boundary conditions and, moreover, given an insight into what is probably the most deceptive aspect of classical electromagnetism. It is felt that making the transition from the study of singly periodic lamellar gratings to crossed lamellar gratings has led us to a more intimate understanding of some of the fundamental concepts of electromagnetism and, more importantly, to regard with due care certain aspects of the subject which were hitherto dismissed as being "fait accompli".

8.4 SPECTRAL PROPERTIES AND SOLAR SELECTIVITY

8.4.1 Spectral characteristics

With the aid of the theory discussed in section (8.2), the spectral characteristics of this type of bi-grating (whose periods are comparable to the wavelength) have been extensively studied. In figure (8.4.1) is shown a spectrum obtained for the crossed lamellar transmission grating. This spectrum, typical of this class of gratings when used in normally incident radiation, can be divided into three characteristic regions:

(i) the transmission region $\lambda/d \leq 1.0$,

(ii) the transition region $1.0 \leq \lambda/d \leq 2.5$,

(iii) the long wavelength filtering region $\lambda/d \geq 2.5$.

(The dotted curve gives the energy transmitted through the grating by the (0,0) order, while the solid curve designates the total energy transmitted through the grid.)

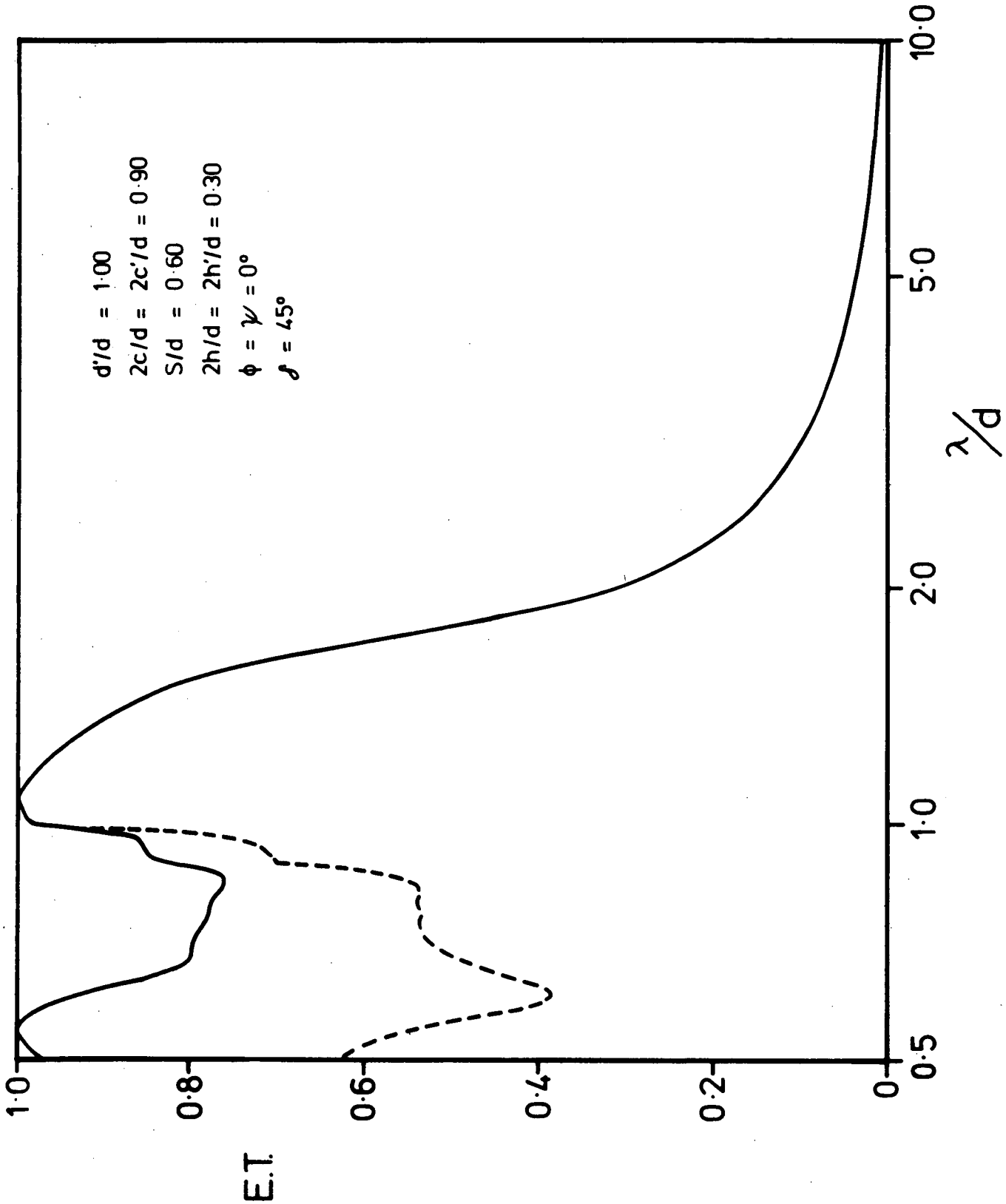


Figure 8.4.1

A typical spectrum of a crossed lamellar transmission grating. The broken curve denotes the transmittance of the order (0,0). (E.T. denotes the transmittance.)

The first two regions are separated by what is termed the $(0,0)^{th}$ Wood anomaly. Such anomalies are due to the passing off of spectral orders and the subsequent redistribution of energy in the remaining propagating orders. In the case under consideration, five orders in both reflection and transmission propagate immediately below the Wood anomaly whereas only the specularly reflected and transmitted orders propagate above it. If the radiation is not normally incident upon the grating, other anomalies occur in the transmission region when higher orders begin or cease to propagate. These anomalies are of less significance as they usually appear only as small variations in efficiency and thus have little effect on the mean transmission value.

Most of the energy is reflected in the long wavelength filtering region. High reflectance can be achieved only if a strong and unperturbed current distribution can be established in the surface metal of the arrays. The magnitude of this reflection is largely dependent upon the current and as such is only weakly dependent upon the mark-space ratio. This is indeed true for the crossed lamellar transmission grating since either or both of the spatially separated arrays can contribute to behaviour characteristic of a singly periodic lamellar transmission grating used in P polarized radiation. The currents induced in the surfaces of the arrays are very strong for such gratings since the flow, or component of the flow, can occur along the length of the grooves in either or both arrays. (That is, the flow is uninterrupted in its propagation by any irregularities such as edges.)

It should be noted that in all efficiency curves the abscissae represent the normalized wavelength (i.e. the ratio of the wavelength of the radiation to the grating period). This is an appropriate choice

since there is no variation in conductivity (in accordance with the assumption of infinite conductivity) and thus no absolute wavelength dependence. Furthermore, the grating equations reveal that it is the normalized quantity λ/d and not the absolute quantity λ which determines the direction of propagation of any order. Hence, predictions made from the curves will apply equally well to any wavelength range by simply rescaling the grating period accordingly. The aperture widths, groove thicknesses and array separation are normalized to the period in a similar manner.

8.4.2 Application in solar selective systems

Since approximately 50% of the incident solar flux is either reflected or absorbed within the atmosphere, it is imperative that the collection efficiency be maximised if solar power is to become a viable source of energy in the future. Although convection and conduction losses may be significantly reduced by thermal insulation, the third loss mechanism, radiation, is more difficult to control. In view of these remarks, it is the aim of this section to discuss the design of the crossed lamellar transmission grating in relation to its viability as a solar selective surface. The system under consideration is assumed to consist of a lossless filter (i.e. the grating) placed in front of a perfect black body which absorbs all the energy transmitted by the grating. The black body is heated by the solar radiation and re-emits infrared radiation (characteristic of the temperature of the absorber) isotropically through the hemisphere of all possible directions. Thus, the function of the grating is to transmit as much as possible of the incident flux while inhibiting the escape of absorber's re-emitted radiation.

In the following subsections, the optimization of the grating parameters with respect to these criteria is discussed. The parameters to be considered are the groove thickness, the aperture width, the array separation and the effect of grating symmetry. (Note that the grating is defined to possess square symmetry if d = d', c = c'.) Finally, the effect of altering the angle of incidence of the solar radiation on the transmitted energy is discussed.

Initially, incident radiation normal to the grating is considered. To find the transmission spectrum for isotropic radiation, it is necessary to integrate over curves corresponding to incident energy arriving from a range of directions (ψ).

8.4.2.1 *The effect of grating symmetry*

Incident solar radiation is largely unpolarized and since wide bandwidth selectivity is required, the grating should not act as a polarizing filter, at least in the case of normal incidence. Hence, the grating parameters in the OX and OZ directions should be identical (d = d', c = c').

It has been found that gratings having square symmetry have a transmittance which does not greatly depend upon the polarization of normally incident radiation. Figure (8.4.2a) shows that the transmission properties of gratings with departures from square symmetry are dependent upon the polarization of the incident beam.

(Curves shown in all the following subsections will possess square symmetry.)

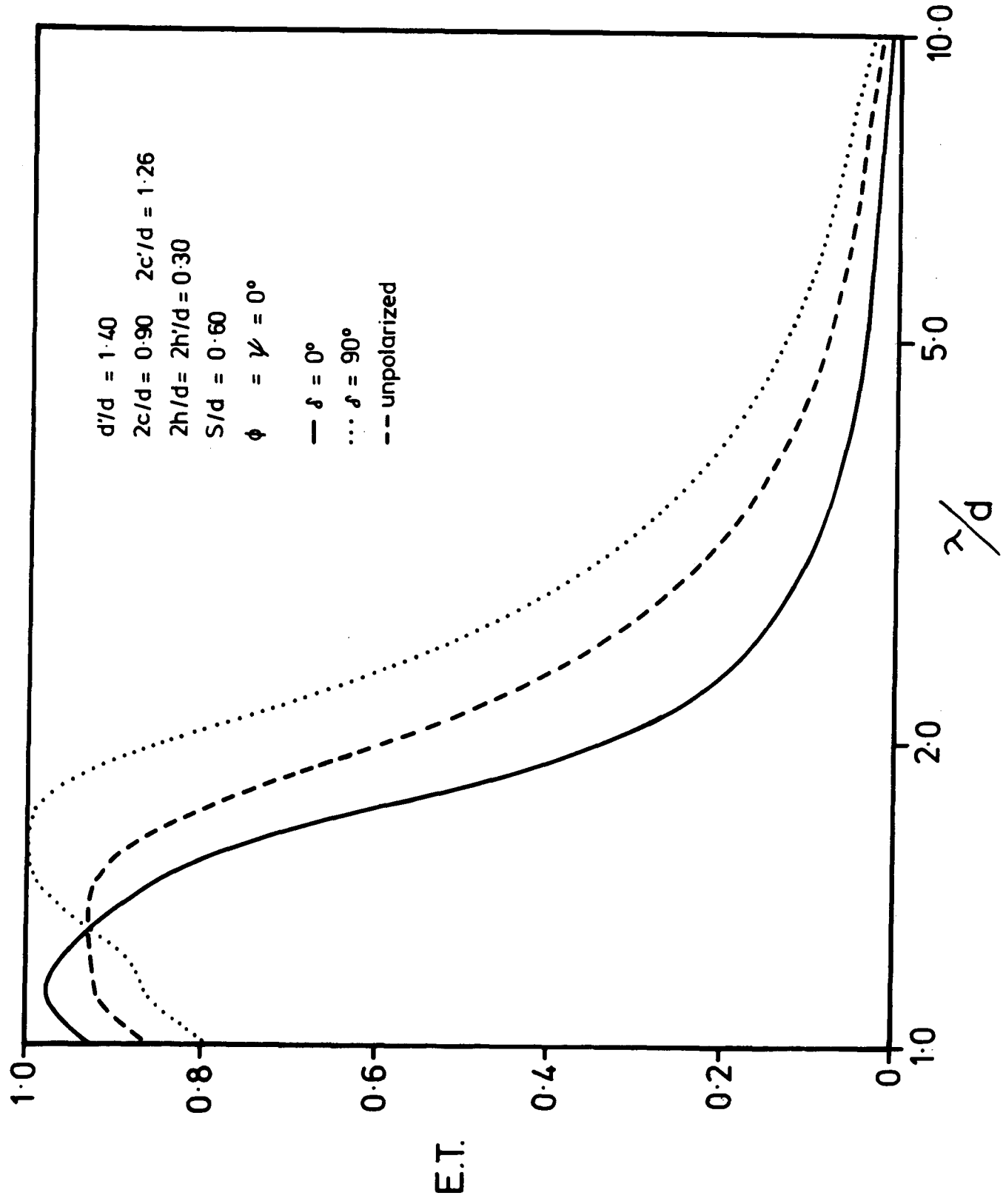


Figure 8.4.2a

The effect of grating symmetry on the transmission properties.

8.4.2.2 *The grating apertures*

The aperture width largely determines the transmittance of a grating. The curves shown in figure (8.4.2b) demonstrate that the wider grating apertures produce greater transmission in the transition region, the same result being observed in the transmission region.

The most appropriate aperture width for solar selective purposes is chosen such that the energy transmission is maximised and thus one deduces that the metal walls between apertures should be as thin as practicable.

8.4.2.3 *The groove thickness*

An examination of the curves in figure (8.4.2c) shows that the groove thickness can be increased to tune the long wavelength filtering tail without substantial effect on the transmission peak at shorter wavelengths. The deeper grooves produce sharper decay from the transmission peak and greater rejection of the infrared wavelengths. This effect is due to the guiding of waves through the grooves.

Long wavelength filtering action does take place only because of the ability of the structure to establish strong uninterrupted flows of surface current in either or both arrays. This is evidenced by the fact that filtering does take place for arrays of infinitesimally thin grooves. Furthermore, such effects are only weakly dependent upon the mark-space ratio of the grooves. However, a more interesting discussion is involved in explaining why filtering action improves with increasing groove depth. For wavelengths in excess of twice the aperture width (i.e. $\lambda > 4c$) all the waveguide modes are evanescent (i.e. ν_m^q is purely imaginary for all values of m and q). As such, the coupling between the tops and the bottoms of the grooves of both arrays decreases with

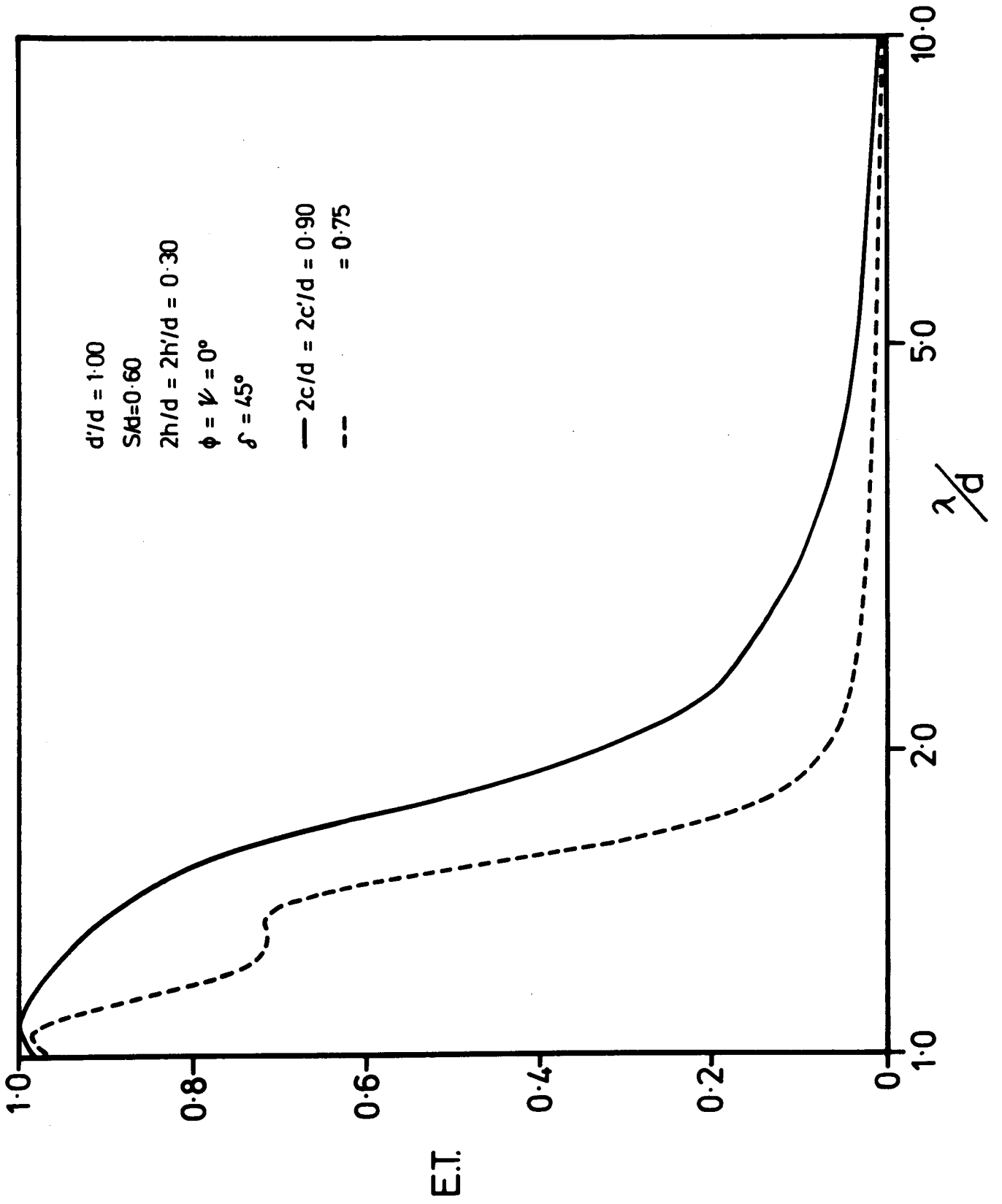


Figure 8.4.2b The transmittance as a function of wavelength for various grating apertures.

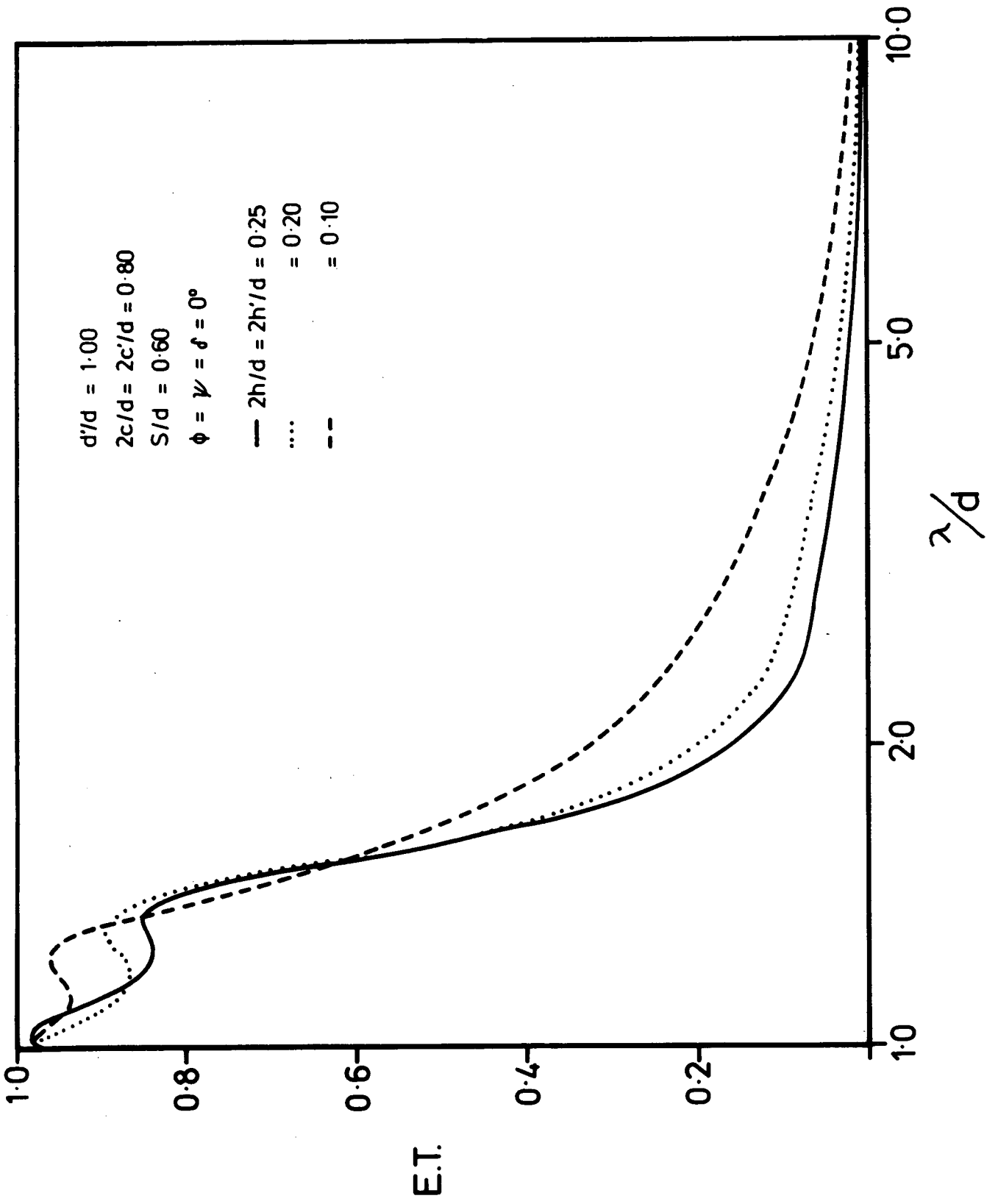


Figure 8.4.2c The effect of groove thickness on the transmittance.

increasing depth. In the limit as h tends to infinity, the coupling vanishes entirely and thus one expects no transmission to be exhibited by the grid after the cut-off wavelength. (For the structure in question, the cut-off wavelength may readily be seen (from figure (8.4.2c)) to be $\lambda/d = 1.6$.) It is also possible to identify the root of this phenomenon with the formalism itself. It may be shown that the y-symmetric and y-antisymmetric modal coefficients become identical in magnitude and differ in phase by $\pi/2$, thereby reducing the total downward flux of energy to zero.

It appears that the optimal groove specification for a grating used as a solar selective surface would be deep grooves with the thinnest possible metal walls. However, it is conceivable that gratings with these dimensions would be extremely difficult and costly to manufacture and it is these considerations which led us to consider gratings with thin walls ($2c/d \simeq 0.9$) and moderately deep grooves ($2h/d \lesssim 0.5$) in assessing the absorptance values and a/e ratios in section (8.4.3.3).

8.4.2.4 *The array separation*

The curves in figures (8.4.2d), obtained by varying the array separation, demonstrate some interesting properties of the crossed lamellar transmission grating. These transmission resonances are similar to those encountered in Fabry-Perot interferometers and have been observed by Casey and Lewis [8.6-7] when studying a system of two parallel wire gratings. Resonances of this type, caused by constructive interference occur at wavelengths defined by

$$\lambda \simeq 2 \times [\text{optical distance}]/n \qquad n = 1, 2, 3, \ldots.$$

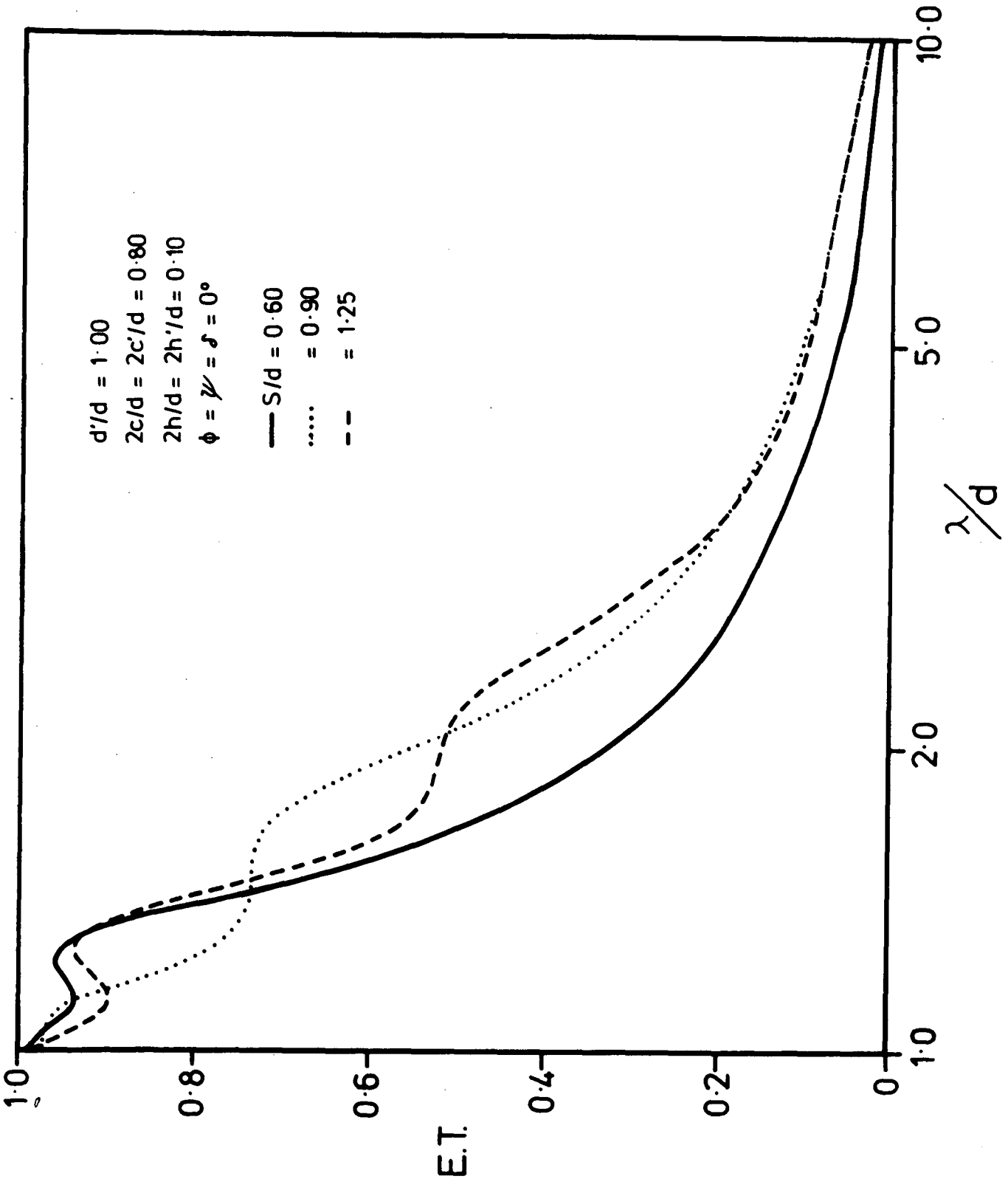


Figure 8.4.2d The transmission properties of the crossed lamellar transmission grating for several array separations.

Hence, for a grating used in normally incident radiation with an array separation S/d = 1.25, one expects a resonance anomaly to occur at wavelengths of $\lambda/d \simeq 2.5$, 1.25, etc. Figure (8.4.2d) confirms that resonances do occur in these regions. Similarly, anomalies occurring in the region of $\lambda/d \simeq 1.8$ for S/d = 0.9 and $\lambda/d \simeq 1.2$ for S/d = 0.6 are exhibited.

Thus, one is lead to deduce that normalized separations in the range 0.5 to 0.6 would be required for gratings used in solar selective systems, since these separations would produce resonance anomalies in the transition region and consequently enhance the transmission there.

8.4.2.5 *Angle of incidence*

The effects of altering the angle of incidence, ϕ, of the incident radiation are illustrated in figures (8.4.2e(a)-(f)). (The broken curve depicts the transmission characteristics of the (0,0) order.) These show that the energy transmitted through the grid decreases as the angle of incidence increases. This is to be expected since, as ϕ tends to 90°, the incident wave propagates almost parallel to the grating surface and cannot adequately couple to the modal field to produce a significant transmitted component of energy. One also notices that the Wood anomaly at $\lambda = 1.0$ for $\phi = 0^{\circ}$ moves to longer wavelengths as ϕ increases. Other localized variations correspond to higher orders becoming evanescent.

From these curves, it appears that such a grating used as a solar selective surface would need to be orientated so that it was always normal to the impinging radiation. Should the sun subtend too large an

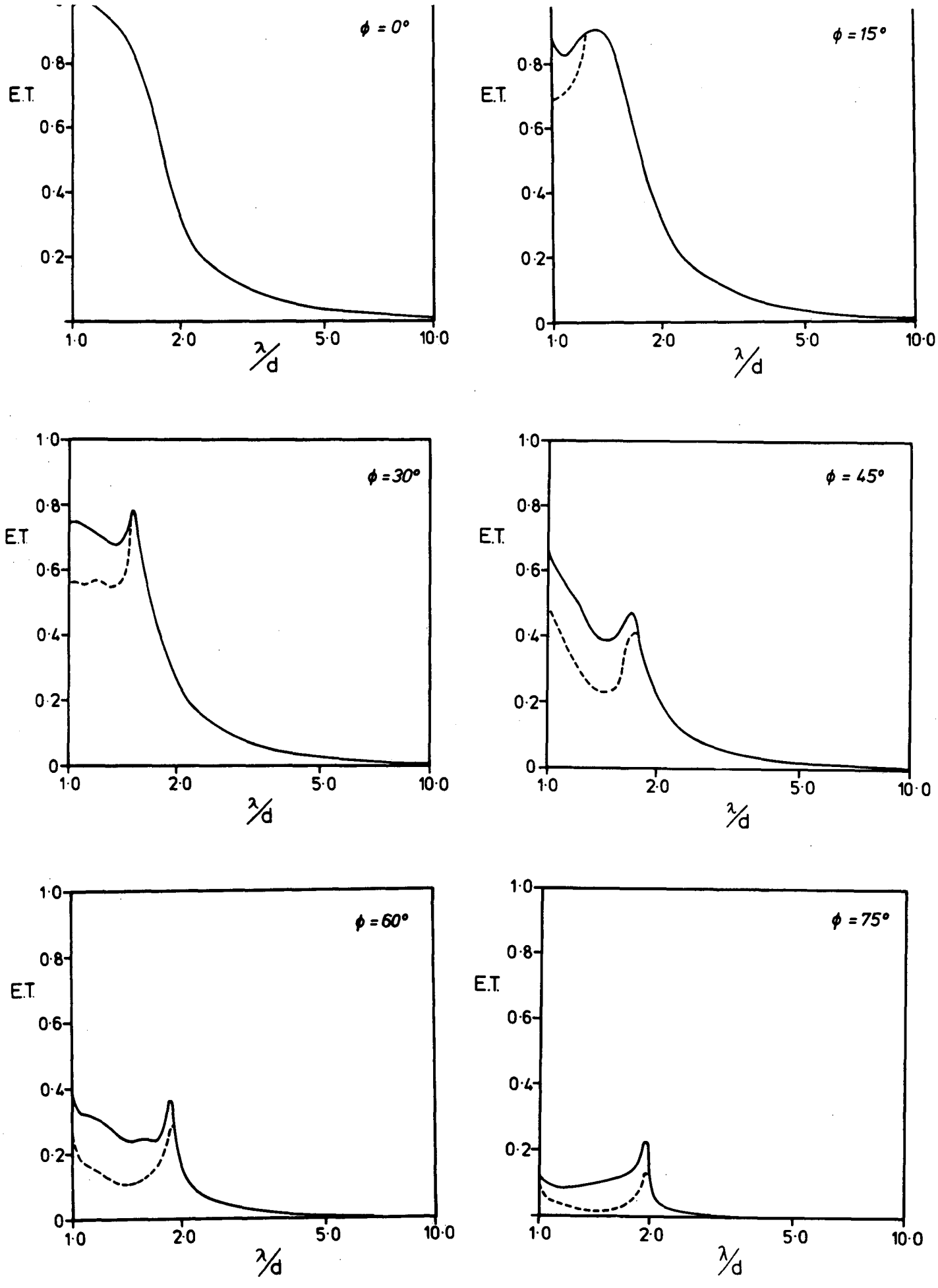


Figure 8.4.2e The effect of angle of incidence on the transmittance of a grating having square symmetry, $2c/d = 0.9$, $S/d = 0.6$, $2h/d = 2h'/d = 0.3$ and $\psi = 0^{o}$.

angle to the grating normal, the short wavelengths will be reflected thereby reducing the integrated absorptance.

8.4.2.6 *Optical geometry*

In summary, one concludes that optimal filtering and passband characteristics for solar selective purposes are obtained for a grating possessing the following properties:

(1) Square symmetry

(2) Thin groove walls

(3) Deep grooves

(4) Normalized array separations in the range 0.5 to 0.6

and (5) Operated in almost normally incident radiation.

8.4.3 Absorptance and emittance

Firstly, the procedures for obtaining the integrated solar absorptance (a) and emittance (e) are to be reviewed, followed by a presentation of some results obtained for this structure. (The review, which has been taken from the report of McPhedran and Maystre [8.20], is included largely because of the desire to present a complete record of the studies.)

8.4.3.1 *The integrated solar absorptance*

Once the grating parameters have been chosen, two factors remain which are of importance in determining the energy transmission through the grating. They are the wavelength spectrum of the incident radiation, which in this case is approximated by the black body distribution characteristic of a temperature of 6000 K, and the angular distribution of the incident radiation. There are two main classes of

incident radiation, namely direct radiation, which is that part of the solar spectrum arriving on Earth from the direction of the solar disc, and diffuse radiation, which is radiation scattered to the ground (by air molecules, dust and clouds) in a semi-isotropic manner from all parts of the sky. The following three cases will be considered in describing the angular distribution function:

(1) The grating is always aligned orthogonal to the incident solar radiation and is illuminated only by direct radiation (i.e. the angular distribution then being described by a Dirac delta function).

(2) The grating is stationary and illuminated by direct solar radiation, the angular distribution being characterized by the Lambertian cosine law ($\cos \phi$).

(3) The grating is stationary and illuminated by totally diffuse solar radiation. (The Lambertian cosine law and dependence upon declination of solid angle imply that the angular distribution function will be described by $2 \sin \phi \cos \phi$.)

In general, the absorptivity will depend on the angles ϕ and ψ, so an integration over ψ values for constant ϕ and λ/d must firstly be performed to find the mean transmission value using an isotropic weighting of ψ of $1/(2\pi)$. Only values of $\psi \in (0,\pi/4)$ and $\phi \in (0,\pi/2)$ need be considered because of symmetry. Then at constant ϕ, the ψ mean transmissions are averaged over λ values, for a chosen grating period using the Planck distribution as the weighting function.

The radiant emittance $W(\lambda,T)$ of a black body at temperature T is given by

$$W(\lambda,T)\, d\lambda = \frac{C_1 \lambda^{-5}}{\exp(C_2/\lambda T) - 1}\, d\lambda$$

where $C_1 = 2\pi hc^2$, $C_2 = hc/K_B$, and h is Planck's constant, c is the speed of light in free space, and K_B is Boltzmann's constant.

These $(\psi,\lambda/d)$ mean transmissions are then averaged over ϕ with the appropriate weighting functions $\delta(\phi)$, cos ϕ and sin 2ϕ introduced above.

8.4.3.2 *The hemispherical emittance*

This quantity is a measure of the fraction of radiation, emitted isotropically by the black body absorber, which escapes through the grating. The angular distribution of the emitted radiation can be described by the normalized function sin(2ϕ). If the black body is assumed to have an equilibrium temperature of 700 K, the peak of the black body curve will be at a value given by the Wien Displacement Law (λ_{max} T = 0.2898 cm K) of λ = 4.14 μm.

The calculation proceeds in stages similar to those in section (8.4.3.1) with the following modifications. The second stage, involving integration of ψ mean transmission values over λ values, now has the black body function of a body at the temperature of the absorber ($\simeq$ 700 K) as a weighting factor. The third stage averages the $(\psi,\lambda/d)$ mean transmission values for $\phi \in (0,\pi/2)$ with a weighting function of the angular distribution sin(2ϕ).

8.4.3.3 *Quantitative results*

An "a/e" calculation was performed for the grating referred to in figure (8.4.1) placed in front of a black body absorber which is assumed

Table 8.4.1

d(μm)	Case 1 a(0,0)%	Case 1 a%	Case 2 a(0,0)%	Case 2 a%	Case 3 a(0,0)%	Case 3 a%	e%
0.5	58	64	25	32	20	26	0.6
0.6	61	71	28	35	23	30	0.8
0.7	61	75	29	38	24	32	1.0
0.8	61	78	30	40	25	34	1.2
0.9	61	79	31	41	26	35	1.5
1.0	61	81	31	42	26	36	1.8
1.5	59	84	32	44	28	38	4.6
2.0	52	86	32	45	28	39	9.2

Absorptance and emittance as a function of the period for a grating having square symmetry, $2c/d = 0.9$, $S/d = 0.6$, $2h/d = 2h'/d' = 0.3$.

to reach an equilibrium temperature of 700 K. Table (8.4.1) lists the absorptivities, a(0,0) for energy passing undeviated through the grating and a for all energy passing through the grating, and emittance e as a function of grating period for the three cases introduced in section (8.4.3.1), i.e.,

(1) a steered grating in direct illumination,

(2) an unsteered grating in direct illumination, and

(3) an unsteered grating in diffuse illumination.

From the accompanying table it can be seen that the grating periods for the most efficient solar selectivity lie in the range 0.8 - 1.0 μm. These results show that diurnal tracking, direct illumination and absorption of all transmitted radiation would be required to give both high values of a and a/e.

Under these conditions, an absorptance of around 80% is possible but in the absence of tracking or in diffuse illumination this value is approximately halved. The emittance, in the range 1 - 2%, was derived for a perfectly conducting grating. However, this assumption is not necessarily valid for many metal coatings used in the visible and conductivity effects would probably raise this value of emittance to 3 - 4%.

8.5 CONCLUSIONS

In this chapter, a new formalism describing the diffraction of a plane wave by an infinitely conducting, crossed lamellar transmission grating has been presented. The theory was formulated by application of the method of moments necessitating the solution of eight series of matrix equations for the unknown modal coefficients. Although the

formalism is appropriate for any wavelength range, the numerical implementation has been restricted to normalized wavelengths greater than 0.5 due to limitations imposed by currently available computing resources. For shorter wavelengths, the series representations of the diffracted field do not converge rapidly, since too many Rayleigh orders are required to specify both the propagating orders and the evanescent field.

Classical energy conservation and reciprocity tests have been used to confirm the validity and physical accuracy of the method. However, as mentioned in section (8.3), energy conservation cannot be used as a means of testing the physical significance of the algorithm since it is analytically satisfied by the grid formalism independently of truncation errors. Instead, it should only be taken as a check of the self-consistency of the method. Thus, a new amplitude constraint for crossed grating structures operated in a Littrow mount has been derived and this has proved itself to be a useful check of the convergence of the formalism.

The crossed lamellar transmission grating has been shown to have great potential as a solar selective surface, capable of achieving an absorptance of approximately 35% and a/e ratios in the range of 15 to 20 when used in a stationary mount and illuminated by totally diffuse illumination. Higher ratios of the order of 30 to 40 and absorptances of about 80% have been shown to be attainable if the grating is used in direct radiation and employs diurnal tracking. The optimal design of a grating for such purposes has been shown to have two identical arrays having deep grooves, thin metal walls and a normalized array separation

of approximately 0.6. The value of the latter parameter is chosen so that a Fabry-Perot resonance is located at approximately $\lambda/d = 1.2$, thereby broadening the transmission bandwidth of the structure.

In summary, one must emphasize that the assumption of a lossless structure has been made throughout. Although entirely valid for metals in the infrared, it may not be a justifiable assumption in the visible. Thus it will be necessary to derive a new class of modal theories capable of handling finitely conducting singly periodic structures before any further advances in the realm of grid formalisms are conceivable.

REFERENCES

[8.1] J. L. Adams, L. C. Botten and R. C. McPhedran, J. Optics (Paris), to appear.

[8.2] J. L. Adams and L. C. Botten, (1977), University of Tasmania, Report DGRG 77/3.

[8.3] A. Esau, E. Ahrens and W. Kebbel, Hoch. Elekt. Akust, 53 (1939) 113.

[8.4] W. Wessel, Hoch. Elekt. Akust, 54 (1939) 62.

[8.5] W. Franz, Z. Agnew, Phys., 1 (1949) 416.

[8.6] E. A. Lewis and J. P. Casey, J. Opt. Soc. Amer., 41 (1951) 360.

[8.7] J. P. Casey and E. A. Lewis, J. Opt. Soc. Amer., 42 (1952) 971.

[8.8] A. Mitsuishi, Y. Otsuka, S. Fujita and H. Yoshinaga, Jap. J. Appl. Phys., 2 (1963) 574.

[8.9] K. F. Renk and L. Genzel, Appl. Optics, 1 (1962) 643.

[8.10] R. Ulrich, K. F. Renk and L. Genzel, I.E.E.E. Trans. MTT, 11 (1963) 363.

[8.11] R. Ulrich, Infrared Physics, 7 (1967) 37.

[8.12] R. Ulrich, Appl. Optics, 7 (1968) 1987.

[8.13] R. Ulrich, Appl. Optics, 8 (1969) 319.

[8.14] A. I. Adodina, Y. S. Komissarov and V. A. Pavlyuk, Radio Engineering and Electronic Physics (U.S.A.), 14 (1969) 946.

[8.15] A. I. Adodina, A. M. Andrusenko, Y. S. Komissarov and V. A. Pavlyuk, Radiophysics and Quantum Electronics, 12 (1969) 1213.

[8.16] C. M. Horwitz, Optics Commun., 11 (1974) 210.

[8.17] C-C. Chen, I.E.E.E. MTT, 18 (1970) 627.

[8.18] C-C. Chen, I.E.E.E. MTT, 19 (1971) 475.

[8.19] C-C. Chen, I.E.E.E. MTT, 21 (1973) 1.

[8.20] R. C. McPhedran and D. Maystre, (1976), University of Sydney, Report SP76/1.

[8.21] R. C. McPhedran and D. Maystre, Appl. Phys., 14 (1977) 1.

[8.22] H. Blok and G. Mur, Appl. Sci. Res., 26 (1972) 389.

[8.23] J. L. Adams, (1976), University of Tasmania, Honours Thesis.

[8.24] D. S. Jones, "Theory of Electromagnetism", Pergamon Press, (1964), New York.

[8.25] J. Meixner, Ann. Phys. (6), 6 (1949) 1.

[8.26] A. Wirgin and R. Deleuil, J. Opt. Soc. Amer., 49 (1969) 1348.

[8.27] A. Wirgin, C. R. Acad. Sci. Paris, 262B (1966) 870.

[8.28] P. Morse and H. Feshbach, "Methods of Theoretical Physics", McGraw-Hill, New York (1963), p. 823.

CHAPTER 9

THE DOUBLE GRATING

The investigations reported in this chapter were conducted jointly by the author and Miss Jenny Adams, whose stirling efforts and stimulating collaboration are very much appreciated. Throughout all stages of this work, the author was privileged to have enjoyed numerous illuminating discussions held with Dr. Graham Derrick and Dr. Ross McPhedran. In particular, they are to be thanked for their valuable guidance in the studies which are presented in section (9.4).

9.1 INTRODUCTION

The discussion in the previous chapter centred on the spectral properties of an inductive grid whose two mutually orthogonal axes of periodicity lay in spatially separated planes. Following upon the success of these studies, my co-worker, Miss Jenny Adams [9.1] then generalized the theory to cope with two lamellar arrays inclined with respect to one another at some arbitrary angle η. The material contained in this chapter was largely motivated by the interesting case arising when η tend to zero. This structure, which is referred to as a double grating, has been the subject of much discussion in the literature over many years.

In 1951, Lewis and Casey [9.2] discussed the use of double grids and double gratings as microwave interference filters. Further work [9.3] by these same authors showed that a double wire grating (operated

in P polarized light) could successfully replace the highly conducting thin films which were conventionally used as the reflecting elements in long wavelength Fabry-Perot interferometers. They demonstrated that the transmission properties of the grating interferometer were far superior to those of the conventional design.

It was not until 1972 that the first rigorous theory of double gratings was published. In that year, Blok and Mur [9.4] considered a double grating composed of perfectly conducting laminae having infinitesimal thickness and derived an integral formalism in their treatment of the problem. Unfortunately, their paper contained very few numerical results and no discussion of the interferometric aspects of the structure.

In this chapter, the results of our study, believed to be the first rigorous theoretical investigation of the properties of grating interferometers is presented. The modal formalism discussed in section (9.2) pertains to a double grating constructed from a pair of singly periodic lamellar transmission gratings (each of arbitrary thickness). The analysis is sufficiently general to accommodate structures composed of lamellar arrays having different periods (provided that the ratio of the two periods is a rational number) and displaced with respect to one another.

The following section is concerned with the interferometric applications of the double grating. Since the design of an interferometer having a high resolving power is dependent upon the use of highly reflecting elements, it is necessary to operate the grating only in P polarized light. It is pleasing to report that in our studies to date, interferometers having very high finesses may be constructed using

deep lamellar arrays having relatively narrow apertures.

In those cases for which the incident radiation is unpolarized, it is essential that a double grid be utilized. Nevertheless, it is felt that the results presented herein would still apply to the double grid. At the time of the preparation of the thesis, this study was still in progress (and in particular an investigation concerning the diffraction properties of the double grid). As such, this chapter contains only a discussion of the features essential to the design of an interferometer. However, it is hoped that in the near future the completed studies can be reported in some detail.

During these investigations, a number of interesting analytic properties, analogous to the conservation relations of nuclear scattering theory, manifested themselves in the results of our computer programs. Thus, in an attempt to understand these properties which are heavily dependent upon the symmetry of the grating, an intensive theoretical investigation (discussed in section (9.4)) was instigated. The search for their origin was conducted on two fronts:

(i) using the conventional integral techniques (which have been discussed in chapters 4 and 8), and

(ii) using the principle of time reversal.

Perhaps the most interesting feature of the results concerns the existence of a detailed energy balance in the mode structure of the upper and lower arrays for:

(i) double gratings possessing both left-right and up-down symmetries operated

  (a) with arbitrary incidence parameters such that only the zero order was in propagation and

(b) in a first order Littrow configuration,

and

(ii) identical lamellar transmission gratings displaced relative to one another and operated in a normal incidence mounting such that the normalized wavelength exceeded unity.

9.2 THE THEORETICAL FORMALISM

9.2.1 Specification of the problem

In this section, only the formalism pertaining to the diffraction of a P polarized wave is presented, largely in the desire not to pre-empt the work of my co-worker in these investigations. Nevertheless, since the main thrust of this chapter lies in the use of the double grating as a Fabry-Perot interferometer, this action would appear to be vindicated.

Consider a P polarized plane wave of wavelength λ impinging upon the structure shown in figure (9.2.1) at some angle θ. The structure shown in the accompanying diagram is composed of a pair of singly periodic lamellar transmission gratings arranged such that the Oy axis is perpendicular to the plane of the grating and the Oz axis is aligned with the grooves of both arrays.

The apertures in the upper array, of width c and depth h, are spaced periodically (in the x-direction), a distance of d apart. Similarly, the geometry of the lower array is specified by the analogous parameters c', h' and d'. The two gratings are separated from one another by a distance $S = (2s - \frac{1}{2}h - \frac{1}{2}h')$ and are displaced relative to one another by an amount δx.

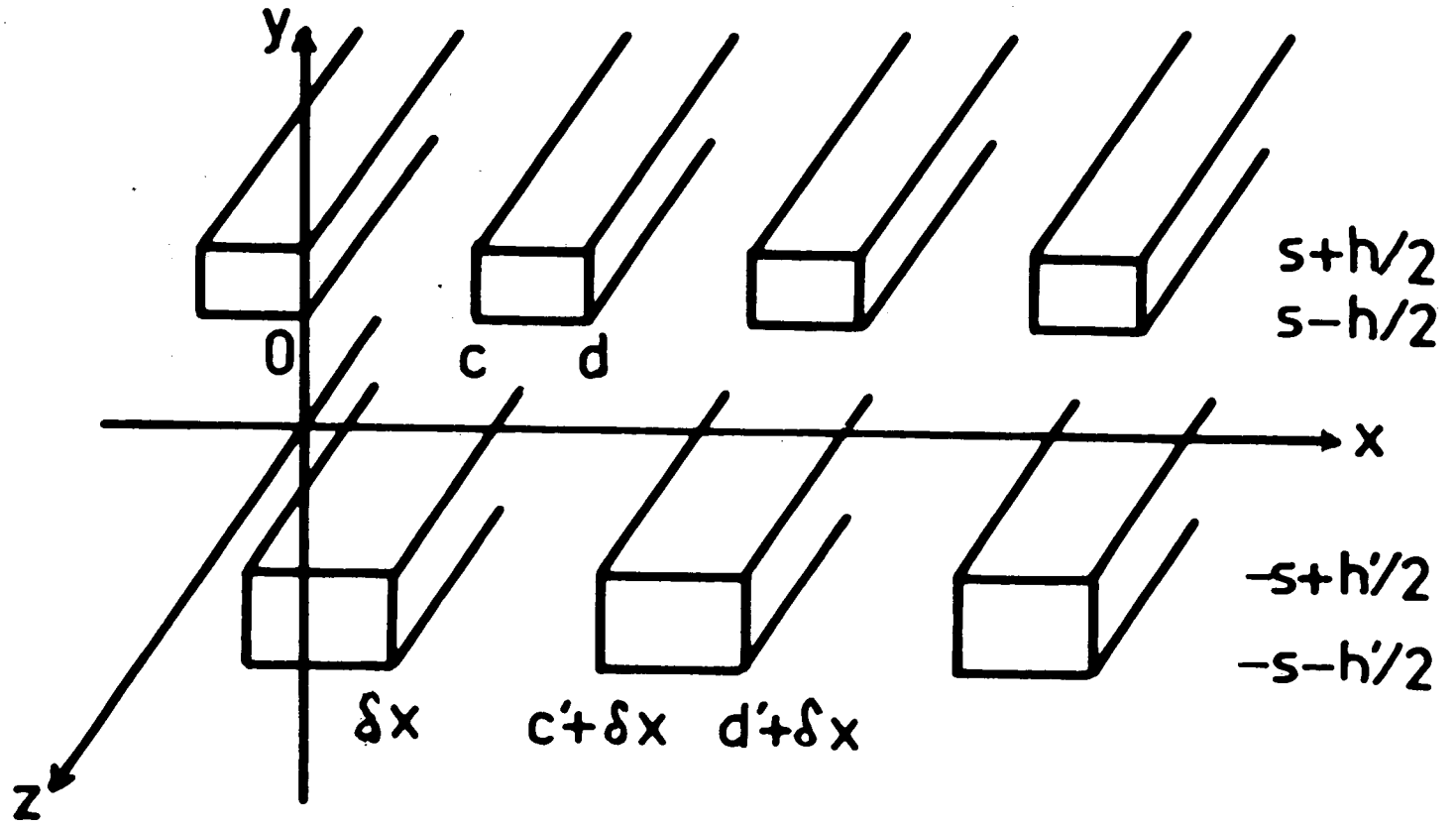


Figure 9.2.1 The geometry of the double grating.

The double grating is then assigned an overall period of d_T such that d_T is the "lowest common multiple" of d and d'. It is interesting to note that if the ratio of d to d' is irrational then d_T becomes infinite. In this case, an unlimited number of "real" orders may propagate giving rise to a continuous spectrum of plane waves.

Now, the incident field $\underline{E}^i$ given by

$$\underline{E}^i = \exp[jk_0(\alpha_0 x - \beta_0 y)]\ \hat{\underline{z}} \tag{9.2.1}$$

where $k_0 = 2\pi/\lambda$

$\alpha_0 = \sin\theta$

and $\beta_0 = \cos\theta$

excites a surface current (in the skin of the perfectly conducting prisms) along the axis of the grating. This in turn gives rise to a diffracted field of the same polarization. In the region $y \geq (s+\frac{1}{2}h)$, the diffracted field may be expressed as a superposition of upward going plane waves of the form

$$R_p(x,y) = \exp[jk_0(\alpha_p x + \beta_p y)] \tag{9.2.2}$$

where
$$\alpha_p = \alpha_0 + p\lambda/d_T \tag{9.2.3}$$

$$\beta_p = \sqrt{1 - \alpha_p^2} \quad \text{if} \quad |\alpha_p| \leq 1$$

$$= j\sqrt{\alpha_p^2 - 1} \quad \text{if} \quad |\alpha_p| > 1\ .$$

On the other hand, in the space $y \leq -(s+\frac{1}{2}h')$, the diffracted (and hence

total) field is specified by an expansion of downward going plane waves of the type

$$\hat{R}_p(x,y) = \exp[jk_0(\alpha_p x - \beta_p y)] \qquad (9.2.4)$$

In the domain $(-s+\frac{1}{2}h') \leq y \leq (s-\frac{1}{2}h)$ the field must be expressed as a superposition of both upward and downward going plane waves (R_p and $\hat{R}_p$).

9.2.2 Boundary conditions

It is not possible to define a single expansion representing the field that is everywhere convergent throughout space since the grating divides the domain into a number of discrete regions. Although Rayleigh expansions (as mentioned above) may be used to represent the fields in the "open" regions (i.e. unbounded in the x-direction), they are incapable of characterizing the field behaviour within the grooves of the arrays. This concept, known as the Rayleigh assumption, has been the subject of much debate in the literature for many years and several notable numerical studies (see for example references [9.5, 9.6]) disputing its validity have been published. Thus, within the apertures, one chooses to represent the field quantities in terms of waveguide expansions, each distinct mode of which obeys the appropriate boundary conditions. Conditions of continuity are then applied at all grating surfaces to match the alternative expansions (where their domains of convergence overlap) and the method of moments used to solve the resulting field problem.

As in the case of the crossed lamellar transmission grating one chooses to impose the continuity of E_x, E_z, H_x and H_z across the free

space-aperture interfaces. However, because of the chosen polarization, both H_z and E_x are trivial. Furthermore, the continuity of H_x reduces to the continuity of $\partial E_z/\partial y$. In addition to this, E_z vanishes on the metal surfaces but $\partial E_z/\partial n$ is not continuous across a metal interface (due to the excitation of a surface current in the skin of the metal).

9.2.3 Rayleigh expansions

In the space $y \geq (s+\frac{1}{2}h)$, the fields are expressed as a superposition of the incident wave together with a series of all upward going plane waves. Similarly, in the region $y \leq -(s+\frac{1}{2}h')$ the fields are represented by an expansion of downward going waves, whilst in the domain $-(s-\frac{1}{2}h') \leq y \leq (s-\frac{1}{2}h)$, an expansion consisting of both upward and downward going waves is used. Thus, one writes

$$\left.\begin{aligned} E_z &= E^i(x,y) + \sum_{p=-\infty}^{\infty} C_p R_p(x,y) && \text{for } y \geq (s+\tfrac{1}{2}h) \\ &= \sum_{p=-\infty}^{\infty} [C_p^+ R_p(x,y) + C_p^- \hat{R}_p(x,y)] && \text{for } -(s-\tfrac{1}{2}h') \leq y \leq (s-\tfrac{1}{2}h) \\ &= \sum_{p=-\infty}^{\infty} \hat{C}_p \hat{R}_p(x,y) && \text{for } y \leq -(s+\tfrac{1}{2}h') \end{aligned}\right\} \quad (9.2.5)$$

9.2.4 Modal expansions

Within each groove of the structure, the field may be specified by waveguide modal expansions, each distinct mode of which must satisfy the appropriate boundary condition on the aperture walls. However, for each period of the entire structure there are $d_T/d(= N+1)$ grooves in the upper array and $d_T/d'(= N'+1)$ grooves in the lower array. Hence, a

separate modal expansion is needed for each of the (N+1) grooves in the upper array and each of the (N'+1) grooves in the lower array. (The prescription of the modes which is detailed in the remainder of this section has been taken from Wirgin [9.7].)

Thus one may specify the electric field in each groove of the upper array by :

$$E_z(x,y) = \sum_{m=1}^{\infty} \{c_m^{\ell} \sin[k_0\mu_m(y-s)] + d_m^{\ell} \cos[k_0\mu_m(y-s)]\} \, MEZU_m^{\ell}(x) \quad (9.2.6)$$

for $(s-\frac{1}{2}h) \leq y \leq (s+\frac{1}{2}h)$,

$\ell d \leq x \leq \ell d + c$

and $\ell \quad \varepsilon \quad [0,N]$.

Here, the mode is given by

$$MEZU_m^{\ell}(x) = \sqrt{\frac{\varepsilon_m}{c}} \sin[\frac{m\pi}{c}(x-\ell d)] \quad (9.2.7)$$

and

$$k_0\mu_m = \sqrt{k_0^2 - (\frac{m\pi}{c})^2} \, . \quad (9.2.8)$$

Similarly within the lower array

$$E_z(x,y) = \sum_{m=1}^{\infty} \{\hat{c}_m^{\ell} \sin[k_0\hat{\mu}_m(y+s)] + \hat{d}_m^{\ell} \cos[k_0\hat{\mu}_m(y+s)] \quad MEZL_m^{\ell}(x) \quad (9.2.9)$$

for $-(s+\frac{1}{2}h') \leq y \leq -(s-\frac{1}{2}h')$,

$\ell d' + \delta x \leq x \leq \ell d' + c' + \delta x$

and $\ell \quad \varepsilon \quad [0,N']$.

In this case the mode is given by

$$MEZL_m^{\ell}(x) = \sqrt{\frac{\epsilon_m}{c'}} \sin\left[\frac{m\pi}{c'}(x-\ell d-\delta x)\right] \tag{9.2.10}$$

and $$k_0\hat{\mu}_m = \sqrt{k_0^2 - \left(\frac{m\pi}{c'}\right)^2}\,. \tag{9.2.11}$$

9.2.5 Solution of the field equations

The initial specification of the problem involves 8 infinite sets of coefficients composed of 4 sets of modal coefficients and 4 sets of Rayleigh coefficients. Boundary conditions are then applied at the surfaces $y = s+\frac{1}{2}h$, $s-\frac{1}{2}h$, $-s+\frac{1}{2}h'$ and $-s-\frac{1}{2}h'$, and the resulting equations are solved using the method of moments.

9.2.5.1 *Electric field equations*

These are derived by applying the continuity of E across the various interfaces.

(1) At $y = s + \frac{1}{2}h$:

$$\left.\begin{aligned} &\exp[jk_0(\alpha_0 x-\beta_0(s+\tfrac{1}{2}h)] + \sum_p C_p \exp[jk_0(\alpha_p x + \beta_p(s+\tfrac{1}{2}h))] \\ &\quad = \sum_{m=1}^{\infty} \left[c_m^{\ell} \sin(k_0\mu_m \tfrac{h}{2}) + d_m^{\ell} \cos(k_0\mu_m \tfrac{h}{2})\right] MEZU_m^{\ell}(x) \\ &\qquad \text{if } x \in [\ell d, \ell d+c] \quad \text{for all} \quad \ell \in [0,N] \\ &\quad = 0 \quad \text{otherwise.} \end{aligned}\right\} \tag{9.2.12}$$

Now applying the method of moments, the above expression is multiplied throughout by $\bar{R}_p(x,0)$ and integrated over the period $x \in [0, d_T]$. From the orthogonality of the basis $\{R_p\}$ expressed by

$$\int_0^{d_T} R_p(x,0)\ \bar{R}_p(x,0)\ dx \;=\; d_T\ \delta_{pP}$$

it follows that

$$\exp[-jk_0\beta_0(s+\tfrac{h}{2})]\ \delta_{p0} + C_p\ \exp[jk_0\beta_p(s+\tfrac{h}{2})] = \sum_{\ell=0}^{N}\sum_{m=1}^{\infty}(c_m^{\ell *}+d_m^{\ell *})\ I_{mp}^{\ell} \tag{9.2.13}$$

where

$$\left.\begin{aligned} c_m^{\ell *} &= c_m^{\ell}\sin(k_0\mu_m\tfrac{h}{2}) \qquad d_m^{\ell *} = d_m^{\ell}\cos(k_0\mu_m\tfrac{h}{2}) \\ I_{mp}^{\ell} &= \frac{1}{d_T}\sqrt{\frac{\varepsilon_m}{c}}\int_{\ell d}^{\ell d+c}\{\exp(-jk_0\alpha_p x)\ \sin[\tfrac{m\pi}{c}(x-\ell d)]\}\ dx\ . \end{aligned}\right\} \tag{9.2.14}$$

(2) At $y = s - \frac{1}{2}h$:

In a similar fashion

$$C_p^{+}\exp[jk_0\beta_p(s-\tfrac{1}{2}h)] + C_p^{-}\exp[-jk_0\beta_p(s-\tfrac{1}{2}h)] = \sum_{\ell=0}^{N}\sum_{m=1}^{\infty}(-\,c_m^{\ell *}+d_m^{\ell *})\ I_{mp}^{\ell}\ . \tag{9.2.15}$$

(3) At $y = -(s - \frac{1}{2}h')$:

$$\left.\begin{aligned} &\sum_p \{C_p^{+}\exp[jk_0(\alpha_p x+\beta_p(-s+\tfrac{1}{2}h'))]+ C_p^{-}\exp[jk_0(\alpha_p x-\beta_p(-s+\tfrac{1}{2}h'))]\ \} \\ &\quad = \sum_{m=1}^{\infty}(\hat{c}_m^{\ell *} + \hat{d}_m^{\ell *})\ \mathrm{MEZL}_m^{\ell}(x) \end{aligned}\right\} \tag{9.2.16}$$

continued ..

if $x \in [\ell d' + \delta x,\ \ell d' + \delta x + c']$ for all $\ell \in [0,N']$ (9.2.16) *continued*

$= 0$ otherwise.

This reduces to the form

$$C_p^+ \exp[jk_0\beta_p(-s+\tfrac{1}{2}h')] + C_p^- \exp[-jk_0\beta_p(-s+\tfrac{1}{2}h')]$$

$$= \sum_{\ell=0}^{N'} \sum_{m=1}^{\infty} (\hat{c}_m^{\ell *} + \hat{d}_m^{\ell *})\ J_{mp}^{\ell} \qquad (9.2.17)$$

where $\hat{c}_m^{\ell *} = \hat{c}_m^{\ell} \sin(k_0\hat{\mu}_m \frac{h'}{2})$ $\qquad \hat{d}_m^{\ell *} = \hat{d}_m^{\ell} \cos(k_0\hat{\mu}_m \frac{h'}{2})$

and $$J_{mp}^{\ell} = \frac{1}{d_T}\sqrt{\frac{\varepsilon_m}{c'}} \int_{\ell d'+\delta x}^{\ell d'+\delta x+c'} \{\exp[-jk_0\alpha_p x] \sin[\frac{m\pi}{c'}(x-\ell d'-\delta x)]\} \qquad (9.2.18)$$

(4) At $y = -(s + \tfrac{1}{2}h')$:

$$\hat{C}_p \exp[-jk_0\beta_p(-s-\tfrac{1}{2}h')]$$

$$= \sum_{\ell=0}^{N'} \sum_{m=1}^{\infty} (-\hat{c}_m^{\ell *} + \hat{d}_m^{\ell *})\ J_{mp}^{\ell}. \qquad (9.2.19)$$

9.2.5.2 *Magnetic field equations*

These constraints may be derived by applying the continuity of $\partial E/\partial y$ across the free space-aperture interfaces.

(1) At $y = s + \frac{1}{2}h$, this is given by

$$- jk_0\beta_0 \exp[jk_0(\alpha_0 x - \beta_0(s+\tfrac{h}{2}))] + \sum_p jk_0\beta_p C_p \exp[jk_0(\alpha_p x + \beta_p(s+\tfrac{h}{2}))]$$

$$= \sum_{m=1}^{\infty} k_0\mu_m [c_m^{\ell *} \cot(k_0\mu_m \tfrac{h}{2}) - d_m^{\ell *} \tan(k_0\mu_m \tfrac{h}{2})] \, MEZU_m^{\ell}(x) \qquad (9.2.20)$$

where $x \in [\ell d, \ell d+c]$ for all $\ell \in [0,N]$.

Upon multiplying both sides of this equation by $MEZU_M^L$ and integrating over the aperture $x \in [Ld, Ld+c]$ the orthonormality of the set of modes ensures that

$$d_T \sum_p j\beta_p \{C_p \exp[jk_0\beta_p(s+\tfrac{h}{2})] - \exp[-jk_0\beta_0(s+\tfrac{h}{2})] \delta_{p0}\} \bar{I}_{Mp}^{L}$$

$$= \mu_M [c_M^{L*} \cot(k_0\mu_M \tfrac{h}{2}) - d_M^{L*} \tan(k_0\mu_M \tfrac{h}{2})] \qquad (9.2.21)$$

for all $M \in [1,\infty)$ and $L \in [0,N]$.

(2) At $y = s - \frac{1}{2}h$, one derives a similar expression of the form

$$d_T \sum_p j\beta_p \{C_p^{+} \exp[jk_0\beta_p(s-\tfrac{h}{2})] - C_p^{-} \exp[-jk_0\beta_p(s-\tfrac{h}{2})]\} \bar{I}_{Mp}^{L}$$

$$= \mu_M [c_M^{L*} \cot(k_0\mu_M \tfrac{h}{2}) + d_M^{L*} \tan(k_0\mu_M \tfrac{h}{2})] \qquad (9.2.22)$$

for all $M \in [1,\infty)$ and $L \in [0,N]$.

(3) At $y = -s + \frac{1}{2}h'$, the continuity of $\partial E/\partial y$ leads to

$$d_T \sum_p j\beta_p \{C_p^+ \exp[jk_0\beta_p(-s+\tfrac{1}{2}H)] - C_p^- \exp[-jk_0\beta_p(-s+\tfrac{1}{2}H)]\} \bar{J}_{Mp}^L$$

$$= \hat{\mu}_M [c_M^{L*} \cot(k_0\hat{\mu}_M \frac{h'}{2}) - \hat{d}_M^{L*} \tan(k_0\hat{\mu}_M \frac{h'}{2})] \qquad (9.2.23)$$

for all $M \in [1,\infty)$ and $L \in [0,N']$.

(4) Finally at $y = -s - \frac{1}{2}h'$, one finds that

$$- d_T \sum_p j\beta_p \hat{C}_p \exp[-jk_0\beta_p(-s-\frac{h'}{2})] \bar{J}_{M_p}^L$$

$$= \hat{\mu}_M [\hat{c}_M^{L*} \cot(k_0\hat{\mu}_M \frac{h'}{2}) + \hat{d}_M^{L*} \tan(k_0\hat{\mu}_M \frac{h'}{2})]. \qquad (9.2.24)$$

Following some simple but tedious manipulation of the four electric and four magnetic equations one may deduce the following four sets of simultaneous linear equations in the four sets of modal unknowns.

$$\sum_{\ell=0}^{N} \sum_{m=1}^{\infty} c_m^{\ell*} [d_T \sum_p j\beta_p I_{mp}^{\ell} \bar{I}_{Mp}^L - \mu_M \cot(k_0\mu_M \frac{h}{2}) \delta_{Mm}\delta_{L\ell}]$$

$$+ \sum_{\ell=0}^{N} \sum_{m=1}^{\infty} d_m^{\ell*} [d_T \sum_p j\beta_p I_{mp}^{\ell} \bar{I}_{Mp}^L + \mu_M \tan(k_0\mu_M \frac{h}{2}) \delta_{Mm}\delta_{L\ell}]$$

$$= 2j\beta_0 d_T \exp[-jk_0\beta_0(s + \frac{h}{2})] \bar{I}_{M0}^L \qquad (9.2.25)$$

where $M \in [1,\infty)$ and $L \in [0,N]$.

$$\sum_{\ell=0}^{N}\sum_{m=1}^{\infty} c_m^{\ell *}\Big[-\sum_p H_p\, I_{mp}^{\ell}\,\bar{I}_{Mp}^{L} - \mu_M \cot\Big(k_0\mu_M \frac{h}{2}\Big)\,\delta_{Mm}\delta_{L\ell}\Big]$$

$$+\sum_{\ell=0}^{N}\sum_{m=1}^{\infty} d_m^{\ell *}\Big[\sum_p H_p\, I_{mp}^{\ell}\,\bar{I}_{Mp}^{L} - \mu_M \tan\Big(k_0\mu_m \frac{h}{2}\Big)\,\delta_{Mm}\delta_{L\ell}\Big]$$

$$+\sum_{\ell=0}^{N'}\sum_{m=1}^{\infty} \hat{c}_m^{\ell *}\Big[-\sum_p G_p\, J_{mp}^{\ell}\,\bar{I}_{Mp}^{L}\Big]$$

$$+\sum_{\ell=0}^{N'}\sum_{m=1}^{\infty} \hat{d}_m^{\ell *}\Big[-\sum_p G_p\, J_{mp}^{\ell}\,\bar{I}_{Mp}^{L}\Big]$$

$$= 0 \qquad (9.2.26)$$

for all $M \in [1,\infty)$ and $L \in [0,N]$

{where $H_p = d_T\beta_p \cot(k_0\beta_p S)$

and $G_p = d_T\beta_p/\sin(k_0\beta_p S)$ }.

$$\sum_{\ell=0}^{N}\sum_{m=1}^{\infty} c_m^{\ell *}\Big[-\sum_p G_p\, I_{mp}^{\ell}\,\bar{J}_{Mp}^{L}\Big]$$

$$+\sum_{\ell=0}^{N}\sum_{m=1}^{\infty} d_m^{\ell *}\Big[\sum_p G_p\, I_{mp}^{\ell}\,\bar{J}_{Mp}^{L}\Big]$$

$$+\sum_{\ell=0}^{N'}\sum_{m=1}^{\infty} \hat{c}_m^{\ell *}\Big[-\sum_p H_p\, J_{mp}^{\ell}\,\bar{J}_{Mp}^{L} - \hat{\mu}_M \cot\Big(k_0\hat{\mu}_M \frac{h'}{2}\Big)\,\delta_{Mm}\delta_{L\ell}\Big]$$

$$+\sum_{\ell=0}^{N'}\sum_{m=1}^{\infty} \hat{d}_m^{\ell *}\Big[-\sum_p H_p\, J_{mp}^{\ell}\,\bar{J}_{Mp}^{L} + \hat{\mu}_M \tan\Big(k_0\hat{\mu}_M \frac{h'}{2}\Big)\,\delta_{Mm}\delta_{L\ell}\Big]$$

$$= 0 \qquad (9.2.27)$$

for all $M \in [1,\infty)$ and $L \in [0,N']$

and finally,

$$\sum_{\ell=0}^{N'} \sum_{m=1}^{\infty} \hat{c}_m^{\ell *} \left[d_T \sum_p j\beta_p J_{mp}^{\ell} \bar{J}_{Mp}^{L} - \hat{\mu}_M \cot\left(k_0 \hat{\mu}_M \frac{h'}{2}\right) \delta_{Mm} \delta_{L\ell} \right]$$

$$+ \sum_{\ell=0}^{N'} \sum_{m=1}^{\infty} \hat{d}_m^{\ell *} \left[- d_T \sum_p j\beta_p J_{mp}^{\ell} \bar{J}_{Mp}^{L} - \hat{\mu}_M \tan\left(k_0 \hat{\mu}_M \frac{h'}{2}\right) \delta_{Mm} \delta_{L\ell} \right]$$

$$= 0 \qquad (9.2.28)$$

for all $M \in [1,\infty)$ and $L \in [0,N']$.

It should be noted that this set of equations cannot be decoupled as was the case for the inductive grid formalism. This feature appears to be characteristic of all bi-grating formalisms. Following the numerical solution of a truncated set of the above equations, the plane wave diffracted field may be reconstructed using the electric field equations of section (9.2.5.1).

9.3 THE DOUBLE GRATING AS A FABRY-PEROT INTERFEROMETER

The design of any Fabry-Perot interferometer relies upon the use of a pair of highly reflecting elements substantially displaced from one another so as to give rise to a multiplicity of interference orders. In the visible, this is achieved with the aid of thin metallic films. However, in the far infrared and millimetric regions of the spectrum, the high losses which occur within the films necessitates the use of an alternative arrangement.

Such an arrangement involves the use of double gratings (for P polarized radiation) or more generally double grids as was originally

suggested by Casey and Lewis [9.2, 9.3]. Provided that the wavelength is substantially in excess of the period of the grating, these structures can provide a very high long wavelength reflectance without incurring the absorption losses which afflict the traditional design.

Figure (9.3.1a) shows a typical interference pattern for such an interferometer computed using the theory described in the previous section. In an attempt to understand the features of these curves, it was decided to adopt a "multiple scattering" treatment to the problem. In the case of a large separation between the two arrays, the only mechanism of "communication" between the pair of lamellar gratings lies in the zeroth orders, the evanescent fields playing no role whatsoever in the coupling of the two arrays.

Now if R_0 and T_0 respectively represent the zeroth order amplitude reflection and transmission coefficients of each of the singly periodic arrays (assumed to be identical), it follows that the overall transmitted amplitude of the double array is given by the geometric series

$$\tau = T_0^2 \exp(i\delta) \left[1 + R_0^2 \exp(2i\delta) + R_0^4 \exp(4i\delta) + \ldots\right]$$

$$= \frac{T_0^2 \exp(i\delta)}{1 - R_0^2 \exp(2i\delta)} \qquad (9.3.1)$$

where

$$\delta = k_0\beta_0(S + h) \; . \qquad (9.3.2)$$

Thus, the total transmitted efficiency is given by

$$|\tau|^2 = [1 + F \sin^2 \xi]^{-1} \tag{9.3.3}$$

where $\xi = \delta + \psi_R$

with $\psi_R = \arg(R_0)$.

Here, the term F is given by

$$F = \frac{4|R_0|^2}{(1 - |R_0|^2)^2} \tag{9.3.4}$$

- a quantity often referred to as the finesse of the interferometer. From equation (9.3.3) it follows that the maxima of the interference pattern are given by

$$\theta_{max} = \arccos \left[\frac{\lambda}{2\pi(S + h)} (\ell\pi - \psi_R)\right] \tag{9.3.5}$$

(where ℓ refers to the order of interference.)

Clearly, it is the presence of the phase term ψ_R, introduced by each grating in the double structure, that displaces the peaks from their geometric limit. This was characterized with the aid of our computer program for a single lamellar transmission grating. The solid curve in figure (9.3.1b) shows the reconstruction of the P polarization transmittance using the output of the lamellar grating program and expression (9.3.3). (On that same graph, shown by the dashed curve, is the transmittance of the single lamellar array.) From an inspection of figures (9.3.1a) and (9.3.1b) it may be readily seen that the agreement between the rigorous theory and the reconstruction is excellent.

Explanation of figures (9.3.1) to (9.3.3)

These figures show the variation of the transmittance of the double grating as a function of the angle of incidence for three different geometries. In each case the interferometer, composed of two identical arrays, is operated in radiation having a normalized wavelength of $\lambda/d = 2.0$. Spectra for the geometries specified by the parameters:

(1) $c/d = 0.8$, $h/d = 0.0$, $\delta x/d = 0.0$, $S/d = 4.0$,

(2) $c/d = 0.5$, $h/d = 0.0$, $\delta x/d = 0.0$, $S/d = 4.0$

and

(3) $c/d = 0.8$, $h/d = 0.5$, $\delta x/d = 0.0$, $S/d = 4.0$

are supplied. On each of these figures are presented two graphs, identified by either an "a" or a "b". Those graphs referenced by an "a" represent calculations made using the rigorous double grating theory of the previous section. On each of these graphs are two curves applying to:

(i) P polarization performance - *denoted by the broken curve*

and

(ii) S polarization performance - *denoted by the solid curve.*

Those graphs marked by a "b" refer to the reconstruction performed using a single lamellar transmission grating theory and the Fabry-Perot interferometer expression (equation 9.3.3-4). Two curves showing:

(i) the transmittance of the single lamellar array - *denoted by the broken curve*

and

(ii) the reconstruction of the double grating performance - *denoted by the solid curve*

are displayed.

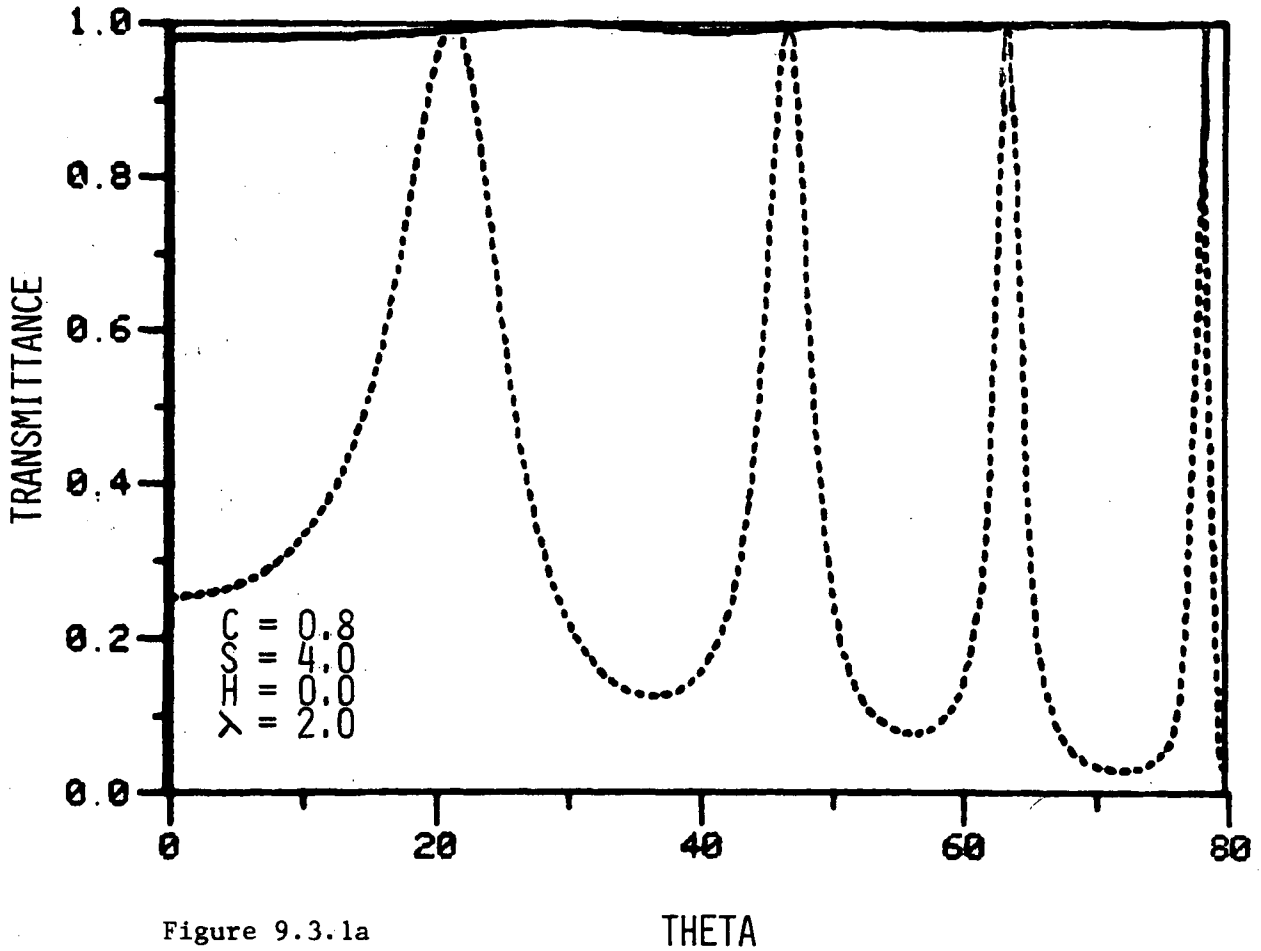


Figure 9.3.1a

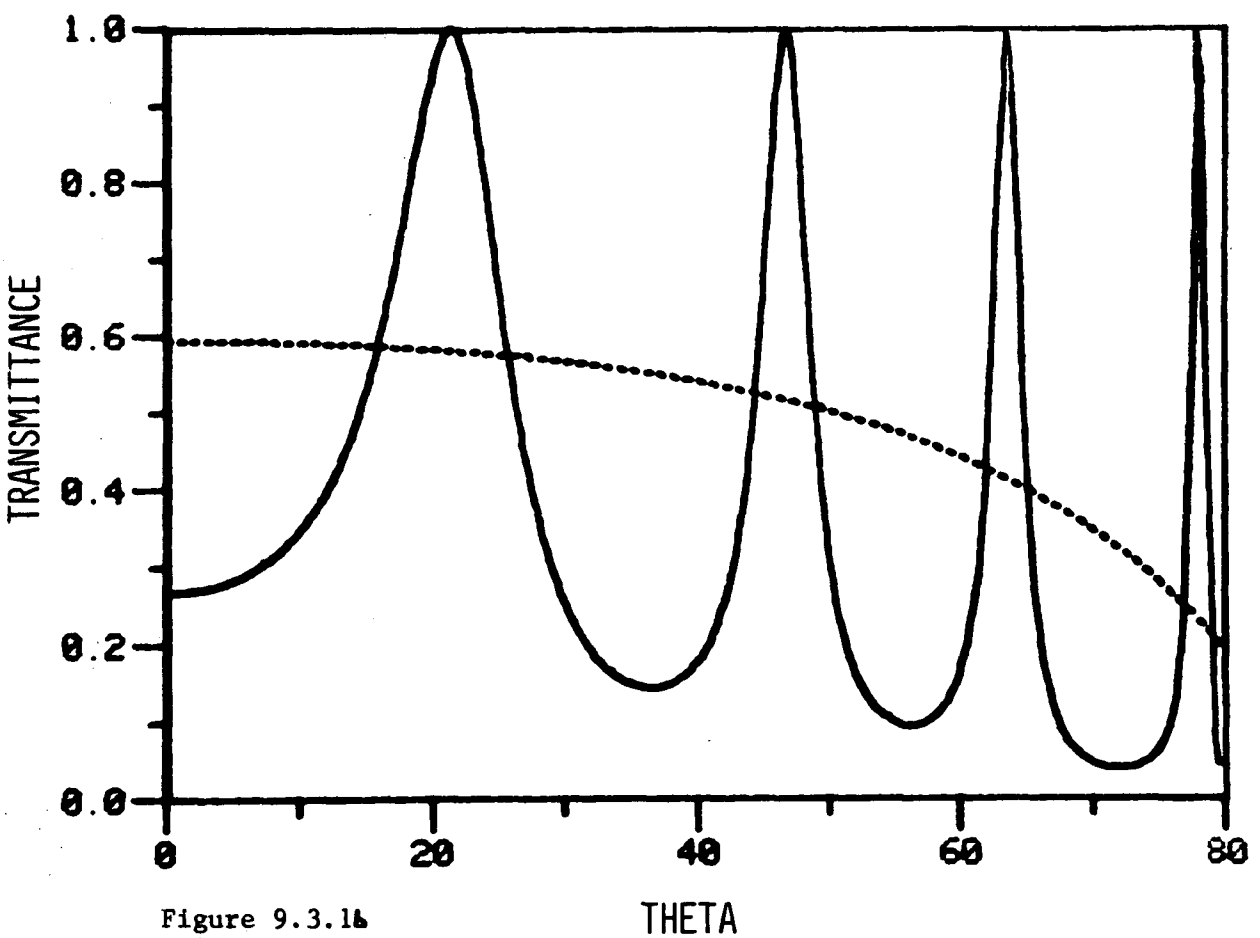


Figure 9.3.1b

Another feature of the P polarization curve shown in both graphs is the enhancement in the resolution of the instrument at high angles of incidence. This can readily be interpreted in terms of the increased reflectance of the single array at such angles, resulting in a greater finesse, which manifests itself as a diminished resonance half-width.

It is also interesting to note that the angular positions of the maxima tend towards the geometrical predictions as the angle of incidence increases. This was readily understood with the aid of a monomodal treatment originally discussed by McPhedran and Maystre [9.8]. For very long wavelengths, the spatial variation across the wavefront is minimal and so it seems reasonable to approximate the waveguide expansion by its first term. This is born out by the numerical results which show that only the fundamental mode carries a significant fraction of the incident flux. For very oblique angles of incidence, this approximation may not be entirely justified, but since the spatial variation is dependent on both the wavelength and the angle of incidence, it is difficult to judge the "cut-off" with any great accuracy. Nevertheless, for the case of infinitesimally thin grooves, it was possible to gain some useful understanding of the problem. Using such an analysis for the case of P polarization, it followed that

$$R_0 \simeq -1 + \frac{\chi_0 |K_0|^2}{\chi_0 |K_0|^2 + \sum_{n\neq 0} \chi_n |K_n|^2} \qquad (9.3.6)$$

where $$K_n = \int_0^c \exp(-jk_0\alpha_n x) \sin(\frac{\pi x}{c})\, dx \quad . \qquad (9.3.7)$$

The quantity χ_n is given by

$$\chi_n = k_0 \cos \theta_n \tag{9.3.8}$$

which means that χ_0 tends to zero by increasing the values of both λ and θ. Thus, R_0 approaches the limit - 1 and hence

$$\arg(R_0) \rightarrow \pi. \tag{9.3.9}$$

Hence, the positions of the interference maxima in P polarized radiation are given by

$$\theta_{max} \simeq \arccos \left[\frac{\ell\lambda}{2(S+h)}\right] \tag{9.3.10}$$

(for any arbitrary integer ℓ), which is the geometrical limit.

In the case of S polarized radiation, it may be shown that

$$R_0 \simeq \frac{\sum_{n\neq 0} \frac{\chi_0}{\chi_n} |L_n|^2}{|L_0|^2 + \sum_{n\neq 0} \frac{\chi_0}{\chi_n} |L_n|^2} \tag{9.3.11}$$

where $L_n = \int_0^c \exp -i\alpha_n x \, dx$. (9.3.12)

(Expression (9.3.11) was derived by applying the monomodal approximation, this time involving only the spatially invariant mode, m = 0.) Thus it follows that $|R_0|$ tends to zero but $\arg(R_0)$ tends to $-\frac{\pi}{2}$ thereby explaining the location of the resonance maxima for S polarization. In this case

$$\theta_{max} \simeq \arccos\left[\frac{(\ell + \tfrac{1}{2})\lambda}{2(S + h)}\right], \qquad (9.3.13)$$

the positions of which correspond to the P polarization minima. Furthermore, since $|R_0|$ tends to zero, the finesse of the instrument diminishes resulting in less distinct interference behaviour. For the orthogonal (P) polarization, exactly the opposite applies, with the minima increasing in strength, consequent upon the increasing reflectance of each of the lamellar gratings.

The previous discussion has centred on a grating having a particularly wide aperture (c/d = 0.8) and infinitesimally thin grooves. Such an arrangement resulted in somewhat smeared resonances, particularly in the case of S polarization. However, it was felt that the finesse of the instrument could be markedly enhanced by either reducing the aperture width or increasing the groove depth. This was indeed found to be the case as evidenced by figures (9.3.2) and (9.3.3) respectively showing transmission spectra for double gratings composed of identical arrays characterized by

(i) $c/d = 0.5$ $\qquad$ $h/d = 0.0$

and

(ii) $c/d = 0.8$ $\qquad$ $h/d = 0.5$.

Note that it is the parameter c which governs the angular position of the maxima. This can be explained by returning to the expression for R_0 (appropriate to P polarization):

$$R_0 \simeq -1 + \frac{\chi_0 |K_0|^2}{\chi_0 |K_0|^2 + \sum_{n \neq 0} \chi_n |K_n|^2} \; .$$

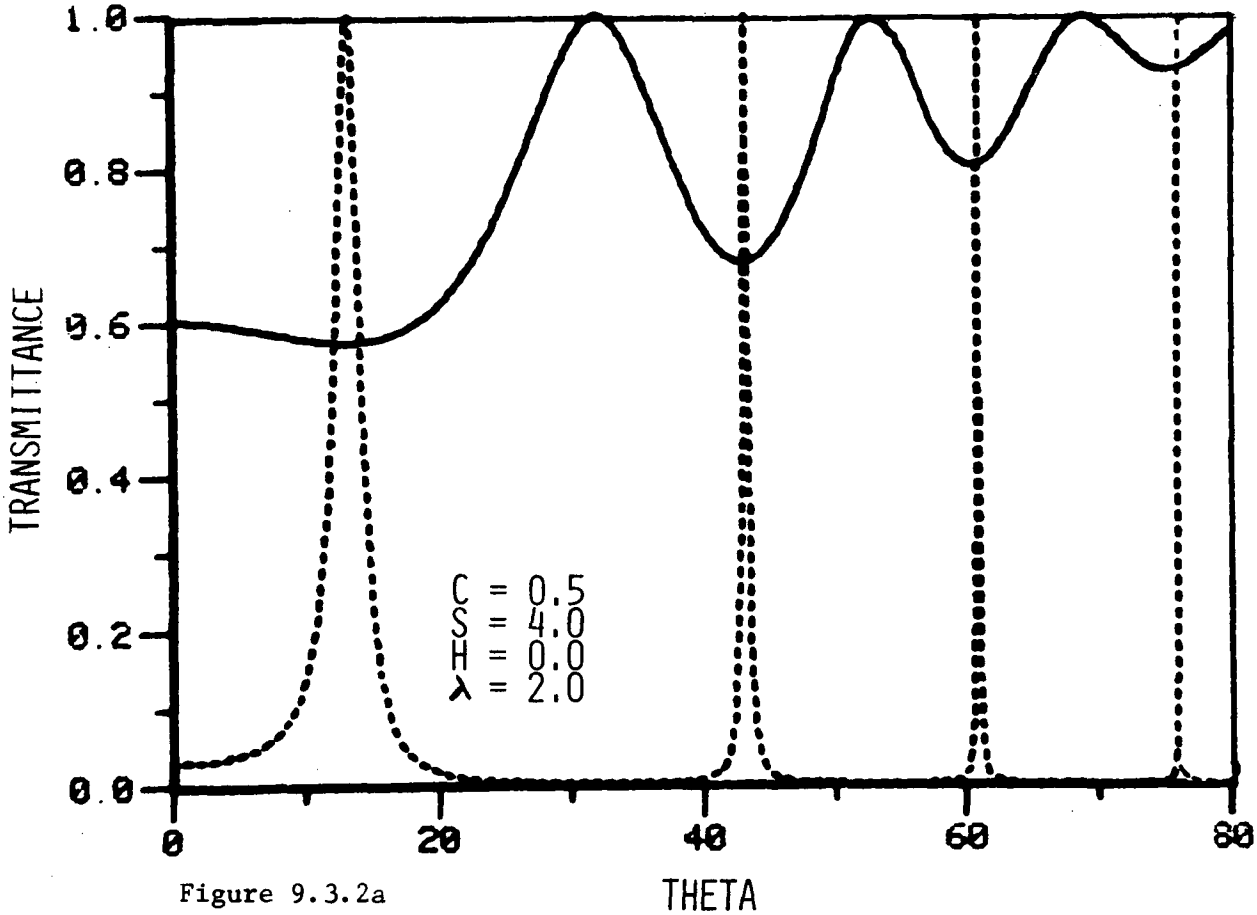


Figure 9.3.2a

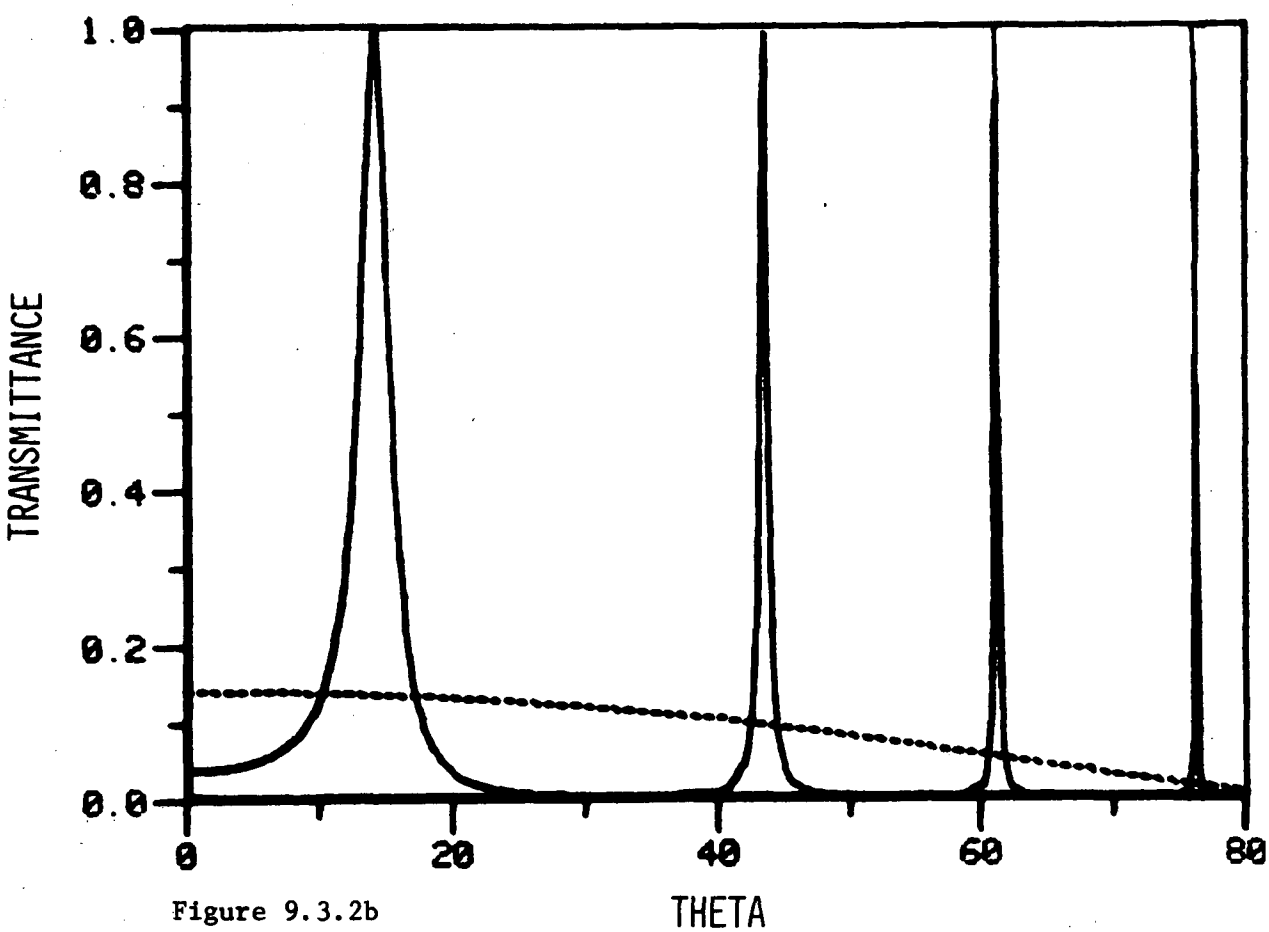


Figure 9.3.2b

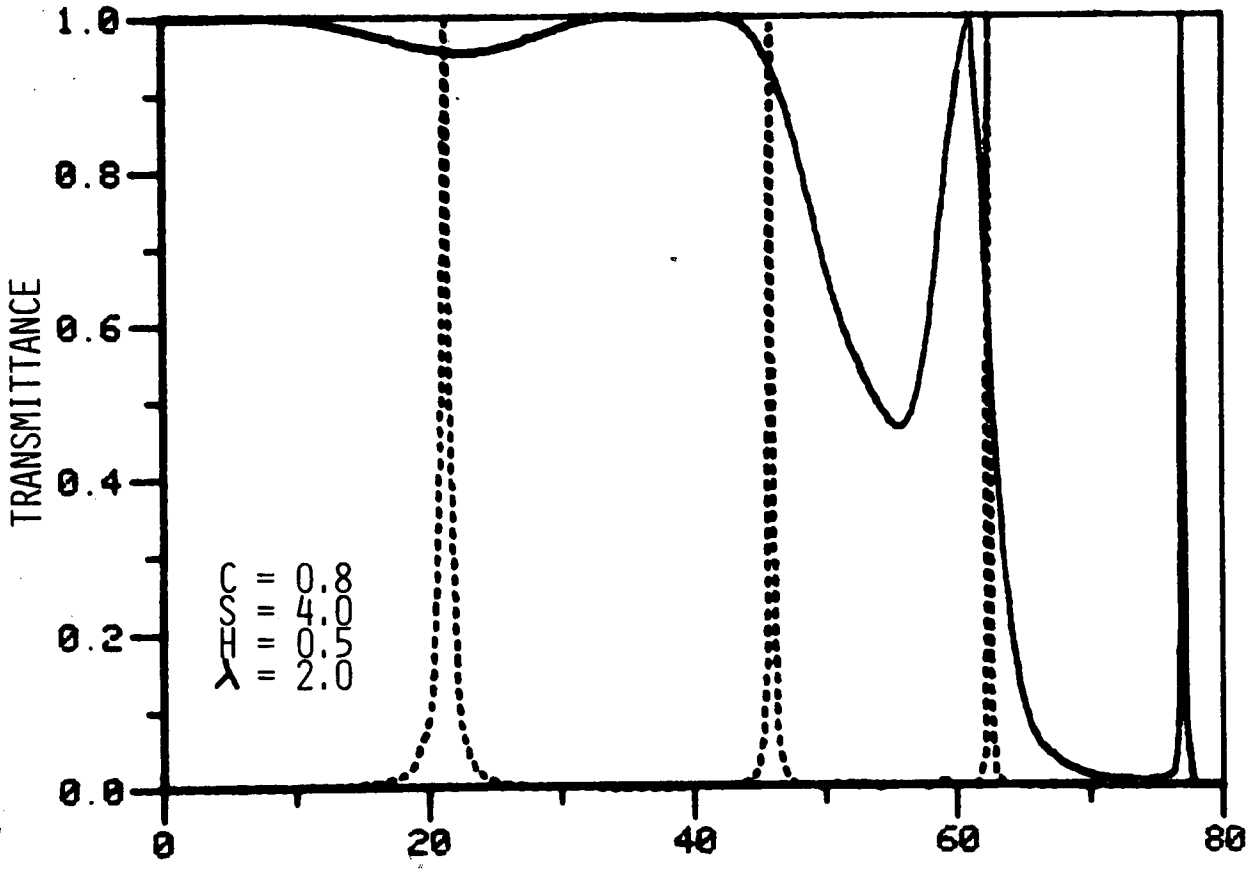


Figure 9.3.3a

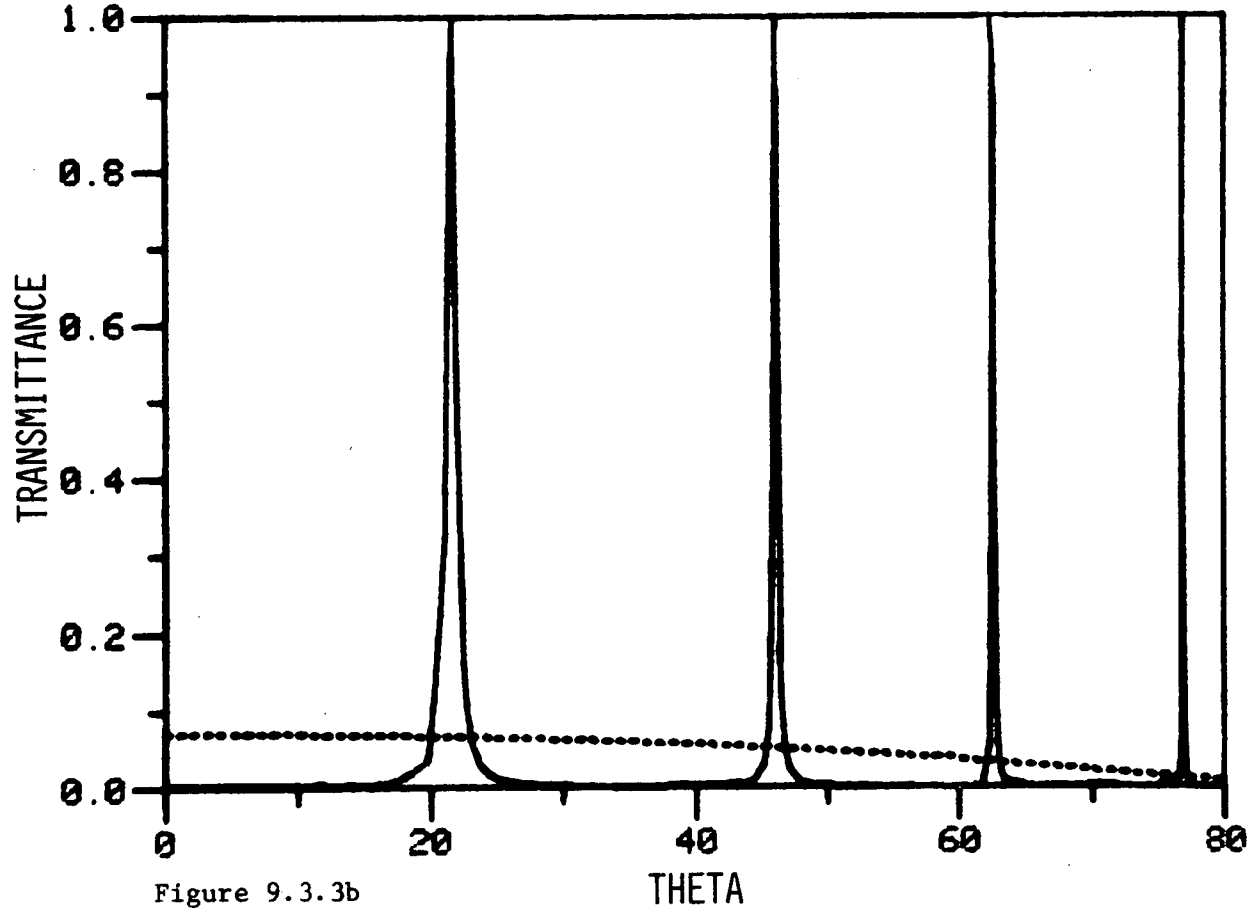


Figure 9.3.3b

Upon evaluation of K_n, it is found that:

$$K_n = -\frac{1}{d}\frac{\pi/c}{k_0^2\alpha_n^2 - \pi^2/c^2}\left[1 + \exp(-jk_0\alpha_n c)\right]$$

which tends to zero for small values of c. Naturally the same remarks apply to S polarization. In the case of finitely thick arrays, the above expression for R_0 is modified only by the presence of additional term in the denominator and so all previous comments are still valid.

The multiple scattering treatment which was adopted has been thoroughly tested and it has been found that only in the case of $S/d \lesssim 0.5$ do any departures from the predictions of the rigorous theory become noticeable. This occurs because of the additional array "coupling" mechanism (in the case of small separations) supplied by the evanescent fields. This is further supported with evidence gained from numerical experiments concerning the effects of the array phasing parameter δx. In the case of large values of S, the multiple scattering treatment predicts no dependence upon the parameter δx, a result which is in agreement with the predictions of the rigorous theory. However, for small array separations ($S/d \lesssim 0.5$) the evanescent coupling becomes quite noticeable. In figures (9.3.4) are shown the effects of the displacement parameter δx on the transmission properties of a double grating composed of two identical arrays arranged such that $S/d = 0.02$. Note that for $\delta x/d = 0.5$, total blocking of P polarized radiation is exhibited but that a noticeable leakage of energy occurs for the orthogonal polarization. It is possible that this may be understood in terms of the directions of current flow appropriate to the two polarizations. However, at this stage, with the

Figure 9.3.4

The following set of graphs demonstrate the effects of the array phasing parameter $\delta x/d$ on the transmittance of the double grating. As mentioned in the text, the transmittance is independent of δx provided that the zeroth order is the only energy carrying mechanism and also that the arrays are substantially displaced from one another. In the following graphs, it was chosen to operate the structure (composed of a pair of identical lamellar arrays) at a normalized wavelength of 0.6. Each of the arrays, specified by $c/d = 0.5$ and $h/d = 0.4$, were separated from one another by an amount $S/d = 0.02$ in order to illustrate the total blocking of P polarized radiation and the substantial leakage of radiation of the orthogonal polarization. Graphs for the following values of $\delta x/d$ are supplied.

(a)	0.00	(e)	0.40	(i)	0.60
(b)	0.10	(f)	0.45	(j)	0.70
(c)	0.20	(g)	0.50	(k)	0.80
(d)	0.30	(h)	0.55	(1)	0.90

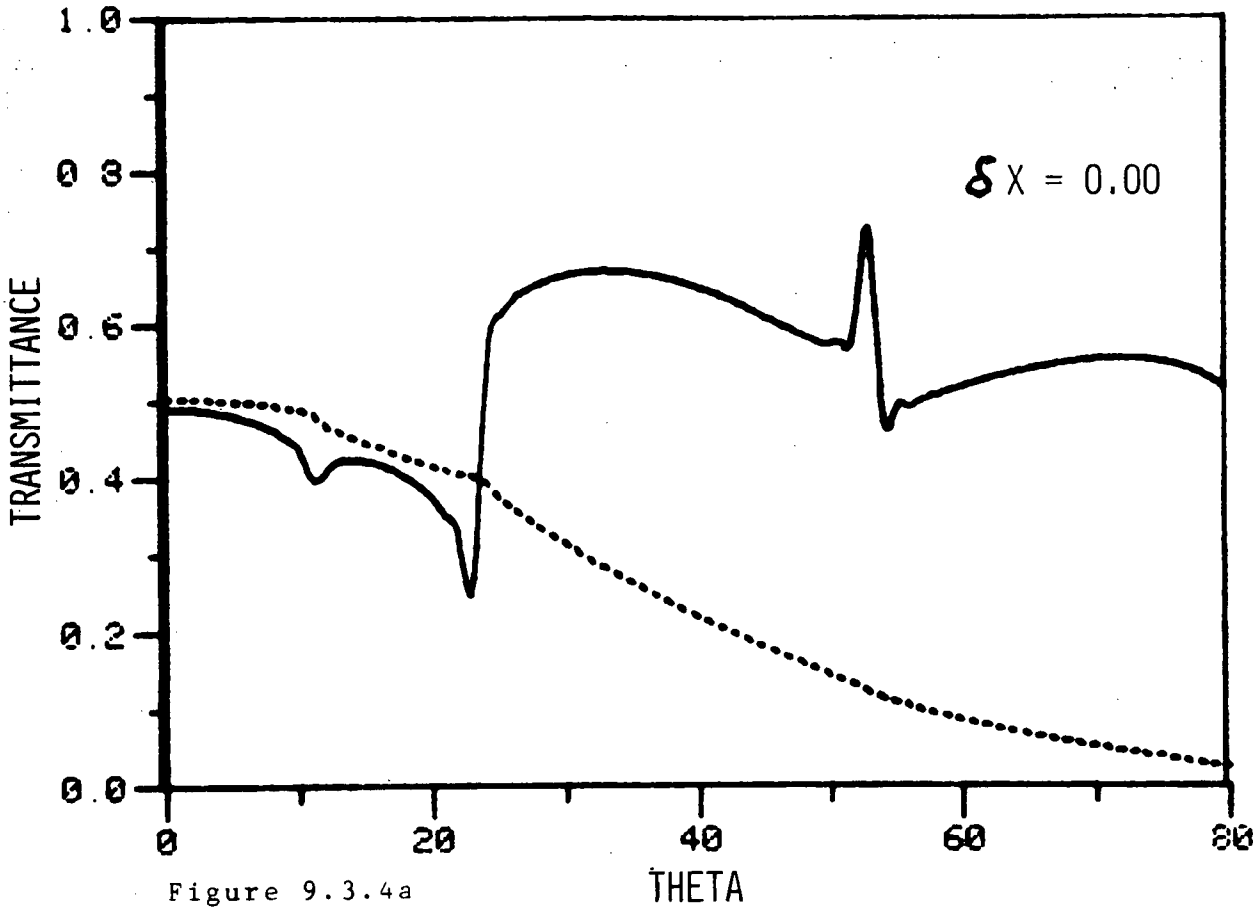


Figure 9.3.4a

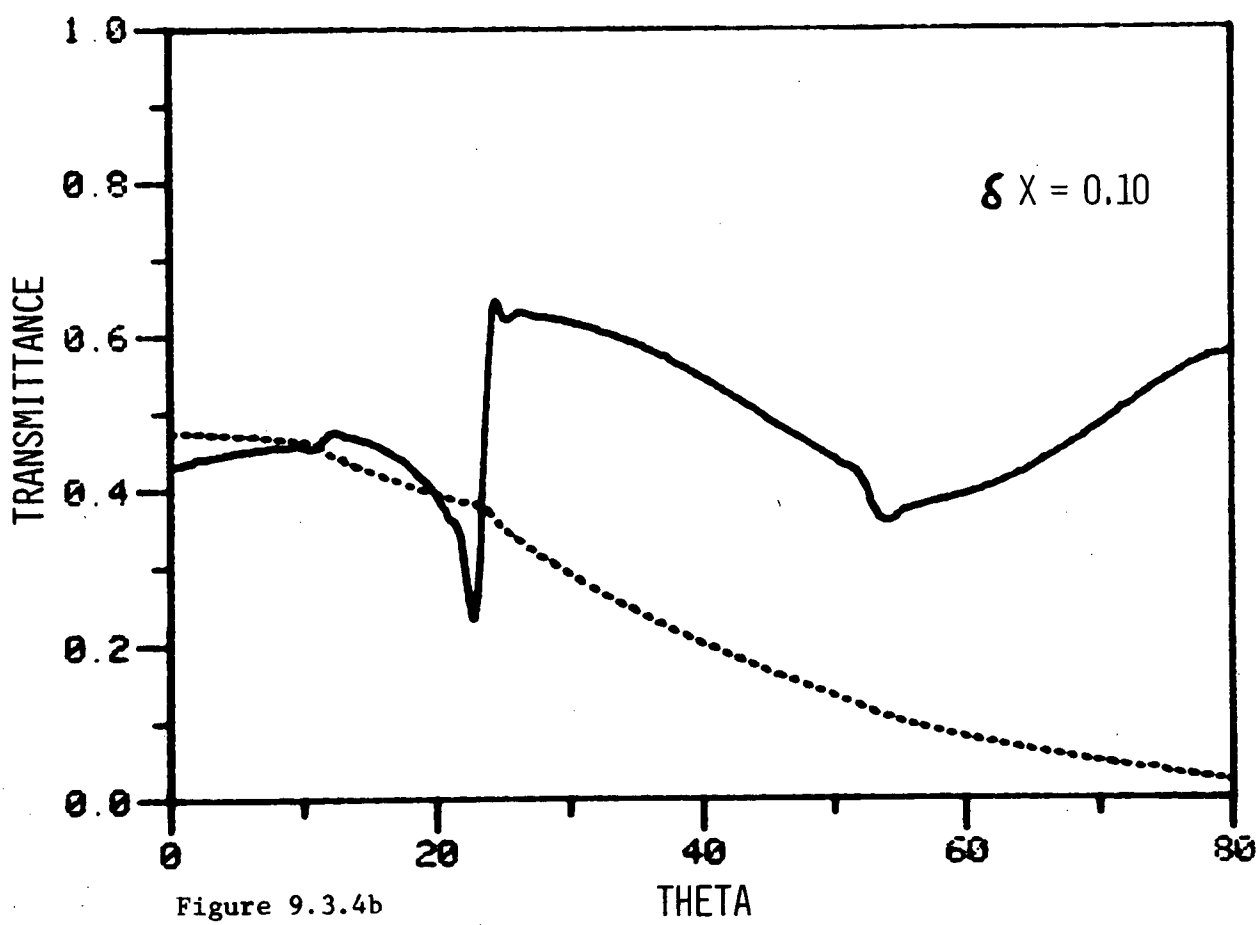


Figure 9.3.4b

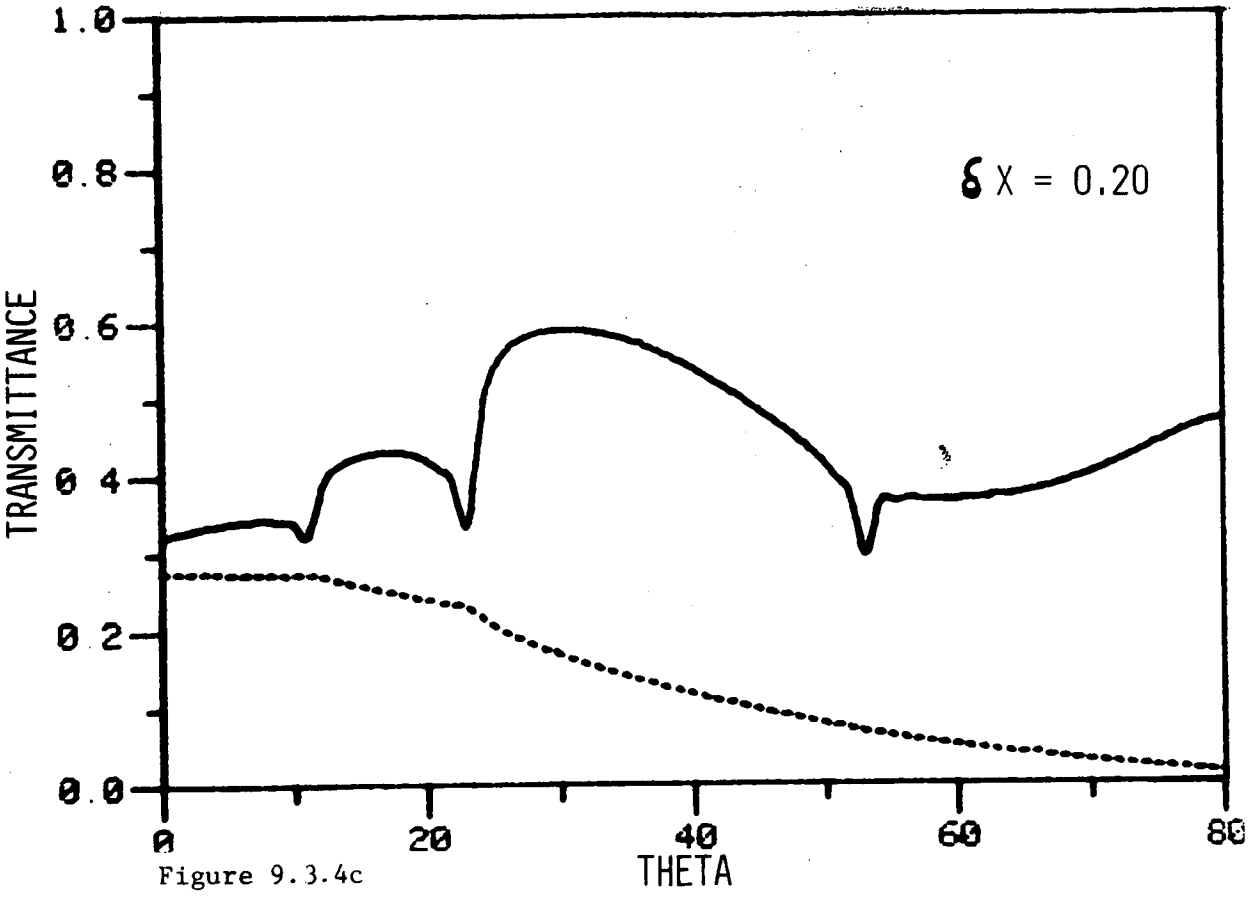


Figure 9.3.4c

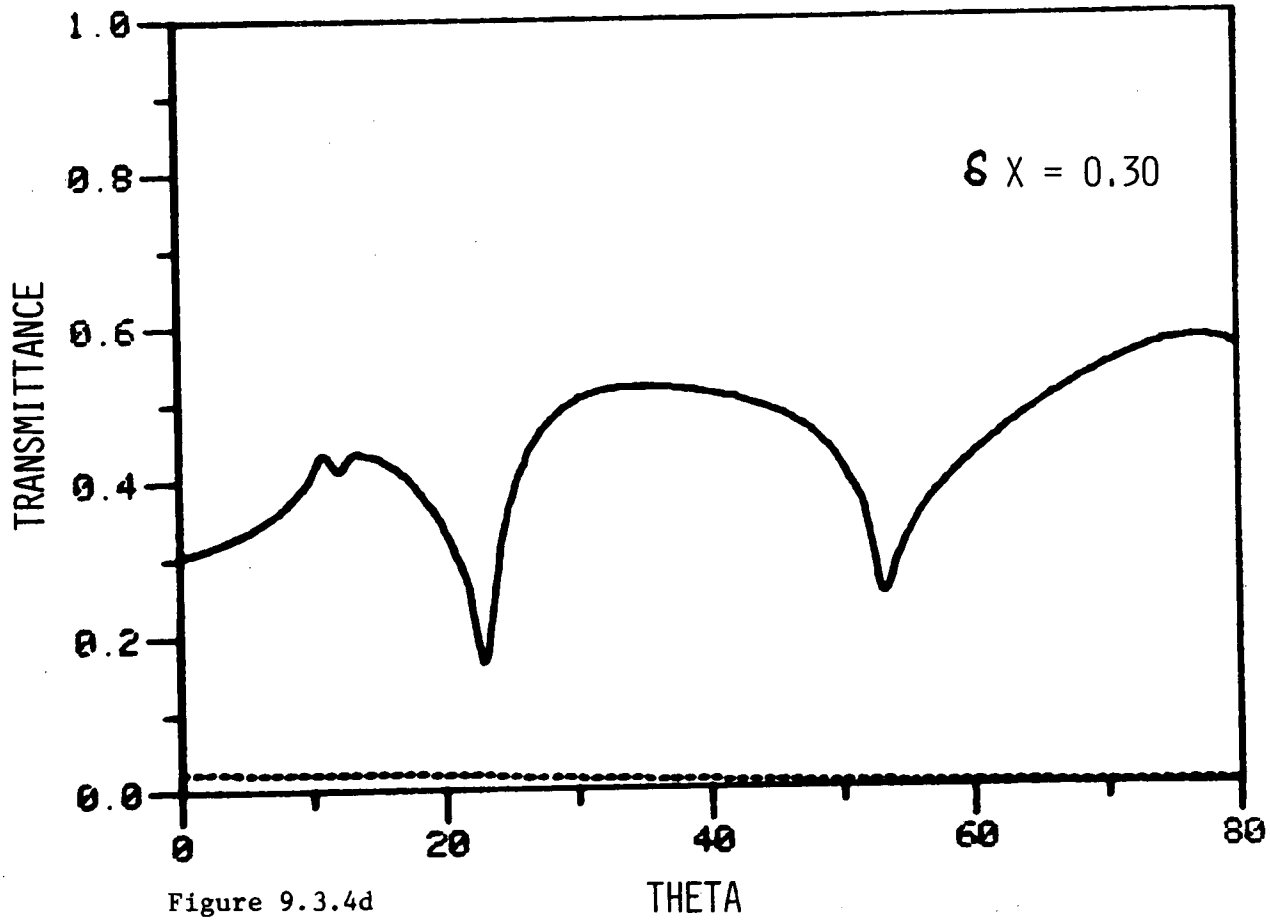


Figure 9.3.4d

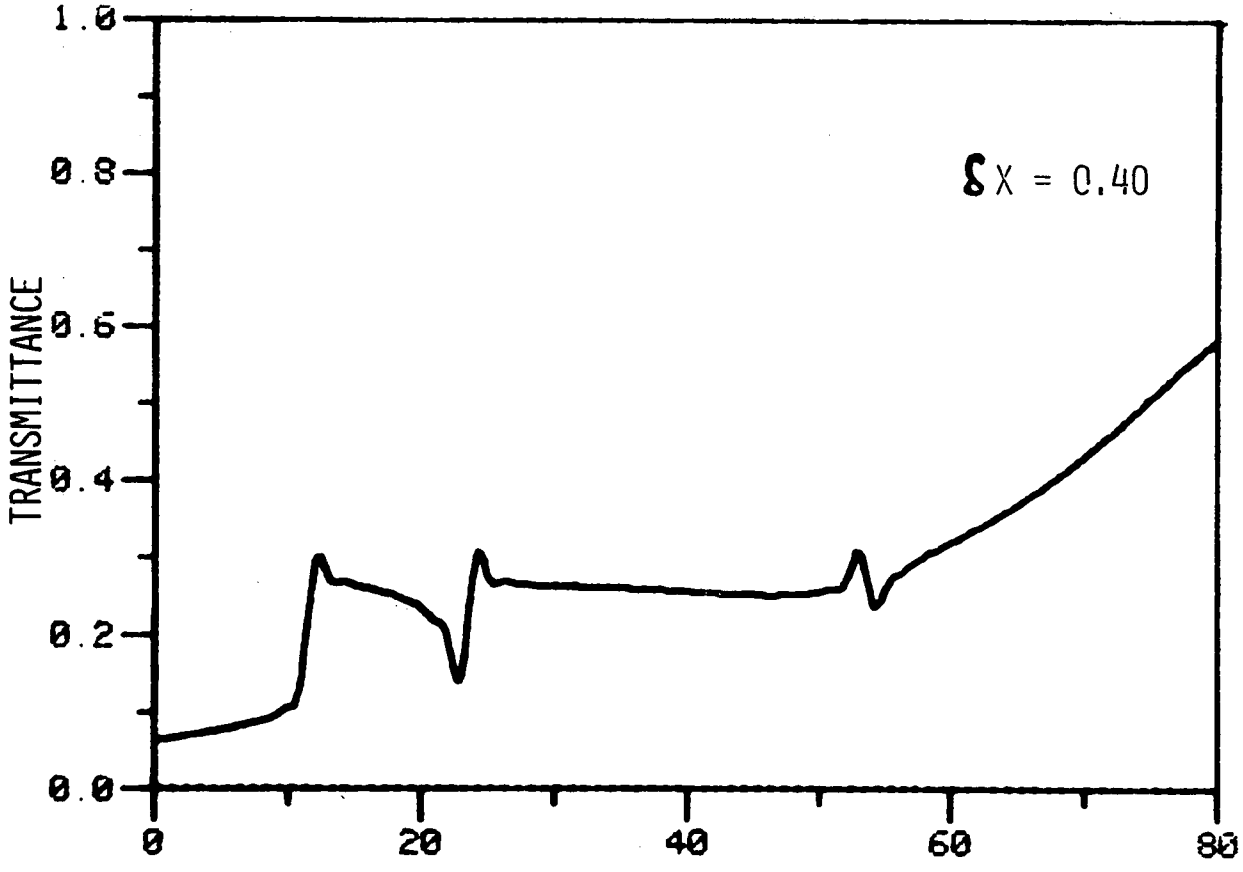


Figure 9.3.4e

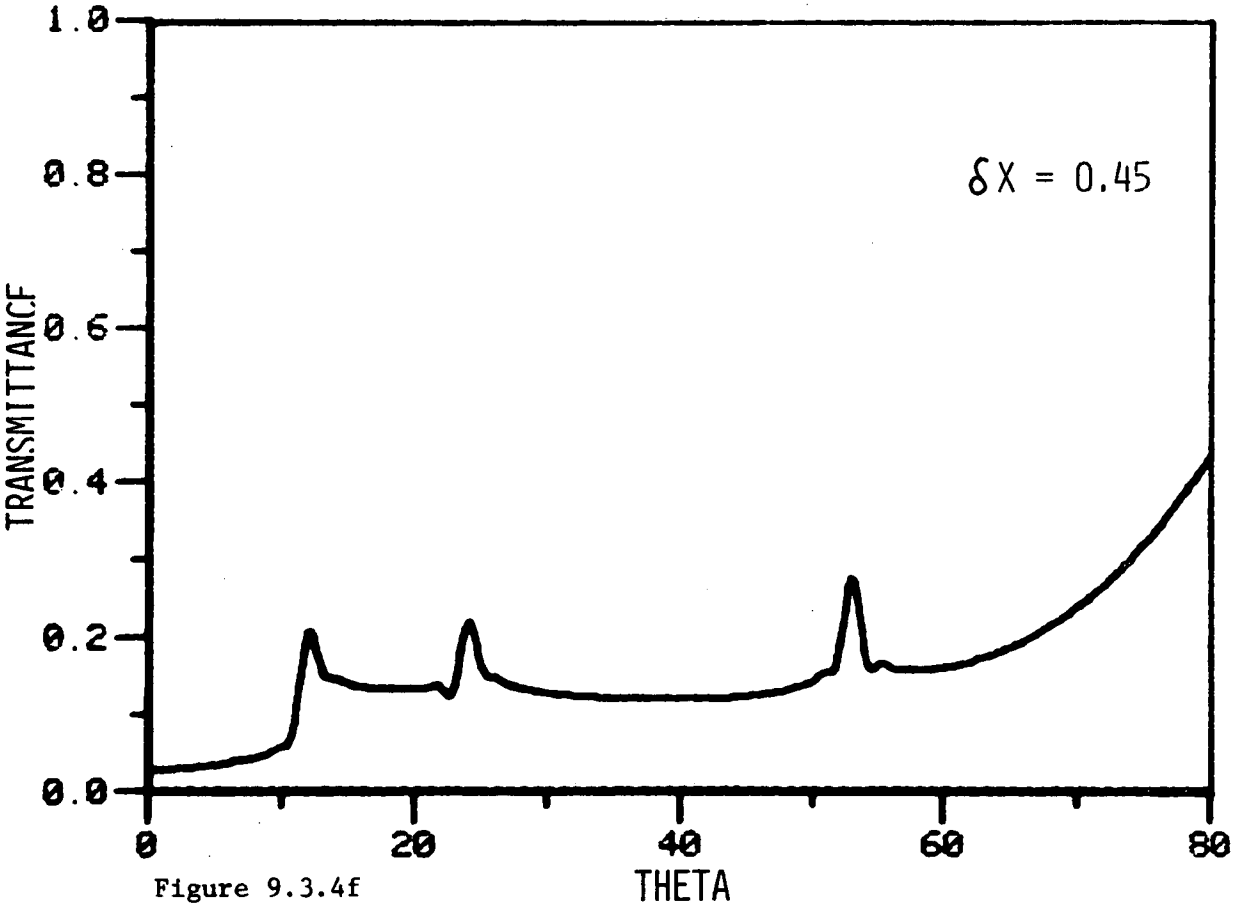


Figure 9.3.4f

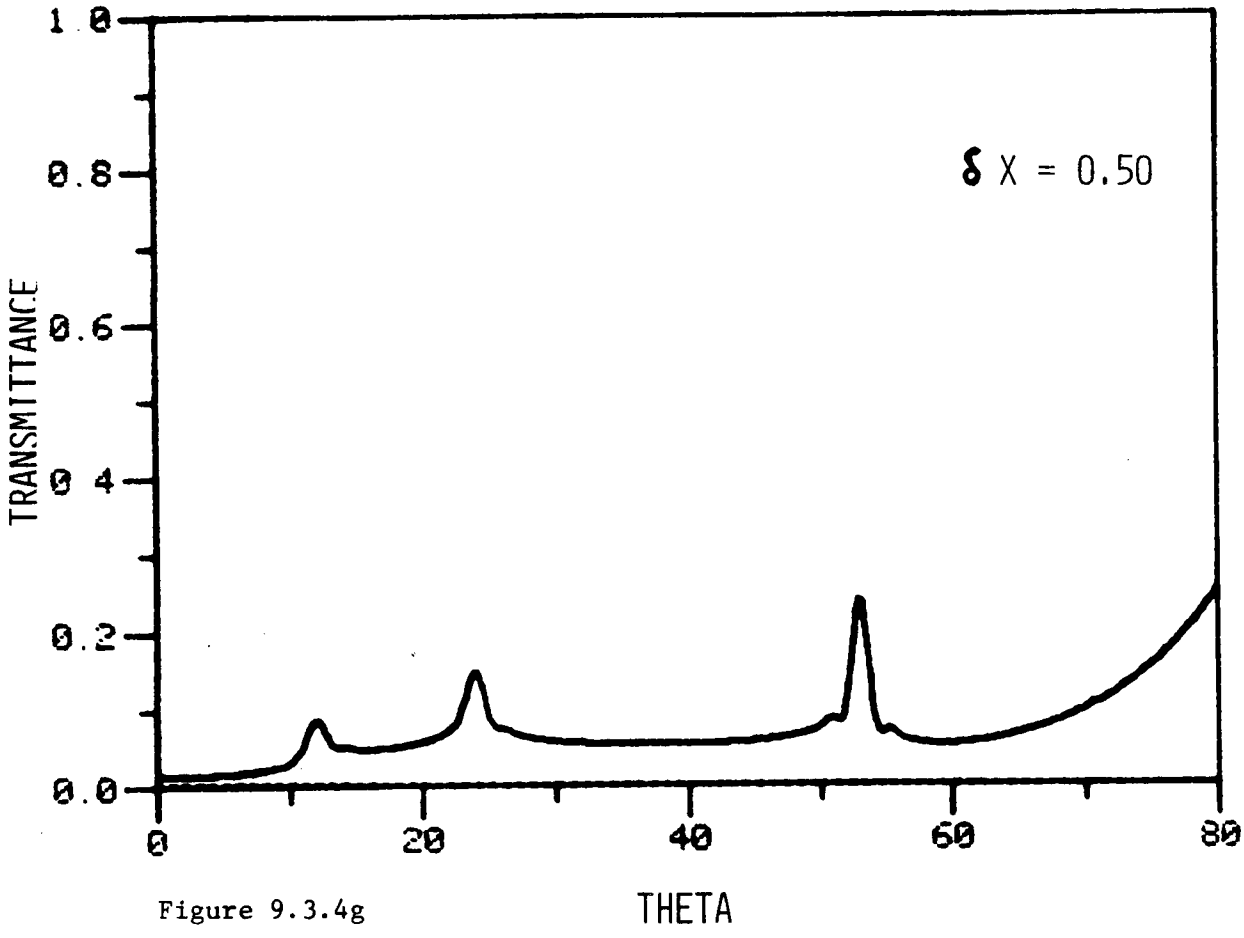


Figure 9.3.4g

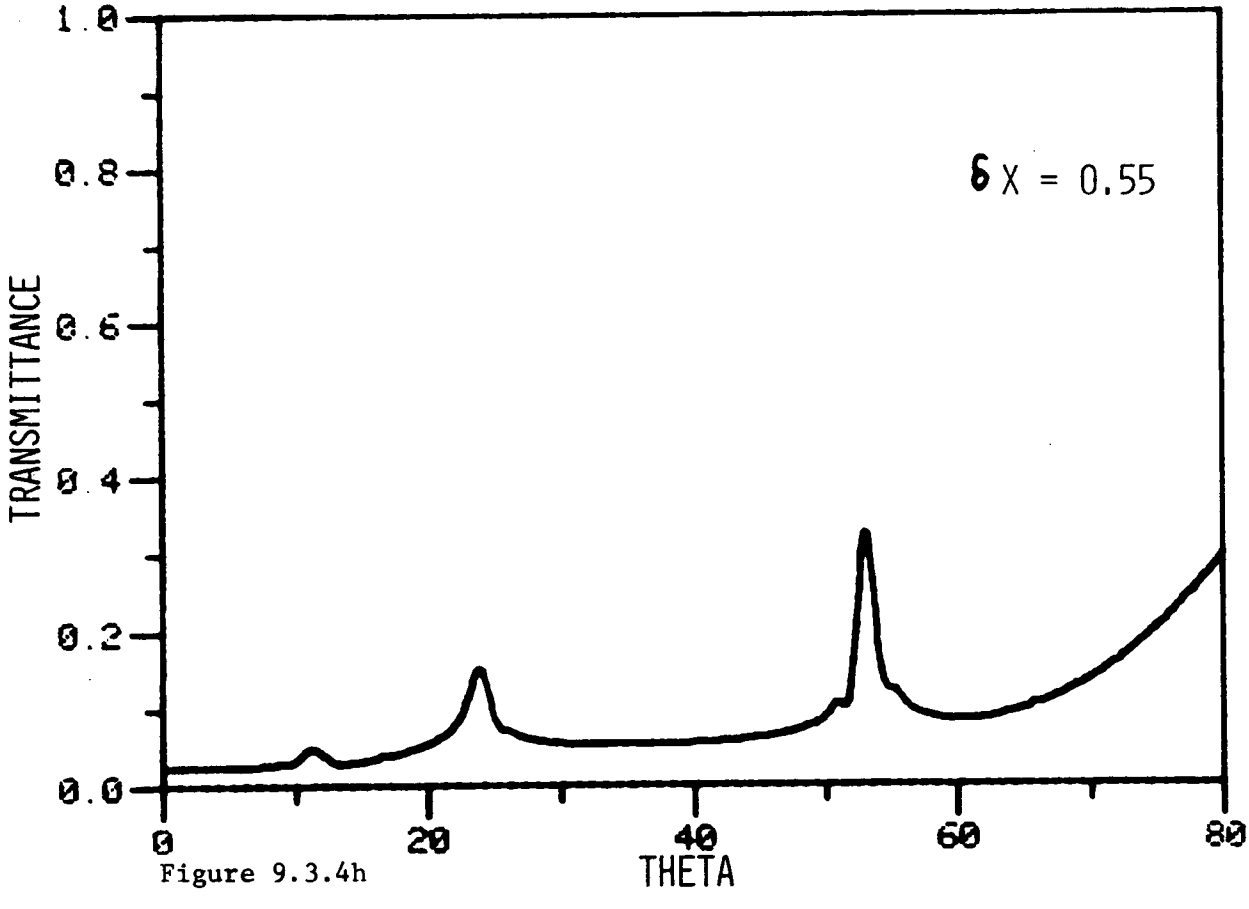


Figure 9.3.4h

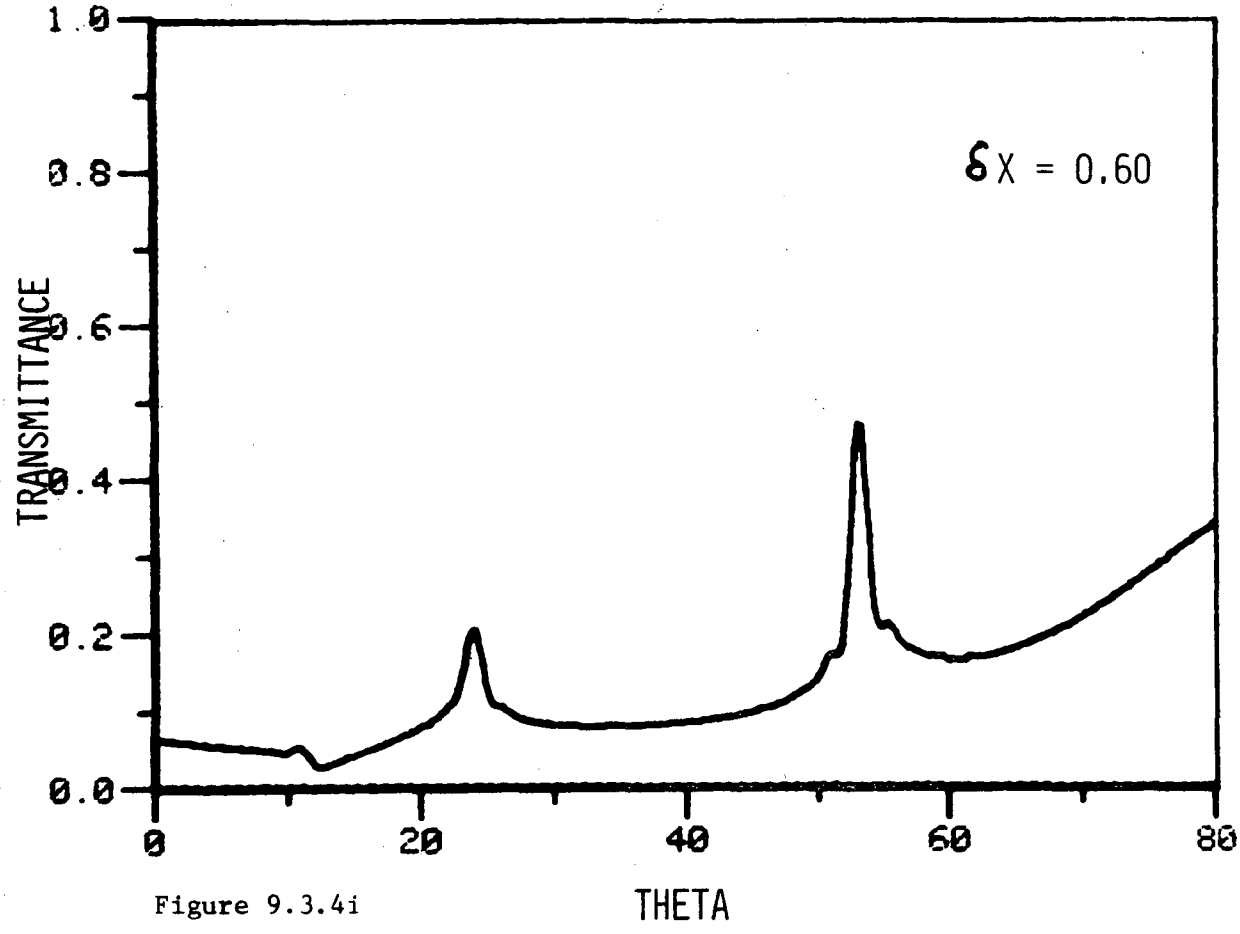


Figure 9.3.4i

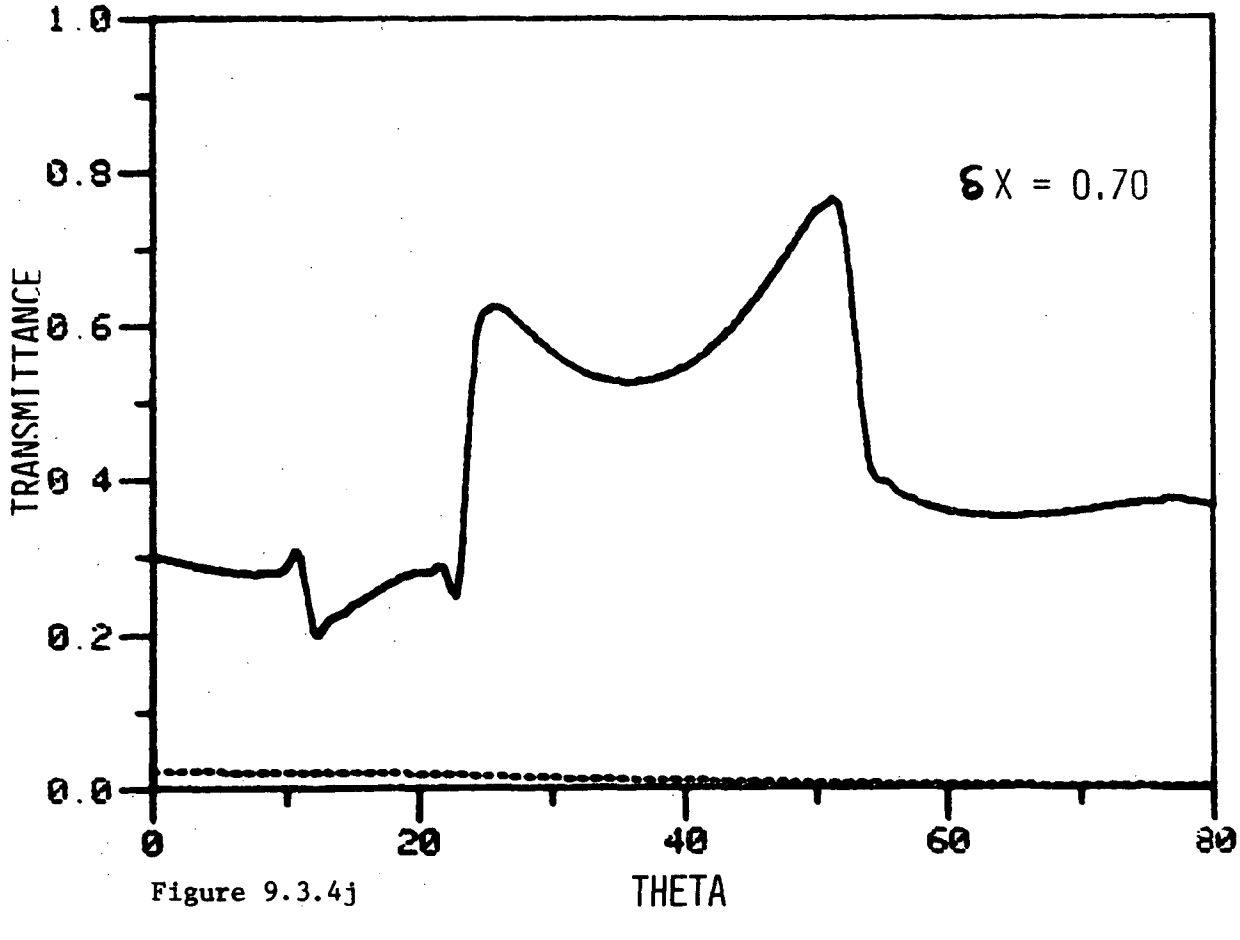


Figure 9.3.4j

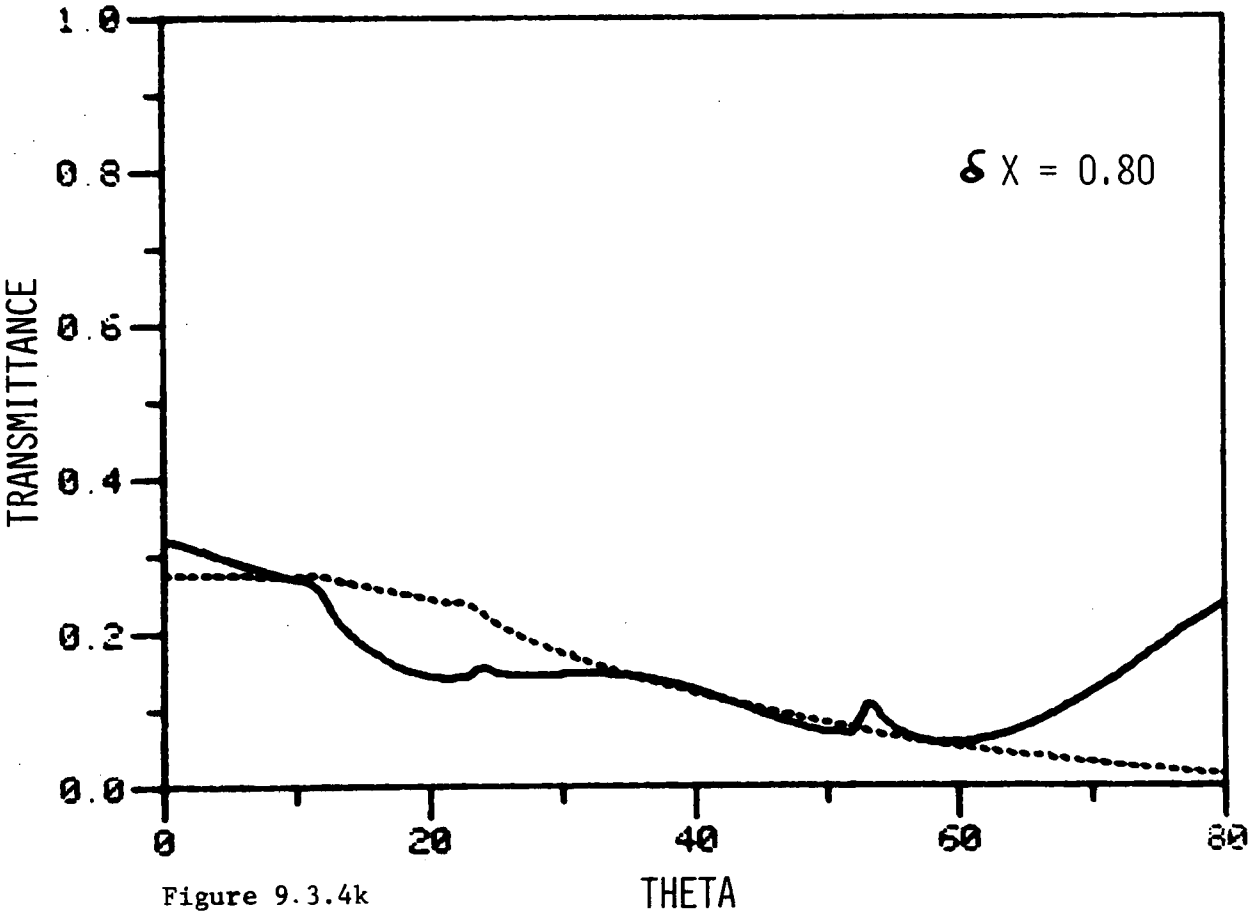


Figure 9.3.4k

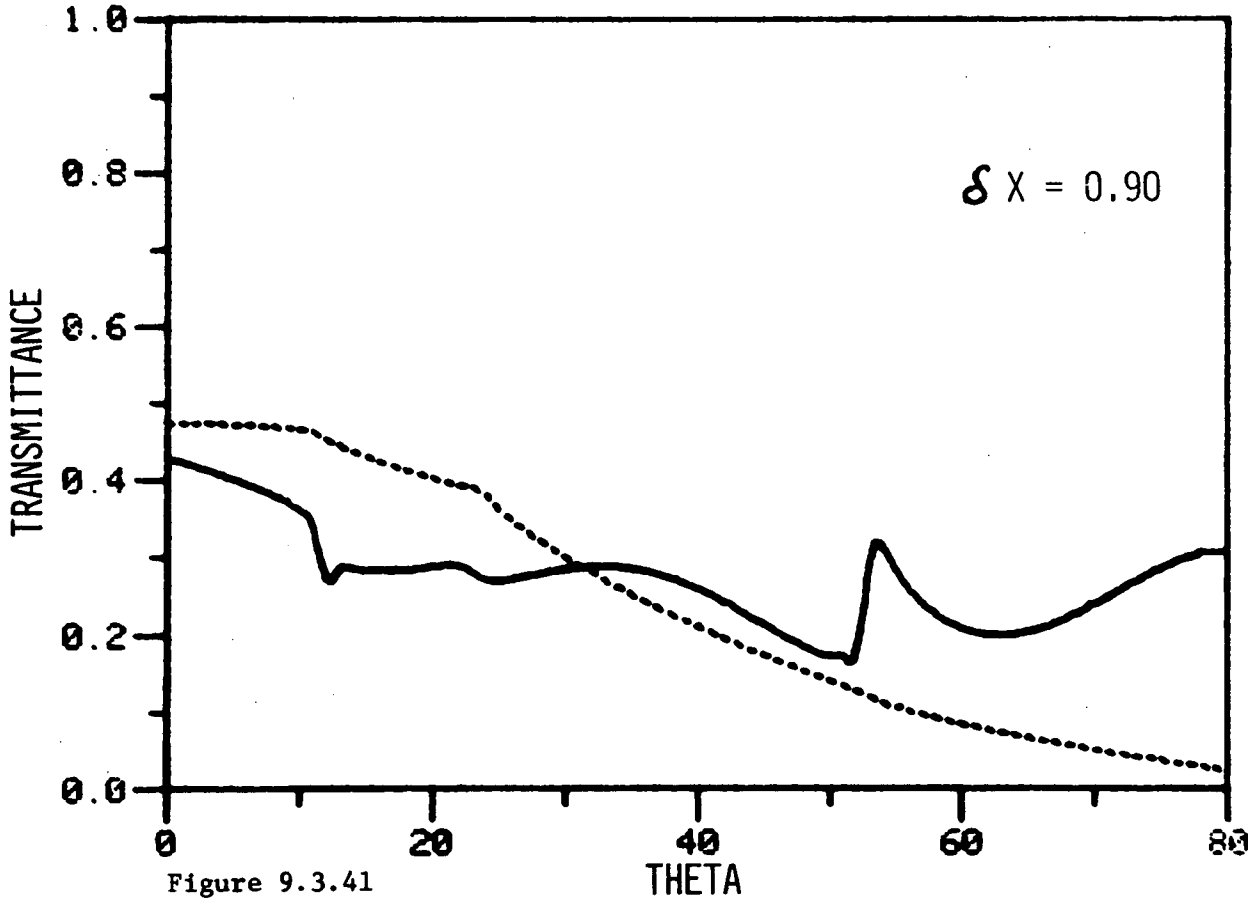


Figure 9.3.41

study still incomplete, it would be unwise to make further speculation.

In summary, it is hoped that the above discussion has adequately elucidated the properties of grating interferometers. From these studies, it is believed that a Fabry-Perot interferometer (for P polarized radiation) having a high finesse may be constructed using a pair of lamellar arrays having narrow apertures and deep grooves. As such, the bulk of the material presented above has centred largely on the performance of structures in P polarized light. (Nevertheless, investigations into the S polarization performance are still being pursued.) However, if the radiation is unpolarized, it is essential to construct the interferometer from a pair of grids - a structure which is currently the subject of an intensive theoretical investigation being undertaken within our group.

9.4 ANALYTIC PROPERTIES OF DIFFRACTION GRATINGS

9.4.1 Introductory Comments

In electromagnetism, and in particular diffraction grating theory, there exist analogues of many of the "conservation theorems" found in nuclear scattering theory. Amongst these properties are conservation of energy and reciprocity, which are independent of the geometry of the structure. Since these have been adequately discussed elsewhere in this thesis, it is not the intention of this section to provide any review of these concepts here.

However, it is possible to establish a number of conservation theorems which rely heavily upon the symmetry of the structure and it is to their derivation that this section is devoted. In line with the

previous theoretical formalisms, it is chosen to elucidate these rules using conventional electromagnetic techniques without resorting to a scattering theory treatment.

It is believed that the best way to approach this problem is to provide some feel for the previous work (e.g. chapter 4) using conventional integral techniques and then move on to a new derivation of these same properties using the concept of time reversibility. Throughout the discussion, particular reference is made to the derivation of conservation relations for the double grating with numerical examples of these properties interspersed in the text. The most interesting of these properties which have been discovered is the existence of a detailed energy balance in the mode structure of the upper and lower arrays for

(i) gratings having both left-right and up-down symmetry operated
  (a) with arbitrary incidence parameters such that only the zeroth orders are propagating, and
  (b) in a first order Littrow configuration; and

(ii) identical lamellar gratings displaced relative to one another (such that neither left-right nor up-down symmetry is displayed) and operated in normally incident light such that $\lambda/d > 1$.

9.4.2 A review of some "conservation relations" derived using conventional integral techniques

9.4.2.1 *Relations for a first order Littrow configuration*

In an earlier chapter, a constraint relating the phases and efficiencies of the reflected and transmitted orders excited by a lossless left-right symmetric transmission grating used in a first order Littrow

mount such that $2/3 < \lambda/d < 2$ was proposed. Although mention has already been made of the importance of symmetry in the derivation of such relations, the significance of the lossless nature of the grating has, until now, been largely omitted. These two properties are of equal importance and should either of them not exist for the particular structure in question, the derivation of these conservation theorems will fail. Throughout the following analysis, clear reference will be made to the implementation of these restrictions.

Now, consider a P polarized plane wave field incident upon the structure shown in figure (9.4.1). Let us postulate that it will excite a field distribution given by the following Rayleigh expansions (outside the grooves of the grating):

$$\text{(i)} \quad E_u(x,y) = \sum_p \left[\delta_{p0} \exp(-jk_0\beta_0 y) + R_p \exp(jk_0\beta_p y)\right] \exp(jk_0\alpha_p x) \text{ in } D_u \tag{9.4.2.1}$$

and

$$\text{(ii)} \quad E_\ell(x,y) = \sum_p \left[T_p \exp(-jk_0\beta_0 y)\right] \exp(jk_0\alpha_p x) \text{ in } D_\ell \ . \tag{9.4.2.2}$$

Since this grating is being operated in a $(-1)^{th}$ order Littrow configuration, the discrete values of α are given by

$$\alpha_p = -\alpha_{-1-p} = (p + \tfrac{1}{2})\ \lambda/d.$$

By now returning an incident field I' along the path of the 0^{th} order (corresponding to a $(+1)^{th}$ order Littrow mount) transmitted wave, a corresponding wave field given by

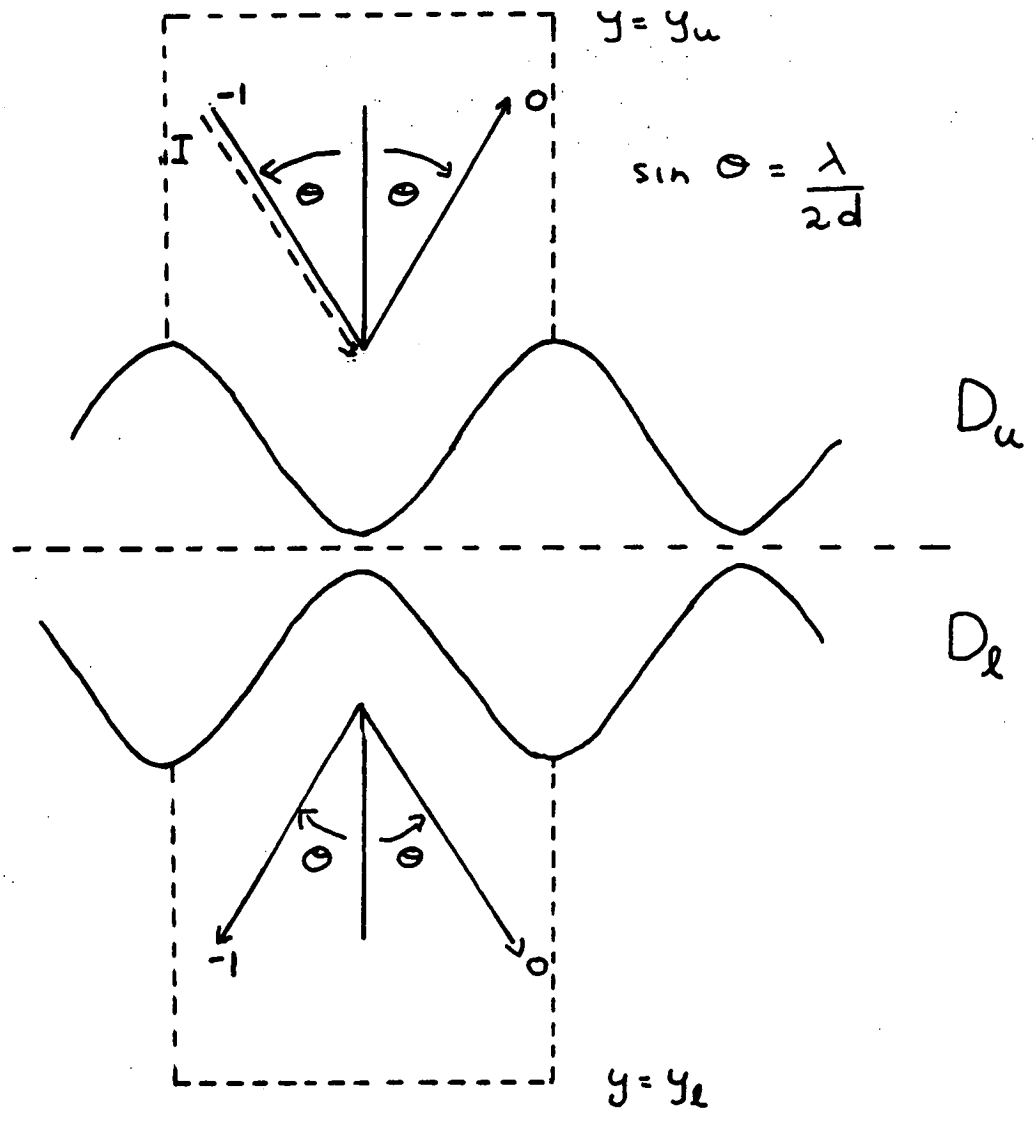


Figure 9.4.1

Integral derivation of the phase constraints for a first order Littrow configuration. By returning an incident field:

(a) along the path traced by the 0^{th} order reflected field of the original problem, it may be shown that

$$\mathrm{Re}(R_0\bar{R}_{-1} + T_0\bar{T}_{-1}) = 0,$$

(b) along the path traced by the 0^{th} order transmitted field of the original problem, it may be shown that

$$\mathrm{Re}(R_0\bar{T}_{-1} + R_{-1}\bar{T}_0) = 0$$

and

(c) along the path traced by the $(-1)^{th}$ order transmitted field of the original problem, it may be shown that

$$\mathrm{Re}(R_0\bar{T}_0 + R_{-1}\bar{T}_{-1}) = 0.$$

$$\text{(i)} \quad E'_u(x,y) = \sum_p [T'_p \exp(jk_0\beta'_p y)] \exp(jk_0\alpha'_p x) \quad \text{in } D_u \tag{9.4.2.4}$$

and

$$\text{(ii)} \quad E'_\ell(x,y) = \sum_p [\delta_{p0} \exp(jk_0\beta'_p y) + R'_p \exp(-jk_0\beta'_p x)] \exp(jk_0\alpha'_p x) \text{ in } D_\ell \tag{9.4.2.5}$$

is established. In this case,

$$\alpha'_p = (p - \tfrac{1}{2})\, \lambda/d = -\alpha_{-p}. \tag{9.4.2.6}$$

Since the structure under consideration possesses both left-right and up-down symmetry it follows that

$$R'_p = R_{-p} \quad \text{and} \quad T'_p = T_{-p}.$$

Now, by defining the superposition of the two fields E and E' in both D_u and D_ℓ to be

$$E^T = E + E',$$

a simple application of Green's Theorem yields

$$\int_0^d [E^T \frac{\partial \bar{E}^T}{\partial y} - \bar{E}^T \frac{\partial E^T}{\partial y}] \Big|_{y=y_u} dx = \int_0^d [E^T \frac{\partial \bar{E}^T}{\partial y} - \bar{E}^T \frac{\partial E^T}{\partial y}] \Big|_{y=y_\ell} dx. \tag{9.4.2.7}$$

(In the above expression, the bar denotes complex conjugation.) It should be noted that such a relation holds only if all media are lossless. Upon integrating this expression explicitly it may be shown that

$$|R_0 + T_{-1}|^2 + |R_{-1} + T_0|^2 = 1. \qquad (9.4.2.8)$$

Since conservation of energy supplies the constraint

$$|R_0|^2 + |R_{-1}|^2 + |T_0|^2 + |T_{-1}|^2 = 1$$

equation (9.4.2.8) may be shown to reduce to the form:

$$\mathrm{Re}(R_0\overline{T}_{-1} + R_{-1}\overline{T}_0) = 0. \qquad (9.4.2.9)$$

(It has already been shown in chapter 8 that for modal formulations such constraints are analytically satisfied provided that the set of Rayleigh orders is truncated symmetrically about the propagating orders. Table (9.4.1) gives evidence supporting this assertion for the case of a double grating.)

A similar discussion presented in chapter 4 centred on a structure possessing only left-right symmetry. The proof given above required the further constraint of up-down symmetry to be imposed. Let us now consider the derivation of a third conservation relation for a structure possessing only up-down symmetry. For this case, it is proposed that a wave be returned along the original path traced by the $(-1)^{th}$ order transmitted ray. Such an incident field would establish a Rayleigh field distribution given by

$$\text{(i)} \quad E'_u(x,y) = \sum_p [T'_p \exp(jk_0\beta_p y)] \exp(jk_0\alpha_p x) \quad \text{in } D_u \qquad (9.4.2.10)$$

$$\text{(ii)} \quad E'_\ell(x,y) = \sum_p [\delta_{p0}\exp(jk_0\beta_0 y) + R'_p \exp(-jk_0\beta_p y)] \exp(jk_0\alpha_p x) \text{ in } D_\ell. \qquad (9.4.2.11)$$

Table 9.4.1

Demonstration of the first order Littrow phase constraints applying to the propagating orders -1 and 0. The following numerical example pertains to a double grating composed of a pair of identical lamellar gratings specified by the parameters c/d = 0.8, h/d = 0.4 separated by S/d = 0.6 displaced, such that $\delta x/d = 0.0$ and used in P polarized light. The grating is operated at a wavelength of $\lambda/d = 1.2001$ such that the order (-1) is in a Littrow configuration. Results obtained from the computer program using 4 waveguide modes and 10 Rayleigh orders were:

$$\rho_0^R = 0.1315478654 \qquad \psi_0^R = -134.37053299^o$$
$$\rho_{-1}^R = 0.4061970027 \qquad \psi_{-1}^R = 44.642000892^o$$
$$\rho_0^T = 0.2323407636 \qquad \psi_0^T = -46.65716855^o$$
$$\rho_{-1}^T = 0.2299143685 \qquad \psi_{-1}^T = -46.66589318^o$$

where ρ and ψ respectively denote the efficiencies and phases of the various orders (denoted by the numerical subscripts) in reflection and transmission (denoted by the superscripts). The above phases have been corrected such that the phase origin lies at x = c/2 - the axis of symmetry for the grating. Using these results one finds that

$$\mathrm{Re}\,[R_0\bar{R}_{-1} + T_0\bar{T}_{-1}] = 4.9 \times 10^{-10}$$
$$\mathrm{Re}\,[R_0\bar{T}_{-1} + R_{-1}\bar{T}_0] = 2.2 \times 10^{-8}$$
$$\mathrm{Re}\,[R_0\bar{T}_0 + R_{-1}\bar{T}_{-1}] = 2.2 \times 10^{-8}$$

These quantities may, effectively, be taken as being zero since they lie within the tolerance of the machine precision (as specified by the difference between the total efficiency sum and unity). The above results were calculated using a "symmetric" distribution of Rayleigh orders

$$\{\, p \mid -1-n \le p \le n \quad \text{for} \quad n \ge 0 \,\}$$

and thus the Littrow expressions are satisfied "analytically", as was discussed in chapter 8.

Because of the up-down symmetry it follows that $R'_p = R_p$ and $T'_p = T_p$. As before the field (E') is superposed upon the original field (E) and this time the spatial invariance of the integral in equation (9.4.2.7) reveals that

$$|R_0 + T_0|^2 + |R_{-1} + T_{-1}|^2 = 1 \qquad (9.4.2.12)$$

which reduces to the form

$$\mathrm{Re}(R_0\bar{T}_0 + R_{-1}\bar{T}_{-1}) = 0.$$

Numerical confirmation of this relation is also presented in table (9.4.1).

9.4.2.2 *A relation for normally incident radiation*

Having discussed the phase constraints appropriate to a first order Littrow mount, let us turn now to discuss the case of normally incident radiation. In this section, consider a lossless up-down symmetric grating operated at normal incidence with light of a sufficiently long wavelength such that only the specularly reflected and transmitted orders are propagating.

Again the primed field (E') is defined to be the result of an incident wave returned along the path of the 0^{th} order transmitted beam. It should be noted that $T'_0 = T_0$ from the reciprocity theorem. However, the equality of R_0 and R'_0 can only be achieved with the application of up-down symmetry. (In the case of non-normal incidence the equality of R_0 and R'_0 can be assured only if left-right symmetry is additionally displayed.)

Now, using the definition,

$$E^T = E + E',$$

the spatial invariance of the integral

$$\int_0^{d_T} [E^T \frac{\partial \bar{E}_T}{\partial y} - \bar{E}^T \frac{\partial E^T}{\partial y}] \, dx$$

shows that

$$|R_0 + T_0|^2 = 1$$

which implies that

$$\arg(R_0) - \arg(T_0) = \tfrac{1}{2}\pi + (\pi)$$

where (π) denotes an arbitrary and indeterminate factor of π.

In summary, this section has attempted to outline some of the phase constraints, derivations of which have been sought theoretically after observing them in our numerical results. In doing so, it is hoped that the consequences of such powerful constraints as conservation of energy and grating symmetry have been clearly elucidated. Let us now turn to a derivation of the same properties using the principle of time reversibility.

9.4.3 The concept of time reversibility

In many of the problems in physics, the assumed direction of time plays no explicit role in their solution. Of course this is true only

for those systems in which energy is conserved. Should one attempt to reverse the temporal dependence of a system which is dissipating energy in the forward time direction, a non-physical solution will result. (In this case, the creation of energy would be witnessed.)

A simple example taken from Blatt and Weisskopf [9.9] involves the motion of the planets around the sun. If the orbit given by the parametric equation:

$$x = f_1(t); \qquad y = f_2(t); \qquad z = f_3(t)$$

satisfies Newton's equations of motion, then the "time-reversed" solution

$$x = f_1(-t); \qquad y = f_2(-t); \qquad z = f_3(-t)$$

is also a physically possible orbit i.e. it also satisfies Newton's equations.

Thus, stated briefly, a system is invariant under time reversal if for every state of that system there exists a time reversed state also satisfying the relevant dynamical equations. Let us now discuss the time reversibility of Maxwell's equations in a lossless medium by drawing largely upon the treatment in Grivet's lucid paper [9.10] on the subject. For a lossless medium, Maxwell's equations state that

$$\text{curl } \underline{H} = \varepsilon \frac{\partial \underline{E}}{\partial t}; \qquad \text{curl } \underline{E} = -\mu \frac{\partial \underline{H}}{\partial t}. \tag{9.4.3.1}$$

By applying the transformation $t' = -t$, these are rewritten in the form

$$\text{curl } \underline{H}' = -\varepsilon \frac{\partial \underline{E}'}{\partial t'}; \qquad \text{curl } \underline{E}' = \mu \frac{\partial \underline{H}'}{\partial t'} \tag{9.4.3.2}$$

where $\underline{E}'$ and $\underline{H}'$ are the time reversed fields. Clearly, Maxwells' equations can be made time invariant by applying either of the following constraints:

$$\underline{H}' = -\underline{H} \qquad \underline{\text{or}} \qquad \underline{E}' = -\underline{E}. \tag{9.4.3.3}$$

(It is necessary to state that in this case the "or" refers to the "exclusive or".) In order to decide between these two, consider the equation of motion of an electron in superposed electric and magnetic fields:

$$m \frac{d^2 \underline{s}}{dt^2} = e(\underline{E} + \frac{d\underline{s}}{dt} \times \underline{B}).$$

Under time inversion, the acceleration is invariant but velocity changes sign. Thus to obtain reciprocity, the chosen transformation appropriate to $t' = -t$ is

$$\underline{E}' = \underline{E} \qquad \text{and} \qquad \underline{H}' = -\underline{H}.$$

Had it been chosen to conserve $\underline{H}$ under time inversion, then a result contravening the reciprocity theorem would have resulted. Of course, such systems are common in nature with a simple example arising in the case of ferromagnetic materials. Should the polarizing field $\underline{H}$ retain its parity under time inversion, then the reversed wave may differ strongly from the incident wave.

Thus, for a lossless medium, it has been established that Maxwell's equations are invariant under the transformation set

$$(t,\underline{E},\underline{H}) \rightarrow (-t, \underline{E}, -\underline{H}). \qquad (9.4.3.4)$$

In the case of a lossy isotropic medium Maxwell's equations are written in the form:

$$\text{curl } \underline{H} = \sigma\underline{E} + \frac{\partial \underline{E}}{\partial t} \rightarrow \text{curl } \underline{H} = -\sigma\underline{E} + \frac{\partial \underline{E}}{\partial t} \qquad (9.4.3.5)$$

which is no longer invariant under transformation (9.4.3.4). Thus it is clear that time reversibility is valid only for lossless media. Intuitively, the reversal involves the creation of energy but, as shown above prevents the invariance of Maxwell's equations.

It is often convenient to express Maxwell's equations in the frequency domain. In this case the necessary transformation is $(t,\underline{E},\underline{H}) \rightarrow (-t,\underline{E},\underline{H})$ followed by complex conjugation.

9.4.4 Derivation of phase properties using the concept of time-reversibility

Let us now turn our attention to the derivation of a number of conservation relations using the concept of time reversibility.

9.4.4.1 *The case of a grating possessing up-down and/or left-right symmetry*

9.4.4.1.1 *Long wavelength formulation*

The following treatment applies only to long wavelengths for which a single undispersed order propagates in both reflection and transmission.

Although the theory presented here assumes the existence of a P polarized wave field, the same derivations apply to the orthogonal polarization.

Consider a plane wave I incident upon the lossless structure, shown in figure (9.4.2), at some angle θ giving rise to the plane wave output energy channels R_0 and T_0 (state 1). Upon application of time reversibility, there are now two input channels $\overline{R}_0$ and $\overline{T}_0$ and a single output channel $\overline{I}$. (Here, the term channel refers to a plane wave mode i.e. a path along which energy can be carried to infinity). Let us now consider the far field behaviour (where the evanescent order terms play no part in the calculation) and at this stage invoke only the left-right symmetry of the structure.

The problem of time reversal is most easily resolved by considering each of the input channels separately as shown in states (2a) and (2b) and by then superposing their resultant fields as shown in stage 2 of figure (9.4.3). In these diagrams each plane wave is represented by a vector associated with which is its amplitude. Next to each of the wave amplitudes is a bracketed quantity supplying a brief explanation as to its derivation. The legend is as follows:

LR - left-right symmetry

RT - Reciprocity Theorem

and NC - no constraint.

In this example, there can be no constraint linking the complex amplitudes R_0' and R_0 unless up-down symmetry is also supplied.

Now, under time reversal there can be only one outgoing wave $\overline{I}$ of amplitude 1 traversing the same path as the original incident beam I.

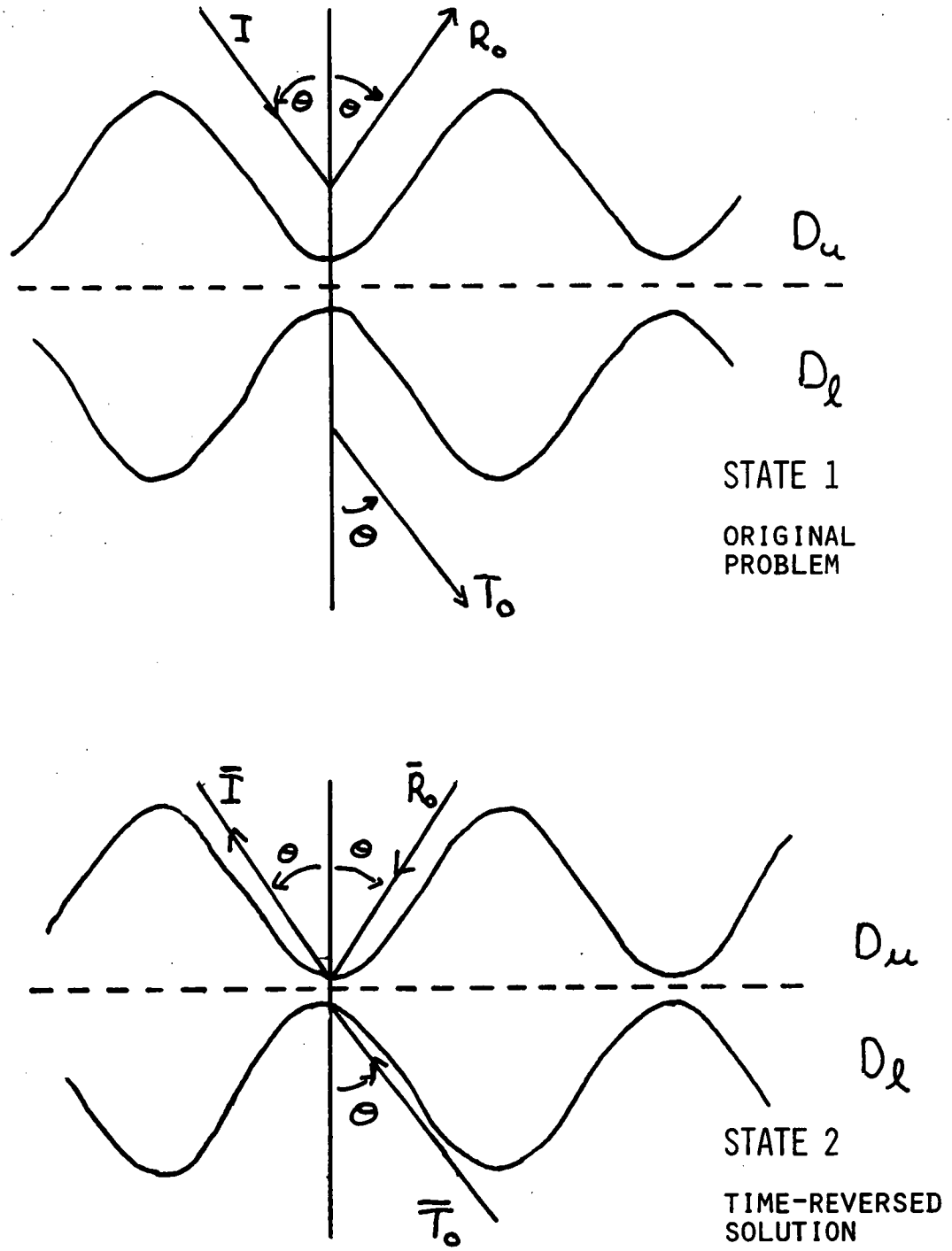


Figure 9.4.2 The diffracting arrangement for the time-reversed solution applying to the long-wavelength formulation.

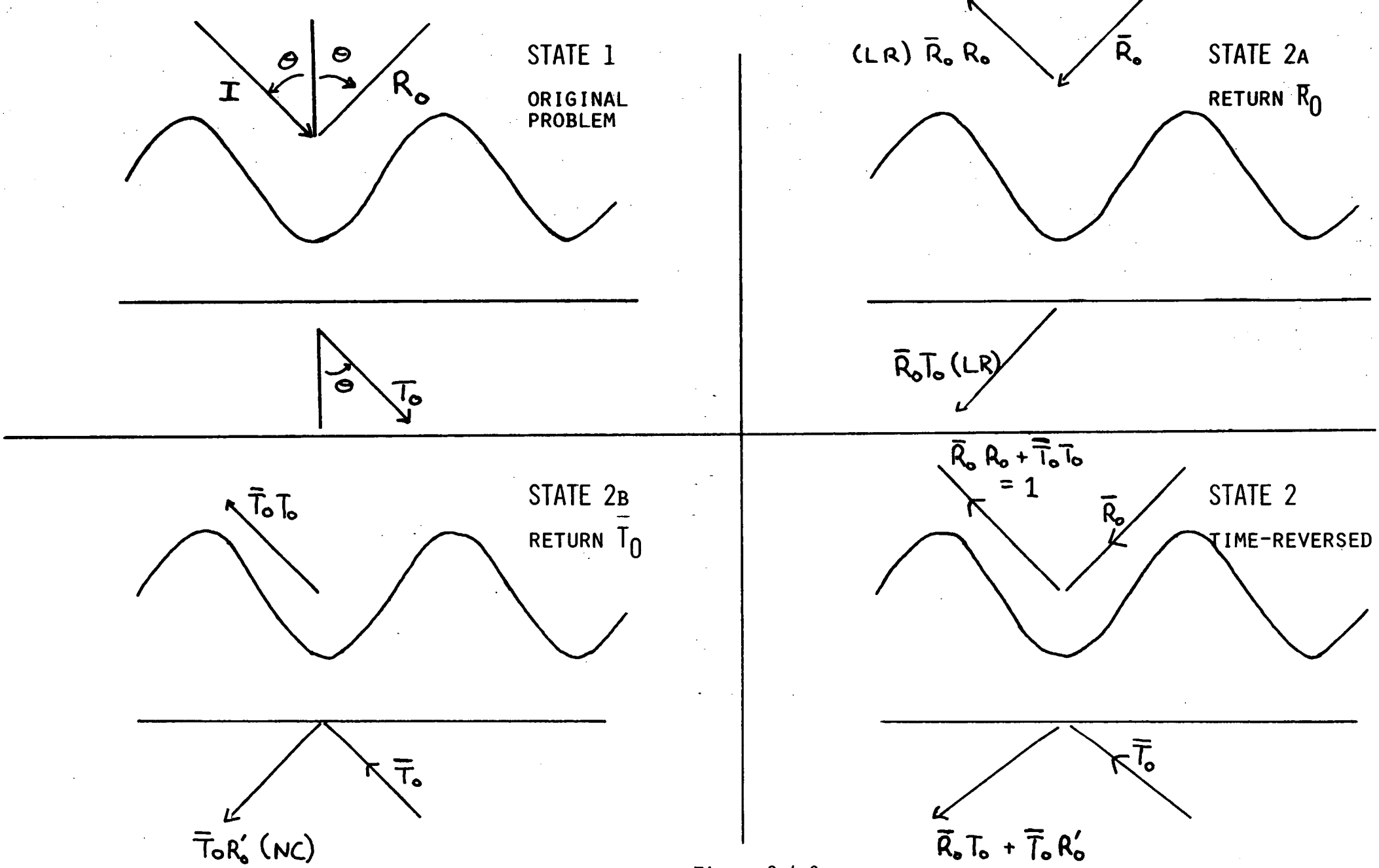

STATE 1
ORIGINAL PROBLEM
I
θ
θ
R₀
T₀
STATE 2A
RETURN R̄₀
(LR) R̄₀ R₀
R̄₀
R̄₀T₀ (LR)
STATE 2B
RETURN T̄₀
T̄₀T₀
T̄₀
T̄₀R′₀ (NC)
STATE 2
TIME-REVERSED
R̄₀ R₀ + T̄₀T₀ = 1
R̄₀
T̄₀
R̄₀T₀ + T̄₀R′₀

Figure 9.4.3

Clearly, from state 2, it can be seen that this is the case since

$$|R_0|^2 + |T_0|^2 = 1 \ . \qquad \text{(conservation of energy)}$$

Furthermore,

$$\bar{R}_0 \, T_0 + \bar{T}_0 \, R_0' = 0$$

which reduces to the form

$$\arg(R_0) - \arg(T_0) = \pi/2 + (\pi)$$

should up-down symmetry be also imposed.

By now invoking the uniqueness condition, these far field results can be extended to all space. (Here one asserts that if two fields diffracted by the same structure exhibit the same far-field behaviour, then their difference has zero far field and is zero everywhere in space.) In the regions D_u and D_ℓ respectively, let us define field distributions $E_u(x,y)$ and $E_\ell(x,y)$ to be excited by the incident field of state 1. Time reversal then implies that the fields in D_u and D_ℓ, excited by the incident fields $\bar{R}_0$ and $\bar{T}_0$ of state 2, are given by $\bar{E}_u$ and $\bar{E}_\ell$ respectively.

In D_u, the field $\bar{E}_u$ is composed of the superposition of

(i) a reflected component due to the incident field $\bar{R}_0$

and

(ii) a transmitted component due to the incident field $\bar{T}_0$.

For the case of (i), the plane wave in the input channel $\bar{R}_0$ is incident upon a left-right symmetric structure at the same angle as the original incident wave I except that it is now propagating in the (-x) direction. Thus it follows that the wave field so established is given by $\bar{R}_0\ E_u(-x,y)$. In the case of part (ii), the transmitted component $\bar{T}_0\ E_\ell(-x,-y)$ follows from application of the Reciprocity Theorem, left-right symmetry and invocation of the uniqueness of the far field behaviour. Thus, for all left-right symmetric structures,

$$\bar{E}_u(x,y) = \bar{R}_0\ E_u(-x,y) + \bar{T}_0\ E_\ell(-x,-y). \qquad (9.4.4.1)$$

On the other hand, a consideration of the time reversed field $\bar{E}_\ell$ in D_ℓ reveals that it is not possible to express this quantity as a linear superposition of E_u and E_ℓ unless up-down symmetry is also imposed. An inspection of state (2b) of figure (9.4.3), showing far field behaviour, demonstrates that only in the case of the additional up-down symmetry can R_0' be equal to R_0, so that the conservation relations are "internal" to this problem only. By the term "internal", it is meant that the time-inverted diffraction problem can be related purely to the problem in the forward time direction without having to involve the results of a further diffracting arrangement. Thus, with this further constraint it may be shown that $\bar{E}_\ell$ is given by a superposition of the following

(i) a reflected component of the field in D_ℓ: $\bar{R}_0\ E_\ell(-x,y)$

and

(ii) a transmitted component of the field in D_u: $\bar{T}_0\ E_u(-x,-y)$.

(Once again these are assured by the symmetry of the structure, reciprocity and the uniqueness of the far field problem.) It then follows that

$$\bar{E}_\ell(x,y) = \bar{T}_0 E_u(-x,-y) + \bar{R}_0 E_\ell(-x,y). \qquad (9.4.4.2)$$

Equations (9.4.4.1-2) are the general constraints on the whole lossless symmetric system. In the following analysis, explicit field representations appropriate to their various domains of convergence will be applied in an attempt to extract a number of explicit phase properties.

9.4.4.1.2 *Application to the case of Rayleigh fields*

For those regions of D_u and D_ℓ not lying within the grooves of the grating the fields E_u and E_ℓ may be expanded in terms of Rayleigh series:

$$\text{viz.} \quad E_u(x,y) = \sum_p [\delta_{p0}\exp(-jk_0\beta_0 y) + R_p\exp(jk_0\beta_p y)] \exp(jk_0\alpha_p x) \qquad (9.4.4.3)$$

$$\text{and} \quad E_\ell(x,y) = \sum_p [T_p \exp(-jk_0\beta_p y)] \exp(jk_0\alpha_p x) \qquad (9.4.4.4)$$

On substituting these field representations into equation (9.4.4.2) and equating the coefficients of like plane wave terms it may be shown that:

(i) the contribution from the real order yields

$$R_0 \bar{T}_0 + \bar{R}_0 T_0 = 0$$

$$\text{i.e. } \arg(R_0) - \arg(T_0) = \pi/2 + (\pi); \qquad (9.4.4.5)$$

and

(ii) evanescent order terms give rise to

$$\bar{T}_p = \bar{R}_0 T_p + \bar{T}_0 R_p. \qquad (9.4.4.6)$$

Similarly from equation (9.4.4.1) it is possible to extract

(i) $|R_0|^2 + |T_0|^2 = 1$ (conservation of energy) (9.4.4.7)

from the real order term ($p = 0$)

and

(ii) $\bar{R}_p = \bar{R}_0 R_p + \bar{T}_0 T_p$ (9.4.4.8)

from the evanescent orders.

The significance of the real order contributions may be readily seen from figure (9.4.4).

Now consider the evanescent orders. By appropriately adding and subtracting equations (9.4.4.6) and (9.4.4.8) it may be shown that for all $p \neq 0$

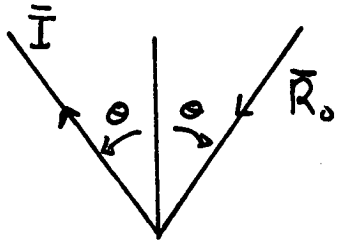
$$A_p/\bar{A}_p = A_0 \quad \text{i.e. } \arg(A_p) = \tfrac{1}{2}\arg(A_0) + (\pi) \qquad (9.4.4.9)$$

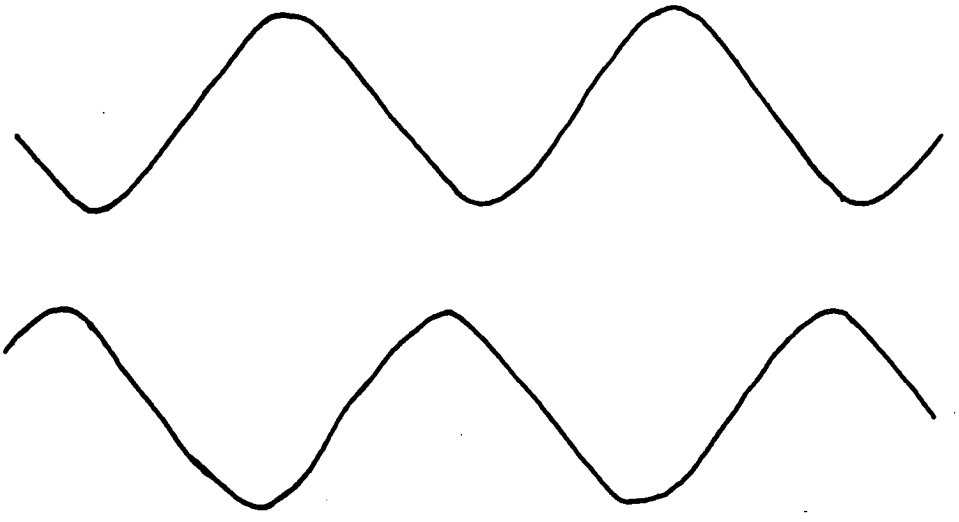
and $$B_p/\bar{B}_p = B_0 \quad \text{i.e. } \arg(B_p) = \tfrac{1}{2}\arg(B_0) + (\pi) \qquad (9.4.4.10)$$

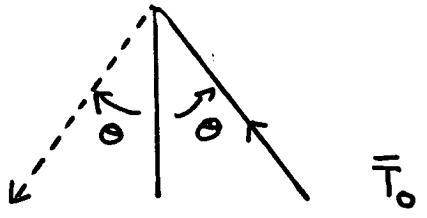
where $A_p = R_p + T_p$

and $B_p = R_p - T_p$.

$$\bar{R}_0 R_0 + \bar{T}_0 T_0 = 1$$







$$\bar{R}_0 T_0 + \bar{T}_0 R_0 = 0$$

I.E. $\text{ARG}(R_0) - \text{ARG}(T_0) = \pi/2 + (\pi)$

Figure 9.4.4

Table 9.4.2

Demonstration of the long wavelength Rayleigh field constraints for a double grating composed of identical lamellar gratings characterized by $c/d = 0.8$, and $h/d = 0.4$ separated from one another by $S/d = 0.6$ and displaced relative to one another by an amount $\delta x/d = 0.0$. The structure was operated in P polarized radiation having a normalized wavelength of 1.8001 impinging on the grating at an angle of 35°. The following calculations were performed with 4 waveguide modes in each array and 11 Rayleigh orders.

(1) Real order - $p = 0$

$\rho_0^R = 0.884018$ $\qquad$ $\psi_0^R = 25.824038^{\circ}$

$\rho_0^T = 0.115982$ $\qquad$ $\psi_0^T = 115.824038^{\circ}$

Thus $\rho_0^R + \rho_0^T = 1$

and $\psi_0^T - \psi_0^R = \pi/2$

(2) Evanescent order - $p = -1$

$|R_{-1}|^2 = 2.049517 \times 10^{1}$ $\qquad$ $\psi_{-1}^R = -164.342390^{\circ}$

$|T_{-1}|^2 = 2.156364 \times 10^{0}$ $\qquad$ $\psi_{-1}^R = -134.369087^{\circ}$

Thus $\arg(A_0) = 45.735190^{\circ}$

$\arg(B_0) = 5.912986^{\circ}$

$\arg(A_{-1}) = -157.132430^{\circ}$

and $\arg(B_{-1}) = -177.043520^{\circ}$

demonstrating that

$$\arg(A_{-1}) = \tfrac{1}{2}\arg(A_0) + \pi, \text{ and}$$

$$\arg(B_{-1}) = \tfrac{1}{2}\arg(B_0) + \pi.$$

These equations may be further reduced by noting that

$$\arg(A_0) = \arg(R_0) \pm \arctan(|T_0/R_0|)$$

$$\arg(B_0) = \arg(R_0) \mp \arctan(|T_0/R_0|).$$

Here, the arbitrary choice of sign arises from the undefined additive factor of π in the expression

$$\arg(R_0) - \arg(T_0) = \pi/2 + (\pi).$$

In the case of the double grating, both up-down and left-right symmetry can only be simultaneously achieved if the arrays are identical and aligned such that the array phasing term $\delta x = 0$. Evidence of these phase relations is presented in table (9.4.2).

9.4.4.1.3 *Application to the modal quantities of a double grating*

Let us now consider the significance of the conservation relations (equations (9.4.4.1-2)) to the modal fields of the double grating. The fields inside the grooves of either the upper or lower array can be expressed as expansions of left-right symmetric and anti-symmetric modes denoted by $E_m(x')$ and $O_m(x')$ such that

$$E_m(x') = E_m(-x')$$

and $$O_m(x') = -O_m(-x').$$

(The origin of the primed co-ordinate system is centred at $x = c/2$ - the

axis of symmetry of the grating

i.e. $x' = x - c/2.$)

In terms of the previous analysis,

$$E_m(x') = \sin(\frac{m\pi}{2}) \cos(\frac{m\pi x'}{c}) \qquad \text{for } m \text{ odd}$$

and $$O_m(x') = \cos(\frac{m\pi}{2}) \sin(\frac{m\pi x'}{c}) \qquad \text{for } m \text{ even.}$$

Now inside the grooves of the upper array, the P polarized wave field may be expanded as follows:

$$E_u(x,y) = \sum_{m=1}^{\infty} [a_m^E \, \text{rsin}(k_0\mu_m y') + b_m^E \, \text{rcos}(k_0\mu_m y')] \, E_m(x')$$

$$+ \sum_{m=1}^{\infty} [a_m^O \, \text{rsin}(k_0\mu_m y') + b_m^O \, \text{rcos}(k_0\mu_m y')] \, O_m(x') \qquad (9.4.4.11)$$

where the function "rsin" is given by

$$\text{rsin}(k_0\mu_m y') = \frac{\sin(k_0\mu_m y')}{\sin(k_0\mu_m h/2)} \qquad \text{where } y' = y - s.$$

(rcos is defined in an analogous fashion.)

This representation has been chosen so as to yield y dependent functions which are purely real (i.e. unaffected by the complex nature of μ_m). Similarly in the grooves of the lower grating, the field is expanded in terms of the following series:

$$E_\ell(x',y'') = \sum_{m=1}^{\infty} [\hat{a}_m^E \, r\sin(k_{0m}\mu_m y'') + \hat{b}_m^E \, r\cos(k_{0m}\mu_m y'')] \, E_m(x')$$

$$+ \sum_{m=1}^{\infty} [\hat{a}_m^O \, r\sin(k_{0m}\mu_m y'') + \hat{b}_m^O \, r\cos(k_{0m}\mu_m y'')] \, O_m(x') \qquad (9.4.4.12)$$

where $y'' = y + s$. It is to be noted that $\hat{\mu}_m = \mu_m$ since both arrays have been assumed to be identical.

By substituting these field quantities into the constraint

$$\bar{E}_\ell(x',y) = \bar{R}_0 \, E_\ell(-x',y) + \bar{T}_0 E_u(-x',-y)$$

and comparing the coefficients of like modes it may be shown that

$$\left.\begin{aligned}
\bar{\hat{a}}_m^E &= \bar{R}_0 \, \hat{a}_m^E - \bar{T}_0 \, a_m^E \\
\bar{\hat{a}}_m^O &= -\bar{R}_0 \, \hat{a}_m^O + \bar{T}_0 \, a_m^O \\
\bar{\hat{b}}_m^E &= \bar{R}_0 \, \hat{b}_m^E + \bar{T}_0 \, b_m^E \\
\hat{b}_m^O &= -\bar{R}_0 \, \hat{b}_m^O - \bar{T}_0 \, b_m^O
\end{aligned}\right\} \qquad (9.4.4.13)$$

Similarly, from

$$\bar{E}_u(x',y) = \bar{R}_0 \, E_u(-x',y) + \bar{T}_0 \, E_\ell(-x',-y)$$

it follows that

$$\left.\begin{aligned}
\bar{a}_m^E &= \bar{R}_0\, a_m^E - \bar{T}_0\, \hat{a}_m^E \\
\bar{a}_m^O &= -\bar{R}_0\, a_m^O + \bar{T}_0\, \hat{a}_m^O \\
\bar{b}_m^E &= \bar{R}_0\, b_m^E + \bar{T}_0\, \hat{b}_m^E \\
\bar{b}_m^O &= -\bar{R}_0\, b_m^O - \bar{T}_0\, \hat{b}_m^O \quad .
\end{aligned}\right\} \qquad (9.4.4.14)$$

Although the above analysis has been quite explicit in the sense that it is appropriate only to the double grating, the method, nevertheless, is relevant to all modal formalisms. It is only necessary to separate the modes into real symmetric and anti-symmetric components for the basic method to still be applicable. In particular, the y dependent terms "rsin" and "rcos", whose form is appropriate only to those structures having vertical groove walls, have been used purely as manifestations of vertical symmetry and anti-symmetry and nowhere has their analytic form been exploited.

Now, after some elementary manipulation of equations (9.4.4.13-14) it may be shown that

$$\frac{(a_m^E + \hat{a}_m^E)}{(\bar{a}_m^E + \bar{\hat{a}}_m^E)} = B_0$$

$$\text{i.e.} \quad \arg(a_m^E + \hat{a}_m^E) = \tfrac{1}{2} \arg(B_0) + (\pi) \qquad (9.4.4.15)$$

Similarly,

$$\arg(a_m^E - \hat{a}_m^E) = \tfrac{1}{2} \arg(A_0) + (\pi). \qquad (9.4.4.16)$$

Naturally, further relations in terms of the other modal quantities may be obtained but these are omitted for the sake of brevity. These expressions are the analogues of the common phase properties of inductive grids [9.8]. Some sample results which we have observed (using our computer program) confirming this property are presented in table (9.4.3).

Although these results may not appear to be of any great importance, they do enable us to explain a startling phenomenon which may be referred to as the "detailed energy balance" in the mode structure of the two arrays.

9.4.4.1.4 *Detailed balance of modal fluxes*

This rather fascinating property was observed numerically and its presence immediately prompted a search for an analytic proof of its existence. Before commencing the mathematical treatment, let us accurately define the result. In the upper and lower arrays, it is possible to calculate a set of mode fluxes $\{F_m^{E/O}\}$ and $\{\hat{F}_m^{E/O}\}$ respectively. Should the double grating possess both left-right and up-down symmetry, then the flux of each mode is array invariant:

$$\text{i.e.} \quad F_m^{E/O} = \hat{F}_m^{E/O} \qquad \text{for all values of m.} \tag{9.4.4.17}$$

(The derivation presented here applies only to the case of a single propagating order in both reflection and transmission). This flux invariance is equivalent to

$$\mathrm{Im}(\hat{a}_m \bar{\hat{b}}_m)\Big|_{E/O} = \mathrm{Im}(a_m \bar{b}_m)\Big|_{E/O}$$

(where the modal quantities can be superscripted with either an E or an *O* depending on their symmetry).

Consider now the derivation of this property for the even modes. Here,

$$\bar{\hat{a}}_m^E = \bar{R}_0 \hat{a}_m^E - \bar{T}_0 a_m^E \qquad (9.4.4.18a)$$

$$\bar{a}_m^E = \bar{R}_0 a_m^E - \bar{T}_0 \hat{a}_m^E \qquad (9.4.4.18b)$$

$$\bar{\hat{b}}_m^E = \bar{R}_0 \hat{b}_m^E + \bar{T}_0 b_m^E \qquad (9.4.4.18c)$$

$$\bar{b}_m^E = \bar{R}_0 b_m^E + \bar{T}_0 \hat{b}_m^E . \qquad (9.4.4.18d)$$

After some elementary manipulation of equations (9.4.4.18a-b) it may be shown that

$$\bar{\hat{a}}_m^E = \frac{\bar{R}_0}{\bar{T}_0 R_0} (a_m^E - R_0 \bar{a}_m^E). \qquad (9.4.4.19)$$

Similarly, from equations (9.4.4.18c-d) one derives

$$\bar{\hat{b}}_m^E = - \frac{\bar{R}_0}{\bar{T}_0 R_0} (b_m^E - R_0 \bar{b}_m^E) \qquad (9.4.4.20)$$

and thus it follows that

$$\mathrm{Im}(\hat{a}_m^E \bar{\hat{b}}_m^E) = \mathrm{Im}\{ \frac{1}{|T_0|^2} [2\mathrm{Re}(\bar{R}_0 a_m^E b_m^E) - |R_0|^2 a_m^E \bar{b}_m^E - \bar{a}_m^E b_m^E]\}$$

$$= \mathrm{Im}(a_m^E \bar{b}_m^E) . \qquad (9.4.4.21)$$

Table 9.4.3

Demonstration of the phase constraints for the modal quantities applying to the long wavelength formulation. Also displayed is the phenomenon referred to as the "detailed balance". The diffracting arrangement is identical to that referred to in table (9.4.2). The following results were computed using 4 waveguide modes and 11 Rayleigh orders.

m	x-symmetry	$\arg(a_m+\hat{a}_m)+(\pi)$	$\arg(a_m-\hat{a}_m)+(\pi)$	F_m	$\hat{F}_m$
1	E	-177.044^o	-157.132^o	0.09095	0.09095
2	O	-177.044^o	-157.132^o	0.00038	0.00038
3	E	-177.044^o	-157.132^o	0.00006	0.00006
4	O	-177.044^o	-157.132^o	0.00000	0.00000
5	E	-177.044^o	-157.132^o	0.00000	0.00000

These results demonstrate the properties

$$\arg(a_m^E + \hat{a}_m^E) = \tfrac{1}{2}\arg(R_0 - T_0) + (\pi)$$

$$\arg(a_m^E - \hat{a}_m^E) = \tfrac{1}{2}\arg(R_0 + T_0) + (\pi)$$

$$\arg(a_m^O + \hat{a}_m^O) = \tfrac{1}{2}\arg(T_0 - R_0) + (\pi)$$

$$\arg(a_m^O - \hat{a}_m^O) = \tfrac{1}{2}\arg(R_0 + T_0) + (\pi)$$

Numerical confirmation can be gained by comparing the above numbers with the quantities $\arg(A_0)$ and $\arg(B_0)$ evaluated in the previous table. These are the singly periodic analogues of the common phase

properties for grids discussed by McPhedran and Maystre [9.8]. It is interesting to note that these properties are analytically satisfied by the formalism independent of the choice of waveguide modes or Rayleigh orders (provided that the set of orders embraces all the propagating waves). This is to be expected in view of the integral derivation presented in section (9.2) which revealed that such results are intimately related to the energy properties of the structure. Note also that the monomodal approximation, used in section (9.3), would appear to be vindicated by the diminished energy fluxes (for higher order modes) shown in the above table.

Naturally, a similar result holds for the anti-symmetric modes. Examples of these properties are shown in table (9.4.3).

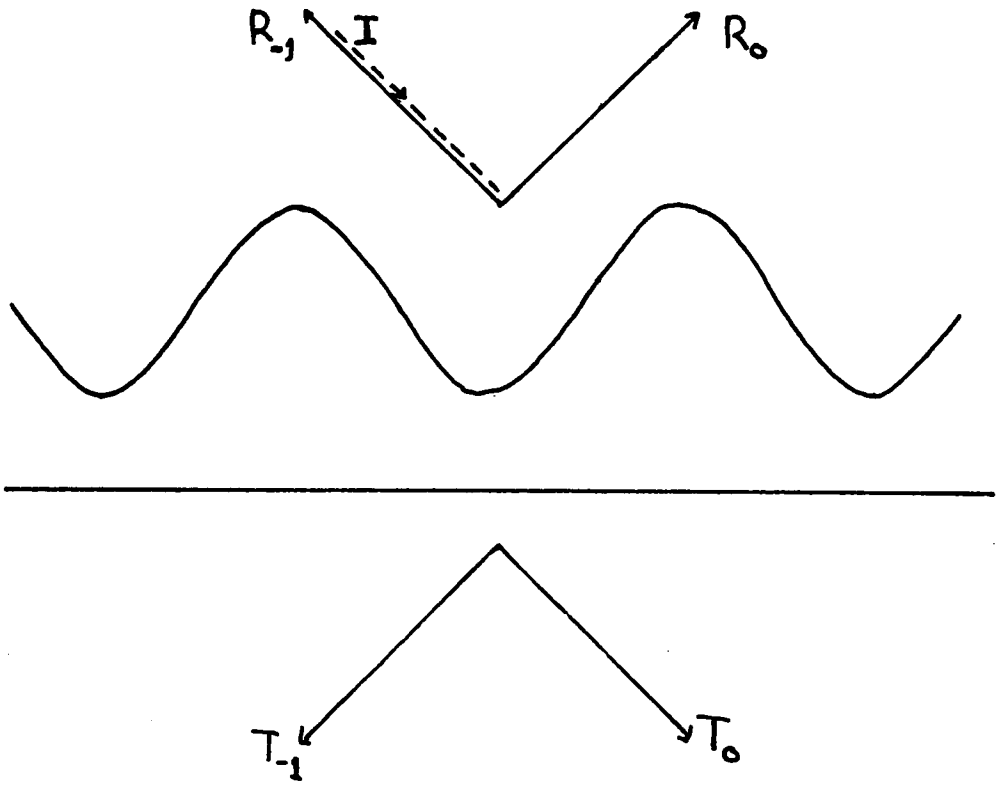
Although these phenomena can be explained in terms of the modal conservation relations, one does not admit to understanding their true physical significance since it is not possible to "bring to mind" any analogue of this system possessing such properties.

9.4.4.1.5 *Application to the case of a first order Littrow mount*

The previous treatment has been concerned only with the case of a single propagating order both in reflection and transmission. This analysis permitted any combination of incidence parameters provided that there was only a pair of outgoing plane wave channels. In general, it is not possible to devise further conservation relations for arbitrary wavelength and angle of incidence combinations. However, there is one very specific combination of these parameters which does permit the derivation of a further set of analytic phase constraints. It is the operation of a lossless symmetric grating in a first order Littrow mount such that only the orders (-1) and (0) are not evanescent. Clearly, the angular symmetry of the far field radiation pattern shown in state (1) of figure (9.4.5) is sufficient to permit a simple analytic treatment of the problem using time reversibility. (It is believed that time reversibility can only be usefully applied to a maximum of two symmetrically displaced rays in either of the spaces D_u and D_ℓ.

Let us now consider the far field distribution in the same manner as before and once again initially impose only left-right symmetry. Now under time inversion, there are four ingoing channels of energy giving rise to a single outgoing wave. Thus in D_u, this means that

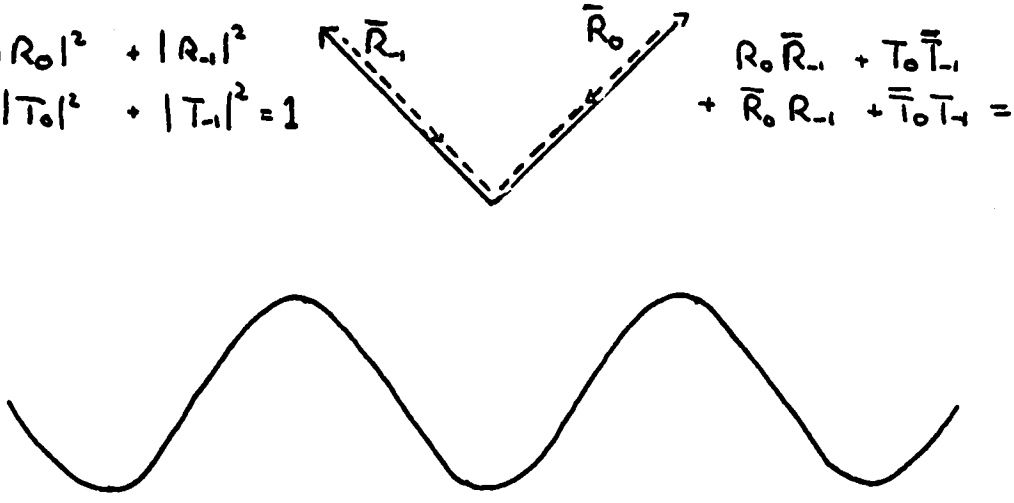
STATE 1 - ORIGINAL PROBLEM



STATE 2 - TIME-REVERSED SOLUTION

$$|R_0|^2 + |R_{-1}|^2 + |T_0|^2 + |T_{-1}|^2 = 1$$

$$R_0\bar{R}_{-1} + T_0\bar{T}_{-1} + \bar{R}_0 R_{-1} + \bar{T}_0 T_{-1} = 0$$



$$\bar{R}_{-1}T_{-1} + \bar{R}_0 T_0 + R'_{-1}\bar{T}_{-1} + R'_0\bar{T}_0 = 0$$

$$\bar{R}_{-1}T_0 + \bar{R}_0 T_{-1} + R'_{-1}\bar{T}_0 + R'_0\bar{T}_{-1}$$

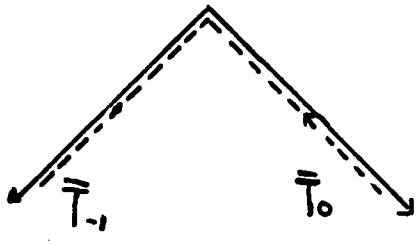


Figure 9.4.5

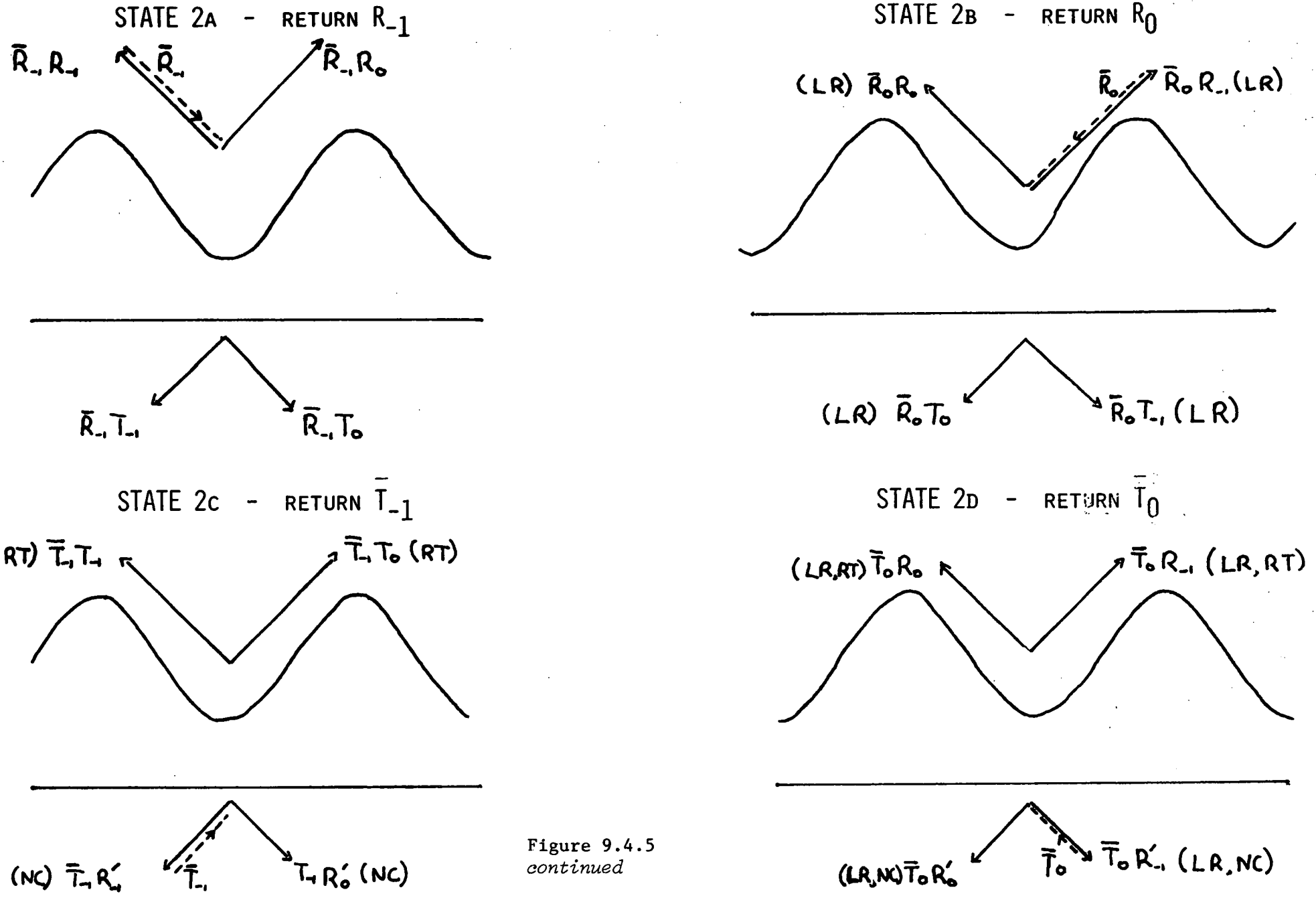


Figure 9.4.5
continued

(i) $|R_{-1}|^2 + |R_0|^2 + |T_{-1}|^2 + |T_0|^2 = 1$ (conservation of energy)

(9.4.4.27)

i.e. the sole outgoing wave has unit amplitude -

and

(ii) $\mathrm{Re}(R_0 \bar{R}_{-1} + T_0 \bar{T}_{-1}) = 0$ (9.4.4.23)

which is in agreement with the result presented in chapter 4.

(Of course, both this and the earlier derivation are essentially the same with both concerned only with far field or energy-carrying terms.)

For fields in D_ℓ however, explicit relations internal to this problem can only be obtained by applying up-down symmetry as well. (This is sufficient to constrain $R'_{-1} = R_{-1}$ and $R'_0 = R_0$ and thus remove the "NC" classification exhibited for waves (in D_ℓ) of states (2c) and (2d) of figure (9.9.5).) Hence, for this case one derives the extra relations:

$$\mathrm{Re}(R_{-1} \bar{T}_{-1} + R_0 \bar{T}_0) = 0 \qquad (9.4.4.24)$$

and $$\mathrm{Re}(R_0 \bar{T}_{-1} + R_{-1} \bar{T}_0) = 0. \qquad (9.4.4.25)$$

Once again, from the uniqueness of the far field behaviour, it is possible to write down relations appropriate to the entire fields. These are:

(i) $\bar{E}_u(x,y) = \bar{R}_{-1}E_u(x,y) + \bar{R}_0E_u(-x,y) + \bar{T}_{-1}E_\ell(x,-y) + \bar{T}_0E_\ell(-x,-y)$

(9.4.4.26)

(which is true for any structure of left-right symmetry only)

and

$$\text{(ii)}\quad \bar{E}_{\ell}(x,y) = \bar{T}_{-1}E_u(x,-y) + \bar{T}_0E_u(-x,-y) + \bar{R}_{-1}E_{\ell}(x,y) + \bar{R}_0E_{\ell}(-x,y) \qquad (9.4.4.27)$$

(which holds only for those structures additionally possessing up-down symmetry).

By now substituting the Rayleigh expansions for E_u and E_{ℓ} into the above equations and equating the coefficients of like plane wave terms, it is possible to extract a number of analytic conservation relations. The constraints for the propagating orders are identical to those derived in the far field treatment. Contributions from the evanescent orders ($p \neq 0,-1$) yield

(i) from equation (9.4.4.26)

$$\bar{R}_{-1-p} = \bar{R}_{-1}R_p + \bar{R}_0R_{-1-p} + \bar{T}_{-1}T_p + \bar{T}_0T_{-1-p} \qquad (9.4.4.28)$$

and

(ii) from equations (9.4.4.27)

$$\bar{T}_{-1-p} = \bar{T}_{-1}R_p + \bar{T}_0R_{-1-p} + \bar{R}_{-1}T_p + \bar{R}_0T_{-1-p}\ . \qquad (9.4.4.29)$$

By defining

$$A_p = R_p + T_p$$

and $$B_p = R_p - T_p,$$

then after some trivial manipulation it may be shown that

$$\left.\begin{aligned} \arg(A_p + A_{-1-p}) &= \tfrac{1}{2}\arg(A_0 + A_{-1}) + (\pi) \\ \arg(B_p + B_{-1-p}) &= \tfrac{1}{2}\arg(B_0 + B_{-1}) + (\pi). \end{aligned}\right\} \quad \forall \quad p \neq 0,-1$$

Let us now turn our attention to the significance of the conservation relations to the modal field quantities for a completely symmetric double grating. In this case, one derives expressions of the form

$$\left.\begin{aligned}
\bar{\hat{a}}_m^E &= \bar{R}^+ \hat{a}_m^E - \bar{T}^+ a_m^E \\
\bar{a}_m^E &= \bar{R}^+ a_m^E - \bar{T}^+ \hat{a}_m^E \\
\bar{\hat{b}}_m^E &= \bar{R}^+ \hat{b}_m^E + \bar{T}^+ \hat{b}_m^E \\
\bar{b}_m^E &= \bar{R}^+ b_m^E + \bar{T}^+ \hat{b}_m^E
\end{aligned}\right\} \qquad (9.4.4.30)$$

where

$$R^+ = R_0 + R_{-1}$$

and $T^+ = T_0 + T_{-1}$.

Naturally, similar relations exist for the x-anti-symmetric modal coefficients but for the sake of brevity it is chosen to reference only the above explicitly. It is to be noted that

$$\begin{aligned}
1 - |T^+|^2 &= 1 - |T_0|^2 - |T_{-1}|^2 - 2\mathrm{Re}(T_0 \bar{T}_{-1}) \\
&= |R_0|^2 + |R_{-1}|^2 + 2\mathrm{Re}(R_0 \bar{R}_{-1}) \\
&= |R^+|^2
\end{aligned}$$

and this together with the similarity of equations (9.4.4.30) and

Table 9.4.4

Demonstration of the "detailed balance" phenomenon for the $(-1)^{th}$ order Littrow mount. The geometry of the grating is identical to that specified in table (9.4.1). These results were computed using 4 waveguide modes in each array and 10 Rayleigh orders for a normalized wavelength of 1.2001.

m	x-symmetry	F_m	$\hat{F}_m$
1	E	0.46346	0.46346
2	0	0.00001	0.00001
3	E	0.00004	0.00004
4	0	0.00000	0.00000

Note also the concentration of the energy flux in the lowest order mode. Since this is the only propagating mode (i.e. μ_1 is real), it follows from the substantial depth of the structure (h/d = 0.4) that the evanescent coupling of the upper and lower edges of either array must be particularly weak - a phenomenon which manifests itself as very low energy fluxes carried by the higher order modes. It should also be noted that these results are analytically preserved within the formalism for a "field symmetric" selection of orders as mentioned in table (9.4.1).

(9.4.4.18) is sufficient to ensure the detailed balance of modal fluxes for a lossless totally symmetric grating operated in a first order Littrow configuration. These results have been confirmed by numerical experiments (see table (9.4.4)).

9.4.4.2 *An asymmetric arrangement of the upper and lower arrays of a double grating*

Prior to this section the discussion has involved only those structures having some geometric symmetry in either one form or another. Thus it is the purpose of this section to discuss an asymmetric arrangement of the upper and lower arrays (both assumed to be identical). In this context, the term "asymmetry" is perhaps a misnomer since there is a distinct regularity in the arrangement of figure (9.4.6). Once again, attention is restricted purely to the long wavelength case for which only the undispersed orders are non-evanescent. Now by invoking the principle of time reversal, incident waves $\bar{R}_0$ and $\bar{T}_0$ are returned as shown in the above configuration. If the point $\{\frac{1}{2}(c+\delta x), 0\}$ is taken to be the phase origin, it becomes evident that the wave returned along the channel $\bar{T}_0$ "sees" a structure almost identical to the original incident wave I. (In fact, it is exactly the same except for the transformation $x \rightarrow -x$). However, from an inspection of the accompanying diagram, it is clear that the same remarks do not apply to a wave returned along the path $\bar{R}_0$.

Only for normally incident radiation can both $\bar{R}_0$ and $\bar{T}_0$ be incident on a structure which is essentially the same for both waves. In this case, the grating, as viewed from either side, is invariant under the parity change $x \rightarrow -x$. Thus, with these introductory comments, the

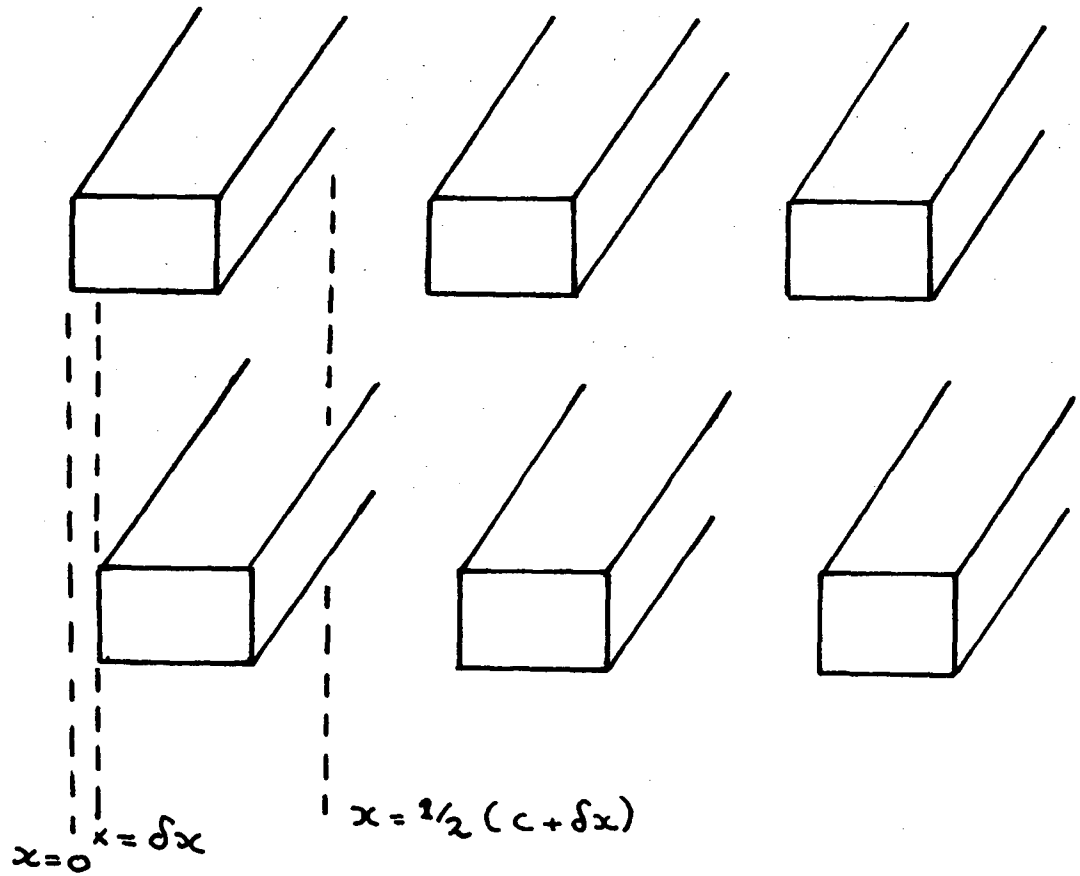


Figure 9.4.6 An asymmetric arrangement of an identical pair of lamellar wire gratings.

general conservation relations for this system are:

$$\bar{E}_u(x,y) = \bar{R}_0 E_u(x,y) + \bar{T}_0 E_\ell(-x,-y) \qquad (9.4.4.31)$$

and $$\bar{E}_\ell(x,y) = \bar{T}_0 E_u(-x,-y) + \bar{R}_0 E_\ell(x,y). \qquad (9.4.4.32)$$

By now substituting the plane wave representations for the fields E_u and E_ℓ into these equations it may be shown that:

(a) for the real order ($p = 0$)

(i) $$|R_0|^2 + |T_0|^2 = 1 \quad \text{(from (9.4.4.31))} \qquad (9.4.4.33)$$

and (ii) $$\mathrm{Re}(R_0 \bar{T}_0) = 0 \quad \text{(from (9.4.4.32))} \qquad (9.4.4.34)$$

i.e. $\arg(R_0) - \arg(T_0) = \pi/2 + (\pi)$.

and

(b) for the evanescent orders ($p \neq 0$)

(i) $$\bar{R}_p = \bar{R}_0 R_{-p} + \bar{T}_0 T_p \quad \text{(from (9.4.4.31))} \qquad (9.4.4.35)$$

and (ii) $$\bar{T}_p = \bar{T}_0 R_p + \bar{R}_0 T_{-p} \quad \text{(from (9.4.4.32))}. \qquad (9.4.4.36)$$

It should be noted that for an array phase shift of $\delta x = 0$, the above expressions revert to their earlier forms derived in section (9.4.4.1.1).

Using the definitions

$$A_p = R_p + T_p$$

and $$B_p = R_p - T_p$$

equations (9.4.4.35-36) may be summarized in the form

Table 9.4.5

Demonstration of the "detailed balance" phenomenon for an asymmetric double grating composed of an identical pair of lamellar gratings specified by the parameters:

$$c/d = 0.8 \qquad \text{and} \qquad h/d = 0.4.$$

The arrays were separated by an amount of $S/d = 0.6$ and displaced relative to one another by $\delta x/d = 0.2$. The grating was irradiated at normal incidence with light having a normalized wavelength of $\lambda/d = 1.8001$. The following calculations were performed with 4 waveguide modes and 11 Rayleigh orders.

m	x-symmetry	F_m	$\hat{F}_m$
1	E	.74810	.74810
2	0	.00000	.00000
3	E	.00093	.00093
4	0	.00000	.00000

$$\left.\begin{aligned} \arg(A_p + A_{-p}) &= \tfrac{1}{2}\arg(A_0) + (\pi) \\ \arg(B_p + B_{-p}) &= \tfrac{1}{2}\arg(B_0) + (\pi). \end{aligned}\right\} \quad \forall\ p \neq 0$$

These properties have also been confirmed by our computer program, evidence of which is given in table (9.4.5). In the case of the waveguide fields in the grooves of both arrays, a treatment analogous to that described earlier yields a set of conservation relations for the modal field coefficients. Since the necessary mathematics is particularly involved, it is felt that its presentation would only detract from line of arguments developed herein and for this reason its inclusion is not warranted. Nevertheless the analysis has once again shown the existence of the detailed balance of modal fluxes in both arrays. This has been exemplified in output obtained from our computer program, a sample result of which is shown in table (9.4.5).

9.4.5 CONCLUDING REMARKS

The discussion in this section has centred on the analytic constraints imposed on the field quantities by grating symmetry. In particular, the concept of time reversibility has shown itself to be a very powerful tool indeed and it is hoped that the material presented above has sufficiently emphasised its relevance and importance in this branch of electromagnetism. It is only with hindsight that one comes to appreciate its immense significance. In particular, it is interesting to ponder upon the fact that reciprocity is a "subset" of time reversal for lossless media but that reciprocity still holds for lossy gratings for which time reversibility cannot be applied.

Although the search for such properties is only of a recent origin in grating theory, such techniques are by now means novel and have long been used in the spheres of nuclear and particle scattering theory. As such, it only points out the relatively unsophisticated state of development presently obtained by this facet of grating theory.

The material contained in this section is of an original nature but it could not even have been attempted without the valued suggestions from those having a far broader appreciation of the concepts of modern physics. Thus, it is with great pleasure that the author should like to thank Dr. Graham Derrick and Dr. Ross McPhedran of the Department of Theoretical Physics, University of Sydney for "opening his eyes" to this fascinating study and also for their many illuminating discussions held throughout the course of these investigations.

It is only with the rapid progress made during the past decade concerning the modelling of electromagnetic scattering processes that theoretical attention has now been turned to the detailed understanding of the actual nature of diffraction gratings. The material contained in this section is part of these current efforts and although having no obvious relevance to the experimenter or grating designer, it does enhance the understanding of the theoretician and as such promotes a far more balanced outlook on the whole subject.

REFERENCES

[9.1] J. L. Adams, (1977), University of Tasmania, Report DGRG 77/7.

[9.2] E. A. Lewis and J. P. Casey, J. Opt. Soc. Amer., 41 (1951) 360.

[9.3] J. P. Casey and E. A. Lewis, J. Opt. Soc. Amer., 42 (1952) 971.

[9.4] H. Blok and G. Mur, Appl. Sci. Res., 26 (1972) 389.

[9.5] R. Petit and M. Cadilhac, C. R. Acad. Sci. Paris, 262 (1966) 468.

[9.6] M. Nevière and M. Cadilhac, Optics Commun., 2 (1970) 235.

[9.7] A. Wirgin, C. R. Acad. Sci. Paris, 262B (1966) 870.

[9.8] R. C. McPhedran and D. Maystre, Appl. Phys., 14 (1977) 1.

[9.9] J. M. Blatt and V. F. Weisskopf, (1952), "Theoretical Nuclear Physics", (Wiley, New York), Chapter 10.

[9.10] P. Grivet, (1967), in "Proceedings of the Symposium on Modern Optics", Polytechnic Institute of Brooklyn, (Polytechnic Press, New York).

CHAPTER 10

FURTHER THEORETICAL STUDIES OF INDUCTIVE GRIDS

The material presented in this chapter has evolved from a study conducted jointly by the author and Dr. Ross McPhedran (University of Sydney). It represents a logical continuation of his earlier work [10.1, 10.2] and as such I wish to thank him for involving me in these further investigations. His stimulating collaboration is gratefully acknowledged. The results of our studies are incorporated in a paper which has been submitted for publication to the journal, "Applied Physics".

10.1 INTRODUCTION

The contents of this chapter, based largely on the material presented in an internal research report [10.3], represents an extension of studies exploring the properties of doubly periodic diffracting structures, with particular emphasis being placed on their potential use in solar selective systems. The first of these investigations [10.1, 10.2] was made into the properties of inductive grids constructed by perforating rectangular apertures in a thick planar sheet of metal, the metal being taken to be perfectly conducting. Such grids were shown to be capable of providing high absorptance to emittance (a/e) ratios provided they were used in direct illumination and tracked the source. Further studies [10.4, 10.5] involving a structure

referred to as the crossed lamellar grating revealed similar solar selective properties.

The bulk of the material presented in this chapter concerns the properties of grids having circular apertures. Chen [10.6] has detailed a theory describing the diffraction of a plane wave by an infinitesimally thin perfectly conducting screen perforated with such apertures. In this paper he also included a discussion of the effects of placing thin dielectric films on either side of the conducting screen. In a further article [10.7] he outlined the way in which the theory could be extended to accommodate a finite thickness of metal.

The theoretical treatment presented in this chapter is based on that of Chen [10.6] but has been extended to deal with thick grids with apertures filled with a dielectric, the grid being "sandwiched" between symmetric films of another dielectric material. Evidence of the validity of the formalism is obtained using the Reciprocity Theorem [10.1, 10.2] and a new Littrow phase constraint [10.4], and naturally by comparison with the published results of Chen [10.6, 10.7].

The formulation has been used to analyse the solar selective properties of this class of grids. It has shown that a closely packed arrangement of apertures gives optimal transmission properties to the grid, but even so the low hole-area fractions attainable limit the performance to values lower than those obtainable with square apertures. The use of appropriate thin films and plugs can be used to produce improvements in transmittance values of up to ten percent.

Also contained in this chapter is a theory describing the diffraction of a plane wave by a rectangular holed inductive grid plugged

with a dielectric material and surrounded by a pair of lossless symmetric thin films. The theoretical treatment is based on a further paper by Chen [10.9] but has been extended in the same manner as mentioned above. This theory augments the earlier study of McPhedran and Maystre [10.1, 10.2] and its inclusion in this thesis was largely motivated by the enhanced performance referred to in the previous paragraph. Although the numerical simulations for this grating failed to yield any improvement in the selective performance of this structure, the theory and the results obtained are still incorporated herein so as to create a balanced record of the study.

10.2 THE THEORETICAL FORMALISM

10.2.1 Notation

Consider a plane wave incident upon a doubly-periodic structure consisting of circular apertures pierced in a plane, perfectly conducting screen. In accordance with the conventional choice of axes for classical diffraction grating problems, the Oy axis of a rectangular Cartesian coordinate system is taken to be orthogonal to the screen, with $y = h/2$ and $y = -h/2$ corresponding to the top and bottom surfaces of the screen. Located symmetrically on either side of the grid are a pair of films each of thickness $s/2$ and of refractive index r_1. The apertures, assumed to be circular, are filled with a further dielectric material of refractive index r_2 and are spaced periodically along the Ox axis with period d. The other axis of periodicity makes an angle η with the Ox axis. The projection of the aperture spacing along the second periodicity axis onto Oz is denoted by d'. A diagram of the grid

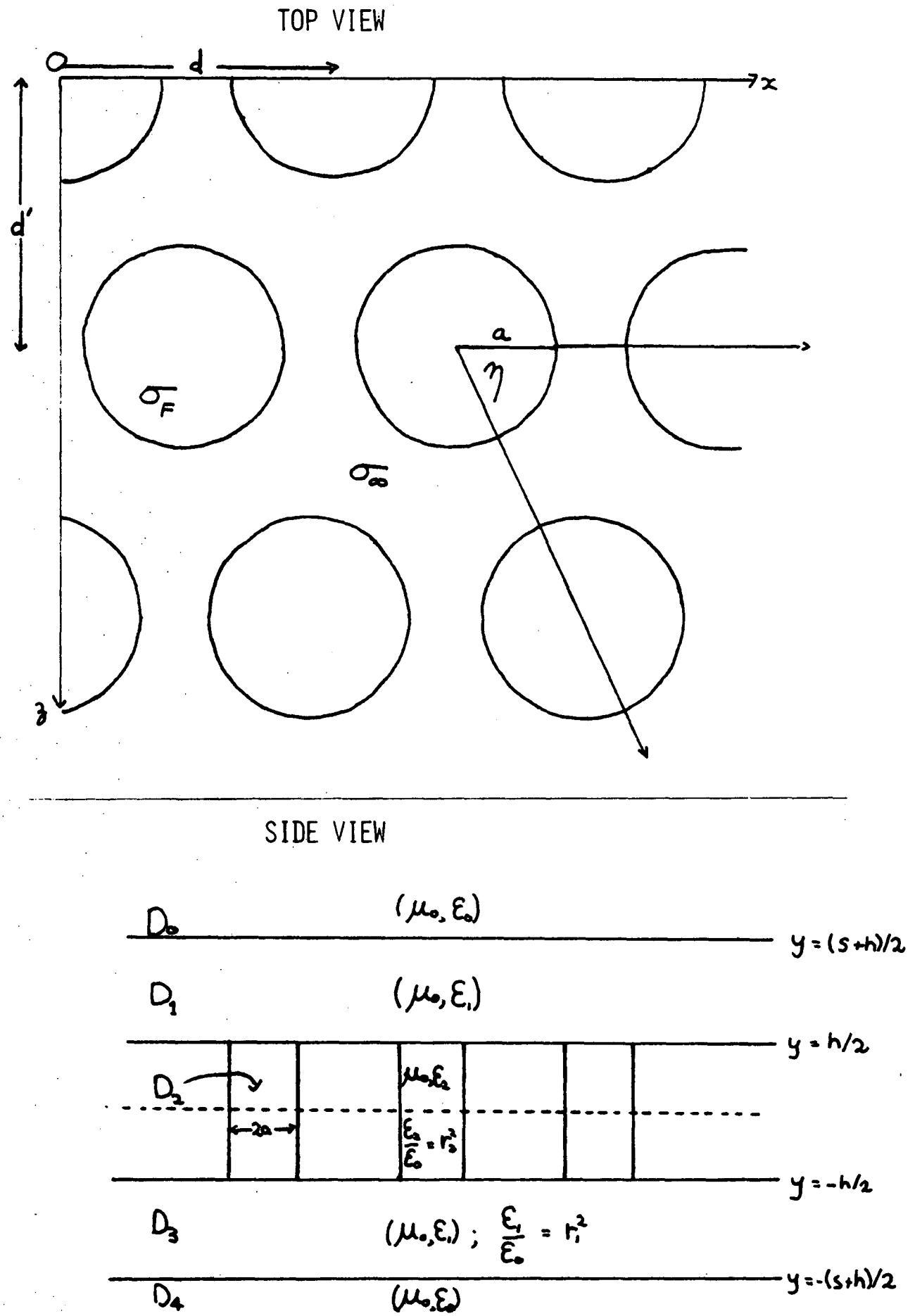


Figure 10.1 The geometry of the round-holed inductive grid.

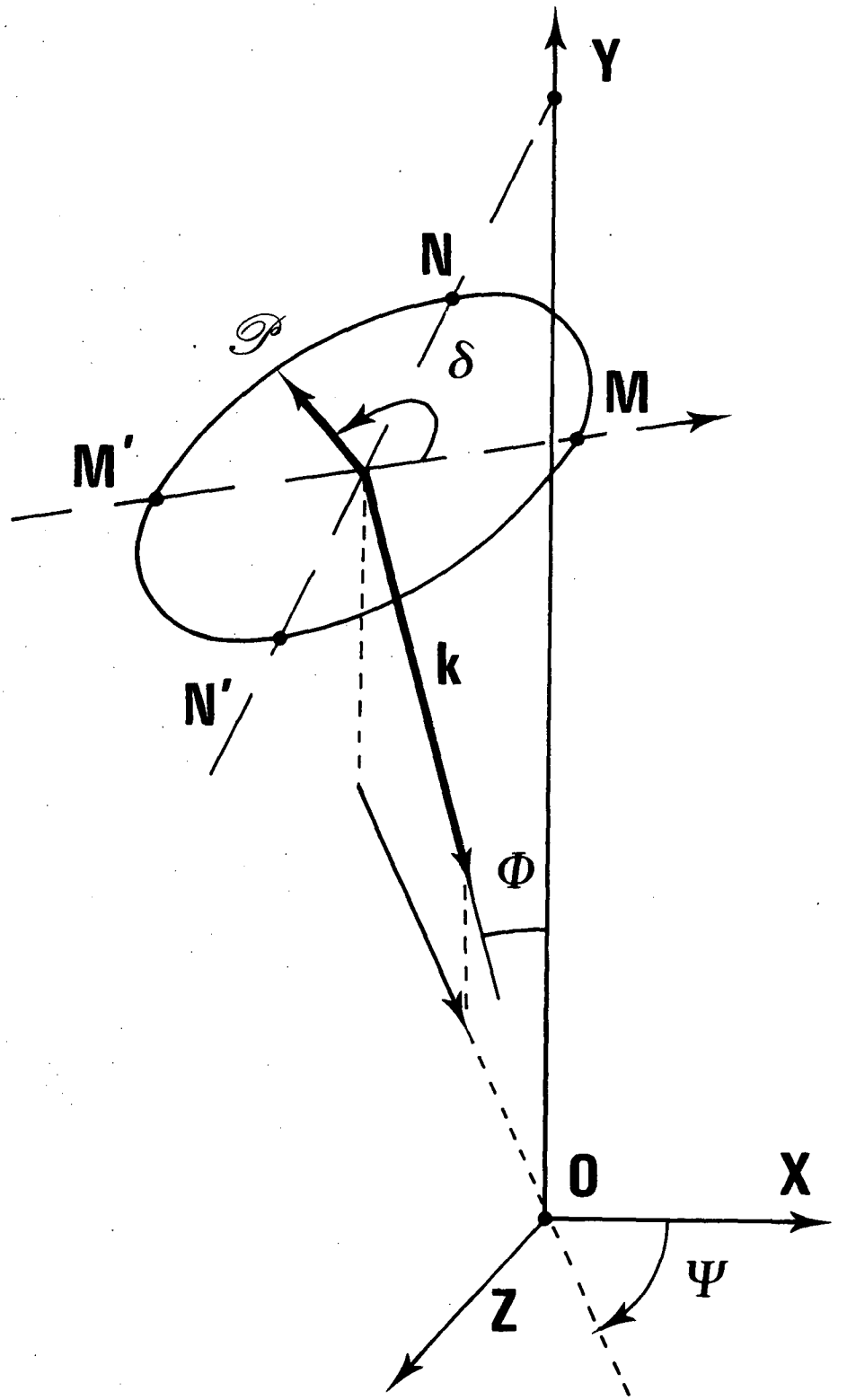


Figure 10.2 Definition of the incidence parameters.

is given in figure (10.1).

The direction of the incoming plane wave is specified by two angles ϕ and ψ (figure 10.2). ϕ is the angle between the incident wave vector $\underline{k}_0$ and the Oy axis. ψ is the angle between the projection of $\underline{k}_0$ onto the Oxz plane and the z axis. Thus, if $k_0 = 2\pi/\lambda$ is the wave number of the incident beam, then

$$\underline{k}_0 = k_0(\alpha,-\beta,\gamma), \tag{10.2.1}$$

where

$$\begin{aligned} \alpha &= \sin\phi\cos\psi, \\ \beta &= \cos\phi, \\ \gamma &= \sin\phi\sin\psi. \end{aligned} \tag{10.2.2}$$

In order to define the polarization of the incident beam, a third angle δ is introduced, such that when $\delta = 90^o$ the vector amplitude of the incident electric field ($\underline{A}$) lies in the plane of incidence. When $\delta = 0^o$, $\underline{A}$ is orthogonal to the plane of incidence. If the magnitude of $\underline{A}$ is normalized to be unity, then

$$\begin{aligned} A_x &= \sin\psi\cos\delta + \sin\delta\cos\phi\cos\psi, \\ A_y &= \sin\delta\sin\phi, \\ A_z &= \cos\phi\sin\psi\sin\delta - \cos\delta\cos\psi. \end{aligned} \tag{10.2.3}$$

Suppressing the time varying term $\exp(-j\omega t)$, the expression for the electric field vector of the incident wave becomes

$$\underline{E}^i = \underline{A} \exp [jk_0(\alpha x - \beta y + \gamma z)]. \tag{10.2.4}$$

If a constant multiplying factor of $1/(j\omega\mu_0)$ is also suppressed, then the magnetic field vector of the incident wave is given by

$$\underline{H}^i = \underline{B} \exp[jk_0(\alpha x - \beta y + \gamma z)], \tag{10.2.5}$$

where

$$\begin{aligned} B_x &= jk_0(\cos\delta \cos\phi \cos\psi - \sin\delta \sin\psi), \\ B_y &= jk_0(\cos\delta \sin\phi), \\ B_z &= jk_0(\cos\delta \cos\phi \sin\psi + \sin\delta \cos\psi). \end{aligned} \tag{10.2.6}$$

10.2.2 The plane wave expansions for the diffracted and transmitted fields

In response to the incident field, the grid gives rise to diffracted fields both above and below it. In the following analysis, quantities associated with the transmitted field (in $y \leq -h/2$) are distinguished from those associated with the reflected field (in $y \geq h/2$) by attaching a superscripted circumflex to the former.

Both the reflected and transmitted fields can be written as superpositions of plane waves. Since the grid is doubly periodic, these plane waves must propagate in discrete directions.

As detailed in reference [10.1], the reflected field in free space must be made up of Cartesian components, each of which is a superposition of plane waves of the form

$$R_{pq}(x,y,z) = \exp[jk_0(\alpha_p x + \beta_{pq} y + \gamma_{pq} z)] \tag{10.2.7}$$

where

$$\alpha_p = \alpha + \frac{pK}{k_0}, \quad (K = 2\pi/d) \tag{10.2.8}$$

$$\gamma_{pq} = \gamma + \frac{qK'}{k_0} - \frac{pK}{k_0 \tan \eta} \quad (K' = 2\pi/d') \tag{10.2.9}$$

and

$$\begin{aligned} \beta_{pq} &= \sqrt{1 - \alpha_p^2 - \gamma_{pq}^2} \quad \text{if} \quad \alpha_p^2 + \gamma_{pq}^2 \leq 1 \\ &= j\sqrt{\alpha_p^2 + \lambda_{pq}^2 - 1} \quad \text{otherwise.} \end{aligned} \tag{10.2.10}$$

For the transmitted field in free space, functions for the form

$$\hat{R}_{pq}(x,y,z) = \exp[jk_0(\alpha_p x - \beta_{pq} y + \gamma_{pq} z)] \tag{10.2.11}$$

are superposed.

Within each element of the sandwich layer the fields are again represented by series of upward-going and downward-going plane waves, this time taking the respective forms

$$R^+_{pq}(x,y,z) = \exp[jk_1(\alpha'_p x + \beta'_{pq} y + \gamma'_{pq} z)] \tag{10.2.12}$$

and

$$R^-_{pq}(x,y,z) = \exp[jk_1(\alpha'_p x - \beta'_{pq} y + \gamma'_{pq} z)] \tag{10.2.13}$$

where

$$k_1 = k_0 r_1 \tag{10.2.14}$$

$$\alpha'_p = \frac{k_0}{k_1} \alpha_p$$

$$\gamma'_{pq} = \frac{k_0}{k_1} \gamma_{pq} \tag{10.2.15}$$

and

$$\beta'_{pq} = \sqrt{1 - \alpha'^2_p - \gamma'^2_{pq}} \quad . \tag{10.2.16}$$

10.2.3 TE and TM modal expansions for the plane wave fields

In problems having dielectric layers placed on either side of the grid, it is advantageous to treat the fields in all regions as sums of orthonormal TE and TM vector modes. In free space, these are respectively

$$\underline{RTE}_{pq}(x,y,z) = \frac{1}{\xi_{pq}\sqrt{dd'}} \{\gamma_{pq} \underline{\hat{x}} - \alpha_p \underline{\hat{z}}\} \; R_{pq}(x,y,z) \tag{10.2.17}$$

and

$$\underline{RTM}_{pq}(x,y,z) = \frac{1}{\xi_{pq}\sqrt{dd'}} \{\alpha_p \underline{\hat{x}} + \gamma_{pq} \underline{\hat{z}}\} \; R_{pq}(x,y,z) \tag{10.2.18}$$

where

$$\xi_{pq} = \sqrt{\alpha_p^2 + \gamma_{pq}^2} \quad . \tag{10.2.19}$$

The transverse modal functions for the transmitted field in free space, $\underline{\hat{RTE}}_{pq}$ and $\underline{\hat{RTM}}_{pq}$ are defined in an analogous fashion.

In the upper and lower dielectric films, regions D_1 and D_3 respectively, the upward-going modal functions are given by

$$\underline{\underline{RTE}}^{+}_{pq}(x,y,z) = \frac{1}{\xi'_{pq}\sqrt{dd'}} \{\gamma'_{pq}\hat{\underline{x}} - \alpha'_{p}\hat{\underline{z}}\} \; R^{+}_{pq}(x,y,z) \tag{10.2.20}$$

$$\underline{\underline{RTM}}^{+}_{pq}(x,y,z) = \frac{1}{\xi'_{pq}\sqrt{dd'}} \{\alpha'_{p}\hat{\underline{x}} + \gamma'_{pq}\hat{\underline{z}}\} \; R^{-}_{pq}(x,y,z). \tag{10.2.21}$$

The downward-going modes RTE^{-}_{pq} and RTM^{-}_{pq} are given by similar expressions.

In region D_0, the transverse resolute of the incident electric field is expressed by

$$\underline{E}^{i}_{t} = E_i \; \underline{\underline{R\hat{T}E}}_{00} + F_i \; \underline{\underline{R\hat{T}M}}_{00} \tag{10.2.22}$$

and for the transverse resolute of the incident magnetic field,

$$\hat{\underline{y}} \times \underline{H}^{i}_{t} = \frac{1}{Z_0} \{\beta E_i \; \underline{\underline{R\hat{T}E}}_{00} + \frac{1}{\beta} F_i \; \underline{\underline{R\hat{T}M}}_{00}\}. \tag{10.2.23}$$

Similarly, the transverse resolutes of the reflected fields in D_0 can be expressed by

$$\underline{E}_{t} = \sum_{p,q} [E_{pq} \; \underline{\underline{RTE}}_{pq} + F_{pq} \; \underline{\underline{RTM}}_{pq}] \tag{10.2.24}$$

and

$$\hat{\underline{y}} \times \underline{H}_{t} = -\frac{1}{Z_0} \sum_{p,q} [\beta_{pq} \; E_{pq} \; \underline{\underline{RTE}}_{pq} + \frac{F_{pq}}{\beta_{pq}} \; \underline{\underline{RTM}}_{pq}]. \tag{10.2.25}$$

The transverse resolutes of the transmitted fields in D_4 are given by

$$\underline{\hat{E}}_t = \sum_{p,q} [\hat{E}_{pq} \underline{\hat{RTE}}_{pq} + \hat{F}_{pq} \underline{\hat{RTM}}_{pq}], \tag{10.2.26}$$

and $$\hat{y} \times \underline{\hat{H}}_t = \frac{1}{Z_0} \sum_{p,q} [\beta_{pq} \hat{E}_{pq} \underline{\hat{RTE}}_{pq} + \frac{\hat{F}_{pq}}{\beta_{pq}} \underline{\hat{RTM}}_{pq}]. \tag{10.2.27}$$

In D_1, corresponding field quantities are defined by:

$$\underline{E}_t^+ = \sum_{p,q} [E_{pq}^+ \underline{RTE}_{pq}^+ + F_{pq}^+ \underline{RTM}_{pq}^+], \tag{10.2.28}$$

$$\hat{y} \times \underline{H}_t^+ = -\frac{1}{Z_1} \sum_{p,q} [\beta'_{pq} E_{pq}^+ \underline{RTE}_{pq}^+ + \frac{F_{pq}^+}{\beta'_{pq}} \underline{RTM}_{pq}^+] \tag{10.2.29}$$

$$\underline{E}_t^- = \sum_{p,q} [E_{pq}^- \underline{RTE}_{pq}^- + F_{pq}^- \underline{RTM}_{pq}^-], \tag{10.2.30}$$

$$\underline{y} \times \underline{H}_t^- = \frac{1}{Z_1} \sum_{p,q} [\beta'_{pq} E_{pq}^- \underline{RTE}_{pq}^- + \frac{F_{pq}^-}{\beta'_{pq}} \underline{RTM}_{pq}^-]. \tag{10.2.31}$$

Similarly, analogous definitions appropriate to D_3 are made by attaching a superscripted circumflex to the field quantities.

In the above, Z_0 denotes the impedance of free space, $Z_1 = Z_0/r_1$, and the summation indices, p and q, run from $-\infty$ to $+\infty$.

10.2.4 Admittance of the dielectric films

It is now necessary to apply boundary conditions at the interface $y = y_0$ between D_0 and D_1 in order to relate the field quantities at y_0 to those at $y = y_1$ {where $y_0 = [s + h]/2$, $y_1 = h/2$}. Define:

$$E_i(y_0) = E_i \exp(-jk_0 \beta y_0), \tag{10.2.32}$$

$$E_{pq}(y_0) = E_{pq} \exp(jk_0 \beta_{pq} y_0), \tag{10.2.33}$$

$$E^{+}_{pq}(y_0) = E^{+}_{pq} \exp(jk_1 \beta'_{pq} y_0), \tag{10.2.34}$$

$$E^{-}_{pq}(y_0) = E^{-}_{pq} \exp(- jk_1 \beta'_{pq} y_0). \tag{10.2.35}$$

Continuity of $\underline{E}_t$ and $(\hat{\underline{y}} \times \underline{H}_t)$ at $y = y_0$ ensure that, for TE modes,

$$E_i(y_0)\, \delta_{p,0}\, \delta_{q,0} + E_{pq}(y_0) = E^{+}_{pq}(y_0) + E^{-}_{pq}(y_0), \tag{10.2.36}$$

and

$$\sigma^{TE}_{pq} [- E_i(y_0)\, \delta_{p,0}\, \delta_{q,0} + E_{pq}(y_0)] = \tau^{TE}_{pq}[E^{+}_{pq}(y_0) - E^{-}_{pq}(y_0)], \tag{10.2.37}$$

where $\sigma^{TE}_{pq} = \beta_{pq}$, $\tau^{TE}_{pq} = r_1 \beta'_{pq}$, $r^{TE}_{pq} = \sigma^{TE}_{pq} / \tau^{TE}_{pq}$. (10.2.38)

Solving (10.2.36) and (10.2.37), it is found that:

$$E^{+}_{pq}(y_1) + E^{-}_{pq}(y_1) = \{\cos (k_1 \beta'_{pq} \tfrac{s}{2}) + j r^{TE}_{pq} \sin(k_1 \beta'_{pq} \tfrac{s}{2})\} E_i(y_0)\, \delta_{p,0}\, \delta_{q,0}$$

$$+ \{\cos (k_1 \beta'_{pq} \tfrac{s}{2}) - j r^{TE}_{pq} \sin (k_1 \beta'_{pq} \tfrac{s}{2})\} E_{pq}(y_0), \tag{10.2.39}$$

and

$$\tau^{TE}_{pq} \{E^{+}_{pq}(y_1) - E^{-}_{pq}(y_1)\} = - \tau^{TE}_{pq} \{r^{TE}_{pq} \cos(k_1 \beta'_{pq} \tfrac{s}{2}) + j \sin(k_1 \beta'_{pq} \tfrac{s}{2})\} E_i(y_0) \times$$

$$\times \delta_{p,0}\, \delta_{q,0} + \tau^{TE}_{pq} \{r^{TE}_{pq} \cos(k_1 \beta'_{pq} \tfrac{s}{2}) - j \sin(k_1 \beta'_{pq} \tfrac{s}{2})\} E_{pq}(y_0)$$

(10.2.40)

An effective incident field for the problem is defined by:

$$E_i^* = \{\cos(k_1 \beta'_{00} \tfrac{s}{2}) + j\, r_{00}^{TE} \sin(k_1 \beta'_{00} \tfrac{s}{2})\}\, E_i(y_0). \qquad (10.2.41)$$

Also, an effective reflected-field coefficient is given by

$$E_{pq}^* = \{\cos(k_1 \beta'_{pq} \tfrac{s}{2}) - j\, r_{pq}^{TE} \sin(k_1 \beta'_{pq} \tfrac{s}{2})\}\, E_{pq}(y_0). \qquad (10.2.42)$$

Furthermore, the TE admittances for the incident and reflected fields:

$$Y_i^{TE} = \frac{\tau_{00}^{TE} \{r_{00}^{TE} \cos(k_1 \beta'_{00} \tfrac{s}{2}) + j \sin(k_1 \beta'_{00} \tfrac{s}{2})\}}{\{\cos(k_1 \beta'_{00} \tfrac{s}{2}) + j\, r_{00}^{TE} \sin(k_1 \beta'_{00} \tfrac{s}{2})\}} \qquad (10.2.43)$$

$$Y_{pq}^{TE} = \frac{\tau_{pq}^{TE} \{r_{pq}^{TE} \cos(k_1 \beta'_{pq} \tfrac{s}{2}) - j \sin(k_1 \beta'_{pq} \tfrac{s}{2})\}}{\{\cos(k_1 \beta'_{pq} \tfrac{s}{2}) - j\, r_{pq}^{TE} \sin(k_1 \beta'_{pq} \tfrac{s}{2})\}} . \qquad (10.2.44)$$

are introduced. Thus, at $y = y_1$, the TE resolutes of the transverse electric and magnetic fields are:

$$\underline{E}_t^{TE} = \sum_{p,q} [E_{pq}^* + E_i^* \delta_{p,0} \delta_{q,0}]\, \underline{RTE}_{pq}(y = 0), \qquad (10.2.45)$$

$$-Z_0\, \hat{\underline{y}} \times \underline{H}_t^{TE} = \sum_{p,q} [Y_{pq}^{TE} E_{pq}^* - Y_i^{TE} E_i^* \delta_{p,0} \delta_{q,0}]\, \underline{RTE}_{pq}(y = 0). \qquad (10.2.46)$$

The TM treatment follows along analogous lines. By defining

$$\sigma_{pq}^{TM} = \frac{1}{\beta_{pq}} \text{ and } \tau_{pq}^{TM} = \frac{r_1}{\beta'_{pq}} , \qquad (10.2.47)$$

equations (10.2.43) and (10.2.44) take exactly the same form for the TM

admittances, and expressions (10.2.45-46) become

$$\underline{E}_t^{TM} = \sum_{p,q} \{F^*_{pq} + F^*_i \delta_{p,0} \delta_{q,0}\} \underline{RTM}_{pq}(y = 0) \qquad (10.2.48)$$

$$-Z_0 \hat{\underline{y}} \times \underline{H}_t^{TM} = \sum_{p,q} \{Y^{TM}_{pq} F^*_{pq} - Y^{TM}_i F^*_i \delta_{p,0} \delta_{q,0}\} \underline{RTM}_{pq}(y = 0). \qquad (10.2.49)$$

For the case of the transmitted field quantities, it may be shown that at $y = y_2 = -y_1$, the transverse electric and magnetic fields are

$$\underline{E}_t = \sum_{p,q} [\hat{E}^*_{pq} \underline{RTE}_{pq}(y = 0) + \hat{F}^*_{pq} \underline{RTM}_{pq}(y = 0)] \qquad (10.2.50)$$

and

$$Z_0 \hat{\underline{y}} \times \underline{H}_t = \sum_{p,q} [Y^{TE}_{pq} \hat{E}^*_{pq} \underline{RTE}_{pq}(y = 0) + Y^{TM}_{pq} \hat{F}^*_{pq} \underline{RTM}_{pq}(y = 0)]. \qquad (10.2.51)$$

Here it has been necessary to invoke the symmetry of the sandwich-structure in order to link output and input admittances.

10.2.5 Fields within the circular apertures

The following discussion of the modal expansions appropriate to the region D_2 (cylindrical apertures) assumes the reader to be familiar with the contents of section (2.3) of the Waveguide Handbook [10.10]. The increased complexity of the problem for grids having circular apertures compared with the problem of grids with rectangular apertures arises in part from the degeneracy of the expressions for the waveguide modes. For this reason, together with the fact that the discussion by

Chen [10.6] was found to be misleading due to its brevity, it is clear that the modes need to be discussed in some detail.

In prescribing the fields within D_2, it is necessary to introduce a cylindrical co-ordinate system. The polar quantities, r and θ, are defined in figure (10.3). From Marcuvitz [10.10] the "horizontally" oriented TE and TM modes are taken to be

$$\underline{MTEH}_{nm}(r,\theta) = g_{nm} \{\frac{na}{\chi'_{nm} r} J_n(\frac{\chi'_{nm} r}{a}) \cos(n\theta)\, \hat{\underline{r}} - J'_n(\frac{\chi'_{nm} r}{a}) \sin(n\theta)\, \hat{\underline{\theta}}\} \quad (10.2.52)$$

and

$$\underline{MTMH}_{nm}(r,\theta) = h_{nm} \{J'_n(\frac{\chi_{nm} r}{a}) \cos(n\theta)\, \hat{\underline{r}} - \frac{na}{\chi_{nm} r} J_n(\frac{\chi_{nm} r}{a}) \sin(n\theta)\, \hat{\underline{\theta}}\} \quad (10.2.53)$$

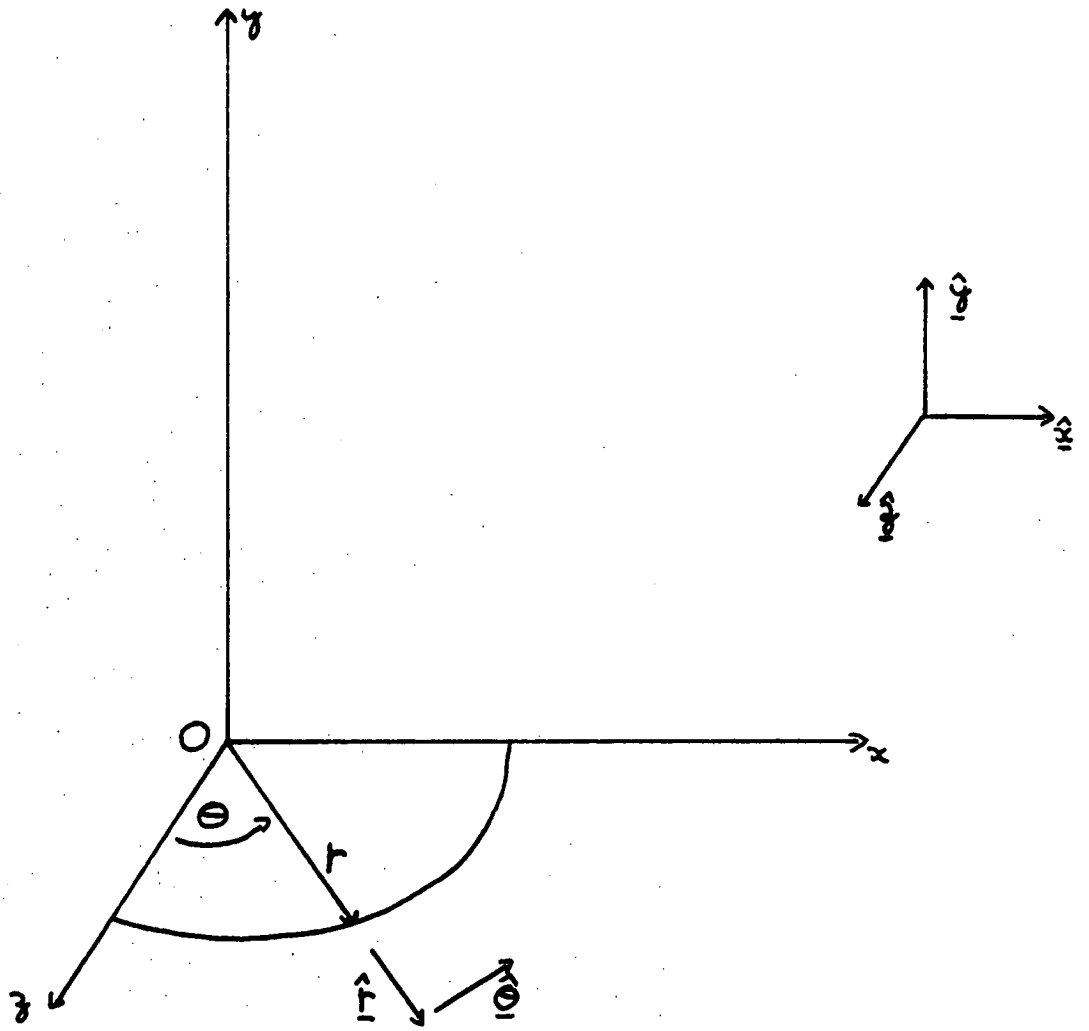
Here χ_{nm} and χ'_{nm} denote the m^{th} non-trivial zeros of the respective functions J_n and J'_n. g_{nm} and h_{nm} are normalization factors which provide orthonomality of the modes:

$$\text{viz.} \quad g_{nm} = \sqrt{\frac{\varepsilon_n}{\pi}} \frac{\chi'_{nm}}{a\, J_n(\chi'_{nm}) \sqrt{\chi'^2_{nm} - n^2}} \quad (10.2.54)$$

and

$$h_{nm} = \sqrt{\frac{\varepsilon_n}{\pi}} \frac{1}{a\, J_{n-1}(\chi_{nm})} \,. \quad (10.2.55)$$

Similarly, the "vertically" oriented modes are given by the following expressions.



$$\hat{r} \times \hat{\theta} = \hat{y}$$

$$\hat{r} \cdot \hat{x} = \sin\theta \qquad \hat{\theta} \cdot \hat{x} = \cos\theta$$

$$\hat{r} \cdot \hat{z} = \cos\theta \qquad \hat{\theta} \cdot \hat{z} = -\sin\theta$$

Figure 10.3 Definition of the polar co-ordinates.

$$\underline{MTEV}_{nm}(r,\theta) = g_{nm} \left\{\frac{na}{\chi'_{nm} r} J_n\left(\frac{\chi'_{nm} r}{a}\right) \sin n\theta \, \hat{\underline{r}} + J'_n\left(\frac{\chi'_{nm} r}{a}\right) \cos n\theta \, \hat{\underline{\theta}}\right\} \quad (10.2.56)$$

and

$$\underline{MTMV}_{nm}(r,\theta) = h_{nm} \left\{J'_n\left(\frac{\chi_{nm} r}{a}\right) \sin n\theta \, \hat{\underline{r}} + \frac{na}{\chi_{nm} r} J_n\left(\frac{\chi_{nm} r}{a}\right) \cos n\theta \, \hat{\underline{\theta}}\right\}. \quad (10.2.57)$$

With the above modal quantities now defined the most general expressions for the transverse resolutes of the fields are

$$\underline{E}_t = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \sum_{\ell=1}^{2} \left[\left\{a_{nm\ell} \sin(k_2 \nu'_{nm} y) + b_{nm\ell} \cos(k_2 \nu'_{nm} y)\right\} \underline{MTE}_{nm\ell}(r,\theta) + \left\{c_{nm\ell} \sin(k_2 \nu_{nm} y) + d_{nm\ell} \cos(k_2 \nu_{nm} y)\right\} \underline{MTM}_{nm\ell}(r,\theta)\right] \quad (10.2.58)$$

$$Z_0 \hat{\underline{y}} \times \underline{H}_t = j r_2 \sum_{n,m,\ell} \left[\nu'_{nm} \left\{a_{nm\ell} \cos(k_2 \nu'_{nm} y) - b_{nm\ell} \sin(k_2 \nu'_{nm} y)\right\} \underline{MTE}_{nm\ell}(r,\theta) + \frac{1}{\nu_{nm}} \left\{c_{nm\ell} \cos(k_2 \nu_{nm} y) - d_{nm\ell} \sin(k_2 \nu_{nm} y)\right\} \underline{MTM}_{nm\ell}(r,\theta)\right] \quad (10.2.59)$$

where $$\nu'_{nm} = \sqrt{1 - \frac{\chi'^2_{nm}}{k_2^2 a^2}} \quad (10.2.60)$$

$$\nu_{nm} = \sqrt{1 - \frac{\chi^2_{nm}}{k_2^2 a^2}} \quad (10.2.61)$$

and $$k_2 = k_0 r_2. \quad (10.2.62)$$

In the above summations (for reasons of brevity), a variable ℓ which runs over horizontal and vertical modes has been included. Its significance is shown

in the following equations:

$$\underline{MTEV}_{nm} = \underline{MTE}_{nm1} \tag{10.2.63}$$

$$\underline{MTEH}_{nm} = \underline{MTE}_{nm2} \tag{10.2.64}$$

$$\underline{MTMH}_{nm} = \underline{MTM}_{nm1} \tag{10.2.65}$$

$$\underline{MTMV}_{nm} = \underline{MTM}_{nm2}. \tag{10.2.66}$$

10.2.6 Field continuity conditions

In the previous two sections, expressions for the plane wave and modal fields in regions D_1, D_3 and D_2 have been obtained. These expansions are now matched at the boundaries of D_2 (i.e. at $y = \pm h/2$) by imposing the continuity of $\underline{E}_t$ and $\hat{y} \times \underline{H}_t$ at these surfaces.

For purposes of clarity it is essential to define the additional notation:

$$\begin{aligned} a^*_{nm\ell} &= 2\, a_{nm\ell} \sin(k_2\, \nu'_{nm}\, h/2) \\ c^*_{nm\ell} &= 2\, c_{nm\ell} \sin(k_2\, \nu_{nm}\, h/2) \\ b^*_{nm\ell} &= 2\, b_{nm\ell} \cos(k_2\, \nu'_{nm}\, h/2) \\ d^*_{nm\ell} &= 2\, d_{nm\ell} \cos(k_2\, \nu_{nm}\, h/2). \end{aligned} \tag{10.2.67}$$

Using equations (10.2.45), (10.2.46), (10.2.48) and (10.2.49) together with (10.2.58) and (10.2.59), the continuity of $\underline{E}_t$ and $\hat{y} \times \underline{H}_t$ at $y = h/2$ is specified by:

$$\sum_{p,q} \left[(E^*_{pq} + E^*_i \, \delta_{p,0} \, \delta_{q,0}) \, \underline{\underline{RTE}}_{pq}(y = 0) + (F^*_{pq} + F^*_i \, \delta_{p,0} \, \delta_{q,0}) \, \underline{\underline{RTM}}_{pq}(y = 0) \right]$$

$$= 0 \qquad \forall \ (x,z) \notin A$$

$$= \tfrac{1}{2} \sum_{n,m,\ell} \left\{ (a^*_{nm\ell} + b^*_{nm\ell}) \, \underline{\underline{MTE}}_{nm\ell}(r,\theta) + (c^*_{nm\ell} + d^*_{nm\ell}) \, \underline{\underline{MTM}}_{nm\ell}(r,\theta) \right\}$$

$$\forall \ (x,z) \in A \qquad (10.2.68)$$

and

$$\sum_{p,q} \left[(Y^{TE}_{pq} \, E^*_{pq} - Y^{TE}_i \, E^*_i \, \delta_{p,0} \delta_{q,0}) \underline{\underline{RTE}}_{pq}(y = 0) + \right.$$

$$\left. + \ (Y^{TM}_{pq} \, F^*_{pq} - Y^{TM}_i F^*_i \, \delta_{p,0} \delta_{q,0}) \, \underline{\underline{RTM}}_{pq}(y = 0) \right]$$

$$= - \tfrac{1}{2} \, j \, r_2 \sum_{n,m,\ell} \left[\nu'_{nm} \left\{ a^*_{nm\ell} \cot(k_2 \, \nu'_{nm} \, h/2) \right. \right.$$

$$\left. - b^*_{nm\ell} \tan(k_2 \, \nu'_{nm} \, h/2) \right\} \underline{\underline{MTE}}_{nm\ell}(r,\theta)$$

$$\left. + \frac{1}{\nu_{nm}} \left\{ c^*_{nm\ell} \cot(k_2 \, \nu_{nm} \, h/2) - d^*_{nm\ell} \tan(k_2 \, \nu_{nm} \, h/2) \right\} \underline{\underline{MTM}}_{nm\ell}(r,\theta) \right]$$

$$\forall \ (x,z) \in A \qquad (10.2.69)$$

where A denotes a typical circular aperture. The corresponding equations at $y = - h/2$ are

$$\sum_{p,q} \left[\hat{E}^*_{pq} \underline{RTE}_{pq}(y = 0) + \hat{F}^*_{pq} \underline{RTM}_{pq}(y = 0)\right]$$

$$= 0 \qquad \forall \quad (x,z) \notin A$$

$$= \tfrac{1}{2} \sum_{nm\ell} \left[(b^*_{nm\ell} - a^*_{nm\ell}) \underline{MTE}_{nm\ell}(r,\theta) + (d^*_{nm\ell} - c^*_{nm\ell}) \underline{MTM}_{nm\ell}(r,\theta)\right]$$

$$\forall \quad (x,z) \in A \qquad (10.2.70)$$

and

$$\sum_{p,q} \left[Y^{TE}_{pq} \hat{E}^*_{pq} \underline{RTE}_{pq}(y = 0) + Y^{TM}_{pq} \hat{F}^*_{pq} \underline{RTM}_{pq}(y = 0)\right]$$

$$= \tfrac{1}{2} j r_2 \sum_{nm\ell} \left[\nu'_{nm} \left\{b^*_{nm\ell} \tan(k_2 \nu'_{nm} h/2) + \right.\right.$$

$$\left. + a^*_{nm\ell} \cot(k_2 \nu'_{nm} h/2)\right\} \underline{MTE}_{nm\ell}(r,\theta)$$

$$\left. + \frac{1}{\nu_{nm}} \left\{d^*_{nm\ell} \tan(k_2 \nu_{nm} h/2) + c^*_{nm\ell} \cot(k_2 \nu_{nm} h/2)\right\} \underline{MTM}_{nm\ell}(r,\theta)\right]$$

$$\forall \quad (x,z) \in A. \qquad (10.2.71)$$

10.2.7 Decoupling of the field quantities

The above continuity equations are solved using the method of moments. The symmetry of the diffracting structure is also exploited in that the four sets of unknowns are decoupled into two pairs according to their y symmetry. Had it been decided to solve the more general problem involving non-identical dielectric films such decoupling would not have been possible, thereby necessitating the solution of a far larger set of equations.

It is necessary to define the following sums and differences of plane wave coefficients:

$$\begin{aligned} ME^*_{pq} &= E^*_{pq} - \hat{E}^*_{pq}, \\ MF^*_{pq} &= F^*_{pq} - \hat{F}^*_{pq}, \\ PE^*_{pq} &= E^*_{pq} + \hat{E}^*_{pq}, \\ \text{and} \quad PF^*_{pq} &= F^*_{pq} + \hat{F}^*_{pq}. \end{aligned} \tag{10.2.72}$$

Thus, by subtracting equations (10.2.68) and (10.2.70), it can be seen that

$$\begin{aligned} &\sum_{p,q} [(ME^*_{pq} + E^*_i \delta_{p,0}\delta_{q,0})\underline{RTE}_{pq}(y = 0) + (MF^*_{pq} + F^*_i \delta_{p,0}\delta_{q,0})\underline{RTM}_{pq}(y = 0)] \\ &\quad = 0 \qquad \forall\ (x,z) \notin A \\ &\quad = \sum_{n,m,\ell} [a^*_{nm\ell}\underline{MTE}_{nm\ell}(r,\theta) + c^*_{nm\ell}\ \underline{MTM}_{nm\ell}(r,\theta)] \\ &\qquad\qquad \forall (x,z) \in A \end{aligned} \tag{10.2.73}$$

and by subtracting equation (10.2.69) from equation (10.2.71), the equivalent magnetic field equation

$$\sum_{p,q} \left[\left(Y_{pq}^{TE} \, ME_{pq}^{*} - Y_{i}^{TE} \, E_{i}^{*} \, \delta_{p,0} \, \delta_{q,0} \right) \underline{RTE}_{pq}(y = 0) \right.$$

$$\left. + \left(Y_{pq}^{TM} \, MF_{pq}^{*} - Y_{i}^{TM} \, F_{i}^{*} \, \delta_{p,0} \, \delta_{q,0} \right) \underline{RTM}_{pq}(y = 0) \right]$$

$$= - j \, r_2 \sum_{n,m,\ell} \left[\nu'_{nm} \, a^{*}_{nm\ell} \cot(k_2 \, \nu'_{nm} \, h/2) \, \underline{MTE}_{nm\ell}(r,\theta) \right.$$

$$\left. + \frac{1}{\nu_{nm}} \, c^{*}_{nm\ell} \cot(k_2 \, \nu_{nm} \, h/2) \, \underline{MTM}_{nm\ell}(r,\theta) \right]$$

$$\forall \, (x,z) \in A \qquad (10.2.74)$$

is derived. In a similar fashion, the addition of the same pairs of equations yield

$$\sum_{p,q} \left[\left(PE^{*}_{pq} + E^{*}_{i} \, \delta_{p,0} \, \delta_{q,0} \right) \underline{RTE}_{pq}(y = 0) + \left(PF^{*}_{pq} + F^{*}_{i} \, \delta_{p,0} \, \delta_{q,0} \right) \underline{RTM}_{pq}(y = 0) \right]$$

$$= 0 \qquad \forall \, (x,z) \notin A$$

$$= \sum_{n,m,\ell} \left[b^{*}_{nm\ell} \, \underline{MTE}_{nm\ell}(r,\theta) + d^{*}_{nm\ell} \, \underline{MTM}_{nm\ell}(r,\theta) \right]$$

$$\forall \, (x,z) \in A \qquad (10.2.75)$$

and

$$\sum_{p,q} \left[\left(Y_{pq}^{TE} \, PE^{*}_{pq} - Y_{i}^{TE} E^{*}_{i} \, \delta_{p,0} \, \delta_{q,0} \right) \underline{RTE}_{pq}(y = 0) \right.$$

$$\left. + \left(Y_{pq}^{TM} \, PF^{*}_{pq} - Y_{i}^{TM} \, F^{*}_{i} \, \delta_{p,0} \, \delta_{q,0} \right) \underline{RTM}_{pq}(y = 0) \right]$$

$$= j\, r_2 \sum_{n,m,\ell} [\nu'_{nm}\, b^*_{nm\ell} \tan(k_2\, \nu'_{nm}\, h/2)\, \underline{\underline{MTE}}_{nm\ell}(r,\theta)$$

$$+ \frac{1}{\nu_{nm}}\, d^*_{nm\ell} \tan(k_2\, \nu_{nm}\, h/2)\, \underline{\underline{MTM}}_{nm\ell}(r,\theta)]$$

$$\forall \quad (x,z) \in A. \qquad (10.2.76)$$

Such an analysis has demonstrated that the y-symmetry of the structure yields a set of decoupled equations.

10.2.8 Reconstitution equations

The four sets of Rayleigh modal inner products are defined by:

$$IEE^{pq}_{nm\ell} = \iint_A \underline{\underline{MTE}}_{nm\ell}(r,\theta) \cdot \overline{\underline{\underline{RTE}}}_{pq}(y = 0)\, dA \qquad (10.2.77)$$

$$IME^{pq}_{nm\ell} = \iint_A \underline{\underline{MTE}}_{nm\ell}(r,\theta) \cdot \overline{\underline{\underline{RTM}}}_{pq}(y = 0)\, dA \qquad (10.2.78)$$

$$IEM^{pq}_{nm\ell} = \iint_A \underline{\underline{MTM}}_{nm\ell}(r,\theta) \cdot \overline{\underline{\underline{RTE}}}_{pq}(y = 0)\, dA \qquad (10.2.79)$$

and $$IMM^{pq}_{nm\ell} = \iint_A \underline{\underline{MTM}}_{nm\ell}(r,\theta) \cdot \overline{\underline{\underline{RTM}}}_{pq}(y = 0)\, dA. \qquad (10.2.80)$$

In the above expressions, a superscripted bar represents complex conjugation. A summary of the derivation of closed forms for these inner products is presented in appendix (A.1).

The orthogonality of the Rayleigh modes is now invoked so as to obtain

$$ME^*_{pq} = - E^*_i \delta_{p,0} \delta_{q,0} + \sum_{n,m,\ell} \{a^*_{nm\ell} IEE^{pq}_{nm\ell} + c^*_{nm\ell} IEM^{pq}_{nm\ell}\} \quad (10.2.81)$$

$$MF^*_{pq} = - F^*_i \delta_{p,0} \delta_{q,0} + \sum_{n,m,\ell} \{a^*_{nm\ell} IME^{pq}_{nm\ell} + c^*_{nm\ell} IMM^{pq}_{nm\ell}\} \quad (10.2.82)$$

$$PE^*_{pq} = - E^*_i \delta_{p,0} \delta_{q,0} + \sum_{n,m,\ell} \{b^*_{nm\ell} IEE^{pq}_{nm\ell} + d^*_{nm\ell} IEM^{pq}_{nm\ell}\} \quad (10.2.83)$$

and

$$PF^*_{pq} = - F^*_i \delta_{p,0} \delta_{q,0} + \sum_{n,m,\ell} \{b^*_{nm\ell} IME^{pq}_{nm\ell} + d^*_{nm\ell} IMM^{pq}_{nm\ell}\} . \quad (10.2.84)$$

By appropriately adding and subtracting equations (10.2.81-84) the final field reconstitution equations are obtained.

i.e.
$$E^*_{pq} = - E^*_i \delta_{p,0} \delta_{q,0} + \tfrac{1}{2} \sum_{n,m,\ell} (a^*_{nm\ell} + b^*_{nm\ell}) IEE^{pq}_{nm\ell} + \tfrac{1}{2} \sum_{n,m,\ell} (c^*_{nm\ell} + d^*_{nm\ell}) IEM^{pq}_{nm\ell} \quad (10.2.85)$$

$$\hat{E}^*_{pq} = \tfrac{1}{2} \sum_{n,m,\ell} (b^*_{nm\ell} - a^*_{nm\ell}) IEE^{pq}_{nm\ell} + \tfrac{1}{2} \sum_{n,m,\ell} (d^*_{nm\ell} - c^*_{nm\ell}) IEM^{pq}_{nm\ell} \quad (10.2.86)$$

$$F^*_{pq} = - F^*_i \delta_{p,0} \delta_{q,0} + \tfrac{1}{2} \sum_{n,m,\ell} (a^*_{nm\ell} + b^*_{nm\ell}) IME^{pq}_{nm\ell} + \tfrac{1}{2} \sum_{n,m,\ell} (c^*_{nm\ell} + d^*_{nm\ell}) IMM^{pq}_{nm\ell} \quad (10.2.87)$$

and

$$\hat{F}^*_{pq} = \tfrac{1}{2} \sum_{n,m,\ell} (b^*_{nm\ell} - a^*_{nm\ell}) IME^{pq}_{nm\ell} + \tfrac{1}{2} \sum_{n,m,\ell} (d^*_{nm\ell} - c^*_{nm\ell}) IMM^{pq}_{nm\ell} . \quad (10.2.88)$$

Thus, if the modal coefficients for the aperture field have been found numerically, it is then a trivial task to deduce the Cartesian field coefficients in regions D_0 and D_4. (A brief discussion of the relationship between the Cartesian and TE-TM field representations is given in appendix (A.2)).

10.2.9 Application of the method of moments in deducing the equations for the modal coefficients

In this section, the method of moments used in deriving the two decoupled pairs of matrix equations relating the waveguide modal coefficients is briefly outlined.

By substituting equations (10.2.81) and (10.2.82) into equation (10.2.74) it then follows that

$$\begin{aligned}
&\sum_{p,q} \Big[Y^{TE}_{pq} \Big\{ -E^*_i \, \delta_{p,0} \, \delta_{q,0} + \sum_{n,m,\ell} \left(a^*_{nm\ell} \, IEE^{pq}_{nm\ell} + c^*_{nm\ell} \, IEM^{pq}_{nm\ell} \right) \Big\} - \\
&\qquad - Y^{TE}_i \, E^*_i \, \delta_{p,0} \, \delta_{q,0} \Big] \; \underline{\underline{RTE}}_{pq} \; (y = 0) \\
&\qquad + \sum_{p,q} \Big[Y^{TM}_{pq} \Big\{ -F^*_i \, \delta_{p,0} \, \delta_{q,0} + \sum_{n,m,\ell} \left(a^*_{nm\ell} \, IME^{pq}_{nm\ell} + c^*_{nm\ell} \, IMM^{pq}_{nm\ell} \right) \Big\} - \\
&\qquad - Y^{TM}_i \, F^*_i \, \delta_{p,0} \, \delta_{q,0} \Big] \; \underline{\underline{RTM}}_{pq} \; (y = 0) \\
= \; &- j \, r_2 \sum_{n,m,\ell} \Big[\nu'_{nm} \cot(k_2 \, \nu'_{nm} \, h/2) \, a^*_{nm\ell} \, \underline{\underline{MTE}}_{nm\ell} \; (r,\theta) + \\
&\qquad + \frac{1}{\nu_{nm}} \cot(k_2 \, \nu_{nm} \, h/2) \, c^*_{nm\ell} \, \underline{\underline{MTM}}_{nm\ell} \; (r,\theta) \Big] \\
&\qquad\qquad \forall \; (x,z) \in A. \qquad\qquad (10.2.89)
\end{aligned}$$

By multiplying this equation by $\overline{\overline{MTE}}_{NML}(r,\theta)$ and integrating over the aperture A, the following linear equation is obtained:

$$\sum_{n,m,\ell} a^*_{nm\ell} \Big[j\, r_2\, \nu'_{NM} \cot(k_2\, \nu'_{NM}\, h/2)\, \delta_{m,M}\, \delta_{n,N}\, \delta_{\ell,L} + \sum_{p,q} \{ Y^{TE}_{pq}\, IEE^{pq}_{nm\ell}\, \overline{\overline{IEE}}^{pq}_{NML} + Y^{TM}_{pq}\, IME^{pq}_{nm\ell}\, \overline{\overline{IME}}^{pq}_{nm\ell} \} \Big]$$

$$+ \sum_{n,m,\ell} c^*_{nm\ell} \Big[\sum_{p,q} \{ Y^{TE}_{pq}\, IEM^{pq}_{nm\ell}\, \overline{\overline{IEE}}^{pq}_{NML} + Y^{TM}_{pq}\, IMM^{pq}_{nm\ell}\, \overline{\overline{IME}}^{pq}_{NML} \} \Big]$$

$$= E^*_i \{ Y^{TE}_{00} + Y^{TE}_i \}\, \overline{\overline{IEE}}^{00}_{NML} + F^*_i \{ Y^{TM}_{00} + Y^{TM}_i \}\, \overline{\overline{IME}}^{00}_{NML}. \qquad (10.2.90)$$

The derivation of the above equation is assured by the complete orthogonality of the waveguide modal set, a subject which is discussed in appendix (A.3).

In a similar fashion a further three equations can be derived.

$$\sum_{n,m,\ell} a^*_{nm\ell} \Big[\sum_{p,q} \{ Y^{TE}_{pq}\, IEE^{pq}_{nm\ell}\, \overline{\overline{IEM}}^{pq}_{NML} + Y^{TM}_{pq}\, IME^{pq}_{nm\ell}\, \overline{\overline{IMM}}^{pq}_{NML} \} \Big]$$

$$+ \sum_{n,m,\ell} c^*_{nm\ell} \Big[\frac{j r_2}{\nu_{NM}} \cot(k_2\, \nu_{NM}\, h/2)\, \delta_{m,M}\, \delta_{n,N}\, \delta_{\ell,L} + \sum_{p,q} \{ Y^{TE}_{pq}\, IEM^{pq}_{nm\ell}\, \overline{\overline{IEM}}^{pq}_{NML} + Y^{TM}_{pq}\, IMM^{pq}_{nm\ell}\, \overline{\overline{IMM}}^{pq}_{NML} \} \Big]$$

$$= E^*_i \{ Y^{TE}_{00} + Y^{TE}_i \}\, \overline{\overline{IEM}}^{00}_{NML} + F^*_i \{ Y^{TM}_{00} + Y^{TM}_i \}\, \overline{\overline{IMM}}^{00}_{NML}. \qquad (10.2.91)$$

The corresponding equations for $\{b^*_{nm\ell}\}$ and $\{d^*_{nm\ell}\}$ can be obtained by substituting (-tan) for (cot) in equations (10.2.90) and (10.2.91). Thus,

$$\sum_{n,m,\ell} b^*_{nm\ell} \Big[-j\, r_2\, \nu'_{NM} \tan(k_2\, \nu'_{NM}\, h/2)\, \delta_{n,N}\, \delta_{m,M}\, \delta_{\ell,L}$$

$$+ \sum_{p,q} \{ Y^{TE}_{pq}\, IEE^{pq}_{nm\ell}\, \overline{IEE}^{pq}_{NML} + Y^{TM}_{pq}\, IME^{pq}_{nm\ell}\, \overline{IME}^{pq}_{NML} \} \Big]$$

$$+ \sum_{n,m,\ell} d^*_{nm\ell} \Big[\sum_{p,q} \{ Y^{TE}_{pq}\, IEM^{pq}_{nm\ell}\, \overline{IEE}^{pq}_{NML} + Y^{TM}_{pq}\, IMM^{pq}_{nm\ell}\, \overline{IME}^{pq}_{NML} \} \Big]$$

$$= E^*_i \{ Y^{TE}_{00} + Y^{TE}_i \}\, \overline{IEE}^{00}_{NML} + F^*_i \{ Y^{TM}_{00} + Y^{TM}_i \}\, \overline{IME}^{00}_{NML} \qquad (10.2.92)$$

and

$$\sum_{n,m,\ell} b^*_{nm\ell} \Big[\sum_{p,q} \{ Y^{TE}_{pq}\, IEE^{pq}_{nm\ell}\, \overline{IEM}^{pq}_{NML} + Y^{TM}_{pq}\, IME^{pq}_{nm\ell}\, \overline{IMM}^{pq}_{NML} \} \Big]$$

$$+ \sum_{n,m,\ell} d^*_{nm\ell} \Big[-\frac{jr_2}{\nu_{NM}} \tan(k_2\, \nu_{NM}\, h/2)\, \delta_{n,N}\, \delta_{m,M}\, \delta_{\ell,L}$$

$$+ \sum_{p,q} \{ Y^{TE}_{pq}\, IEM^{pq}_{nm\ell}\, \overline{IEM}^{pq}_{NML} + Y^{TM}_{pq}\, IMM^{pq}_{nm\ell}\, \overline{IMM}^{pq}_{NML} \} \Big]$$

$$= E^*_i \{ Y^{TE}_{00} + Y^{TE}_i \}\, \overline{IEM}^{00}_{NML} + F^*_i \{ Y^{TM}_{00} + Y^{TM}_i \}\, \overline{IMM}^{00}_{NML} \,. \qquad (10.2.93)$$

Equations (10.2.90) and (10.2.91) and also equations (10.2.92) and (10.2.93) are pairs of decoupled linear equations which are solved numerically using standard techniques.

10.3 CONFIRMATION OF THE THEORY

As discussed in [10.1], conservation of energy cannot be used as a check on the accuracy of numerical results provided by the formulation since this property is analytically satisfied by virtue of the technique

used to solve the field equations. Accordingly, the two-dimensional form of the Reciprocity Theorem [10.1], and a new amplitude constraint for the Littrow mount [10.4, 10.5, 10.8] are used to demonstrate the rigour of the theory and also the accuracy of its implementation.

10.3.1 The reciprocity test

From [10.2], the form of the Reciprocity Theorem for a doubly-periodic grating is

$$\beta_{00}(\underline{A} \cdot \underline{U}') = \beta'_{00}(\underline{A}' \cdot \underline{U}). \qquad (10.3.1)$$

Here $\underline{U}$ is the vectorial amplitude of the electric field of the (p,q)th order of diffraction excited by an incident field of vectorial amplitude $\underline{A}$. Those symbols carrying a prime refer to the second diffraction problem obtained by returning an incident beam of the same wavelength along the same path as that traversed by the (p,q)th reflected order of the first problem.

This test was performed using a grating characterized by x and z periods of 1.0 and 2.0 respectively, the periodicity axes being inclined at an angle of 70°. The cylindrical apertures of radius 0.45 and height 0.40 were filled with a plug of dielectric having a refractive index 2.0. The symmetrical sandwich was composed of elements each of thickness 0.1 and refractive index 1.5. The first diffraction problem was defined by the incidence parameters $\lambda = 0.95$, $\phi = 30^{\circ}$, $\psi = 0^{\circ}$, $\delta = 90^{\circ}$. The second diffraction problem was obtained by returning the order (-1, +1), resulting in the incidence parameters

$$\lambda' = 0.95, \quad \phi' = 69.3964^{o}, \quad \psi' = -61.2656^{o}, \quad \delta' = 0^{o}.$$

The calculations in both cases used 20 waveguide modes, and 225 plane waves. The left and right-hand sides of equations (10.3.1.) were respectively (using polar notation):

$$(0.0185481, \quad 109.617^{o}) \quad \text{and} \quad (0.0186180, \quad 109.732^{o}).$$

10.3.2 The Littrow amplitude constraint

In chapter 8, the derivation of the following expression was presented.

$$\sum_{(p,q)\in\Omega} \left[\frac{1}{\beta_{pq}} \{\alpha_p \gamma_{pq}(C_{pq} \bar{B}_{FG} + B_{pq} \bar{C}_{FG} + \hat{C}_{pq} \hat{B}_{FG} + \hat{B}_{pq} \hat{C}_{FG}\right.$$

$$\left. + (1-\alpha_p^2)(C_{pq} \bar{C}_{FG} + \hat{C}_{pq} \bar{\hat{C}}_{FG}) + (1-\gamma_{pq}^2)(B_{pq} \bar{B}_{FG} + \hat{B}_{pq} \bar{\hat{B}}_{FG})\}\right] = 0 \qquad (10.3.2)$$

where Ω is the set of propagating orders (p,q). This expression applies when the grid is operated in a Littrow mount in the order (f,g), and F = f-p, G = g-q. The above constraint is in general not precisely satisfied by the computed field amplitudes, and thus can be used as a convergence test for the algorithm. However, when using a set of Rayleigh orders truncated symmetrically about the Littrow order, equation [10.3.2] is analytically satisfied. Evidence of these assertions is presented in table (10.1).

TABLE 10.1

β_{tol}	Number of Plane Waves used	L.H.S. of Littrow Expression
1.0	4	-2.499×10^{-7}
2.0	12	-6.650×10^{-8}
20.0	225	8.248×10^{-5}

Demonstration of the Littrow phase expression for a grid defined by $d = d' = 1.0$, $a = 0.45$, $h = 0.40$, $s = 0.24$ plugged with a dielectric of refractive index 2.0 and sandwiched between symmetric dielectric layers of refractive index 1.5. The grid was operated in a $(-1,-1)^{th}$ order Littrow configuration for a normalized wavelength of $\lambda/d = 1.2001$ ($\phi = 58.0596^{\circ}$, $\psi = 45.0^{\circ}$, $\delta = 0^{\circ}$). The algorithm selects those orders (p,q) for which $|\beta_{pq}| < \beta_{tol}$ for inclusion in the calculation. Should the number of plane waves included be an even number, the algorithm produces results in accordance with symmetric truncation rules mentioned earlier. This is evidenced by the first two results, for which the L.H.S. of the Littrow expression is more than one order of magnitude less than the significance of the input data. Only in the case of $\beta_{tol} = 20.0$ does the algorithm not accommodate a symmetric set of orders, thereby giving rise to a result which can be taken as a useful convergence test.

10.3.3 Convergence testing and choice of modes

The discussion is now turned to consider the way in which the choice of waveguide and free space modes is made, in the context of a grid destined to be evaluated for use in a solar selective system. In this context, the important quantities to be calculated are the total energy transmission, E.T., and the energy transmitted in the order (0,0), E.T.(0,0). These must be determined, in general, for two orthogonal polarizations (e.g. $\delta = 0^{o}$ and $\delta = 90^{o}$) for the performance in unpolarized light to be known [10.1].

In order to determine the optimal set of waveguide modes, calculations are made for a representative set of wavelengths covering the various regions of the spectrum. These are carried out using the first twenty modes (10 horizontally polarized ($\equiv$ 6 TE and 4 TM) and 10 vertically polarized ($\equiv$ 6 TE and 4 TM)). Modes are included in accordance with a system of ordering based upon their cut-off wavelengths. The computer program calculates the energy carried by each mode from the formula derived in appendix (A.4). These energies are used to determine the most physically appropriate set of modes. Note that in circumstances where there are no mode selection rules operating, the optimal set may be wavelength dependent.

For situations possessing field symmetry the choice of the most appropriate modal set is simplified by the modal selection rules which then operate. For example with normally incident light only those modes having n = 1 can couple to the incident field (i.e. can have a non-zero inner product with the incident field term). Thus it is necessary to include only those modes having n = 1. Furthermore for a

grid having square symmetry, so that it is only necessary to carry out the calculations for normally incident light for one polarization angle, selection rules operate concerning the horizontal and vertical modes. Only those modes having a non-zero radial component of $\underline{E}$ parallel to the incident field can couple to it. Thus calculations for the particular case of normal incidence can be carried out quite economically using only five modes. For off-axis incidence the situation becomes more complex and quite large sets of modes are needed to adequately specify the spatial dependence of the modal fields. The choice of this set then becomes critical. It is believed that the results presented hereafter are accurate to within ± 2%. Note that this figure refers only to localized regions of bad convergence and that in general the accuracy is better than ± 2%. Of course, for properties such as absorptance and emittance obtained by integration over the entire spectrum such localized errors are of little consequence.

10.4 SOLAR EVALUATION OF GRIDS WITH CIRCULAR APERTURES

In this section are presented the results of a study involving the use of such grids as solar selective elements placed in front of a black body absorber. In reference [10.2] a thorough analysis of the spectral properties of grids having rectangular apertures was expounded. Although this section does not attempt nearly as comprehensive an examination, it will confirm the applicability (to grids with circular apertures) of the properties found in [10.2]. Furthermore, a detailed examination of the improvements gained by sandwiching the structure between a dielectric film pair and inserting a dielectric plug is given.

10.4.1 Grids without plugs and films

It has been shown [10.2] that the solar transmissivity of grids is closely related to the hole-area fraction (HAF). For the case of circular apertures whose diameter is 90% of the fundamental grid period the hole-area fraction corresponding to orthogonal periodicity axes is 63.6%. If the apertures are arranged to lie on the vertices of an equilateral triangle ($\eta = 60^{o}$) the HAF can be raised to 73.5%. Figures (10.4a) and (10.4b) show transmission curves for grids having aperture arrangements as described above and in table (10.2) are presented the solar absorptance values for the grids as functions of their periods. The superiority of the grid having a more closely packed arrangement of apertures is evident. Note that the equilateral triangle arrangement has two axes of the periodicity parallelogram at right angles to one another and so its transmission is independent of the polarization angle (δ) for normally incident radiation.

In figure (10.4c) is shown a transmission curve for a grid having the same HAF as the grid of figure (10.4a), but with $\eta = 60^{o}$ rather than 90^{o}. The solar transmission values for the two grids are very similar, confirming that it is the HAF rather than the placement of apertures which is important. (Note that the grid having $\eta = 60^{o}$ puts about 3% more of the transmitted radiation into the order (0,0) than does the grid having $\eta = 90^{o}$.)

Shown in figure (10.4d) are the spectral characteristics of a grid identical to that of figure (10.4b) except for a slightly smaller thickness. Their transmission curves are similar as are their solar transmission values. This grid exhibits weaker anomalies and a diminished

Figure 10.4

The spectral performance (in normally incident radiation - $\phi = 0^{o}$, $\psi = 90^{o}$ and $\delta = 90^{o}$) of four grids characterized by the following parameters:

(a) RGRID1 - d = 1.0, d' = 1.0, a = 0.45, h = 0.40, $\eta = 90^{o}$,

(b) RGRID2 - d = 1.0, d' = 0.866, a = 0.45, h = 0.40, $\eta = 60^{o}$,

(c) RGRID3 - d = 1.0, d' = 0.866, a = 0.419, h = 0.40, $\eta = 60^{o}$

and

(d) RGRID4 - d = 1.0, d' = 0.866, a = 0.45, h = 0.30, $\eta = 60^{o}$.

Shown on the following graphs is the energy transmission as a function of the normalized wavelength (λ/d). The solid curve denotes the total transmittance while the broken curve shows the transmittance of the order (0,0). A comparison of figures (10.4a) and (10.4b) reveals that it is the "aperture packing" which is the major factor in determining the transmission bandwidth of the grid. Note that the filtering action exhibited by figure (10.4d) is noticeably weaker than that shown in figure (10.4b). This is a consequence of the smaller grid thickness which manifests itself as a stronger evanescent coupling between the upper and lower surfaces of the aperture.

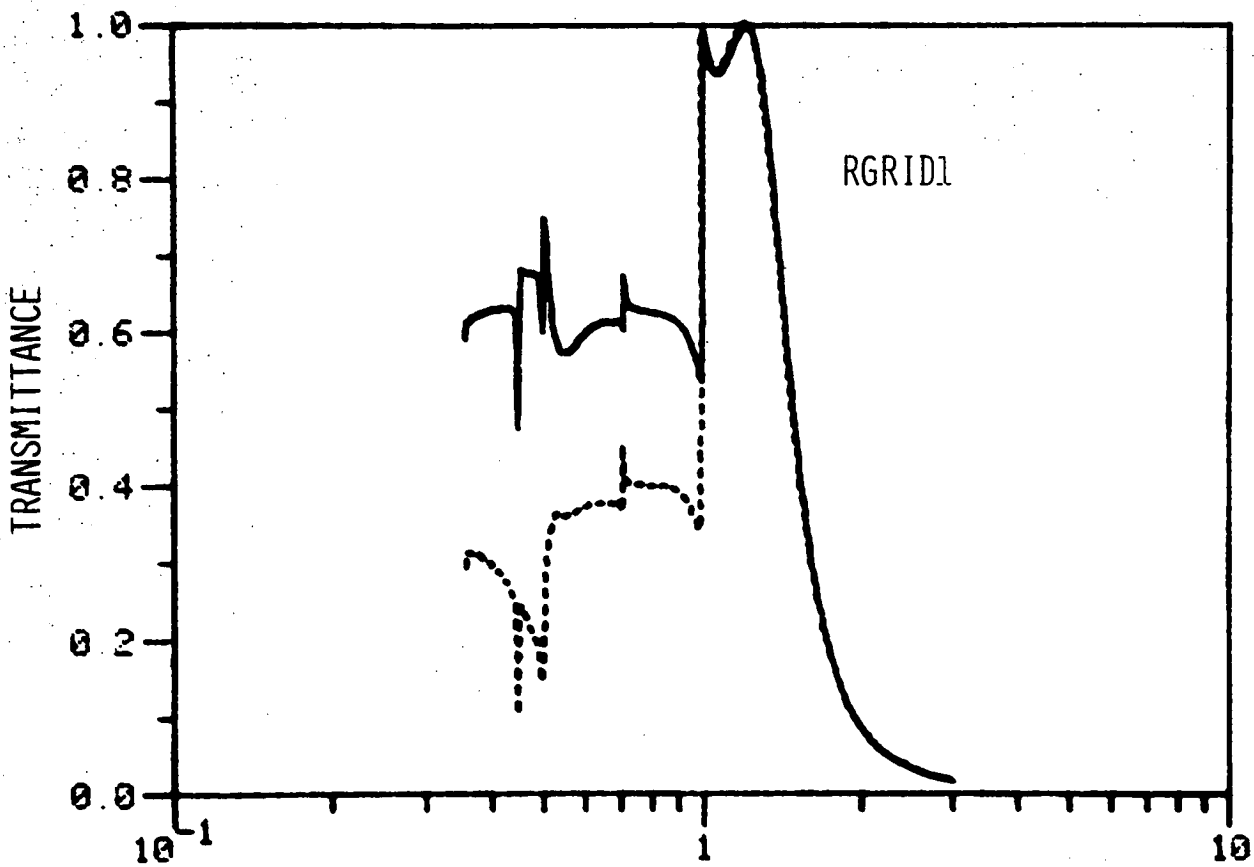


Figure 10.4a

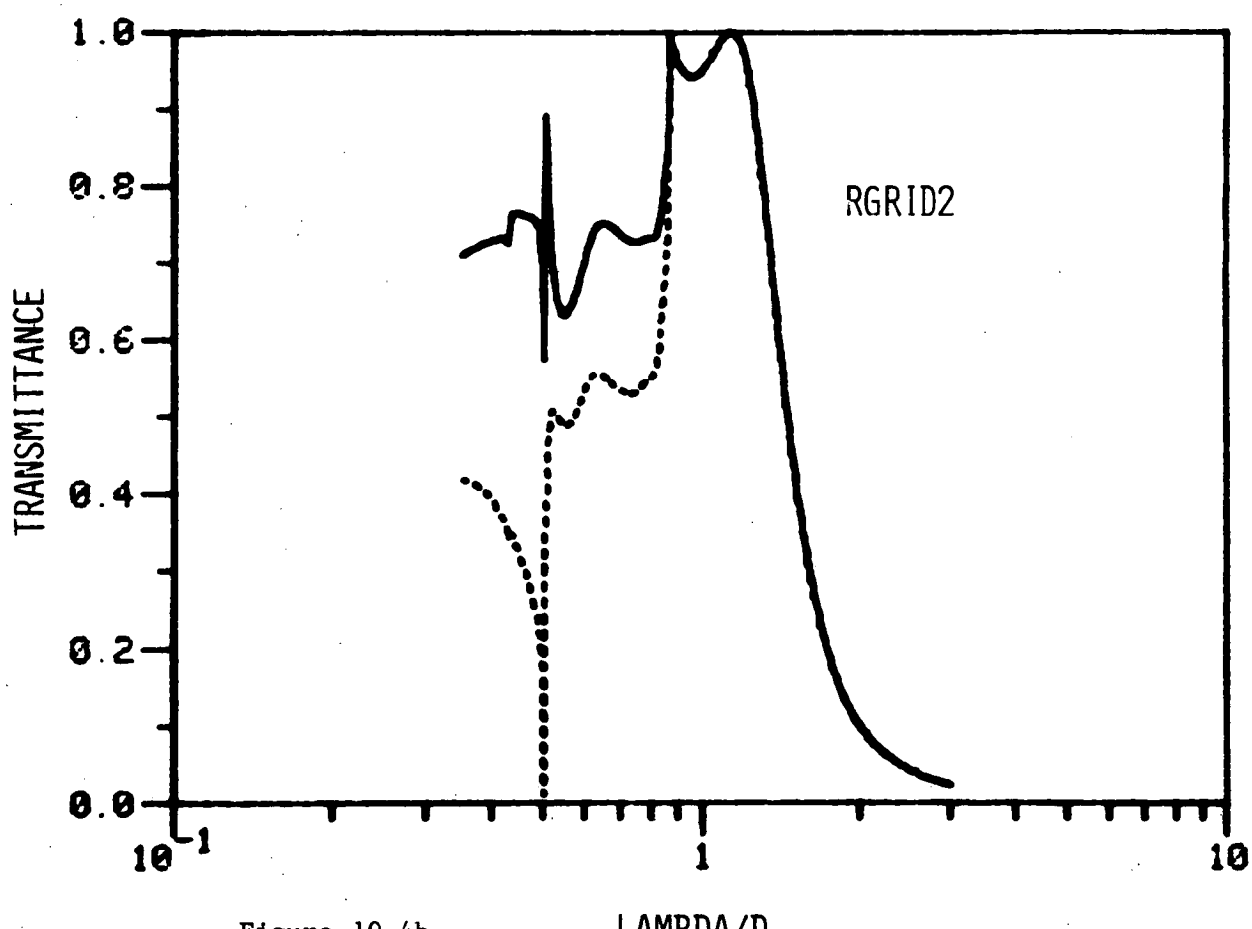


Figure 10.4b

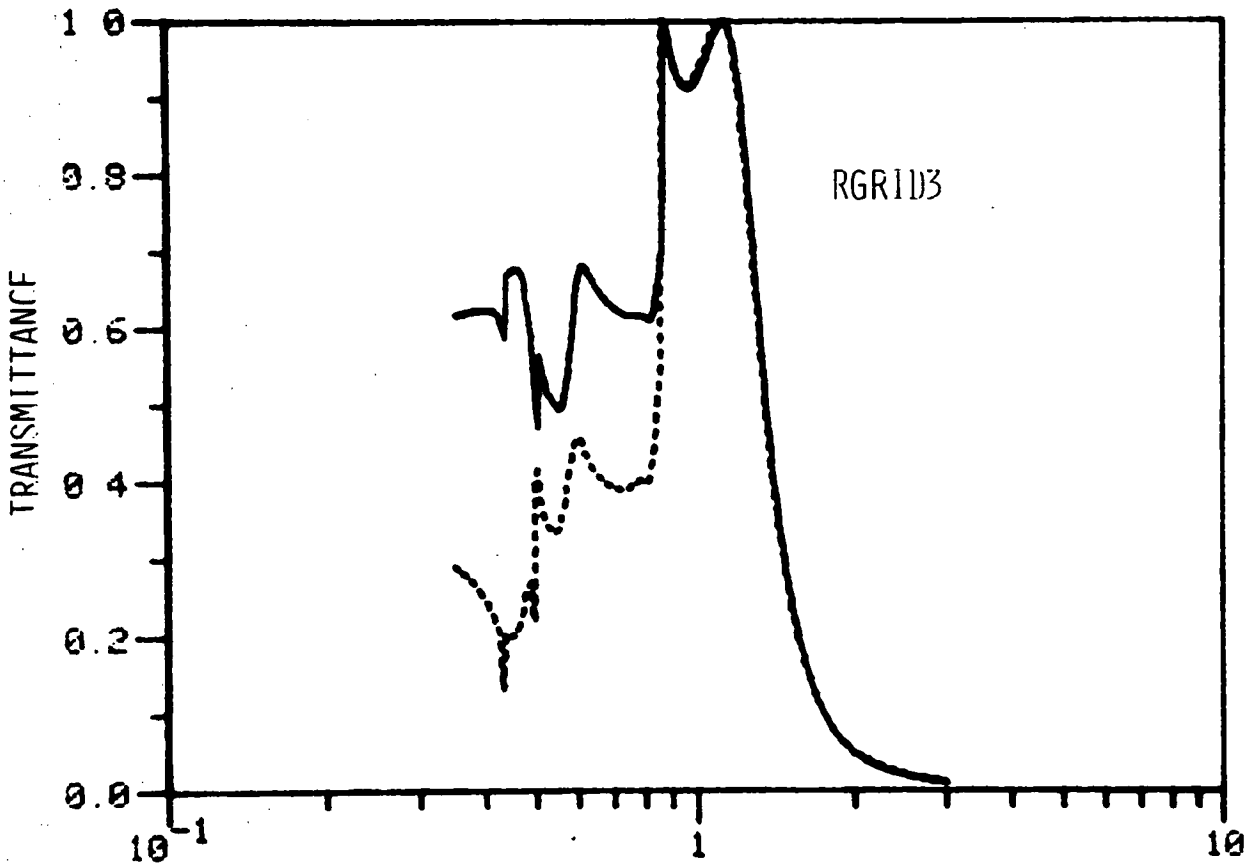


Figure 10.4c

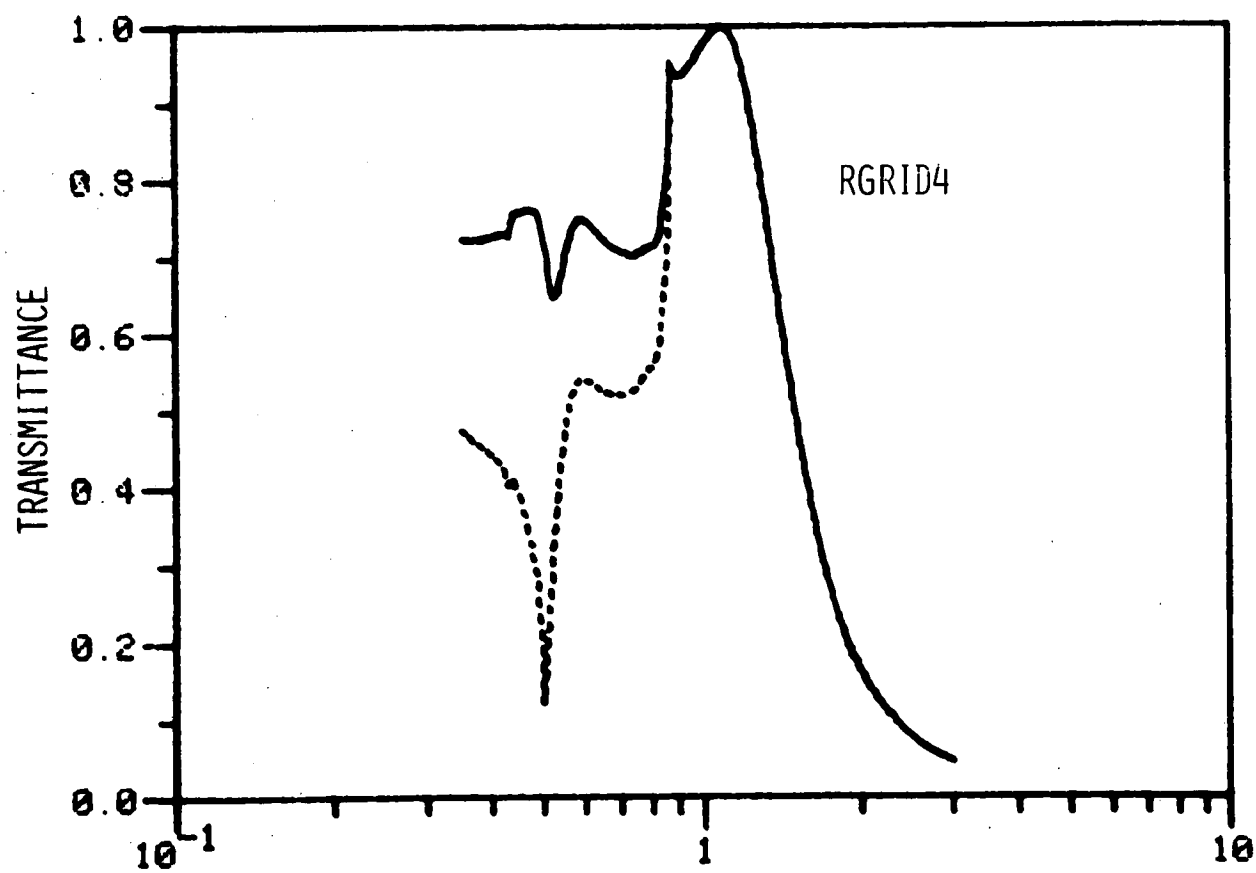


Figure 10.4d

TABLE 10.2

Grid Description	Rescaling Factor	Integrated Transmittance %
RGRID1 - No Plug	0.6	57.1
- No Film	0.8	61.9
	1.0	64.1
(Figure (10.4a))	1.2	64.1
RGRID3 - No Plug	0.6	55.7
	0.8	61.5
	0.9	63.3
	1.0	64.2
	1.1	64.4
(Figure (10.4c))	1.2	64.1
RGRID4 - No Plug	0.6	64.9
	0.8	70.0
	0.9	72.9
	1.0	74.0
	1.1	74.5
	1.2	74.7
(Figure (10.4d))	1.3	74.8

Grid Description	Rescaling Factor	Integrated Transmittance %
RGRID2 - No Film	0.6	63.4
- No Plug	0.8	70.4
- Normal Incidence	1.0	73.7
(Figure (10.4b))	1.2	74.5
RGRID2 - No Film	0.6	69.9
- Plug - thickness = 0.40μm	0.8	69.9
- r.i. = 1.50	1.0	70.7
(Figure (10.6b))		
RGRID2 - Film - thickness = 0.24μm	0.4	80.4
- r.i. = 1.50	0.5	83.5
- Plug - thickness = 0.40μm	0.6	81.5
- r.i. = 1.50	0.8	78.6
- Normal Incidence	1.0	75.9
(Figure (10.8f))	1.2	64.1
RGRID2 - Film - thickness = 0.18μm	0.4	77.4
- r.i. = 1.50	0.5	81.0
- Plug - thickness = 0.40μm	0.6	80.1
- r.i. = 1.50	0.8	78.7
- Normal Incidence	1.0	77.6
(Figure (10.8d))	1.2	75.9
RGRID2 - No Film	0.8	39.0
- No Plug	1.0	42.3
- $\phi = 60^{\circ}$ - unpolarized	1.1	42.5
(Figure (10.5c))	1.2	42.2

Figure 10.5

Shown in this set of three graphs is the performance of RGRID2 ($d = 1.0$, $d' = 0.866$, $a = 0.45$, $h = 0.40$, $n = 60^{o}$) in an off-axis configuration. Figures (10.5a-c) exhibit transmission spectra for the following incidence parameters:

(a) $\phi = 60^{o}$, $\psi = 0^{o}$, $\delta = 0^{o}$,

(b) $\phi = 60^{o}$, $\psi = 0^{o}$, $\delta = 90^{o}$, and

(c) $\phi = 60^{o}$, $\psi = 0^{o}$, unpolarized radiation.

Note the marked polarization spectra of these spectra, a feature not existing for normally incident radiation for this grid.

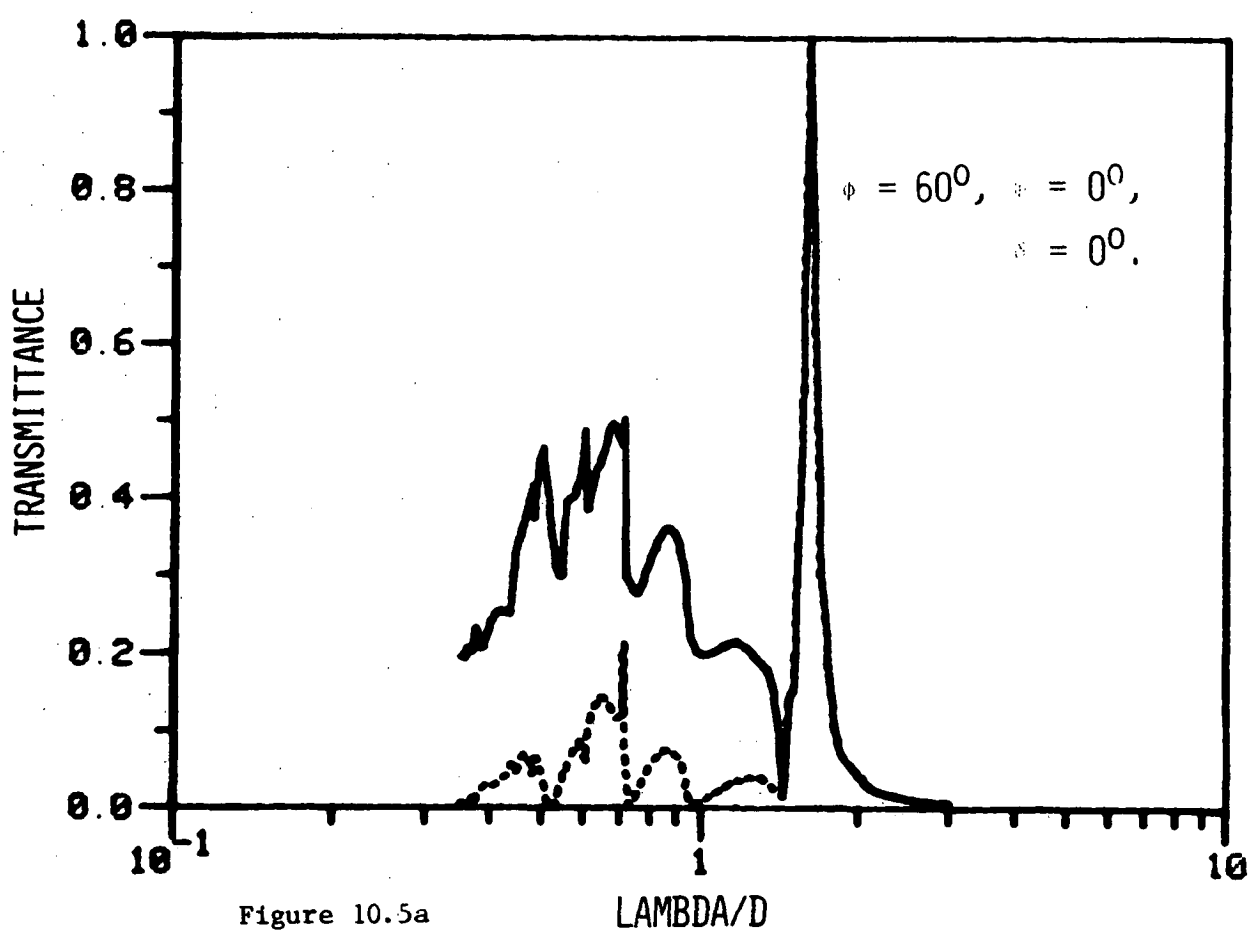


Figure 10.5a

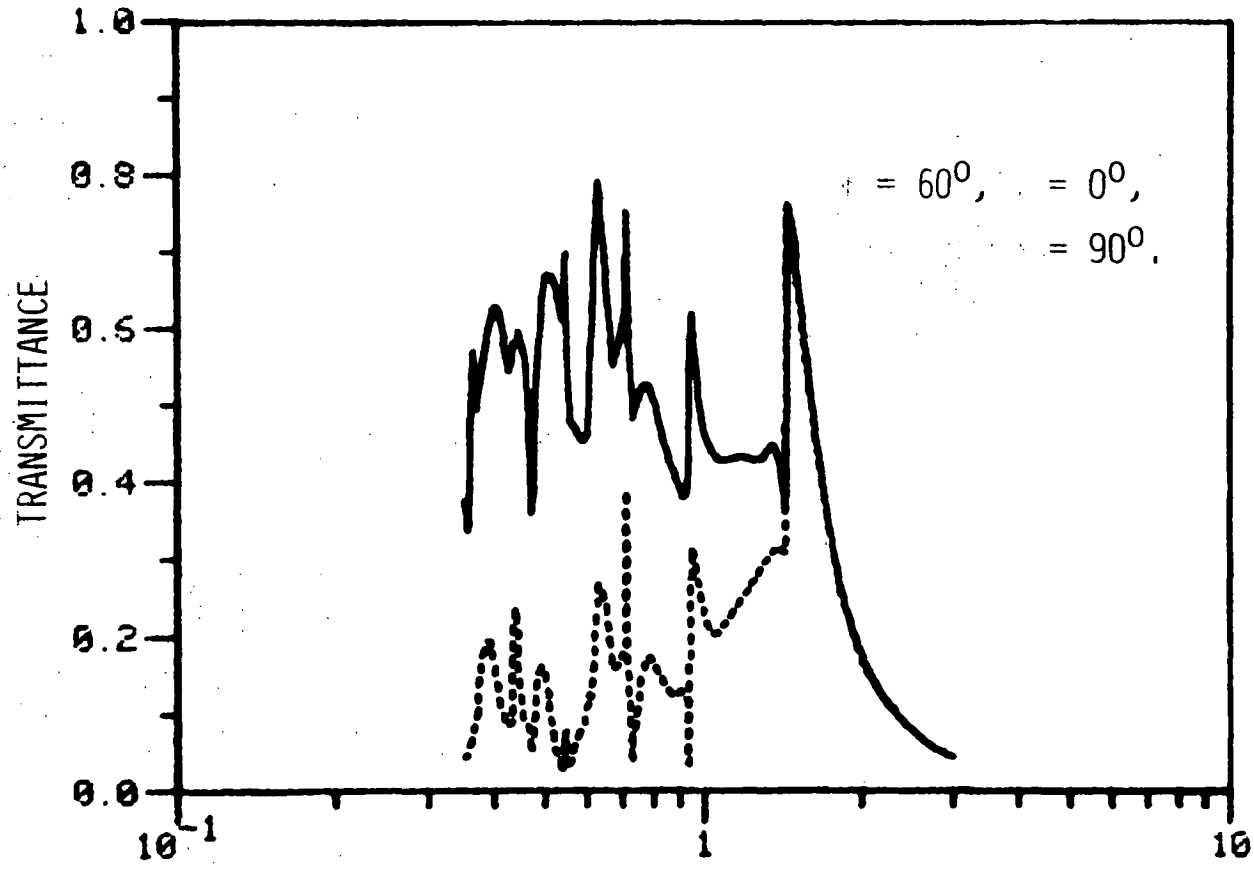


Figure 10.5b

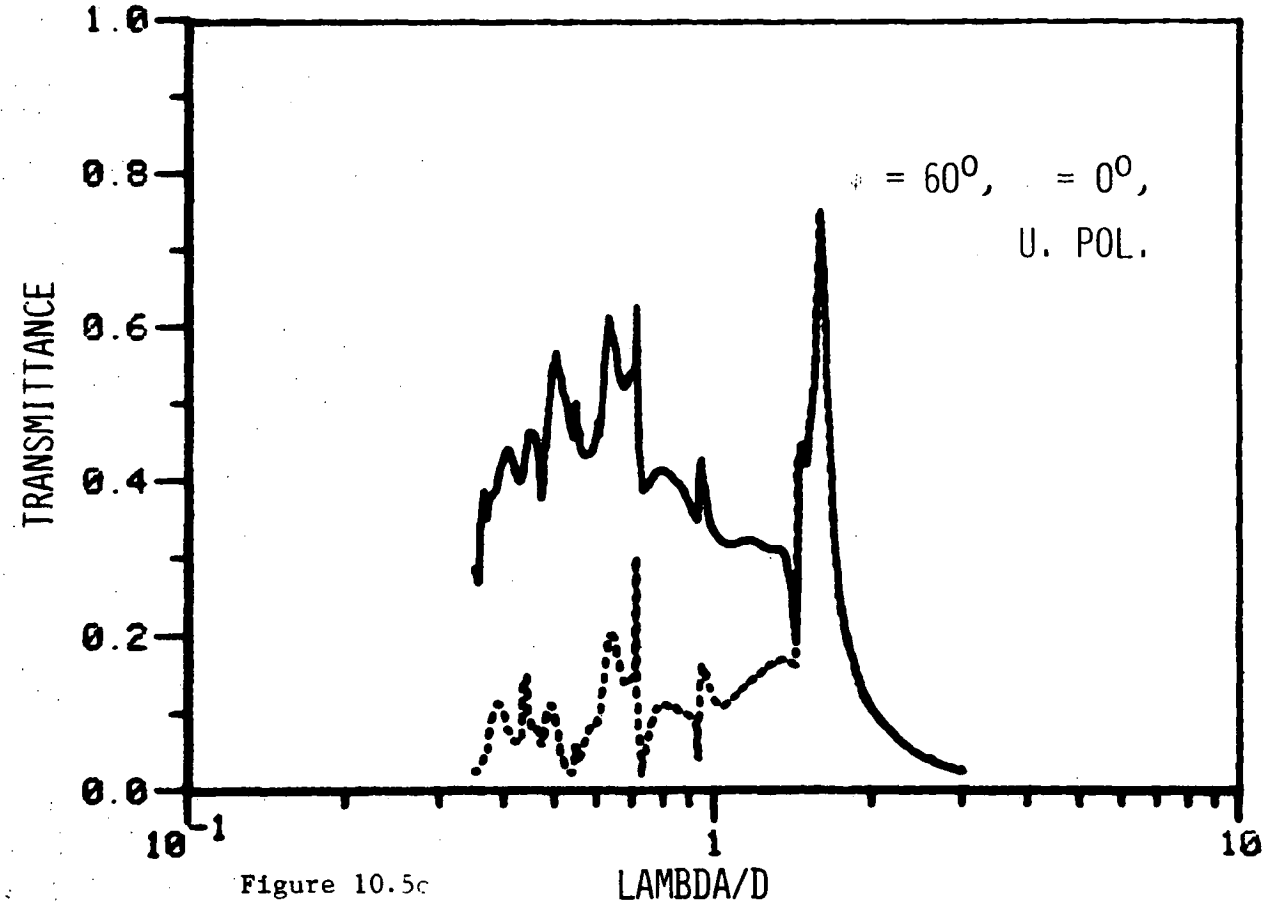


Figure 10.5c

long wavelength filtering action because of its smaller thickness (as was found for grids with square apertures [10.2]).

Finally figures (10.5a-c) demonstrates the behaviour of the optimal grid (of figure (10.4b)) when operated with an angle of incidence $\phi = 60^{o}$. The strong polarization dependence of the transmittance is at once evident. The effect of increasing ϕ to 60^{o} is to lower the transmittance from around 70% to approximately 40%. As in [10.2] one concludes that it is imperative to use these grids in direct illumination and to have their surface directed normally towards the sun.

10.4.2 Grids with plugs only

Let us now consider the effect on the performance of the optimal grid of introducing plugs of refractive index 1.5 in place of air in the circular apertures. As can be seen by comparison of figures (10.6a-b) the presence of the plug has the effect of broadening and complicating the resonance region of the transmission curve. The single resonance peak of figure (10.6a) for $\frac{\lambda}{d} > 1.0$ is split into a pair of resonance maxima. As is shown in figure (10.6c), these move away from the Wood anomaly at $\frac{\lambda}{d} = 0.866025$ towards longer wavelengths with increasing optical thickness of the plug. Further increases in the optical depth would be expected to result in the creation of additional grid resonances.

Another aspect of the performance illustrated by figures (10.6b-c) is the frequency and strength of the diffraction anomalies occurring in the transmission region, $\frac{\lambda}{d} < 0.866025$. This is attributed to the greater number of waveguide modes capable of propagating (i.e. ν_{nm} real) for a given wavelength. The frequency of the anomalies makes difficult

Figure 10.6

The accompanying graphs show the changes in grid performance brought about by the introduction of a lossless dielectric plug. The basic geometry of the structures examined is that of RGRID2. Shown in the following three graphs are spectra for:

(a) RGRID2 (for comparison purposes),

(b) RGRID2 with a plug of refractive index 1.5 + i 0.0

and

(c) a structure similar to RGRID2 but with a plug of normalized thickness 0.6 and refractive index 1.5 + i 0.0.

Note that the presence of the plug tends to split the single resonance peak and hence broadens the transmission bandwidth of the grid.

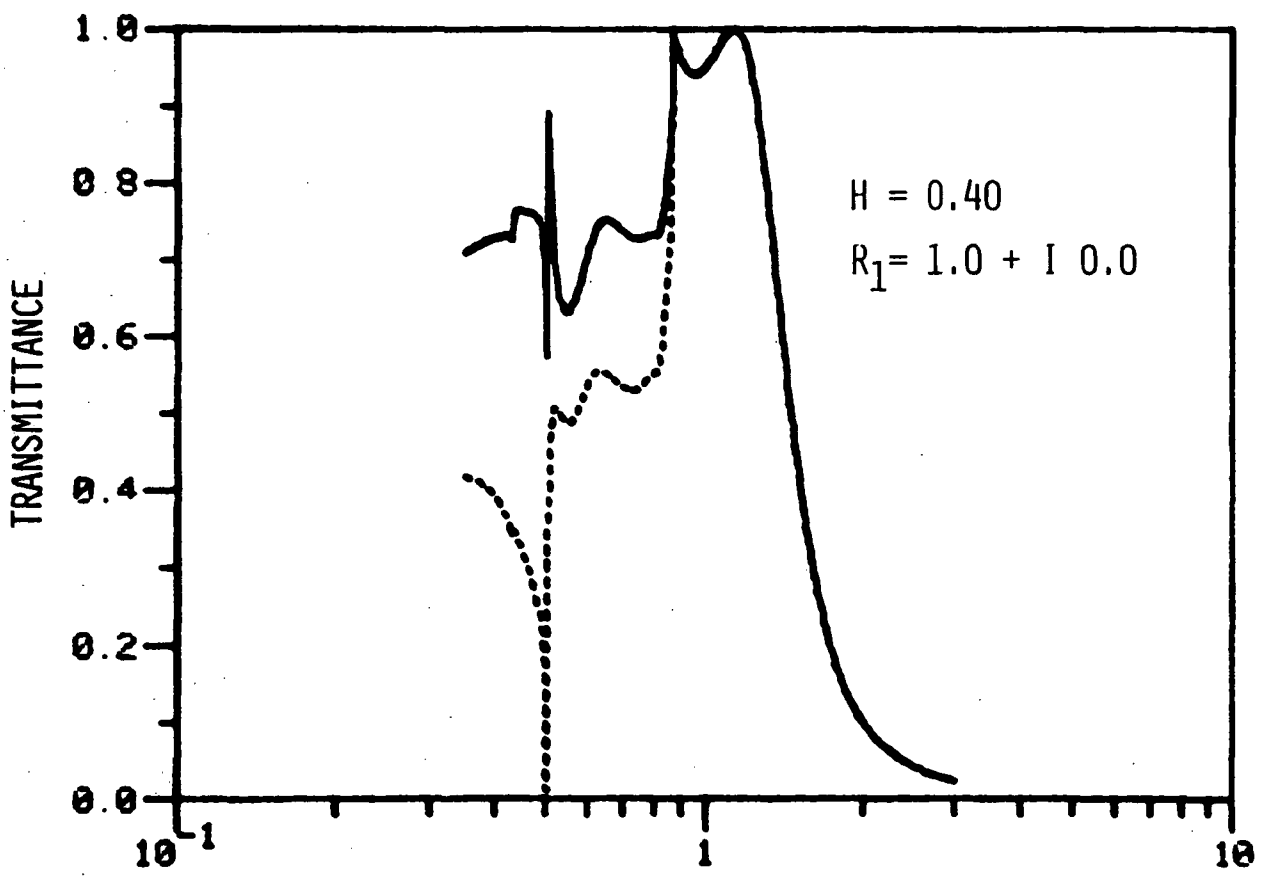


Figure 10.6a

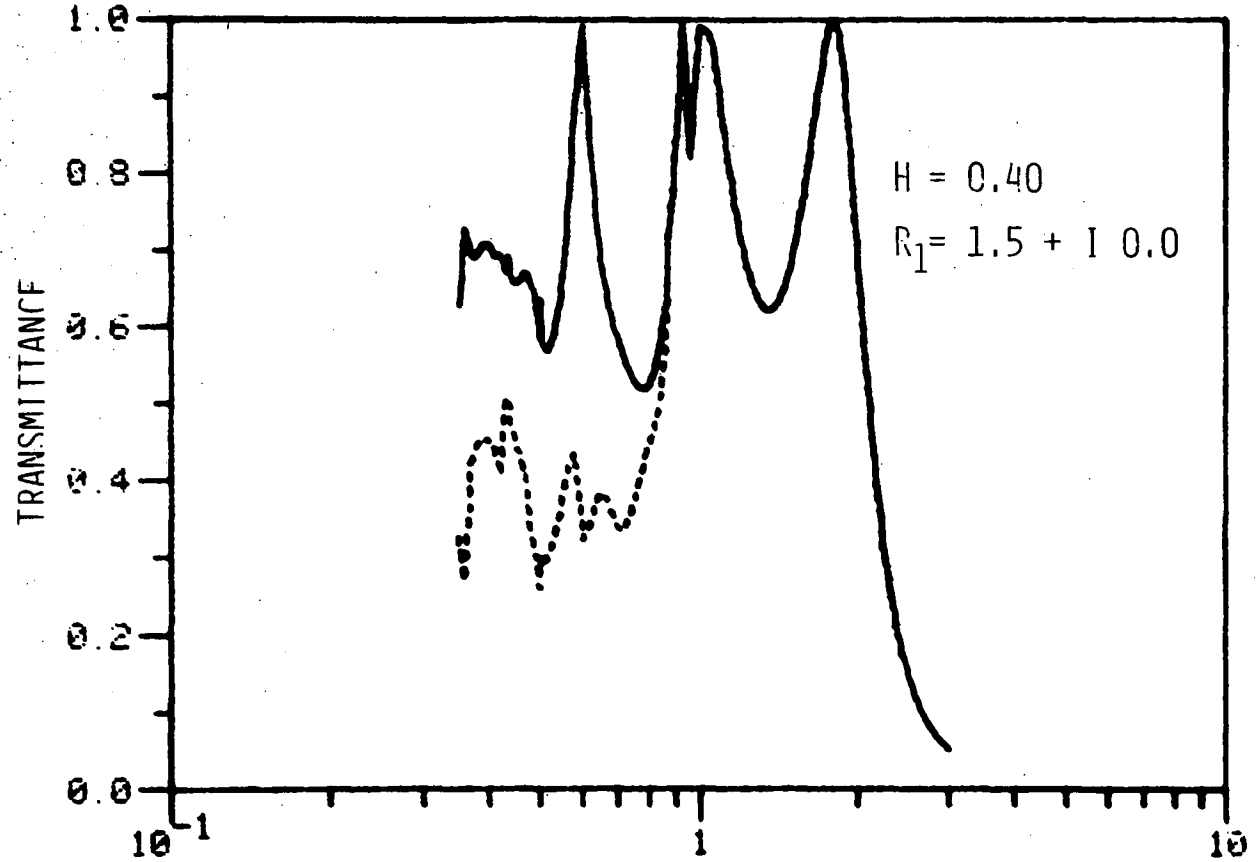


Figure 10.6b

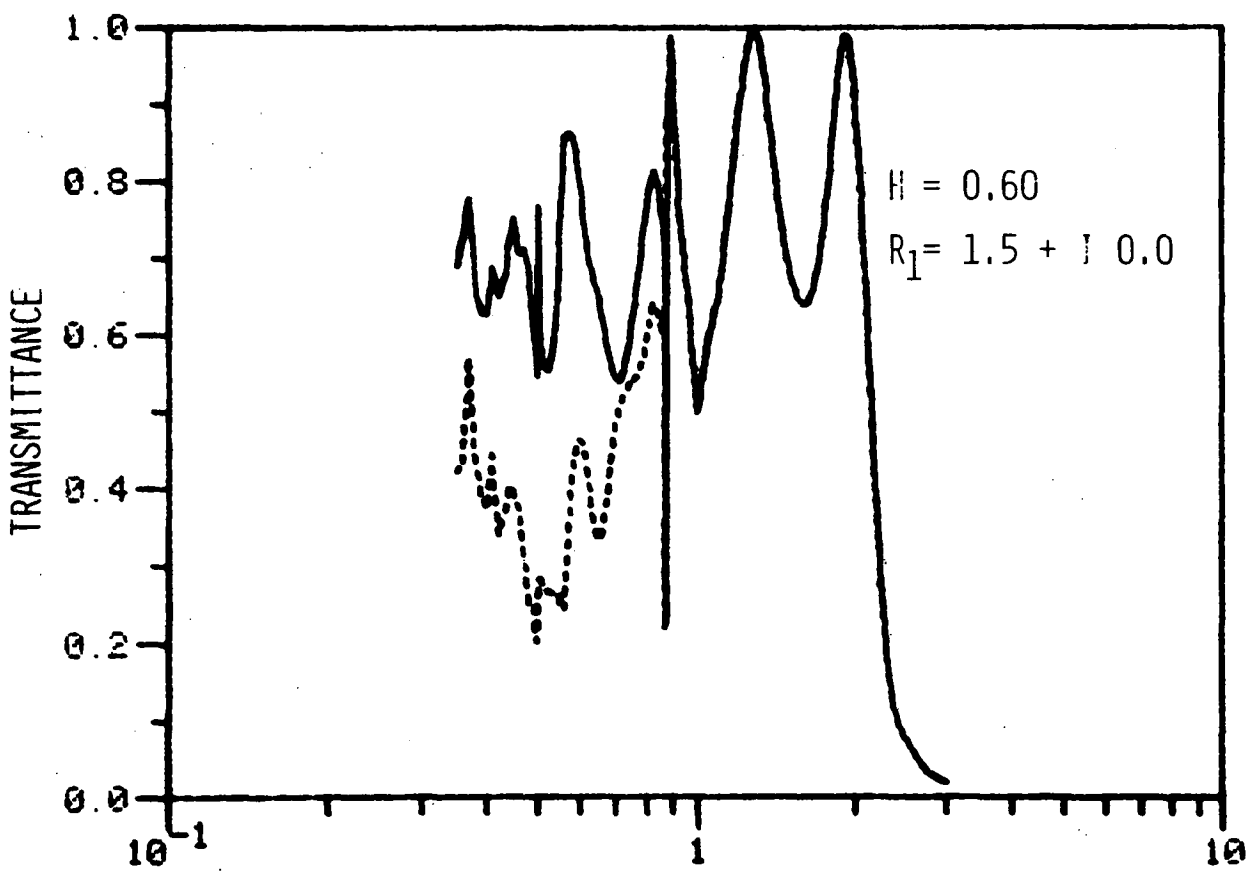


Figure 10.6c

the characterization of convergence over the entire spectrum using only a restricted set of sample points. However this is not of great importance as the anomalies would tend to be greatly smoothed in any experimental situation because of inevitable imperfections in manufacture and measurement.

The solar absorptance of the grids shown in figures (10.6b-c) is several percent lower than that of figure (10.6a). Nevertheless, it has been noted that the insertion of plugs has had the potentially beneficial effect of widening the resonance region.

10.4.3 Grids with films only

This section examines the behaviour of the optimal grid with a symmetric sandwich of dielectric material (refractive index of 1.5) surrounding it. The performance for sandwich thicknesses of 0, 0.08 and 0.12 of the grid period is shown in figures (10.7a-c). Note that non-zero sandwich thickness leads to lower performance in the transmission region, and also to frequent and strong anomalies there. These anomalies have three sources:

(a) Wood anomalies - arising when propagating orders in the sandwich become evanescent. Because of the refractive index of the film, it is possible to support a higher number of propagating orders than in air with the consequence of more frequent anomalies.

(b) Grid resonance anomalies - arising within the aperture regions which are now leaky resonant cavities. The presence of the sandwich layer bounding the air filled cavity represents an additional impedance mismatch with consequent reflection losses.

Figure 10.7

The effects of sandwiching the basic structure, "RGRID2", between symmetric dielectric layers of a material having a refractive index of 1.5 + i 0.0. Graphs for normalized film thicknesses of

(a) 0.00 (c) 0.12

(b) 0.08 (d) 0.24

are presented. The most noticeable feature of these graphs is the profusion of anomalous features occurring within the transmission region introduced by the presence of the lossless films. It is these anomalies which generally degrade the overall performance of the basic grid shown in figure (10.7a). Note the virtual elimination of the resonance dip occurring just on the long wavelength side of the Wood anomaly at $\lambda/d = 1.00$.

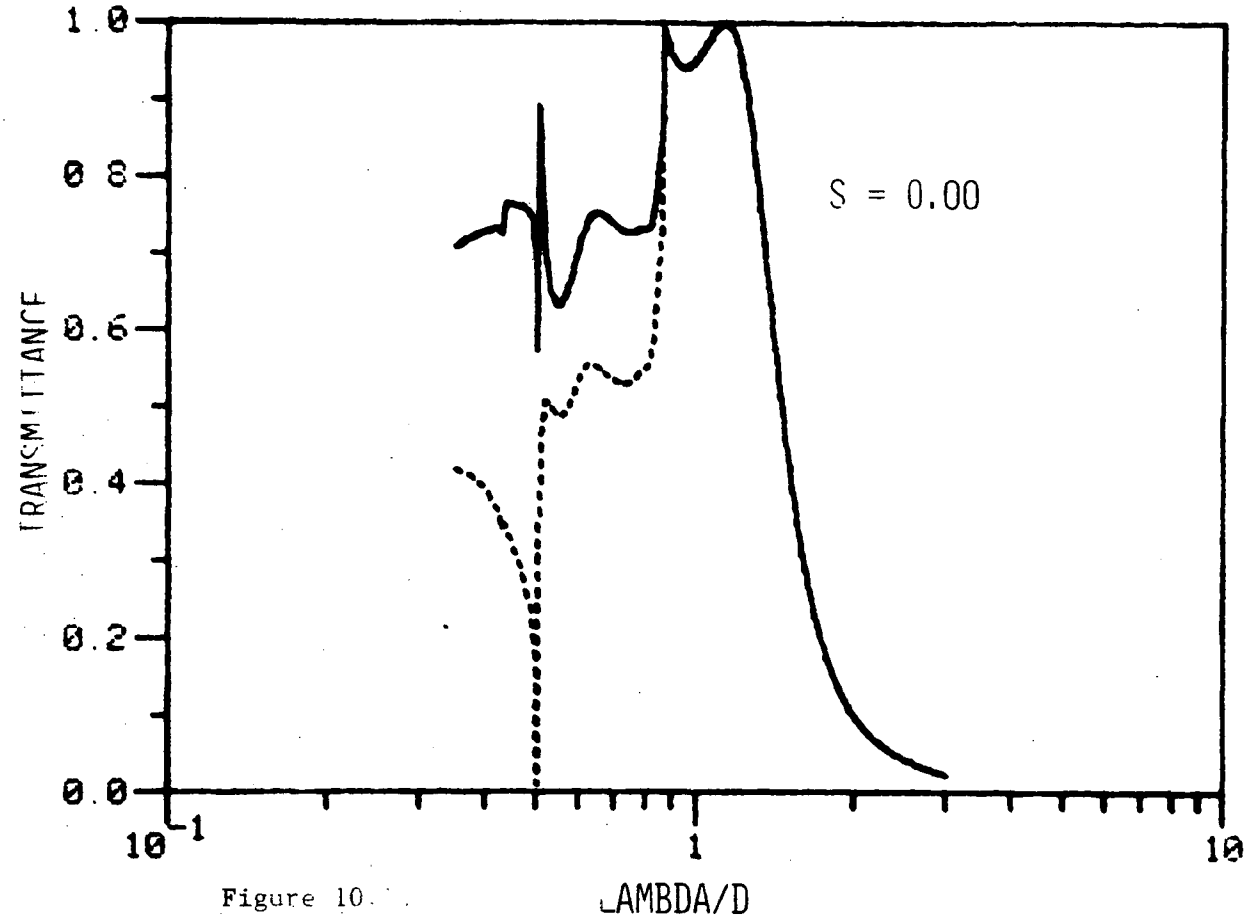


Figure 10.

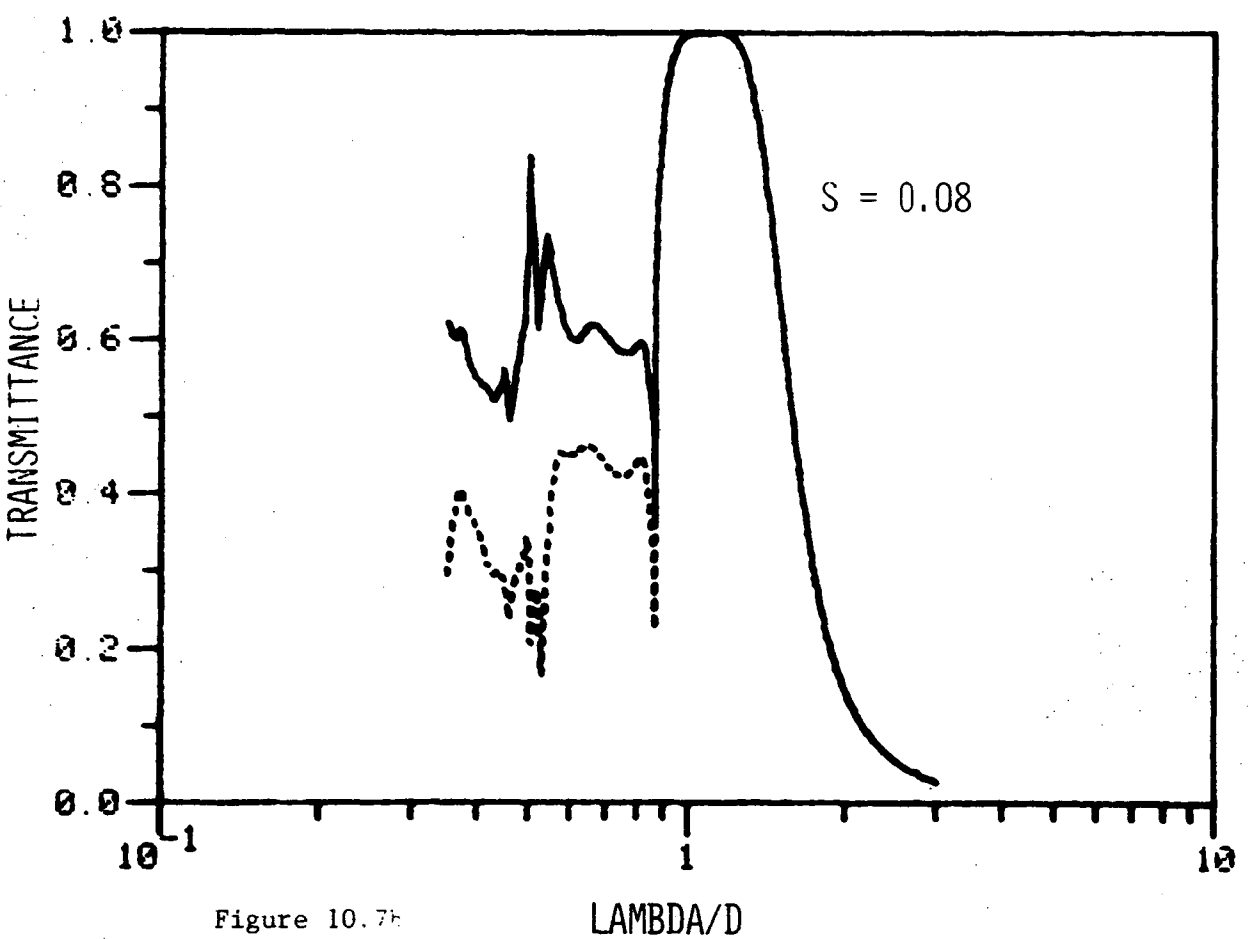


Figure 10.7b

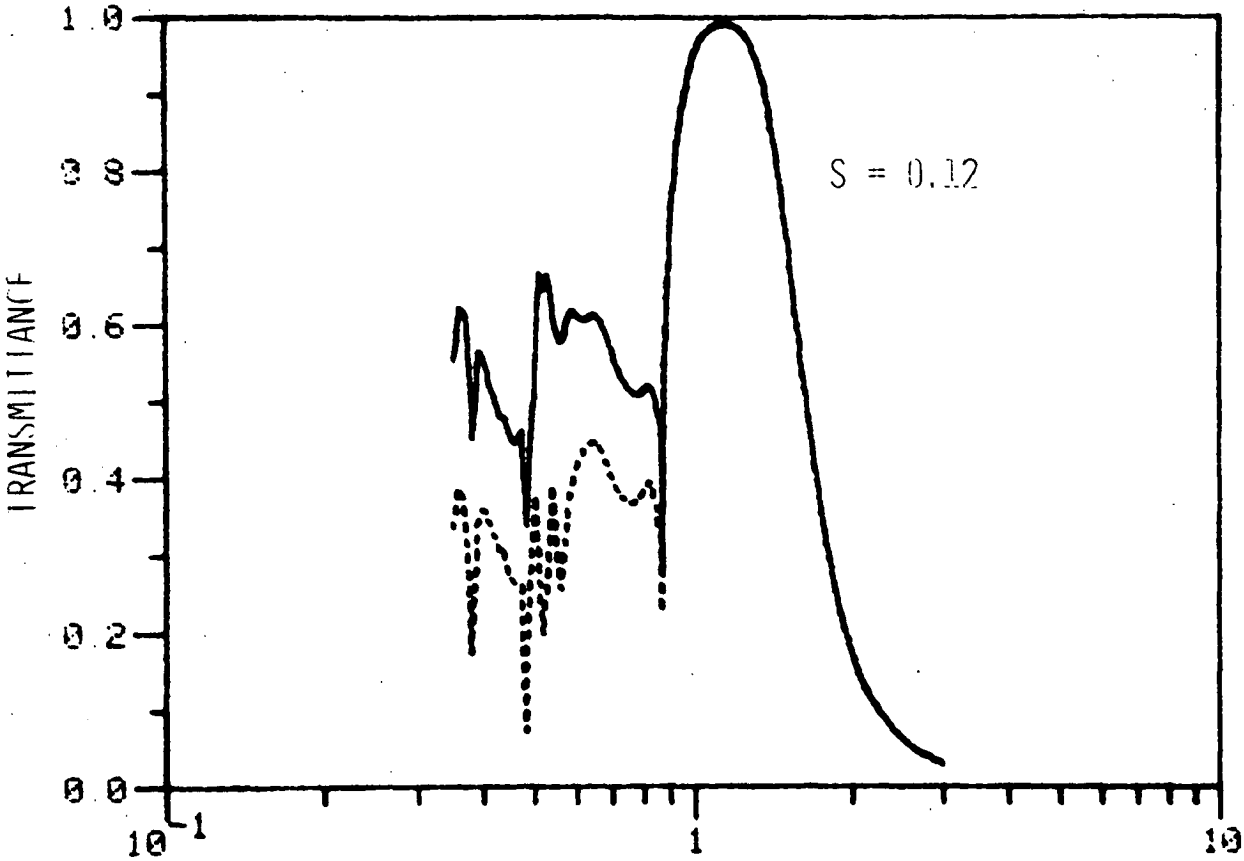


Figure 10.7c

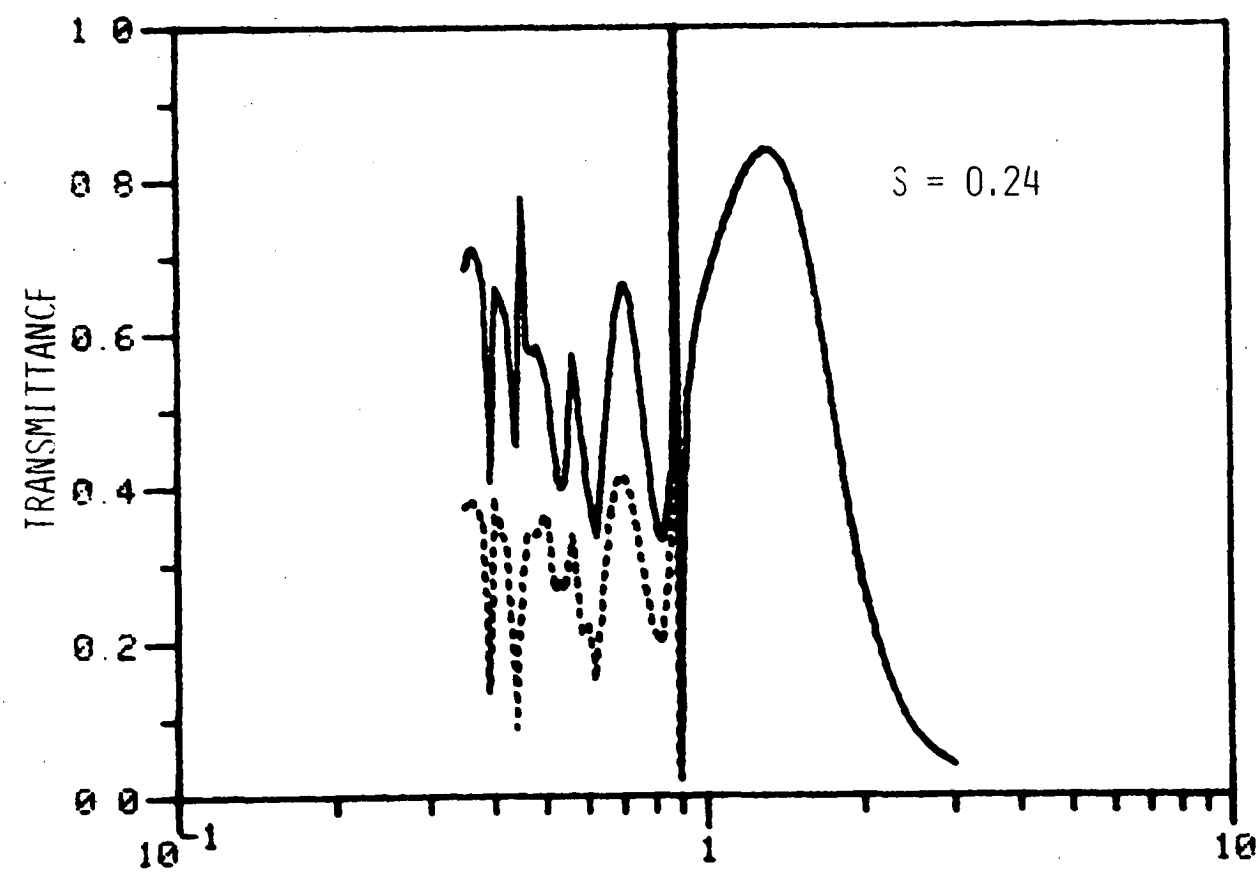


Figure 10.7d

It is these losses which are responsible for the lowering of the mean transmission values in this region.

(c) Thin film interference fringes - become more evident as the optical thickness of the sandwich increases.

The encouraging feature evident in figures (10.7) is that the introduction of the sandwich structure has effectively enabled the elimination of the resonance dip occurring just after the final Wood anomaly. It was this dip which lowered the transmission in the resonance region for grids with plugs only. Consequently, it was deduced that a combination of the plug and sandwich structures would lead to a performance possessing an enhanced bandwidth.

10.4.4 Grids with both films and plugs

In figures (10.8) are presented a comprehensive set of curves illustrating the changes in grid behaviour with varying sandwich thickness for a fixed plug thickness. The refractive index of the plug has been chosen to be equal to that of the sandwich in order to minimise reflection losses at their interface. The performance figures shown in table (10.2) for two of these grids illustrate the great improvement in solar transmittance which has been achieved by the modification of the grid geometry suggested above. It should be noted that the optimal performance is achieved for grid periods around 0.5μm, whereas for the unmodified grid it was approximately 1.0μm. This change has come about because the resonance region has been incorporated into the transmission region. This is the first time for which it has been possible to surpass the geometrical optics limit (the hole-area fraction) on the transmittance.

Figure 10.8

In this series of graphs the effects on the transmission properties of the basic structure "RGRID2" caused by introducing both dielectric plugs and symmetric film layers (having refractive indices of 1.5 + i 0.0) are demonstrated. Figure (10.8a) showing the performance of "RGRID2" is once again included for comparison purposes. Figures (10.8b-h) show transmission spectra for normalized film thicknesses of 0.08, 0.12, 0.18, 0.21, 0.24, 0.27 and 0.30 respectively. Note the marked improvement in the solar selective properties which may be achieved using the suggested modifications.

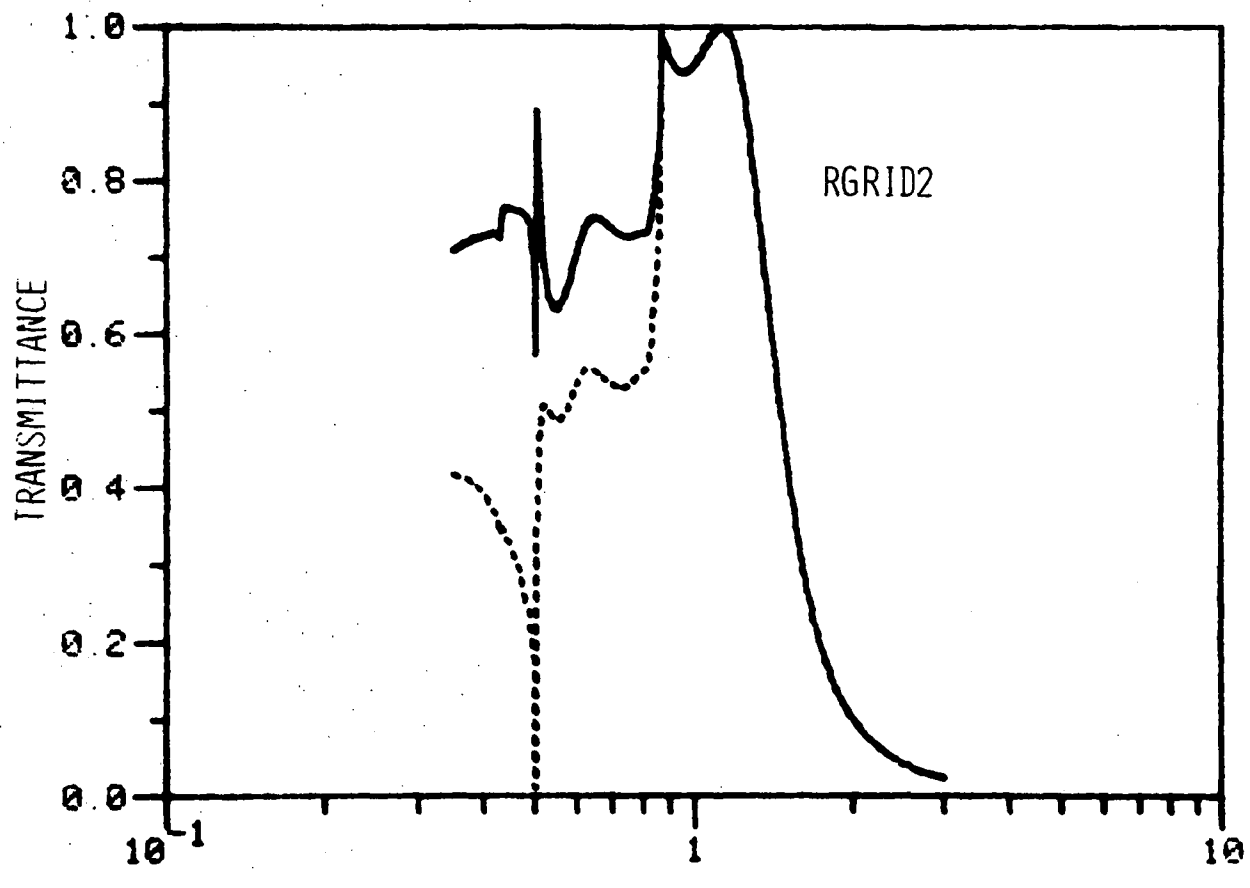


Figure 10.8a

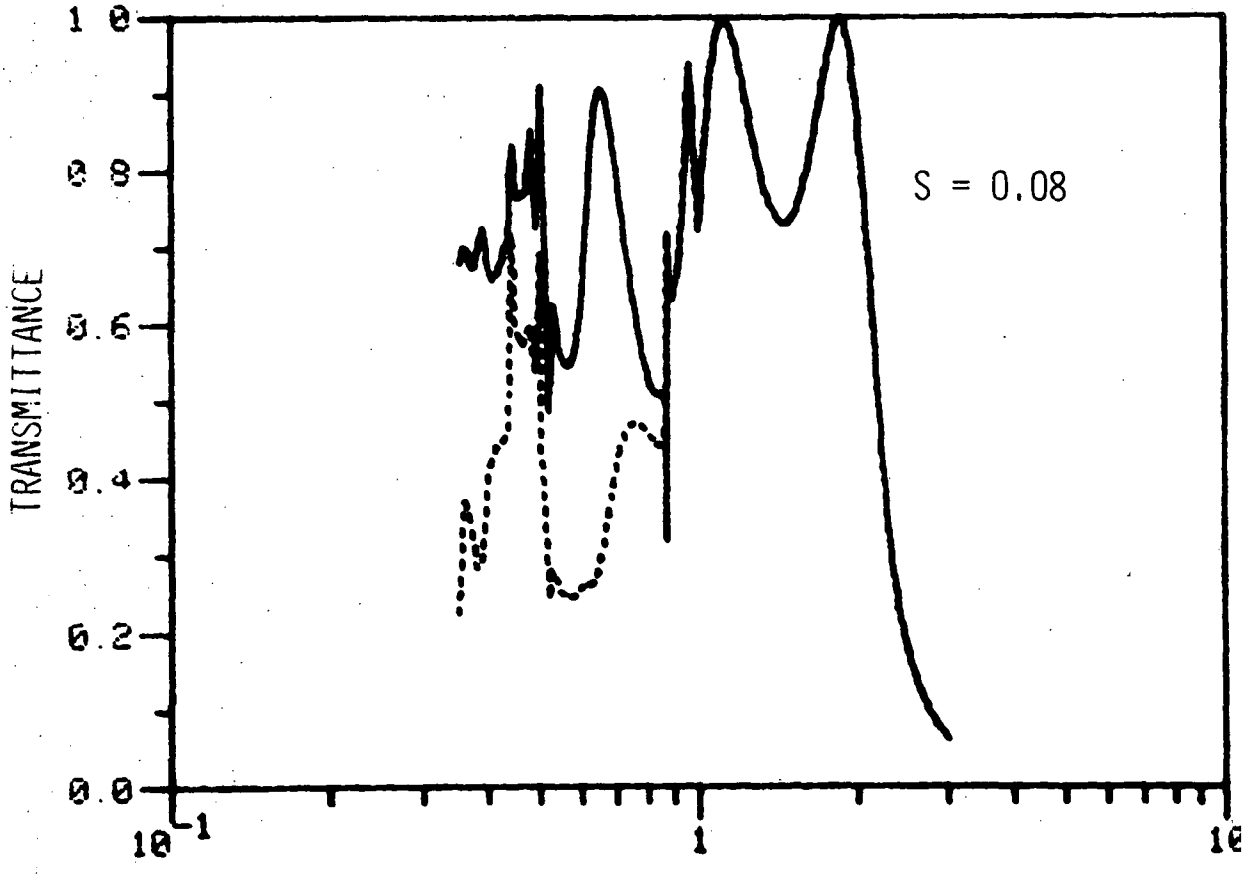


Figure 10.8b

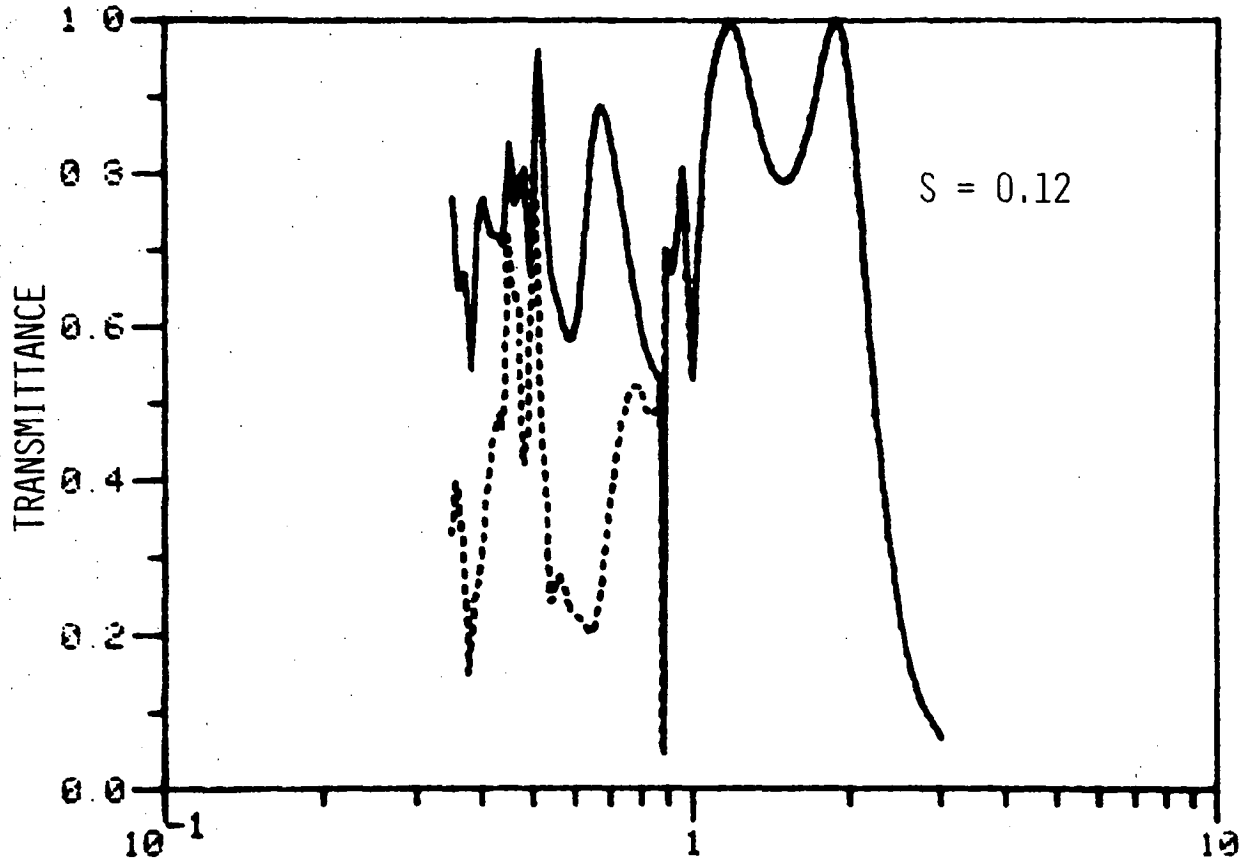


Figure 10.8c

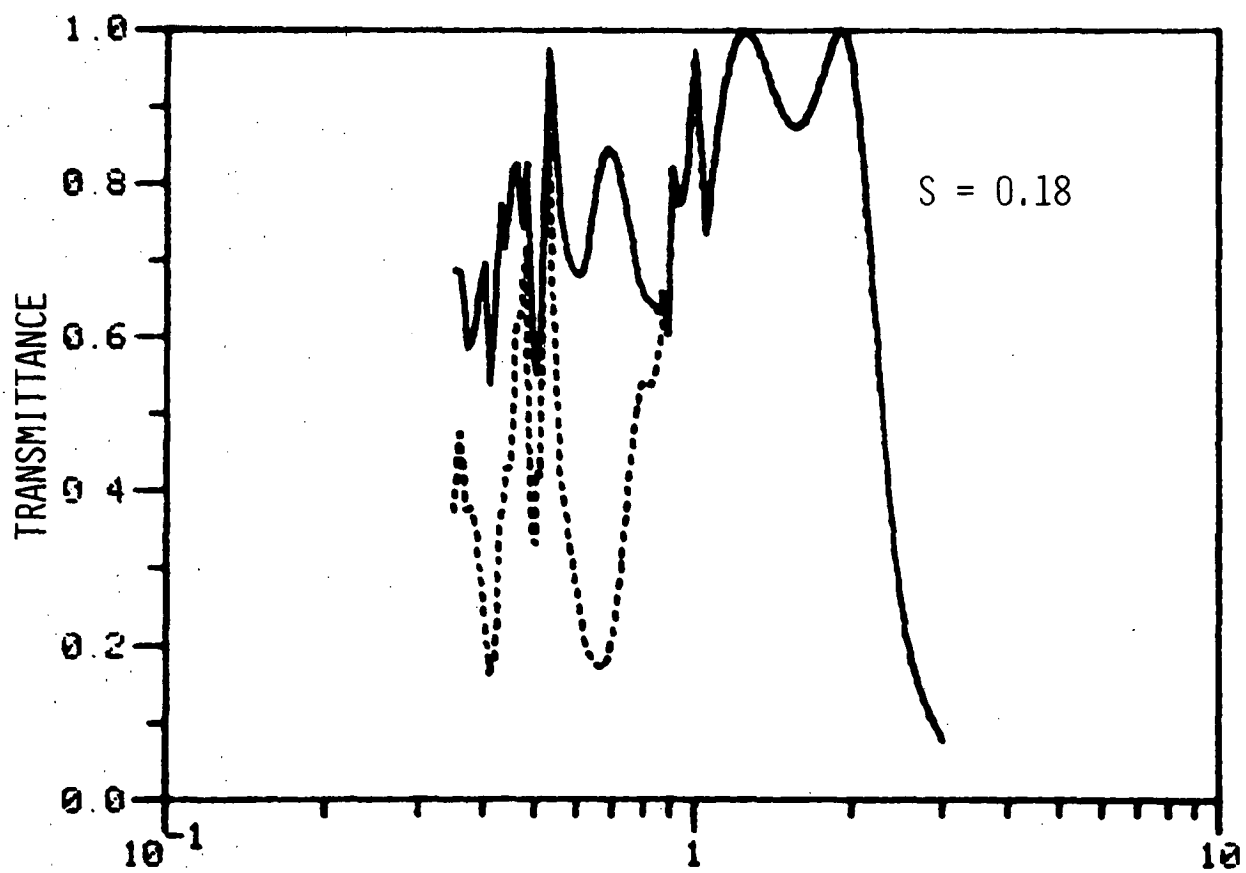


Figure 10.8d

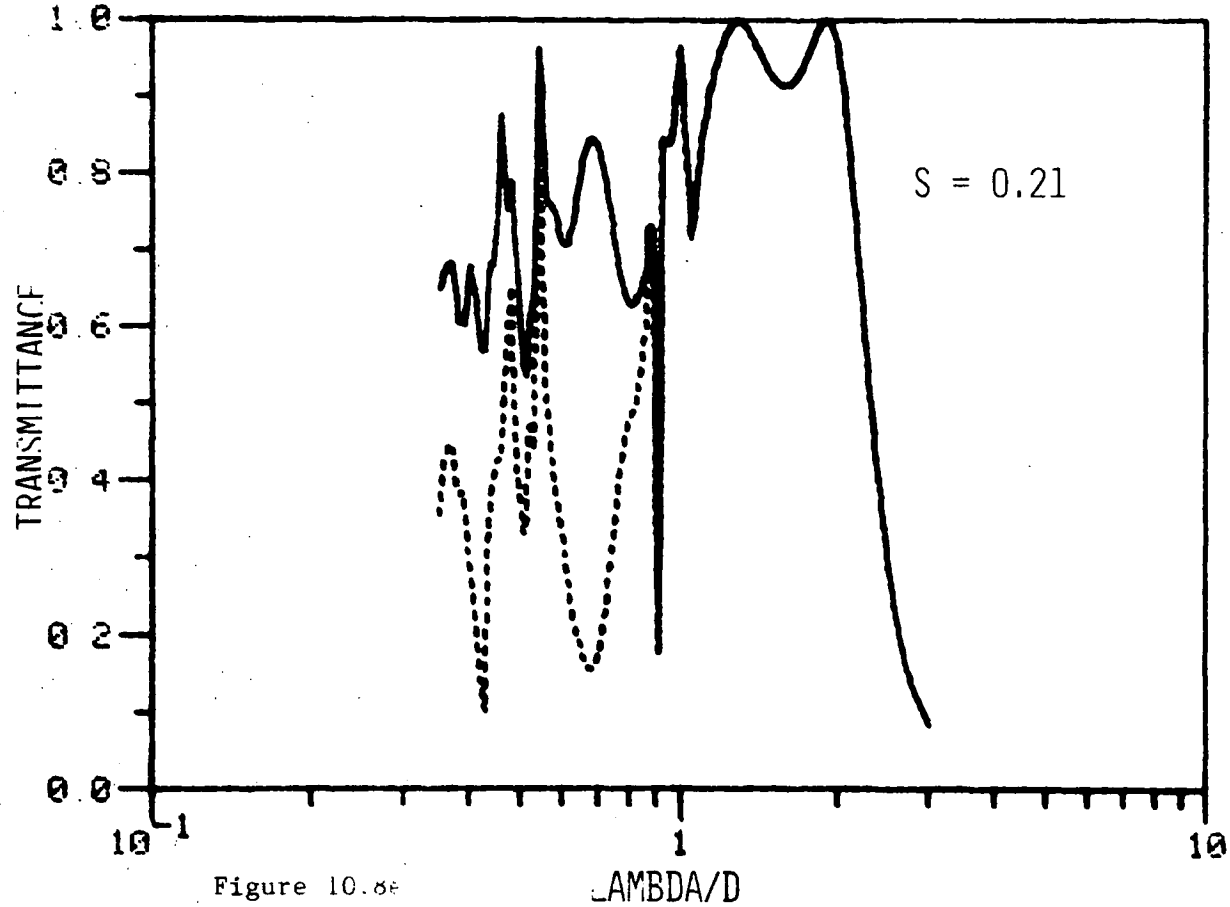


Figure 10.8e

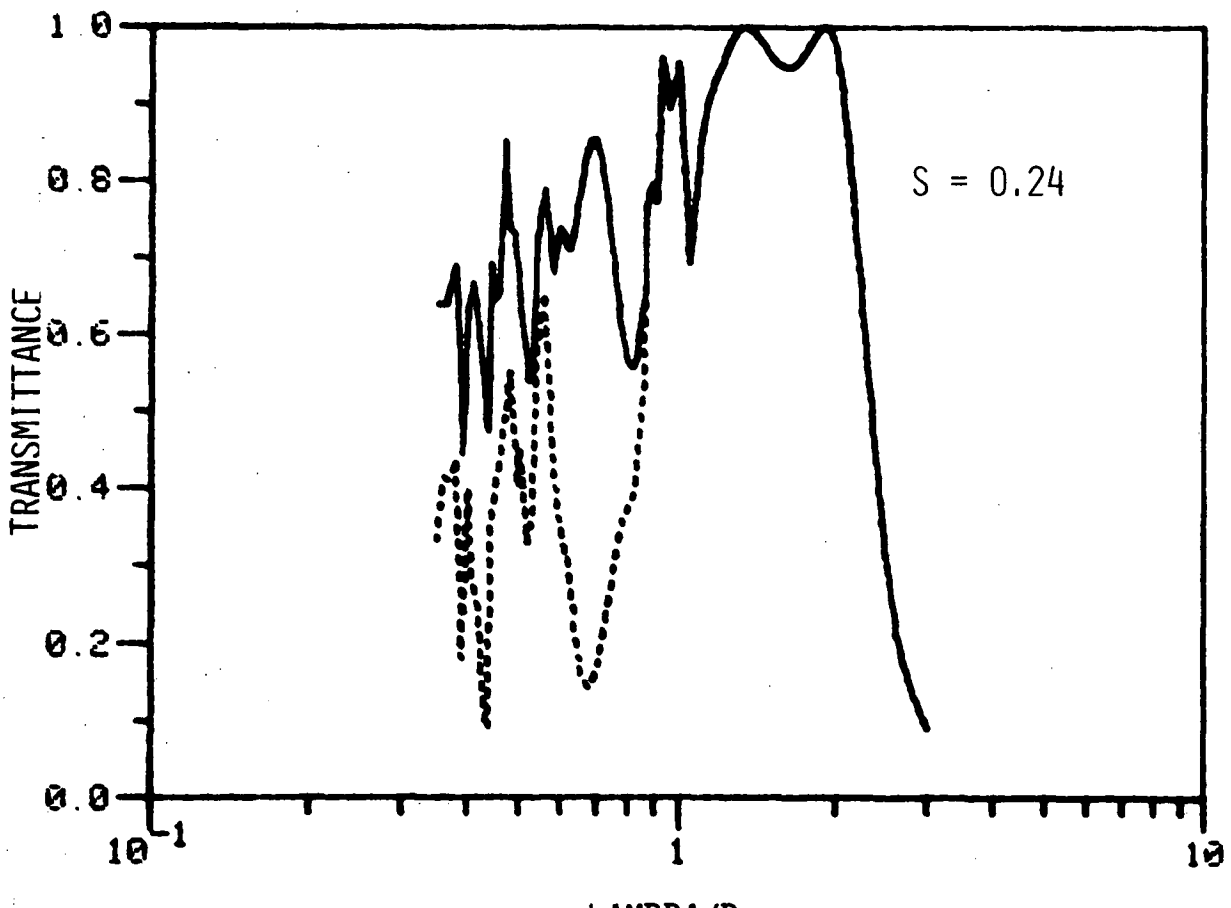


Figure 10.8f

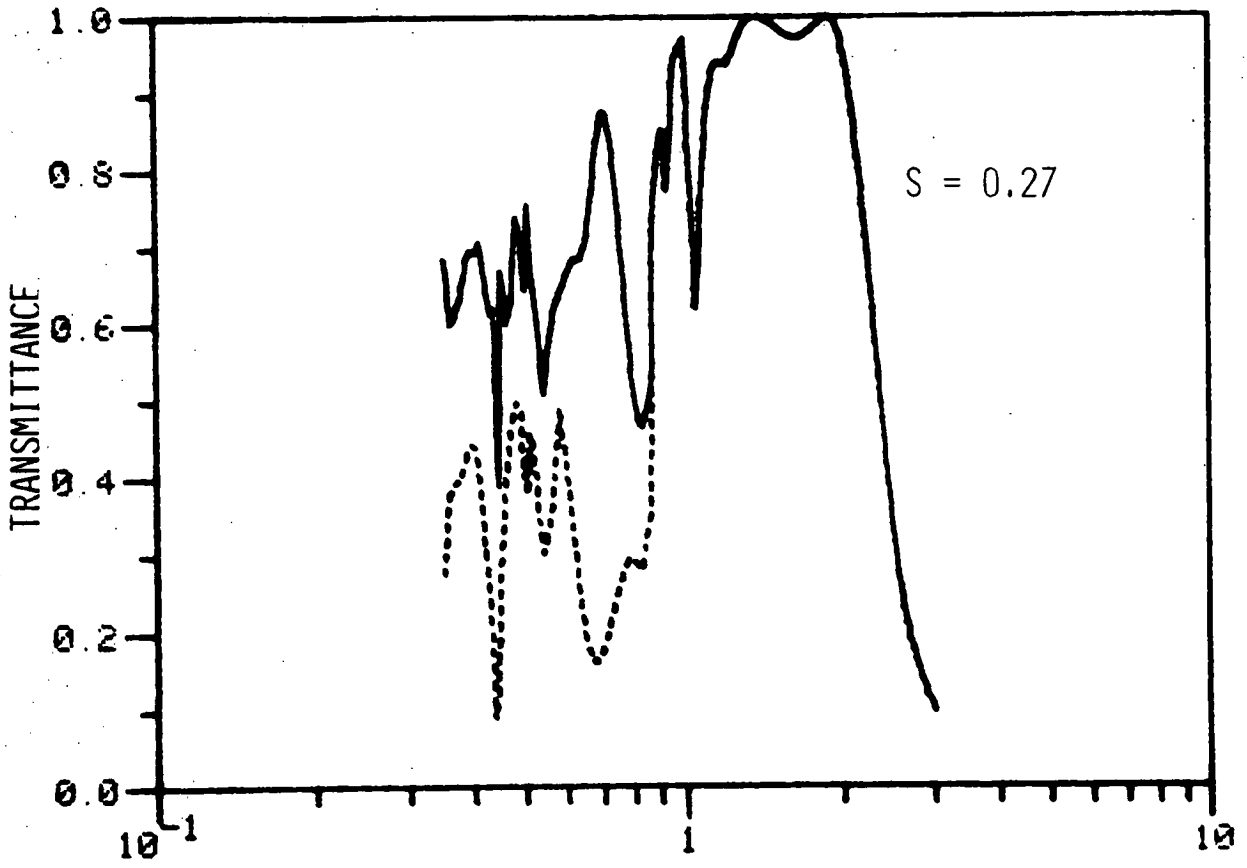


Figure 10.8g

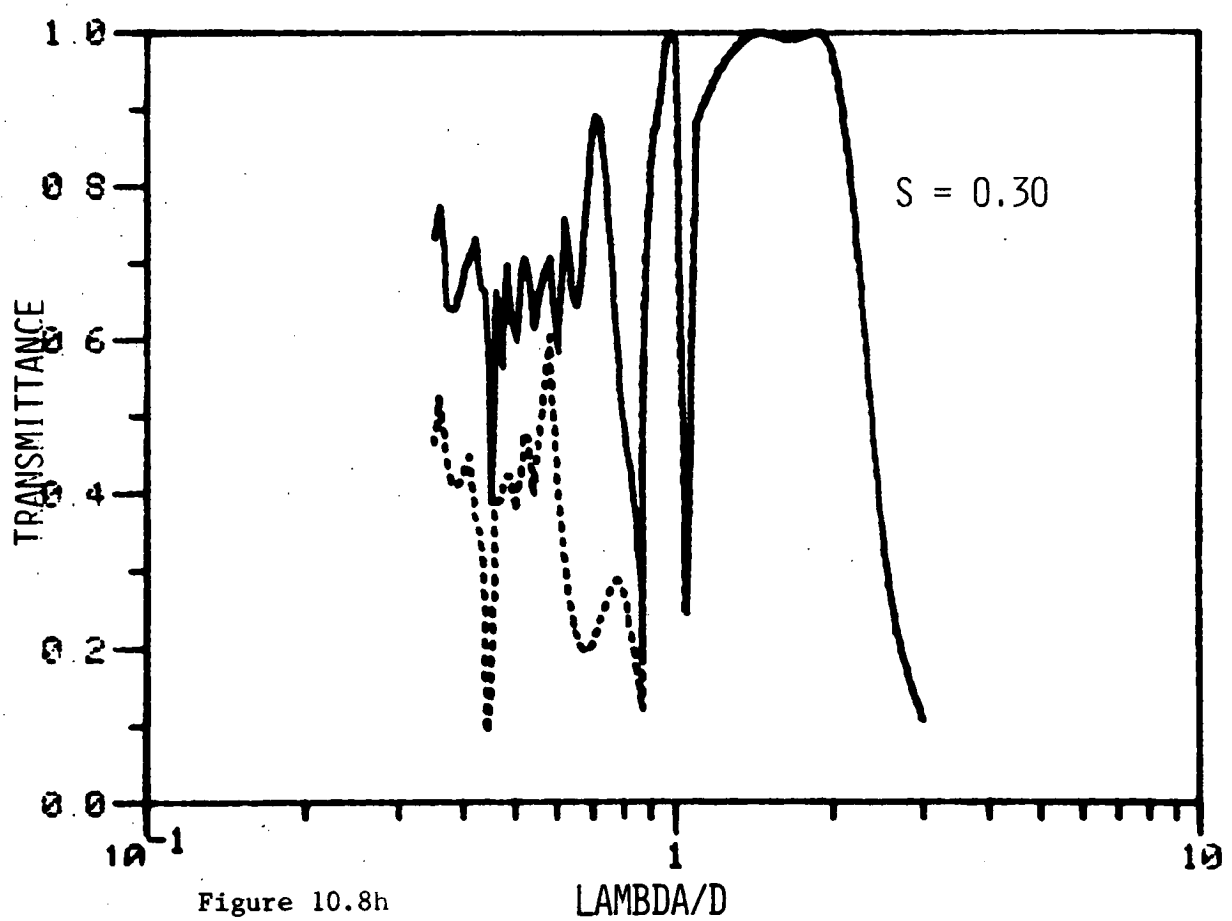


Figure 10.8h

10.5 MODIFICATION TO THE THEORY FOR THE CASE OF RECTANGULAR APERTURES

Following upon the success of the previous study (section (10.4)) it was decided to augment the earlier work of McPhedran and Maystre [10.1, 10.2] and consider a grid of rectangular apertures plugged with a dielectric and surrounded by a symmetric pair of thin films. This theory is entirely similar to that derived in section (10.2) except for the specification of the waveguide modal fields within the apertures. Thus, to avoid needless repetition, the section is concerned only with the derivation of the TE and TM modal fields appropriate to this geometry.

From reference [10.2] the normalized Cartesian modal functions are given by

$$MEX_{nm}(x,z) = \sqrt{\frac{\varepsilon_n \varepsilon_m}{c\ c'}} \cos\left(\frac{n\pi x}{c}\right) \sin\left(\frac{m\pi z}{c'}\right) \qquad (10.5.1)$$

$$MEZ_{nm}(x,z) = \sqrt{\frac{\varepsilon_n \varepsilon_m}{c\ c'}} \sin\left(\frac{n\pi x}{c}\right) \cos\left(\frac{m\pi z}{c'}\right) \qquad (10.5.2)$$

(Here c and c' are the dimensions of the rectangular aperture.)
It is to be noted that on the walls x = 0,c

$$MEZ_{nm} = 0 \quad \text{and} \quad \frac{\partial}{\partial x} MEX_{nm} = 0$$

whereas on the walls z = 0,c'

$$MEX_{nm} = 0 \quad \text{and} \quad \frac{\partial}{\partial z} MEZ_{nm} = 0.$$

Although both of the above partial derivatives of the single modes vanish on their respective walls, this must not be taken to mean that

$$\text{(i)} \qquad \partial E_x/\partial x = 0 \qquad \text{on} \qquad x = 0, c$$

$$\text{or} \quad \text{(ii)} \qquad \partial E_z/\partial z = 0 \qquad \text{on} \qquad z = 0, c'.$$

In either case, the modal expansions for E_x and E_z exhibit divergent behaviour as the walls $x = 0, c$ and $z = 0, c'$ are respectively approached and hence the term by term differentiation of their series cannot be justified.

Now, from these Cartesian expressions for the modes it is possible to define an orthonormal set of TE/TM modes characterizing the field. Let these be defined by

$$\underline{\underline{MTE}}_{nm}(x,z) = A_{nm} \cos\left(\frac{n\pi x}{c}\right) \sin\left(\frac{m\pi z}{c'}\right)\hat{\underline{x}} + B_{nm} \sin\left(\frac{n\pi x}{c}\right) \cos\left(\frac{m\pi z}{c'}\right)\hat{\underline{z}} \qquad (10.5.3)$$

and

$$\underline{\underline{MTM}}_{nm}(x,z) = C_{nm} \cos\left(\frac{n\pi x}{c}\right) \sin\left(\frac{m\pi z}{c'}\right)\hat{\underline{x}} + D_{nm} \sin\left(\frac{n\pi x}{c}\right) \cos\left(\frac{m\pi z}{c'}\right)\hat{\underline{z}} \qquad (10.5.4)$$

$$\text{But,} \quad \text{div}\,(\underline{\underline{MTE}}_{nm}) = 0 \qquad (10.5.5)$$

$$\text{and} \quad [\text{curl}\,(\underline{\underline{MTM}}_{nm})]_y = 0, \qquad (10.5.6)$$

the definitions of TE and TM fields, and so it can be shown that

$$\left.\begin{aligned} \frac{n\pi}{c} A_{nm} + \frac{m\pi}{c'} B_{nm} &= 0 \\ \text{and} \quad \frac{m\pi}{c'} C_{nm} - \frac{n\pi}{c} D_{nm} &= 0 \end{aligned}\right\} \qquad (10.5.7)$$

After appropriate normalization of the modes such that

$$\iint_A \underline{MTE}_{nm} \cdot \underline{MTE}_{NM} \, dA = \delta_{n,N} \, \delta_{m,M}$$

and $$\iint_A \underline{MTM}_{nm} \cdot \underline{MTM}_{NM} \, dA = \delta_{n,N} \, \delta_{m,M}$$

it follows that

$$\underline{MTE}_{nm}(x,z) = g_{nm} \left\{ \frac{m\pi}{c'} \cos\left(\frac{n\pi x}{c}\right) \sin\left(\frac{m\pi z}{c'}\right) \hat{\underline{x}} - \frac{n\pi}{c} \sin\left(\frac{n\pi x}{c}\right) \cos\left(\frac{m\pi z}{c'}\right) \hat{\underline{z}} \right\} \qquad (10.5.8)$$

and

$$\underline{MTM}_{nm}(x,z) = g_{nm} \left\{ \frac{n\pi}{c} \cos\left(\frac{n\pi x}{c}\right) \sin\left(\frac{m\pi z}{c'}\right) \hat{\underline{x}} + \frac{m\pi}{c'} \sin\left(\frac{n\pi x}{c}\right) \cos\left(\frac{m\pi z}{c'}\right) \hat{\underline{z}} \right\} \qquad (10.5.9)$$

where

$$g_{nm} = \sqrt{\frac{\varepsilon_n \, \varepsilon_m}{c \, c'}} \left(\left(\frac{n\pi}{c}\right)^2 + \left(\frac{m\pi}{c'}\right)^2 \right)^{-\frac{1}{2}\ell} . \qquad (10.5.10)$$

From these expressions it is at once clear that the mode $\underline{MTE}_{00}$ is trivial and that the modes of the form $\underline{MTM}_{n0}$ and $\underline{MTM}_{0m}$ are physically insignificant for all values of $n,m \in [0,\infty)$ Note that these modes exhibit none of the degenerate behaviour shown for circular waveguide expansions. In this case, the summation over the parameter ℓ (i.e. horizontally and

vertically polarized modes) is no longer warranted.

Having derived the above modal functions, the TE and TM resolutes of the electric and magnetic fields may be written as:

$$\underline{E}_t^{TE} = \sum_{n,m} [a_{nm} \sin(k_2 \nu_{nm} y) + b_{nm} \cos(k_2 \nu_{nm} y)] \; \underline{MTE}_{nm}(x,z) \qquad (10.5.11)$$

and

$$\underline{E}_t^{TM} = \sum_{n,m} [c_{nm} \sin(k_2 \nu_{nm} y) + d_{nm} \cos(k_2 \nu_{nm} y)] \; \underline{MTM}_{nm}(x,z) \qquad (10.5.12)$$

where the value of ν_{nm} is fixed by the Helmholtz equation -

$$\text{i.e.} \quad k_2 \nu_{nm} = \left(k_2^2 - (\tfrac{n\pi}{c})^2 - (\tfrac{m\pi}{c'})^2\right)^{\frac{1}{2}} . \qquad (10.5.13)$$

$$\text{Now,} \quad \hat{\underline{y}} \times \underline{H}_t^{TE} = - \frac{1}{j\omega \mu_0} [\frac{\partial E_{t,x}^{TE}}{\partial y} \hat{\underline{x}} + \frac{\partial E_{t,z}^{TE}}{\partial y} \hat{\underline{z}}]$$

and thus

$$Z_0(\hat{\underline{y}} \times \underline{H}_t^{TE}) = jr_2 \sum_{n,m} \nu_{nm} [a_{nm} \cos(k_2 \nu_{nm} y) - b_{nm} \sin(k_2 \nu_{nm} y)] \; \underline{MTE}_{nm} . \qquad (10.5.14)$$

For the case of the transverse resolute of the magnetic field, the following relation is used

$$\hat{\underline{y}} \times \underline{H}_t^{TM} = \frac{1}{j\omega \mu_0} [(\frac{\partial E_y^{TM}}{\partial x} - \frac{\partial E_{t,x}^{TM}}{\partial y})\hat{\underline{x}} - (\frac{\partial E_{t,z}^{TM}}{\partial y} - \frac{\partial E_y^{TM}}{\partial z})\hat{\underline{z}}] \; .$$

The field component E_y^{TM} is found by solving the differential equation

$$\mathrm{div}(\underline{E}^{TM}) = 0,$$

from which the solution

$$E_y^{TM} = \sum_{n,m} \frac{1}{k_2 \nu_{nm}} [d_{nm} \sin(k_2 \nu_{nm} y) - c_{nm} \cos(k_2 \nu_{nm} y)] \times \sqrt{\frac{\varepsilon_n \varepsilon_m}{c \times c'}} \times$$

$$\times \{(\frac{n\pi}{c})^2 + (\frac{m\pi}{c'})^2\} \sin(\frac{n\pi x}{c}) \sin(\frac{m\pi z}{c'}) \qquad (10.5.15)$$

is obtained. After some simple manipulation the final result becomes

$$Z_0(\hat{y} \times \underline{H}_t^{TM}) = jr_2 \sum_{n,m} \frac{1}{\nu_{nm}} [c_{nm} \cos(k_2 \nu_{nm} y) - d_{nm} \sin(k_2 \nu_{nm} y)] \underline{MTM}_{nm}. \qquad (10.5.16)$$

Since equations (10.5.11,12,15,16) are identical to those derived in section (10.2), the reduction of the diffraction problem using the method of moments is completely unaltered except for the re-calculation of the inner products IEE, IEM, IME and IMM. This formalism has been implemented numerically and rigorously tested using the same criteria as outlined in section (10.3). In the following section, the results of the computer simulations of the performance of this class of grids are discussed.

10.6 THE EFFECTS OF DIELECTRIC FILMS AND PLUGS ON THE PERFORMANCE OF GRIDS WITH SQUARE APERTURES

A full account of the solar selective properties of bare grids perforated with rectangular apertures has been given by McPhedran and Maystre [10.1, 10.2]. Their study revealed that optimal performance

was obtained from thick meshes having very thin metal walls and used in normally incident radiation. The spectral characteristics of such a grid characterized by the parameters $c/d = 0.9$ and $h/d = 0.4$ is shown in figure (10.9a). From a comparison of this graph and figure (10.4b) for the optimal bare grid having circular apertures, the enhanced transmission bandwidth exhibited by the square holed grid is immediately evident. This results from the more closely packed arrangement of apertures which can be achieved with rectangular holes.

The prominent dip in the transmission curve occurring in the range $0.7 \lesssim \lambda/d \lesssim 1.0$ is the only feature degrading the otherwise excellent performance shown in figure (10.9a). The presence of a similar feature is also apparent for RGRID2 (figure 10.4b). Thus, in view of the success of the study outlined in section (10.4), one was led to speculate as to whether the use of appropriate plugs and films could improve the selectivity of square holed grids. After performing a number of computer simulations it was found that any improvements were only of a marginal nature. In figure (10.9b) is shown the transmission spectrum for a grid of similar geometry to "GRID6" except for the inclusion of a dielectric plug of refractive index 1.5 and a pair of symmetric thin films each of thickness $\frac{1}{2}s/d = 0.12$ and refractive index 1.5. As in the case of the round holed grid, the most noticeable feature is the profusion of anomalous effects (which was discussed in section (10.4)).

Also, note the splitting of the resonance peak caused by the inclusion of a dielectric plug. This splitting tends to broaden the transmission bandwidth and shift the cut-off wavelength further into the

Figure 10.9

Shown in the accompanying graphs are

(a) the spectral characteristics of the optimum bare square-holed grid (GRID6) whose geometry is specified by the parameters $c/d = 0.9$ and $h/d = 0.4$.

and

(b) the performance of the same structure, this time having its apertures filled with a lossless dielectric of refractive index $1.5 + i\ 0.0$ and coated with a symmetric pair of thin films (having the same refractive index) whose total thickness is given by $s/d=0.24$.

Note the general improvement in the transmission bandwidth of the modified structure. However, this broadening, restricted to the transition region ($1.0 \lesssim \lambda/d \lesssim 2.5$), is offset by a degraded performance at shorter wavelengths and thus the overall performance is only marginally improved. This degradation is caused by the presence of additional anomalous features characteristic of the plug and the films. It is interesting to note that the broadening gives rise to optimum performance at a grid period of 0.5 μm instead of 1.0 μm for the "bare" structure. The following graphs are for normally incident radiation ($\phi = 0^{o}$) with $\psi = 90^{o}$ and $\delta = 90^{o}$.

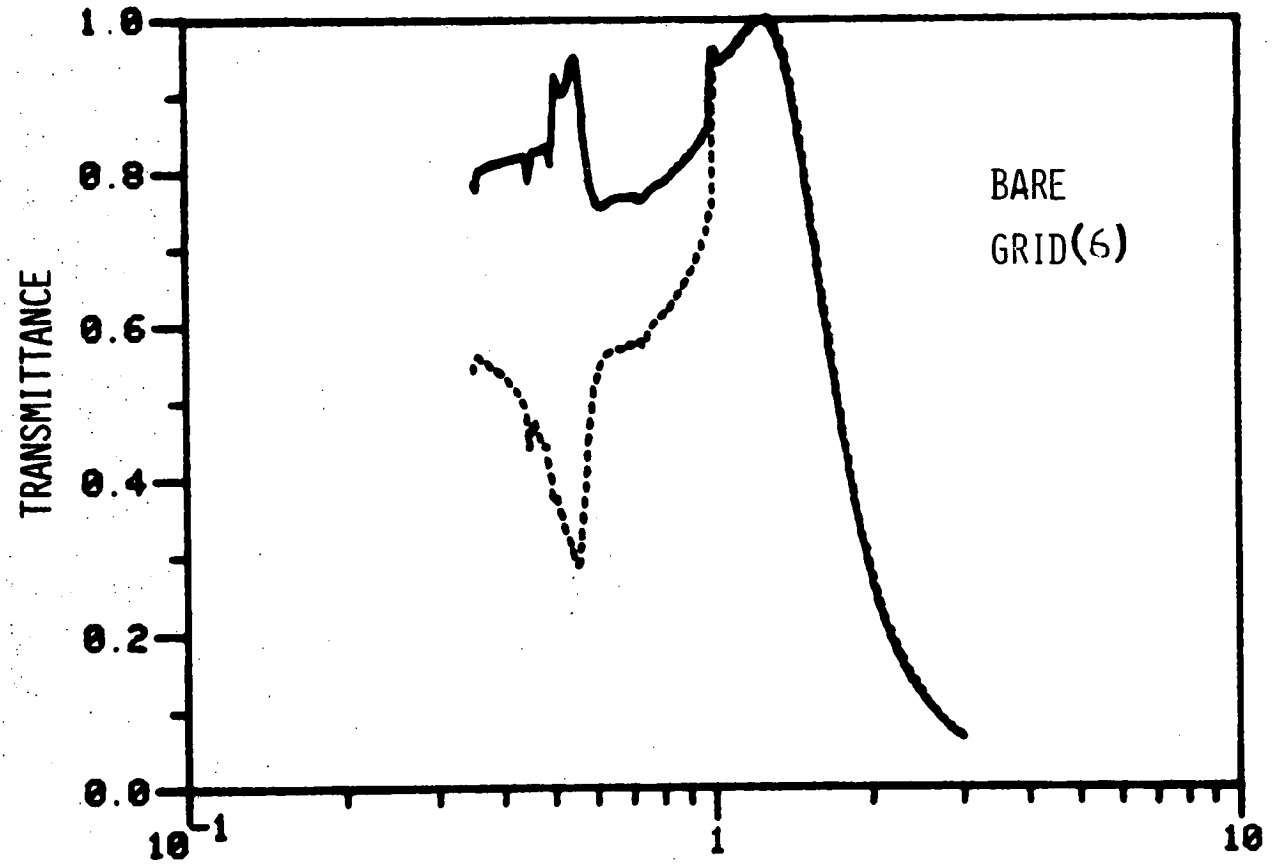


Figure 10.9a

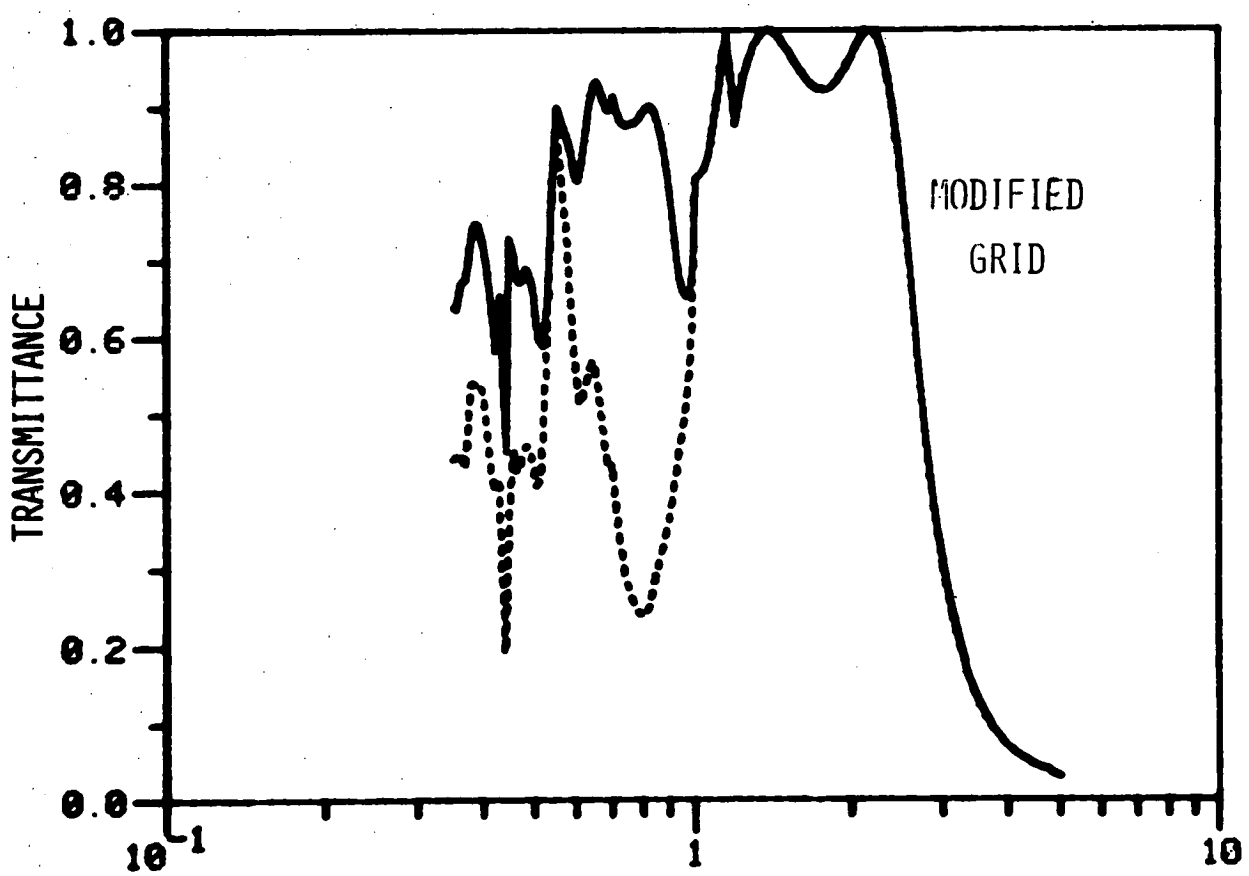


Figure 10.9b

infrared. Thus, to obtain optimal selectivity, the grid period needs to be reduced from 1.0μm (as in the case of figure (10.9a)) to approximately 0.5μm. However, even though the bandwidth has been markedly widened, the presence of the anomalous features, and in particular the deep dip in the total energy curve at $\lambda/d \simeq 0.9$, results in a transmittance of 84% - only a marginal improvement on the value of 81% achieved for "GRID6".

Figures (10.10) show the variation of transmissivity with grid thickness. Note that the resonance at $\lambda/d \simeq 0.9$ exists for all these graphs and is strongest in the case of the thickest grating ($h/d = 0.6$, figure (10.10c)). It appears to weaken with decreasing thickness but is offset by the creation of further anomalous features such as those shown in figure (10.10a). Finally, in figures (10.11) the effect of film thickness is demonstrated. Here, it is apparent that the choice of $s/d = 0.24$ (figure 10.11c) is most appropriate.

In summary, it is disappointing to report that no significant improvement in the solar selective performance of this class of grids can be achieved by incorporating dielectric plugs and films. From the experience gained from this study and that outlined in section (10.4), it is believed that this technique is only appropriate to those geometries whose transmission bandwidth is low. The inclusion of the plug splits and broadens the resonance, whilst the thin film pair is used to rectify the resultant deep resonances created by the former. However, these introduce numerous small anomalies, which, in the case of square holed grids, detract from those gains brought about by increasing the bandwidth. Only in the case of grids having circular

Figure 10.10

The accompanying three graphs demonstrate the effect of grid thickness upon the spectral performance of the modified grid. In each case the geometry of the structure is specified by the parameters c/d = 0.9 and s/d = 0.24. The entire grid is "plugged" and "filmed" with a dielectric material of refractive index 1.5 + i 0.0. The incidence parameters were given by $\phi = 0^{\circ}$, $\psi = 90^{\circ}$ and $\delta = 90^{\circ}$. Graphs for the following normalized thicknesses:

(a) h/d = 0.3,

(b) h/d = 0.4 and

(c) h/d = 0.6

are presented. Note the presence of the insidious resonance centred on $\lambda/d = 0.9$. It is the presence of this feature that is the main reason for the marginal improvement in performance over that obtained using the "bare" grid. Our numerical experiments revealed that it was impossible to "correct" this degrading feature by adjusting only the film thickness. However, it is conceivable that by appropriate choices of film and plug constants this situation could possibly be rectified.

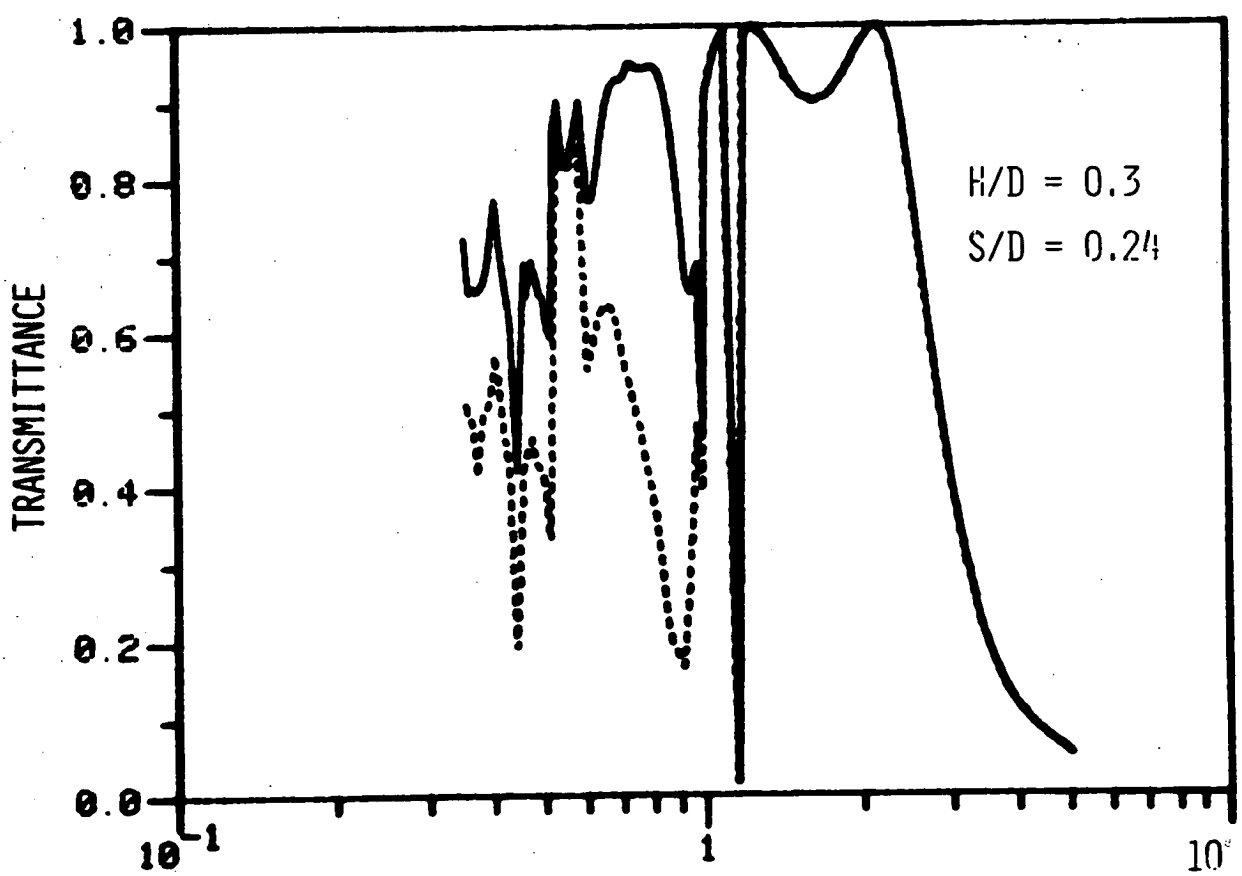


Figure 10.10a

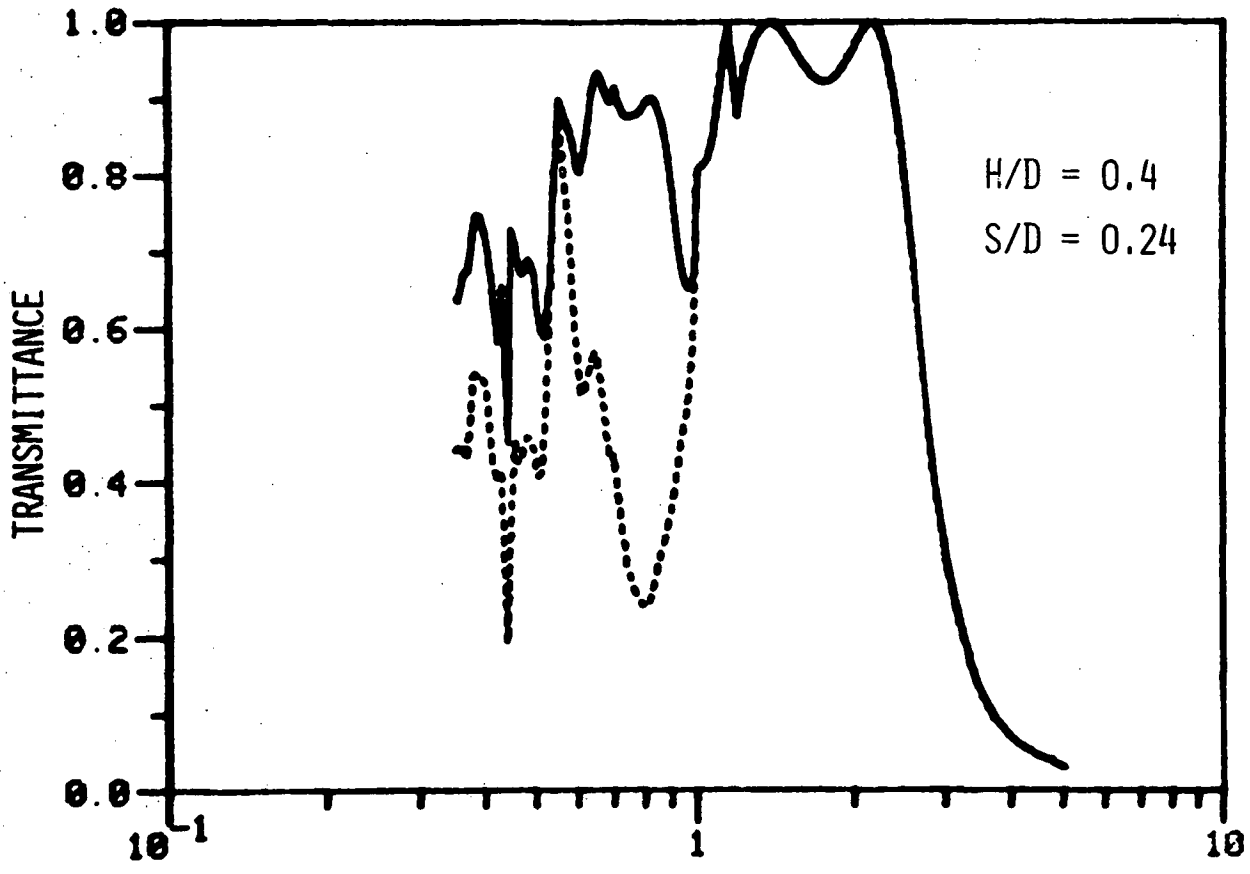


Figure 10.10b

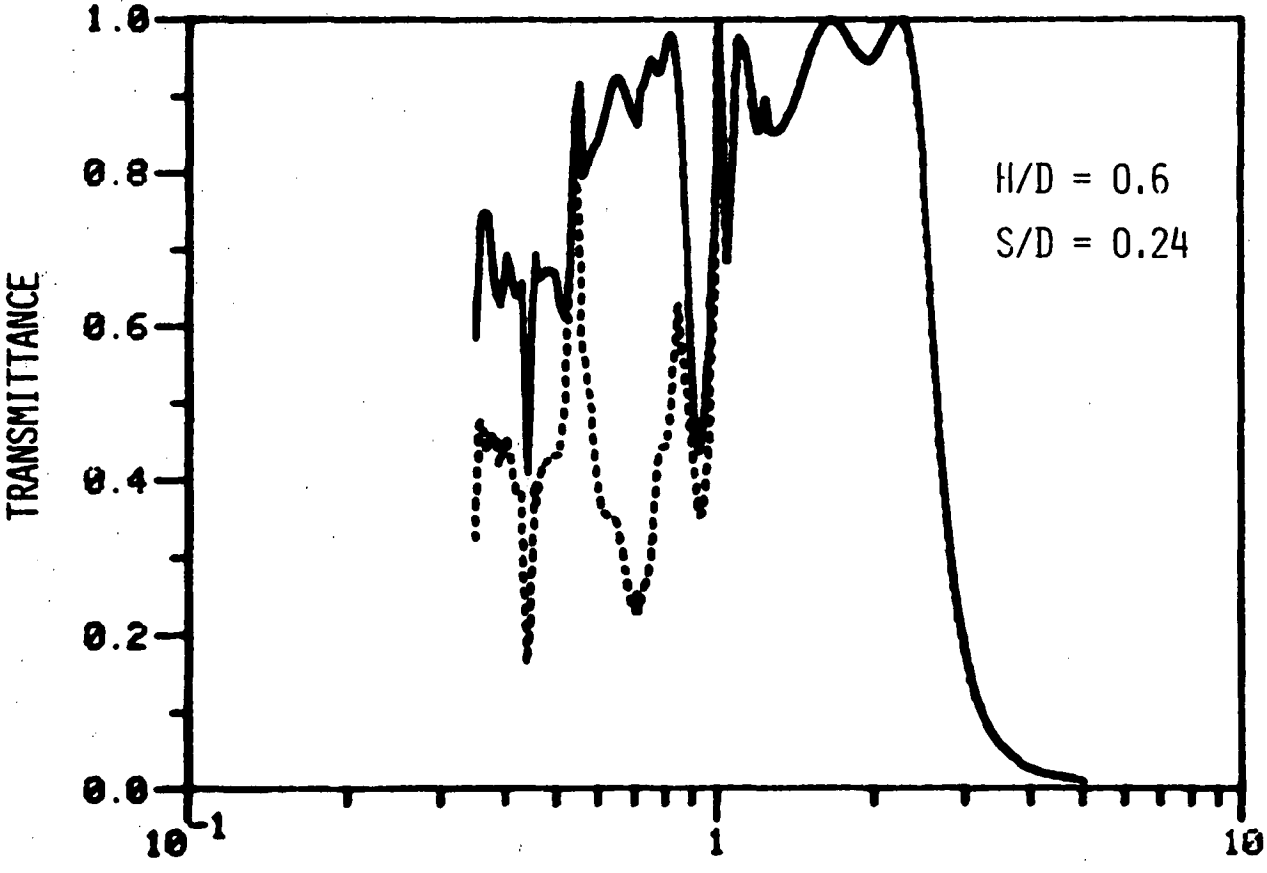


Figure 10.10c

Figure 10.11

The effect of film thickness of the performance of the modified grid, characterized by the parameters c/d = 0.9 and h/d = 0.4, and coated with a dielectric material of refractive index 1.5 + i 0.0. The structure was operated in normally incident radiation as specified by the angles $\phi = 0^{\circ}$, $\psi = 90^{\circ}$ and $\delta = 90^{\circ}$. Graphs for normalized film thicknesses of

(a) s/d = 0.12,

(b) s/d = 0.18 and

(c) s/d= 0.24

are supplied. These graphs show the partial correction of the "split" in the transmission spectrum introduced by the dielectric plug - a feature which successfully broadened the bandwidth.

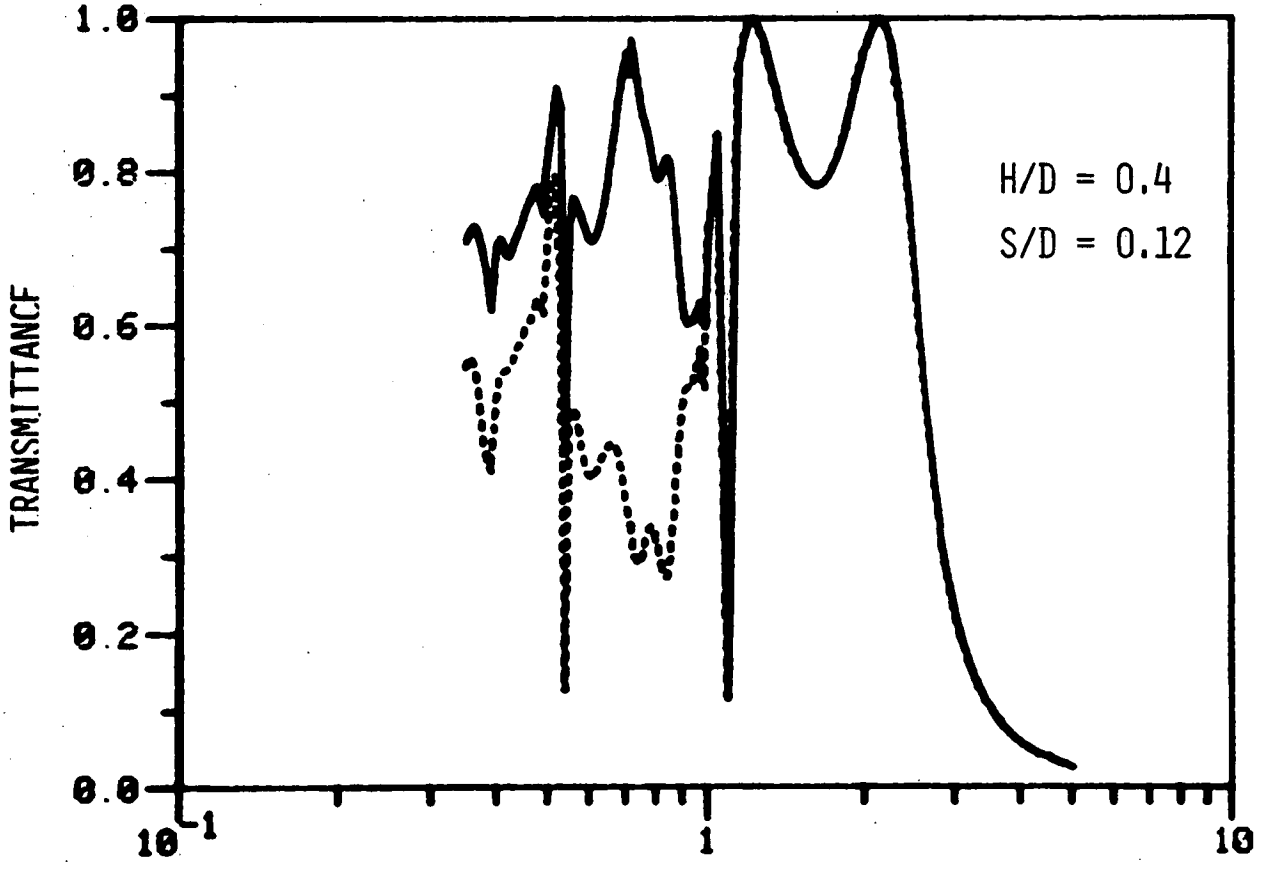


Figure 10.11a

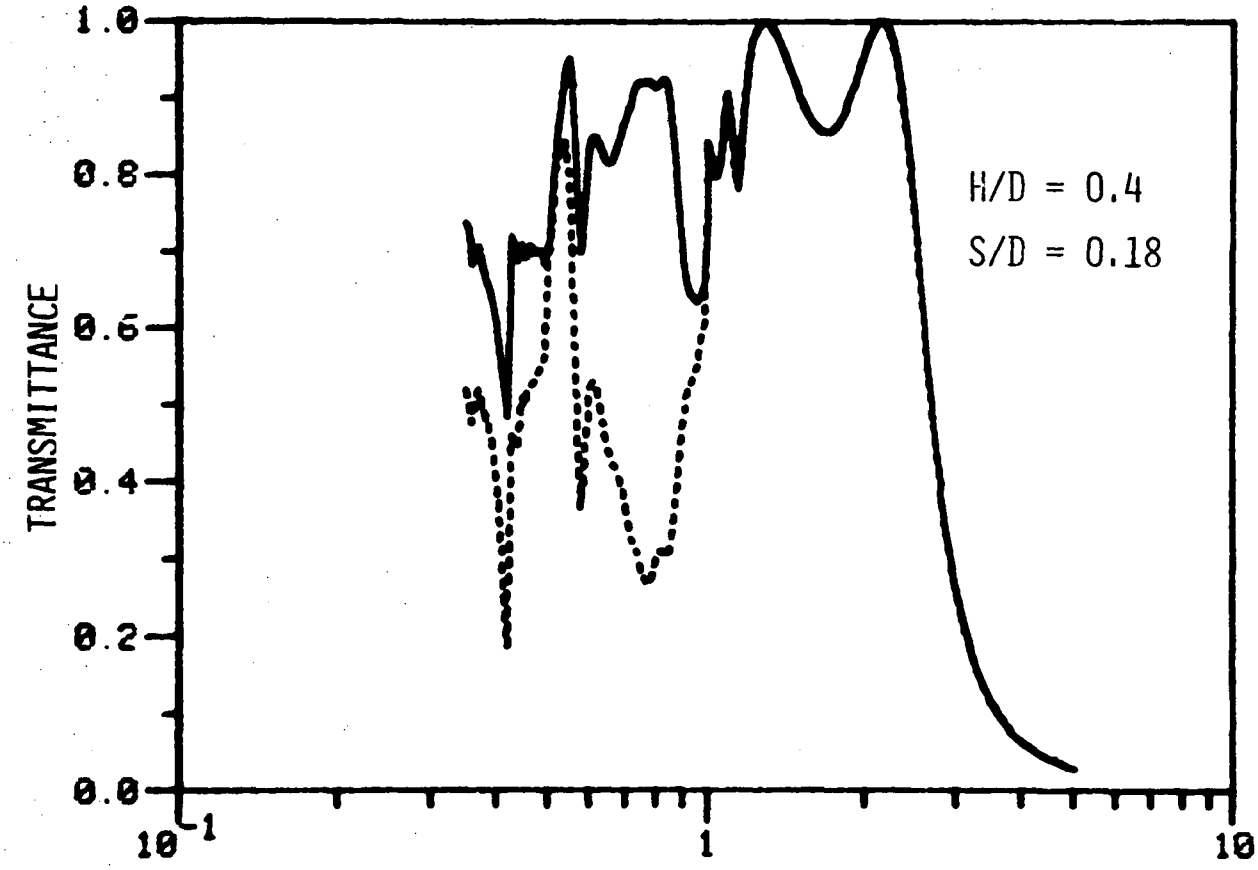


Figure 10.11b

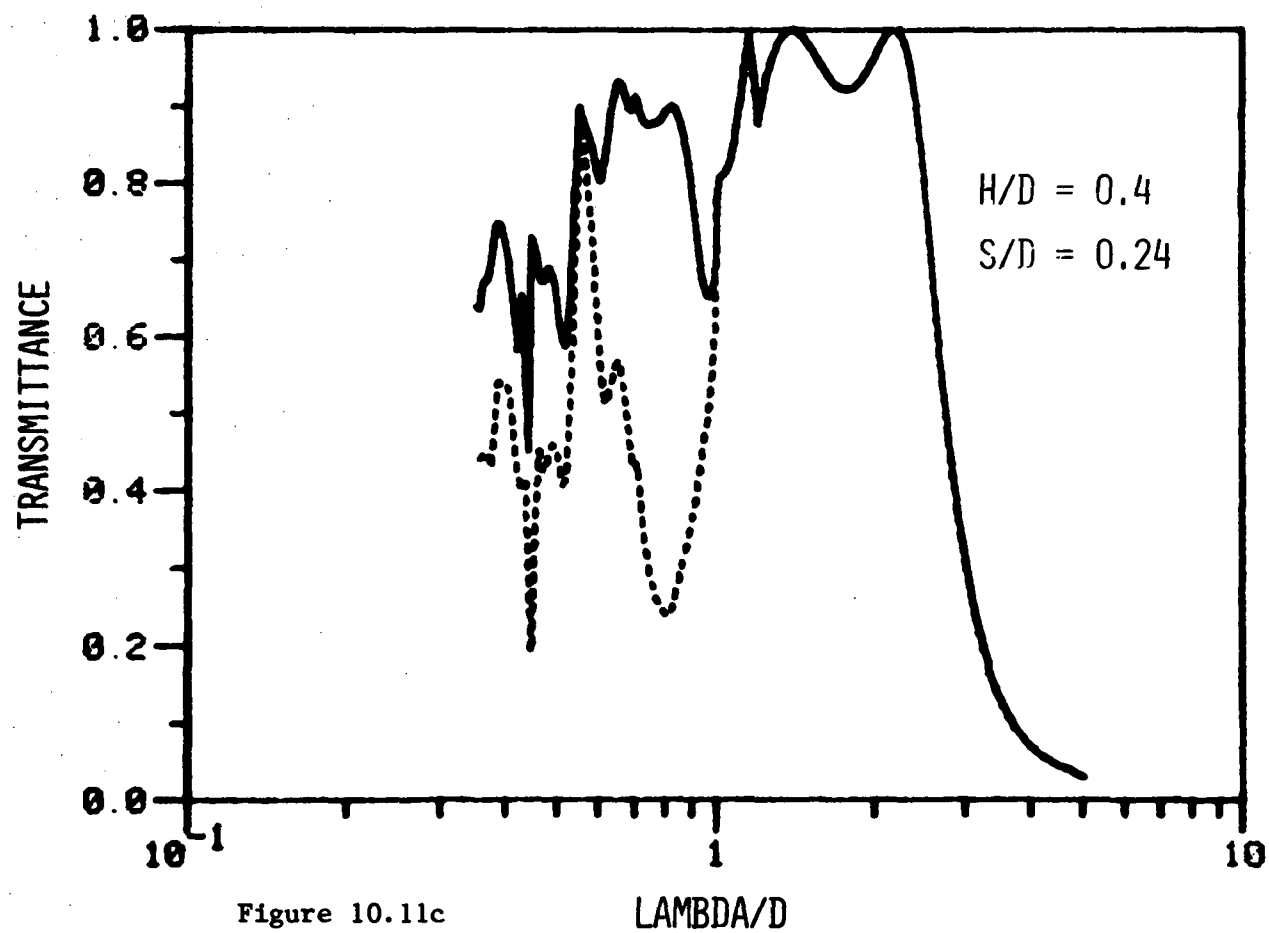


Figure 10.11c

apertures, for which the available hole-area fraction severely limits the transmission bandwidth, can this technique be successfully applied.

10.7 CONCLUSIONS

Two theoretical formulations capable of solving the diffraction problem involving inductive grids having circular and rectangular apertures have been presented. They represent a significant advance on the previous work of Chen [10.6, 10.7] in that they are capable of taking into account the effects of a dielectric sandwich and a dielectric plug on the performance of thick grids. Both theories have been implemented on the Control Data Cyber 72 computer of the University of Sydney and the results obtained have been in excellent agreement with the Reciprocity Theorem and the Littrow phase constraint.

In the case of grids having circular apertures, the solar evaluation has shown that the conclusions derived for grids with square apertures in reference [10.2] are also applicable to them. However the smaller hole-area fractions obtainable with round apertures make them less suited to applications in solar selective systems. It is only by using both plugs and a sandwich film of suitable thickness that the solar absorptance of the optimal grid of round holes can be made equal to that attainable with a bare grid of square holes. Following upon the success of the study for round-holed grids it had been hoped that significant improvements in the spectral response of the square holed grid could be achieved. However the results obtained from this further work failed to reveal any worthwhile enhancements in selectivity and in fact strongly suggested that for this case the upper limit on the

transmittance was closely related to the hole-area fraction (i.e. the geometrical optics limit).

Currently, all diffraction theories for doubly periodic gratings are severely restricted by the assumption that the grid is perfectly conducting. Of course, this is justifiable in the far infrared and microwave sections of the electromagnetic spectrum but is of dubious validity in the visible region. However, imposing such an assumption does permit a modal treatment of the problem, even if only very simple aperture geometries can be accommodated.

At this stage, no attempt has been made to formulate analytic expressions for the modes of finitely conducting structures. Nevertheless, it is anticipated that the difficulties that will be encountered in the development of a finitely conducting grid formalism will be at least "an order of magnitude" greater than those involved in infinite conductivity theories. Considering the complexity and the algebraic inelegance of the current formalisms, it will, initially, be necessary to select a geometry of sufficient simplicity which will readily enable the construction of finite-conductivity modes. It is felt that this structure may well be the round holed inductive grid, and in view of the practical results (concerning the identical performance attainable from both round and square holed grids) there may be no real need to consider more complex geometries for which analytic expressions for the modes cannot easily be derived. However, before launching into the realm of finite conductivity, this choice of aperture geometry could be adequately tested by the formulation of an infinite conductivity capacitive grid theory.

REFERENCES

[10.1] R. C. McPhedran and D. Maystre, (1976), University of Sydney, Report SP76/1.

[10.2] R. C. McPhedran and D. Maystre, Appl. Phys., 14 (1977) 1.

[10.3] R. C. McPhedran and L. C. Botten, (1977), University of Sydney, Report SP77/5.

[10.4] J. L. Adams, L. C. Botten and R. C. McPhedran, accepted for publication by J. Optics (Paris).

[10.5] J. L. Adams and L. C. Botten, (1977), University of Tasmania, Report DGRG77/3.

[10.6] C-C. Chen, I.E.E.E. MTT-19, 19 (1971) 475.

[10.7] C-C. Chen, I.E.E.E. MTT-21, 21 (1973) 1.

[10.8] L. C. Botten, accepted for publication by Optica Acta.

[10.9] C-C. Chen, I.E.E.E. MTT-18, 18 (1970) 627.

[10.10] N. Marcuvitz, "Waveguide Handbook", McGraw-Hill, New York, 1951.

[10.11] N. Amitay and V. Galindo, I.E.E.E. MTT-16, 16 (1968) 265.

[10.12] G. N. Watson, "A Treatise on the Theory of Bessel Functions", Cambridge Univ. Press, Cambridge, 1958.

APPENDIX

The material contained in the appendix solely applies to chapter 10. Contained herein are numerous mathematical derivations necessary to the round-hole grid formalism but whose direct presence in that chapter would detract from the physical arguments developed there.

A.1 INNER PRODUCTS

It is the purpose of this section to elaborate on the evaluation of the inner product terms defined in equations [10.77-80]. These definitions involve the calculations of eight separate inner products. Throughout this appendix, extensive reference is made to the works of Amitay and Galindo [10.11] and of Watson [10.12] so as to circumvent any unnecessary presentation of mathematical identities.

Firstly, consider the quantity

$$IEE^{pq}_{nml} = \iint_A \underline{MTEV}_{nm}(r,\theta) \cdot \overline{\underline{RTE}}_{pq}(y = 0)\, dA$$

$$= \frac{g_{nm}}{\xi_{pq}\sqrt{dd'}} \iint_A \left(\frac{na}{\chi'_{nm} r} J_n\left(\frac{\chi'_{nm} r}{a}\right) \sin(n\theta)\, (\gamma_{pq} \sin\theta - \alpha_p \cos\theta) \right.$$

$$\left. + J'_n\left(\frac{\chi'_{nm} r}{a}\right) \cos(n\theta)\, (\gamma_{pq} \cos\theta + \alpha_p \sin\theta) \right)$$

$$\times \exp\{-j k_0 (\alpha_p x + \gamma_{pq} z)\}\, dA . \qquad A.1.2$$

To remove the (x,z) dependence from the plane wave term, it is necessary to invoke a generating function for Bessel functions. Thus

$$\exp\{-j\, k_0\, (\alpha_p x + \gamma_{pq} z)\} = \sum_{\ell=0}^{\infty} \varepsilon_\ell\, (-j)^\ell\, J_\ell(k_0\, \xi_{pq} r)\, \cos\, \ell\, (\theta - \theta'_{pq}) \tag{A.1.2}$$

where $\tan \theta'_{pq} = \alpha_p/\gamma_{pq}$,

$$\xi_{pq} = \sqrt{\alpha_p^2 + \gamma_{pq}^2}$$

and ε_ℓ is the Neumann symbol.

Now since,

$$\frac{na}{\chi'_{nm} r}\, J_n\, (\frac{\chi'_{nm} r}{a}) = \tfrac{1}{2}\{J_{n-1}\, (\frac{\chi'_{nm} r}{a}) + J_{n+1}\, (\frac{\chi'_{nm} r}{a})\}$$

and $$J'_n\, (\frac{\chi'_{nm} r}{a}) = \tfrac{1}{2}\{J_{n-1}\, (\frac{\chi'_{nm} r}{a}) - J_{n+1}\, (\frac{\chi'_{nm} r}{a})\}$$

expression (A.1.1) reduces to

$$IEE^{pq}_{nm1} = \frac{g_{nm}}{2\sqrt{dd'}} \iint_A [J_{n-1}\, (\frac{\chi'_{nm} r}{a})\, \cos\{(n-1)\, \theta + \theta'_{pq}\}$$

$$- J_{n+1}\, (\frac{\chi'_{nm} r}{a})\, \cos\{(n+1)\, \theta - \theta'_{pq}\}]$$

$$\times \sum_{\ell=0}^{\infty} \varepsilon_\ell\, (-j)^\ell\, J_\ell\, (k_0\, \xi_{pq} r)\, \cos\, \{\ell(\theta - \theta'_{pq})\} dA\,. \tag{A.1.3}$$

The orthogonality of the azimuthally dependent terms enables us to write

$$IEE^{pq}_{nm1} = \frac{\pi g_{nm}}{\sqrt{dd'}} (-j)^{n-1} \cos(n\,\theta'_{pq}) \int_o^a [J_{n-1}(\frac{\chi'_{nm} r}{a})\, J_{n-1}(k_0\,\xi_{pq} r) + J_{n+1}(\frac{\chi'_{nm} r}{a})\, J_{n+1}(k_0\,\xi_{pq} r)]\, r\, dr.$$

After utilizing an identity presented on pages (133-135) of Watson [10.12] and performing some tedious manipulations one arrives at the final result:

$$IEE^{pq}_{nm1} = \sqrt{\frac{4\,\varepsilon_n \pi}{dd'}}\, \frac{(-j)^{n-1}\ a \cos(n\,\theta'_{pq})}{\sqrt{(\chi'^{\,2}_{nm} - n^2)\,\{1 - k_0^2\,\xi_{pq}^{\,2}\,a^2/\chi'^{\,2}_{nm}\}}}\, J'_n(k_0\,\xi_{pq} a). \quad (A.1.4)$$

A similar analysis of the inner product

$$IEE^{pq}_{nm2} = \iint_A \underline{MTEH}_{nm}(r,\theta) \cdot \overline{\underline{RTE}}_{pq}(y = 0)\, dA \quad (A.1.5)$$

yields the rather simple result

$$IEE^{pq}_{nm2} = -\tan(n\,\theta'_{pq})\, IEE^{pq}_{nm1}. \quad (A.1.6)$$

Now, consider the evaluation of

$$IME^{pq}_{nm1} = \iint_A \underline{MTEV}_{nm}(r,\theta) \cdot \overline{\underline{RTM}}_{pq}(y = 0)\, dA$$

$$= \frac{g_{nm}}{2\sqrt{dd'}} \iint_A [J_{n-1}(\frac{\chi'_{nm} r}{a}) \sin\{(n-1)\,\theta + \theta'_{pq}\} + J_{n+1}(\frac{\chi'_{nm} r}{a}) \sin\{(n+1)\,\theta - \theta'_{pq}\}]$$

$$\times \sum_{\ell=0}^{\infty} \varepsilon_\ell\, (-j)^\ell\, J_\ell(k_0\,\xi_{pq} r) \cos\{\ell(\theta - \theta'_{pq})\}\, r\, dr\, d\theta. \quad (A.1.7)$$

By noting that

$$\int_0^{2\pi} \sin\{(n-1)\ \theta + \theta'_{pq}\}\ \cos(\ell\ \theta - \ell\ \theta'_{pq})d\theta \ = \ \frac{2\pi}{\varepsilon_\ell}\ \delta_{\ell,|n-1|} \sin(n\ \theta'_{pq}) \quad (A.1.8)$$

and

$$\int_0^{2\pi} \sin\{(n+1)\ \theta - \theta'_{pq}\}\ \cos(\ell\ \theta - \ell\ \theta'_{pq})d\theta \ = \ \pi\ \delta_{\ell,n+1}\ \sin(n\ \theta'_{pq})$$

it follows that

$$IME^{pq}_{nm1} = \frac{\pi\ g_{nm}\ (-j)^{n-1}}{\sqrt{dd'}} \sin(n\ \theta'_{pq}) \int_0^a [J_{n-1}\ (\frac{\chi'_{nm} r}{a}) J_{n-1}(k_0\ \xi_{pq} r)$$

$$- \ J_{n+1}\ (\frac{\chi'_{nm} r}{a})\ J_{n+1}(k_0\ \xi_{pq} r)]\ r\ dr\ . \quad (A.1.9)$$

By now using a result presented in Watson [10.12] expression (A.1.9) reduces to

$$IME^{pq}_{nm1} = \sqrt{\frac{4\varepsilon_n\ \pi}{dd'}}\ \frac{n \sin(n\ \theta'_{pq})\ (-j)^{n-1}}{\sqrt{(\chi'^{\,2}_{nm} - n^2)}}\ \frac{J_n(k_0\ \xi_{pq} a)}{k_0\ \xi_{pq}}\ . \quad (A.1.10)$$

Similarly, it can be shown that

$$IME^{pq}_{nm2} = \iint_A \underline{MTEH}_{nm}(r,\theta)\ .\ \overline{\underline{RTM}}_{pq}(y = 0)\ dA \ = \ \cot(n\ \theta'_{pq})\ IME^{pq}_{nm1}. \quad (A.1.11)$$

The evaluation of the final two inner product pairs follows in a relatively easy manner and thus only the results are presented.

$$IEM^{pq}_{nm1} = 0 \qquad (A.1.12)$$

$$IEM^{pq}_{nm2} = 0 \qquad (A.1.13)$$

$$IMM^{pq}_{nm1} = \sqrt{\frac{4\pi\varepsilon_n}{dd'}} \frac{\cos(n\theta'_{pq})(-j)^{n-1}}{\{k_0^2 \xi_{pq}^2 - \chi_{nm}^2/a^2\}} k_0 \xi_{pq} J_n(k_0 \xi_{pq} a) \qquad (A.1.14)$$

and

$$IMM^{pq}_{nm2} = \tan(n\theta'_{pq})\, IMM^{pq}_{nm1} . \qquad (A.1.15)$$

This completes the discussion of the inner products.

A.2 CARTESIAN - TE/TM CONVERSIONS

In section (10.2) expressions for the TE and TM basis vector functions $\underline{RTE}_{pq}$ and $\underline{RTM}_{pq}$ were given.

viz. $$\underline{RTE}_{pq} = \frac{1}{\sqrt{dd'}\,\xi_{pq}} (\gamma_{pq}\hat{\underline{x}} - \alpha_p\hat{\underline{z}})\, R_{pq} \qquad (A.2.1)$$

and $$\underline{RTM}_{pq} = \frac{1}{\sqrt{dd'}\,\xi_{pq}} (\alpha_p\hat{\underline{x}} + \gamma_{pq}\hat{\underline{z}})\, R_{pq} \qquad (A.2.2)$$

(together with similar expressions for the transmitted field basis functions). It then follows from the equivalence of the expressions for the transverse resolute of the electric field that

$$(\underline{E}_t)_{pq} = (B_{pq},\ C_{pq})\ R_{pq}$$

$$= E_{pq}\ \underline{RTE}_{pq} + F_{pq}\ \underline{RTM}_{pq} \qquad \text{(A.2.3)}$$

and hence $$B_{pq} = \frac{1}{\sqrt{dd'}} \frac{(\gamma_{pq}\ E_{pq} + \alpha_p\ F_{pq})}{\xi_{pq}} \qquad \text{(A.2.4)}$$

$$C_{pq} = \frac{1}{\sqrt{dd'}} \frac{(\gamma_{pq}\ F_{pq} - \alpha_p\ E_{pq})}{\xi_{pq}} \qquad \text{(A.2.5)}$$

Similarly, $$\hat{B}_{pq} = \frac{1}{\sqrt{dd'}} \frac{(\gamma_{pq}\ \hat{E}_{pq} + \alpha_p\ \hat{F}_{pq})}{\xi_{pq}} \qquad \text{(A.2.6)}$$

$$\hat{C}_{pq} = \frac{1}{\sqrt{dd'}} \frac{(\gamma_{pq}\ \hat{F}_{pq} - \alpha_p\ \hat{E}_{pq})}{\xi_{pq}} . \qquad \text{(A.2.7)}$$

Naturally, there exist similar expressions for the incident fields.

Here, $$(\underline{E}^i_t) = (A_x,\ A_z)\ \hat{R}_{00}$$

$$= E_i\ \underline{\hat{RTE}}_{00} + F_i\ \underline{\hat{RTM}}_{00} \qquad \text{(A.2.8)}$$

from which it follows that

$$E_i = \sqrt{dd'}\ \frac{(\gamma\ A_x - \alpha\ A_z)}{\xi_{00}} \qquad \text{(A.2.9)}$$

$$F_i = \sqrt{dd'}\ \frac{(\alpha\ A_x + \gamma\ A_z)}{\xi_{00}} . \qquad \text{(A.2.10)}$$

A.3 ORTHOGONALITY OF THE WAVEGUIDE MODES

It is the intention of this appendix to clarify the total orthogonality of the modal set. The discussion commences by redefining the four modal basis sets.

$$\underline{\underline{MTEV}}_{nm}(r,\theta) = g_{nm}\left\{\frac{na}{\chi'_{nm}r} J_n\left(\frac{\chi'_{nm}r}{a}\right) \sin n\theta\ \hat{\underline{r}} + J'_n\left(\frac{\chi'_{nm}r}{a}\right) \cos n\theta\ \hat{\theta}\right\} \quad (A.3.1)$$

$$\underline{\underline{MTMV}}_{nm}(r,\theta) = h_{nm}\left\{J'_n\left(\frac{\chi_{nm}r}{a}\right) \sin n\theta\ \hat{\underline{r}} + \frac{na}{\chi_{nm}r} J_n\left(\frac{\chi_{nm}r}{a}\right) \cos n\theta\ \hat{\theta}\right\} \quad (A.3.2)$$

$$\underline{\underline{MTEH}}_{nm}(r,\theta) = g_{nm}\left\{\frac{na}{\chi'_{nm}r} J_n\left(\frac{\chi'_{nm}r}{a}\right) \cos n\theta\ \hat{\underline{r}} - J'_n\left(\frac{\chi'_{nm}r}{a}\right) \sin n\theta\ \hat{\underline{\theta}}\right\} \quad (A.3.3)$$

$$\underline{\underline{MTMH}}_{nm}(r,\theta) = h_{nm}\left\{J'_n\left(\frac{\chi_{nm}r}{a}\right) \cos n\theta\ \hat{\underline{r}} - \frac{na}{\chi_{nm}r} J_n\left(\frac{\chi_{nm}r}{a}\right) \sin n\theta\ \hat{\underline{\theta}}\right\}. \quad (A.3.4)$$

It is immediately clear that the $\{\underline{\underline{MTEV}}_{nm}\}$ and $\{\underline{\underline{MTEH}}_{nm}\}$ are orthonormal sets (by their definitions). Similar remarks apply to the vertical and horizontal resolutes of the TM modal functions.

Now let us prove that the $\{\underline{\underline{MTMV}}_{nm}(r,\theta)\}$ are orthogonal to the other sets of modal functions. It is immediately clear that they are orthogonal to both $\{\underline{\underline{MTEH}}_{nm}(r,\theta)\}$ and $\{\underline{\underline{MTMH}}_{nm}(r,\theta)\}$since

$$\int_0^{2\pi} \sin(n\theta)\ \cos(m\theta)\ d\theta = 0\ \forall n,m.$$

Thus the only remaining question concerns the inner product

$$\int_0^a \int_0^{2\pi} \underline{\underline{MTMV}}_{nm}(r,\theta) \cdot \underline{\underline{MTEV}}_{NM}(r,\theta)\, r\, dr\, d\theta \qquad (A.3.5)$$

$$= h_{nm}\, g_{NM} \left[\frac{Na}{\chi'_{NM}} \int_0^a J'_n\left(\frac{\chi_{nm} r}{a}\right) J_N\left(\frac{\chi'_{NM} r}{a}\right) dr \int_0^{2\pi} \sin(n\theta)\, \sin(N\theta)\, d\theta \right.$$

$$\left. + \frac{na}{\chi_{nm}} \int_0^a J_n\left(\frac{\chi_{nm} r}{a}\right) J'_N\left(\frac{\chi'_{NM} r}{a}\right) dr \int_0^{2\pi} \cos(n\theta)\, \cos(N\theta)\, d\theta \right]$$

$$= -\pi\, \delta_{n,N}\, h_{nm}\, g_{nM} \left[\frac{na^2}{\chi_{nm}\, \chi'_{NM}} J_n\left(\frac{\chi_{nm} r}{a}\right) J_n\left(\frac{\chi'_{nM} r}{a}\right) \Big|_0^a \right.$$

$$- \int_0^a \frac{na}{\chi_{nm}} J_n\left(\frac{\chi_{nm} r}{a}\right) J'_n\left(\frac{\chi'_{nM} r}{a}\right) dr$$

$$\left. + \int_0^a \frac{na}{\chi_{nm}} J_n\left(\frac{\chi_{nm} r}{a}\right) J'_n\left(\frac{\chi'_{nM} r}{a}\right) dr \right]$$

$$= 0 \qquad \forall\ n,m,\ N,M.$$

In a similar manner it can be shown that

$$\int_A \int \underline{\underline{MTEH}}_{nm}(r,\theta) \cdot \underline{\underline{MTMH}}_{NM}(r,\theta)\, dA = 0 \qquad \forall\ n,m,\ N,M.$$

Thus, the modal fields can be specified by a complete set of linearly independent modal functions comprising the four sets

$$\{\underline{\underline{MTEV}}_{nm}\},\quad \{\underline{\underline{MTEH}}_{nm}\},\quad \{\underline{\underline{MTMV}}_{nm}\} \text{ and } \{\underline{\underline{MTMH}}_{nm}\}.$$

A.4 ENERGY FLUX THROUGH THE APERTURES

In this appendix, it is aimed to derive an expression for the energy carried by each mode within the cylindrical waveguide. The downward flux of energy through the guide is denoted by the symbol EF.

Thus,
$$EF = -\iint_A S_y \, dA \qquad (A.4.1)$$

where
$$S_y = \hat{\underline{y}} \cdot \underline{S}$$
$$= \tfrac{1}{2} \text{ Re } \{\hat{\underline{y}} \cdot (\underline{E} \times \bar{\underline{H}})\}$$
$$= -\tfrac{1}{2} \text{ Real } \{\underline{E}_t \cdot (\hat{\underline{y}} \times \bar{\underline{H}}_t)\} \qquad (A.4.2)$$

Thus

$$EF = \tfrac{1}{2} \text{ Re } \Big[\iint_A \sum_{nm\ell} \Big[\{b_{nm\ell} \cos k_2 \nu'_{nm} y + a_{nm\ell} \sin k_2 \nu'_{nm} y\} \underline{\underline{MTE}}_{nm\ell}$$
$$+ \{d_{nm\ell} \cos k_2 \nu_{nm} y + c_{nm\ell} \sin k_2 \nu_{nm} y\} \underline{\underline{MTM}}_{nm\ell}\Big]$$
$$\cdot -\frac{j r_2}{Z_0} \sum_{nm\ell} \Big[\overline{\nu'_{nm} \{a_{nm\ell} \cos k_2 \nu'_{nm} y - b_{nm\ell} \sin k_2 \nu'_{nm} y\} \underline{\underline{MTE}}_{nm\ell}}$$
$$+ \overline{\frac{1}{\nu_{nm}} \{c_{nm\ell} \cos k_2 \nu_{nm} y - d_{nm\ell} \sin k_2 \nu_{nm} y\} \underline{\underline{MTM}}_{nm\ell}}\Big] dA\Big]. \qquad (A.4.3)$$

The modal orthogonality equations show that in equation (A.4.3) the only terms to make a non-zero contribution arise when the indices in the first and second sums are equal. Furthermore, no cross terms arise between TE and TM modes.

Thus $$EF = \sum_{nm\ell} [EFE_{nm\ell} + EFM_{nm\ell}] \qquad (A.4.4)$$

where $$EFE_{nm\ell} = \frac{r_2}{4Z_0} \frac{\nu'_{nm}}{\sin(k_2 \nu'_{nm} h)} |a^*_{nm\ell}||b^*_{nm\ell}| \sin(\psi^*_b - \psi^*_a) \qquad (A.4.5)$$

and $$EFM_{nm\ell} = \frac{r_2}{4Z_0} \frac{1}{\nu_{nm} \sin(k_2 \nu_{nm} h)} |c^*_{nm\ell}||d^*_{nm\ell}| \sin(\psi^*_d - \psi^*_c) \qquad (A.4.6)$$

In the last two equations

$$\psi^*_g = \arg(g^*_{nm\ell})$$

for $g_{nm\ell}$ representing any of the $\{a_{nm\ell}\}$, $\{b_{nm\ell}\}$, $\{c_{nm\ell}\}$ or $\{d_{nm\ell}\}$.

The flux of the incident field is $\frac{1}{2Z_0}$ dd' β.

The normalized modal energy fluxes are obtained by dividing this quantity into expressions (A.4.5) and (A.4.6).