1974

Double reflection dips - from grating ruled semiconductors

Le-Fu Teng

Department: Physics

DOUBLE REFLECTION DIPS - FROM GRATING RULED SEMICONDUCTORS

by

Le Fu Teng, 1934-

A DISSERTATION

Presented to the Faculty of the Graduate School of the

UNIVERSITY OF MISSOURI-ROLLA

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

PHYSICS

1974

Ralph W. Alexander Jr.
Advisor

Leonard L. Levenson

Harold L. Tyler

W. J. James

Robert J. Bell

ABSTRACT

The double dips, previously observed by Fischer, Anderson, et al., which appeared in the reflection spectra of grating surfaces on the Te-doped semiconductors GaAs and InSb around the plasmon and phonon frequencies have been measured in more detail. In the plasmon region, several possible explanations of the phenomenon are discussed. A simple equation checked with the rigorous theory is proposed and is shown to fit the data well. In the phonon region, the extra dip is calculated and identified as being caused by surface phonons.

ACKNOWLEDGEMENT

I am very indebted to Drs. Alexander and Bell, my advisors, for their suggestion of the subject, constant and patient direction, correction, encouragement and support. Without these, it would have been impossible for me to finish any of this work.

I am especially appreciative to Dr. G. Kovenor. He always tried patiently to listen, understand, and answer any of my questions. Without his anxious and intelligent help, I could have done nothing at all.

I am also indebted to Dr. K. Mayhan, for permitting me to use his expensive instrument, the Beckman IR-12 spectrometer; and to Dr. B. Fischer, for lending me the GaAs and InSb samples

Thanks are also given to all of my committee members, for their kindly support and advice; to Ms. Carolyn Ward, for her help with the computer programming; to Mr. J. K. Wong, for his assistance in drawing the figures; to Mr. J. Smart and to all the people in our IR group, for their unforgettable friendship.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (cont.)

LIST OF TABLES

I. INTRODUCTION

In the derivation of the reflectivity of light from solids, Maxwell's equations show that the reflection minima occur very close to those frequencies making the real part of the dielectric function equal to 1. Usually for single excitation modes, these reflection minima are also very close to the resonance frequency, near which the real part of the dielectric function equals zero. The Te-doped GaAs and InSb semiconductors which are being considered here in the far infrared region have only one important phonon mode plus one plasmon mode. When the TO-phonon frequency, ω_{TO} (or the LO-phonon frequency, ω_{LO}), is much larger or much less than the plasmon frequency, ω_p, there is no coupling between the two modes. In this case the frequencies of the minima or dips occurring in the reflection curve are considered to be at ω_{LO} and ω_p, or at ω_{TO} and ω_p.

If the surface of the sample is roughened in some manner, a new kind of mode which is called a surface mode[2] can be excited. A thin film guided wave is also classified as a surface mode. Surface modes, and hence thin film guided waves, can be excited in several ways. A grating, the most usual type of periodically regular roughness on a sample, is employed in one of the two most important guided wave coupling methods used in integrated optics[3].

Due to their importance in spectroscopy, gratings have been well studied. Much effort was concentrated on Wood's anomalies of metal gratings. Ritchie[4] was the first to show that Wood's anomalies are indeed due to a surface plasmons excited by a grating effect.

Surface plasmons, surface phonons, and other surface polaritons on doped semiconductors have been investigated by many people using the prism technique.[5] Fischer[6] is the only one so far to use successfully heavily Te-doped InSb and GaAs semiconductors to study surface plasmons by the grating technique. Anderson, et al.[7] used samples of lower concentrations and spark cut gratings to look at the coupling effect of surface plasmons and surface phonons. From the shift of the reflection dips of the bulk plasmon and bulk phonon frequencies, they concluded that coupling of surface phonons and surface plasmons was observed. However, since the shift did not depend on the polarization of the incident light or the incident angle as it should have, there is doubt that Anderson observed the usual kind of surface phonon or surface plasmon. Also after the spark cut layer was etched off with the grating still left on the surface, the reflection result reduces to the bulk case as shown in Fig. 1. Thus we conclude that Anderson's result is not due to the grating effect; rather, it may be due to the damage layer from the spark cutting on the surface.

Both Anderson, et al. and Fischer found unexplained phenomena in their experiments. In Anderson's case, Fig. 1, he found a double dip around the phonon region. A maximum between these two dips is near the phonon frequency. This effect does not depend on the polarization of the incident light or the incident angle. After the spark layer was cut on the surface, only a single dip was seen. There is no difference between the reflection from a smooth surface and the surface with the grating still on it but with the spark cut layer etched off. Anderson, et al. observed the double dip only for

samples of InSb with an impurity concentration of 3.96×10^{17} Te-atoms/cm^3 and 2.6×10^{17} Te-atoms/cm^3. The effect in the latter is weaker. For the third sample he used, $n=1.43 \times 10^{17}$ Te-atoms/cm^3 and only one dip was found.

In Fischer's experiments,[8] double dips occurred in both the phonon and plasmon regions. Contrary to Anderson's result, they strongly depended on the polarization of incident light and the grating constants. He found double dips occurring on samples of GaAs, with an impurity concentration $n=4.2 \times 10^{18}$ Te/cm^3, grating constant d = 10μm; InSb, with $n=1.1 \times 10^{18}$ Te/cm^3, d = 10μm; and GaAs, with $n=4.2 \times 10^{18}$ Te/cm^3, d = 20μm. Only the first one has double dips in the reflectivity in the phonon region. Double dips occurred in the plasmon region for all of the above three samples. Fischer studied more than ten samples, but no special mention was made for the samples other than those listed above. According to Fischer, double dips of the phonon region occur only when the incident light is polarized with the electric field perpendicular to the grooves of the grating, while the double dips of the plasmon region occur when the electric field is parallel to the grooves of the grating. However, after more careful examination, I have found that this statement should be modified.

My work has been to perform more experiments on the double dip problem and to formulate a reasonable explanation for this phenomenon. In the phonon region, it seems that Fischer's data can be explained as the excitation of surface phonons by a grating. This comparatively less complicated work will be described only in the last chapter. I

have done nothing for Anderson's case. The electron concentrations of the InSb samples I used were so high that no phonon experiment could be made. This can be seen from my reflection spectra as well as from my calculated curves - no phonon dips appear. For this reason most of my work concentrated on the double dips around the plasmon region. At first we thought these double dips were due to a kind of plasmon and phonon emission effect. This is a possible mechanism that can convert a reflection minimum into a maximum separating two dips. Different kinds of possible emission were assumed but none explained the experiments. The emission idea will be discussed in detail in a later chapter. I found that a grating on a surface can also produce the double dip around the plasmon region. In this work rigorous grating theory accompanied with a reasonable assumption was used to fit the data. The periodic surface rulings (i.e. the grating) did not alter the reflection spectra in the non resonant frequency ranges. Only around the "plasma edge" did the reflection dip split in two due to the inhomogeneous distribution of the electrons inside the ruled area.

II. BASIC THEORIES

In order to discuss the experimental results, some basic theories about the grating and surface waves are described below.

1. The criterion of a smooth or a rough surface: There is no real, perfectly smooth surface. Optically a surface being considered as smooth or rough depends on the wavelength of the incident light and the problem of interest. The Rayleigh criterion[9] for a smooth surface is defined by the rms roughness height $h < (\lambda/8\cos\theta)$, where λ and θ are the wavelength and the incident angle of the light, respectively. The roughness height h is shown in Fig. 2(a). Fig. 2(b) is for a grating case. If the groove area is small compared to other portions of the surface, a grating may be considered smooth when the grating constant d is much larger or smaller than the wavelength λ. The usual reflection criterion may not be satisfied in some special reflection cases. If there exists a surface polariton or thin film guided wave coupled with the evanescent wave in the material, a grating surface which is smooth in the usual sense shows its roughness in allowing excitation of the evanescent wave.
2. Conservation of momentum amplitude for the light scattered from a grating: The momentum amplitude of the light is conserved no matter at what angle it is scattered from a medium. This can be easily verified from the conservation of energy and momentum of the whole system. Since the recoil velocity of the medium is so much smaller than the

velocity of a photon, it turns out that the momentum amplitude of the light is not changed.

3. The momentum component parallel to the grating surface of the scattered light: Letting x be the direction along the grating surface, then the x-component of the momentum of the scattered light is $k_x = k_o \sin\theta + 2n\pi/d$, where $n=0,\pm1,\pm2\cdots$, k_o is the wave vector (also momentum) of the incident light, and d is the grating constant. When light with momentum of x-component $k_o \sin\theta$ is incident on a single groove element of a grating, the element scatters this incident light into many different modes. After superposition of the contributions from all individual elements, most modes will interfere destructively. Only those modes with x-momentum equaling $k_o \sin\theta + 2n\pi/d$ add together constructively by matching the phase and hence have significant amplitude. The relation $k_x = k_o \sin\theta + 2n\pi/d$ can also be easily obtained by multiplying both sides of the familiar grating formula $d(\sin\theta^{(d)} - \sin\theta) = n\lambda$ by $2\pi/\lambda d$ and noting that $k_o = 2\pi/\lambda$, $k_o \sin\theta^{(d)} = k_x$, where $\theta^{(d)}$ is the angle at which the diffracted light is scattered.

4. Evanescent waves produced by a grating: Since the momentum amplitude of the scattering wave is unchanged, the z-component of the momentum (perpendicular to the grating surface) is $k_z = \sqrt{k_o^2 - k_x^2}$. It is possible that for sufficiently large absolute values of n such that $n > (k_o d/2\pi)(1-\sin\theta)$ or $n < -(k_o d/2\pi)(1+\sin\theta)$, $|\sin\theta^{(d)}|$ will be larger than 1 and

$k_x^2 > k_o^2$. In this case, k_z is imaginary, indicating that an evanescent wave has been produced by the diffraction. It is this property that makes the grating a useful tool to excite a surface or a thin film guided wave.[10] If no suitable surface polariton or thin film guided wave exists in a medium, the evanescent wave will not take any energy, and the reflection of light will appear as from a smooth surface provided that $\lambda >> d$.

5. Band effect of the grating: Ritchie[11] has shown that there exist band gaps for the surface wave at $k_s = n\pi/d$ if this wave is excited by a grating. Ritchie,[12] Cowan,[13] et al. also observed the gaps in their surface plasmon experiments using aluminum-coated gratings. There was no prediction about the gap width. The band gap is analogous to the band effect of solids in one dimension, or it is just the Bragg condition in the X-ray diffraction theorem. A wave cannot propagate through a periodic medium when the Bragg condition is satisfied, because under this condition, a standing wave is formed inside the periodic region and no net energy is transmitted. This fact has lead to the grating structure inside a medium being used as a band pass filter in integrated optics.[3]

6. Non-existence of the TE surface modes: Surface waves propagate along the interface of two different media, but with the amplitude decreasing exponentially perpendicular to the interface into both media. If both media are

non-magnetic, there will be no TE surface modes. To see this, let the interface be the z=0 plane and let one of the media be air occupying the $z > 0$ space, with the other medium occupying the $z < 0$ space. The TE mode surface waves are of the form

$$B = (B_x\hat{x} + B_z\hat{z})e^{ik_x x - \alpha z} \qquad \text{(in air)}$$

$$B' = (B'_x\hat{x} + B'_z\hat{z})e^{ik'_x x + \alpha' z} \qquad \text{(in medium)}$$

It is clear that α and α' must be both real and positive for surface waves. The boundary conditions for H_x and B_z continuous at z=0 yield $B_x = B'_x/\mu$, $k_x = k'_x$ and $B_z = B'_z$. Inserting all these into the expressions of B and B', substituting them into the equations $\nabla \cdot B = 0$ and $\nabla \cdot B' = 0$, one finally gets the relation between α and α' as $\alpha/\alpha' = -\mu$, where μ is the permeability of the medium.

Now, for non-magnetic active material $\mu=1$ and then $\alpha = -\alpha'$. This obviously contradicts the assumption for surface waves. Also if α and α' are not imaginary, the wave amplitude will become infinite on either side of the interface as $|z| \to \infty$. For TM surface mode, the argument is just the same except $B \to D$, $H \to E$, $\mu \to \varepsilon$ etc. then $\alpha/\alpha' = -\varepsilon$. Since the effective dielectric function ε can be negative, a surface wave can propagate along the surface.

7. Dispersion relations for non-radiative surface waves: The dispersion relation of a surface wave has been derived by

many people.[14] For a non-radiative mode,[15] it is $k_s = \frac{\omega}{c}\sqrt{\frac{\varepsilon}{1+\varepsilon}}$, and hence on the air side the wave vector component perpendicular to the surface is $k_z = \sqrt{(\frac{\omega}{c})^2 - k_s^2} = \frac{\omega}{c}\sqrt{\frac{1}{1+\varepsilon}}$. The parameter k_z must be imaginary, and therefore $\varepsilon \leq -1$ if a surface wave is to exist. Then, $k_s > (\omega/c)$ means that the surface wave momentum is always larger than the momentum of a photon. Plotting the relation between ω and k_s, it is shown that at small value of ω, $\varepsilon << -1$ and $k_s \approx \omega/c$. As ω reaches the asymptotic value ω_s, $\varepsilon = -1$ and $k_s \to \infty$. A photon in air cannot match the momentum of the surface wave, and hence no coupling is possible between them. The prism or the grating technique basically serves to produce an evanescent wave of the same form as the surface wave, with momentum greater than a free photon, so that coupling to the surface wave is possible. Bell, et al.[16] have shown that for complex ε the form $k_s = \frac{\omega}{c}\sqrt{\frac{\varepsilon}{1+\varepsilon}}$ remains unchanged. However, for complex k_s and real ω the real part of k_s will not go to infinity and will have a bend back towards smaller k as ω approaches ω_s. Complex ε also yields complex k_z. In this case, the surface wave must have a somewhat generalized definition.

8. Radiative surface modes: Ritchie[4] and Ferrell[17] have shown that for a sample of finite thickness the non-retarded dispersion relation of surface waves in a metal has the form $\omega_s^{\pm}(k) = (\omega_p/\sqrt{2})(1 \pm e^{-ka})^{1/2}$, where a is the thickness of the metal slab. As $a \to \infty$, $\omega_s^{\pm} = \sqrt{2}\omega_p$, just the non-retarded

limit ($k\to\infty$) of the non-radiative mode in (7). When the phase velocity of the surface wave is not much smaller than the velocity of light, retardation must be considered. The above equation will split into many modes. According to Kliewer and Fuchs,[18] $\omega_s^+ = (\omega_p/\sqrt{2})(1+e^{-ka})^{1/2}$ is called a tangential mode in which the electrons oscillate tangent to the metal surface. It is non-radiative, and its splitting when retardation is included resembles the non-radiative mode described in (7). The mode with frequency $\omega_s^- = (\omega_p/\sqrt{2})\cdot(1-e^{-ka})^{1/2}$ is the one in which the electrons oscillate normally between the two surfaces. This mode splits into many branches when retardation is taken into account, and some of these modes have phase velocity $v > c$. These modes are radiative with complex frequency[19] essentially equal to ω_p.[19] They exist only for small thickness a.

9. Dielectric function and the reflection formula: For an isotropic material with one phonon and one plasmon, the dielectric function (in the long wavelength limit) is given by[20]

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{\delta\varepsilon}{1 - (\omega/\omega_o)^2 - i\omega\Gamma/\omega_o^2} - \frac{4\pi Ne^2}{m^*(\omega^2 + i\omega/\tau)}$$

where ε_∞ is the high frequency dielectric constant; ω_o is considered to be the transverse optical mode frequency ω_{To} when Γ, the damping parameter of the ions, is small; $\delta\varepsilon$ is the strength of the mode, $\omega_p^2 = 4\pi Ne^2/m^*\varepsilon_\infty$; N is the free

carrier concentration with electron charge e and effective mass m^*; and τ is the collision time of this plasmon mode. For the electric vector perpendicular to the plane of incidence, the reflectivity of the light is given by[1]

$$R = [(\cos\theta - a)^2 + b^2]/[(\cos\theta + a)^2 + b^2],$$

where

$$2\begin{Bmatrix} a^2 \\ b^2 \end{Bmatrix} = [(\varepsilon' - \sin^2\theta)^2 + \varepsilon''^2]^{1/2} \pm (\varepsilon' - \sin^2\theta)$$

with $\varepsilon' = \mathrm{Re}[\varepsilon(\omega)]$

and $\varepsilon'' = \mathrm{Im}[\varepsilon(\omega)]$

For InSb, McMahon has achieved an excellent data fit. As for my samples, especially for GaAs, the concentration is much higher, and the slope of the calculated curve at just below the plasmon frequency seems higher than the experimental data.

III. EXPERIMENT

The samples were obtained from Fischer and their physical parameters are listed in Table I. They are n-type Te-doped semiconductors, GaAs and InSb. Gratings were ruled on the optically polished samples by a diamond ruling engine. The grooves are about 1μm deep and wide. The size of the samples is about 1cm in diameter.

Experiments were performed in the wavelength region $350cm^{-1}$-$900cm^{-1}$ using a Beckman IR-12 infrared spectrometer, with a resolution of about $0.25cm^{-1}$. A IGP-225 polarizer (Molectron Corporation) was used. The degree of polarization is 99%, and the transmission is 70% within the region of interest. From the gratings incorporated in the spectrometer, Wood's anomalies appear for the polarized light. Two peaks or two dips appear at the wavenumbers $270cm^{-1}$ and $570cm^{-1}$ (Fig. 3). When the polarized field is parallel to the grating grooves inside the instrument, the anomalies are two peaks at the above wavelengths. Also, the intensity increases slightly as the wavenumber increases. When the polarized field is perpendicular to the grooves of the grating, the behavior is just opposite to the previous case. This behavior was also described by Stewart and Gallaway.[21] It seems that the p- and s-anomalies should compensate each other to produce a flat intensity curve at some direction of the polarized field. In fact, the anomalies do disappear at a certain orientation of the polarizer. But surprisingly, only a small rotation of the polarizer is required to make the transmission spectrum through the polarizer flat. This is very sensitive to the angle of rotation, as can be seen from Fig. 4.

Table I. Samples

Sample		Concentration	Grating Constant
A10	GaAs	4.2×10^{18} Te-atoms/cm^3	10μm
A20	GaAs	4.2×10^{18} Te-atoms/cm^3	20μm
A30	GaAs	4.2×10^{18} Te-atoms/cm^3	30μm
A40	GaAs	4.2×10^{18} Te-atoms/cm^3	40μm
B10	InSb	*1.1×10^{18} Te-atoms/cm^3	10μm
B20	InSb	*1.1×10^{18} Te-atoms/cm^3	20μm
B30	InSb	*1.1×10^{18} Te-atoms/cm^3	30μm
C10	InSb	*4.7×10^{17} Te-atoms/cm^3	10μm
C20	InSb	*4.7×10^{17} Te-atoms/cm^3	20μm
C30	InSb	*4.7×10^{17} Te-atoms/cm^3	30μm

*Concentrations were calculated by fitting the normal "plasmon edge". Values originally marked by Fischer seem too high.

In order to do reflection measurements for the samples, an ATR unit must be placed behind the polarizer. In this case, the orientation of the polarizer must be changed slightly to obtain a flat background for different incident angles. Therefore, the experiments were performed by first using a mirror at the sample position, slightly turning the orientation angle of the polarizer until a flat background was obtained, and then replacing the mirror with the sample. Since the direction of the polarized field is only slightly different from the normal position, the sample holder needs only a little modification.

The angle of the light incident on the sample mounted on the ATR unit can be varied from 15° to larger than 70°. However, not much data were taken with this regular arrangement, since the size of the samples is so small ($\approx 1 cm^2$) that it was very hard to obtain enough signal. An arrangement something like the Littrow Mount has been used to overcome this difficulty. As shown in Fig. 5, collimated light was focused at the sample by a concave mirror. The light then coming from the concave mirror is thus collimated. As the beam width is small compared to the focal length of the concave mirror, the divergence angle of the beam is small. The incident angle in this arrangement is limited by the size of the concave mirror to less than 10°.

Double beam spectra were taken. A comb screen in the reference arm was used to adjust the light to a convenient intensity, hence the reflectivity showing in the spectra is of an arbitrary scale. The experimental purpose is to observe the anomalous double dips with

different semiconductors, gratings, free carrier concentrations, incident angles, field polarizations, relative orientations among the field, incident plane, and grating grooves, higher order diffractions, band pass incident waves, etc.

A summary of the data obtained is listed as Table II.

Among all samples, A10 shows the double dip effect most strongly. Fig. 6 is a comparison of the reflectance of the electric field perpendicular [curve (a)] and parallel [curve (b)] to the incident plane for sample A10. Also shown in the figure is the reflectance from the smooth surface [curve (c)] which resembles the E parallel case. Fig. 7 makes the same comparisons for samples B10 and C10. Both samples are sensitive to the polarization of the incident field. The higher reflectivity for the E parallel field is partly due to the instrument. Fig. 8 shows the behavior with different grating constants (d = 10μm, 20μm, and 30μm) for the samples labeled A10, A20, and A30. Fig. 9 shows the double dips of the standard reference case [curve (a), same curve from Fig. 6, curve (a)] compared to the case of larger incident angle [curve (b)], different grating groove orientation [curve (c)], and a narrower band of incident wavelengths [curve (d)] obtained by using a band pass filter cutting out the frequency beyond $1000cm^{-1}$ [curve (d')]. It seems that nothing has changed much in any case. The essential feature remains the same except the sizes of the two dips differ slightly. An attempt was made to get the higher order diffraction spectra for d = 20μm, 30μm, and 40μm, but not enough signal was detected. For d = 10μm, the grating constant is smaller than the wavelength in the double

Table II. Observed Reflectivity Minima near the Plasma Frequency. The Frequency of the Maximum between the Minima is also listed.

		$*E_{pp}$ or $\overset{*}{E}_{sp}$			E_{ps}
	Double	Max. & Min. freq. (cm^{-1})			Min. freq.
Sample	Dips?	min.	max.	min.	(cm^{-1})
A10	$\overset{*}{**}$ S	$\overset{*}{**}$560	672	740	$\overset{*}{**}$680
A20	* W	* 640	672	725	-----
A30	$\overset{*}{**}$ N	$\overset{*}{**}$720			$\overset{*}{**}$720
A40	$\overset{*}{*}$ N	$\overset{*}{*}$720			$\overset{*}{*}$720
B10	*S, $\overset{*}{*}$Y	$\overset{*}{**}$620	670	770	$\overset{*}{*}$740
B20	$\overset{*}{*}$ N	$\overset{*}{*}$770			$\overset{*}{*}$750
B30	$\overset{*}{*}$ N	$\overset{*}{*}$760			$\overset{*}{*}$760
C10	$\overset{*}{*}$ W	$\overset{*}{*}$420	480	540	$\overset{*}{*}$450
C20	$\overset{*}{*}$ N	$\overset{*}{*}$450			$\overset{*}{*}$450
C30	$\overset{*}{*}$ N	$\overset{*}{*}$450			$\overset{*}{*}$450

Notes: * - Fischer's data, $\overset{*}{*}$ - current data

S = yes,strong, W = yes,very weak, Y = yes,not strong, N = no

E_{ps} = E parallel to incident plane and E perpendicular to grating grooves.

$*E_{pp}$ = E parallel to incident plane and E parallel to grating grooves. (By Fischer)

$\overset{*}{*}E_{sp}$ = E perpendicular to incident plane and E parallel to grating grooves. (By Teng)

dip region, so no diffracted order exists.

Fischer's experimental procedure differed from mine. He did the experiments in the p-polarized orientation, i.e. he fixed the orientation at which the electric field E was parallel to the incident plane, then rotated the grating sample with respect to the incident field. He found double dips occurred when the grating grooves were rotated parallel to the field but only one dip appeared when they were perpendicular to each other. In these experiments, I rotated the field direction with respect to the incident plane. I always got the double dips when E was perpendicular to the incident plane for all orientations of the grating grooves with respect to the incident plane. When E was parallel to the incident plane, no double dips appeared for the case that the grating grooves were perpendicular to the electric field E. This agrees with Fischer's result. Combining Fischer's and my results, the only common orientation for which double dips did not appear is that with E parallel to the incident plane and perpendicular to the grating grooves. For all the other orientations, double dips were obtained either by Fischer or by me.

IV. DISCUSSION OF THE RESULTS

A summary of the experimental results is as follows:

1. A double dip occurs in some Te-doped semiconductors on both sides of the plasmon frequency ω_p. The frequency of the maximum is located around the middle point between these two dips, and may be larger or smaller than ω_p. (Figs. 6, 7)
2. For an appropriate grating sample, a double dip occurs at all orientations except one of the electric field E, the incident plane, and the grating grooves. The exception for which a double dip does not appear is with E parallel to the incident plane and perpendicular to the grating grooves, which is the case for which a surface wave may be excited. This suggests that a surface wave may have the effect of eliminating the double dips. (Figs. 6, 7 and 9)
3. As a special case of point (2) above, the double dip occurs almost independent of the orientation of the grating grooves if the electric field E is perpendicular to the incident plane. Except for some change of the shape and the relative depth of the two dips, the positions of the two dips and the maximum remain essentially the same. [Fig. 9, curve (c)]
4. The double dip occurs only for samples with relatively smaller grating constants. Except for one sample with $d = 20\mu m$ (which shows a much smaller effect) all samples with $d > 10\mu m$ show no double dips. This suggests that the effect is dependent on the grating constants. (Fig. 8)

5. The double dip occurs only for samples having a grating on the surface. It does not show up on a smooth or an irregularly rough surface.

The double dips do not depend on the incident angle of the light. [Fig. 9, curve (b)]. No systematic effects were found from the spectra for various kinds of doped semiconductors or different carrier concentrations.

An emission effect around the plasmon frequency was first considered in order to explain the double dip phenomenon. This concept included the excitation and subsequent emission of the radiative and non-radiative surface plasmon modes. The non-radiative surface modes cannot couple with photons directly. With the help of the grating on the sample surface, however, these modes can emit light as well as be excited by it. In fact, Wood's anomalies appeared as dips for some frequencies and peaks for others. In the experiments of surface waves on Al-coated gratings, Ritchie, et al. observed peaks at the surface plasmon frequencies instead of dips. Both radiative and non-radiative modes can radiate light at the frequencies consistent with our data but failed in other important conditions: (a) the surface plasmon can only couple with light polarized parallel to the incident plane; (b) the peak position depends on the incident angle of light; (c) according to Ferrell,[17] the photon intensity and the line breadth of the emitting light from a radiative plasmon should have a characteristic angular dependence, $I(\theta) \sim \cos\theta/[1 + (\theta_o/\theta)^2]$, where θ_o is a constant; (d) the sample must be a thin film, as thin as the order of wavelength; (e) in the

non-radiative case, it should depend on the orientation of the grating grooves;[22] (f) the shape of peaks appearing in all experiments performed by Ritchie, etc. are high and sharp.[11,12] All of the above features [(a) through (e)] have not shown up in our double dip spectra, and therefore surface plasma radiation must be ruled out as explaining the effects.

Another type of wave transformation process may be considered as an emission process. When an incident wave is coupled with a medium it is possible for a secondary wave to be generated. Light of different frequency may be emitted. Bianconi and Iannuzzi[23] were able to observe the frequency transformation through thin silver films (thickness 200-400A). They first shone non-polarized light of frequency $\sim 2\omega_p$ on the film, then excited a plasmon with a p-polarized light at the plasma frequency ω_p. By measuring the reflectance at $2\omega_p$ before and after the plasmons were excited, they found that the p-polarized light at frequency ω_p had caused the reflectance to be increased at $2\omega_p$. In other words, a fraction of wave energy at ω_p had been transformed into the frequency $2\omega_p$ by the thin silver film. This fraction is of the order of 10^{-2}-10^{-1}, about the same order as the peak in the double dip effect. If this process can be reversed, i.e. the light at frequency $2\omega_p$ can be transformed into frequency ω_p through excitation of the plasmon, this could explain the double dips. In order to test this, an experiment was performed as described in the preceding chapter. The plasma frequency ω_p of sample A is $680cm^{-1}$, so $2\omega_p = 1360cm^{-1}$. A silicon filter was used to sharply cut-off the incident light beyond $1000cm^{-1}$.

Thus, with the filter there is no light at $2\omega_p$ shining on the sample. [see Fig. 9, curves (d) and (d')] If the double dip is caused by the above wave transformation process, it should disappear under this condition. Nothing other than the overall intensity changed in this experiment, meaning that such a transformation process cannot be responsible for the effect.

A band effect can also cause double dips around the surface plasmon frequency. Dispersion relations were calculated for the samples. Sample A10 has $k/2\pi = 976cm^{-1}$ at the suspected emission frequency $\omega_p = 680cm^{-1}$; it is approximately equal to $k/2\pi = 1000cm^{-1}$ which is calculated from the gap condition $k = n\pi/d$ for $d = 10\mu$ and $n=2$. This is just coincidence because to have a band gap, a surface wave must have been excited. It is unreasonable to be exciting a surface wave by s-polarized but not by p-polarized light. Furthermore, if a band gap has been observed at n=2 for A10, it should be more easily seen at n=1, also at n=1,2 for sample A20 etc. But nothing was observed in the spectra for those cases.

There are two types of Wood's anomalies. The first type[24] was identified as a surface plasmon resonance and required that the electric field be parallel to the incident plane. The second type, with the electric field perpendicular to the incident plane, has been commonly called "P-anomalies"[25] referring to the fact that the electric field is parallel to the grating grooves. In order to avoid confusion with the "P-polarized" light, the second type shall be called E-mode anomalies and the first type H-mode anomalies. The E-mode anomalies were also discovered by Wood,[26] but experiments

were made in more detail by Palmer.[25] The two most important points among the nine characteristics summarized by Palmer are: (a) the E-mode anomalies occur only when the grating grooves are sufficiently deep compared to the wavelength; (b) both E- and H-mode anomalies may have the same appearance or may be entirely different. To explain the phenomenon, Hessel and Oliner[27] presented a theory based on the guided wave approach. They suggested that the E-mode anomalies are due to the guided wave propagating inside the grating grooves provided that the grooves are deep enough to support the wave. They estimated the groove depth h in the range $(n\lambda_g/2) > h > (2n-1)(\lambda_g/4)$, $n=1,2,\cdots$ where λ_g is the wavelength in the guide. For our double dips, the s-polarized light would be consistent with the E-mode anomalies. But the depth of the grating grooves as calculated from the above relation should be at least $h > 3\mu m$, which is not the case. The groove depth is about $1\mu m$ as estimated from photomicrographs. Furthermore, the fact that the double dip does not depend on the orientation of the grating cannot be understood with this theory. However, the possibility of E-mode anomalies does not have to be given up, and a rigorous grating theory is presented in the next chapter. The guided wave effect will be automatically included should it have any contribution to the double dips.

An explanation given by Fischer[8] is that the double dip is due to a fraction of bulk plasmon resonance adding to the "plasma edge" formed by the individual electron oscillation. If it is so, then the bulk plasmon excitation will appear as double dips in many cases. The fact in Fischer's explanation that the bulk plasmon resonance

in these experiments would only be excited by p-polarized light cannot be ignored and casts doubts. Also, the role of a grating in surface plasmons is to produce its evanescent wave; there is no need of an evanescent wave, and hence no need of a grating in exciting the bulk plasmon. The data show that no double dips appeared in the spectrum from a smooth surface, and the double dips do not depend on the orientation of the grating. All these cannot be explained by the bulk plasmon resonance.

V. A RIGOROUS GRATING THEORY

The grating theory is a very old subject in optics. A variety of theories were worked out by many people using different techniques to solve the problem. But we can say that up to the present there is still lack of a completely satisfactory one.

The first grating theory to explain Wood's anomalies was derived by Rayleigh.[28] He used an expansion of the form $\sum_n C_n e^{i\beta z} e^{i(k\sin\theta+2\pi n/d)x}$ and a periodic boundary condition for a completely conducting material (metal). Fano[29] followed his method but assumed some dielectric influence on the surface to avoid the singularities at the Rayleigh wavelengths $\lambda_R = (d/n)(1-\sin\theta)$. Artmann[30] then made a further improvement for the singularities. Their theories qualitatively explain the H-mode anomalies quite well but predict no E-mode anomalies. After the E-mode anomalies were discovered for gratings with deeper grooves, Lippmann[31] showed the Rayleigh theory to be incomplete. He pointed out that both Rayleigh and Fano had neglected the incoming waves inside the grating groove region. This is valid only for shallow gratings. Sellberg[32] calculated the Rayleigh expansion fields with a computer. He found that in most circumstances, the summation of the expansion does not approach a limit, but either oscillates between two values or diverges. For example, it converges only when the sum of the two base angles is less than 60° for triangular grating grooves. Nonetheless, the Rayleigh expansion can still be useful in some cases.

Twersky,[33] et al. developed another approach which is based on

multiple-scattering point of view. The total scattered field is expressed in terms of the multiple scattering amplitude of one element of the grating grooves. This expression then specifies completely all the coupling effects from individual elements. For appropriate sets of parameters, Wood's anomalies are obtained. Also, the theory satisfactorily predicts the location and shape of the anomalies for specific forms of gratings. Unfortunately, it is limited to gratings of relatively shallow groove depth, because the restriction to scatterers is equivalent to requiring that the depth must be small compared to the wavelength. Therefore, these multiple-scattering results have not exhibited the E-mode anomalies.

There are also other theories based on small-perturbation arguments. For example, Harris, et al.[34] employed a WKB approximation to solve the problem. Some people (Peng,[35] et al.) confined the problem to special grating shapes, such as sinusoidal and triangular. All these are obviously not a general treatment and cannot be valid in all cases.

Oliner and Hessel's[27] guided wave idea, mentioned in the last chapter, is to use a periodic impedance on the boundary surface. The periodic impedance contains all information and effects inside the grating and grooves, including multi-scattering to match the phase, guided wave resonance due to appropriate groove depth, etc. Hence no other special conditions, like incoming wave, groove shape and depth, or any approximation need be considered directly. Theoretically, it may be called a rigorous theory. The problem is how to determine the relations between the Fourier components of the

impedance and the grating shape or the dielectric function inside the grating. It seems very difficult. Even for the most simple case, as appeared in his examples, it is extremely tedious.

A rigorous grating theory has recently been worked out by Petit, et al.[36] It is not ideal in some cases, because the problem can only be solved numerically. Before the result has been plotted by the computer, it is hard to gain any knowledge simply from the equations. However, it is practical and useful to calculate all possible classical grating effects in a spectrum.

The following derivation follows that of Petit with a slight modification. The numerical method is similar to the finite-difference technique used by Kalhor, et al.[37]

The main idea of this theory is to use a periodic dielectric function inside the groove area. The dielectric function of a material is obtained from 9, Chapter II, then the Fourier components of this periodic function can be easily calculated for any grating shape. Unlike Hessel and Oliner who solved the problem by periodic boundary conditions for the surface impedance which contains the details, this theory treats the boundary planes between air and the grating surface, the grating surface and the substrate in the usual non-periodic way, but leaves the periodic part inside the grooves to be automatically solved by numerical methods.

Figure 10 is a grating in a Cartesian coordinate system. The grating surface is at the z=0 plane with grooves parallel to the y-axis. A rectangular grating is assumed although the theory can be applied to any groove shape.

The electromagnetic wave in the space without free charge is

$$\Delta\psi - \frac{\mu\varepsilon}{c^2}\frac{\partial^2\psi}{\partial t^2} = 0 \qquad (1)$$

where $\mu=1$ for non-magnetic materials and ε is the effective dielectric function including all current effects inside the medium. ψ can be any component of the vector potential, scalar potential, electric field vector or magnetic field vector.

The equation is separated by setting $\psi(x,y,z,t) = W(x,y,z)T(t)$. The $T(t)$ equation takes the form $\frac{\partial^2 T}{\partial t^2} + \omega^2 T = 0$ and has the solution $T(t)\sim e^{i\omega t}$. The frequency ω will be assumed to have the same value in all expressions, and hence $e^{i\omega t}$ is a common factor in all ψ's to be ignored. To consider s-polarized plane waves only, i.e. the electric vector is polarized in the y-direction, ψ is taken to be $E_y = E$ and $E_x = E_z = 0$. Let the incident plane be perpendicular to the y-axis, then the field E will be independent of y.

Now divide the system into three regions. Region (1) with $z \geq h$, (where h is the height of the grating grooves) and region (3) with $z \leq 0$ are homogeneous; so, the wave equation for the electric field E has the Helmholtz form.

$$z \geq h \text{ or } z \leq 0, \quad \frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial z^2} + \alpha E = 0; \quad \alpha \equiv k_o^2\varepsilon = \frac{\omega^2}{c^2}\varepsilon. \qquad (2)$$

For the inhomogeneous region (2), where $0 \leq z \leq h$, ε and hence α is a function of x and z but periodic with respect to x as can be seen from the figure.

$$\alpha = k_o^2 = \frac{\omega^2}{c^2} \quad \text{if} \quad z > f(x)$$
$$= k_o^2 \varepsilon_g \quad \text{if} \quad z < f(x) \tag{3}$$

and $$\alpha(x,z) = \alpha(x+d,z) \tag{4}$$

where d is the grating constant. ε_g is taken to be different from ε because of the distortion of electron distribution inside the groove region. The function f(x) describes the grating profile.

The Helmholtz equation then takes the inhomogeneous form.

$$0 \leq z \leq h, \quad \frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial z^2} + \alpha(x,z)E = 0; \quad \alpha(x,z) = \alpha(x+d,z). \tag{5}$$

There is in general no exact solution for this differential equation, but as is well known in solid state physics, it has the Bloch form in one dimension:

$$E = u(x,z)e^{ik_x t} \quad \text{with} \quad u(x,z) = u(x+d,z) \tag{6}$$

The function u is periodic in x with period d, so it can be taken as a Fourier expansion of the form

$$u(x,z) = \sum_{n=-\infty}^{\infty} E_n(z)e^{inKx}, \quad K = \frac{2\pi}{d} \text{ and } n=0,\pm1,\pm2 \cdots \tag{7}$$

Then, inserting Eq. (7) into Eq. (6)

$$0 \leq z \leq h, \quad E(x,z) = \sum_{n=-\infty}^{\infty} E_n(z)e^{i\gamma_n x} \tag{8}$$

where $\gamma_n \equiv k_x + nK = k_o \sin\theta + nK$. k_x has been set equal to $k_o \sin\theta$ in order to have the proper behavior as $d \to \infty$ and k_o and θ are the wave vector and the incident angle, respectively, of the incident wave.

The solutions in regions (1) and (3) are of course the Rayleigh expansion. The x-component of all wave vectors must be γ_n to satisfy the boundary conditions, and the magnitudes of the wave vector are k_o for the air side and $\sqrt{\varepsilon}\, k_o$ for the medium side because of the conservation of wave vector magnitudes in the grating scattering process. So, the z-components of the wave vectors are

$$\beta_{n,1} = \sqrt{k_o^2 - \gamma_n^2} \quad \text{if } k_o > \gamma_n$$
$$= i\sqrt{\gamma_n^2 - k_o^2} \quad \text{if } k_o < \gamma_n \qquad \text{in air} \tag{9}$$

and $$\beta_{n,3} = \sqrt{\varepsilon\, k_o^2 - \gamma_n^2} \quad \text{in the medium of region (3).} \tag{10}$$

The evanescent waves are produced in the air side in the case $k_o < \gamma_n$. Since the dielectric function ε in the medium is usually complex, the evanescent wave for $\beta_{n,3}$ on the medium side is not explicitly expressed.

The Rayleigh expansions in regions (1) and (3) are then

$$h \leq z,\ E(x,z) = e^{i(k_o \sin\theta \cdot x - k_o \cos\theta \cdot z)} + \sum_{n=-\infty}^{\infty} B_n e^{i\beta_{n,1} z} e^{i\gamma_n x} \tag{11}$$

and

$$0 \geq z,\ E(x,z) = \sum_{n=-\infty}^{\infty} C_n e^{-i\beta_{n,3} z} e^{i\gamma_n x} \tag{12}$$

where the first term in Eq. (11) represents the incident wave with amplitude equaling 1.

The boundary condition require $E(x,z)$ and $\frac{dE(x,z)}{dz}$ [i.e. $H_y(x,z)$] continuous at z=h and z=0. Equating Eqns. (11) and (8) for z= h, yields

$$e^{i(k_o \sin\theta \cdot x - k_o \cos\theta \cdot h)} + \sum_{n=-\infty}^{\infty} B_n e^{i\beta_{n,1}} e^{i\gamma_n x} = \sum_{n=-\infty}^{\infty} E_n(h) e^{i\gamma_n x} \qquad (13)$$

Since the relation holds for any value of x, the coefficients of $e^{i\gamma_n x}$ for each value of n must vanish. Therefore

$$e^{-ik_o \cos\theta \cdot h} + B_o e^{i\beta_{0,1} h} = E_o(h) \qquad (14)$$

and

$$B_n e^{i\beta_{n,1} h} = E_n(h) \text{ for } n \neq 0.$$

For the derivatives of Eqns. (11) and (8) with respect to z at z=h, one has

$$-ik_o \cos\theta e^{-ik_o \cos\theta \cdot h} e^{ik_o \sin\theta \cdot x} + \sum_{n=-\infty}^{\infty} B_n i_{n,1} e^{i\beta_{n,1} h} e^{i\gamma_n x} = \sum_{n=-\infty}^{\infty} \left.\frac{dE_n}{dz}\right|_{z=h} e^{i\gamma_n x} \qquad (15)$$

again, letting the coefficients of $e^{i\gamma_n x}$ for each value of n equal zero,

$$-ik_o\cos\theta e^{-ik_o\cos\theta\cdot h} + B_o i\beta_{0,1} e^{i\beta_{0,1}h} = \left.\frac{dE_o}{dy}\right|_{z=h} \tag{16}$$

and

$$B_n i\beta_{n,1} e^{i\beta_{n,1}h} = \left.\frac{dE_n}{dz}\right|_{z=h} \quad \text{for } n\neq 0.$$

Following the same procedure, one can do the same thing for Eqns. (12) and (8) at z=0 and for Eqns. (11) and (8) at z=h. The result is, for E(x,0) continuous:

$$C_n = E_n(0) \tag{17}$$

and for $\left.\frac{dE}{dz}\right|_{z=0}$ continuous:

$$-C_n\beta_{n,3} = \left.\frac{dE_n}{dz}\right|_{z=0} . \tag{18}$$

By cancelling B_n from Eqns. (14) and (16), and also C_n from Eqns. (17) and (18), one gets the following conditions

$$\left.\frac{dE_n}{dz}\right|_{z=0} + i\beta_{n,3}E_n(0) = 0, \tag{19}$$

$$\left.\frac{dE_o}{dz}\right|_{z=h} - i\beta_{0,1}E_o(h) = -2ik_o\cos\theta e^{-ik_o h\cos\theta}, \tag{20}$$

and

$$\left.\frac{dE_n}{dz}\right|_{z=h} - i\beta_{n,1}E_n(h) = 0 \text{ for } n\neq 0.$$

Equations (17) and (18) are actually the boundary conditions for the wave function inside the groove at the surfaces z=0 and z=h. These can be written in matrix form

$$\left.\frac{d[E]}{dz}\right|_{z=0} + [L_o][E(0)] = 0 \tag{21}$$

$$\left.\frac{d[E]}{dz}\right|_{z=h} + [L_h][E(h)] = [G] \tag{22}$$

where $[E]$, $[L_o]$, $[L_h]$ and $[G]$ are all matrices. Explicitly,

$$[E] = \begin{bmatrix} \vdots \\ E_{-3} \\ E_{-2} \\ E_{-1} \\ E_0 \\ E_1 \\ E_2 \\ E_3 \\ \vdots \end{bmatrix} \tag{23}$$

$$[L_o] = \begin{bmatrix} \ddots & & & & & & & & \\ & i\beta_{-3,3} & 0 & 0 & 0 & 0 & 0 & 0 & \\ & 0 & i\beta_{-2,3} & 0 & 0 & 0 & 0 & 0 & \\ & 0 & 0 & i\beta_{-1,3} & 0 & 0 & 0 & 0 & \\ & 0 & 0 & 0 & i\beta_{0,3} & 0 & 0 & 0 & \\ & 0 & 0 & 0 & 0 & i\beta_{1,3} & 0 & 0 & \\ & 0 & 0 & 0 & 0 & 0 & i\beta_{2,3} & 0 & \\ & 0 & 0 & 0 & 0 & 0 & 0 & i\beta_{3,3} & \ddots \end{bmatrix} \tag{24}$$

$$[L_h] = \begin{bmatrix} \ddots & & & & & & & & \\ & -i\beta_{-3,1} & 0 & 0 & 0 & 0 & 0 & 0 & \\ & 0 & -i\beta_{-2,1} & 0 & 0 & 0 & 0 & 0 & \\ & 0 & 0 & -i\beta_{-1,1} & 0 & 0 & 0 & 0 & \\ & 0 & 0 & 0 & -i\beta_{0,1} & 0 & 0 & 0 & \\ & 0 & 0 & 0 & 0 & -i\beta_{1,1} & 0 & 0 & \\ & 0 & 0 & 0 & 0 & 0 & -i\beta_{2,1} & 0 & \\ & 0 & 0 & 0 & 0 & 0 & 0 & -i\beta_{3,1} & \ddots \end{bmatrix} \tag{25}$$

$$[G] = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ 0 \\ -ik_o h\cos\theta \\ -2ik_o \cos\theta e \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix} \tag{26}$$

It can be seen that both $[L_o]$ and $[L_h]$ are diagonal matrices. [G] is related to the incident wave. If no source is applied to the system, [G] is zero.

We now return to the inhomogeneous Helmholtz Eq. (5) in the region $0 \leq z \leq h$. Since $\alpha(x,z)$ is periodic in x with period d, it can be expanded in a Fourier series

$$\alpha(x,z) = \sum_{n=-\infty}^{\infty} \alpha_n(z) e^{inKx} \tag{27}$$

with

$$\alpha_n(z) = \frac{1}{d} \int_0^d d(x,z) e^{-inKx} dx \tag{28}$$

Substituting Eqns. (27) and (8) into Eq. (5), then yields

$$0 \leq z \leq h, \frac{d^2 E_n(z)}{dz^2} - \gamma_n^2 E_n(z) + \sum_{m=-\infty}^{\infty} \alpha_{n-m}(z) E_m(z) = 0 \tag{29}$$

Writing Eq. (29) in matrix form:

$$\frac{d^2[E(z)]}{dz^2} = [A(z)][E(z)] \tag{30}$$

where [A(z)] is a matrix with elements

$$a_{n,m} = -\alpha_{n-m} + \gamma_n^2 \delta_{nm}. \tag{31}$$

For the rectangular grating which is considered in Fig. 10, the dielectric function inside the grooves is

$$\varepsilon(x) = \varepsilon(x+d)$$

$$\varepsilon(x) = 1 \qquad 0 < x < a$$

$$= \varepsilon_g \qquad a < x < d$$

Therefore the Fourier components are

$$\varepsilon_n = \frac{1}{d}\int_0^d \varepsilon(x) e^{-inKx} dx$$

$$= \frac{1}{d}\left[\int_0^a e^{-inKx} dx + \int_a^d \varepsilon_g e^{-inKx} dx\right]$$

$$= \frac{i}{nKd}\left[(1-\varepsilon_g)e^{-inKa} + {}_g e^{-inKd} - 1\right] \text{ for } n \neq 0$$

and

$$\varepsilon_o = \frac{a + \varepsilon_g b}{d} .$$

Then $\alpha_n = k_o^2 \varepsilon_n$ is independent of z, and so is [A(z)]. Therefore [A(z)] in Eq. (30) becomes a constant matrix for this special grating shape.

The problem now is to solve the matrix differential equation

$$\frac{d^2[E(z)]}{dz^2} = [A][E(z)] \qquad (30)'$$

with boundary conditions

$$\left.\frac{d[E(z)]}{dz}\right|_{z=0} + [L_o][E(0)] = 0 \qquad (21)$$

and $$\left.\frac{d[E(z)]}{dz}\right|_{z=h} + [L_h][E(h)] = [G] \qquad (22)$$

To develop a numerical method to solve this problem, let's divide h into N equal intervals, $\Delta = h/N$. For sufficiently large N the integration identities

$$\int_{I\Delta}^{(I+1)\Delta} \frac{dE}{dz}\, dz = E\{(I+1)\Delta\} - E\{I\Delta\}$$

and $$\int_{I\Delta}^{(I+1)\Delta} \frac{d^2E}{dz^2}\, dz = \left.\frac{dE}{dz}\right|_{z=(I+1)\Delta} - \left.\frac{dE}{dz}\right|_{z=I\Delta}$$

can be approximated by

$$E\{(I+1)\Delta\} - E\{I\Delta\} = \left.\frac{dE}{dz}\right|_{z=I\Delta} \cdot \Delta$$

and $$\left.\frac{dE}{dz}\right|_{z=(I+1)\Delta} - \left.\frac{dE}{dz}\right|_{z=I\Delta} = \left.\frac{d^2E}{dz^2}\right|_{z=I\Delta} \cdot \Delta$$

where I is any positive integer between 0 and N, and all E and $\frac{dE}{dz}$ in the above equations denote the matrices [E] and $\frac{d[E]}{dz}$ as in Eqns. (21), (22) and (30)'. For simplicity, let

$$E\{I\Delta\} \equiv [E(I)],\ E\{(I+1)\Delta\} \equiv [E(I+1)]$$

and $$\left.\frac{dE}{dz}\right|_{z=I\Delta} \equiv [F(I)],\ \left.\frac{dE}{dz}\right|_{z=(I+1)\Delta} \equiv [F(I+1)].$$

Then Eq. (31) becomes

$$\begin{aligned} [E(I+1)] &= [E(I)] + [F(I)]\cdot\Delta \\ [F(I+1)] &= [A]\cdot\Delta\cdot[E(I)] + [F(I)] \end{aligned} \tag{32}$$

where Eq. (30)' has been used to substitute $\left.\frac{d^2E}{dz^2}\right|_{z=I\Delta}$ with $[A][E(I)]$. Equation (32) has the matrix form

$$\begin{bmatrix} [E(I+1)] \\ [F(I+1)] \end{bmatrix} = \begin{bmatrix} [1] & [1]\Delta \\ [A]\Delta & [1] \end{bmatrix} \begin{bmatrix} [E(I)] \\ [F(I)] \end{bmatrix} \tag{33}$$

Since $[A]$ is constant, $[E(I)]$, and $[F(I)]$ can be expressed by $[E(I-1)]$, and $[F(I-1)]$, having the same form as Eq. (33). Doing this $(I+1)$ times, one finally gets to the expression

$$\begin{bmatrix} [E(I+1)] \\ [F(I+1)] \end{bmatrix} = \begin{bmatrix} [1] & [1]\Delta \\ [A]\Delta & [1] \end{bmatrix}^{I+1} \begin{bmatrix} [E(0)] \\ [F(0)] \end{bmatrix} \tag{34}$$

From Eq. (21), $[F(0)] = -[L_o](E(0)]$ and taking $I+1 = N$, and noting that $[E(N)] \equiv [E(h)]$ and $[F(N)] \equiv \left.\frac{d[E]}{dz}\right|_{z=h}$, then Eq. (34) becomes

$$\begin{bmatrix} [E(h)] \\ \left.\frac{d[E]}{dz}\right|_{z=h} \end{bmatrix} = \begin{bmatrix} [1] & [1]\Delta \\ [A]\Delta & [1] \end{bmatrix}^{N} \begin{bmatrix} [1] & 0 \\ -[L_o] & 0 \end{bmatrix} \begin{bmatrix} [E(0)] \\ [E(0)] \end{bmatrix} \tag{35}$$

The matrix

$$\begin{bmatrix} [1] & [1]\Delta \\ [A]\Delta & [1] \end{bmatrix}^{N} = \begin{bmatrix} [M_{11}] & [M_{12}] \\ [M_{21}] & [M_{22}] \end{bmatrix}$$

in Eq. (35) is a constant and all $[M_{ij}]$'s can be calculated numerically. By setting $[M_1] = [M_{11}] - [M_{12}][L_o]$ and $[M_2] = [M_{21}] - [M_{22}][L_o]$, one has

$$[E(h)] = [M_1][E(0)] \tag{36}$$

and

$$\left.\frac{d[E(z)]}{dz}\right|_{z=h} = [M_2][E(0)]$$

Now substituting Eq. (36) into Eq. (22) to obtain relation between [E(0)] and [G], one has

$$[M][E(0)] = [G] \tag{37}$$

with $[M] = [M_2] + [L_h][M_1]$.

Then

$$[E(0)] = [M]^{-1}[G] \tag{38}$$

and substituting Eq. (38) into Eq. (36) to find [E(h)] and $\left.\frac{d[E(z)]}{dz}\right|_{z=h}$

$$[E(h)] = [M_1][M]^{-1}[G] \tag{39}$$

and

$$\left.\frac{d[E(z)]}{dz}\right|_{z=h} = [M_2][M]^{-1}[G]. \tag{40}$$

The reflectance R at the zeroth order spectrum is equal $|B_o|^2$ where

B_o in Eq. (14) has the value

$$B_o = \frac{E_o(h) - e^{-ik_o \cos\theta \cdot h}}{e^{i\beta_{0,1} h}} . \qquad (41)$$

Setting $[T] \equiv [M_1][M]^{-1}$, then from Eq. (39)

$$E_o(h) = T_{oo}(-2ik_o \cos\theta e^{-ik_o \cos\theta \cdot h}).$$

Therefore, from Eq. (41)

$$R = B_o B_o^* = \left|2ik_o T_{oo} \cos\theta + 1\right|^2 \qquad (42)$$

The significance of the matrix [M] can be seen from Eqns. (38) to (40). If the determinant of [M] equals zero, the equations have solutions only for [G] = 0, i.e. no source applied to the system. This means a resonance condition occurs inside the grating. In this TE-mode geometry, resonance is possible only by exciting a guided wave inside the groove area for suitable depth of the grooves.

In the next chapter, we consider numerical calculations using these results.

VI. NUMERICAL CALCULATIONS

A. Computation of the Rigorous Theory

Using the rigorous grating theory presented in the last chapter, the reflectance was calculated for the samples discussed previously (see Table I). The parameters used are from Willardson and Beer's[38] book, and McMahon's thesis.[39] They are listed in Table III.

Only seven terms starting from the central n=0, going to both positive and negative integers and terminating at n=±3, were taken for the summations of the infinite series presented in the last chapter. This series has been checked and found to converge rapidly for the grating constants used here. The N in Eq. (35) of the last chapter was taken to be N=10, which is large enough for h to be in the range of micrometers. The possibility of a guided wave inside the groove region has been checked but found not to exist since the absolute value of the determinant of [M] is still large when the groove height is up to 5μm.

The results of the calculation depend upon what is assumed for the dielectric functions.

Case 1: Assume the dielectric function inside the lanes (the region labeled g in Fig. 10) of the grating region is the same as the bulk value of the material, i.e. $\varepsilon_g = \varepsilon$. The results of this assumption are no double dips in the reflectivity as shown in Fig. 11, curve (a). The reflection in this case is almost the same as if it were from a smooth surface. Here the wavelength is larger than the grating constant, and no diffraction effect appeared.

Table III. Parameters Used in Calculations

	W_{To}	m*/m	ε_∞	$\delta\varepsilon$	$1/\tau$	Γ
InSb	181.5	0.015	15.68	2.05	16.2	4.2
GaAs	268.0	0.08	10.9	1.63	15.0	1.9

Symbols have been defined in 9, Chapter II.

Case 2: Assume $\varepsilon_g > \varepsilon$. This difference is due to the distortion of the electron density inside the grating region; so, the concentration of electrons as well as damping (and probably the force constant) inside this region is not the same as the bulk value. In this case, the double dip appears. By choosing suitable ω_{pg} (the plasma frequency in region g) for ε_g, the dip positions can be fitted. In Fig. 11, curve (b), $W_{pg} = 520cm^{-1}$ is used for sample A10. The first and the second dips are at $590cm^{-1}$ and $735cm^{-1}$ respectively. Compared to the experimental values of $570cm^{-1}$ and $740cm^{-1}$, the discrepancies are less than 5%. Nevertheless, the calculated intensity, especially around the first dip position, is much higher than the experimental value.

Case 3: R. W. Gamon and E. D. Palik[40] in their recent paper pointed out that the carrier density, and hence the plasma frequency, near the surface is not the same as the bulk value for doped semiconductors InSb and GaAs. The variation is about ±7%. They did not mention how deep from the surface this variation exists; it might be only in the order of angstroms. However, because of the damage inside the groove region, and also because of the distortion of electron density from the neighboring lanes, it is reasonable to make a further assumption that the dielectric function in the region S of Fig. 10(b) is not ε but $\varepsilon_s < \varepsilon$. Fig. 11, curve (c) is a plot for sample A10 based on this assumption by taking $W_{pg} = 520cm^{-1}$ for ε_g and $W_{ps} = 720cm^{-1}$ for ε_s, assuming a layer thickness of 1μm on the groove surface. The position of dips fits the experimental data. The intensity is still too high, although better than the previous case.

The discrepancy of the calculated intensity with the experimental result might be partly explained by considering their spectra in the smooth surface case. Taking a=0 or h=0 in Fig. 10(b), the grating theory reduces to the formula of the smooth surface case - the same expression as in 9, Chapter II. Although McMahon found an excellent fit for his low concentration samples of InSb, in these higher concentration samples, it can be seen from Fig. 12, curves (a) and (b), that the reflectance of the frequency region between W_{To} and W_p from experiment is much lower than that from calculation. The largest discrepancy, which is about 25%, is around the first dip position. If the calculated results in Fig. 11, curves (b) and (c) are reduced by the same fraction, the curve will be closer to the experimental result although it is still higher. [Fig. 12, curve (c)]. Actually the depth of the dips depends on the height which is chosen to have the dielectric function ε_s. By choosing a height of 5μm [Fig. 12, curve (d)], the effect from the dielectric function ε becomes small and the intensity minima of the double dips is even lower than the experimental data. But the intensity in the maximum position between the dips is still too high to fit the data.

The calculated results of samples B10 and C10 are similar to that of A10 described above. Just by choosing suitable values of W_{pg} and W_{ps}, the double dip positions can always be fitted within an accuracy of 5%. The intensities are higher than the experimental value. For large grating constants, the dielectric function ε_g inside the lane region approaches the bulk value ε because of less

distortion of the electron densities, so the double dips will disappear and will reduce to a usual plasma dip.

The calculation was only for the special case in which grating grooves are parallel to the electric field E and perpendicular to the incident plane. For E parallel to the incident plane, or other orientations of the grating grooves, the situation will be much more complicated. The model described in the next section will cover these cases with a very simple argument.

B. A Simple Formula to Explain the Observations

The double dips, if they are due to the distortion of electron densities near the grating surface, can be always imagined as arising from the combination without interference of two reflectances. One reflectance has a dip higher, and the other has a dip lower than the bulk plasma frequency. Looking at Fig. 10(b) and as a result of an average effect from the change of electron densities, damping forces, and force constants inside the lane portion of the grating, let W'_{pg} be the new plasma frequency, τ_g the new damping constant at the region g, and W'_{ps}, τ_s the corresponding quantities at the groove portion s. Then from 9, Chapter II, the dielectric function ε'_g and the reflectance R_1 from the regions g can be calculated with the data W'_{pg} and τ_g. A similar calculation uses ε'_g to find R_2 from the region s with data W'_{ps} and τ_p. If interference is neglected because of the small grating constants compared to the wavelength, the total reflectance appearing in the spectrum can be written as

$$R = \frac{R_1 A_1 + R_2 A_2}{A_1 + A_2} \tag{43}$$

where A_1 is the surface area of the lane portion and A_2 that of the groove portion. Fig. 13, curve (a) is a plot of R by choosing $A_1:A_2 = 7:1$, $W'_{pg} = 525cm^{-1}$, $W'_{ps} = 710cm^{-1}$, and $1/\tau_g = 1/\tau_s = 1/\tau = 15cm^{-1}$ for sample A10. The τ's do not affect the dips much, and they were all taken to be equal for the sake of simplicity. Fig. 13, curve (b) is the experimental data for comparison. The dips and the maximum between them are almost coincident with the calculated values. The intensity is consistent too. Again, for samples B10 and C10, the fit can also be obtained just by choosing suitable values of W'_{pg} and W'_{ps}.

Not only the position and intensity of maximum and dips can be fitted well by this simple formula, but also all properties as summarized in the beginning of Chapter IV can be explained with the following arguments.

1. The plasma frequency inside the lane region is less than the bulk value, while inside the groove region it is larger than the bulk value. So the plasma frequency W_p can be at any position between the two dips. It can be near the midpoint, (Fig. 6), near the lower [Fig. 7(b)], or the higher dip, [Fig. 7(a)]. However, the frequency of the maximum intensity in these three figures is always near the midpoint of the two dips.
2. The only orientation that does not show double dips is for E parallel to the incident plane and perpendicular to the

grating grooves, which is the case that is supposed to have a surface wave excited by the grating. The absence of the double dips may be because any surface wave excitation causing a third dip lying between the double dips washes out the maximum between or it might be because of the following reasons. Referring to Fig. 10(b) in a bulk wave case, the oscillation of electrons at g is only a response to the light shining on g, and that at s is only a response of light shining on s. If the dielectric functions are different between points g and s, and if interference effects can be ignored, the combination of the reflectance from these two points will appear as two dips. In case a surface wave is excited, even if it is not strong, the oscillation of electrons at point g is not only the response of the light shining on g, but also a response of the longitudinal surface wave coming from points s. Therefore, the difference of the dielectric function between those points will be averaged out.

3. From Eq. (43) it can be seen that the double dips only depend on the ratio of the areas $A_1:A_2$. This is why they do not depend on the orientation of the grating grooves. As the direction of the grooves is turned, the damping and the plasma frequency on the lane and groove portion might change a little and so might the position of dips and their intensities. A change of intensity appeared in the spectrum, but any change of the positions is too small to

be observed.

4. It is also seen from the Eq. (43) that the appearance of the double dips must depend on the relative area of the lanes and grooves inside the grating. If $A_1:A_2$ is too large, R_2 will have only a very little effect on R_1. Also, the surface becomes closer to the smooth case and every parameter will return to the bulk value. So this explains why the double dips occur only for smaller grating constants. For d = 20μm, 30μm ··· etc., the grating constant is larger than the wavelength of interest, so that interference becomes important and Eq. (43) is no longer valid. However, it is still true that ε_g will approach ε as the grating constant becomes larger and larger. Then the two dips become closer and closer and finally become one dip at the bulk frequency. (see Fig. 8).
5. The important idea for the double dips is that the dielectric function must be in the form of exactly two major groups. For a smooth surface or a surface that does not vary in a regular manner, this condition obviously does not hold. This is why the effect has only been observed for the grating surface so far.

C. Connection Between the Simple Formula and the Grating Theory

The formula given in the previous section which has been described by a very simple model explains the experimental phenomenon. It is worthwhile to make the connection between the simple formula

and the grating theory given in the preceding chapter. Case 3 in section (A) gave results very similar to the simple formula in section (B). Both showed the double dips in the correct positions with almost the same value of parameters. The discrepancy is that the result from the simple model can also fit the relative intensity of the double dips and the position of the maximum, while the grating theory failed to do this. This may mean that the grating theory or the calculation assumptions are not completely accurate because the factors affecting them are so complicated.

However, another important characteristic appearing in the calculation result of the grating theory is disquieting. Examining Fig. 12, curve (b) and (c), one sees that there is only a small difference in the spectra between the Cases 2 and 3. This means that for a thin layer with dielectric function ε_s, the part contributed by ε_s to the reflectance is negligibly small. Actually the calculations show that only when the layer thickness reaches 4 to 5μm [see Fig. 12, curve (d)] does the contribution of ε_s to R become comparable to that of ε_g, the dielectric function in the lane position. Only in this case can R be considered as the superposition of the reflectance from the lane and the groove portion of the grating, and then the model described in section (B) is correct. This thickness is hard to verify, and is also hard to believe.

If the layer thickness is not large enough, the model of section (B) is not valid but Eq. (43) can still be used to describe the double dip phenomenon. In this case, $F_1 = A_1/(A_1 + A_2)$ and

$F_2 = A_2/(A_1 + A_2)$ only represent some fraction factor, and then F_1R_1 and F_2R_2 are the effective reflectances from the materials with dielectric functions ε_g and ε, respectively. The double dips are still formed from two groups of reflectance. With this idea, the experimental properties can still be explained in a similar way.

VII. DOUBLE DIPS AROUND THE PHONON REGION

Fischer found around the phonon region another dip at the frequency $W = 305cm^{-1}$ for the sample A10 (GaAs, $N = 4.2x10^{18}$ Te-atoms/cm^3, $d = 10\mu m$) besides the phonon-plasmon coupling minimum at $265cm^{-1}$. The depth of these two dips is approximately the same, but the extra one occurs only when E is parallel to the incident plane and perpendicular to the grating grooves, which is the case for which it is possible to excite a surface wave inside the material. This phenomenon is absent in other orientations. Fischer suggested an explanation that this is related to a coupled mode of surface plasmon and Reststrahlen-phonon.

However, since the grating constant d equals $10\mu m$, the wavenumber k is always larger than $6280cm^{-1}$. At this large k, no surface plasmon can be excited around the phonon region. So if this extra dip is a result of surface wave absorption, it can only be a surface phonon.

The longitudinal phonon frequency of GaAs[38] is $W_L = 297cm^{-1}$; then for large k, the surface phonon frequency (from Anderson's thesis) is $W_{sL} \approx (1+1/\varepsilon_\infty)^{-1/2} W_L$, so that $W_{sL} = 284cm^{-1}$ for $\varepsilon_\infty = 10.9$. To find the reflection dip from this eigenfrequency, the exact approach involves grating theory. But instead, we will make estimations following Fischer's very simple method.[6]

Ritchie, et al.[11] have shown that the excitation of a photon-surface plasmon-photon system is equivalent to a typical Lorentzian oscillator. Fischer assumed Lorentzian oscillators for his surface plasmons and got a fit to the dispersion curve. In

addition he also surprisingly found that all parameters except the eigenfrequency used in the oscillator are almost the same as the bulk values. So, from this assumption, the Lorentzian oscillator for the surface phonon contributing to the dielectric function is then

$$\varepsilon(W) = \varepsilon_\infty + \frac{W_{sL}^2 \delta\varepsilon}{(W_{sL}^2 - W^2) - i\Gamma W} \tag{44}$$

Assuming nothing couples to this surface wave, the reflection minimum will occur at the frequency W_{dip}, that makes the real part of the refraction index equal to unity, i.e.

$$\text{Real}(\sqrt{\varepsilon(W_{dip})}) = 1 \tag{45}$$

Using the data $W_{sL} = 284\text{cm}^{-1}$, and the bulk values for $\varepsilon_\infty = 10.9$, $\delta\varepsilon = 1.63$ and $\Gamma = 1.9$ in Eqns. (44) and (45), the dip position was found by numerical method to be

$$W_{dip} = 307\text{cm}^{-1}$$

This fits the experimental value of 305cm^{-1}. The present experiment in this region was performed with a Twyman-Green interferometer, which has been described by Anderson in his thesis. The results agree with Fischer's (Fig. 14).

From the above argument, it can be shown that samples A20($d = 20\mu m$) and A30($d = 30\mu m$) should have the same position of surface phonon dips at 305cm^{-1} and A40($d = 40\mu m$) should have a surface phonon dip almost coincident at the position 265cm^{-1}, which

is the common coupling phonon dip for all cases. These dips were not clearly observed in my experiment.

In the last chapter, in order to explain the non-double dip case in the plasmon region, a surface plasmon close to the bulk frequency has been supposed to be excited for this orientation. It seems unreasonable to have assumed nothing couples to the surface phonon. If the coupling is included, the calculated value of W_{dip} should be lowered by 10 to 20 cm^{-1}.

VIII. CONCLUSION

The best fit to the data is made by assuming that in Fig. 10(b), the dielectric function ε_s ($\varepsilon_s < \varepsilon$) extends to a depth of a few micrometers; and that there is no coupling to surface phonons. Since the character of a semiconductor usually changes dramatically with a few defects in the material, e.g. the huge change of conductivity with a small concentration of impurities, it is not unreasonable for the grating ruled in the surface to cause the results found in this study.

What was done in this research has been to make experimental measurements and to propose a simple model to explain almost all the experimental phenomena with an uncomplicated assumption. To confirm the idea and to obtain the actual distortion of the electron density by a grating with small grating constants, more evidence and more work from other aspects are of course still required.

Figure 1

Reflection spectra of Te-doped InSb ($n=3.96x10^{17}Te/cm^3$) crystal with a spark cut grating with different etch times.

Curve (a) without etch

Curve (b) etch time = 6 sec.

Curve (c) etch time = 12 sec.

Curve (d) etch time = 57 sec.

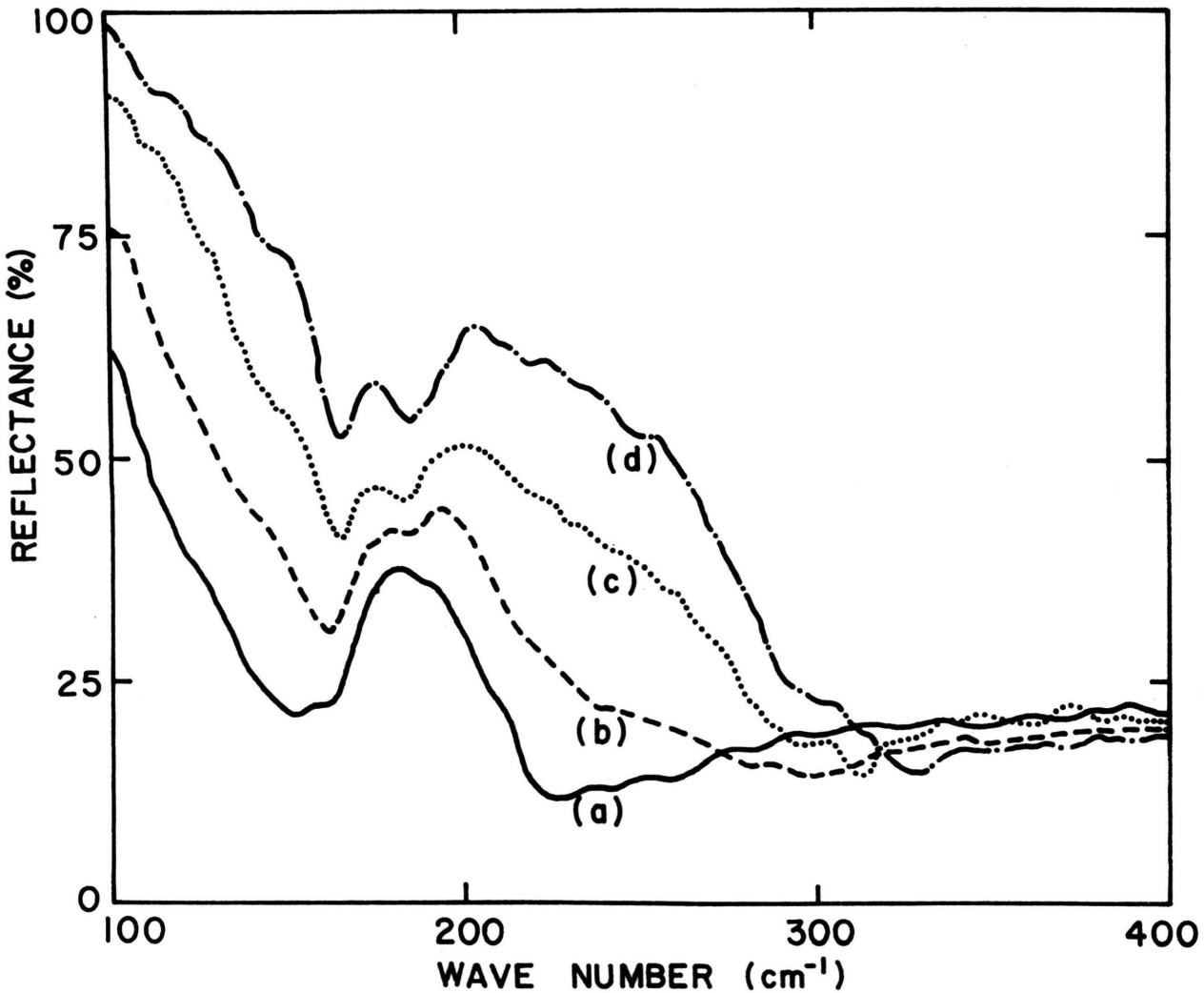
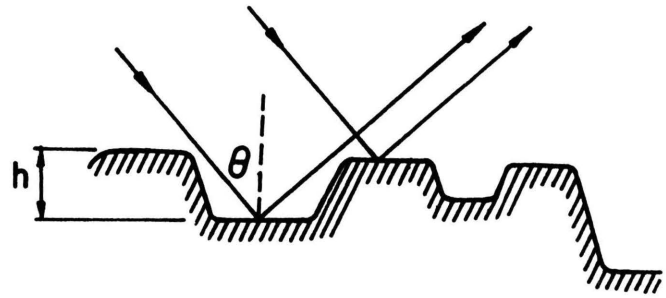

100
75
50
25
0
REFLECTANCE (%)
100
200
300
400
WAVE NUMBER (cm⁻¹)
(a)
(b)
(c)
(d)

Fig. 1

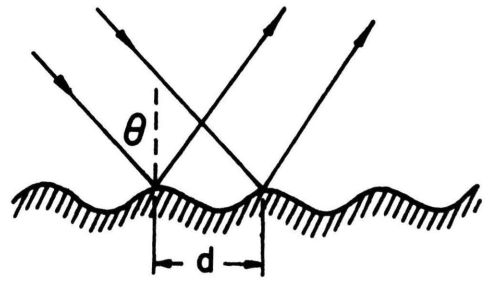
Figure 2

(a) a rough surface

(b) a grating surface



(a)



(b)

Fig. 2

Figure 3

Transmission spectra of the Beckman IR-12 Spectrometer through a IGP-225 polarizer.

Curve (a) E polarized vertically

Curve (b) E polarized horizontally

Curve (c) without polarizer

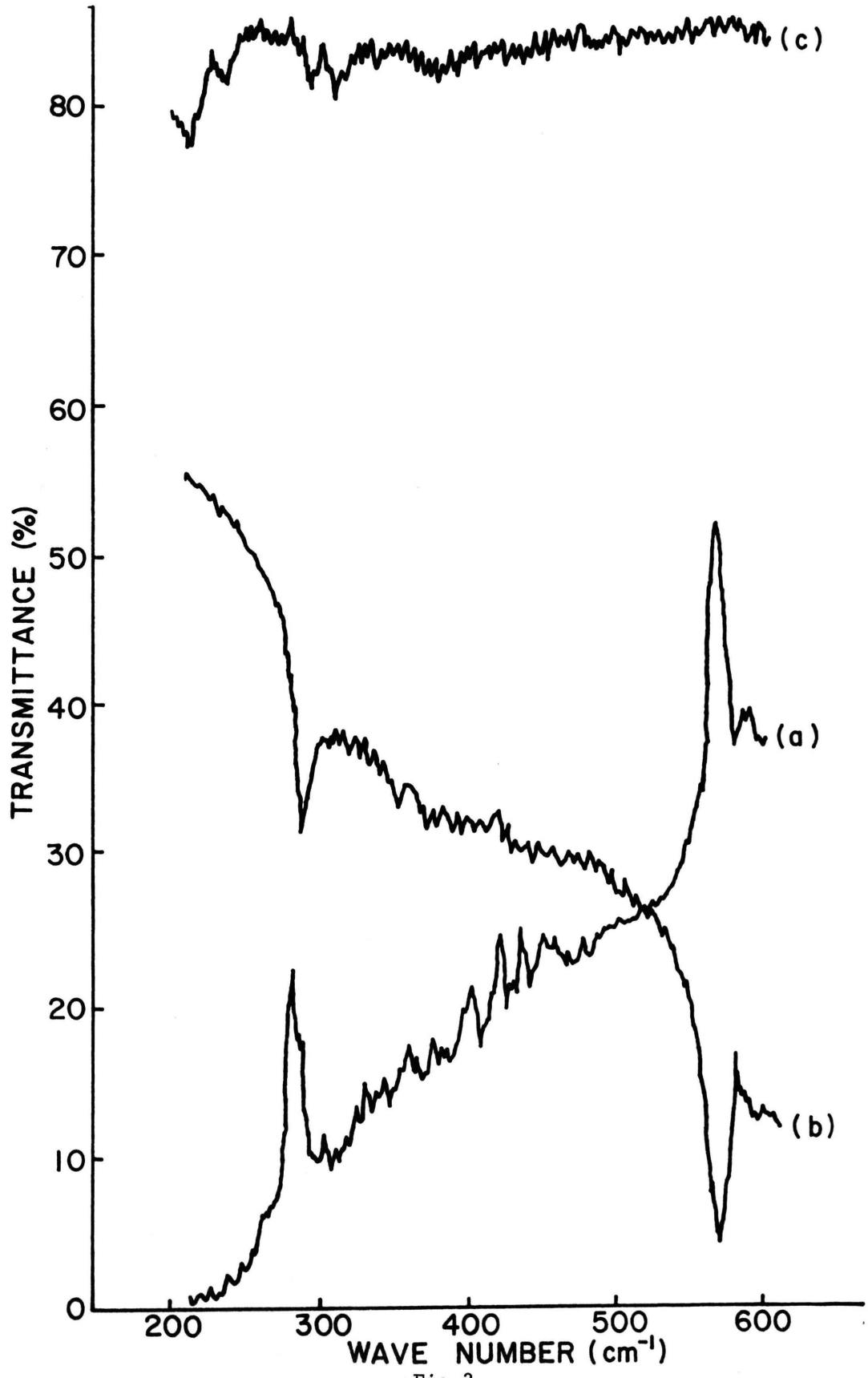

(c)
(a)
(b)
TRANSMITTANCE (%)
0
10
20
30
40
50
60
70
80
200
300
400
500
600
WAVE NUMBER (cm⁻¹)

Fig. 3

Figure 4

Change of the transmission spectra through different orientations of the IGP-225 polarizer.

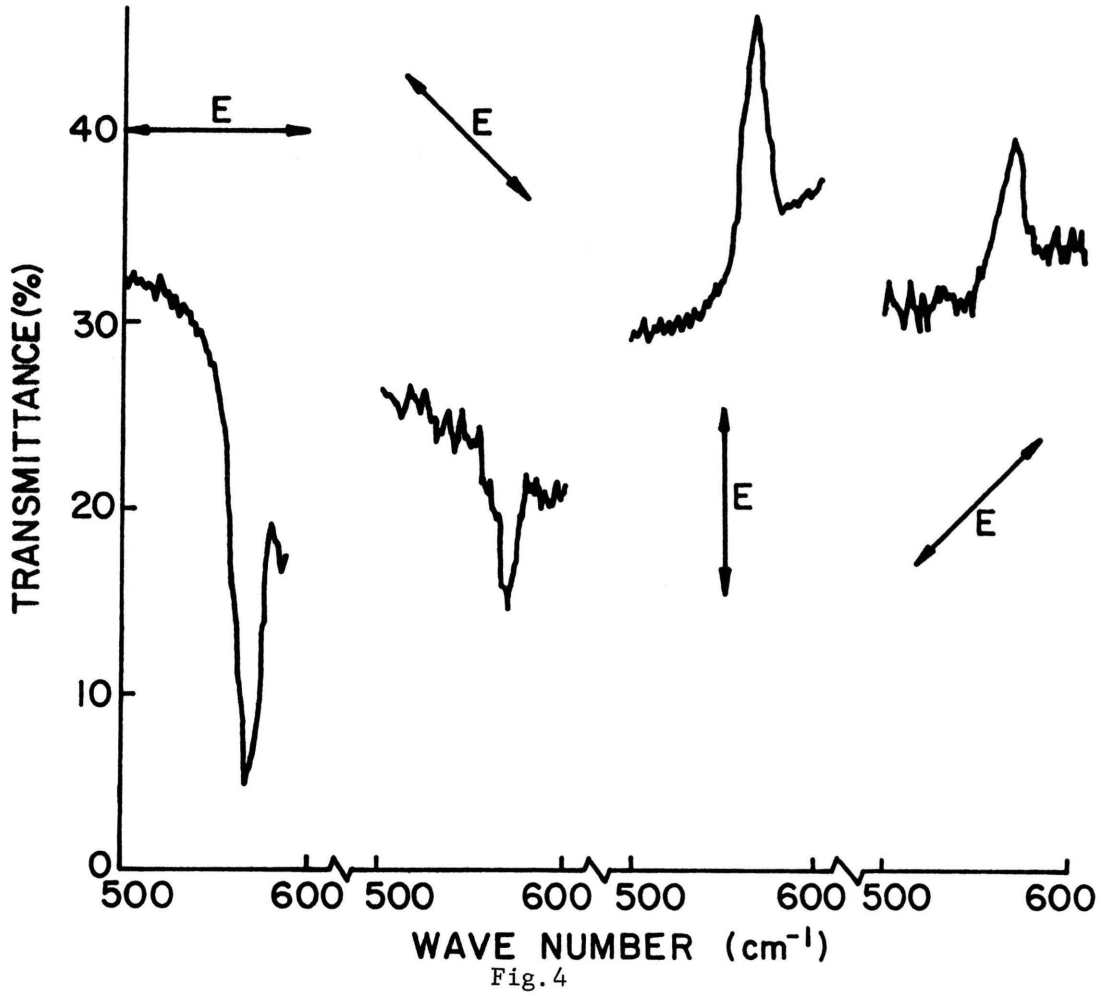

E
E
E
E
E
40
30
20
10
0
TRANSMITTANCE(%)
500
600
500
600
500
600
500
600
WAVE NUMBER (cm⁻¹)

Fig.4

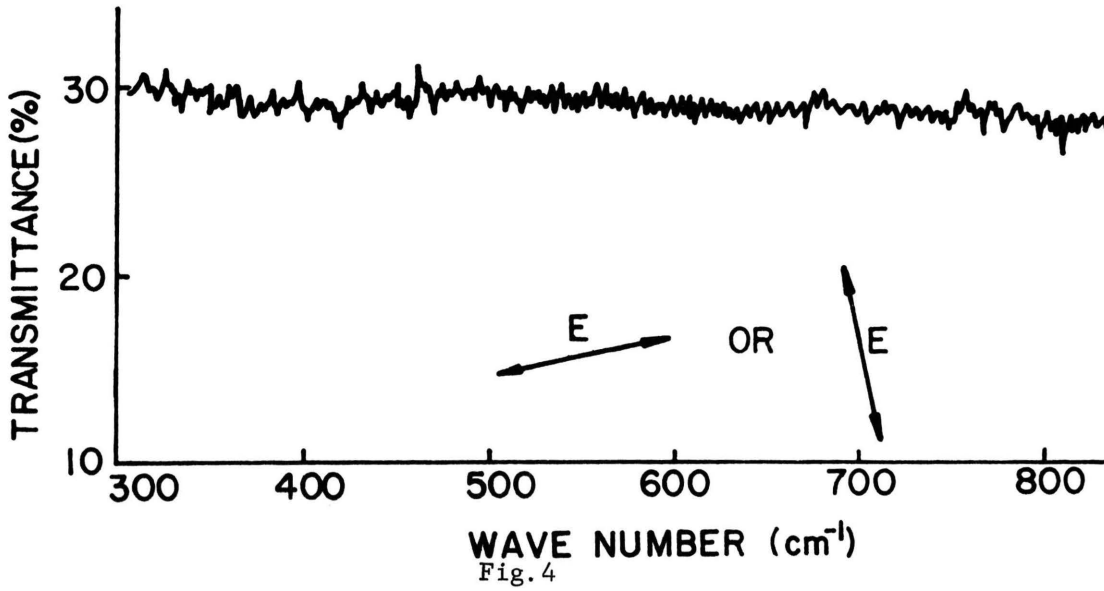
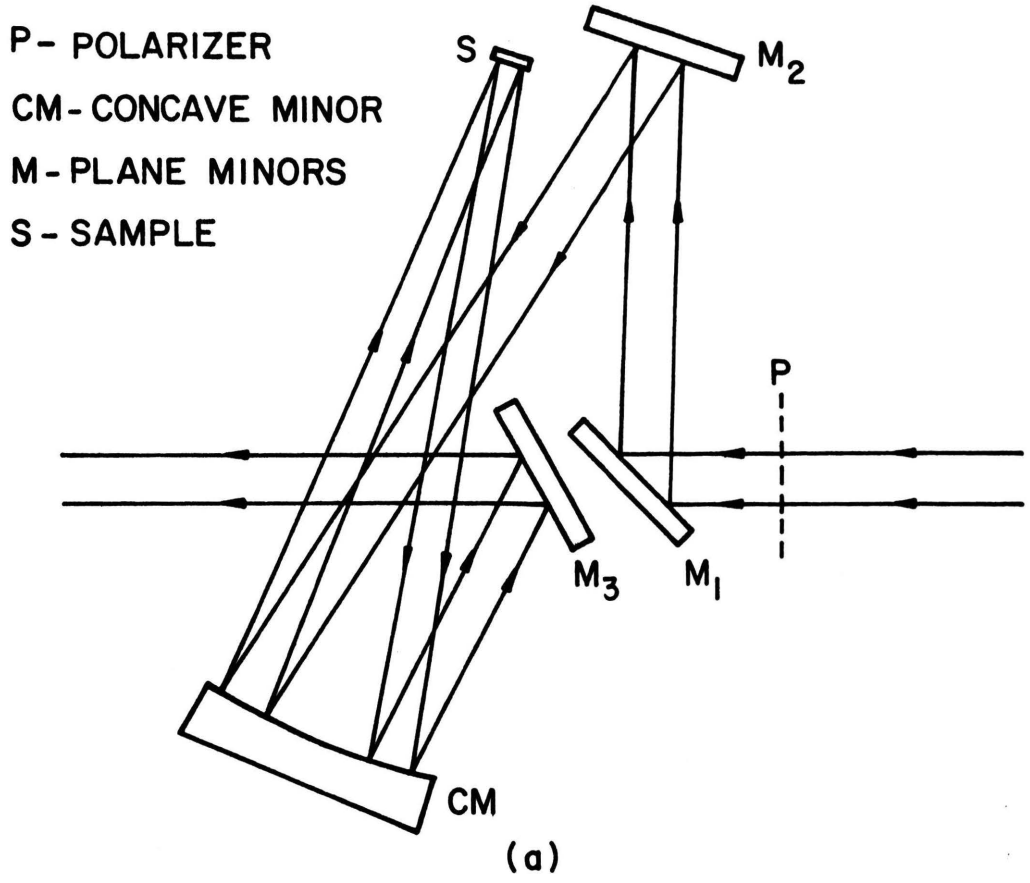

30
20
10
TRANSMITTANCE(%)
E
OR
E
300
400
500
600
700
800
WAVE NUMBER (cm⁻¹)

Fig.4

Figure 5

(a) Littrow mount arrangement for the samples.

(b) The side view of sample-concave mirror system.


P- POLARIZER
CM- CONCAVE MINOR
M- PLANE MINORS
S- SAMPLE
S
M2
P
M3
M1
CM
(a)

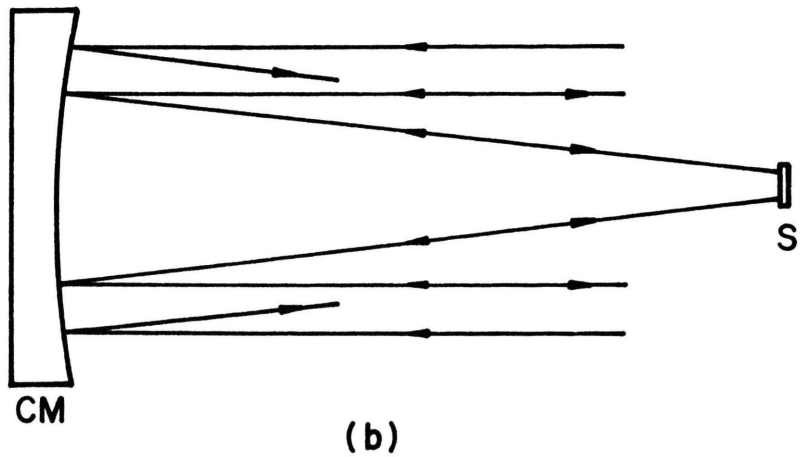

S
CM
(b)

Fig 5

Figure 6

Reflection spectra of sample A10 (GaAs, $n=4.2\times10^{18}Te/cm^3$, d = 10μm) incident angle ≈ 6°.

Curve (a) E perpendicular to the incident plane and parallel to the grating grooves.

Curve (b) E parallel to the incident plane and perpendicular to the grating grooves.

Curve (c) Spectrum of the sample with smooth surface (E perpendicular to the incident plane).

Curve (d) Spectrum of a mirror background (E perpendicular to the incident plane.

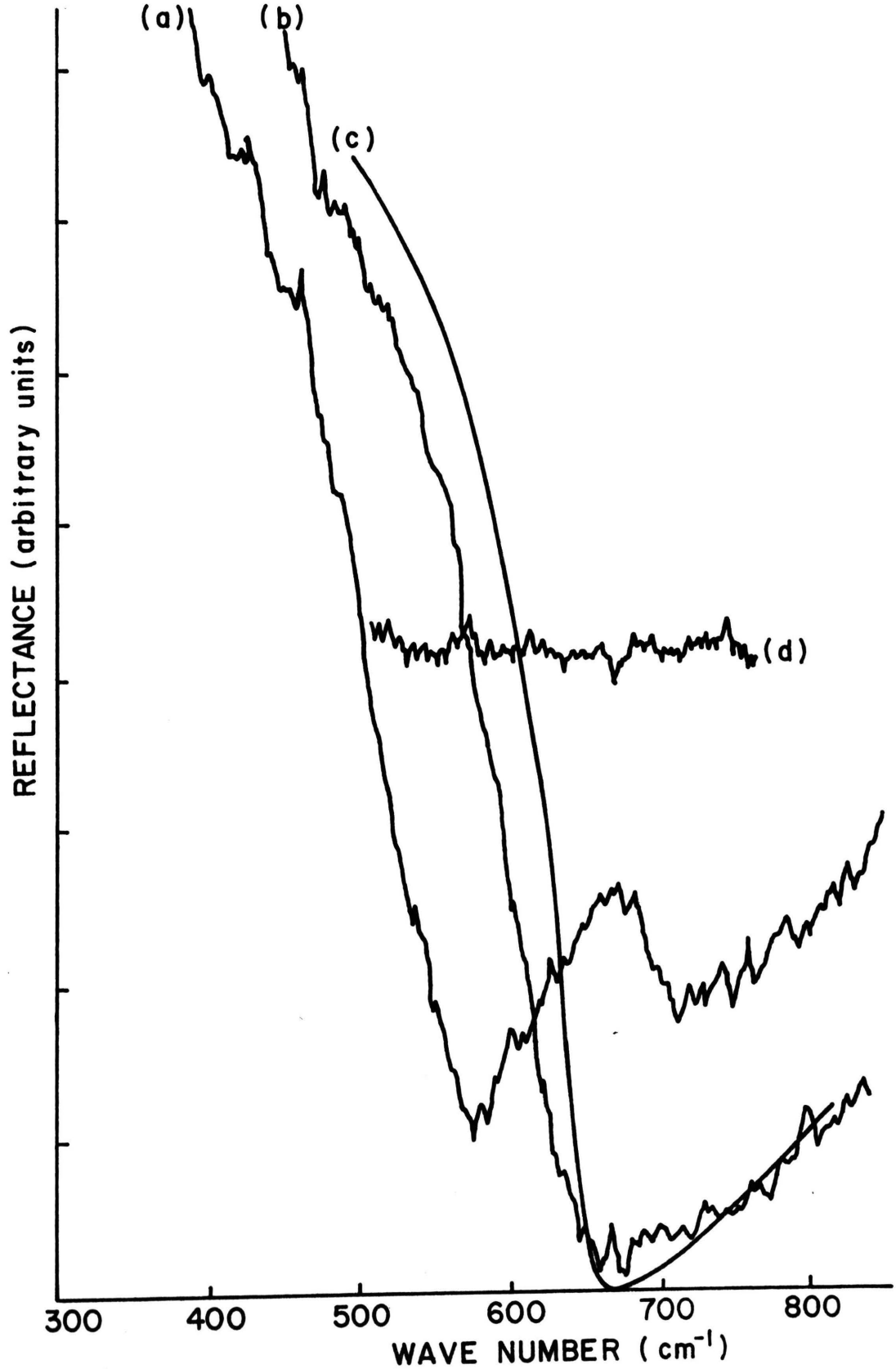

(a)
(b)
(c)
(d)
REFLECTANCE (arbitrary units)
300
400
500
600
700
800
WAVE NUMBER (cm⁻¹)

Fig 6

Figure 7

Reflection spectra of samples.

(a) B10 (InSb, n=1.1x10^{18}Te/cm^{3}, d = 10μm)
orientation same as 6(a).

(b) C10 (InSb, n=4.7x10^{17}Te/cm^{3}, d = 10μm)
orientation as in 6(b).

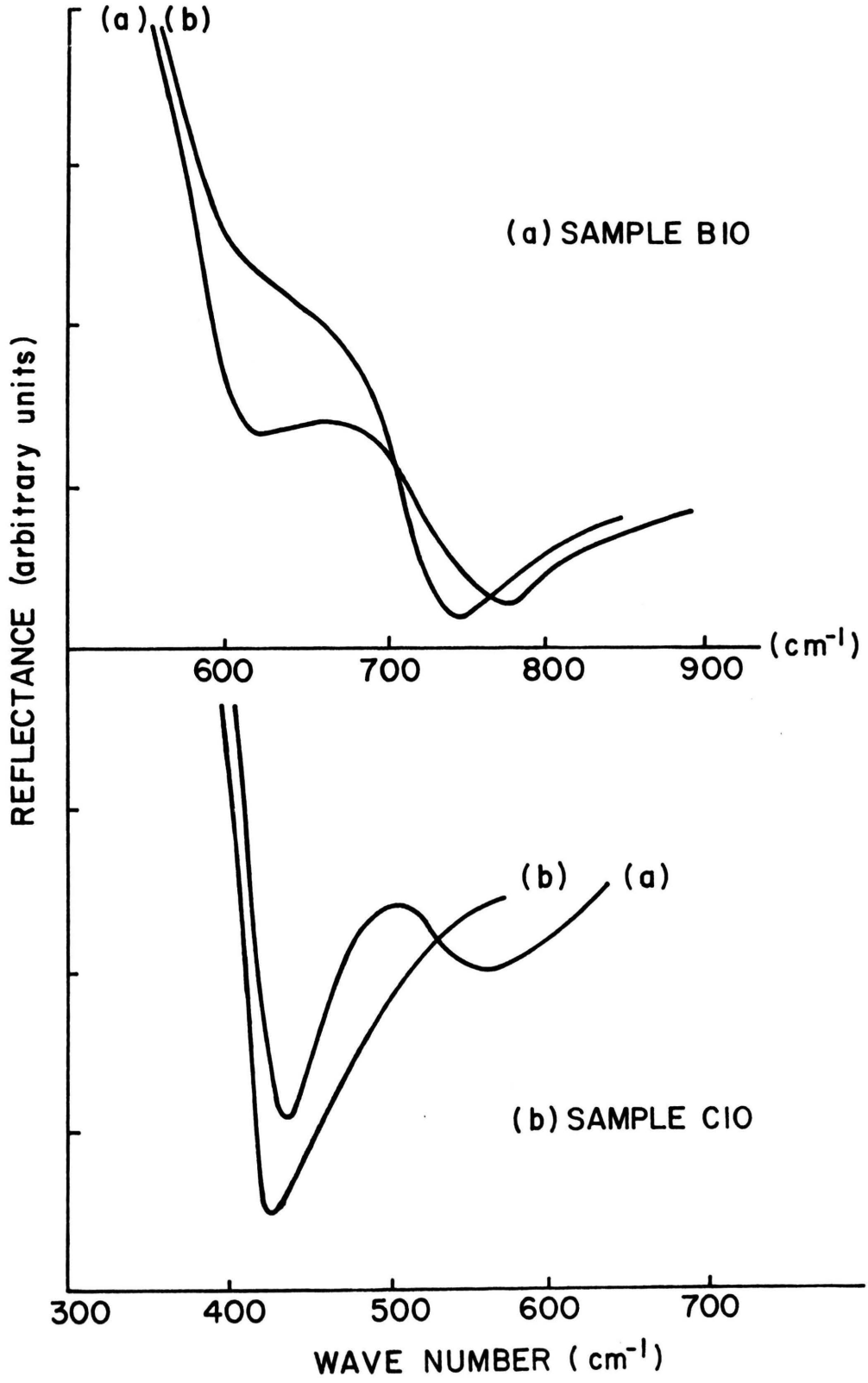

(a) (b)
(a) SAMPLE B10
REFLECTANCE (arbitrary units)
600
700
800
900
(cm⁻¹)
(b)
(a)
(b) SAMPLE C10
300
400
500
600
700
WAVE NUMBER (cm⁻¹)

Fig 7

Figure 8

Reflection spectra for different grating constants of samples A (GaAs, $n=4.2 \times 10^{18} Te/cm^3$) with E parallel to the incident plane and parallel to the grating grooves.

Curve (a) d = 10μm

Curve (b) d = 20μm

Curve (c) d = 30μm

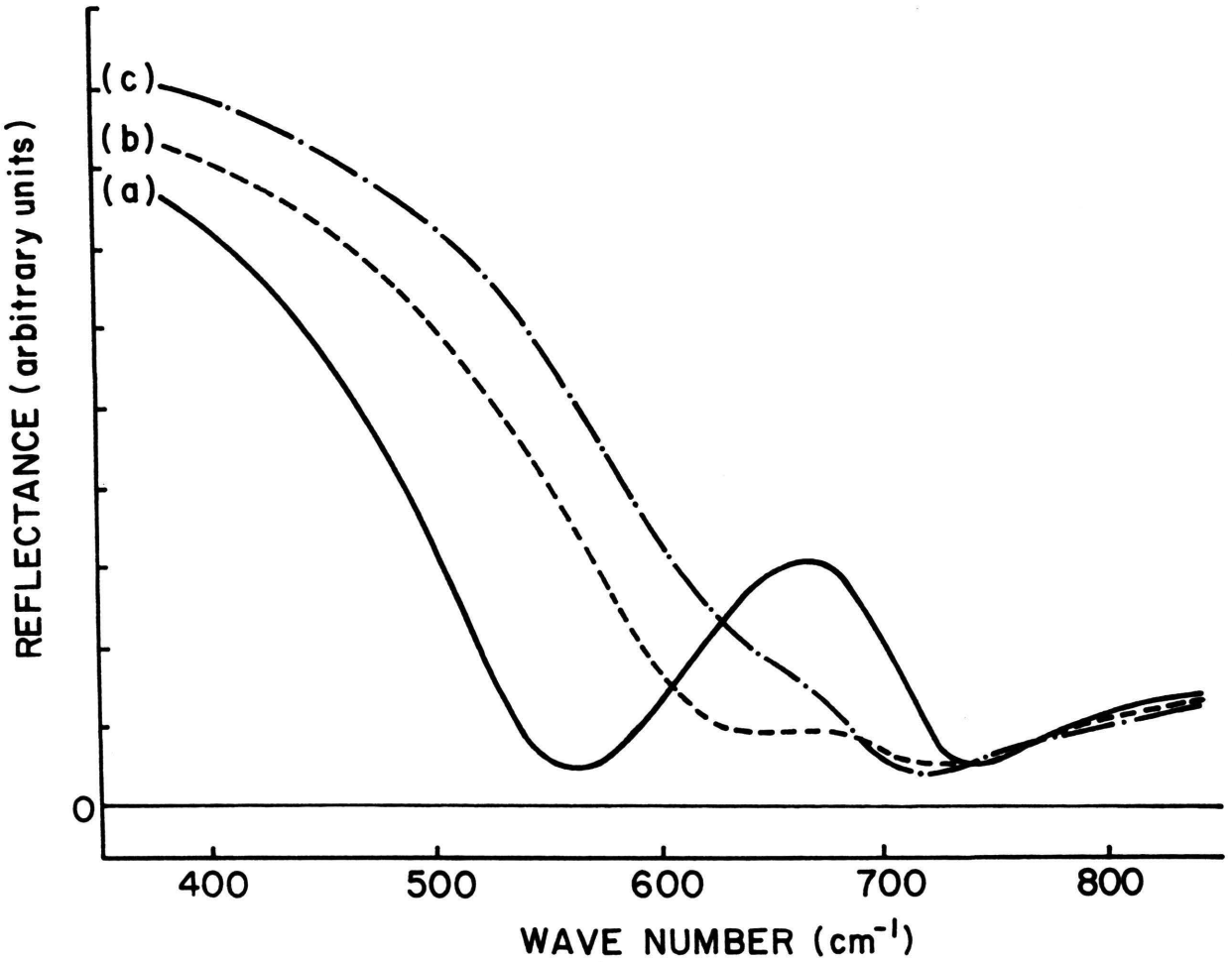

(c)
(b)
(a)
REFLECTANCE (arbitrary units)
0
400
500
600
700
800
WAVE NUMBER (cm⁻¹)

Fig 8

Figure 9

Reflection spectra from different sample conditions, A10 (GaAs, $n=4.2\times10^{18}$ Te/cm^3, d = 10μm)
E perpendicular to the incident plane and parallel to the grating grooves.
Curve (a) same curve as Fig. 6, curve (a)
Curve (b) incident angle ≈ 45°
Curve (c) grating grooves rotated an angle ≈ 60° with respect to curve (a)
Curve (d) a band pass filter placed in front of the sample
Curve (d') spectrum of the filter used for curve (d) only
(Note: Cut off the intensity beyond 1000cm^{-1}.)

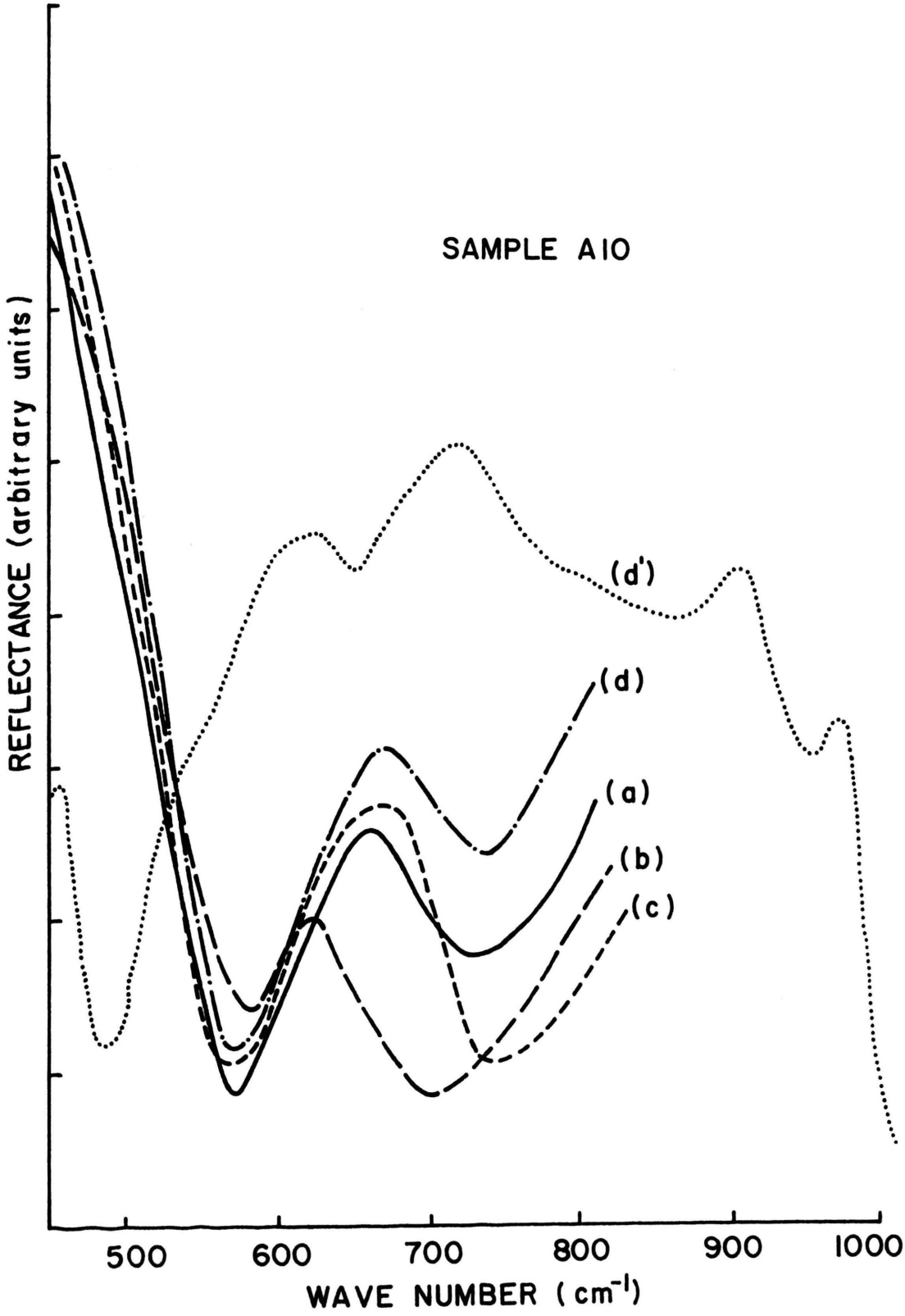
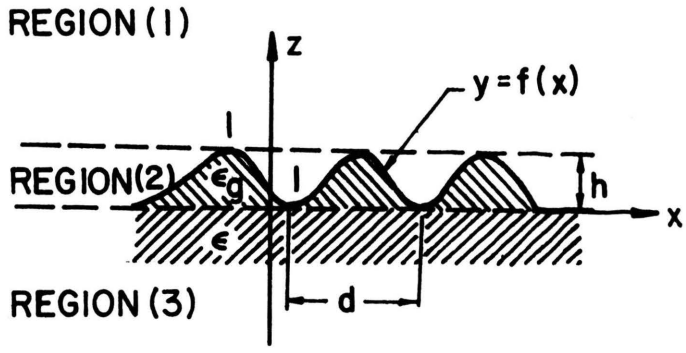

SAMPLE A10
REFLECTANCE (arbitrary units)
(d')
(d)
(a)
(b)
(c)
500
600
700
800
900
1000
WAVE NUMBER (cm⁻¹)

Fig 9

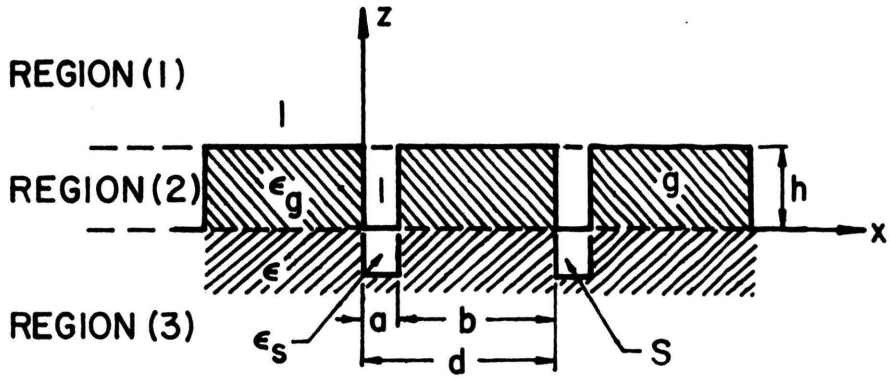
Figure 10

(a) a grating of arbitrary shape divided into three regions.

(b) same as (a), but with rectangular grooves.


REGION (1)
z
y = f(x)
l
REGION(2)
ϵ_g
l
h
x
ϵ
REGION (3)
d

(a)


z
REGION (1)
l
REGION (2)
ϵ_g
l
g
h
x
ϵ
REGION (3)
ϵ_s
a
b
d
S

(b)

Fig 10

Figure 11

Calculated reflection spectra from the rigorous grating theory.

Curve (a) $\varepsilon_g = \varepsilon_s = \varepsilon$

Curve (b) $\varepsilon_g > \varepsilon$, $\varepsilon_s = \varepsilon$

Curve (c) $\varepsilon_g > \varepsilon$, $\varepsilon_s < \varepsilon$

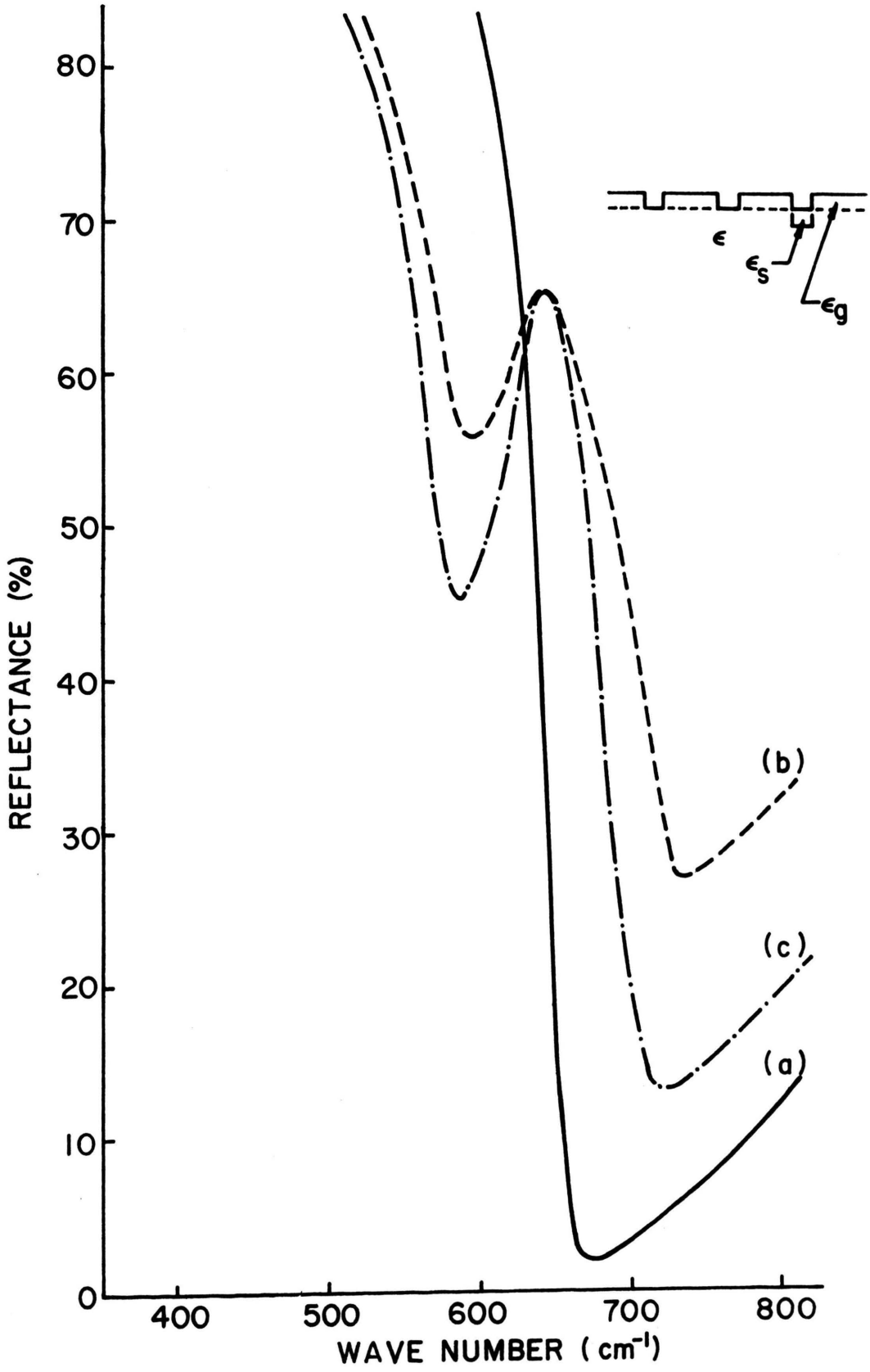

80
70
60
50
40
30
20
10
0
REFLECTANCE (%)
400
500
600
700
800
WAVE NUMBER (cm⁻¹)
ϵ
ϵs
ϵg
(b)
(c)
(a)

Fig 11

Figure 12

Some comparison of the reflection spectrum curves.

Curve (a) calculated from a smooth surface.

Curve (b) experimental spectrum from a smooth surface.

Curve (c) a curve obtained from reducing curve
Fig. 11, curve (c) by a fraction of
F = [Fig. 12, curve (b)]/[Fig. 12, curve (c)].

Curve (d) $\varepsilon_g > \varepsilon$, $\varepsilon_s < \varepsilon$, h = 5μm.

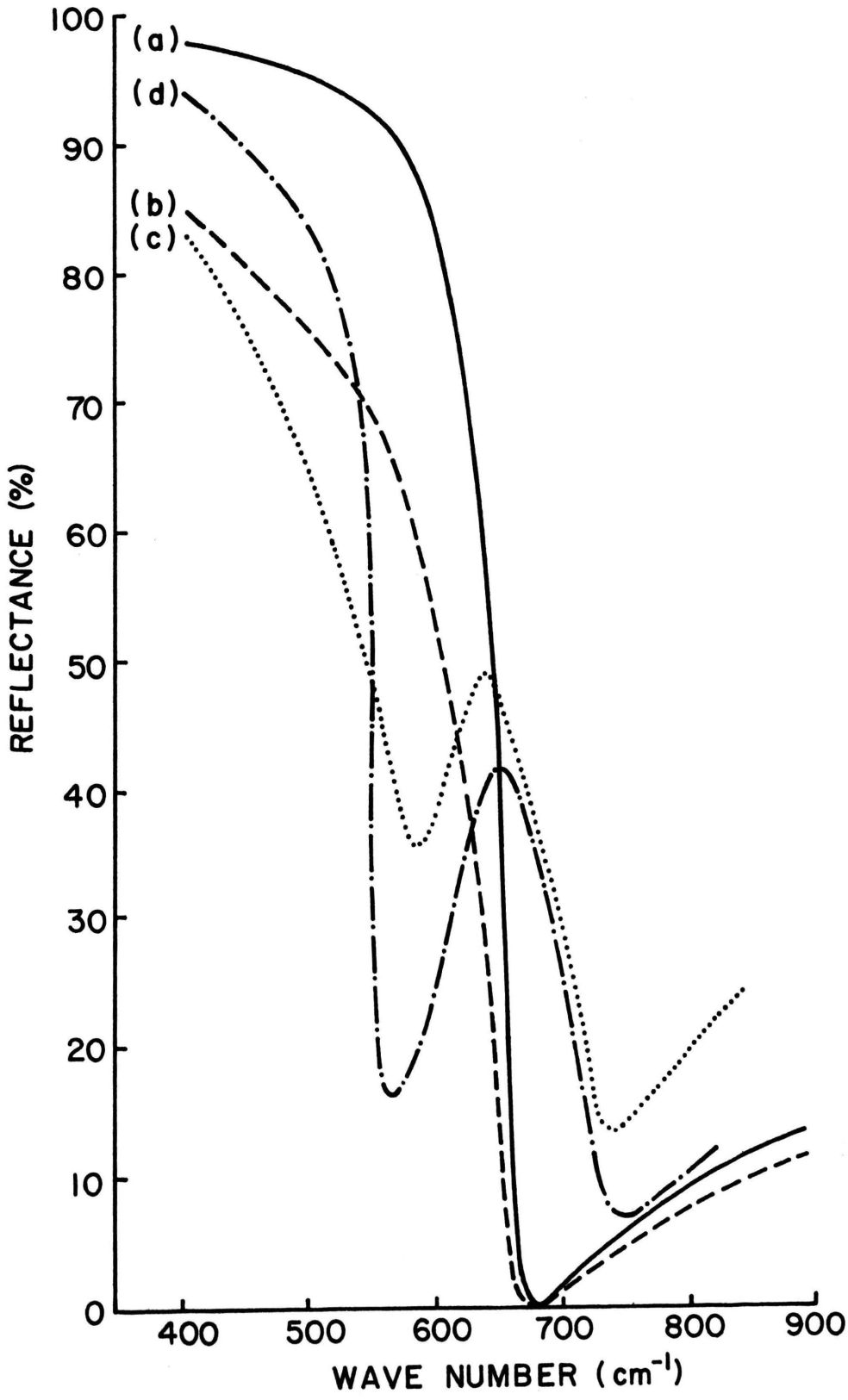

100
90
80
70
60
50
40
30
20
10
0
(a)
(d)
(b)
(c)
REFLECTANCE (%)
400
500
600
700
800
900
WAVE NUMBER (cm⁻¹)

Fig 12

Figure 13

A reflection spectrum calculated from the simple formula Eq. (43) compared with the experimental curve.

Curve (a) calculated spectrum.

Curve (b) experimental spectrum.

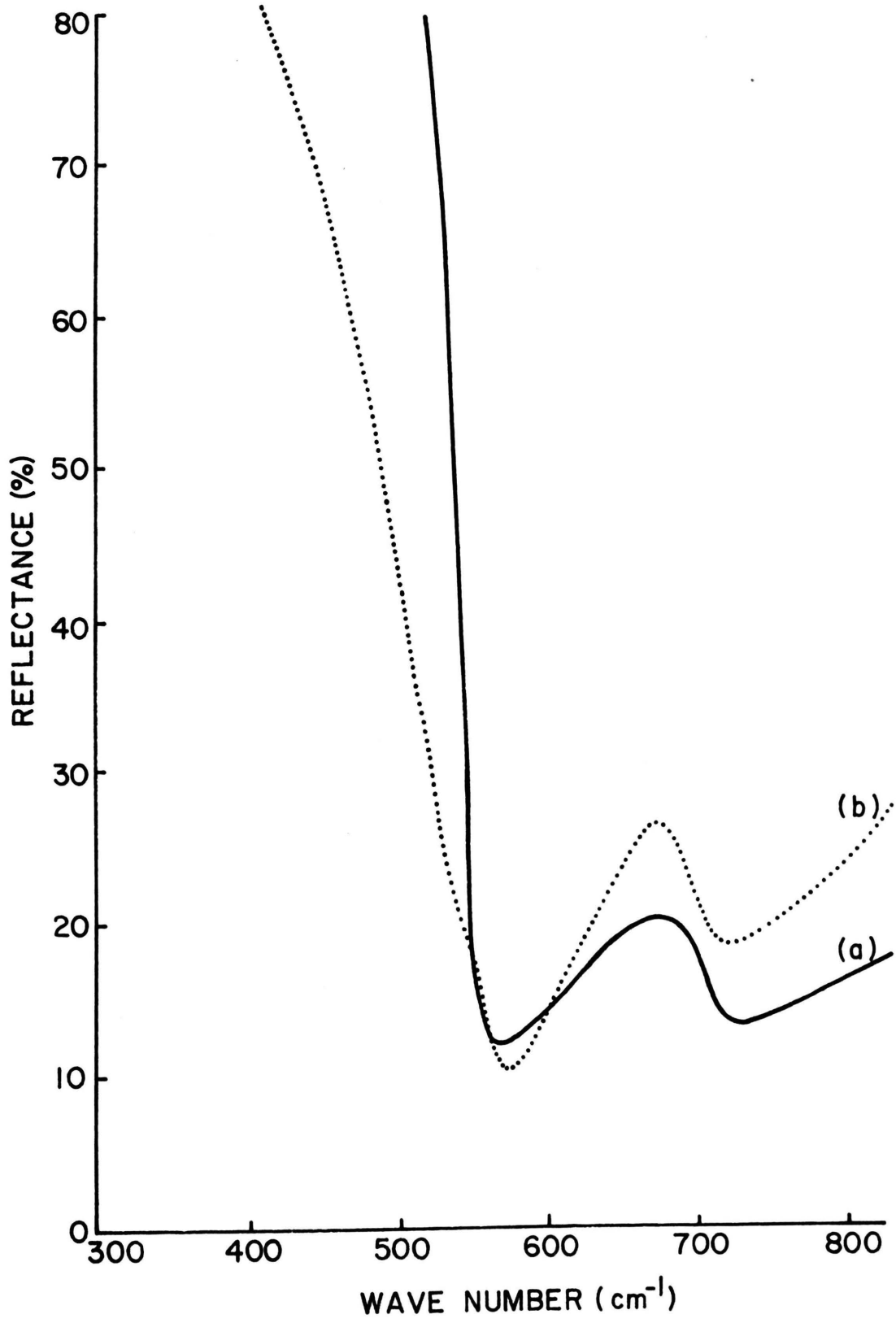

80
70
60
50
40
30
20
10
0
REFLECTANCE (%)
300
400
500
600
700
800
WAVE NUMBER (cm⁻¹)
(b)
(a)

Fig 13

Figure 14

Reflection spectra of sample A10 in the phonon region from a Fourier transform Michelson spectrometer. Incident angle ≈ 20°.
Curve (a) E perpendicular to the incident plane and parallel to the grating grooves.
Curve (b) E parallel to the incident plane and perpendicular to the grating grooves.

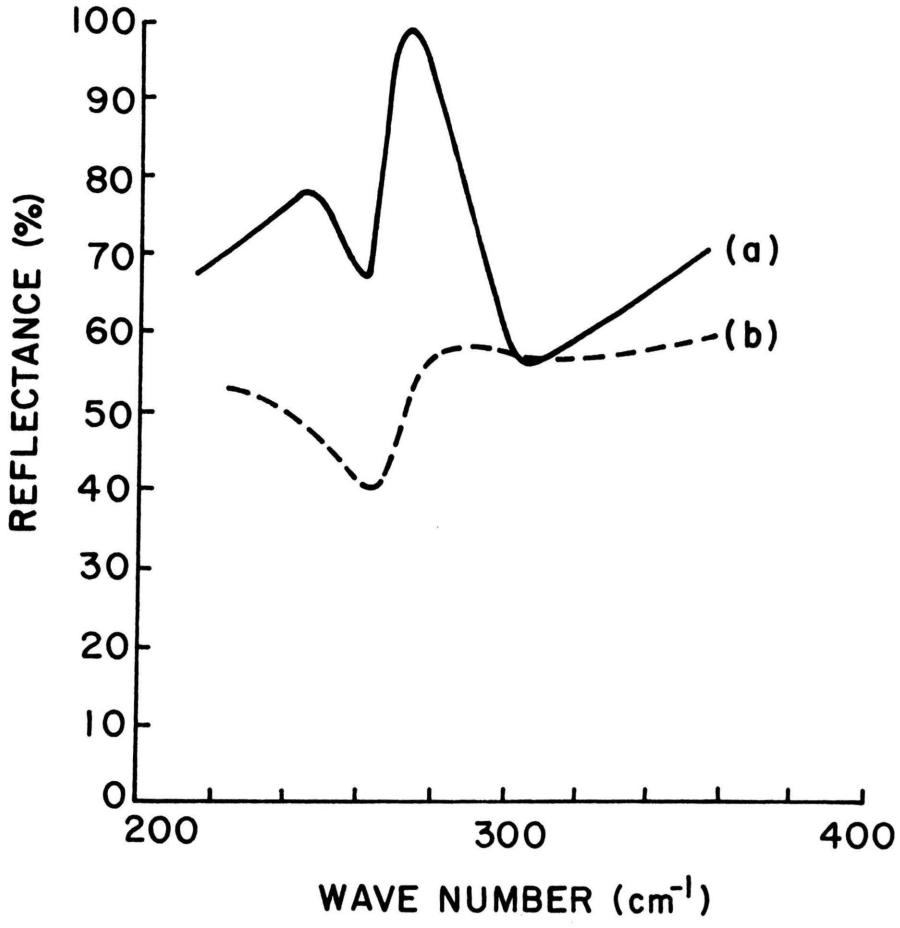

100
90
80
70
60
50
40
30
20
10
0
REFLECTANCE (%)
200
300
400
WAVE NUMBER (cm⁻¹)
(a)
(b)

Fig 14

BIBLIOGRAPHY

1. T. J. McMahon and R. J. Bell, Phys. Rev. 182, 526 (1969).

2. J. M. Elson and R. H. Ritchie, Phys. Rev. B, 4, 4129 (1971).

3. W. S. C. Chang, "Periodical Structure and Its Application to Integrated Optics", to be published.

4. R. H. Ritchie, Phys. Rev. 106, 874 (1957).

5. A. Otto, Z. Phys. 216, 398 (1968).

6. N. Marschall, B. Fischer, and H. J. Queisser, Phys. Rev. Lett. 27, 95 (1971).

7. W. E. Anderson, R. W. Alexander, Jr., and R. J. Bell, Phys. Rev. Lett. 27, 1057 (1971).

8. B. Fischer and N. Marschall, Surface Sci. 34, 845 (1973).

9. P. Beckmann and A. Spizzichino, The Scattering of Electromagnetic Waves from Rough Surfaces (Pergamon Press, New York, 1963).

10. J. H. Harris, R. K. Winn, and D. G. Dalgoutte, Appl. Optics II, 2234 (1972).

11. J. J. Cowan, E. T. Arakwa, R. H. Ritchie: The surface plasmon resonance in grating diffraction. Oak Ridge National Laboratory ORNL-TM-2615.

12. R. H. Ritchie, E. T. Arakawa, J. J. Cowan, and R. N. Hamm, Phys. Rev. Lett. 22, 1530 (1968).

13. J. J. Cowan, E. T. Arakawa, Z. Physik, 235, 97 (1970).

14. For example, R. H. Ritchie, Surface Sci. 34, 1 (1973).

15. W. F. Parks, as quoted in W. E. Anderson's Thesis.

16. R. J. Bell, R. W. Alexander, Jr., W. F. Parks and G. Kovener, Opt. Commun. 8, 147 (1973).

17. R. A. Ferrell, Phys. Rev. 111, 1214 (1958).

18. K. L. Kliewer and R. Fuchs, Phys. Rev. 153, 498 (1967).

19. W. Steinmann, Phys. Stat. Sol., 28, 437 (1968).

20. R. J. Bell, T. J. McMahon and D. G. Rathbun, Appl. Phys. 39, 48 (1968).

BIBLIOGRAPHY (cont.)

21. J. E. Stewart and W. S. Gallaway, Applied Optics, 1, 421 (1962).

22. R. H. Bgork, Y. Y. Teng, and A. S. Karakashian, Phys. Rev. Lett. 37A, 27 (1971).

23. A. Bianconi and M. Iannuzzi, Phys. Rev. Lett. 27, 932 (1971).

24. R. W. Wood, Phil. Mag. 4, 396 (1902).

25. C. H. Palmer, J. Opt. Soc. Am. 42. 269 (1952).

26. R. W. Wood, Phys. Rev. 48, 928 (1935).

27. A. Hessel and A. A. Oliner, Applied Optics, 4, 1275 (1965).

28. Lord Rayleigh, Proc. Roy. Soc. (London) A79, 399 (1907).

29. U. Fano, J. Opt. Soc. Am. 31, 213 (1941).

30. K. Artmann, Z. Phys. 119, 529 (1942).

31. B. A. Lippmann and A. Oppenheim, Tech. Res. Group N. Y. (1954).

32. J. Hagglund and F. Sellberg, J. Opt. Soc. Am. 56, 1031 (1966).

33. V. Twersky, J. Opt. Soc. Am. 52, 145 (1962).

34. J. H. Harris, R. K. Winn and D. G. Dalgoutte, Appl. Optics 11, 2234 (1972).

35. S. T. Peng, T. Tamir and H. L. Bertoni, Electronics Letters 9, 150 (1973).

36. M. Neviere, R. Petit and M. Cadilhac, Opt. Commun. 8, 113 (1973).

37. H. A. Kalhor and M. K. Moaveni, J. Opt. Soc. Am. 63, 1584 (1973).

38. R. K. Willardson and A. C. Beer, Semiconductors and Semimetals, Vol. 3 (Academic Press, New York, 1967).

39. T. J. McMahon, 1969 (unpublished).

40. R. W. Gamon and E. D. Palik, Surface polariton dispersion relation and damping rates, to be published.

VITA

The author, Le-Fu Teng, was born on December 4, 1934, in Canton, China. He received his primary education in Canton and secondary education in Taiwan. He received a Bachelor of Electrical Engineering Degree in June 1959 and a Master of Science Degree in Physics in June 1966 from the National Taiwan University in Taiwan, China.

He has been enrolled in the Graduate School of the University of Missouri-Rolla since January 1971. He was a research assistant since September 1972.