REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 10, 2003	FinalReport 08/01/99-01/31/03

4. TITLE AND SUBTITLE
Optical Interactions at Randomly Rough Surfaces

5. FUNDING NUMBERS
DAAD 19-99-1-0321

6. AUTHOR(S)
Professor Alexei A. Maradudin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Irvine

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
40364-PH

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
This project is concerned with theoretical and experimental studies of the scattering of light from and its transmission through randomly rough surfaces. The most significant results obtained during the reporting period include the design of one-dimensional random surfaces that transmit light in a specified fashion; the demonstration that the scattering of light from the random surface of an amplifying (gain) medium increases the height of and narrows the enhanced backscattering peak; the prediction of multiple-scattering effects in the second harmonic generation of light in reflection from a clean random metal surface and in transmission through a rough metal film in the Kretchmann attenuated total reflection geometry; the theoretical study of the wavelength dependence of the reflectivity and total scattered energy when p-polarized light is scattered from a rough dielectric film deposited on a metallic substrate; the development of a new perturbation theory of rough surface scattering in which the small parameter is the dielectric contrast between the medium of incidence and the scattering medium; and the observation of spectral changes (Wolf shifts) obtained from a real image of a point source.

14. SUBJECT TERMS
Random surfaces, amplifying media, surface electromagnetic waves, enhanced backscattering, spectral shifts, reflectivity, transmissivity, second harmonic generation

15. NUMBER OF PAGES
191

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

Enclosure 1

Table of Contents

Statement of Problem Studied…….3

Summary of Most Important Results……..4

List of Participating Scientific Personnel……..15

Listing of All Publications and Technical Reports……...16
***remaining publications and reports to be sent when received**

Standard Form 298……...191

4. Statement of the Problem Studied

This project is concerned with theoretical and experimental studies of the scattering of light from and its transmission through randomly rough surfaces. The problems studied include the design of one-dimensional random surfaces that transmit light in a specified fashion; the scattering of light from the random surface of an amplifying (gain) medium multiple-scattering effects in the second harmonic generation of light in reflection from a clean random metal surface and in transmission through a rough metal film in the Kretschmann attenuated total reflection geometry; the theoretical study of the wavelength dependence of the reflectivity and total scattered energy when p-polarized light is scattered from a rough dielectric film deposited on a metallic substrate; the development of a new perturbation theory of rough surface scattering in which the small parameter is the dielectric contrast between the medium of incidence and the scattering medium; and the investigation of spectral changes (Wolf shifts) obtained from a real image of a point source.

PROJECT DESCRIPTION

5. Scientific Progress and Accomplishments

(*The numbers in brackets refer to the list of papers published under ARO sponsorship during this reporting period)

A. Designer Surfaces

Optical devices that give rise to a scattered intensity that is proportional to the cosine of the scattering angle are frequently used in the optical industry, e.g. for calibrating scatterometers. Such diffusers have the property that their radiance or luminance is the same in *all* scattering directions. Due to this angular dependence such devices are often referred to as *Lambertian* diffusers. In the visible region of the optical spectrum volume disordered media, e.g. compacted powdered barium sulfate and freshly smoked magnesium oxide, are used as Lambertian diffusers. However, this type of diffuser is inapplicable in the infrared region due to its strong absorption and the presence of a specular component in the scattered light, in this frequency range. The design of a random surface that acts as a Lambertian diffuser, especially in the infrared region of the optical spectrum, is therefore a desirable goal. In a recent paper [1] a method was described for designing a one-dimensional, random surface that acts as a Lambertian diffuser. The method was tested using rigorous computer simulations, and was shown to yield the desired scattering pattern.

In a subsequent paper [2] a method for designing a two-dimensional random surface that acts as a Lambertian diffuser was described. The idea underlying the approach used in this work was that if it is desired to have a circularly symmetric distribution of the intensity of light scattered from a two-dimensional random surface when it is illuminated at normal incidence, the surface itself should be circularly symmetric. Thus, a general approach to the design of a two-dimensional random surface that has circular symmetry was presented in this work. A method of fabricating such surfaces on photoresist was described. A longer, more detailed presentation of this work was presented in Ref. [3].

An improved method for designing one-dimensional random surfaces that scatter light with a specified angular dependence of the intensity of the scattered light was presented inn Ref. [4]. A method for fabricating such surfaces on photoresist was described. A longer, more detailed presentation of this method was presented in Ref. [5].

The phenomenon of enhanced backscattering in the scattering of light from a randomly rough surface is the presence of a well-defined peak in the retroreflection direction in the angular dependence of the intensity of the light scattered diffusely from the surface. A striking feature of this phenomenon is that it occurs for any angle of incidence. Suppose, however, that one would like to have a random surface that displays enhanced backscattering for only a single, specified, angle of incidence. Such a surface could be useful, for example, in situations where one wishes to position a source (and hence the detector) at a specified

direction with respect to the site at which the scattering surface is located. In [6] it was shown how a one-dimensional random surface can be generated that produces an enhanced backscattering peak for only one specified angle of incidence when illuminated by p-polarized light whose plane of incidence is perpendicular to the generators of the surface. Numerical calculations confirmed this angular dependence of the intensity of the light scattered diffusely from this surface.

In a recent paper [7] a method was proposed for designing two-dimensional random surfaces that scatter light uniformly within a rectangular region and produce no scattering outside that region. The method was first tested by means of computer simulations. Then a procedure for fabricating such surfaces on photoresist was described, and light scattering measurements with the fabricated samples were presented. The results validated the design procedure, and showed that the fabrication method is feasible.

B. Band-Limited Diffusers in Transmission

The transmission of light through a structure consisting of vacuum in the region $x_3 > \zeta(x_1)$, a dielectric film characterized by a real, positive, dielectric constant ϵ in the region $-D < x_1 < \zeta(x_1)$, and vacuum in the region $x_3 < -D$ was studied theoretically [8]. The surface profile function $\zeta(x_1)$ was assumed to be a single-valued function of x_1, that is differentiable, and constitutes a random process. This structure was illuminated from the region $x_3 > \zeta(x_1)$ by s-polarized light whose plane of incidence was the x_1x_3-plane. By the use of the geometrical optics limit of phase perturbation theory it was shown how to design the surface profile function $\zeta(x_1)$ in such a way that the mean differential transmission coefficient has a prescribed form within a specified range of the angle of transmission, and vanishes outside this range. In particular, the case in which the transmitted intensity was a constant within a specified range of the angle of transmission , and vanished outside it, was considered. Rigorous numerical simulation calculations showed that the transmitted intensity indeed had this property.

In a subsequent paper [9] the same problem was studied for a more general system consisting of a dielectric medium characterized by a dielectric constant ϵ_1 in the region $x_3 > H$, a second dielectric medium characterized by a dielectric constant ϵ_2 in the region $\zeta(x_1) < x_3 < H$, and vacuum in the region $x_3 < \zeta(x_1)$. This structure was illuminated from the region $x_3 > H$ by s-polarized light whose plane of incidence was the x_1x_3-plane. The surface profile function $\zeta(x_1)$ was designed in such a way that the mean differential transmission coefficient had a specified form as a function of the angle of transmission. The particular case where the incident light was incident normally on this structure, and the intensity of the transmitted light was a constant within a specified range of the angle of transmission and vanished outside this range, was considered. Rigorous numerical simulations showed that the scattered light also had a constant intensity within a certain (different) range of the scattering angle and vanished outside this range.

C. Angular Intensity Correlation Functions

An experimental study was carried out of the angular intensity correlation function of far-field speckle patterns scattered in the double passage of waves through a one-dimensional random phase screen [10]. The theoretical analysis of the symmetry of speckle patterns around the backscattering direction, and the motion of the speckles as the source was moved, carried out by Escamilla *et al.* [Appl. Opt. **32**, 2734 (1993)], were verified.

In two recent theoretical studies (B. Shapiro, Phys. Rev. Lett. **83**, 4733 (1999), S. E. Skipetrov and R. Maynard, Phys. Rev. B**62**, 886 (2000)) of angular intensity correlation functions of light scattered from volume-disordered systems, a new contribution, in addition to the known $C^{(1)}$, $C^{(10)}$, $C^{(1.5)}$, $C^{(2)}$ and $C^{(3)}$ correlations, was predicted. This new long-range correlation function was labeled the C_0 correlation function and, in contrast with the $C^{(1)}$, $C^{(10)}$, $C^{(1.5)}$, $C^{(2)}$ and $C^{(3)}$ correlation functions, is non-universal: its value depends on the specific type of disorder in the random medium. It describes the correlation in the immediate vicinity of the source and/or of the detector, and provides information on the details of the disorder in these regions. In [11] the contribution C_0 to the normalized angular intensity correlation function for the scattering of p- and s-polarized light from the weakly rough one-dimensional random interface between a medium characterized by a dielectric constant ϵ_0 (the medium of incidence) and a medium characterized by a dielectric constant ϵ (the scattering medium) was determined. It was shown that the C_0 contribution comes from the correlations of the elastic initial or final single-scattering processes, and can be the leading contribution to the long-range correlation function $C^{(2)}$. Experimental conditions under which the $C^{(1)}$ and $C^{(10)}$ correlation functions vanish, and C_0 is much larger than the $C^{(1.5)}$, $C^{(2)}$, and $C^{(3)}$ correlations, were indicated. Under these conditions C_0 is the dominant contribution to the normalized angular intensity correlation functions. These conditions should serve as a guide to experimentalists seeking to observe this new correlation function.

In subsequent theoretical work [12] two computer simulation studies of the speckle correlations in the light scattered from a volume disordered dielectric medium consisting of a random array of dielectric spheres were carried out. In both studies, light was treated in the scalar wave approximation, and the wavelength of the light was assumed to be much greater than the radius of the spheres. In one study the scattering medium was formed by placing the spheres randomly in space. The spheres occupied space uniformly under the provision that no two spheres overlap. In the second study the scattering medium was formed by placing the spheres randomly on the vertices of a simple cubic lattice, so that a fixed fraction of the vertices was occupied by the spheres. The lattice constant of the simple cubic lattice was taken to be of the order of magnitude of the wavelength of light in vacuum. In both studies the volume filling fraction was the same and the region outside the spheres was vacuum. The equation for the wave field was solved numerically to determine the scattered field, and the result was used to calculate the angular intensity correlation function. Particular attention was paid to regions of angles of incidence and scattering angles in which either the $C^{(1)}$ or $C^{(10)}$ contributions dominate the correlation function. In the case where the dielectric spheres were placed randomly on the sites of a simple cubic lattice, so that the resulting system was periodic on average rather than uniform, structure was observed in the

$C^{(1)}$ and $C^{(10)}$ contributions that was caused by scattering processes involving one or more Bragg reflections.

The statistical properties of the scattering matrix $S(q|k)$ were studied in the scattering of light of frequency ω from a randomly rough one-dimensional surface whose surface profile function constituted a zero-mean, stationary, Gaussian random process [13]. This was done by studying the effects of $S(q|k)$ on the angular intensity correlation function $C(q,k|q',k') = \langle I(q|k)I(q'|k')\rangle - \langle I(q|k)\rangle\langle I(q'|k')\rangle$, where the intensity $I(q|k)$ is defined in terms of $S(q|k)$ by $I(q|k) = L_1^{-1}(\omega/c)|S(q|k)|^2$, with L_1 the length of the x_1-axis covered by the random surface. Attention was focused on the $C^{(1)}$ and $C^{(10)}$ correlation functions, which are the contributions to $C(q,k|q',k')$ that are proportional to $\delta(q-k-q'+k')$ and $\delta(q-k+q'-k')$, respectively. The existence of both of these correlation functions is consistent with the amplitude of the scattered field obeying complex Gaussian statistics in the limit of a long surface and in the presence of weak surface roughness. It was shown that the deviation of the statistics of the scattering matrix from complex circular Gaussian statistics and the $C^{(10)}$ correlation function are determined by exactly the same statistical moment of $S(q|k)$. As the random surface becomes rougher, the amplitude of the scattered field no longer obeys complex Gaussian statistics but obeys complex circular Gaussian statistics instead. In this case the $C^{(10)}$ correlation function should therefore vanish. This result is confirmed by numerical simulation calculations.

D. The Scattering of Light from the Random Interface between Two Dielectric Media with Low Contrast.

The coherent and incoherent scattering of p- and s-polarized light incident from a dielectric medium characterized by a real, positive, dielectric constant ϵ_0 onto its one-dimensional, randomly rough, interface with a dielectric medium characterized by a real, positive, dielectric constant ϵ, has been calculated by means of a perturbation theory with a new small parameter, namely the dielectric constant $\eta = \epsilon_0 - \epsilon$ between the medium of incidence and the scattering medium [14]. The proper self-energy entering the expression for the reflectivity was obtained as an expansion in powers of η though the second order in η, and the reducible vertex function in terms of which the scattering intensity is expressed was obtained as an expansion in powers of η through the fourth. The roughness-induced shifts of the Brewster angle (in p-polarization) and of the critical angle for total internal reflection ($\epsilon_0 > \epsilon$) were obtained. The angular dependence of the intensity of the incoherent component of the scattered light was found to display an enhanced backscattering peak, which is due to the coherent interference of multiply-scattered lateral waves supported by the interface and their reciprocal partners. Analogues of the Yoneda peaks observed in the scattering of x-rays from solid surfaces are also present. The results of this small-contrast perturbation theory were found to be in good agreement with those obtained in computer simulation studies.

E. Spectral Changes of Light on its Scattering from a Random Surface

Coherence theory predicts that the correlation in the fluctuations of a source distribution can cause frequency shifts in the spectrum of the emitted radiation, even when the source is at rest relative to the observer. Recently, the Wolf effect, or frequency shifts from a real image of a point source, was measured [15], and the coherent interference effect of a finite-band source was verified.

In a subsequent experimental study [16] the angular spectrum redistribution, or frequency shifts, in the scattering of light from 1D and 2D surfaces illuminated by a laser source was measured, and an additional verification of the spectral and spatial coherent interference effect of a finite-band source was provided.

F. Enhanced Backscattering at Grazing Angles

The backscattering signal at small grazing angles is very important for vehicle re-entry and missile tracking applications. It is also useful in FTIR grazing angle microscopy. Recently, an experimental study was performed of the far-field scattering at small grazing angles, especially the enhanced backscattering at grazing angles [17]. For a weakly rough random dielectric film on a reflecting metal substrate, an enhanced backscattering peak that is higher than the background at its position by a factor significantly larger than two was observed.

G. Wavelength Dependence of the Light Scattered from a Dielectric Film Deposited on a Metal Substrate.

The scattering of p-polarized light from a system consisting of a dielectric film deposited on a semi-infinite metal (silver) was studied theoretically with an emphasis on the wavelength dependence of the total integrated scattering and angle resolved scattering from this system [18]. In particular, the reflectivity $R(\lambda)$ and the total scattered energy $U(\lambda)$ were determined as functions of the wavelength of the incident light by a rigorous numerical simulation approach. Both surfaces of the dielectric film could be one-dimensional random surfaces. It was found that the total scattered energy was reduced the most by making the vacuum-film interface rough while at the same time keeping the film-metal interface planar.

H. An Effective Impedance Boundary Condition for the Coherent Scattering of Light from a One-Dimensional Randomly Rough Surface.

As a continuation of our studies of impedance boundary conditions that can be useful in theoretical studies of the scattering of light from random surfaces, an impedance boundary condition was derived [19] for a system of vacuum separated from a metal by a one-dimensional interface defined by $x_1 = \zeta(x_1)$, that relates the mean value $\langle L(x_1|\omega)\rangle$ to the mean value

$\langle H(x_1|\omega)\rangle$, where $L(x_1|\omega) = [\partial H_2^>(x_1, x_3|\omega)/\partial x_3]_{x_3=0}$ $(= [\partial E_2^>(x_1, x_3|\omega)/\partial x_3]_{x_3=0})$ and $H(x_1|\omega)$ $= [H_2^>(x_1, x_3|\omega)]_{x_3=0} (= [E_2^>(x_1, x_3|\omega)]_{x_3=0})$, where $H_2^>(x_1, x_3|\omega)$ $(E_2^>(x_1, x_3|\omega))$ is the total magnetic (electric) field in the vacuum region in the case of a p-polarized (s-polarized) electromagnetic field whose plane of incidence is the x_1x_3-plane. The angle brackets here denote an average over the ensemble of realizations of the surface profile function $\zeta(x_1)$. The result was used to calculate the reflectivity of the metal surface in both polarizations, and yielded results in reasonable agreement with those obtained by the use of phase perturbation theory.

I. Enhanced Backscattering of Light from the Random Surface of an Amplifying Medium

An experimental study was carried out of the enhanced backscattering of light from a randomly rough surface through a laser dye-doped polymer [20]. The sample was a slice of a pyrromethene-doped polymer coupled with a two-dimensional rough gold layer with a large slope. When the sample was illuminated with an s-polarized He-Ne laser and was pumped by a cw argon-ion laser, amplified scattering was observed. The enhanced backscattering peak increased sharply, and its width narrowed for a sample with a small (negative) value of the imaginary part of its dielectric constant.

In a second experimental study of the scattering of light from the same system [21], the satellite peaks supported by the scattering structure were increased as well as the enhanced backscattering peak.

In [22] the scattering of s-polarized light from a rough dielectric film deposited on the plane surface of a semi-infinite perfect conductor was investigated by numerical simulations. The dielectric film was allowed to be either active or passive, situations that were modeled by assigning negative and positive values, respectively, to the imaginary part ϵ_2 of the dielectric constant of the film. The reflectance $\mathcal{R}$ and the total scattered energy $\mathcal{U}$ for the system were studied as functions of both ϵ_2 and the angle of incidence of the light. In addition, the positions and widths of the enhanced backscattering and satellite peaks were studied. It was found that these peaks become narrower and higher when the amplification of the system is increased, and that their widths are linear functions of ϵ_2. The position of the backscattering peak was found to be independent of ϵ_2, while a weak dependence of the positions of the satellite peaks on ϵ_2 was found.

J. Multiple-Scattering Effects in the Generation of Second Harmonic Light in Scattering from or Transmission Through a Random Metal Surface.

Perturbative calculations of the second harmonic light generated in the transmission of p-polarized light through a thin metal film with a one-dimensional random surface in the Kretschmann attenuated total reflection geometry have been carried out [23]. The metal film is deposited on the planar surface of a prism through which the light is incident. The back surface of the film is a one-dimensional random surface whose generators are perpendicular

to the plane of incidence. It is in contact either with a semi-infinite vacuum or a semi-infinite nonlinear crystal (quartz). It was found that when the random surface separates the metal film from vacuum so that the nonlinearity of the film surfaces gives rise to the harmonic light, for a general angle of incidence a dip appears in the angular dependence of the intensity of the transmitted harmonic light in the direction normal to the mean surface. When the second harmonic generation is due to the nonlinearity of the crystal in contact with the metal film, a peak in the angular dependence of the intensity of the transmitted harmonic light occurs in this direction. These dips and peaks are multiple-scattering effects. However, when the angle of incidence is optimal for the excitation of surface plasmon polaritons at the film-vacuum/nonlinear crystal interface, the nonlinear mixing of the incident light and the backward propagating surface plasmon polariton leads to an intense peak in the angular dependence of the intensity of the transmitted harmonic light in the direction normal to the mean surface. This peak is already present in the single-scattering approximation.

In [24] a computer simulation approach was used to study the generation of second harmonic light in reflection and transmission in the Kretschmann attenuated total reflection geometry in the case where the back, randomly rough, face of the metal film was in contact with vacuum. The nonlinearity responsible for the second harmonic generation was assumed to arise at the prism-metal and metal-vacuum interfaces, and thus entered the problem only through the boundary conditions at these interfaces at the harmonic frequency. The source terms entering these boundary conditions were obtained from the solutions of the corresponding scattering and transmission problems at the fundamental frequency. It was found that a peak in the angular dependence of the intensity of both the transmitted and reflected second harmonic light occurs in the directions normal to the mean scattering surface, in addition to an enhanced backscattering peak in the retroreflection direction. The enhanced transmission peak occurs in the non-radiative region, and therefore cannot be observed in the far field.

K. Surface Electromagnetic Waves

In a recent theoretical study of the scattering of a surface plasmon polariton by a circularly symmetric protuberance or indentation on an otherwise planar metal surface in contact with vacuum (A. V. Shchegrov, I. V. Novikov, and A. A. Maradudin, Phys. Rev. Lett. **78**, 4269-4272 (1997)), it was found that the angular dependence of the intensity of the volume electromagnetic waves scattered into the vacuum region possesses a maximum in the plane of incidence at a polar scattering angle of approximately 28°. This suggests that if a p-polarized volume electromagnetic field in the form of a beam of finite width is incident on the same surface defect, the efficiency of exciting a surface plasmon polariton will be greatest for a polar angle of incidence close to 28°. To test this hypothesis the latter problem was studied theoretically [25]. The reduced Rayleigh equations for the amplitudes of the p- and s-polarized components of the scattered field were reduced to a set of one-dimensional integral equations by exploiting the circular symmetry of the surface defect, which was assumed to have a Gaussian form. The efficiency of exciting surface plasmon polaritons in this fashion

was calculated as a function of the polar angle of incidence, and was found to be maximal when this angle is close to 28°.

The possibility of the strong (Anderson) localization of surface plasmon polaritons propagating along a metal surface, a finite part of which is randomly rough, was studied theoretically [26]. The surface roughness assumed was such that the roughness-induced conversion of a surface plasmon polariton propagating on it into volume electromagnetic waves in the vacuum above the surface was suppressed. Estimates of the Anderson localization length of the surface plasmon polariton were obtained analytically and numerically.

By means of Green's second integral identity in the plane, and a parameterization of the surface profile function, the dispersion relation for surface electromagnetic waves guided by a straight channel cut into the otherwise planar surface of a solid in contact with vacuum was determined numerically [27]. The solid could be either a metal or a dielectric medium, and was characterized by an isotropic, real, frequency-dependent dielectric function $\epsilon(\omega)$ that is negative in some range of frequencies. The resulting propagating bound electromagnetic modes were named *channel polaritons*.

L. Left-Handed Media

Recently a physical medium was fabricated (D. R. Smith *et al.*, Phys. Rev. Lett. **84**, 4184 (2000)) in which both the effective permittivity $\epsilon(\omega)$ and the effective permeability $\mu(\omega)$ are negative over a restricted frequency range. Thus, in this frequency range, such a medium is "left-handed," and is characterized by a negative refractive index. A left-handed medium should exhibit unusual phenomena associated with the propagation and scattering of electromagnetic waves. Thus, in a recent paper [28] the scattering of p- and s-polarized electromagnetic waves from a weakly rough one-dimensional random surface of a left-handed medium was studied theoretically. It was assumed that the surface profile function was a single-valued function of the coordinate in the mean plane of the surface that is normal to its grooves and ridges, and constitutes a zero-mean, stationary, Gaussian random process. It was shown that in contrast to nonmagnetic media with a negative dielectric function, the planar surface of a left-handed medium can support both p- and s-polarized surface electromagnetic waves. The reflectivity of such a surface as a function of the angle of incidence displays structure associated with the existence of a Brewster angle and a critical angle for total internal reflection in both polarizations. The angular distribution of the intensity of the light that has been scattered incoherently displays an enhanced backscattering peak, and Yoneda bands, for both polarizations of the incident light.

M. Review Articles.

During the reporting period four major review articles were written .

The first [29] was devoted to a survey of recent theoretical work on multiple-scattering phenomena in the scattering of light from randomly rough surfaces. The phenomena discussed were enhanced backscattering, enhanced transmission, satellite peaks, changes in the

spectrum of light scattered from a random surface, and the design of random surfaces that scatter light uniformly within a specified range of scattering angles and produce no scattering outside this range.

The second article [30] was devoted to a survey of recent theoretical work on multiple-scattering effects in the second harmonic light generated at clean, randomly rough, one-dimensional metal surfaces. Both weakly rough and strongly corrugated surfaces were considered, and comparisons between theoretical and existing experimental results were presented. The agreement was quite good.

A more extensive treatment of this problem in which second harmonic generation in the Kretschmann attenuated total reflection geometry was also studied, was presented in the third review article [31].

The fourth article [31] dealt with theoretical and experimental aspects of the scattering of light from randomly rough surfaces and from random phase screens.

6. List of publications and technical reports supported under this grant.

[**1**] A. A. Maradudin, I. Simonsen, T. A. Leskova, and E. R. Méndez, "Design of one-dimensional Lambertian diffusers of light," Waves in Random Media **11**, 529-533 (2001).

[**2**] A. A. Maradudin, T. A. Leskova, and E. R. Méndez, "Two-dimensional random surfaces that act as circular diffusers," Opt. Lett. **28**, 72-74 (2003).

[**3**] A. A. Maradudin, E. R. Méndez, and T. A. Leskova, "Two-Dimensional Random Surfaces that Act as Circular Diffusers," SPIE **4780**, 51-60 (2002).

[**4**] E. R. Méndez, E. E. García-Guerrero, T. A. Leskova, A. A. Maradudin, J. Muñoz-Lopez, and I. Simonsen," The design of one-dimensional random surfaces with specified scattering properties," Appl. Phys. Lett. **81**, 798-800 (2002).

[**5**] E. R. Méndez, E. E. García-Guerrero, T. A. Leskova, A. A. Maradudin, J. Muñoz-Lopez, and I. Simonsen, "The Design of One-Dimensional Random Surfaces with Specified Scattering Properties," SPIE **4780**, 40-50 (2002).

[**6**] M. Ciftan, T. A. Leskova, and A. A. Maradudin, "The fabrication of one-dimensional random surfaces that display enhanced backscattering for only one specified angle of incidence," SPIE **4447**, 17-23 (2001).

[**7**] E. R. Méndez, E. E. García-Guerrero, H. M. Escamilla, A. A. Maradudin, T. A. Leskova, and A. V. Shchegrov, "Photofabrication of random achromatic optical diffusers for uniform illumination," Appl. Opt. **40**, 1098-1108 (2001).

[**8**] T. A. Leskova, A. A. Maradudin, E. R. Méndez, and I. Simonsen, "Band-limited uniform diffuser in transmission," SPIE **4100**, 113-123 (2000).

[**9**] A. A. Maradudin, E. R. Méndez, T. A. Leskova, and I. Simonsen, "A band-limited uniform diffuser in transmission. II. A multilayer system," SPIE **4447**, 130-139 (2001).

[**10**] Zu-Han Gu, "Speckle correlations around the backscattering direction in the double-passage configuration," Appl. Opt. **39**, 4684-4689 (2000).

[**11**] T. A. Leskova and A. A. Maradudin, "The angular intensity correlation function C_0 for the scattering of light from a one-dimensional randomly rough surface," Phil. Mag. B**81**, 1289-1302 (2001).

[**12**] A. R. McGurn and A. A. Maradudin, "Computer simulation studies of the speckle correlations of light scattered from a random array of scatterers: scalar wave approximation," Phys. Rev. B**64**, 165204-1 - 165204-14 (2001).

[**13**] T. A. Leskova, I. Simonsen, and A. A. Maradudin, "The angular intensity correlation functions $C^{(1)}$ and $C^{(10)}$ for the scattering of light from randomly rough dielectric and metal surfaces, "Waves in Random Media **12**, 307-319, (2002).

[**14**] T. A. Leskova, A. A. Maradudin, and I. V. Novikov, "The scattering of light from the random interface between two dielectric media with low contrast," J. Opt. Soc. Am. A**17**, 1288-1300 (2000).

[**15**] Zong-Qi Lin and Zu-Han Gu, "Wolf effect from a real image as a secondary source," SPIE **4100**, 78-83 (2000); Zu-Han Gu, "Angular spectrum redistribution from rough surface scattering," SPIE **4447**, 109-114 (2001).

[**16**] Zu-Han Gu, "Spectral coherence from rough surface scattering with a secondary source," SPIE **4447**, 115-121 (2001).

[**17**] Zu-Han Gu and M. Ciftan, "Enhanced backscattering at grazing angle," SPIE **4447**, 1-5 (2001).

[**18**] I. Simonsen, T. A. Leskova, A. A. Maradudin, and O. Hunderi, "Wavelength dependence of the light scattered from a dielectric film deposited on a metal substrate," SPIE **4100**, 65-77 (2000).

[**19**] T. A. Leskova and A. A. Maradudin, "An effective impedance boundary condition for the coherent scattering of light from a one-dimensional randomly rough surface," SPIE **4100**, 46-56 (2000).

[**20**] Zu-Han Gu and G. D. Peng, "Amplification of enhanced backscattering from a dye-doped polymer bounded by a rough surface," Opt. Lett. **25**, 375-377 (2000).

[**21**] G. D. Peng and Zu-Han Gu, "Amplified backscattering from dye-doped polymer bounded with 1-D rough metal film," SPIE **4100**, 103-112 (2000).

[**22**] I. Simonsen, T. A. Leskova, and A. A. Maradudin, "Light scattering from an amplifying medium bounded by a randomly rough surface: a numerical study," Phys. Rev. B**64**, 035425-1 - 035425-7, (2001).

[**23**] T. A. Leskova, M. Leyva-Lucero, E. R. Méndez, A. A. Maradudin, and I. V. Novikov, "The surface enhanced second harmonic generation of light from a randomly rough metal surface in the Kretschmann geometry," Opt. Commun. **183**, 529-545 (2000).

[**24**] I. V. Novikov, A. A. Maradudin, T. A. Leskova, E. R. Méndez, and M. Leyva-Lucero, "Second harmonic generation of light in the Kretschmann attenuated total reflection geometry in the presence of surface roughness," Waves in Random Media **11**, 183-231 (2001).

[**25**] M. Kretschmann, T. A. Leskova, and A. A. Maradudin, "Excitation of surface plasmon polaritons by the scattering of volume electromagnetic beam from a circularly symmetric defect on a planar metal surface," SPIE **4447**, 24-33 (2001).

[**26**] A. A. Maradudin, I. Simonsen, T. A. Leskova, and E. R. Méndez, "Localization of surface plasmon polaritons on a random surface," Physica B**296**, 85-97 (2001).

[**27**] I. V. Novikov and A. A. Maradudin, "Channel Polaritons," Phys. Rev. B**66**, 035403-1 - 035403-13 (2002).

[**28**] T. A. Leskova, A. A. Maradudin, and I. Simonsen, "Scattering of electromagnetic waves from the random surface of a left-handed medium," SPIE **4447**, 6-16 (2001).

[**29**] E. I. Chaikina, E. E. Garcia-Guerrero, Zu-Han Gu, T. A. Leskova, A. A. Maradudin, E. R. Méndez, and A. V. Shchegrov, "Multiple-scattering phenomena in the scattering of light from randomly rough surfaces," In *Frontiers of Laser Physics and Quantum Optics*," eds. Z. Xu, S. Xie, S. Y. Zhu, and M. O. Scully (Springer-Verlag, New York, 2000), pp. 225-259.

[**30**] E. R. Méndez, M. Leyva-Lucero, T. A. Leskova, and A. A. Maradudin, "Multiple-Scattering effects in the second harmonic light generated at randomly rough metallic interfaces," in *Frontiers of Laser Physics and Quantum Optics*, eds. Z. Xu, S. Xie, S. Y. Zhu, and M. O. Scully (Springer-Verlag, New York, 2000) pp. 149-171.

[**31**] A. A. Maradudin and E. R. Méndez, "Scattering by Surfaces and Phase Screens," In *Scattering*, eds. R. Pike and P. Sabatier (Academic Press, New York, 2002), pp. 864-894.

[**32**] T. A. Leskova, A. A. Maradudin, and E. R. Méndez, "Multiple-scattering phenomena in the second harmonic generation of light reflected from and transmitted through randomly rough metal surfaces," in *Optical Properties of Nanostructured Random Media*, ed. V. M. Shalaev (Springer-Verlag, New York, 2002) pp. 359-442.

[**33**] A. A. Maradudin, "Comment on 'Radiative transfer over small distances from a heated metal'," Opt. Lett. **26**, 479 (2001).

7. Scientific personnel supported by this project

Dr. T. A. Leskova

Dr. W. Zierau

Dr. Zu-Han Gu

Dr. E. R. Méndez

Mr. I. Simonsen (received the Ph.D. degree)

M. M. Kretschmann

INSTITUTE OF PHYSICS PUBLISHING WAVES IN RANDOM MEDIA

Waves Random Media **11** (2001) 529–533 PII: S0959-7174(01)18801-4

Design of one-dimensional Lambertian diffusers of light

A A Maradudin[1], I Simonsen[1,2], T A Leskova[3] and E R Méndez[4]

[1] Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, CA 92697, USA
[2] Department of Physics, The Norwegian University of Science and Technology, N-7491, Trondheim, Norway
[3] Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Russia
[4] División de Física Aplicada, CICESE, Ensenada, Baja CA 22800, Mexico

Received 9 November 2000, in final form 10 May 2001
Published 21 September 2001
Online at stacks.iop.org/WRM/11/529

Abstract

We describe a method for designing a one-dimensional random surface that acts as a Lambertian diffuser. The method is tested using rigorous computer simulations and is shown to yield the desired scattering pattern.

Optical devices that give rise to a scattered intensity that is proportional to the cosine of the scattering angle are frequently used in the optical industry, e.g. for calibrating scatterometers [1]. Such diffusers have the property that their radiance or luminance is the same in *all* scattering directions. Due to this angular dependence such devices are often referred to as *Lambertian diffusers*. In the visible region of the optical spectrum volume disordered media, e.g. compacted powdered barium sulfate, and freshly smoked magnesium oxide [2] are used as Lambertian diffusers. However, this type of diffuser is inapplicable in the infrared region due to its strong absorption and the presence of a specular component in the scattered light, in this frequency range.

The design of a random surface that acts as a Lambertian diffuser, especially in the infrared region of the optical spectrum, is therefore a desirable goal, and one that has been regarded as difficult to achieve [3]. In this paper we present a solution to this problem that is based on an approach used in several recent papers to design one-dimensional random surfaces with specified scattering properties [4–6], and to fabricate them in the laboratory [5, 7]. The design of a two-dimensional random surface that acts as a Lambertian diffuser will be described elsewhere [8].

To motivate the calculations that follow we begin by considering the scattering of s-polarized light of frequency ω from a one-dimensional, randomly rough, perfectly conducting surface defined by $x_3 = \zeta(x_1)$. The region $x_3 > \zeta(x_1)$ is vacuum, the region $x_3 < \zeta(x_1)$ is the perfect conductor (figure 1). The plane of incidence is the x_1x_3-plane. The surface profile function $\zeta(x_1)$ is assumed to be a single-valued function of x_1 that is differentiable, and to constitute a random process.

 © 2001 IOP Publishing Ltd Printed in the UK

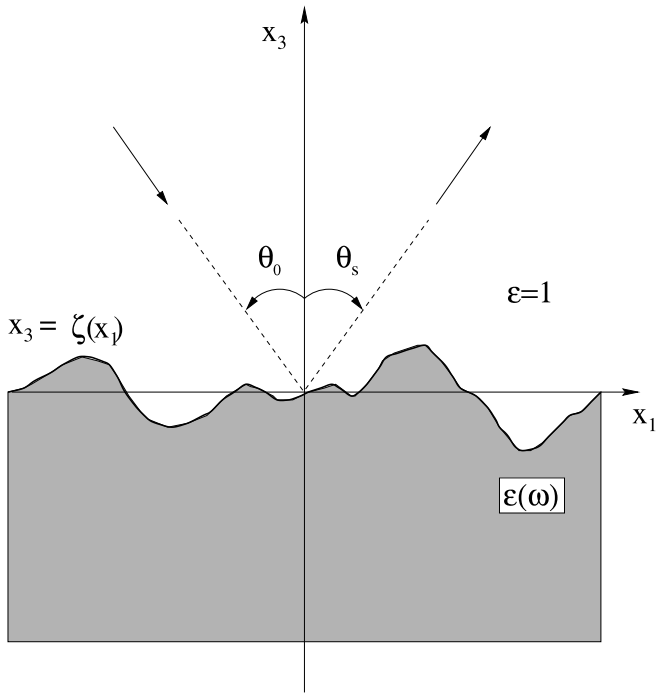

Figure 1. The scattering geometry assumed in this paper.

The mean differential reflection coefficient $\langle \partial R/\partial\theta_s \rangle$, where the angle brackets denote an average over the ensemble of realizations of the surface profile function, is defined such that $\langle \partial R/\partial\theta_s \rangle \, \mathrm{d}\theta_s$ is the fraction of the total time-averaged flux incident on the surface that is scattered into the angular interval $(\theta_s, \theta_s + \mathrm{d}\theta_s)$ in the limit as $\mathrm{d}\theta_s \to 0$. In the geometrical optics limit of the Kirchhoff approximation it is given by [5]

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle = \frac{1}{L_1} \frac{\omega}{2\pi c} \frac{1}{\cos\theta_0} \left[\frac{1 + \cos(\theta_0 + \theta_s)}{\cos\theta_0 + \cos\theta_s} \right]^2 \int_{-\infty}^{\infty} \mathrm{d}x_1 \int_{-\infty}^{\infty} \mathrm{d}u \, \exp[\mathrm{i}(q-k)u] \times \langle \exp[\mathrm{i}au\zeta'(x_1)] \rangle. \tag{1}$$

In this expression L_1 is the length of the x_1-axis covered by the random surface, θ_0 and θ_s are the angles of incidence and scattering, respectively, $a = (\omega/c)(\cos\theta_0 + \cos\theta_s)$, and $q = (\omega/c)\sin\theta_s$, $k = (\omega/c)\sin\theta_0$. In the following, we will restrict ourselves to the case of normal incidence ($\theta_0 = 0^\circ$), in which case equation (1) simplifies to

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle = \frac{1}{L_1} \frac{\omega}{2\pi c} \int_{-\infty}^{\infty} \mathrm{d}x_1 \int_{-\infty}^{\infty} \mathrm{d}u \, \exp \mathrm{i}qu \langle \exp[\mathrm{i}au\zeta'(x_1)] \rangle \tag{2}$$

where a is now given by $a = (\omega/c)(1 + \cos\theta_s)$.

We wish to find a surface profile function $\zeta(x_1)$ for which the mean differential reflection coefficient has the form

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle = \frac{1}{2} \cos\theta_s. \tag{3}$$

To this end we write $\zeta(x_1)$ in the form [5]

$$\zeta(x_1) = \sum_{\ell=-\infty}^{\infty} c_\ell \, s(x_1 - \ell 2b). \tag{4}$$

Here the c_ℓ are independent, positive, random deviates, b is a characteristic length, and the function $s(x_1)$ is defined by [5]

$$s(x_1) = \begin{cases} 0 & x_1 \leqslant -(m+1)b \\ -(m+1)bh - hx_1 & -(m+1)b \leqslant x_1 \leqslant -mb \\ -bh & -mb \leqslant x_1 \leqslant mb \\ -(m+1)bh + hx_1 & mb \leqslant x_1 \leqslant (m+1)b \\ 0 & (m+1)b \leqslant x_1 \end{cases} \tag{5}$$

where m is a positive integer. Such trapezoidal grooves can be generated experimentally [5,7].

Since the c_ℓ are positive random deviates, their probability density function (PDF) $f(\gamma) = \langle \delta(\gamma - c_\ell) \rangle$ is non-zero only for positive values of γ.

It has been shown [5] that when the surface profile function is given by equations (4) and (5), the expression (2) for the mean differential reflection coefficient becomes

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle = \frac{1}{4h} \left(1 + \tan^2 \frac{\theta_s}{2} \right) \left[f\left(-\frac{1}{h} \tan \frac{\theta_s}{2} \right) + f\left(\frac{1}{h} \tan \frac{\theta_s}{2} \right) \right]. \tag{6}$$

Thus, we find that in the geometrical optics limit of the Kirchhoff approximation the mean differential reflection coefficient is determined by the PDF of the coefficients c_ℓ entering the expansion (4), and is independent of the wavelength of the incident light. If we make the change of variable $\tan(\theta_s/2) = \gamma h$, $0 \leqslant \gamma h \leqslant 1$, so that $\frac{1}{2}\cos\theta_s = \frac{1}{2}(1 - \gamma^2 h^2)/(1 + \gamma^2 h^2)$, on combining equations (3) and (6) we find that the equation determining $f(\gamma)$ is

$$f(-\gamma) + f(\gamma) = 2h \frac{1 - \gamma^2 h^2}{(1 + \gamma^2 h^2)^2}. \tag{7}$$

It follows that

$$f(\gamma) = 2h \frac{1 - \gamma^2 h^2}{(1 + \gamma^2 h^2)^2} \theta\left(\frac{1}{h} - \gamma \right) \theta(\gamma). \tag{8}$$

The preceding results were obtained in the geometrical optics limit of the Kirchhoff approximation for a perfectly conducting surface. However, our earlier experience in designing surfaces with specified scattering properties [4–6] shows that when a surface designed on the basis of these assumptions is ruled on a lossy metal, the results of rigorous scattering calculations show that the resulting scattering pattern retains the form prescribed in the approximate, single-scattering calculations. We now demonstrate that such a result is obtained in the context of the present problem.

From the form of $f(\gamma)$ given in equation (8) a long sequence of c_ℓ was generated by applying the rejection method [9], and the resulting surface profile function $\zeta(x_1)$ was generated using equations (4) and (5). We found from numerical experiments that in order to have a surface that acts as a Lambertian diffuser in reflection the parameter b had to be large. Physically, this means that the grooves $\zeta(x_1)$ have to be wide.

In figure 2 we present the results of rigorous numerical Monte Carlo simulations [10] for the angular dependence of the mean differential reflection coefficient $\langle \partial R/\partial \theta_s \rangle$ for s-polarized incident light of wavelength $\lambda = 612.7$ nm scattered from a randomly rough silver surface of the type described above (noisy curve). The value of the dielectric constant of silver at this wavelength is $\epsilon(\omega) = -17.2 + \mathrm{i}0.5$. The surface was characterized by the parameters $b = 80\lambda = 49\ \mu$m, $h = 0.2$ and $m = 1$, and its length used in the simulation was $L_1 = 164\lambda = 100\ \mu$m. Furthermore, the plot in figure 2 was obtained by averaging the results for $N_\zeta = 35\,000$ realizations of the surface profile function $\zeta(x_1)$. Such a large number of surface realizations was needed in order to reduce the noise level sufficiently. The reason

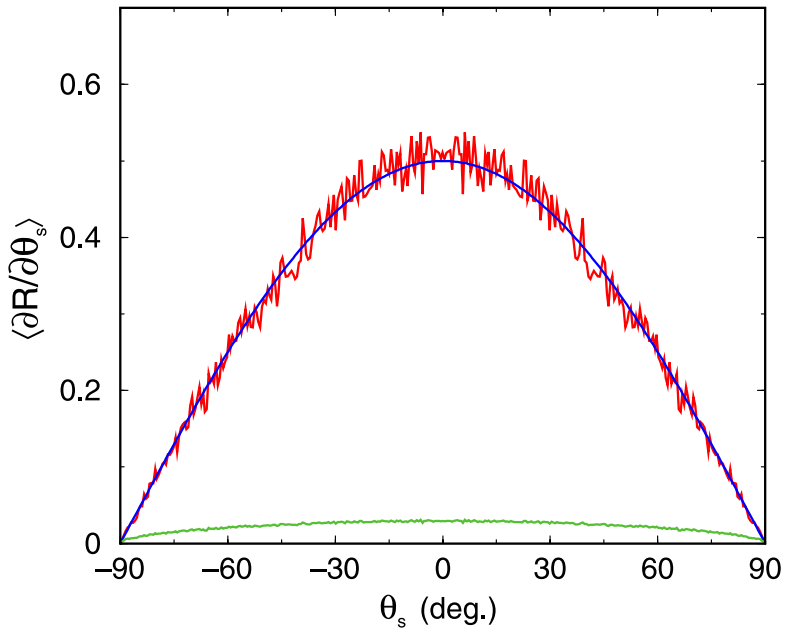

Figure 2. The noisy curve is $\langle \partial R/\partial \theta_s \rangle$ calculated by a numerical simulation approach for a random silver surface defined by equations (4) and (5) with $b = 80\lambda$, $h = 0.2$, $m = 1$, and the PDF (8), when s-polarized light of wavelength $\lambda = 612.7$ nm ($\epsilon(\omega) = -17.2 + i0.5$) is incident normally on it. The upper full curve is $\langle \partial R/\partial \theta_s \rangle$ given by equation (3). The lower full curve is the error in the calculated mean differential reflection coefficient as measured by its standard deviation.

(This figure is in colour only in the electronic version)

for the slow convergence of the mean DRC with increasing N_ζ we believe is due to the large value of b used in the simulations. Without compromising the spatial discretization used in the numerical calculation ($\Delta x_1 = 0.164\lambda$) needed in order to resolve the oscillations of the incident field, only a few grooves $s(x_1)$ could be included for each realization in the sum (4) defining the surface.

The lower smooth curve represents an estimate of the error in the calculated $\langle \partial R/\partial \theta_s \rangle$ due to the use of a finite number of surface realizations for its calculation. This error is obtained as the standard deviation of the mean differential reflection coefficient (see [10] for details).

The upper smooth full curve in figure 2 represents the geometrical optics limit of the Kirchhoff approximation, equation (3). As can be readily observed from this figure, the agreement between the geometrical optics limit of the Kirchhoff approximation for a random perfectly conducting surface and the result of rigorous numerical simulations for a real random silver surface is excellent within the noise level. This is indeed the case for all scattering angles θ_s, which we find somewhat surprising, since one might have expected the geometrical optics approximation to break down for the largest scattering angles. That this is not observed in our simulation results is probably an indication that multiple scattering processes are of minor importance in the scattering taking place at the random surface even for the largest scattering angles.

Simulations (results not shown) were also performed where the wavelength of the incident light was changed by plus and minus 10% from its original value of $\lambda = 612.7$ nm. Such changes did not affect the Lambertian nature of the scattered light in any significant way.

This weak wavelength sensitivity is consistent with our earlier experience in designing surfaces with specified scattering properties [4–6]. Surfaces generated on the basis of different b parameters have also been considered. We found that the scattered intensity showed little sensitivity to this parameter as long as it is large.

Acknowledgments

The work of AAM and TAL was supported by Army Research Office Grant DAAD 19-99-1-0321. IS would like to thank the Research Council of Norway (contract no 32690/213) and Norsk Hydro ASA for financial support. The work of ERM was supported by CONACYT Grant 3804P-A. This work has also received support from the Research Council of Norway (Program for Supercomputing) through a grant of computing time.

References

[1] Stover J C 1995 *Optical Scattering* (Bellingham WA: SPIE)
[2] Grum F and Luckey G W 1968 Optical sphere paint and a working standard of reflectance *Appl. Opt.* **7** 2295–300
[3] Baltes H P 1980 Progress in inverse optical problems *Inverse Scattering Problems in Optics* ed H P Baltes (New York: Springer) pp 1–13
[4] Leskova T A, Maradudin A A, Novikov I V, Shchegrov A V and Méndez E R 1998 Design of one-dimensional band-limited uniform diffusers of light *Appl. Phys. Lett.* **73** 1943–5
[5] Méndez E R, Martinez-Niconoff G, Maradudin A A and Leskova T A 1998 Design and synthesis of random uniform diffusers *SPIE* **3426** 2–13
[6] Maradudin A A, Simonsen I, Leskova T A and Méndez E R 1999 Random surfaces that suppress single scattering *Opt. Lett.* **24** 1257–9
[7] Chaikina E I, García-Guerrero E E, Gu Z-H, Leskova T A, Maradudin A A, Méndez E R and Shchegrov A V 2000 Multiple-scattering phenomena in the scattering of light from randomly rough surfaces *Frontiers of Laser Physics and Quantum Optics* ed Z Xu, S Xie, S Y Zhu and M O Scully (New York: Springer) pp 225–59
[8] Maradudin A A, Méndez E R and Shchegrov A V (unpublished work)
[9] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in Fortran* 2nd edn (New York: Cambridge University Press) pp 281–2
[10] Maradudin A A, Michel T, McGurn A R and Méndez E R 1990 Enhanced backscattering of light from a random grating *Ann. Phys., NY* **203** 255–307

Two-dimensional random surfaces that act as circular diffusers

A. A. Maradudin and T. A. Leskova

Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, Irvine, California 92697

E. R. Méndez

División de Física Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, Apartado Postal 2732, Ensenada, B.C. 22800, Mexico

Received August 6, 2002

We propose a method for designing a two-dimensional random Dirichlet surface that, when it is illuminated at normal incidence by a scalar plane wave, scatters the wave with a circularly symmetric distribution of intensity. The method is applied to the design of a surface that acts as a Lambertian diffuser. The method is tested by computer simulations, and a procedure for fabricating such surfaces on photoresist is described.
© 2003 Optical Society of America
OCIS code: 240.5770.

Optical diffusers that produce a scattered intensity proportional to the cosine of the polar scattering angle are frequently used in the optical industry, e.g., for calibrating scatterometers.[1] Such diffusers have the property that their radiance or luminance is the same in all scattering directions. For this reason they are often referred to as Lambertian diffusers. In the visible region of the optical spectrum volume-disordered media, such as compacted powdered barium sulfate and freshly smoked magnesium oxide,[2] are used as Lambertian diffusers. However, this type of diffuser is inapplicable in the medium- or far-infrared regions because of its strong absorption and the presence of a specular peak in the scattered light. A different realization of such a diffuser is therefore required.

The design of a two-dimensional random surface that acts as a Lambertian diffuser, especially in the medium- or far-infrared region of the optical spectrum, is therefore a desirable goal. In this Letter we propose a method for designing a two-dimensional randomly rough Dirichlet surface that, when it is illuminated at normal incidence by a scalar plane wave, scatters the wave with a prescribed circularly symmetric distribution of intensity. We illustrate this method by applying it to the design of a surface that acts as an achromatic Lambertian diffuser.

The physical system that we consider consists of vacuum in the region $x_3 > \zeta(\mathbf{x}_{||})$, where $\mathbf{x}_{||} = (x_1, x_2, 0)$ is a position vector in the plane $x_3 = 0$, and the scattering medium in the region $x_3 < \zeta(\mathbf{x}_{||})$. The surface profile function $\zeta(\mathbf{x}_{||})$ is assumed to be a single-valued function of $\mathbf{x}_{||}$ that is differentiable with respect to x_1 and x_2 and constitutes a random process. The surface $x_3 = \zeta(\mathbf{x}_{||})$ is illuminated at normal incidence by a scalar plane wave of frequency ω, and the Dirichlet boundary condition is satisfied on it.

The mean differential reflection coefficient $\langle \partial R/\partial \Omega_s \rangle$, where the angle brackets denote an average over the ensemble of realizations of the surface profile function, is defined such that $\langle \partial R/\partial \Omega_s \rangle \mathrm{d}\Omega_s$ is the fraction of the total time-averaged incident flux that is scattered into the element of solid angle $\mathrm{d}\Omega_s$ about a given scattering direction. In the geometrical optics limit of the Kirchhoff approximation the coefficient is given by

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{1}{S}\left(\frac{\omega}{2\pi c}\right)^2 \int \mathrm{d}^2 u_{||} \exp(-i\mathbf{q}_{||} \cdot \mathbf{u}_{||}) \times \int \mathrm{d}^2 x_{||} \langle \exp[ia\mathbf{u}_{||} \cdot \nabla \zeta(\mathbf{x}_{||})] \rangle , \quad (1)$$

where S is the area of the $x_1 x_2$ plane covered by the random surface, $\mathbf{q}_{||} = (\omega/c)\sin\theta_s(\cos\phi_s, \sin\phi_s, 0)$, (θ_s, ϕ_s) are the polar and azimuthal angles of scattering, and $a = (\omega/c)(1 + \cos\theta_s)$.

The idea underlying our approach to the design of a two-dimensional random surface that scatters in a circularly symmetric fashion a scalar plane wave that is normally incident on it is that the surface itself should have circular symmetry. Consequently, we assume that the surface profile function $\zeta(\mathbf{x}_{||})$ is a function of the radial coordinate $r = |\mathbf{x}_{||}|$ alone, $\zeta(\mathbf{x}_{||}) = H(r)$, and we choose $H(r)$ to have the form

$$H(r) = a_n r + b_n , \quad nb \le r \le (n+1)b , \quad n = 0, 1, 2, \ldots , \quad (2)$$

where $\{a_n\}$ are independent identically distributed random deviates, b is a characteristic length, and $\{b_n\}$ are determined in such a way as to make $H(r)$ a continuous function of r:

$$b_n = b_0 + (a_0 + a_1 + \cdots + a_{n-1} - n a_n)b , \quad n \ge 1 . \quad (3)$$

We seek the probability density function (pdf) of a_n, $f(\gamma) = \langle \delta(\gamma - a_n) \rangle$, such that the surface profile defined by Eqs. (2) and (3) yields a mean differential reflection coefficient of a prescribed form. The average over the ensemble of realizations of the surface

0146-9592/03/020072-03$15.00/0 © 2003 Optical Society of America

profile function in the definition of $f(\gamma)$ is equivalent to an average over the ensemble of realizations of $\{a_n\}$ and the independence of $\{a_n\}$, and the fact that they are identically distributed has the consequence that this pdf is independent of n.

For the form of the surface profile function that we have chosen, Eq. (1) becomes

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{2\pi}{a} \left(\frac{\omega}{2\pi c} \right)^2 \frac{1}{q_{||}} f\left(\frac{q_{||}}{a} \right). \tag{4}$$

If we make the change of variable $\gamma = q_{||}/a = \tan(\theta_s/2)$, we find that $f(\gamma)$ is given by

$$f(\gamma) = 8\pi \frac{\gamma}{(1+\gamma^2)^2} \left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle (\gamma), \quad \gamma \geq 0. \tag{5}$$

We seek to design a surface that acts as a Lambertian diffuser, for which

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{1}{\pi} \cos\theta_s, \quad 0 \leq \theta_s \leq \pi/2. \tag{6}$$

We find from Eq. (5) that

$$f(\gamma) = 8 \frac{\gamma(1-\gamma^2)}{(1+\gamma^2)^3} \theta(1-\gamma)\theta(\gamma), \tag{7}$$

where $\theta(z)$ is the Heaviside unit step function. From the result given by Eq. (7), a long sequence of $\{a_n\}$ is obtained, e.g., by the rejection method,[3] and the surface profile function $H(r)$ is constructed on the basis of Eqs. (2) and (3).

Because rigorous calculations of scattering from two-dimensional random surfaces are time consuming, we illustrate this approach to the design of random surfaces that scatter light in a circularly symmetric fashion by performing the calculations of the mean differential reflection coefficient in the Kirchhoff approximation. In this approximation we have that

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{1}{S} \left(\frac{\omega}{2\pi c} \right)^2 \langle |r(\mathbf{q}_{||})^2| \rangle, \tag{8}$$

where

$$r(\mathbf{q}_{||}) = 2\pi \sum_{n=0}^{N-1} \exp(-iab_n) \times \int_{nb}^{(n+1)b} \mathrm{d}r r J_0(q_{||} r) \exp(-iaa_n r), \tag{9}$$

$J_0(z)$ is a Bessel function, $a = (\omega/c)(1+\cos\theta_s)$, $q_{||} = (\omega/c)\sin\theta_s$, and the area S is $S = \pi(Nb)^2$. A large number N_p of realizations of the random surface is now generated, and for each realization $|r(\mathbf{q}_{||})|^2$ is calculated. The N_p results for $|r(\mathbf{q}_{||})|^2$ that are obtained are summed, and we then divide the sum by N_p to yield the average indicated in Eq. (8). The average $\langle |r(\mathbf{q}_{||})|^2 \rangle$ must increase quadratically with N to produce a mean differential reflection coefficient that is independent of N. This is indeed found to be the case for the example studied.

In Fig. 1 we plot a segment of one realization of the surface profile function $H(r)$ and its derivative calculated by the approach proposed here. The parameters used in generating these functions were $b = 200\lambda$ and $b_0 = 0$. Because of the form of $f(\gamma)$ given by Eq. (7), each slope a_n is positive, and as a result $H(r)$ increases monotonically with increasing r. The average slope of the surface is $\langle a_n \rangle = 0.4292$, independently of n, with a standard deviation of 0.2078.

A plot of $\langle \partial R/\partial \Omega_s \rangle$ as a function of θ_s, calculated for the parameters used in obtaining the surface profile function plotted in Fig. 1, is presented in Fig. 2. The results obtained for $N_p = 15,000$ realizations of the surface profile function were used in obtaining the average in Eq. (8). The random surface generated by our approach is seen to act as a Lambertian diffuser: The mean differential reflection coefficient follows the cosine law Eq. (6) very closely.

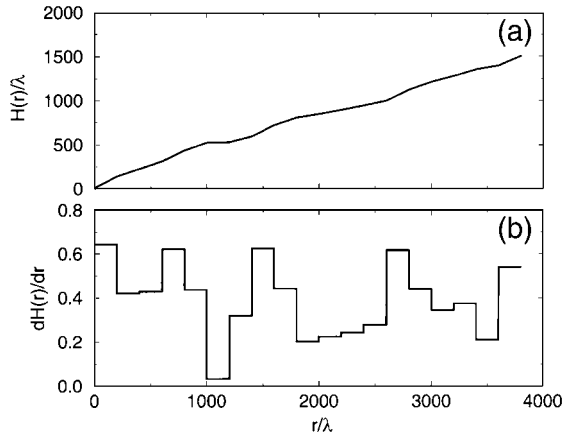

Fig. 1. Segment of a single realization of (a) a numerically generated surface profile function $H(r)$ and (b) its derivative for a two-dimensional surface that acts as a Lambertian diffuser. The parameters are $b = 200\lambda$ and $b_0 = 0$.

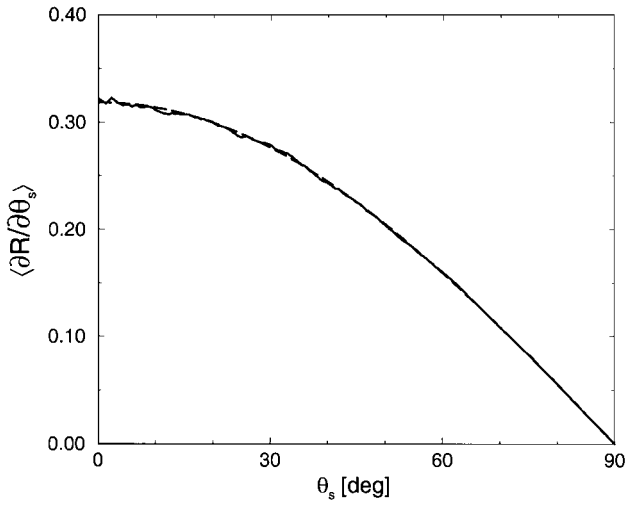

Fig. 2. Mean differential reflection coefficient for scattering from a two-dimensional Lambertian diffuser estimated from $N_p = 15,000$ realizations of the surface profile function for normal incidence. The parameters are $b = 200\lambda$, $b_0 = 0$, and $N = 200$. The dashed curve is a plot of the function given by Eq. (6).

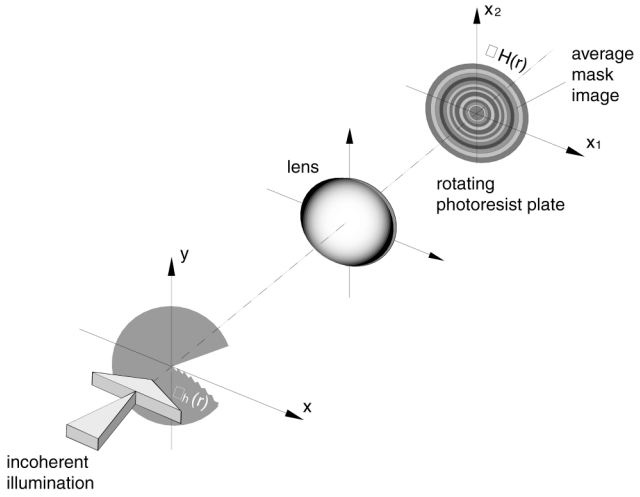

Fig. 3. Schematic diagram of the setup proposed for the fabrication of a circularly symmetric random surface.

Several comments need to be made concerning the calculations that led to Fig. 2. If the coefficients $\{a_n\}$ are obtained by a straightforward application of the rejection method to the pdf $f(\gamma)$ given by Eq. (7), a fraction of the surfaces that are generated possess regions in which several consecutive a_n are very small ($<10^{-2}$). The presence of such practically horizontal regions on the surface gives rise to a peak in the mean differential reflection coefficient in the immediate vicinity of the specular direction $\theta_s = 0°$. To eliminate this peak, as a particular realization of the function $H(r)$ was being generated, we calculated the corresponding differential reflection coefficient at $\theta_s = 0°$, and if its value exceeded $2\max\langle\partial R/\partial\Omega_s\rangle$ that surface was discarded. Although this procedure alters the statistics of the resulting surface from the statistics defined by the pdf $f(\gamma)$, the altered statistics produce a surface with the desired scattering properties and can be used in fabricating surfaces with these properties. The coherent interference of waves scattered from symmetric regions of the surface also contributes a peak in the specular direction. We can eliminate this contribution by making the surface large enough. In addition, a fraction of the surfaces generated possess regions in which several consecutive a_n are almost equal to unity. The presence of such regions on the surface gives rise to a peak in the immediate vicinity of the grazing directional. This peak, however, can also be eliminated by use of larger surfaces.

The form of $f(\gamma)$ given by Eq. (7) allows some segments of the surface to have large enough slopes that when they are illuminated at normal incidence they give rise to scattered rays that, because of the conical bowl shape of the surface, strike the surface a second time before being reflected back into the vacuum region. Such double-scattering processes, which are not captured by the Kirchhoff approximation, shadow grazing-angle scattering and in principle should cause the mean differential reflection coefficient to depart from the desired Lambertian form by decreasing more rapidly to zero as the scattering angle θ_s approaches 90°. However, estimates of the double-scattering contribution to $\langle\partial R/\partial\Omega_s\rangle$, to be described in the future, show that this departure is negligibly small.

Surfaces of the type considered here can be fabricated on photoresist in the following way: Consider the function $h(\rho)$ that constitutes a single realization of the random process constructed according to Eqs. (2), (3), and (7) and satisfies the requirement that the corresponding differential reflection coefficient at $\theta_s = 0°$ does not exceed $2\max\langle\partial R/\partial\Omega_s\rangle$. We define the function $\theta_h(\rho) = -\kappa h(\rho)$ that, with an appropriate choice of the constant κ, can be interpreted as an angle. A disk-shaped mask is then constructed, whose transparent section is defined by the condition $\theta_h(\rho) < \theta < \theta_g$, where θ_g is a constant greater than $[\theta_h(\rho)]_{\max}$.

A well-corrected optical system forms a demagnified image of the mask on a photoresist-coated plate, as shown in Fig. 3. The rotating photoresist plate is exposed for a time T_e, during which it executes a large integer number of revolutions. The total exposure of the plate will then be circularly symmetric, with a radial dependence of the form $E(r) = KT_e[\theta_g - \theta_H(r)]/2\pi$, where $\theta_H(r)$ is a scaled version of the function $\theta_h(\rho)$, and K is a constant related to the sensitivity of the photoresist and the intensity of the illumination. The exposure of the plate can be put in the final form $E(r) = E_0 + \alpha H(r)$, where E_0 and α are constants that one can adjust by varying the shape of the mask and the exposure time. Assuming that the relation between exposure and resulting height on the plate is linear, the developed surface will have the desired properties. Moreover, the characterized nonlinearities of the photoresist response can be taken into account by prescribed deformations of the mask in Fig. 3.

In conclusion, we have proposed a method for generating a two-dimensional random surface that, when it is illuminated at normal incidence by a scalar plane wave, scatters the wave with a circularly symmetric distribution of intensity. We illustrated this method by designing a surface that acts as a Lambertian diffuser. We validated this method by computer simulations and described a procedure for fabricating such surfaces on photoresist.

The research of A. A. Maradudin (aamaradu@uci.edu) and T. A. Leskova (leskova@duey.ps.uci.edu) was supported in part by U.S. Army Research Office grant DAAD 19-99-1-0321. The research of E. R. Méndez (emendez@cicese.mx) was supported in part by the Consejo Nacional de Ciencia Y Tecnología (Mexico) through grant 32254-E.

References

1. J. C. Stover, *Optical Scattering* (SPIE, Bellingham, Wash., 1995).
2. F. Grum and G. W. Luckey, Appl. Opt. **7,** 2295 (1968).
3. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge U. Press, New York, 1992), p. 281.

Two-Dimensional Random Surfaces that Act as Circular Diffusers

A. A. Maradudin[a], E. R. Méndez[b], and T. A. Leskova[a]

[a] Department of Physics and Astronomy
and Institute for Surface and Interface Science
University of California
Irvine, CA 92697, U.S.A.
[b] División de Física Aplicada, Centro de Investigación
Científica y de Educación Superior de Ensenada
Apdo. Postal 2732, Ensenada, B.C. 22800, México

ABSTRACT

We propose a method for designing a two-dimensional random Dirichlet surface which, when illuminated normally by a scalar plane wave, scatters it with a prescribed circularly symmetric distribution of intensity. The method is applied to the design of a surface that scatters light uniformly within a circular region and produces no scattering outside that region, and a surface that acts as a Lambertian diffuser. The method is tested by computer simulations, and a procedure for fabricating such surfaces on photoresist is described.

Keywords: random surfaces, scattering, band-limited uniform diffusers, Lambertian diffusers, Kirchhoff approximation

In several recent articles the present authors and their colleagues have addressed the question of how to design random surfaces that scatter[1–5] or transmit[6,7] light in a specified manner. The majority of these studies have dealt with one-dimensional random surfaces.[1–7] The only study of two-dimensional random surfaces in this context dealt with surfaces that scatter light uniformly within a rectangular region and produce no scattering outside this region.[8] Such surfaces are called band-limited uniform diffusers, and can be useful in projection systems where one wishes to illuminate a screen uniformly, but does not wish to waste light by illuminating outside the boundaries of the screen.

It seems desirable to extend such calculations to the design of two-dimensional random surfaces that produce a more general specified angular distribution of the intensity of the light scattered from them in more general regions of scattering angles. Thus, in this paper we propose a method for designing a two-dimensional randomly rough Dirichlet surface which, when illuminated at normal incidence by a scalar plane wave, scatters it with a prescribed circularly symmetric distribution of intensity. This method differs from the one employed in Refs. 1-8, and is simpler to implement than the latter, both theoretically and experimentally. We illustrate it by applying it to the design of a two-dimensional random surface that scatters a scalar plane wave with a uniform intensity within a circular region and produces no scattering outside this region (a band-limited uniform diffuser), and one that acts as a Lambertian diffuser. We validate the results obtained by computer simulations, and describe a procedure for fabricating such surfaces on photoresist.

The physical system we consider in this work consists of vacuum in the region $x_3 > \zeta(\mathbf{x}_\parallel)$, where $\mathbf{x}_\parallel = (x_1, x_2, 0)$ is a position vector in the plane $x_3 = 0$, and the scattering medium in the region $x_3 < \zeta(\mathbf{x}_\parallel)$. The surface profile function $\zeta(\mathbf{x}_\parallel)$ is assumed to be a single-valued function of $\mathbf{x}_\parallel$, that is differentiable with respect to x_1 and x_2, and constitutes a random process, but not necessarily a stationary one.

The surface $x_3 = \zeta(\mathbf{x}_\parallel)$ is illuminated from the vacuum side by a scalar plane wave, $\psi(\mathbf{x}, t)_{inc} = \hat{\psi}(\mathbf{x}|\omega)_{inc} \exp(-i\omega t)$, where

$$\hat{\psi}(\mathbf{x}|\omega)_{inc} = \exp[-i\mathbf{k}_\parallel \cdot \mathbf{x}_\parallel - i\alpha_0(k_\parallel)x_3]. \tag{1}$$

email: aamaradu@uci.edu,emendez@cicese.mx,leskova@duey.ps.uci.edu

In this expression $\mathbf{k}_\parallel = (k_1, k_2, 0)$ is a two-dimensional wave vector, and $\alpha_0(k_\parallel) = [(\omega/c)^2 - k_\parallel^2]^{\frac{1}{2}}$, with $Re\alpha_0(k_\parallel) > 0$, $Im\alpha_0(k_\parallel) > 0$. The scattered field in the far field, $\psi(\mathbf{x}, t)_{sc} = \hat{\psi}(\mathbf{x}|\omega)_{sc} \exp(-i\omega t)$, is a superposition of outgoing waves,

$$\hat{\psi}(\mathbf{x}|\omega)_{sc} = \int \frac{d^2 q_\parallel}{(2\pi)^2} R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) \exp[i\mathbf{q}_\parallel \cdot \mathbf{x}_\parallel + i\alpha_0(q_\parallel)x_3] \tag{2}$$

in the region $x_3 > \zeta(\mathbf{x}_\parallel)_{max}$. the total field in the vacuum region is assumed to satisfy the Dirichlet boundary condition

$$[\psi(\mathbf{x}, t)_{inc} + \psi(\mathbf{x}, t)_{sc}]|_{x_3 = \zeta(\mathbf{x}_\parallel)} = 0 \tag{3}$$

on the surface $x_3 = \zeta(\mathbf{x}_\parallel)$.

The differential reflection $\partial R/\partial\Omega_s$ is defined in such a way that $\partial R/\partial\Omega_s d\Omega_s$ is the fraction of the total time-averaged incident flux that is scattered into the element of solid angle $d\Omega_s$ about a given scattering direction. It is given in terms of the scattering amplitude

$$\frac{\partial R}{\partial \Omega_s} = \frac{1}{S}\left(\frac{\omega}{2\pi c}\right)^2 \frac{\cos^2\theta_s}{\cos\theta_0} |R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)|^2. \tag{4}$$

where S is the area of the plane $x_3 = 0$ covered by the random surface. In this expression the wave vectors $\mathbf{q}_\parallel$ and $\mathbf{k}_\parallel$ must be expressed in terms of the polar and azimuthal angles of scattering (θ_s, ϕ_s) and incidence (θ_s, ϕ_0), respectively, according to

$$\mathbf{q}_\parallel = (\omega/c)\sin\theta_s(\cos\phi_s, \sin\phi_s, 0) \tag{5a}$$

$$\mathbf{k}_\parallel = (\omega/c)\sin\theta_0(\cos\phi_0, \sin\phi_0, 0), \tag{5b}$$

As we are concerned with the scattering of a wave from a random surface, it is the mean differential reflection coefficient that we must calculate. It is given by

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{1}{S}\left(\frac{\omega}{2\pi c}\right)^2 \frac{\cos^2\theta_s}{\cos\theta_0} \langle |R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)|^2 \rangle, \tag{6}$$

where the angle brackets denote an average over the ensemble of realizations of the surface profile function.

The application of Green's second integral identity to the region $x_3 > \zeta(\mathbf{x}_\parallel)$ yields the following expression for the scattering amplitude $R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)$[8]

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) = \frac{1}{2i\alpha_0(q_\parallel)} \int d^2 x_\parallel \exp[-i\mathbf{q}_\parallel \cdot \mathbf{x}_\parallel - i\alpha_0(q_\parallel)\zeta(\mathbf{x}_\parallel)] F(\mathbf{x}_\parallel), \tag{7}$$

where the source function $F(\mathbf{x}_\parallel)$ is defined by

$$F(\mathbf{x}_\parallel) = \left(-\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_1}\frac{\partial}{\partial x_1} - \frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_2}\frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3}\right) \times \left[\hat{\psi}(\mathbf{x}|\omega)_{inc} + \hat{\psi}(\mathbf{x}|\omega)_{sc}\right]\Big|_{x_3 = \zeta(\mathbf{x}_\parallel)}. \tag{8}$$

In the Kirchhoff approximation, which we adopt here for simplicity, the source function is[8]

$$F(\mathbf{x}_\parallel) = 2\left(-\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_1}\frac{\partial}{\partial x_1} - \frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_2}\frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3}\right)\hat{\psi}(\mathbf{x}|\omega)_{inc}\Big|_{x_3 = \zeta(\mathbf{x}_\parallel)} \tag{9a}$$

$$= 2\left(-ik_1\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_1} - ik_2\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_2} - i\alpha_0(k_\parallel)\right)\exp[i\mathbf{k}_\parallel \cdot \mathbf{x}_\parallel - i\alpha_0(k_\parallel)\zeta(\mathbf{x}_\parallel)]. \tag{9b}$$

Substitution of Eq. (9b) into Eq. (6) yields the result that

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) = -\frac{1}{\alpha_0(q_\parallel)}\int d^2x_\parallel \left(k_1\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_1}+k_2\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_2}+\alpha_0(k_\parallel)\right) \times\exp[-i(\mathbf{q}_\parallel-\mathbf{k}_\parallel)\cdot\mathbf{x}_\parallel-i(\alpha_0(q_\parallel)+\alpha_0(k_\parallel))\zeta(\mathbf{x}_\parallel)]. \quad (10)$$

After an integration by parts this expression becomes[8]

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) = -\frac{(\omega/c)^2+\alpha_0(q_\parallel)\alpha_0(k_\parallel)-\mathbf{q}_\parallel\cdot\mathbf{k}_\parallel}{\alpha_0(q_\parallel)[\alpha_0(q_\parallel)+\alpha_0(k_\parallel)]} \times\int d^2x_\parallel\exp[-i(\mathbf{q}_\parallel-\mathbf{k}_\parallel)\cdot\mathbf{x}_\parallel-i(\alpha_0(q_\parallel)+\alpha_0(k_\parallel))\zeta(\mathbf{x}_\parallel)]. \quad (11)$$

It follows from Eqs. (5) and (10) that the mean differential reflection coefficient can be written in the form

$$\left\langle\frac{\partial R}{\partial\Omega_s}\right\rangle = \frac{1}{S}\left(\frac{\omega}{2\pi c}\right)^2\frac{[1+\cos\theta_0\cos\theta_s-\sin\theta_0\sin\theta_s\cos(\phi_0-\phi_s)]^2}{\cos\theta_0(\cos\theta_0+\cos\theta_s)^2} \times\int d^2x_\parallel\int d^2x'_\parallel\exp[-i(\mathbf{q}_\parallel-\mathbf{k}_\parallel)\cdot(\mathbf{x}_\parallel-\mathbf{x}'_\parallel)] \times\langle\exp[-i(\omega/c)(\cos\theta_0+\cos\theta_s)(\zeta(\mathbf{x}_\parallel)-\zeta(\mathbf{x}'_\parallel))]\rangle. \quad (12)$$

The result (11) for $R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)$ was obtained on the assumption that the total field in the vacuum region satisfies the Dirichlet boundary condition (3) on the surface $x_3=\zeta(\mathbf{x}_\parallel)$. If, instead, it is assumed that the total field in the vacuum region satisfies the Neumann boundary condition

$$\left[-\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_1}\frac{\partial}{\partial x_1}-\frac{\partial\zeta(\mathbf{x}_\parallel)}{\partial x_2}+\frac{\partial}{\partial x_3}\right][\psi(\mathbf{x},t)_{inc}+\psi(\mathbf{x},t)_{sc}]|_{x_3=\zeta(\mathbf{x}_\parallel)}=0 \quad (13)$$

on the surface $x_3=\zeta(\mathbf{x}_\parallel)$, the corresponding scattering amplitude $R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)$ is given by Eq. (11), but with the opposite sign. Since $R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)$ enters the mean differential reflection coefficient $\langle\partial R/\partial\Omega_s\rangle$ only in the form $|R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)|^2$, it follows that the result given by Eq. (12) is valid for a two-dimensional Neumann surface as well as for a two-dimensional Dirichlet surface.

The result given by Eq. (12) simplifies greatly in the case of normal incidence, $\theta_0=0$, $\mathbf{k}_\parallel=0$.

$$\left\langle\frac{\partial R}{\partial\Omega_s}\right\rangle = \frac{1}{S}\left(\frac{\omega}{2\pi c}\right)^2\int d^2x_\parallel\int d^2x'_\parallel\exp[-i\mathbf{q}_\parallel\cdot(\mathbf{x}_\parallel-\mathbf{x}'_\parallel)] \times\langle\exp[-ia(\zeta(\mathbf{x}_\parallel)-\zeta(\mathbf{x}'_\parallel))]\rangle, \quad (14)$$

where $a=(\omega/c)(1+\cos\theta_s)$, and we will restrict ourselves to this case in what follows.

The geometrical optics limit of the Kirchhoff approximation is obtained by making the change of variables $\mathbf{x}'_\parallel=\mathbf{x}_\parallel-\mathbf{u}_\parallel)$ in Eq. (12), expanding the difference $\zeta(\mathbf{x}_\parallel)-\zeta(\mathbf{x}_\parallel-\mathbf{u}_\parallel$ in powers of $\mathbf{u}_\parallel$, and keeping only the linear term. In this way we obtain the result

$$\left\langle\frac{\partial R}{\partial\Omega_s}\right\rangle = \frac{1}{S}\left(\frac{\omega}{2\pi c}\right)^2\int d^2u_\parallel\exp(-i\mathbf{q}_\parallel\cdot\mathbf{u}_\parallel) \times\int d^2x_\parallel\langle\exp[-ia\mathbf{u}_\parallel\cdot\nabla\zeta(\mathbf{x}_\parallel)]\rangle. \quad (15)$$

The idea underlying our approach to the design of a two-dimensional random surface that scatters in a circularly symmetric fashion a scalar plane wave incident normally on it, is that the surface itself should have

circular symmetry. Consequently, we assume that the surface profile function $\zeta(\mathbf{x}_\parallel)$ is a function of the radial coordinate $r = |\mathbf{x}_\parallel|$ alone, $\zeta(\mathbf{x}_\parallel) = H(r)$, and we choose $H(r)$ to have the form

$$H(r) = a_n r + b_n \qquad nb \le r \le (n+1)b, \qquad n = 0, 1, 2 \ldots \tag{16}$$

where the $\{a_n\}$ are independent random deviates, b is a characteristic length, and the $\{b_n\}$ are determined in such a way as to make $H(r)$ a continuous function of r,

$$b_n = b_0 + (a_0 + a_1 + \cdots + a_{n-1} - n a_n) b \qquad n \ge 1. \tag{17}$$

We seek the probability density function (pdf) of a_n, $f(\gamma) = \langle \delta(\gamma - a_n) \rangle$ such that the surface profile defined by Eqs. (5) and (6) yields a mean differential reflection coefficient of a prescribed form. Due to the independence of the $\{a_n\}$ the pdf $f(\gamma)$ is independent of n.

For the form of the surface profile function we have chosen, the gradient of the surface profile function is given by

$$\nabla \zeta(\mathbf{x}_\parallel) = \hat{\mathbf{x}}_\parallel a_n \qquad nb < r < (n+1)b. \tag{18}$$

The double integral in Eq. (15) therefore becomes

$$\begin{aligned}
&\int d^2 u_\parallel \exp(-i\mathbf{q}_\parallel \cdot \mathbf{u}_\parallel) \int d^2 x_\parallel \langle \exp[-ia\mathbf{u}_\parallel \cdot \nabla \zeta(\mathbf{x}_\parallel)] \rangle \\
&= \int d^2 u_\parallel \exp(-i\mathbf{q}_\parallel \cdot \mathbf{u}_\parallel) \sum_{n=0}^{N-1} \int_{nb}^{(n+1)b} dx_\parallel x_\parallel \int_{-\pi}^{\pi} d\phi_x \langle \exp[-iaa_n u_\parallel \cos(\phi_u - \phi_x)] \rangle \\
&= \int d^2 u_\parallel \exp(-i\mathbf{q}_\parallel \cdot \mathbf{u}_\parallel) \sum_{n=0}^{N-1} \pi[(n+1)^2 b^2 - n^2 b^2] \langle J_0(a a_n u_\parallel) \rangle \\
&= \int d^2 u_\parallel \exp(-i\mathbf{q}_\parallel \cdot \mathbf{u}_\parallel) \sum_{n=0}^{N-1} \pi[(n+1)^2 b^2 - n^2 b^2] \int_{-\infty}^{\infty} d\gamma f(\gamma) J_0(a\gamma u_\parallel) \\
&= S \int_{-\infty}^{\infty} du_\parallel u_\parallel \int_{-\pi}^{\pi} d\phi_u \exp[-iq_\parallel u_\parallel \cos(\phi_q - \phi_u)] \int_{-\infty}^{\infty} d\gamma f(\gamma) J_0(q\gamma u_\parallel) \\
&= 2\pi S \int_{-\infty}^{\infty} d\gamma f(\gamma) \int_0^{\infty} du_\parallel u_\parallel J_0(q_n u_n) J_0(a\gamma u_\parallel) \\
&= 2\pi S \int_{-\infty}^{\infty} d\gamma f(\gamma) \frac{1}{q_\parallel} \delta(q_\parallel - a\gamma) \\
&= S \frac{2\pi}{a} \frac{1}{q_\parallel} f\left(\frac{q_\parallel}{a}\right),
\end{aligned} \tag{19}$$

where $S = \pi(Nb)^2$. On combining Eqs. (15) and (19) we find that the mean differential reflection is given by

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{2\pi}{a} \left(\frac{\omega}{2\pi c}\right)^2 \frac{1}{q_\parallel} f\left(\frac{q_\parallel}{a}\right). \tag{20}$$

Thus, the mean differential reflection coefficient is given in terms of the pdf of a_n. If we make the change of variable $\gamma = q_\parallel / a = \tan(\theta_s/2)$, we find that $f(\gamma)$ is given by

$$f(\gamma) = 8\pi \frac{\gamma}{(1+\gamma^2)^2} \left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle (\gamma) \qquad \gamma \ge 0. \tag{21}$$

We first seek to design a surface that acts as a band-limited uniform diffuser within the circular region $0 \le \theta_s \le \theta_m$ and $0 \le \phi_s \le 2\pi$. For such a surface

$$\begin{aligned}\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle &= \frac{\theta(\theta_m - \theta_s)\theta(\theta_s)}{4\pi \sin^2(\theta_m/2)} \\ &= \frac{\theta(\tan(\theta_m/2) - \tan(\theta_s/2))\theta(\tan(\theta_s/2))}{4\pi \sin^2(\theta_m/2)} \\ &= \frac{1}{4\pi}\frac{1+\gamma_m^2}{\gamma_m^2}\theta(\gamma_m - \gamma)\theta(\gamma), \end{aligned} \tag{22}$$

where $\theta(z)$ is the Heaviside unit step function, and $\gamma_m = \tan(\theta_m/2)$. It follows from Eq. (21) that $f(\gamma)$ in this case is given by

$$f(\gamma) = 2\frac{1+\gamma_m^2}{\gamma_m^2}\frac{\gamma}{(1+\gamma^2)^2}\theta(\gamma_m - \gamma)\theta(\gamma). \tag{23}$$

Similarly, if we seek to design a surface that acts as a Lambertian diffuser, for which

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{1}{\pi}\cos\theta_s \qquad 0 \le \theta_s \le \pi/2, \tag{24}$$

we find from Eq. (21) that

$$f(\gamma) = 8\frac{\gamma(1-\gamma^2)}{(1+\gamma^2)^3}\theta(1-\gamma)\theta(\gamma). \tag{25}$$

From the results given by Eqs. (23) and (25) a long sequence of $\{a_n\}$ is obtained, e.g. by the rejection method,[9] and the surface profile function $H(r)$ constructed on the basis of Eqs. (16) and (17).

Because rigorous calculations of scattering from two-dimensional random surfaces are time consuming, we illustrate this approach to the design of random surfaces that scatter light in a circularly symmetric fashion by performing the calculations of the mean differential reflection coefficient in the Kirchhoff approximation, but without going to the geometrical optics limit of it. From Eqs. (11) and (15) we find that the scattering amplitude $R(q_\parallel|\mathbf{0})$ in the Kirchhoff approximation can be written as

$$R(\mathbf{q}_\parallel|\mathbf{0}) = -\frac{1}{\cos\theta_s} r(\mathbf{q}_\parallel), \tag{26}$$

where

$$\begin{aligned} r(\mathbf{q}_\parallel) &= \int d^2x_\parallel \exp(-i\mathbf{q}_\parallel \cdot \mathbf{x}_\parallel) \exp[-iaH(r)] \\ &= \sum_{n=0}^{N-1} \int_{nb}^{(n+1)b} dr r \int_{-\pi}^{\pi} d\phi_x \exp[-iq_\parallel r\cos(\phi_q - \phi_x)] \exp[-ia(a_n r + b_n)] \\ &= 2\pi \sum_{n=0}^{N-1} \int_{nb}^{(n+1)b} dr r \exp[-ia(a_n r + b_n)] J_0(q_\parallel r) \\ &= 2\pi \sum_{n=0}^{N-1} \exp(-iab_n) \int_{nb}^{(n+1)b} dr\, r J_0(q_\parallel r) \exp(-iaa_n r), \end{aligned} \tag{27}$$

where $J_0(z)$ is a Bessel function, $a = (\omega/c)(1 + \cos\theta_s)$, and $q_{\|} = (\omega/c)\sin\theta_s$. The mean differential reflection coefficient is then given by

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{1}{S}\left(\frac{\omega}{2\pi c}\right)^2 \langle |r(\mathbf{q}_{\|})|^2 \rangle, \tag{28}$$

where the area S is $S = \pi(Nb)^2$.

The procedure now is to generate a large number N_p of realizations of the random surface, and for each realization to calculate the corresponding values of $|r(\mathbf{q}_{\|})|^2$. The N_p values of $|r(\mathbf{q}_{\|})|^2$ so obtained are summed and the sum then divided by N_p to yield the average indicated in Eq. (28). It should be noted that the average $\langle |r(\mathbf{q}_{\|})|^2 \rangle$ must increase quadratically with N in order to produce a mean differential reflection coefficient that is independent of N. This is indeed found to be the case for the examples studied.

An example of the surface profile function $H(r)$ and its derivative calculated by the approach proposed here is plotted in Fig. 1 for the case of a circular uniform diffuser defined by Eq. (22). The parameters used in generating these functions were $\theta_m = 20°$, $b = 200\lambda$, and $N = 20$. In addition, we set $b_0 = 0$ in Eq. (17). Due to the form of $f(\gamma)$ given by Eq. (23) each slope a_n is positive, with the result that the surface increases monotonically with increasing r.

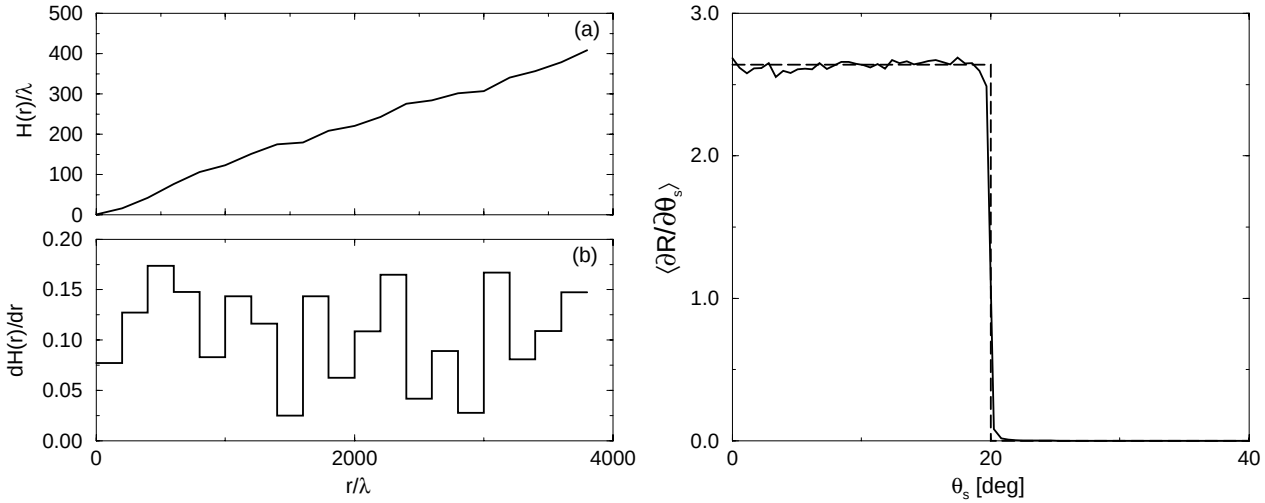

Figure 1. Numerically generated surface profile function $H(r)$ (a) and its derivative (b) for a two-dimensional surface that acts as a circular band-limited uniform diffuser. The parameters employed are $\theta_m = 20°$, $b = 200\lambda$, $b_0 = 0$, and $N = 20$.

Figure 2. The mean differential reflection coefficient for scattering from a circular uniform diffuser estimated from $N_p = 8000$ realizations of the surface profile function for the case of normal incidence. The parameters employed are $\theta_m = 20°$, $b = 200\lambda$, $b_0 = 0$, and $N = 200$. The dashed curve is a plot of the function given by Eq. (22).

In Fig. 2 we present a plot of $\langle \partial R/\partial \Omega_s \rangle$ as a function of θ_s for this circular uniform diffuser obtained by the approach proposed in this paper. The parameters used in obtaining the result plotted in this figure were the same as those assumed in obtaining the surface profile function plotted in Fig. 1. The results obtained for $N_p = 8000$ realizations of the surface profile function were used in calculating the average indicated in Eq. (28). We see from this figure that the random surface generated by our approach indeed acts as a band-limited uniform diffuser. There is no scattered intensity for scattering angles larger than 20°, and the cutoff is very sharp. For values of θ_s in the interval $0 \leq \theta_s < 20°$ the scattered intensity is very closely uniform.

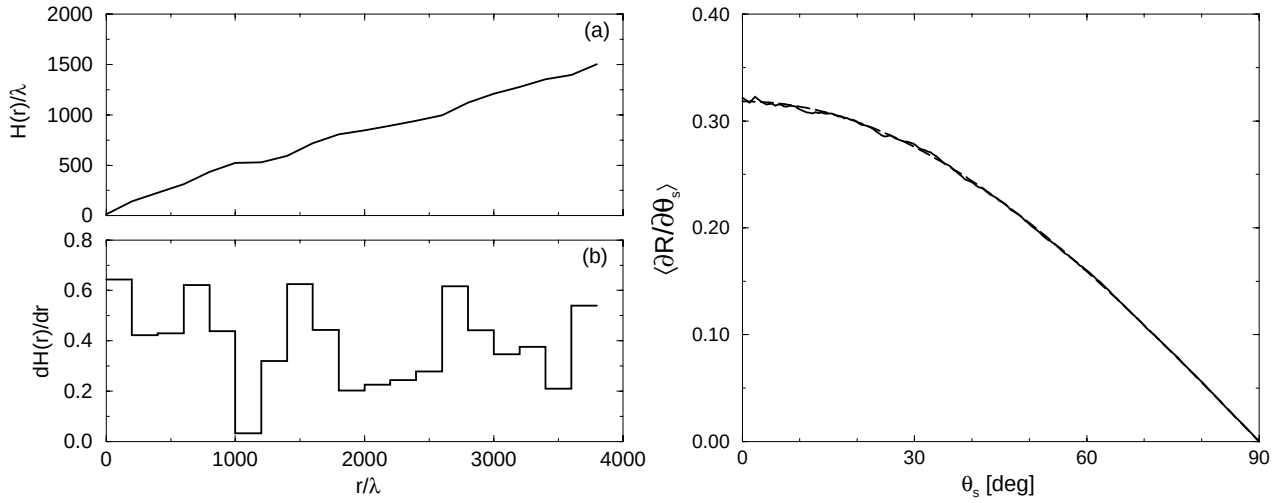

Figure 3. Numerically generated surface profile function $H(r)$ (a) and its derivative (b) for a two-dimensional surface that acts as a Lambertian diffuser. The parameters employed are $b = 200\lambda$, $b_0 = 0$, and $N = 20$.

Figure 4. The mean differential reflection coefficient for scattering from a two-dimensional Lambertian diffuser estimated from $N_p = 15000$ realizations of the surface profile function for the case of normal incidence. The parameters employed are $b = 200\lambda$, $b_0 = 0$, and $N = 200$. The dashed curve is a plot of the function given by Eq. (24).

In Fig. 3 we have plotted an example of the surface profile function $H(r)$ and its derivative calculated by the approach proposed here for the case of a Lambertian diffuser defined by Eq. (24). The parameters used in generating these functions were $b = 200\lambda, b_0 = 0$, and $N = 20$. Again, due to the form of $f(\gamma)$ given by Eq. (25), each slope a_n is positive, and as a result the surface increases monotonically with increasing r. A plot of $\langle \partial R/\partial \Omega_s \rangle$ as a function of θ_s for the Lambertian diffuser obtained by the approach proposed in this paper is presented in Fig. 4. The parameters used in obtaining the result plotted in this figure were the same as those assumed in obtaining the surface profile function plotted in Fig. 3. The results obtained for $N_p = 15000$ realizations of the surface profile function were used in obtaining the average appearing in Eq. (28). The random surface generated by our approach is seen to act as a Lambertian diffuser: the mean differential reflection coefficient follows the cosine law given by Eq. (24) very closely.

Several comments need to be made concerning the calculations that led to Figs. 2 and 4. If the coefficients $\{a_n\}$, and hence the surface profile function $H(r)$, are obtained by a straightforward application of the rejection method to the pdfs $f(\gamma)$ given by Eqs. (23) and (25), a fraction of the surfaces generated possess regions in which several consecutive a_n are very small, namely smaller than 10^{-2}. The presence of such practically horizontal regions on the surface gives rise to a peak in the mean differential reflection coefficient in the immediate vicinity of the specular direction $\theta_s = 0°$. To eliminate this peak, as a particular realization of the function $H(r)$ was being generated, the differential reflection coefficient at $\theta_s = 0°$ was calculated, and if its value exceeded $2\max\langle \partial R/\partial \Omega_s \rangle$ that surface was discarded. Although this procedure alters the statistics of the resulting surface from the statistics defined by the pdf $f(\gamma)$, the altered statistics produces a surface with the desired scattering properties, and can be used in fabricating surfaces on photoresist with these properties. The coherent interference of the waves scattered from symmetric regions of the surface also contributes to the peak in the specular direction. This contribution can be eliminated by making the surface large enough. In addition, in the case of the Lambertian diffuser, a fraction of the surfaces generated possess regions in which several consecutive a_n are almost equal to unity. The presence of such regions on the surface gives rise to a peak in the mean differential relection coefficient in the immediate vicinity of the grazing direction. This peak, however, can also be eliminated by using larger surfaces.

To fabricate surfaces of the type considered here, we begin with the function $h(\rho)$ that constitutes a realization of the required random process, constructed according to Eqs. (16), (17), and (21). We define the function

$$\theta_h(\rho) = -\kappa h(\rho). \tag{29}$$

With an appropriate choice of the constant κ, $\theta_h(\rho)$ can be interpreted as an angle [i. e. κ must have suitable dimensions and be chosen in such a way that $\theta_h(\rho) < 2\pi$].

A disk-shaped mask is constructed as depicted in Fig. 5. The transparent sector is defined by the condition

$$\theta_h(\rho) < \theta < \theta_g, \tag{30}$$

where θ_g is a constant greater than $[\theta_h(\rho)]_{\max}$.

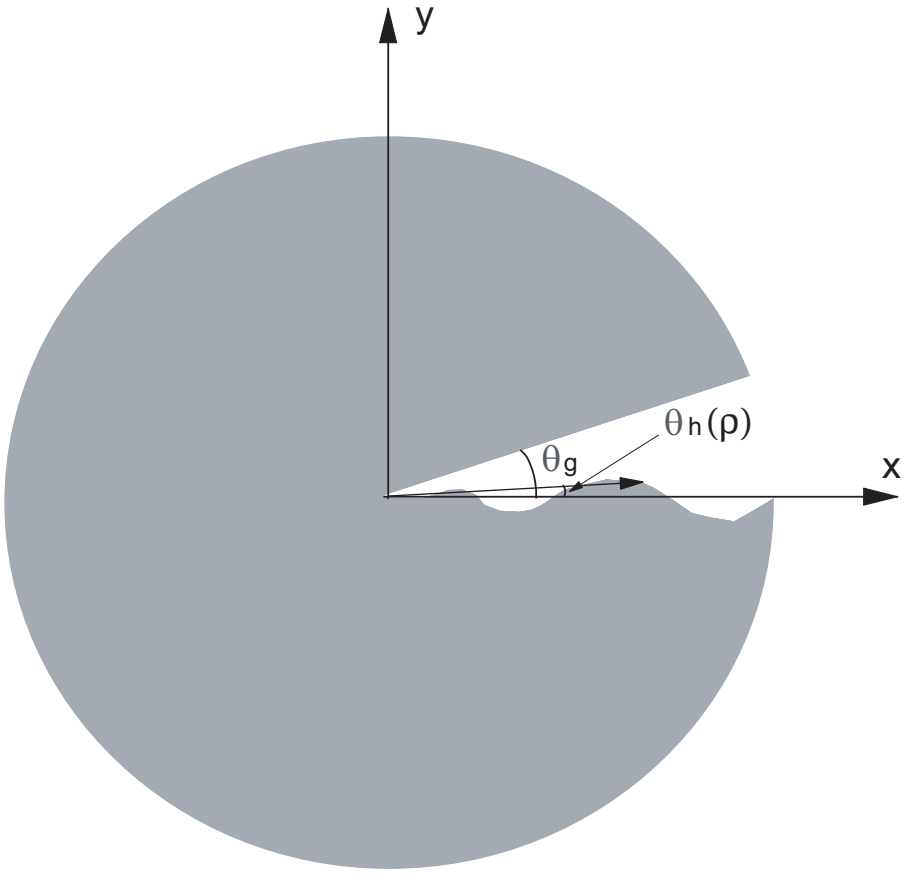

Figure 5: The disk-shaped mask employed in the fabrication of a circularly symmetric random surface.

To fabricate the circularly-symmetric surface, a well-corrected optical system is employed to form a demagnified image of the mask on a photoresist-coated plate, as shown in Fig. 6. It is assumed that the radial structure of the mask is resolvable.

The rotating photoresist plate is exposed for a time T_e, during which it executes a large integer number of revolutions. The total exposure of the plate should be rotationally symmetric, with a radial dependence of the form

$$E(r) = KT_e \frac{\theta_g - \theta_H(r)}{2\pi}, \tag{31}$$

where $\theta_H(r)$ is a scaled version of the function $\theta_h(\rho)$, and K is a constant related to the sensitivity of the photoresist and the intensity of the illumination.

The exposure of the plate can be put in the final form

$$E(r) = E_0 + \alpha H(r), \tag{32}$$

where E_0 and α are constants that can be adjusted with the shape of the mask and the exposure time.

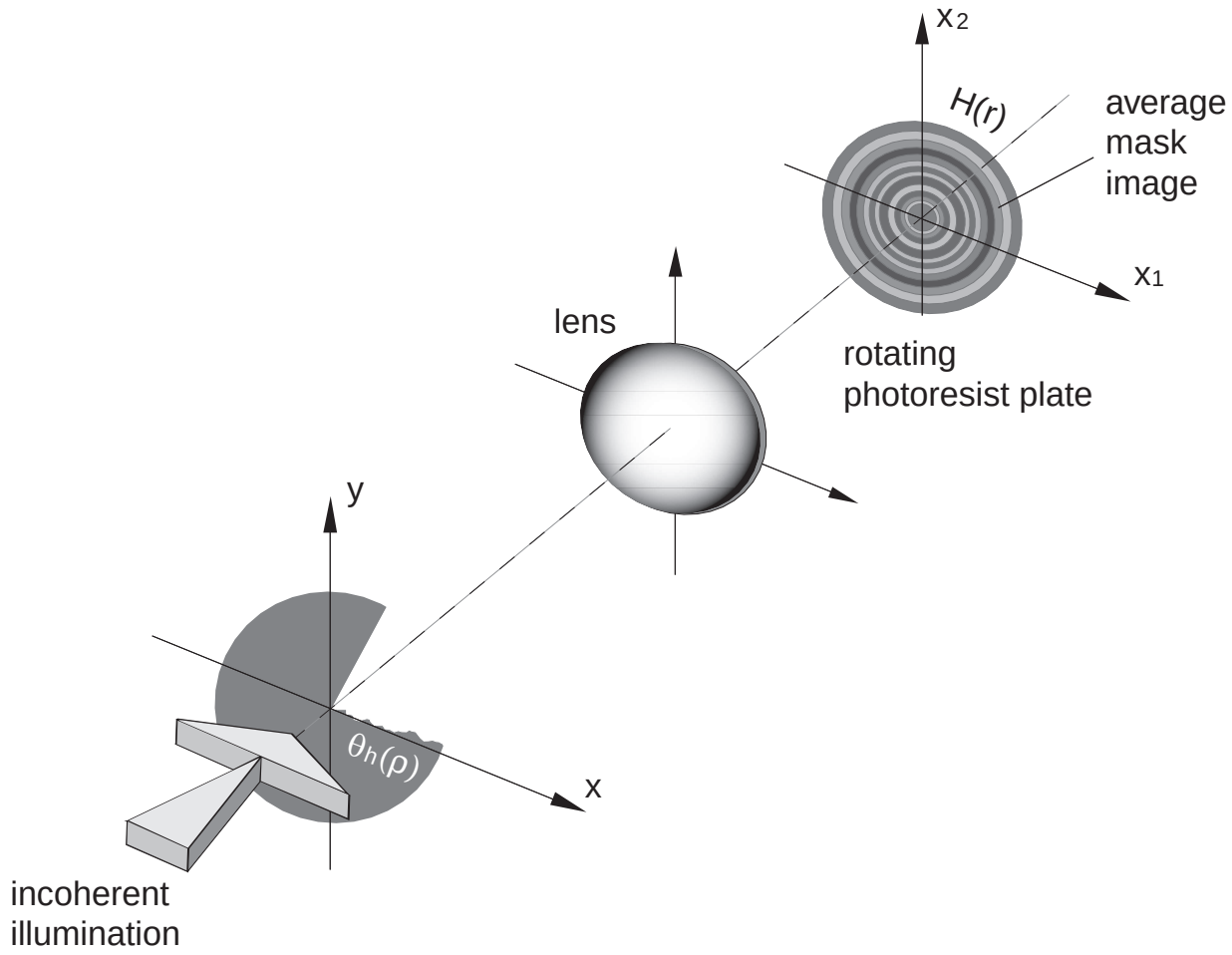

Figure 6: Schematic diagram of the setup proposed for the fabrication of a circularly symmetric random surface.

With this arrangement, the total radial exposure is proportional to a scaled version of the original random profile $h(\rho)$ and, thus, has suitable statistical properties. Assuming that the relation between exposure and resulting height on the surface is linear, the developed surface will have the desired properties. Moreover, characterized nonlinearities of the photoresist response may be taken into account by prescribed deformations of the mask illustrated in Fig. 5.

In conclusion, we have proposed a method for generating a two-dimensional random Dirichlet or Neumann surface that when illuminated normally by a scalar plane wave scatters it with a prescribed circularly symmetric distribution of intensity. We have illustrated this method by applying it to the design of a surface that acts as a band-limited uniform diffuser for scattering angles within a circular region, as well as a two-dimensional surface that acts as a Lambertian diffuser. We have validated this method by computer simulations, and have described a procedure for fabricating such surfaces on photoresist.

Acknowledgements

The research of T. A. L. and A. A. M. was supported in part by Army Research Office grant DAAD 19-99-1-0321. The research of E.R.M. was supported in part by CONACYT (México).

REFERENCES

1. T. A. Leskova, A. A. Maradudin, I. V. Novikov, A. V. Shchegrov, and E. R. Méndez, Appl. Phys. Lett. **73**, 1943 (1998).
2. E. R. Méndez, G. Martinez-Niconoff, A. A. Maradudin, and T. A. Leskova, SPIE **3426**, 2 (1998).
3. A. A. Maradudin, I. Simonsen, T. A. Leskova, and E. R. Méndez, Opt. Lett. **24**, 1257 (1999).
4. A. A. Maradudin, T. A. Leskova, E. R. Méndez, and A. V. Shchegrov, Physics of the Solid State **41**, 835 (1999).
5. I. Simonsen, T. A. Leskova, A. A. Maradudin, and E. R. Méndez, Waves in Random Media **11**, 529 (2001).
6. A. A. Maradudin, T. A. Leskova, E. R. Méndez, and I. Simonsen, SPIE **4100**, 113 (2000).
7. A. A. Maradudin, E. R. Méndez, T. A. Leskova, and I. Simonsen, SPIE **4447**, 130 (2001).
8. E. R. Méndez, E. E. García-Guerrero, H. M. Escamilla, A. A. Maradudin, T. A. Leskova, and A. V. Shchegrov, Appl. Opt. **40**, 1098 (2001).
9. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran, 2nd Edition* (Cambridge University Press, New York, 1992), pp. 281-282.

Design of one-dimensional random surfaces with specified scattering properties

E. R. Méndez and E. E. García-Guerrero
División de Física Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, Apdo. Postal 2732, Ensenada, B.C. 22800, México

T. A. Leskova, A. A. Maradudin, and J. Muñoz-López
Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, California 92697

I. Simonsen
The Nordic Institute for Theoretical Physics–NORDITA, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

(Received 12 April 2002; accepted for publication 3 June 2002)

We propose a method for designing a one-dimensional random perfectly conducting surface which, when illuminated by a plane wave, scatters it with a prescribed angular distribution of intensity. The method is applied to the design of a surface that scatters light uniformly within a specified range of scattering angles, and produces no scattering outside this range. It is tested by computer simulations, and a procedure for fabricating such surfaces on photoresist is described. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1495900]

In a recent series of papers the present authors and their colleagues have presented a method for designing one-dimensional random surfaces that scatter light in a specified manner.[1,2] This method is based on expressing the random surface profile function as a superposition of equally spaced trapezoidal grooves, whose statistically independent random amplitudes are drawn from a probability density function (PDF) that is determined in such a way that the mean intensity of the scattered light has the specified angular distribution. It has been shown that such surfaces can be fabricated on photoresist,[1,2] to produce the specified angular distribution of the intensity of the scattered light.[2,3]

In this letter we present a method for designing a one-dimensional random surface that scatters light in a prescribed manner, that is simpler to implement, both theoretically and experimentally than the method of Refs. 1 and 2. Like the latter it is based on the geometrical optics limit of the Kirchhoff approximation for the scattering of s-polarized light incident normally on a perfectly conducting surface. This method is illustrated by applying it to the determination of a surface that scatters light uniformly within a specified range of scattering angles, and produces no scattering outside this range (a band-limited uniform diffuser). It is tested by rigorous computer simulation calculations. Finally, we indicate how the kinds of surfaces generated by our approach can be fabricated on photoresist.

The physical system we initially consider consists of vacuum in the region $x_3 > \zeta(x_1)$ and a perfect conductor in the region $x_3 < \zeta(x_1)$. The surface profile function $\zeta(x_1)$ is assumed to be a single-valued function of x_1 that is differentiable, and constitutes a random process. The surface $x_3 = \zeta(x_1)$ is illuminated from the vacuum region by an s-polarized plane wave of frequency ω, whose plane of incidence is the x_1x_3-plane.

Our starting point is the geometrical optics limit of the Kirchhoff approximation for the mean differential reflection coefficient which, in the case of normal incidence, is given by[1]

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle = \frac{1}{L_1} \frac{\omega}{2\pi c} \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} du \exp(iqu) \times \langle \exp(iau\zeta'(x_1)) \rangle, \tag{1}$$

where L_1 is the length of the x_1-axis covered by the random surface, $q = (\omega/c)\sin\theta_s$, where θ_s is the angle of scattering measured clockwise from the x_3-axis, and $a = (\omega/c)(1 + \cos\theta_s)$. The angle brackets denote an average over the ensemble of realizations of $\zeta(x_1)$. We now define a set of equally-spaced points along the x_1-axis by $x_n = nb$, where b is a characteristic length and $n = 0, \pm 1, \pm 2, \ldots$. The surface profile function $\zeta(x_1)$ is then given by

$$\zeta(x_1) = a_n x_1 + b_n; \quad nb \leqslant x_1 \leqslant (n+1)b, \tag{2}$$

where the $\{a_n\}$ are independent random deviates. Therefore, the probability density function of a_n,

$$f(\gamma) = \langle \delta(\gamma - a_n) \rangle, \tag{3}$$

is independent of n. In order that the surface be continuous at $x_1 = (n+1)b$, the relation

$$b_{n+1} = b_n - (n+1)(a_{n+1} - a_n)b \tag{4}$$

must be satisfied. From this recurrence relation the $\{b_n\}$ can be determined from a knowledge of the $\{a_n\}$, provided that an initial value, for example, that of b_0, is specified. It is convenient to choose $b_0 = 0$, and we will do so in what follows.

The double integral in Eq. (1) can now be evaluated, with the result

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle = \frac{1}{1 + \cos\theta_s} f\left(\frac{-\sin\theta_s}{1 + \cos\theta_s} \right). \tag{5}$$

 © 2002 American Institute of Physics
Downloaded 26 Mar 2003 to 128.200.81.38. Redistribution subject to AIP license or copyright, see http://ojps.aip.org/aplo/aplcr.jsp

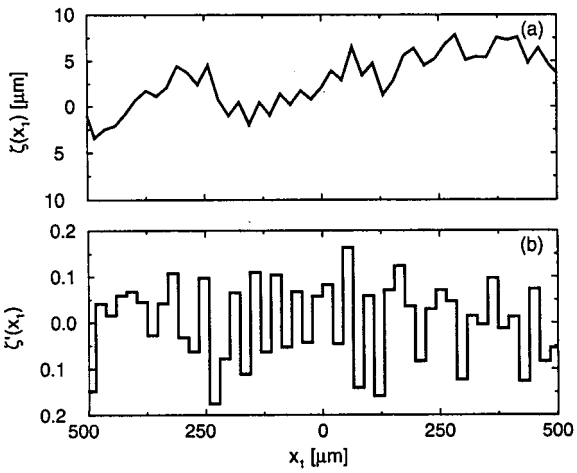

FIG. 1. (a) A segment of the surface profile function $\zeta(x_1)$ for the case of a band-limited uniform diffuser. (b) The derivative $\zeta'(x_1)$ of this surface profile function.

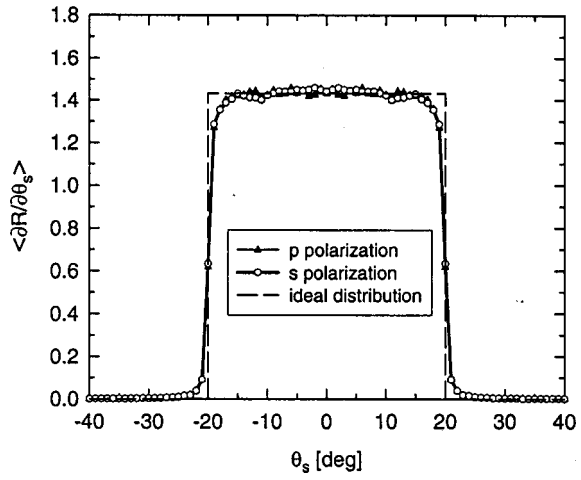

FIG. 2. The mean differential reflection coefficient $\langle \partial R/\partial\theta_s \rangle$ estimated from $N_p = 20\,000$ realizations of perfectly conducting surface profiles: $\theta_m = 20°$, $b = 22\ \mu$m, $\lambda = 632.8$ nm.

Thus, we find that the mean differential reflection coefficient is expressed in terms of the PDF of the random deviate a_n. We now make the change of variable $\sin\theta_s/(1+\cos\theta_s) = \tan(\theta_s/2) = \gamma$ and obtain from Eq. (5) that $f(\gamma)$ is given by

$$f(\gamma) = \frac{2}{1+\gamma^2}\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle (-\gamma). \tag{6}$$

It should be noted that this result does not depend explicitly on the wavelength of the incident light. A long sequence of $\{a_n\}$ is then generated, for example, by the rejection method,[4] and the corresponding sequence of $\{b_n\}$ is obtained from Eq. (4). The surface profile function $\zeta(x_1)$ is then constructed on the basis of Eq. (2).

We illustrate this approach by applying it to the design of a random surface that gives rise to a mean differential reflection coefficient that is a constant in the angular interval $|\theta_s| < \theta_m < \pi/2$, and vanishes for $|\theta_s| > \theta_m$ (a band-limited uniform diffuser),

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle = \frac{\theta(\theta_m - |\theta_s|)}{2\theta_m} = \frac{\theta(\gamma_m - |\gamma|)}{4\tan^{-1}\gamma_m}, \tag{7}$$

where $\theta(z)$ is the Heaviside unit step function and $\gamma_m = \tan(\theta_m/2)$. We find from Eq. (6) that the PDF of a_n is given by

$$f(\gamma) = \frac{1}{2\tan^{-1}\gamma_m}\,\frac{\theta(\gamma_m - |\gamma|)}{1+\gamma^2}. \tag{8}$$

A segment of the surface profile function $\zeta(x_1)$ and its derivative calculated by the approach proposed here are plotted in Figs. 1(a) and 1(b), respectively. The parameters used in generating these functions were $\theta_m = 20°$ and $b = 22\ \mu$m.

Although the derivation of the PDF $f(\gamma)$ given by Eq. (6) was based on a single-scattering approximation and the assumption of s-polarization, the random surfaces generated by its use retain the scattering properties for which they were designed when multiple scattering is taken into account and the incident light is p-polarized. To demonstrate this, we present in Fig. 2 plots of $\langle \partial R/\partial\theta_s \rangle$ as functions of θ_s for the scattering of both p- and s-polarized light from a one-dimensional perfectly conducting random surface designed to act as a band-limited uniform diffuser with $\theta_m = 20°$, obtained by means of rigorous computer simulation calculations[5] that take into account multiple-scattering processes of all orders. We see that the surface displays the scattering property for which it was designed for both polarizations of the incident light.

In Fig. 3 we present the results of a rigorous computer simulation calculation of $\langle \partial R/\partial\theta_s \rangle$ as a function of θ_s for s-polarized light incident normally on a one-dimensional random silver surface designed to act as a band-limited uniform diffuser with $\theta_m = 20°$. Results are presented for three wavelengths of the incident light: (a) $\lambda = 632.8$ nm (He-Ne laser); (b) $\lambda = 532$ nm (the second harmonic of the YAG laser); (c) $\lambda = 442$ nm (He-Cd laser). These wavelengths cover the entire visible region of the optical spectrum. We see that even when the scattering medium is a finitely conducting metal, a surface ruled on it in accordance with Eqs. (2), (4),

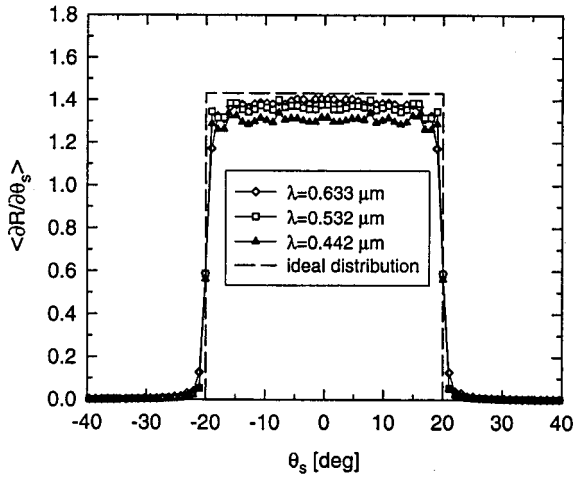

FIG. 3. The mean differential coefficient $\langle \partial R/\partial\theta_s \rangle$ estimated from $N_p = 40\,000$ realizations of metallic surface profiles in s-polarization: $\theta_m = 20°$, $b = 22\ \mu$m.

Downloaded 26 Mar 2003 to 128.200.81.38. Redistribution subject to AIP license or copyright, see http://ojps.aip.org/aplo/aplcr.jsp

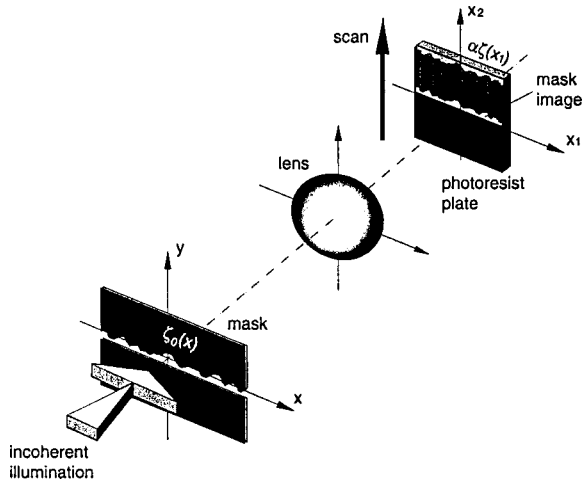

FIG. 4. Schematic diagram of the proposed experimental arrangement for the fabrication of surfaces with $f(\gamma)$ given by Eq. (6). The mask is imaged on the photoresist plate, which is then scanned along x_2.

and (8) still acts as a band-limited uniform diffuser, although with a weak dependence of the scattering pattern on the wavelength of the incident light. We attribute this dependence to the strong wavelength dependence of the dielectric function of silver in the range of wavelengths considered.

We conclude this paper by describing the manner in which one-dimensional surfaces of this kind can be fabricated on photoresist. First, a single realization of a profile function $\zeta_0(x)$, generated in accordance with Eqs. (2), (4), and (6), is used to fabricate a slit of variable width in the manner shown in Fig. 4. A good quality optical system is used to form an incoherent, demagnified image of the slit on the photoresist plate. Assuming that the object is resolvable, we express the intensity image on the photoresist plate as

$$I(x_1,x_2)=I_0\,\theta(x_2+d)\,\theta[\zeta(x_1)-x_2], \tag{9}$$

where I_0 is a constant, the coordinates x_1 and x_2 are fixed on the plate, and $\zeta(x_1)$ is a scaled version of the mask profile $\zeta_0(x)$. The plate is then scanned in the x_2 direction at speed v, producing a total exposure of the form

$$E(x_1)=K\int_{-T/2}^{T/2} I(x_1,x_2+vt)\,dt, \tag{10}$$

where K is a constant related to the sensitivity of the photoresist, $T=L_2/v$ is the time it takes to execute the scan, and the total scan length L_2 is assumed to be greater than the physical size of the plate. Then, the limits in Eq. (10) can be extended to infinity. Substitution of Eq. (9) into Eq. (10) gives

$$E(x_1)=E_0+\alpha\zeta(x_1), \tag{11}$$

where $E_0=KI_0d/v$ and $\alpha=KI_0/v$. This expression shows that the exposure has a linear dependence on the heights of the numerically generated realization of the surface profile function. Assuming that the relation between exposure and height is linear, the surface on the developed plate will have the desired properties. Note also that the vertical scale of the resulting profile can be adjusted through I_0 and the speed of the scan.

Thus, in this letter we have presented a method for generating a one-dimensional random surface that scatters light incident normally on it in such a way that the angular dependence of the intensity of the scattered light has a prescribed form. This method is simpler to implement than the one used in our earlier studies of the same problem.[1,2] Although it is based on a single-scattering approximation, for scattering from a perfectly conducting surface, we have shown by rigorous numerical simulation calculations that the surfaces generated by this approach yield the desired scattering pattern for both *p* and *s* polarizations of the incident light, when multiple scattering is taken into account, and when the perfect conductor is replaced by a finitely conducting metal as well. In addition, we have shown that the resulting scattering patterns are virtually independent of the wavelength of the incident light over a broad spectral range. Finally, we have described a procedure for fabricating such surfaces on photoresist.

The research of three of the authors (T.A.L., A.A.M., and E.R.M.) was supported in part by Army Research Office Grant No. 19-99-1-0321, and that of two of the authors (J.M.-L. and E.E.G.-G.) was supported by scholarships from CONACYT (México).

[1] E. R. Méndez, G. Martinez-Niconoff, A. A. Maradudin, and T. A. Leskova, Proc. SPIE **3426**, 2 (1998).
[2] E. R. Méndez, E. E. García-Guerrero, H. M. Escamilla, A. A. Maradudin, T. A. Leskova, and A. V. Shchegrov, Appl. Opt. **40**, 1098 (2001).
[3] E. I. Chaikina, E. E. García-Guerrero, Zu-Han Gu, T. A. Leskova, A. A. Maradudin, E. R. Méndez, and A. V. Shchegrov, in *Frontiers of Laser Physics and Quantum Optics*, edited by Z. Xu, S. Xie, S. Y. Zhu, and M. O. Scully (Springer, New York, 2000), p. 225.
[4] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd Edition (Cambridge University Press, New York, 1992), p. 281.
[5] A. A. Maradudin, T. Michel, A. R. McGurn, and E. R. Méndez, Ann. Phys. (N.Y.) **203**, 255 (1990).

Downloaded 26 Mar 2003 to 128.200.81.38. Redistribution subject to AIP license or copyright, see http://ojps.aip.org/aplo/aplcr.jsp

Photofabrication of random achromatic optical diffusers for uniform illumination

Eugenio R. Méndez, E. Efrén García-Guerrero, Héctor M. Escamilla, Alexei A. Maradudin, Tamara A. Leskova, and Andrei V. Shchegrov

We propose a method of designing two-dimensional random surfaces that scatter light uniformly within a specified range of angles and produce no scattering outside that range. The method is first tested by means of computer simulations. Then a procedure for fabricating such structures on photoresist is described, and light-scattering measurements with the fabricated samples are presented. The results validate the design procedure and show that the fabrication method is feasible. © 2001 Optical Society of America

OCIS codes: 290.0290, 030.5770, 050.1970.

1. Introduction

In many practical situations it is desirable to have optical diffusers with specific light-scattering properties. In particular, a nonabsorbing diffuser that scatters light uniformly within a range of angles and produces no scattering outside this range constitutes an attractive and useful optical element. We call this a band-limited uniform diffuser. Such an element would have applications, for example, in projection systems, in which it is important to produce even illumination without wasting light. Among other things, a band-limited uniform diffuser can also be useful in microscope illumination systems, the fabrication of displays and projection screens, and Fourier transform holography.

When this research was performed, E. R. Méndez (emendez@cicese.mx), E. E. García-Guerrero, and H. M. Escamilla were with the División de Física Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, Apartado Postal 2732, Ensenada, B. C. 22800, México. A. A. Maradudin was with the Department of Physics and Astronomy and the Institute for Surface and Interface Science, University of California, Irvine, Irvine, California 92697. T. A. Leskova was with the Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow Region 142092, Russia. A. V. Shchegrov was with the Rochester Theory Center for Optical Science and Engineering and the Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627-0171. He is now with KLA—Tencor Corporation, 50 Rio Robles M/S-K, San Jose, California 95134.

Received 17 July 2000; revised manuscript received 5 December 2000.

0003-6935/01/071098-11$15.00/0

© 2001 Optical Society of America

The design of band-limited uniform diffusers has been considered by several authors in the past. Using a two-level random height distribution on a surface, Kurtz[1] proposed a method for producing quasi-band-limited diffusers. An improvement was reported by Kurtz *et al.*,[2] who used a method based on the fabrication of deterministic microelement arrays with the appropriate slope distribution. A few years later, Nakayama and Kato[3] considered the case of pseudorandom diffusers, and Kowalczyk[4] reported some theoretical considerations on the more general random case.

More recently, diffractive optical elements that scatter light uniformly over specified angular regions have become commercially available.[5] These elements, however, are not random and possess the desired characteristics over only a relatively narrow range of wavelengths. In other words, they are not achromatic.

Another kind of diffuser, whose design is based on a randomized microlenslet concept, is also produced commercially.[6] Although these holographic light-shaping diffusers are achromatic and posses characteristics that approximate the desired ones, the scattering distribution they produce is not uniform, and they do not have a well-defined maximum angle of scattering.

Despite the interest in the subject, until recently there were no clear procedures reported in the literature for designing and fabricating random surfaces that behave as band-limited uniform diffusers, and it was unclear what kind of surface statistics were required for the production of such an optical element. Following previous work with one-dimensional sur-

faces,[7] in this paper we address these questions for a particular case of two-dimensional surfaces and propose an optical technique for fabricating diffusers with the desired characteristics. For simplicity, in the theoretical treatment we illustrate the ideas by considering first the scattering of a scalar wave from a random, impenetrable surface on which the Dirichlet boundary condition is satisfied. Initially the analysis is based on the Kirchhoff approximation, but we later adopt the thin-phase-screen model to extend the applicability of the analysis to other situations, such as that of the transmission of light through a dielectric diffuser. Working within this framework and further simplifying the results by taking the geometrical-optics limit, we propose a method for designing and fabricating achromatic, random optical diffusers with special characteristics. The results are validated by means of computer simulations as well as by the experimental realization of such diffusers.

The paper is organized as follows. First, in Section 2, we review some basic concepts of scattering theory in the Kirchhoff approximation and the thin-phase-screen model, taking the geometrical-optics limit. In agreement with other authors,[2] we find that the original problem of fabricating random band-limited uniform diffusers can be recast as that of generating random surface profiles with a rectangular probability density function (PDF) of slopes. Then, in Section 3, we address the problem of generating random profiles with the desired properties. The approach is tested first by means of the numerical simulations that are presented in Section 4. The proposed fabrication procedure is described in Section 5, in which we also present representative experimental results. Finally, in Section 6, we present our main conclusions.

2. Kirchhoff Approximation and the Thin-Phase-Screen Model

We approach the scattering problem in a formal way, by considering first the scattering of a scalar wave from a two-dimensional random surface that is defined by the equation $x_3 = \zeta(\mathbf{x}_\parallel)$, where $\mathbf{x}_\parallel = \hat{x}_1 x_1 + \hat{x}_2 x_2$, with $\hat{x}_1$ and $\hat{x}_2$ unit vectors in the x_1 and the x_2 directions, respectively, is a position vector in the plane $x_3 = 0$. The region $x_3 > \zeta(\mathbf{x}_\parallel)$ is vacuum, and the region $x_3 < \zeta(\mathbf{x}_\parallel)$ is an impenetrable medium. It is assumed that the Dirichlet boundary condition is satisfied on the surface $x_3 = \zeta(\mathbf{x}_\parallel)$. The surface-profile function is assumed to constitute a random process, but not necessarily a stationary one. The surface is illuminated from the vacuum side by an incident field given by

$$\psi(\mathbf{x})_{\mathrm{inc}} = \psi_0 \exp[i\mathbf{k}_\parallel \cdot \mathbf{x}_\parallel - i\alpha_0(k_\parallel)x_3], \tag{1}$$

where $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{k}_\parallel = (k_1, k_2, 0)$, $\alpha_0(k_\parallel) = [(\omega/c)^2 - k_\parallel{}^2]^{1/2}$, and ω is the frequency of the incident field. A time dependence of the form $\exp(-i\omega t)$ is assumed, but explicit reference to it is suppressed.

The scattered field in the far field can be written in the form

$$\psi(\mathbf{x})_{\mathrm{sc}} = \int_{-\infty}^{\infty} \frac{\mathrm{d}^2 q_\parallel}{(2\pi)^2} R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)\exp[i\mathbf{q}_\parallel \cdot \mathbf{x}_\parallel + i\alpha_0(q_\parallel)x_3], \tag{2}$$

where the two-dimensional wave vector $\mathbf{q}_\parallel = (q_1, q_2, 0)$, $\alpha_0(q_\parallel) = [(\omega/c)^2 - q_\parallel{}^2]^{1/2}$, and the scattering amplitude $R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)$ is given by

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) = \frac{1}{2i\alpha_0(q_\parallel)} \int_{-\infty}^{\infty} \mathrm{d}^2 x_\parallel{}' \exp[-i\mathbf{q}_\parallel \cdot \mathbf{x}_\parallel{}' - i\alpha_0(q_\parallel)\zeta(\mathbf{x}_\parallel{}')]F(\mathbf{x}_\parallel{}'). \tag{3}$$

Infinite limits are used to indicate that the region of integration is large in comparison with the wavelength and the lateral scale of the irregularities of the surface. In the Kirchhoff approximation, which we adopt now for simplicity, the source function $F(\mathbf{x}_\parallel)$ is given by

$$F(\mathbf{x}_\parallel) = 2\left(-\frac{\partial\zeta}{\partial x_1}\frac{\partial}{\partial x_1} - \frac{\partial\zeta}{\partial x_2}\frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3}\right)\psi(\mathbf{x})_{\mathrm{inc}}\Big|_{x_3=\zeta(\mathbf{x}_\parallel)}. \tag{4}$$

Then,

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) = \frac{-\psi_0}{\alpha_0(q_\parallel)} \int_{-\infty}^{\infty} \mathrm{d}^2 x_\parallel{}' \left[\frac{\partial\zeta}{\partial x_1}k_1 + \frac{\partial\zeta}{\partial x_2}k_2 + \alpha_0(k_\parallel)\right] \times \exp\{-i(\mathbf{q}_\parallel - \mathbf{k}_\parallel)\cdot\mathbf{x}_\parallel{}' - i[\alpha_0(q_\parallel) + \alpha_0(k_\parallel)]\zeta(\mathbf{x}_\parallel{}')\}. \tag{5}$$

After an integration by parts we may write[8]

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) = \psi_0 F_3(\mathbf{q}_\parallel, \mathbf{k}_\parallel) \int_{-\infty}^{\infty} \mathrm{d}^2 x_\parallel{}' \exp[-i(\mathbf{q}_\parallel - \mathbf{k}_\parallel)\cdot\mathbf{x}_\parallel{}' - ik_a(q_\parallel, k_\parallel)\zeta(\mathbf{x}_\parallel{}')], \tag{6}$$

where

$$F_3(\mathbf{q}_\parallel, \mathbf{k}_\parallel) = -\frac{(\omega/c)^2 + \alpha_0(q_\parallel)\alpha_0(k_\parallel) - \mathbf{q}_\parallel\cdot\mathbf{k}_\parallel}{\alpha_0(q_\parallel)[\alpha_0(q_\parallel) + \alpha_0(k_\parallel)]} \tag{7}$$

denotes an angular prefactor[8] and $k_a(q_\parallel, k_\parallel) = [\alpha_0(q_\parallel) + \alpha_0(k_\parallel)]$. Note that the term $\phi(\mathbf{x}_\parallel{}') = k_a(q_\parallel, k_\parallel)\zeta(\mathbf{x}_\parallel{}')$ may be interpreted as the phase acquired by the wave on reflection from the surface.

We also note that, for small angles of incidence and scattering, $F_3(\mathbf{q}_\parallel, \mathbf{k}_\parallel) \approx -1$ and $k_a(q_\parallel, k_\parallel)$ is practically independent of $\mathbf{q}_\parallel$ and $\mathbf{k}_\parallel$. Although it is not really necessary for the analysis, we use the small-angle approximation and assume that $k_a(q_\parallel, k_\parallel) \approx 2(\omega/c)$, which simplifies the resulting expressions and facilitates their interpretation. An analysis valid for larger angles of scattering can be found elsewhere.[9] With this simplification, our result takes the form assumed by the thin-phase-screen model,[10] which provides an approximate relationship between the surface-height variations and the phase acquired by the reflected wave front. The thin-phase-screen

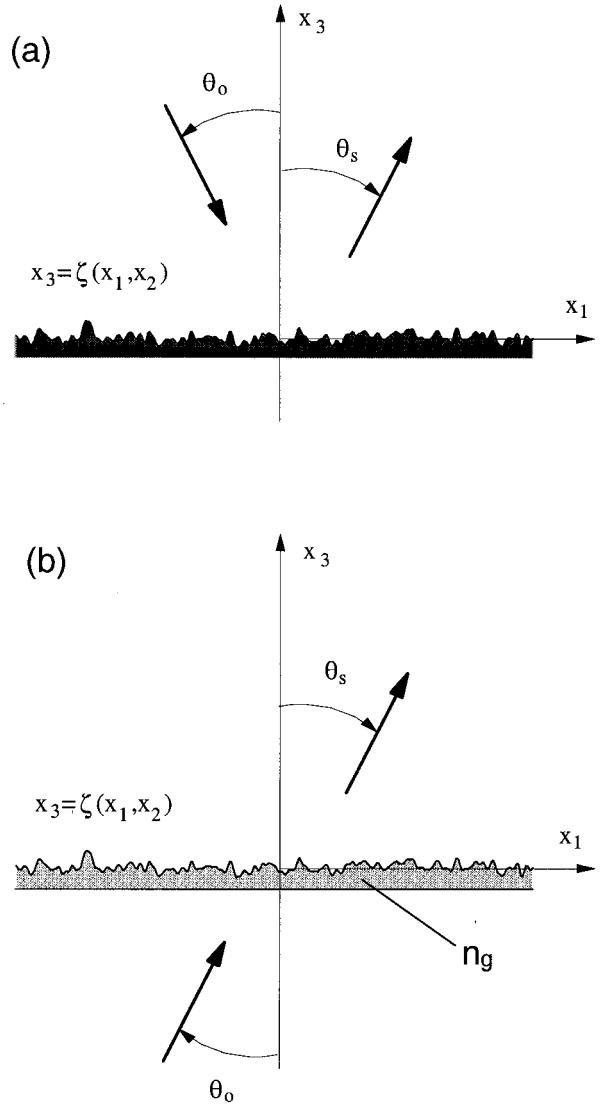

Fig. 1. Schematic diagram of the scattering geometry for (a) reflection by a surface and (b) transmission through a dielectric diffuser.

model represents a good approximation when the surface has only lateral features that are larger than the wavelength and does not produce significant amounts of multiple scattering. Thus, although approximate, our theoretical treatment is not limited to the reflection geometry considered and can be applied to other geometries and boundary conditions. So, restricting our attention to the region of applicability of the thin-phase-screen model, we can consider a variety of situations. In particular, our results can be applied to the case of reflection by a metallic surface or transmission through a slab of transparent dielectric material with one rough interface (see Fig. 1). For small angles of incidence and scattering, the phase introduced by the diffuser in the thin-phase-screen approximation is given by[11]

$$\phi(\mathbf{x}_\parallel') = k_a \zeta(\mathbf{x}_\parallel') = a\,\frac{\omega}{c}\,\zeta(\mathbf{x}_\parallel'), \tag{8}$$

with

$$a = \begin{cases} 2 & \text{in reflection} \\ (n_0 - n_g) & \text{in transmission} \end{cases}, \tag{9}$$

and where n_g is the refractive index of the material of which the slab is made and n_0 is that of the surrounding medium (normally air). Within this model, we can write

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) \approx \psi_0 \kappa_0 \int_{-\infty}^{\infty} \mathrm{d}^2 x_\parallel' \exp[-i(\mathbf{q}_\parallel - \mathbf{k}_\parallel)\cdot \mathbf{x}_\parallel' - ik_a \zeta(\mathbf{x}_\parallel')], \tag{10}$$

where κ_0 is a constant that accounts for the average reflectivity or transmissivity of the sample. This expression represents our basic starting equation that relates the surface profile with the angular spectrum of the scattered field.

At this stage, we restrict our attention to the case in which $\zeta(\mathbf{x}_\parallel) = \zeta_1(x_1) + \zeta_2(x_2)$, where $\zeta_1(x_1)$ and $\zeta_2(x_2)$ are statistically independent random processes. Then we can write

$$R(\mathbf{q}_\parallel|\mathbf{k}_\parallel) \approx \psi_0 \kappa_0 R_1(q_1|k_1) R_2(q_2|k_2), \tag{11}$$

with

$$R_j(q_j|k_j) = \int_{-\infty}^{\infty} \mathrm{d}x_j \exp[-i(q_j - k_j)x_j - ik_a \zeta_j(x_j)], \tag{12}$$

for $j = 1, 2$.

The mean differential reflection coefficient $\langle \partial R/\partial \Omega_s \rangle$ is defined as the fraction of the total time-averaged incident flux that is scattered into the element of solid angle $\mathrm{d}\Omega_s$ about a given scattering direction. In terms of $R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)$, it is given by

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = \frac{1}{|\psi_0|^2} \frac{1}{L_1 L_2} \frac{\omega/c}{(2\pi)^2} \frac{\alpha_0^2(q_\parallel)}{\alpha_0(k_\parallel)} \langle |R(\mathbf{q}_\parallel|\mathbf{k}_\parallel)|^2 \rangle, \tag{13}$$

where L_1 and L_2 are the lengths of the surface along the x_1 and the x_2 directions, respectively. In our small-angle approximation, the mean differential reflection coefficient may be written as

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle \approx |\kappa_0|^2 \frac{I_1(q_1|k_1)}{\lambda L_1} \frac{I_2(q_2|k_2)}{\lambda L_2}, \tag{14}$$

where we have defined

$$I_j(q_j|k_j) = \langle |R_j(q_j|k_j)|^2 \rangle. \tag{15}$$

By noting that the resulting expressions for $I_1(q_1|k_1)$ and $I_2(q_2|k_2)$ are formally identical for the two independent directions, we focus our attention on integrals of the form

$$I_j(q_j|k_j) = \int_{-\infty}^{\infty} \mathrm{d}x_j \int_{-\infty}^{\infty} \mathrm{d}x_j' \exp[-i(q_j - k_j)(x_j - x_j')] \times \langle \exp\{-ik_a[\zeta_j(x_j) - \zeta_j(x_j')]\} \rangle. \tag{16}$$

To simplify the notation, we omit the subindices in these generic expressions. With the change of variables $x' = x + u$, Eq. (16) becomes

$$I(q|k) = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}u \exp[i(q-k)u] \times \langle \exp\{-ik_a[\zeta(x) - \zeta(x+u)]\}\rangle. \quad (17)$$

Let us assume, for the moment, that the surface is statistically stationary. The average in Eq. (17) is then independent of x, and we can write

$$I(q|k) = L \int_{-\infty}^{\infty} \mathrm{d}u \exp[i(q-k)u] g(u), \quad (18)$$

where we have defined

$$g(u) = \langle \exp\{-ik_a[\zeta(x) - \zeta(x+u)]\}\rangle. \quad (19)$$

We are interested, for example, in having a uniform scattering distribution in the range $-\theta_m < \theta_s - \theta_0 < \theta_m$. In that case, we would seek an $I(q|k)$ of the form

$$I(q|k) = K \frac{1}{2\theta_m} \mathrm{rect}\left(\frac{\theta_s - \theta_0}{2\theta_m}\right) \approx K \frac{1}{2q_m} \mathrm{rect}\left(\frac{q-k}{2q_m}\right), \quad (20)$$

where $q_m = (\omega/c)\sin\theta_m$, K represents a constant with appropriate dimensions, rect represents the rectangle function,[12] and we have used the assumption that the angles of incidence and scattering are small. It is now clear, from Eq. (18), that the desired intensity distribution requires $g(u)$ to be a sinc function.

We obtain the geometrical-optics limit by expressing the difference $\zeta(x) - \zeta(x+u)$ in Eq. (19) in powers of u and retaining only the leading nonzero term. We find that

$$g(u) = \langle \exp[-ik_a u \zeta'(x)]\rangle, \quad (21)$$

which leads to

$$I(q|k) = L \int_{-\infty}^{\infty} \mathrm{d}u \langle \exp[-ik_a u \zeta'(x)]\rangle \exp[i(q-k)u] = \frac{2\pi L}{k_a} P_{\zeta'}\left(\frac{q-k}{k_a}\right), \quad (22)$$

where $P_{\zeta'}(x)$ represents the PDF of slopes on the surface. For small angles of incidence and scattering we have that

$$I(q|k) \approx \frac{2\pi L}{k_a} P_{\zeta'}\left(\frac{\theta_s - \theta_0}{a}\right). \quad (23)$$

These results show that the approximate form of Eq. (20) would be satisfied if the PDF of slopes had a rectangular shape. Because the random process that results by differentiation of a Gaussian random process is itself Gaussian, the statistics of the required profile cannot be Gaussian, at least within the geometrical-optics approximation. Furthermore, if we assume that $\zeta(x)$ is a Gaussian random process, the average entering Eq. (19) may be written as

$$g(u) = \exp\{-\delta^2 k_a{}^2[1 - c(u)]\}, \quad (24)$$

where δ represents the standard deviation of heights on the surface and $c(u)$ is the normalized height correlation function. We note that $g(u)$ is always positive and, thus, even with a physical optics treatment, Eq. (20) cannot be satisfied when $\zeta(x)$ is a Gaussian random process.

The result given by Eq. (22) shows that, in the geometrical-optics limit, our original problem can be restated as that of generating a random process, $\zeta(x)$, whose PDF of slopes is a rectangle function.

3. Design of the Diffuser

With this motivation, we consider non-Gaussian surfaces and address the problem of generating a profile with a rectangular PDF of slopes. Also, for generality, we relax the restriction of stationarity of the surface. Returning to Eq. (17) and taking the geometrical-optics limit, represented by Eq. (21), we write

$$I(q|k) \cong \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}u \exp[i(q-k)u] \times \langle \exp[-ik_a u \zeta'(x)]\rangle. \quad (25)$$

Because $\zeta(x)$ has been assumed to be nonstationary, we cannot assume that $\zeta'(x)$ is a stationary random process. The average $\langle \exp[-iau\zeta'(x)]\rangle$ therefore has to be assumed to be a function of x, and we cannot carry out the integral over x to yield L, as we did when $\zeta(x)$ was assumed to be stationary.

To evaluate the average in relation (25), we begin by expressing the surface-profile function in the form

$$\zeta(x) = \sum_{j=-\infty}^{\infty} c_j s(x - 2jb), \quad (26)$$

where the $\{c_j\}$ are positive, independent, random deviates. The function $s(x)$ is defined as (see Fig. 2)

$$s(x) = \begin{cases} 0 & \text{for } x \le -(n+1)b \\ -(n+1)H - mx & \text{for } -(n+1)b < x < -nb \\ -H & \text{for } -nb \le x \le nb \\ -(n+1)H + mx & \text{for } nb < x < (n+1)b \\ 0 & \text{for } (n+1)b \le x \end{cases}, \quad (27)$$

where n is a nonnegative integer, H represents a height, b denotes a characteristic length, and $m = H/b$ is a characteristic slope. We note that the function $s(x)$ defined by Eq. (27) is the negative of the function $s(x)$ defined in Ref. 7. Another difference is that in Ref. 7 the $\{c_j\}$ were chosen as zero-mean random deviates, whereas in this paper we assume that these numbers are positive random deviates. The present choices are determined by the fabrication process, which is described in Section 5.

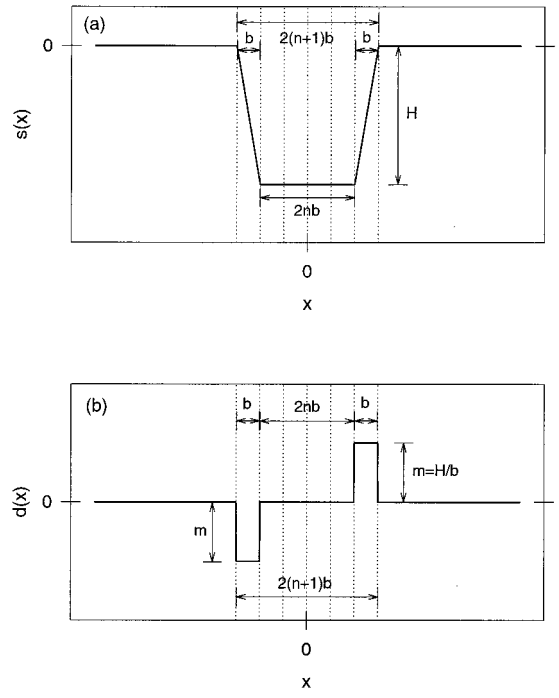

Fig. 2. Functions $s(x)$ and $d(x)$ for the case $n = 2$.

The derivative of the surface-profile function $\zeta'(x)$ is then given by

$$\zeta'(x) = \sum_{j=-\infty}^{\infty} c_j d(x - 2jb), \tag{28}$$

where

$$d(x) = \begin{cases} 0 & \text{for } x \le -(n+1)b \\ -m & \text{for } -(n+1)b < x < -nb \\ 0 & \text{for } -nb \le x \le nb \\ m & \text{for } nb < x < (n+1)b \\ 0 & \text{for } (n+1)b \le x \end{cases}, \tag{29}$$

The functions $s(x)$ and $d(x)$ are shown in Fig. 2.

When the PDF of the c_j, $f(\gamma)$, is known, a long sequence of the random numbers $\{c_j\}$ can be generated by, e.g., the rejection method.[13] Then, from this sequence, the surface-profile function $\zeta(x)$ and its derivative $\zeta'(x)$ can be obtained according to Eqs. (26) and (28), respectively.

The average $\langle \exp[-ik_a u\zeta'(x)] \rangle$ can now be written as

$$\begin{aligned} \langle \exp[ik_a u\zeta'(x)] \rangle &= \left\langle \exp\left[ik_a u \sum_{j=-\infty}^{\infty} c_j d(x - 2jb) \right] \right\rangle \\ &= \left\langle \prod_{j=-\infty}^{\infty} \exp[ik_a u c_j d(x - 2jb)] \right\rangle \\ &= \prod_{j=-\infty}^{\infty} \langle \exp[ik_a u c_j d(x - 2jb)] \rangle, \end{aligned} \tag{30}$$

where we have used the independence of the $\{c_j\}$ in the last step. The crucial point in the present derivation is that, with our choice of $d(x)$, for any value of x, only one factor in the infinite product on the right-hand side of Eq. (30) is different from unity. This is a direct consequence of the form chosen for the function $d(x)$, a situation that is illustrated in Fig. 3.

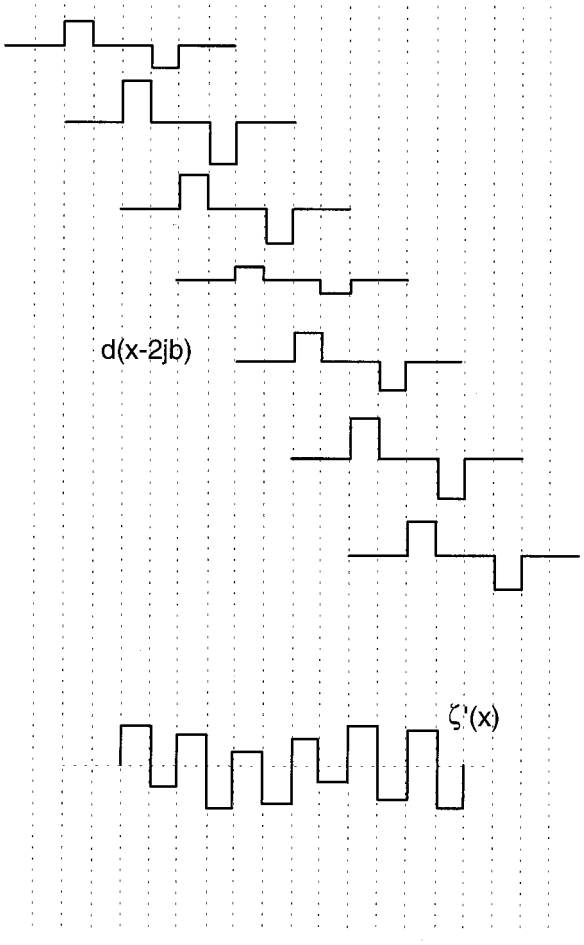

Fig. 3. Illustration of the sum represented by Eq. (28) with $n = 1$. At any one of the segments of length b shown in the figure there is only one nonzero contribution to $\zeta'(x)$.

Indeed, we find that, in the intervals $2jb < x - nb < (2j + 1)b$,

$$\begin{aligned} \langle \exp[ik_a u\zeta'(x)] \rangle &= \langle \exp(ik_a u m c_j) \rangle \\ &= \int_{-\infty}^{\infty} f(\gamma) \exp(ik_a u m \gamma) \mathrm{d}\gamma, \end{aligned} \tag{31}$$

whereas in the intervals $(2j - 1)b < x + nb < 2jb$,

$$\begin{aligned} \langle \exp[ik_a u\zeta'(x)] \rangle &= \langle \exp(-ik_a u m c_j) \rangle \\ &= \int_{-\infty}^{\infty} f(\gamma) \exp(-ik_a u m \gamma) \mathrm{d}\gamma. \end{aligned} \tag{32}$$

It follows from Eqs. (31) and (32) that relation (25) can be rewritten as

$$I(q|k) = \sum_j \int_{2jb}^{(2j+1)b} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}u \exp[i(q-k)u] \times \int_{-\infty}^{\infty} f(\gamma)\exp(ik_a u m\gamma)\mathrm{d}\gamma + \sum_j \int_{(2j-1)b}^{2jb} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}u \exp[i(q-k)u] \times \int_{-\infty}^{\infty} f(\gamma)\exp(-ik_a u m\gamma)\mathrm{d}\gamma. \quad (33)$$

Because the integrand in each integral over x is independent of x, the sum over j of each integral over x gives $L/2$. Thus we obtain

$$I(q|k) = \frac{L}{2}\int_{-\infty}^{\infty} \mathrm{d}u \exp[i(q-k)u] \int_{-\infty}^{\infty} f(\gamma)[\exp(ik_a\gamma m u) + \exp(-ik_a\gamma m u)]\mathrm{d}\gamma. \quad (34)$$

When the orders of integration are changed, Eq. (34) becomes

$$I(q|k) = \pi L \int_{-\infty}^{\infty} \mathrm{d}\gamma f(\gamma)[\delta(q-k+k_a\gamma m) + \delta(q-k-k_a\gamma m)] = \frac{\pi L}{k_a m}\left[f\left(\frac{k-q}{k_a m}\right) + f\left(\frac{q-k}{k_a m}\right)\right]. \quad (35)$$

When the results given by Eqs. (35) are substituted into approximation (14), we find that, for small angles of incidence and scattering, the mean differential reflection coefficient is given by

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = |\kappa_0|^2 \frac{1}{2am_1}\left[f_1\left(\frac{k_1-q_1}{k_a m_1}\right) + f_1\left(\frac{q_1-k_1}{k_a m_1}\right)\right] \times \frac{1}{2am_2}\left[f_2\left(\frac{k_2-q_2}{k_a m_2}\right) + f_2\left(\frac{q_2-k_2}{k_a m_2}\right)\right], \quad (36)$$

where the subindices (1, 2) indicate quantities along the directions $\hat{x}_1$ and $\hat{x}_2$. Thus we have obtained the result that, in the geometrical-optics limit, the mean differential reflection coefficient is determined by the PDF $f(\gamma)$ of the coefficients c_j entering the expansions (26) and (28). For small angles of incidence and scattering we can write

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = |\kappa_0|^2 \frac{1}{2am_1}\left[f_1\left(\frac{\theta_{10}-\theta_{1s}}{am_1}\right) + f_1\left(\frac{\theta_{1s}-\theta_{10}}{am_1}\right)\right] \times \frac{1}{2am_2}\left[f_2\left(\frac{\theta_{20}-\theta_{2s}}{am_2}\right) + f_2\left(\frac{\theta_{2s}-\theta_{20}}{am_2}\right)\right], \quad (37)$$

where θ_{10} and θ_{20} are the angles that the incident wave vector makes with the x_1 and the x_2 axes, respectively, and θ_{1s} and θ_{2s} are those of the scattered wave vector.

From this result we find that if we wish a constant value for $\langle \partial R_s/\partial\Omega_s \rangle$ for $-\theta_{1m} < \theta_{1s} - \theta_{10} < \theta_{1m}$ and $-\theta_{2m} < \theta_{2s} - \theta_{20} < \theta_{2m}$, subject to the restriction that $f(\gamma)$ is nonzero for $\gamma > 0$, so that each c_j is a positive number, we must choose

$$f(\gamma) = \mathrm{rect}\left(\gamma - \frac{1}{2}\right), \quad (38)$$

because with this choice for $f(\gamma)$ we obtain

$$\left\langle \frac{\partial R}{\partial \Omega_s} \right\rangle = |\kappa_0|^2 \frac{1}{2am_1}\left[\mathrm{rect}\left(\frac{\theta_{10}-\theta_{1s}}{am_1} - \frac{1}{2}\right) + \mathrm{rect}\left(\frac{\theta_{1s}-\theta_{10}}{am_1} - \frac{1}{2}\right)\right] \times \frac{1}{2am_2}\left[\mathrm{rect}\left(\frac{\theta_{20}-\theta_{2s}}{am_2} - \frac{1}{2}\right) + \mathrm{rect}\left(\frac{\theta_{2s}-\theta_{20}}{am_2} - \frac{1}{2}\right)\right] \quad (39)$$

$$= |\kappa_0|^2 \left[\frac{1}{2\theta_{1m}}\mathrm{rect}\left(\frac{\theta_{10}-\theta_{1s}}{2\theta_{1m}}\right)\right] \times \left[\frac{1}{2\theta_{2m}}\mathrm{rect}\left(\frac{\theta_{20}-\theta_{2s}}{2\theta_{2m}}\right)\right], \quad (40)$$

where we have made the identifications $\theta_{1m} = am_1$ and $\theta_{2m} = am_2$.

4. Numerical Simulations

The theory presented in the preceding sections was first tested by means of a computer simulation approach in which perfectly conducting surfaces were used. The simulation is based on the application of the Kirchhoff boundary conditions, which result in the scattering amplitude expressed by our Eq. (5). Randomly rough surface profiles and their derivatives were generated numerically, from the procedure represented by Eqs. (26) and (28), with the functions $s(x)$ and $d(x)$ given by Eqs. (27) and (29) and the PDF of the random numbers c_j given by Eq. (38). In Fig. 4 we show, as an example, a realization of a one-dimensional sample profile and its derivative. The mean differential reflection coefficient was estimated from Eqs. (5) and (13) with a large number of such profiles.

For simplicity, we consider first the scattering by one-dimensional surfaces; that is, surfaces that present height variations along x_1 and are invariant along x_2. In Fig. 5, we show an example of an estimated mean differential reflection coefficient for the

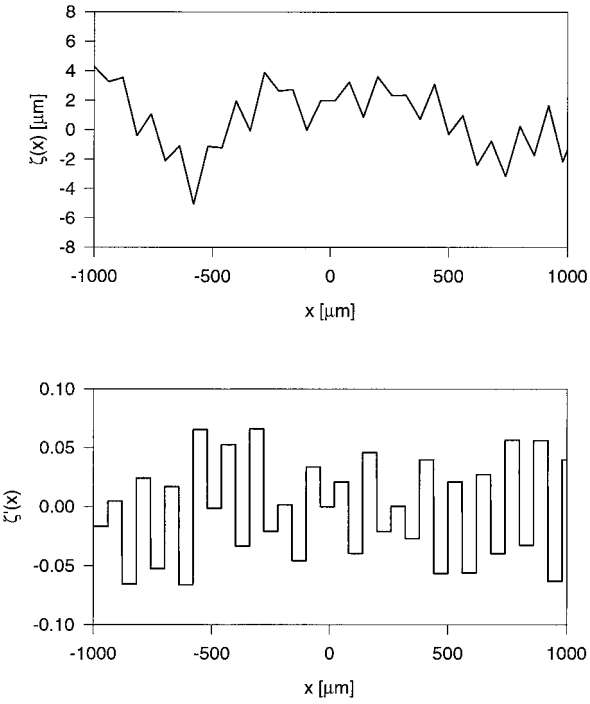

Fig. 4. Numerical generation of a surface profile and its derivative. The parameters used are $b = 60$ μm, $n = 1$, and $\theta_m = 2h = 8°$.

case of normal incidence, averaged over 6000 realizations of the surface-profile function. It can be seen that the scattering distribution is close to the desired result. There is almost no light outside the range $-\theta_m < \theta_s < \theta_m$, and, apart from a small peak in the direction $\theta_s = 0$, the distribution is fairly uniform. It should be pointed out that the peak is part of the incoherent component; the coherent component was also estimated with the simulation and is negligible in this case. As we discuss below, the peak is a diffraction effect that is not present in our theoretical results because our analysis is based on a geometrical-optics approximation.

Let us consider the result expressed by Eq. (39).

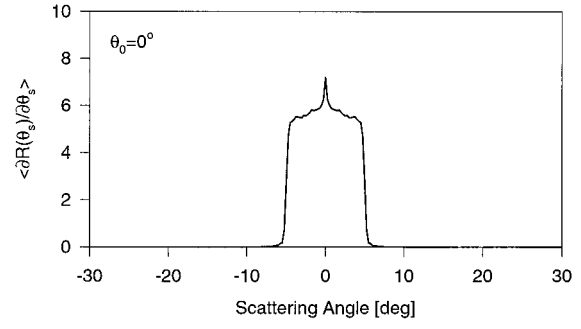

Fig. 5. Mean differential reflection coefficient estimated from $N_p = 6000$ realizations of the surface-profile function for the case of normal incidence and a wavelength $\lambda = 0.6328$ μm. The parameters used are $b = 60$ μm, $n = 1$, and $\theta_m = 2m = 5°$. The sampling on the surface was $\Delta x = \lambda/5$, and the size of the surface was $L = 2000\lambda$.

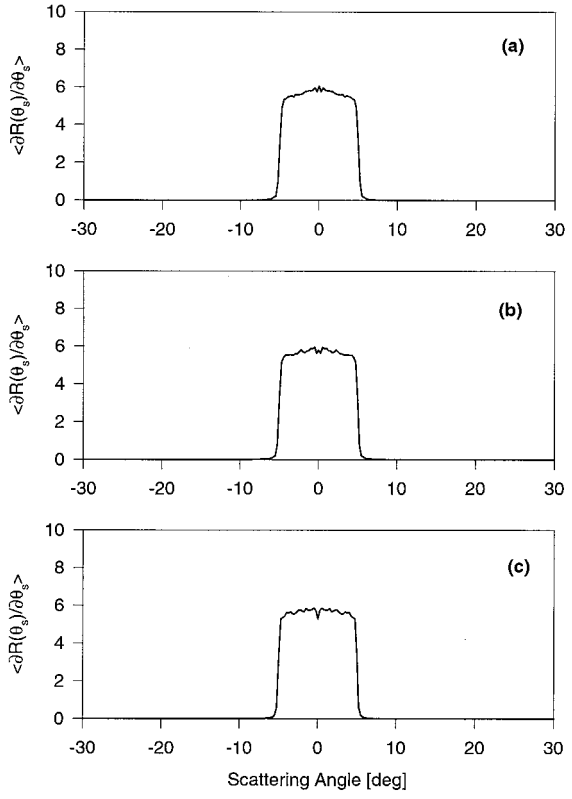

Fig. 6. Same as Fig. 5 but with random deviates c_j for the generation of the surface along x_1 drawn from the distribution given by Eq. (41) with $\epsilon = 0.01$. The illumination wavelengths are (a) $\lambda = 0.6328$ μm, (b) $\lambda = 0.532$ μm, and (c) $\lambda = 0.442$ μm.

In the geometrical-optics limit, the scattering distribution consists of two rectangular distributions, and it is clear that diffraction effects will smooth these two contributions to the scattering pattern. The peak observed along $\theta_s = 0$ in the scattering distribution shown in Fig. 5 is due to the overlap of the tails of the two distributions predicted on the basis of geometrical optics. This unwanted effect can be corrected by use of random numbers c_j for the generation of the surface, drawn from a distribution of the kind

$$f(\gamma) = \mathrm{rect}\left[\gamma - \left(\frac{1}{2} + \epsilon\right)\right], \tag{41}$$

where ϵ is a small constant. In our approximation, the scattering distribution is then given by

$$\left\langle \frac{\partial R}{\partial \theta_{1s}} \right\rangle = \frac{1}{2\theta_{1m}} \left\{ \mathrm{rect}\left[\frac{\theta_{10} - \theta_{1s}}{\theta_{1m}} - \left(\frac{1}{2} + \epsilon\right) \right] + \mathrm{rect}\left[\frac{\theta_{1s} - \theta_{10}}{\theta_{1m}} - \left(\frac{1}{2} + \epsilon\right) \right] \right\}. \tag{42}$$

This scattering pattern consists then of two contributions of rectangular form [nonzero in the ranges $-(\theta_{1m} + \epsilon\theta_{1m}) \le \theta_{1s} - \theta_{10} \le -\epsilon\theta_{1m}$ and $\epsilon\theta_{1m} \le \theta_{1s} - \theta_{10} \le (\theta_{1m} + \epsilon\theta_{1m})$], separated by a gap of angular width $2\epsilon\theta_{1m}$ about the specular direction.

In Fig. 6, we present computer simulation results

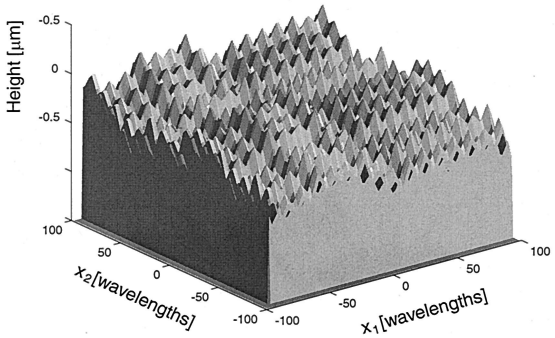

Fig. 7. Realization of a two-dimensional surface profile, $\zeta(\mathbf{x}_\parallel) = \zeta_1(x_1) + \zeta_2(x_2)$, with $b_1 = b_2 = 60$ μm, $n_1 = n_2 = 1$, and $\theta_{1m} = \theta_{2m} = 5°$.

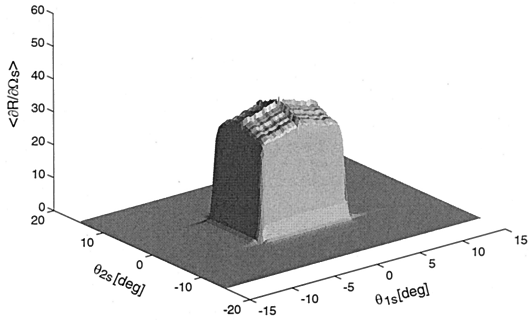

Fig. 8. Mean differential reflection coefficient estimated from $N_p = 2000$ realizations of the surface-profile function for the case of normal incidence and a wavelength $\lambda = 0.6328$ μm. The parameters used are $b_1 = b_2 = 60$ μm, $n_1 = n_2 = 1$, and $\theta_{1m} = \theta_{2m} = 5°$. The sampling on the surface was $\Delta x_1 = \Delta x_2 = \lambda/5$, and the size of the surface was $L_1 = L_2 = 2000\lambda$.

for the estimated mean differential reflection coefficients for surfaces we generated by using the PDF of the random numbers c_j given by Eq. (41) with $\epsilon = 0.01$, $\theta_m = 5°$, and three different wavelengths. Let us consider first the results presented in Fig. 6(a), which correspond to a wavelength $\lambda = 0.6328$ μm. The main difference between the distribution given by Eq. (42) and the results shown in Fig. 6(a) is that in the numerical results the two sections of the scattering distribution are not separated, as the theoretical analysis predicts. The overlap between the distributions is due to diffraction, and its relative importance is determined by the characteristic length b, the wavelength λ, and the value of ϵ. Therefore it is always possible to choose an appropriate ϵ to obtain an approximately flat scattering curve, as has been done in this case. Also in Fig. 6, we present calculations for the wavelengths $\lambda = 0.532$ μm [Fig. 6(b)] and $\lambda = 0.442$ μm [Fig. 6(c)]. These results show that the optimal ϵ does not change much as a function of the wavelength and that the diffusers are fairly achromatic. This fact constitutes an advantage of this kind of refractive (or reflective) diffusers over diffractive ones.

An example of a realization of a two-dimensional profile, $\zeta(\mathbf{x}_\parallel) = \zeta_1(x_1) + \zeta_2(x_2)$, with $\zeta_j(x_j)$ generated on the basis of Eq. (26), the function $s(x)$ given by Eq. (27), the PDF of the random numbers c_j given by Eq. (41), $\theta_{1m} = \theta_{2m} = 5°$, and $\epsilon = 0.01$, is shown in Fig. 7. The mean differential reflection coefficient estimated from $N_p = 2000$ realizations of such surfaces (only larger) is shown in Fig. 8 as a function of the two scattering angles θ_{1s} and θ_{2s} for the case of normal incidence. It can be seen that the scattering distribution is close to the desired result. There is almost no light outside the range $-\theta_{1m} < \theta_{1s} < \theta_{1m}$, $-\theta_{2m} < \theta_{2s} < \theta_{2m}$ and, inside it, the distribution is fairly uniform. These results show that, at least under the Kirchhoff approximation, our design approach is a sound one.

5. Experimental Details and Results

In this section, we provide some details of the experimental technique proposed for the fabrication of uniform band-limited diffusers and present our main experimental results.

We fabricated the surfaces by etching spin-coated photoresist films deposited on glass substrates through their selective exposure to blue light. A schematic diagram of the optical system used for exposing the plates is shown in Fig. 9. The illumination is provided by a He–Cd laser ($\lambda = 0.442$ μm). A cylindrical condenser system is used to concentrate the light passed through a rotating glass diffuser on a slit, providing an illumination that is effectively incoherent. An incoherent (normally demagnified) image of the slit is formed by a microscope objective on the photoresist-coated plate. In most cases we used a ×10 (N.A. = 0.25) microscope objective, but we also present results obtained with ×3 (N.A. = 0.1) and ×1 (N.A. = 0.05) objectives. The slit has an approximate width of $l = 180$ μm, and its incoherent image, of width l^*, has a nearly (smoothed by diffraction effects) rectangular shape. The intensity image, centered on x_0, is approximately given by the expression

$$I_s(x - x_0) \approx I_0 \,\mathrm{rect}\left(\frac{x - x_0}{l^*}\right), \tag{43}$$

where I_0 represents a constant.

Our first task is to produce a groove with a trapezoidal shape, such as the one shown in Fig. 2(a), on the photoresist-coated plate. For this, the plate is exposed while a constant speed scan of length $b = l^*/(2n + 1)$ is executed. To illustrate the procedure, we take the image of the slit centered on the origin.

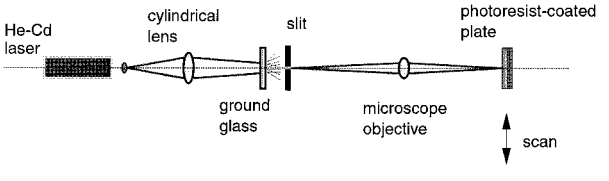

Fig. 9. Schematic diagram of the experimental arrangement used for the fabrication of the diffusers.

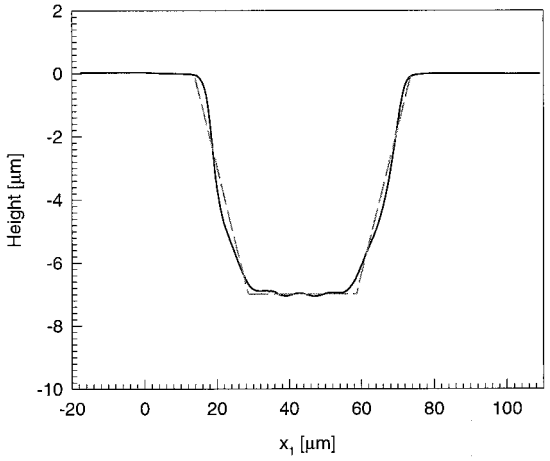

Fig. 10. Measured profile that illustrates the experimental realization of the function $s(x)$. The profile was estimated by means of a Dektakst mechanical profilometer and has parameters $H \approx 7$ μm, $b \approx 15$ μm, and $n = 1$. The ideal profile is shown by the dashed curve.

The exposure of the plate that is due to this distribution and the motion of the plate, of length b and time T, may be written as

$$E(x) = \int_{-T/2}^{T/2} I_s(x - vt)\mathrm{d}t, \tag{44}$$

where $v = b/T$ represents the speed of the scan. This equation may be rewritten as

$$E(x) = \frac{E_0}{b}\int_{-\infty}^{\infty} \mathrm{rect}\left(\frac{s}{b}\right)\mathrm{rect}\left[\frac{x - s}{(2n + 1)b}\right]\mathrm{d}s, \tag{45}$$

where we have defined $E_0 = I_0 T$ and used approximation (43) and the change of variable $s = vt$. Expression (45) can be readily evaluated, giving

$$E(x) = \begin{cases} 0 & \text{for } x \le -(n + 1)b \\ -(n + 1)E_0 - (E_0/b)x & \text{for } -(n + 1)b < x < -nb \\ -E_0 & \text{for } -nb \le x \le nb \\ -(n + 1)E_0 + (E_0/b)x & \text{for } nb < x < (n + 1)b \\ 0 & \text{for } (n + 1)b \le x \end{cases}. \tag{46}$$

The similarity between this function and the function $s(x)$, defined by Eq. (27), is evident. Therefore, if a linear response of the photoresist (depth proportional to exposure) is assumed, this procedure should produce a groove whose shape constitutes a good approximation to the desired trapezoidal one. To test and optimize this procedure, we fabricated several grooves by exposing photoresist-coated plates in this fashion. The plates were mounted on a computer-controlled motorized stage and moved in steps of 0.1 μm for the required length b. The plates were developed, dried, and the resulting surface profiles were subsequently estimated by a Dektakst mechanical profilometer fitted with a submicrometer stylus tip. An example of such a measurement is shown in Fig. 10, in which we present the estimated profile of a groove with $H \approx 7$ μm, $b \approx 15$ μm, and $n = 1$. For comparison, we show the ideal trapezoidal profile, indicated by the dashed curve. We can see that the corners of the fabricated profile are not as sharp as the ones in the ideal shape and that some of what should be straight-line segments are slightly curved (the lateral segments) or present oscillations (the bottom segment). It is difficult to tell how close the approximate and the ideal profiles need to be, but, from the results obtained with the fabricated surfaces, we can say that approximate profiles of the kind shown in Fig. 10 constitute acceptable approximations.

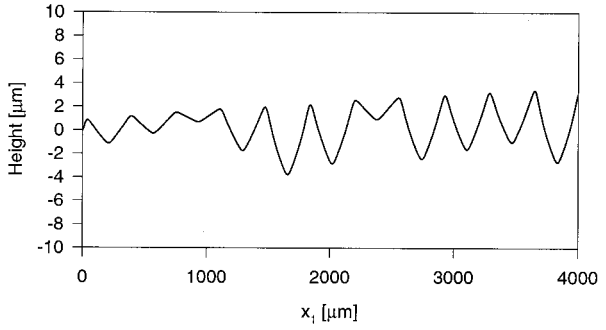

Fig. 11. Measured segment of the surface profile of a fabricated sample. The parameters are $b \approx 60$ μm and $n = 0$.

With the basic building block of the function $s(x)$, we can generate random profiles by superposing a sequence of shifted versions of it with random weights, as indicated by Eq. (26). Again, several profiles were generated and characterized to test and optimize the procedure. In Fig. 11, we present an example of a profilometric trace of a section of one of the samples fabricated according to Eq. (26). The faceted nature of the surface is clearly visible in the figure. For this particular example, we chose $n = 0$, which resulted in a function $s(x)$ of triangular shape, rather than a trapezoidal one. The symmetric triangular indentations can be clearly appreciated in the figure.

The profilometric data is useful for refining and establishing the fabrication procedure but, ultimately, the important characteristic of a given sample is its light-scattering properties. For this reason, angle-resolved light-scattering measurements with the fabricated samples were taken with a standard scatterometer (see descriptions in, e. g., Refs. 14 and 15). For convenience, the measurements were taken in a transmission geometry at normal incidence. An example is shown in Fig. 12, in which we show the scattering distribution pattern produced by a one-dimensional photoresist surface fabricated with the method described here. This particular sample has $b \approx 6$ μm and $n = 1$. From the figure we can see that the scattering distribution is significant only in the range $-10° < \theta_s < 10°$,

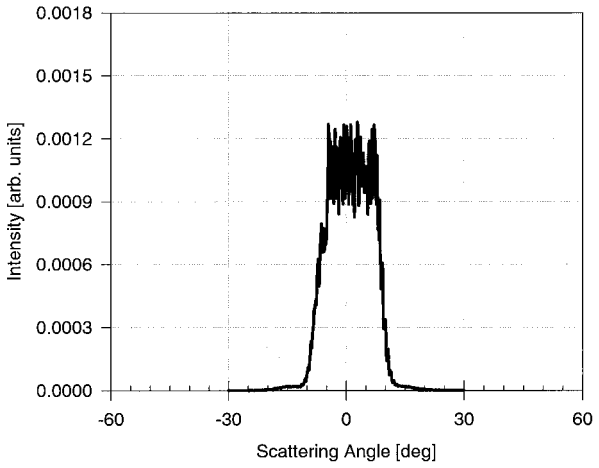

Fig. 12. Mean intensity (unnormalized differential reflection coefficient) produced by a sample with $b \approx 6$ μm and $n = 1$. The experiment was done in transmission, and the sample was illuminated at normal incidence.

approximately. In this range, the pattern presents some random oscillations with what appears to be a fairly uniform mean value. These oscillations are due to speckle noise, smoothed out by integration at the detector. However, because for the lengths of surface that we have fabricated only a few hundred independent random numbers were used, and not all of the facets provide a contribution to a point in the far field, the speckle statistics of the resulting speckle pattern could be non-Gaussian,[16] leading to larger fluctuations. It is expected that the use of longer surfaces, with more facets, will reduce the amplitude of these fluctuations.

Typically, the fabricated samples produce scattering patterns that are fairly well confined to a given angular range and they present some fluctuations in the region in which a constant intensity would be expected. The fluctuations, however, seem to have a fairly constant average value. These results validate the proposed design and fabrication methods for generating random band-limited uniform diffusers.

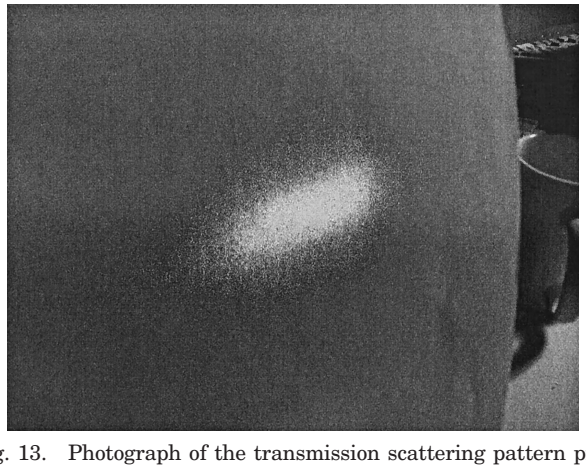

Fig. 13. Photograph of the transmission scattering pattern produced by a two-dimensional diffuser with $b_1 = b_2 \approx 6$ μm and $n_1 = n_2 = 1$.

A photograph of the scattering pattern obtained with a two-dimensional diffuser is shown in Fig. 13. Because the response of the detector of the electronic camera used to take this picture is not necessarily linear, the photograph provides only qualitative information on the scattering distribution. Nevertheless, a region of approximately rectangular shape, surrounded by a less intense and more isotropic diffuse halo, can be appreciated in the photograph. The halo is due to imperfections in the coating of the photoresist plate and to a small roughness component introduced by stray light during its exposure.

6. Summary and Conclusions

In this paper, we have studied the problem of designing and fabricating achromatic nonabsorbing diffusers that scatter uniformly within a range of angles and produce no scattering outside that range. It has been shown that, in the particular case of our separable geometry, the original problem can be recast as that of generating a random surface-profile function $\zeta(x_1, x_2) = \zeta_1(x_1) + \zeta_2(x_2)$ that, differentiated along the x_1 and the x_2 directions, obeys a rectangular PDF. A procedure for numerically generating surfaces with these properties has been proposed. Tests conducted through computer simulations showed that the proposed design is feasible and confirmed the expectation that the generated diffusers would be fairly achromatic. The technique is also well suited for the generation of such surfaces in photoresist, and we have implemented a procedure for their optical fabrication. The generated surfaces were characterized with a mechanical profilometer, and their scattering properties were measured with a scatterometer. The results indicate that good approximations to the desired uniform band-limited diffusers can be fabricated with the proposed method.

Although other methods of fabricating uniform band-limited diffusers have been proposed and are now in use for the production of light-shaping diffusers, we believe that the design procedure studied here is the only one that can produce achromatic diffusers whose scattering patterns show good uniformity within the specified angular range and, at the same time, can direct all of the available energy into that range.

We express our gratitude to Fabián Alonso and Juan A. Peralta for help in various aspects of this study. This research was supported in part by Consejo Nacional de Ciencia y Tecnología (Mexico) and by U.S. Army Research Office grants DAAH 04-96-1-0187 and DAAG55-98-C-0034.

References and Note

1. C. N. Kurtz, "Transmittance characteristics of surface diffusers and the design of nearly band-limited binary diffusers," J. Opt. Soc. Am. **62,** 929–989 (1972).

2. C. N. Kurtz, H. O. Hoadley, and J. J. DePalma, "Design and synthesis of random phase diffusers," J. Opt. Soc. Am. **63,** 1080–1092 (1973).
3. Y. Nakayama and M. Kato, "Linear recording of Fourier transform holograms using a pseudorandom diffuser," Appl. Opt. **21,** 1410–1418 (1982).
4. M. Kowalczyk, "Spectral and imaging properties of uniform diffusers," J. Opt. Soc. Am. A **1,** 192–200 (1984).
5. See, e.g., Thorlabs Inc. Newton, New Jersey 07860-0366.
6. Physical Optics Corp., Torrance, Calif. 90501-1821.
7. T. A. Leskova, A. A. Maradudin, I. V. Novikov, A. V. Shchegrov, and E. R. Méndez, "Design of one-dimensional band-limited uniform diffusers of light," Appl. Phys. Lett. **73,** 1943–1945 (1998).
8. P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (Pergamon, London, 1963), p. 20.
9. A. A. Maradudin, I. Simonsen, T. A. Leskova, and E. R. Méndez are preparing a manuscript to be called "Design of one-dimensional Lambertian diffusers of light."
10. W. T. Welford, "Optical estimation of statistics of surface roughness from light scattering measurements," Opt. Quantum Electron. **9,** 269–387 (1977).
11. Z. H. Gu, H. M. Escamilla, E. R. Méndez, A. A. Maradudin, J. Q. Lu, T. Michel, and M. Nieto-Vesperinas, "Interaction of two optical beams at a symmetric random surface," Appl. Opt. **31,** 5878–5889 (1992).
12. J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996), p. 13.
13. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge U. Press, New York, 1992), pp. 281–282.
14. K. A. O'Donnell and E. R. Méndez, "An experimental study of scattering from characterized random surfaces," J. Opt. Soc. Am. A **4,** 1194–1205 (1987).
15. R. E. Luna, E. R. Méndez, J. Q. Lu, and Z. H. Gu, "Enhanced backscattering due to total internal reflection at a dielectric–air interface," J. Mod. Opt. **42,** 257–269 (1995).
16. E. Jakeman and J. G. McWhirter, "Non-Gaussian scattering by a random phase screen," Appl. Phys. B **26,** 125–131 (1981).

Computer simulation studies of the speckle correlations of light scattered from a random array of scatterers: Scalar wave approximation

A. R. McGurn[1] and A. A. Maradudin[2]
[1]*Department of Physics, Western Michigan University, Kalamazoo, Michigan 49008*
[2]*Department of Physics and Astronomy, University of California, Irvine, California 92697*
(Received 7 July 2000; revised manuscript received 2 May 2001; published 5 October 2001)

Two computer simulation studies of the speckle correlations in the light scattered from a volume disordered dielectric medium consisting of a random array of dielectric spheres are made. In both studies light is treated in the scalar wave approximation, and the wavelength of the light is taken to be much greater than the radius of the dielectric spheres. In one study, the scattering medium is formed by placing dielectric spheres of radius R and dielectric constant ϵ randomly in space. The spheres occupy space uniformly, under the provision that no two spheres overlap. In a second study, the scattering medium is formed by placing dielectric spheres of radius R and dielectric constant ϵ randomly on the vertices of a simple cubic lattice so that a fixed fraction of the vertices is occupied by the spheres. The lattice constant of the simple cubic lattice is taken to be of the order of magnitude of the wavelength of light in vacuum. In both studies the the volume filling fraction is the same, and the region outside the spheres is vacuum. The field equations are integrated numerically to determine the scattered fields, and these fields are used to calculate the speckle correlation function defined by $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')=\langle[I(\vec{q}|\vec{k})-\langle I(\vec{q}|\vec{k})\rangle][I(\vec{q}'|\vec{k}')-\langle I(\vec{q}'|\vec{k}')]\rangle$. Here $I(\vec{q}|\vec{k})$ is proportional to the differential scattering coefficient for the elastic scattering of light of wave vector $\vec{k}$ into light of wave vector $\vec{q}$, and $\langle\ \rangle$ indicates an average over an ensemble of random systems. Results are presented for $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ with particular attention paid to regions of $\vec{k}$ space in which either the $C^{(1)}$ or $C^{(10)}$ contributions dominate the correlator.

DOI: 10.1103/PhysRevB.64.165204 PACS number(s): 71.55.Jv, 78.90.+t, 78.20.Bh

I. INTRODUCTION

In this paper we present computer simulation studies of features in the speckle pattern of light scattered by volume disordered media. Of specific interest are the correlations that exist between the differential scattering cross sections for two different sets of angles of incidence and scattering. These correlations are measured by the speckle correlator[1–10] $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')=\langle[I(\vec{q}|\vec{k})-\langle I(\vec{q}|\vec{k})\rangle][I(\vec{q}'|\vec{k}')-\langle I(\vec{q}'|\vec{k}')\rangle]\rangle$, where $I(\vec{q}|\vec{k})$ is proportional to the differential scattering coefficient for the elastic scattering of light of wave vector $\vec{k}$ into light of wave vector $\vec{q}$, and $\langle\ \rangle$ denotes an average over a statistical ensemble of random configurations. Two studies are presented, one for the light scattered by a homogeneous randomly disordered medium and one for the light scattered by a random medium that is periodic on average. The media we study are formed by placing dielectric spheres of radius R and dielectric constant ϵ in space at a fixed volume filling fraction. The spheres, which are not allowed to overlap, have a radius much smaller than the wavelength of the light in vacuum, and the region outside the spheres is taken to be vacuum.

The features in the speckle correlator, $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$, of the scattered light that are of interest are the short-range contributions $C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ (Refs. 1–9) and $C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$.[10–12] These features dominate $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$, so that to an excellent approximation we study $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')\approx C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')+C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$. The contribution $C^{(10)}$ is a feature in the speckle correlator for volume scattering that has recently been predicted on the basis of a diagrammatic perturbation theory for a homogeneous random disordered medium.[10–12] The general features of $C^{(1)}$ have been studied both theoretically and experimentally in other types of volume and surface disordered optical systems.[1–9]

The present computer simulation studies are intended to complement and support the Green's-function study of $C^{(10)}$ and to provide a comparison of this term to the $C^{(1)}$ term occurring in scattering from the same statistical system. In addition, different features of these correlators, associated with the periodic on average system, are presented. The computer simulation yields an essentially exact solution for the speckle correlator, and is not subject to the approximations of a diagrammatic perturbation-theory treatment. Since the general properties of the correlator are faithfully represented by simulation results, a meaningful comparison of $C^{(10)}$ with $C^{(1)}$ in the same bulk randomly disordered optical system can be made. An interesting aspect of the present set of simulation studies is the comparison of the results for the homogeneous random disordered and the periodic on average random system. The $C^{(1)}$ and $C^{(10)}$ terms of the periodic on average system are found to exhibit interesting features not observed in these terms for the uniformly random system.

Theoretical work on speckle correlations is concerned with computing the features in the speckle correlator arising from a variety of different types of scattering processes and classifying the contributions of these processes to the general speckle correlator $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$. Short-, long-, and infinite-range terms denoted by $C^{(1)},C^{(2)},C^{(3)}$, respectively, were first shown to contribute to the total speckle correlator of the

 ©2001 The American Physical Society

scattered light and to arise from distinctly different types of scattering processes.[1–12] The magnitude of the contribution of each of these terms, $C^{(i)}$ for $i=1,2,3$, was found to decrease rapidly with increasing i. More recently, additional features arising from different types of scattering processes have been predicted in the speckle correlator.[10–12] These are the $C^{(10)}$ short-range and the $C^{(1.5)}$ long-range terms. The $C^{(10)}$ contribution is found to be of the same order of magnitude as the $C^{(1)}$ contribution, but the magnitude of the $C^{(1.5)}$ term is found to be intermediate between those of the $C^{(1)}$ and $C^{(10)}$ terms and the much smaller $C^{(2)}$ term. For the systems we treat in this paper the $C^{(1.5)}$, $C^{(2)}$, $C^{(3)}$ contribution are masked by the statistical noise in our simulations.

$C^{(1)}$ first occurs in the lowest order in the perturbative expansion of the speckle correlator in powers of the volume disorder. $C^{(1)}$ contains phase coherent peaks in wave-vector space known as the memory and time-reversed memory effects in the angular speckle correlator.[1–9] $C^{(10)}$ first occurs in the same order in the expansion of the speckle correlator in powers of the volume disorder[10–12] as does $C^{(1)}$. Because of the scattering geometries involved, the $C^{(10)}$ term, however, does not display phase-coherent effects. $C^{(10)}$ has been studied by us in the scattering of light from randomly rough surfaces, and in the context of a Green's-function perturbation theory in the scattering from volume disordered media.[10]

The remaining $C^{(1.5)}$, $C^{(2)}$, and $C^{(3)}$ terms occur as increasingly higher-order terms in the perturbation-theory expansion of C in the random disorder. $C^{(1.5)}$ and $C^{(2)}$ contain a number of features that can be related to multiple-scattering processes.[10–12] $C^{(3)}$, which is the weakest contribution to the speckle correlator, is a smoothly varying function in wave-vector space.

The general speckle correlator is the sum of the five contributions mentioned above. This paper will concentrate on the $C^{(1)}$ and $C^{(10)}$ contributions. These dominate the speckle correlator, and are the only contributions to the speckle correlator observed in our simulation studies. A comparison of $C^{(1)}$ and $C^{(10)}$ is made for several values of the dielectric constant in the limit that the wavelength of the incident light is much larger than the radius of the dielectric spheres. A comparison is also made between $C^{(1)}$ and $C^{(10)}$ for a homogeneous random system and a random system that is periodic on average.

The outline of this paper is as follows. In Sec. II the model for bulk elastic scattering is presented, and the calculation of the scattering cross section in this model is discussed. In Sec. III the definition of the speckle correlator $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ is presented, and the $C^{(1)}$, $C^{(10)}$ contributions to it are discussed. Numerical results for the scattering cross section and the angular speckle correlator in the light elastically scattered are then presented and discussed. A discussion and comparison is given of the simulation results for the $C^{(1)}$ and $C^{(10)}$ contributions to the speckle correlator. In Sec. IV our conclusions are presented.

II. MODEL AND DIFFUSE SCATTERING

We consider a bulk random dielectric medium composed of an array of spheres of dielectric constant ϵ and radius R that are randomly distributed in space. The region outside the spheres is vacuum, and a volume filling fraction ρ of space is occupied by the spheres. Two different systems are studied. In the first system (homogeneous random system) the spheres randomly occupy space uniformly with the condition that they do not overlap. In the second system (periodic on average random system) the spheres randomly occupy the sites of a simple cubic lattice of lattice constant $a \gg R$. In both cases the wavelength of the light in vacuum is taken to be much larger than the radius of the spheres, i.e., $\lambda \gg R$. This is the fundamental restriction in our simulation and it is made to simplify the mathematical treatment of the scattering cross section of the electromagnetic waves from the individual spheres in the array of scatterers [see Eqs. (9)–(16) below].

Both of these systems are described by a position-dependent dielectric constant of the form

$$\epsilon(\vec{r}) = 1 + \delta\epsilon(\vec{r}), \tag{1}$$

where

$$\delta\epsilon(\vec{r}) = (\epsilon - 1)\sum_{l} S[\vec{r} - \vec{r}(l)], \tag{2}$$

$$S(\vec{r}) = \begin{cases} 1, & |\vec{r}| \leqslant R \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

and $\{\vec{r}(l)\}$ are the position vectors of the three-dimensional array of spheres. In the computer simulation studies presented below, a finite volume of the system described by Eq. (2) is considered. The far-field scattering from the finite volume of randomly arrayed spheres is computed and used to compute the speckle correlator. An average is made of the differential scattering coefficient and the speckle correlator over a large number of realizations of the random system. The resulting averages are taken as approximations of these quantities for the infinite system.

To simplify the mathematics, the propagation of light will be treated in the scalar wave approximation.[13] In this approximation the scalar wave field, $\psi(\vec{r},t) = \psi(\vec{r})\exp(-i\omega t)$, is determined by

$$\left[\Delta - \frac{\epsilon(\vec{r})}{c^2}\frac{\partial^2}{\partial t^2}\right]\psi(\vec{r},t) = 0, \tag{4}$$

the field energy density is given by

$$\rho(\vec{r}) = \frac{1}{8\pi}\left(\frac{\epsilon(\vec{r})}{c^2}\left|\frac{\partial\psi(\vec{r},t)}{\partial t}\right|^2 + |\nabla\psi(\vec{r},t)|^2\right), \tag{5}$$

and the energy current is given by

$$\vec{J}(\vec{r}) = -\frac{1}{8\pi}\left(\frac{\partial\psi^*(\vec{r},t)}{\partial t}\nabla\psi(\vec{r},t) + \frac{\partial\psi(\vec{r},t)}{\partial t}\nabla\psi^*(\vec{r},t)\right). \tag{6}$$

The solution to the scattering problem defined by Eqs. (1)–(6) can be formally written in terms of the Green's func-

tion of the Helmholtz equation for the propagation of the scalar wave field in vacuum. This Green's function is defined as the solution of

$$\left[\Delta+\left(\frac{\omega}{c}\right)^2\right]G(\vec{r},\vec{r'})=-4\pi\delta(\vec{r}-\vec{r'}), \tag{7}$$

subject to an outgoing wave boundary condition at infinity. The well-known solution is

$$G(\vec{r}|\vec{r'})=\frac{e^{i(\omega/c)|\vec{r}-\vec{r'}|}}{|\vec{r}-\vec{r'}|}. \tag{8}$$

In terms of $G(\vec{r}|\vec{r'})$ the scattering solution of Eq. (4) is

$$\psi(\vec{r})=\psi_{inc}(\vec{r})+\frac{1}{4\pi}\left(\frac{\omega}{c}\right)^2\int d^3r'G(\vec{r}|\vec{r'})\,\delta\epsilon(\vec{r'})\psi(\vec{r'}), \tag{9}$$

where $\psi_{inc}(\vec{r})$ is the incident field, which we shall take to be a plane wave. Substituting Eqs. (2) and (3) into Eq. (9), we find

$$\psi(\vec{r})=\psi_{inc}(\vec{r})+\frac{1}{4\pi}\left(\frac{\omega}{c}\right)^2(\epsilon-1)\sum_l\int d^3r'G(\vec{r}|\vec{r'}) \times S[\vec{r'}-\vec{r}(l)]\psi(\vec{r'}). \tag{10}$$

If the radius R of the spheres is small compared to the wavelength of the light in vacuum, $\psi(\vec{r'})$ varies slowly over the volume of the individual dielectric spheres, and we can rewrite Eq. (10) as

$$\psi(\vec{r})=\psi_{inc}(\vec{r})+\frac{\epsilon-1}{4\pi}\left(\frac{\omega}{c}\right)^2\sum_l\psi[\vec{r}(l)] \times\int_V d^3uG[\vec{r}|\vec{r}(l)+\vec{u}], \tag{11}$$

where V is the volume of a sphere of radius R centered at the origin of coordinates.

The differential scattering cross section for the scattering of scalar waves from the random volume disorder is obtained from the ratio of the scattered scalar wave current and the incident current. The scattering solution of Eq. (4) as represented in Eq. (9) is of the form

$$\psi(\vec{r})=\psi_{inc}(\vec{r})+\psi_{sc}(\vec{r}), \tag{12}$$

where $\psi_{sc}(\vec{r})$ is the scattered wave. For an incident plane wave of the form

$$\psi_{inc}(\vec{r})=\exp(i\vec{k}\cdot\vec{r}), \tag{13}$$

the scattered wave is given in the far field by

$$\psi_{sc}(\vec{r})=f(\hat{q},\hat{k})\frac{\exp(ik_0r)}{r}. \tag{14}$$

Here $k_0=\omega/c$, and $\vec{k}$ and $\vec{q}$ are the wave vectors of the incident and scattered fields. The scattering cross section is then

$$\frac{d\sigma}{d\Omega}(\vec{q}|\vec{k})=|f(\hat{q},\hat{k})|^2. \tag{15}$$

From Eq. (11) the field scattered from the array of spheres is given in the far field by

$$\psi_{sc}(\vec{r})\approx\frac{\epsilon-1}{4\pi}\left(\frac{\omega}{c}\right)^2 3V\frac{j_1(k_0R)}{k_0R}\frac{e^{ik_0r}}{r}\sum_{k,l}e^{-ik_0\hat{r}\cdot\vec{r}(k)} \times\left[\overleftrightarrow{I}-\frac{\epsilon-1}{4\pi}\left(\frac{\omega}{c}\right)^2\overleftrightarrow{M}\right]^{-1}_{\vec{r}(k),\vec{r}(l)}\psi_{inc}[\vec{r}(l)], \tag{16}$$

where the matrix elements of $\overleftrightarrow{M}$ at $\vec{r}(k),\vec{r}(l)$ are

$$M_{\vec{r}(k),\vec{r}(l)}=\int_V d^3uG[\vec{r}(k)|\vec{r}(l)+\vec{u}]. \tag{17}$$

In Eqs. (16) and (17) $\hat{r}=(\sin\theta_s\cos\phi_s,\sin\theta_s\sin\phi_s,\cos\theta_s)$, for polar and azimuthal scattering angles θ_s, ϕ_s, and $j_1(x)$ is the spherical Bessel function of order 1.

III. SPECKLE CORRELATIONS

The angular correlator for the scattering of light measures the correlations in the speckle patterns arising from two different scatterings of light from the same randomly disordered medium. The angular intensity correlator is defined in terms of the fluctuations in the differential scattering cross section given by Eq. (15) from its value averaged over many realizations of the random system. These fluctuations are

$$\delta I(\vec{q}|\vec{k})=\left[\frac{\partial\sigma}{\partial\Omega}(\vec{q}|\vec{k})-\left\langle\frac{\partial\sigma}{\partial\Omega}(\vec{q}|\vec{k})\right\rangle\right], \tag{18}$$

where $\langle\,\rangle$ represents an average over the ensemble of realizations of the random system and $I(\vec{q}|\vec{k})=(\partial\sigma/\partial\Omega)(\vec{q}|\vec{k})$. In terms of $\delta I(\vec{q}|\vec{k})$, the angular speckle correlator is

$$C(\vec{q},\vec{k}|\vec{q}',\vec{k}')=\langle\delta I(\vec{q}|\vec{k})\,\delta I(\vec{q}'|\vec{k}')\rangle =\left[\left\langle\frac{\partial\sigma}{\partial\Omega}(\vec{q}|\vec{k})\frac{\partial\sigma}{\partial\Omega}(\vec{q}'|\vec{k}')\right\rangle-\left\langle\frac{\partial\sigma}{\partial\Omega}(\vec{q}|\vec{k})\right\rangle \times\left\langle\frac{\partial\sigma}{\partial\Omega}(\vec{q}'|\vec{k}')\right\rangle\right]. \tag{19}$$

The scattering geometries involved in defining the correlator in Eq. (19) are shown in Fig. 1.

In the studies below results for the homogeneous random systems are obtained from runs involving 125 spheres, and results for the periodic on average studies are obtained from runs on systems of 225 dielectric spheres. The volume filling fraction in both of these studies is fixed at $\rho=0.000\,247$. The spheres occupy a cubic volume that is centered at the origin of spatial coordinates, with faces parallel to the $x-y$, $y-z$,

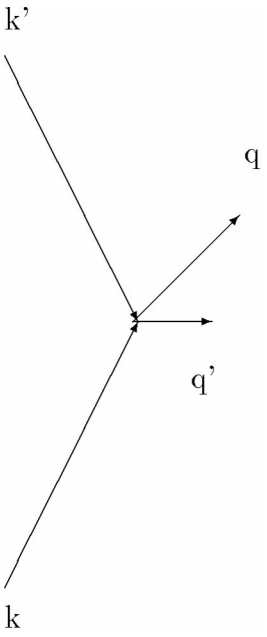

FIG. 1. Schematic representation of the wave-vector geometry for the correlation between two elastic scattering processes in three-dimensional space. One process involves the scattering of an incident wave with wave vector $\vec{k}$ into an outgoing wave with wave vector $\vec{q}$, and the other process involves the scattering of an incident wave with wave vector $\vec{k}'$ into an outgoing wave with wave vector $\vec{q}'$. The scattering system is at the vertex.

and $x-z$ planes. The radius of the spheres is taken to be $R/\lambda=0.064$, where λ is the wavelength of light in vacuum. For the majority of runs $\epsilon=-9$, but additional runs for $\epsilon=-2$, 2, 9 were also made. We have considered the cases of negative dielectric constants because metal particles can have negative dielectric constants. While the physical effects discussed below are observed for particles with both negative and positive dielectric constants, the effects show up best in systems of particles with negative dielectric constants. This should be a point of interest to those who would be interested in observing the effects predicted in the computer simulations studies presented below. For the periodic on average system the spheres are taken to occupy randomly the sites of a simple cubic lattice with lattice constant $a/\lambda \approx 1$. The speckle correlators are obtained by averaging over results from 500 or 1000 realizations of the disordered system.

For *s*-wave scattering from the spheres in our model, the solutions of the Helmholtz equation inside the spheres are proportional to the spherical Bessel function $j_0[(\epsilon)^{1/2}(\omega/c)r]$ for positive ϵ and the modified spherical Bessel function $i_0[(-\epsilon)^{1/2}(\omega/c)r]$ for negative ϵ. For positive dielectric spheres $j_0[(\epsilon)^{1/2}(\omega/c)R]/j_0(0)=\sin[(\epsilon)^{1/2}(\omega/c)R]/[(\epsilon)^{1/2}(\omega/c)R]$ is a measure of the validity of the approximation in going from Eq. (10) to Eq. (11). For $\epsilon=9$ this ratio is 0.77 and for $\epsilon=2$ this ratio is 0.95. For negative dielectric spheres $i_0[(-\epsilon)^{1/2}(\omega/c)R]/i_0(0)=\sinh[(-\epsilon)^{1/2}(\omega/c)R]/[(-\epsilon)^{1/2}(\omega/c)R]$ is a measure of the validity of the approximation made in going from Eq. (10) to Eq. (11). For $\epsilon=-9$ this ratio is 1.26 and for $\epsilon=-2$ this ratio is 1.05. The $\epsilon=\pm 9$ results are near the limits of the validity of the approximation but should give reliable representations of the speckle patterns from experimental realizations of these systems.

A. Scattering cross sections

We first present the differential scattering cross sections for the two different models. In Fig. 2(a) the differential scattering cross section is presented for the homogeneous random system. The wave vectors of the incident and scattered waves are, respectively, $\vec{k}=k_0(1,0,0)$ and $\vec{q}=k_0(\cos\phi_s,\sin\phi_s,0)$, and results are shown for several values of ϵ. In general the diffuse scattering cross sections are seen to be fairly isotropic in space. This is due to the small ratio of the dielectric sphere radius to the wavelength of light. This favors *s*-wave scattering from the individual spheres which, along with the uniform random distribution of spheres in space, gives rise to an isotropic cross section.

For comparison, in Fig. 2(b) the differential scattering cross section for several values of the dielectric constant is shown for a random system that is periodic on average. In these calculations we have taken $a/\lambda=1$. Consequently, the *s*-wave scattering from an individual dielectric sphere can be affected by the average periodicity of the system to give phase coherent peaks in the cross section. In general the scattering cross sections have a small isotropic component and four regularly spaced peaks. The peaks that are observed at $\phi_s=0$, 90, 270, and 360° arise from residual Bragg scattering caused by the average periodicity. The residual Bragg scattering peaks remain in the cross section when we change the wavelength of the light to slightly detune the light from the Bragg scattering condition. This is illustrated by the results plotted in Fig. 2(c) for a random system with $a/\lambda=0.9$ that is periodic on average.

B. $C^{(1)}$ and $C^{(10)}$ contributions to the angular speckle correlator

The contributions $C^{(1)}$ and $C^{(10)}$ to the speckle correlator are computed as statistical averages involving the differential scattering cross sections and their products. These contributions dominate $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ since they occur in the lowest order scattering processes yielding a contribution to $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$. A comparison of $C^{(1)}$ and $C^{(10)}$ is made for the two types of model disorder and in some cases for several values of ϵ.

In homogeneous random systems, the $C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ contribution has been shown to be nonzero only for processes that satisfy $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$, and the $C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ contribution has been shown to be nonzero only for processes that satisfy the wave-vector condition $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$. We first make some simulation studies that identify $C^{(1)}$ and $C^{(10)}$ as the dominant contributions to the speckle correlator in the homogeneous random system, and verify that $C^{(1)}$ and $C^{(10)}$ occur for wave vectors that satisfy these two conditions. To do this computer simulation runs

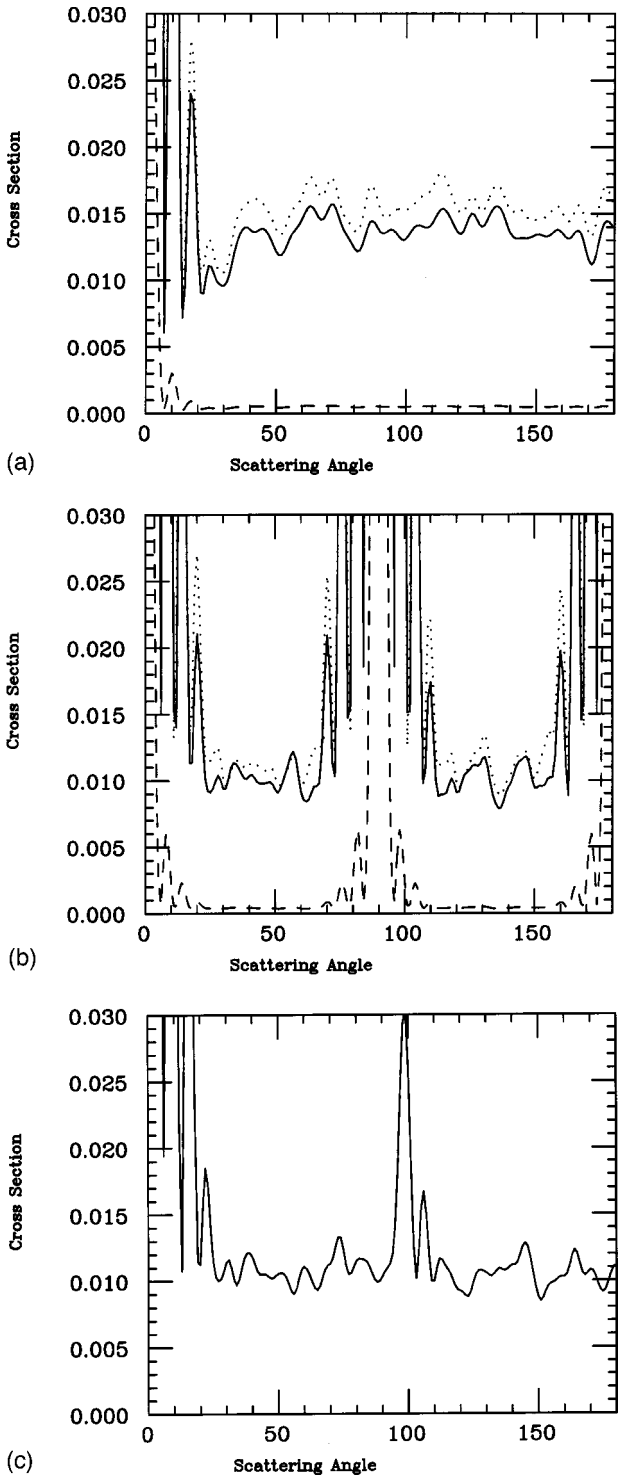

FIG. 2. (a) Differential scattering cross section versus azimuthal scattering angle ϕ_s for a uniformly random system. Results are shown for systems with $R/\lambda=0.064$, $\epsilon=-9$ (solid line), 2 (dashed line), and 9 (dotted). Note that the dotted line results have been multiplied by 0.1 to fit on the scale of the figure. (b) Differential scattering cross section for the periodic on average system with $a/\lambda=1$ and $R/\lambda=0.064$. The notation is the same as in (a). (c) Differential scattering cross section for the periodic on average system for $a/\lambda=0.9$, $R/\lambda=0.064$ for $\epsilon=-9$.

are made in which $\vec{q}$, $\vec{k}$ are fixed and a series of values of $\vec{q}'$ and $\vec{k}'$ are scanned through. During the scan, maxima are observed in the speckle correlator when either or both of the two wave-vector conditions given above are satisfied. These maxima correspond to the $C^{(1)}$ and $C^{(10)}$ contributions to the speckle correlator. Similar scans are made on the periodic on average system. The periodic on average system is found to exhibit additional peaks in $C^{(1)}$ and $C^{(10)}$ for $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'\neq 0$ and for $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'\neq 0$ that are not found in these functions for the homogeneous random system. The origins of these peaks will be discussed in Sec. III C.

In Fig. 3 results are presented of a scan on $\vec{q}'$, $\vec{k}'$ for fixed $\vec{q}$, $\vec{k}$ for $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ in the homogeneous disordered system. Plots for a number of different positive and negative values of ϵ are shown. The wave vectors of the incident and scattered light are taken to be in the $x-y$ plane and are of the form

$$\vec{k}=k_0(1,0,0), \tag{20}$$

$$\vec{k}'=k_0(\cos\phi,\sin\phi,0), \tag{21}$$

$$\vec{q}=k_0(0,1,0), \tag{22}$$

$$\vec{q}'=k_0(\sin\phi,-\cos\phi,0). \tag{23}$$

Here the azimuthal angle ϕ runs from 0 to 360°. The plots have maxima at $\phi=90°$, corresponding to wave vectors that satisfy $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$ for a nonzero $C^{(10)}$. A second set of maxima are found at $\phi=270°$ and satisfy $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$ for nonzero $C^{(1)}$. Figure 3 clearly display, the contributions of both $C^{(1)}$ and $C^{(10)}$ to the speckle correlator. It is interesting to note that the results for $|\epsilon|=9$ exhibit peaks in the correlator scans that are similar to the results for $|\epsilon|=2$. The intensities of these correlators, however, are quite different.

In Fig. 4 results are presented as in Fig. 3, but for a random system that is periodic on average. The wave vectors and the angle ϕ are defined as in Eqs. (20)–(23), and the average periodicity is defined such that $a/\lambda=1$. Unlike the homogeneous random system, in the scan over ϕ the periodic on average system displays four peaks instead of two peaks. The two additional peaks at $\phi=180°$ and 360° arise from the average periodicity of the array of spheres, and correspond to scattering sequences in which at least one Bragg reflection is involved. A further discussion of the origin of of these peaks will be given below in subsection C. Well defined peaks are observed in the $|\epsilon|=9$ plots, but a more complicated behavior involving a central peak with side dips is found in the $|\epsilon|=2$ plots. The widths and side peaks observed in Figs. 3 and 4 are for angular scans passing though, not along, the envelopes of $C^{(1)}$ and $C^{(10)}$, and arise from finite-size effects. For an infinite sample the central peaks would become delta functions and the side peaks would disappear. For finite-sized samples, the side peaks arise from diffractive effects due to the finite sample size.

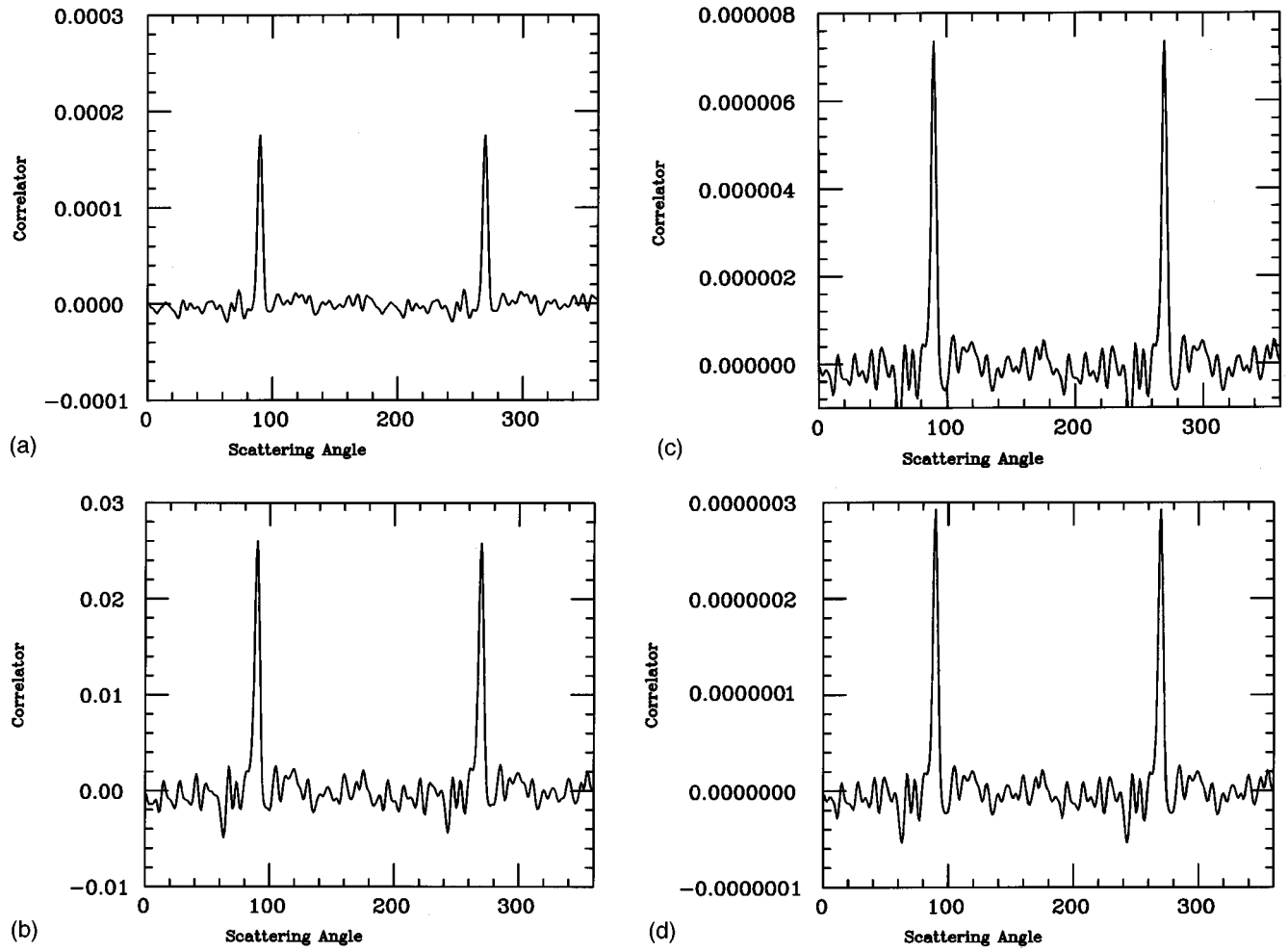

FIG. 3. $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ versus the scattering angle ϕ for the uniformly disordered system. Results are shown for $\epsilon=$: (a) -9; (b) 9; (c) -2; and (d) 2.

Figures 3 and 4 exhibit the existence of the $C^{(1)}$ and $C^{(10)}$ functions, which are nonzero when $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$ and $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$, respectively. We now look at the envelopes of these functions. The envelopes are obtained as $\vec{q}$, $\vec{k}$, $\vec{q}'$, $\vec{k}'$ are varied so that either or both of the conditions $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$, $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$ are always satisfied.

Results are first presented for the envelope function of $C^{(1)}$. We account for the $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$ restriction on $C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ by writing

$$\vec{k}=k_0(\sin\theta,0,\cos\theta), \tag{24}$$

$$\vec{k}'=k_0(\sin\theta,0,-\cos\theta), \tag{25}$$

$$\vec{q}=k_0(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta), \tag{26}$$

$$\vec{q}'=k_0(\sin\theta\cos\phi,\sin\theta\sin\phi,-\cos\theta). \tag{27}$$

In this representation $\vec{q}-\vec{k}$ and $\vec{q}'-\vec{k}'$ are parallel vectors that are sent into one another by reflection through the $x-y$ plane. They lie on chords of the sphere of radius k_0 centered at the origin of coordinates in wave-vector space. This is a natural way to choose vectors satisfying $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$.

In Fig. 5 the speckle correlator of the homogeneous random system is shown as a function of the polar angle θ for fixed azimuthal angles $\phi=0$, 20, and 90°. For the choice of wave vectors in Eqs. (24)–(27), the correlator is symmetric under reflection through the $x-y$ plane so that results are shown only for $0\leqslant\theta\leqslant 90°$. For $\phi=0°$, peaks are observed at $\theta=0°$ and $\theta=90°$. These correspond to light that is scattered along time reversed $[\vec{q}=\vec{k}=-\vec{q}'=-\vec{k}'=k_0(0,0,1)]$ and same sequence scattering $[\vec{q}=\vec{k}=\vec{q}'=\vec{k}'=k_0(1,0,0)]$ paths, respectively. It is expected in Fig. 5 and the other figures presented in this paper that time reversed and same sequence scattering process may lead to enhanced correlations. In the case of same sequence scattering, two waves that undergo exactly the same scattering process in traversing a disordered medium should undergo the same change in phase from the initial to the final scattering. This retention of the phase difference before and after scattering should contribute to a constructive interference in the intensity which would tend to increase the intensity correlator when same

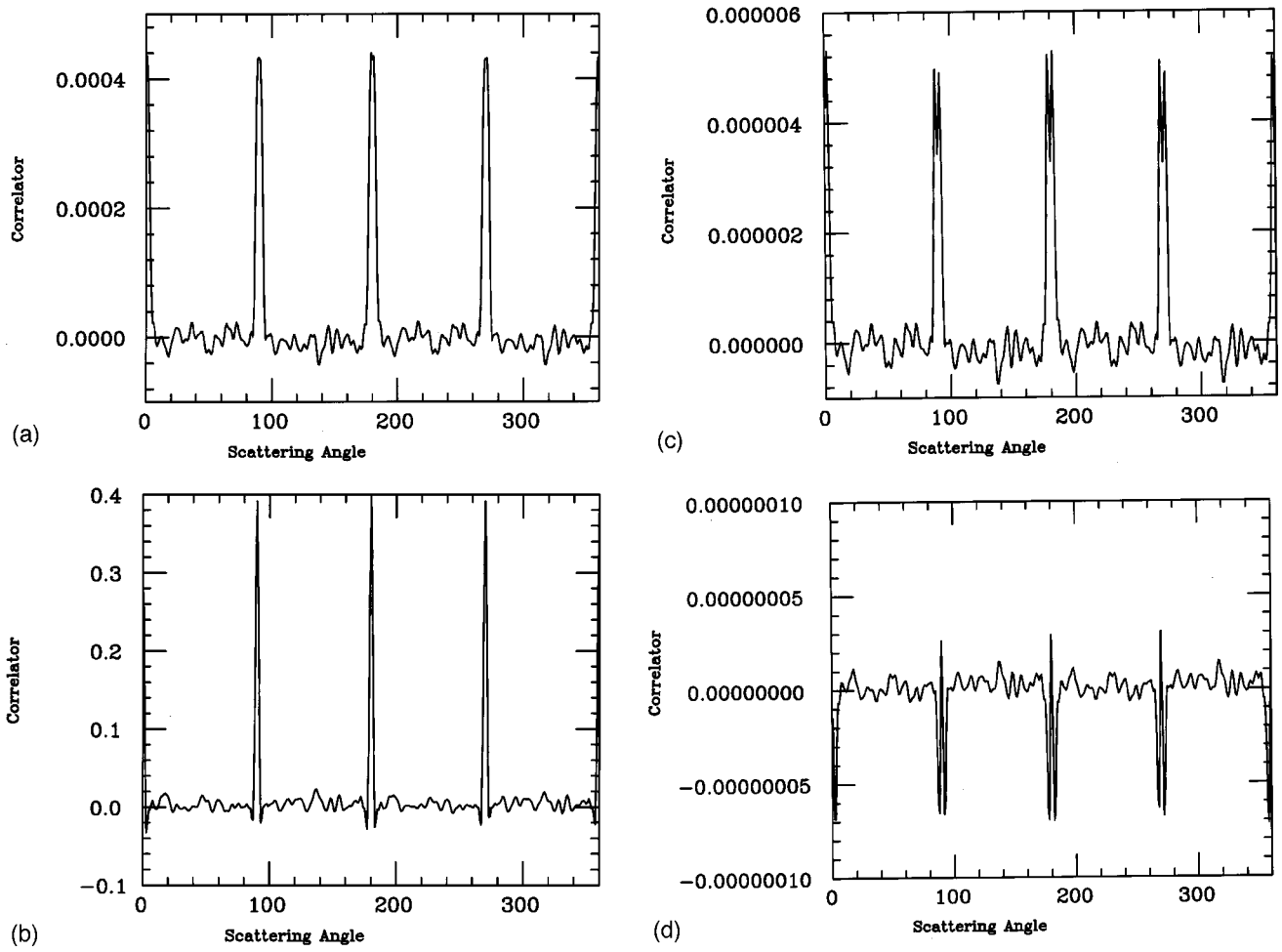

FIG. 4. $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ versus the scattering angle ϕ for the periodic on average system. Results are shown for $\epsilon=$: (a) -9; (b) 9; (c) -2; and (d) 2.

sequence scattering process are possible. Same sequence scattering processes can only occur for a restricted range of $\vec{q}$, $\vec{q}'$, $\vec{k}$, and $\vec{k}'$ values. Likewise, for time reversed processes two waves encounter the scatterers of the medium in an exactly reversed sequence, and the retention of the relative phase difference of the two waves is expected to enhance the intensity correlations between the processes scattering the two waves. Once again time reversed processes can only occur for a restricted range of $\vec{q}$, $\vec{q}'$, $\vec{k}$, and $\vec{k}'$. For $\phi=20°$, a small peak is observed at $\theta=0°$ corresponding to the time reversed sequence $[\vec{q}=\vec{k}=-\vec{q}'=-\vec{k}'=k_0(0,0,1)]$, but no peak is found associated with the $\theta=90°$ same sequence scattering processes. For $\phi=90°$ a peak structure is observed near the time reversed $\theta=0$ sequence. Otherwise, the $\phi=90°$ correlator is a smoothly varying function of θ. It is noted in these plots that both $C^{(1)}$ and $C^{(10)}$ are nonzero at general ϕ for $\theta=0°$ and at $\phi=90°$ for $\theta=90°$. At these points the correlator has contributions from both the $C^{(1)}$ and $C^{(10)}$ processes.

To study the $C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ contribution to the angular speckle correlator, the $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$ restriction on the wave vectors $\vec{q},\vec{k},\vec{q}',\vec{k}'$ is accounted for by writing

$$\vec{k}=k_0(\sin\theta,0,\cos\theta), \tag{28}$$

$$\vec{k}'=k_0(-\sin\theta,0,\cos\theta), \tag{29}$$

$$\vec{q}=k_0(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta), \tag{30}$$

$$\vec{q}'=k_0(-\sin\theta\cos\phi,-\sin\theta\sin\phi,\cos\theta). \tag{31}$$

The vectors $\vec{q}-\vec{k}$ and $\vec{q}'-\vec{k}'$ are now antiparallel vectors in a plane of constant polar angle of the sphere of radius k_0 centered at the origin of coordinates in wave-vector space, and occupy parallel chords of the sphere. This is again a natural representation for wave vectors that satisfy $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$. Here θ is the polar angle of the plane containing $\vec{q}-\vec{k}$ and $\vec{q}'-\vec{k}'$, and ϕ is the azimuthal angle in this plane.

The results for $C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ are plotted in Fig. 6 versus θ for $\phi=0$, 20, and 90°. These results are, again, invariant under reflection in the $x-y$ plane so that results are shown only for $0\leqslant\theta\leqslant90°$. For the case in which $\phi=0°$ peaks are observed for $\theta=0°$ corresponding to $\vec{q}=\vec{k}=\vec{q}'$

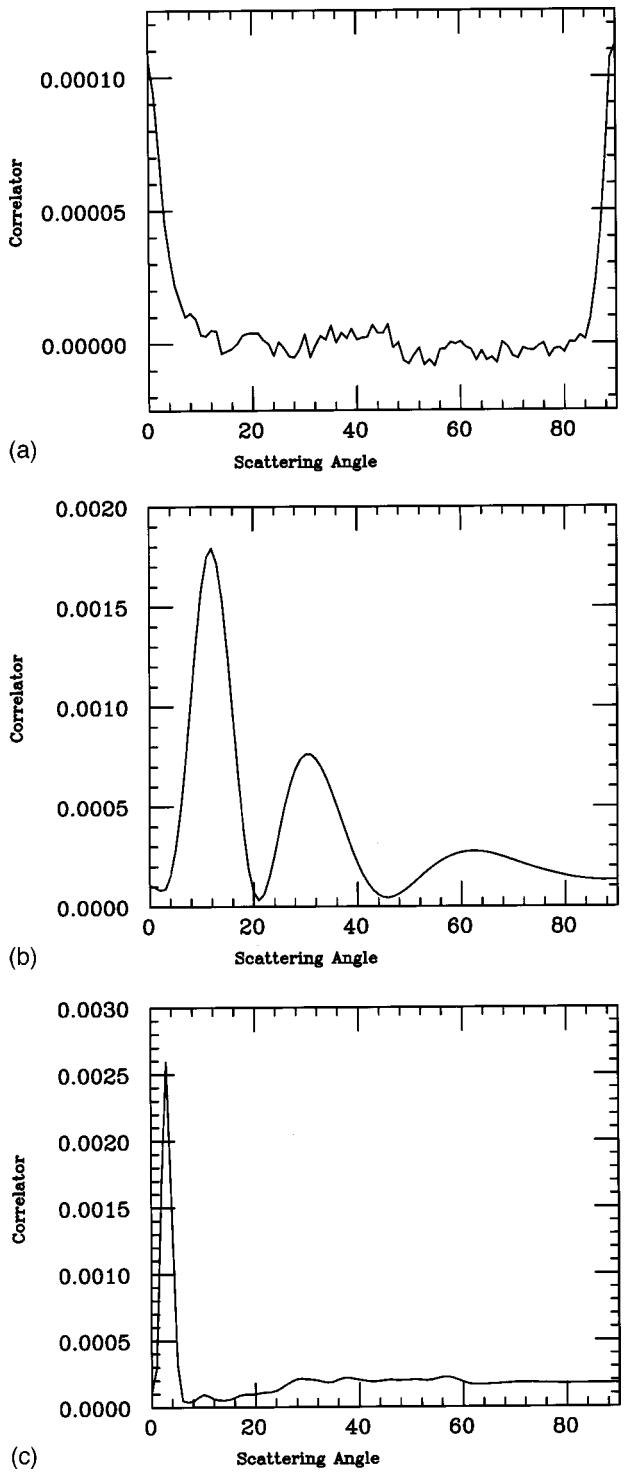

FIG. 5. $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ versus θ for the uniformly disordered system. Results are shown for $\phi=$: (a) 0°; (b) 20°; and (c) 90°.

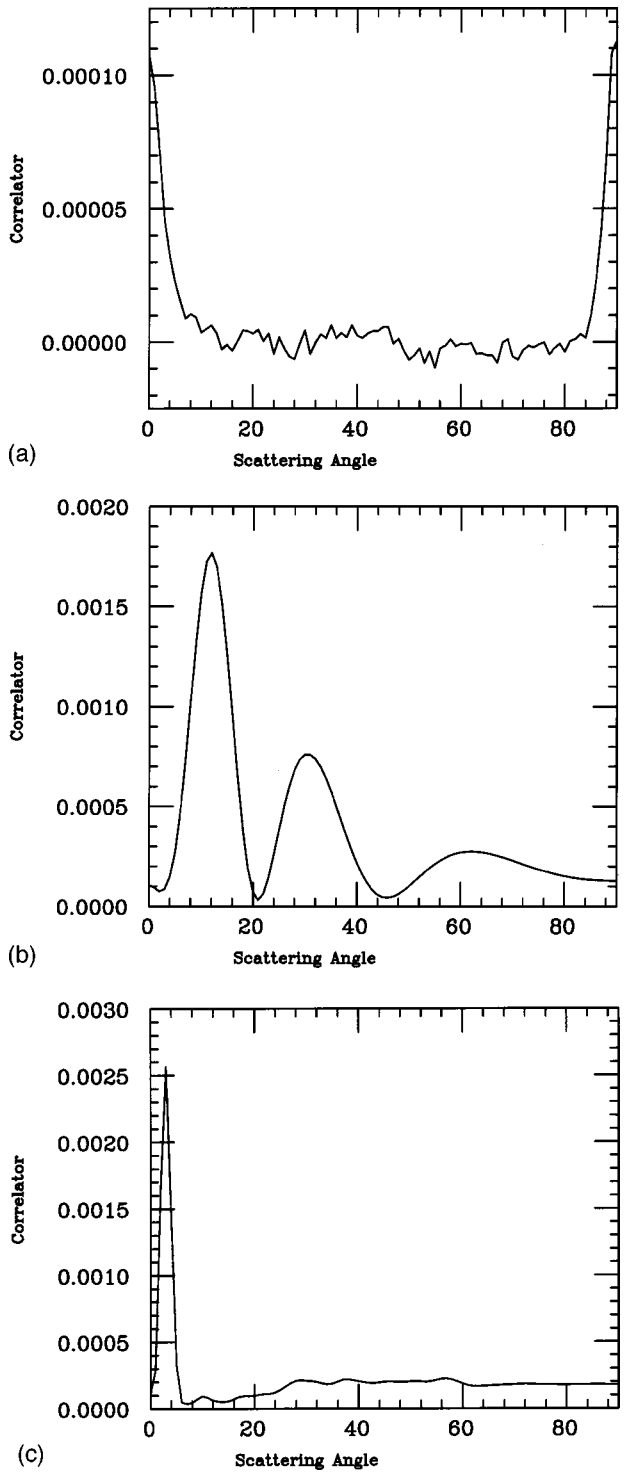

FIG. 6. $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ versus θ for the uniformly disordered system. Results are shown for $\phi=$: (a) 0°; (b) 20°; and (c) 90°.

$=\vec{k}'=k_0(0,0,1)$ (same sequence scattering), and for $\theta=90°$ corresponding to $\vec{q}=\vec{k}=-\vec{q}'=-\vec{k}'=k_0(1,0,0)$ (time reversed scattering). At these points, in addition to the phase coherence of the same sequence and time reversed scattering processes, as per our discussion of Fig. 5, both $C^{(1)}$ and $C^{(10)}$ scattering processes make nonzero contributions to $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$, and the increase in the number of types of correlated scattering processes contributes to an enhancement of $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$. For general ϕ and $\theta=0°$ both $C^{(1)}$ and $C^{(10)}$ are nonzero and arise from same sequence scattering paths. Peaks are generally observed in Figs. 6 at $\theta=0°$.

In Figs. 7 and 8 we present results for the periodic on average system corresponding to those in Figs. 5 and 6, respectively, for the homogeneous random system. In these

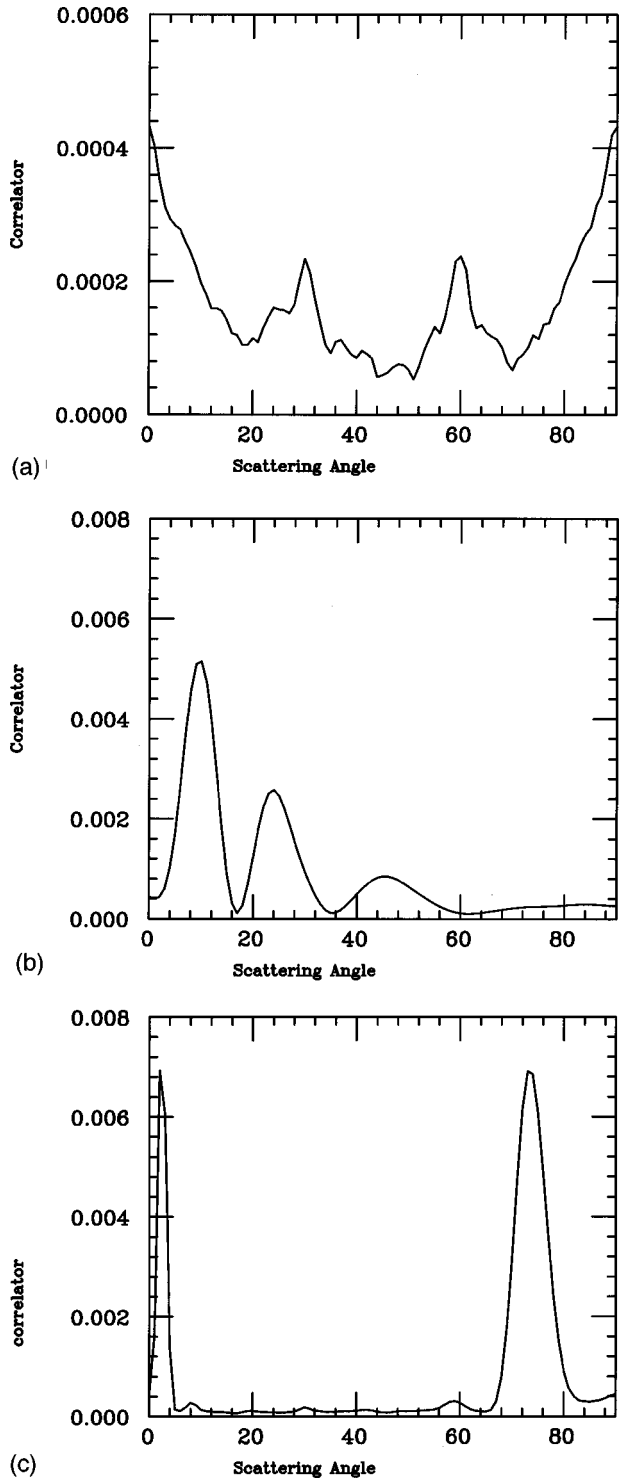

FIG. 7. Results as in Fig. 5 but for the periodic on average system.

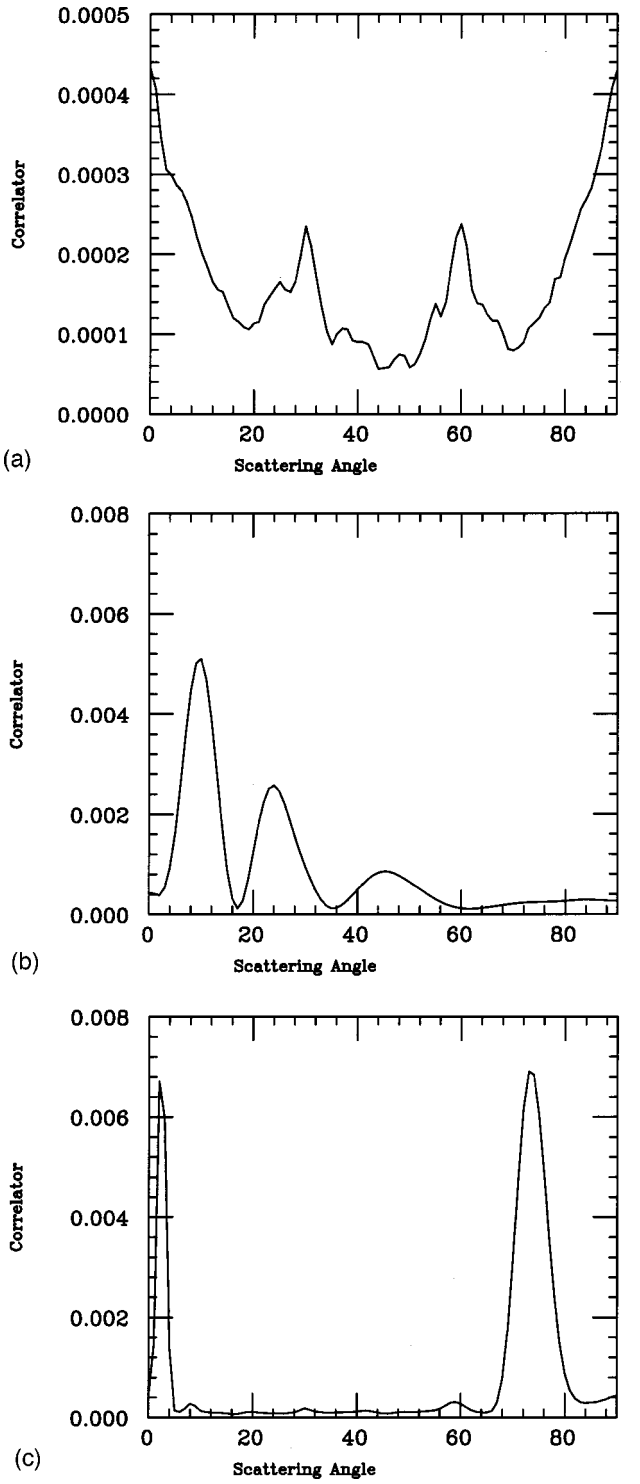

FIG. 8. Results as in Fig. 6 but for the periodic on average system.

plots either or both of the conditions $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$ or $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$ are satisfied. The plots are very similar to the results for the uniform isotropic system. An exception is the $\phi=90°$ results for both $C^{(1)}$ and $C^{(10)}$. For this case peaks are observed for $C^{(1)}$ along a same sequence path and for $C^{(10)}$ along paths of antiparallel incident and scattering beams. Associated with these peaks are two large side peaks arising from residual Bragg scattering effects.

Finally, we note that in plotting the envelope functions in Figs. 5–8 we fixed the azimuthal angle, ϕ, of the vectors $\vec{q}$, $\vec{k}$, $\vec{q}'$, $\vec{k}'$ and scanned over the polar angle θ. As a results of this the differences $\vec{q}-\vec{k}$ and $\vec{q}'-\vec{k}'$ change with changing

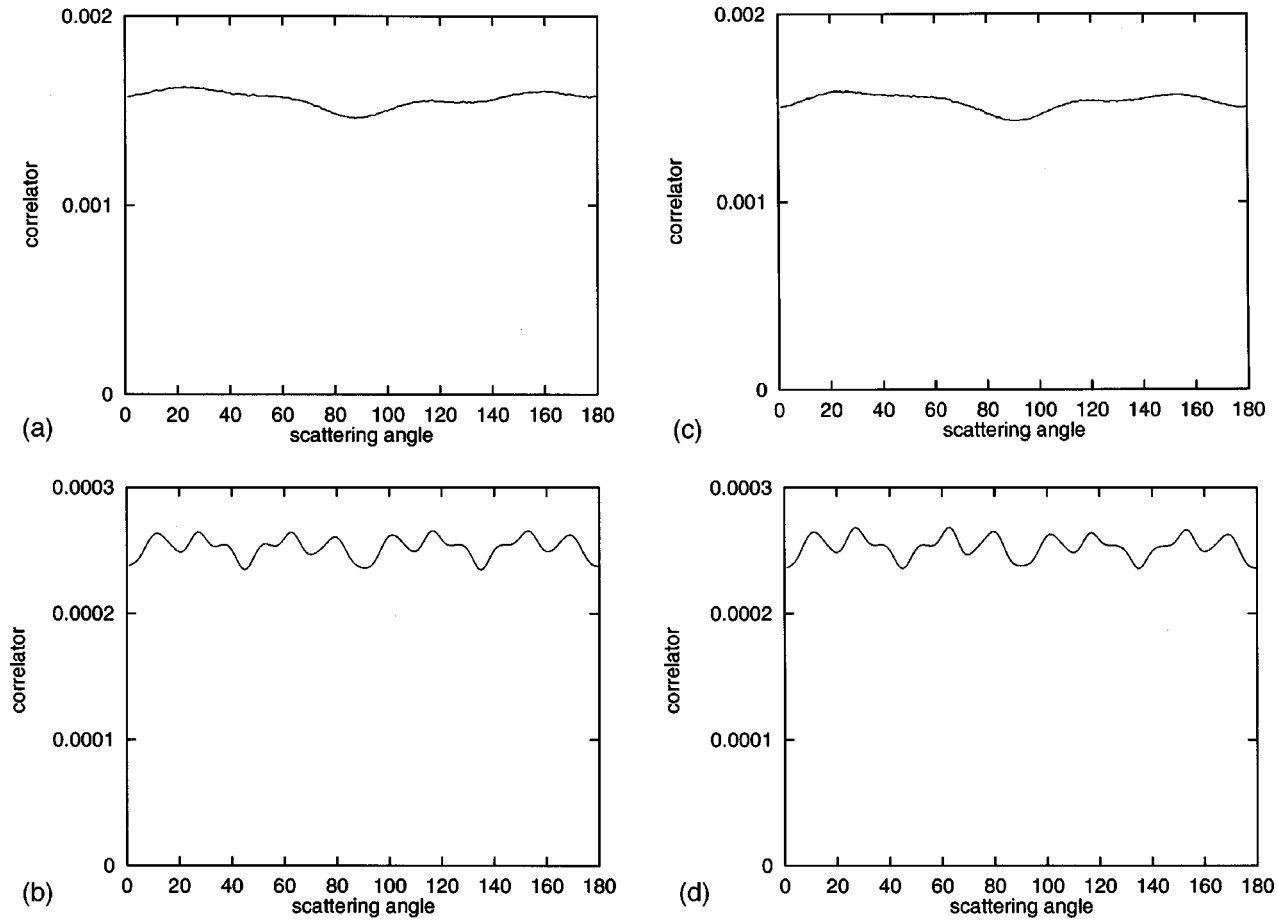

FIG. 9. Plot of the angular speckle correlator for the system of Fig. 4(a). Plots as functions of ϕ are presented for the wave vectors defined in Eqs. (32)–(35) for: (a) $\theta=10°$; (b) $\theta=45°$; and for the wave vectors defined in Eqs. (36)–(39) for: (c) $\theta=10°$; (d) $\theta=45°$.

polar angle. While this samples a wide range of behaviors of the correlators in $\vec{q}$, $\vec{k}$, $\vec{q}'$, $\vec{k}'$, it is interesting to study the correlators for scans in wave-vector space which involve fixed differences of $\vec{q}-\vec{k}$ and $\vec{q}'-\vec{k}'$. If θ is fixed and ϕ is allowed to vary in the vectors defined by

$$\vec{q}=k_0(\sin\theta,0,\cos\theta), \tag{32}$$

$$\vec{k}=k_0(\sin\theta,0,-\cos\theta), \tag{33}$$

$$\vec{q}'=k_0(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta), \tag{34}$$

$$\vec{k}'=k_0(\sin\theta\cos\phi,\sin\theta\sin\phi,-\cos\theta), \tag{35}$$

a plot of the envelope of $C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ for $\vec{q}-\vec{k}=\vec{q}'-\vec{k}'=k_0(0,0,2\cos\theta)$ is obtained. In Figs. 9(a) and 9(b) we present results for $C^{(1)}$ versus ϕ from the periodic on average system with $\epsilon=-9$ for scattering processes in which $\theta=10$ and 45°. In general, the correlator is seen to be a smooth slowly varying function of ϕ for both polar angles. If we fix θ and vary ϕ for vectors defined by

$$\vec{q}=k_0(\sin\theta,0,\cos\theta), \tag{36}$$

$$\vec{k}=k_0(\sin\theta,0,-\cos\theta), \tag{37}$$

$$\vec{q}'=k_0(\sin\theta\cos\phi,\sin\theta\sin\phi,-\cos\theta), \tag{38}$$

$$\vec{k}'=k_0(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta), \tag{39}$$

a plot of the envelope of $C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ for $\vec{q}-\vec{k}=-(\vec{q}'-\vec{k}')=k_0(0,0,2\cos\theta)$ is obtained. In Figs. 9(c) and 9(d) we present results for $C^{(10)}$ versus ϕ from the periodic on average system with $\epsilon=-9$ for scattering processes in which $\theta=10$ and 45°. Again, the correlator is found to be relatively independent of ϕ. Similar behaviors are observed in the random system that is not periodic on average.

C. Discussion of additional peaks arising from an average periodicity

An explanation based on analytic techniques can be given for the new peaks observed in the results for the periodic on average system shown in Fig. 4. (Specifically, these are the peaks occurring at $\phi=0$ and 180°.) This is done by applying diagrammatic Green's-function methods to a model system[10] which, though different from the one described in Eqs. (1)–(3), exhibits a periodic on average randomness that is easier to handle analytically and which is qualitatively similar to that of the periodic on average system discussed in Sec. II.

We consider a form for $\epsilon(\vec{r})$ given by

$$\epsilon(\vec{r}) = u(\vec{r}) + \epsilon_0(\vec{r}), \tag{40}$$

where $u(\vec{r})$ is a periodic function and $\epsilon_0(\vec{r})$ is drawn from a set of Gaussian random functions,[10,14–16] $\{\epsilon_0(\vec{r})\}$. The scattering properties of the system are expressed in terms of $\epsilon(\vec{r})$, and the average scattering properties are then determined by averaging over the set of Gaussian random functions, $\{\epsilon_0(\vec{r})\}$. In Eq. (40), the periodic function $u(\vec{r})$ is represented by

$$u(\vec{r}) = \sum_{\vec{G}} u_{\vec{G}} e^{i\vec{G}\cdot\vec{r}}, \tag{41}$$

where the sum is over the set of reciprocal-lattice vectors $\{\vec{G}\}$ of the average periodic lattice, and the Gaussian random function $\epsilon_0(\vec{r})$ satisfies

$$\langle \epsilon_0(\vec{r}) \rangle = 0, \tag{42}$$

$$\langle \epsilon_0(\vec{r}) \epsilon_0(\vec{r'}) \rangle = \sigma^2 g(|\vec{r} - \vec{r'}|). \tag{43}$$

In Eqs. (42) and (43) $\langle\ \rangle$ indicates an average over the set of Gaussian random functions, and the function $g(|\vec{r}-\vec{r'}|)$ is such that $g(0)=1$ and $\lim_{r\to\infty} g(r) \to 0$.

Following a well-known treatment,[10–12,14–16] it can be shown that the electromagnetic scattering cross section of the system in Eqs. (40)–(43) is proportional to a two-particle Green's function, i.e.,

$$\frac{\partial\sigma(\vec{q},\vec{k})}{\partial\Omega} \propto \langle |G_R(\vec{q},\vec{k})|^2 \rangle. \tag{44}$$

Here the Green's function of the random system, $G_R(\vec{q},\vec{k})$, satisfies the Dyson equation

$$G_R(\vec{q},\vec{k}) = (2\pi)^3 \delta(\vec{q}-\vec{k}) G(\vec{k}) + G(\vec{q}) \int \frac{d^3p}{(2\pi)^3} V(\vec{q}|\vec{p}) G_R(\vec{p},\vec{k}), \tag{45}$$

where $G(\vec{k})$ is the Fourier transform of Eq. (8), and the scattering potential for the dielectric system described in Eqs. (40)–(43) is given by

$$V(\vec{p}|\vec{k}) = v(\vec{p}|\vec{k}) \hat{\epsilon}(\vec{p}-\vec{k}). \tag{46}$$

In Eq. (46) $\hat{\epsilon}(\vec{q}) = \int d^3r e^{-i\vec{q}\cdot\vec{r}} \epsilon(\vec{r})$, so that from Eqs. (40) and (41)

$$\hat{\epsilon}(\vec{q}) = \sum_{\vec{G}} u_{\vec{G}} (2\pi)^3 \delta(\vec{q}-\vec{G}) + \hat{\epsilon}_0(\vec{q}), \tag{47}$$

where $\hat{\epsilon}_0(\vec{q}) = \int d^3r e^{-i\vec{q}\cdot\vec{r}} \epsilon_0(\vec{r})$, $u_{\vec{G}} = \int_{plc} d^3r e^{-i\vec{G}\cdot\vec{r}}$ and plc indicates that the integral is over a primitive lattice cell. The coefficient $v(\vec{p}|\vec{k})$ in Eq. (46), whose general form is not important to our discussions, contains details of the interaction of the electromagnetic fields with the system which do not affect the gross geometric features of the speckle correlator.

In terms of the Green's function $G_R(\vec{q}|\vec{k})$, it has been shown[1,10–12] that the angular speckle correlator $C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ satisfies

$$C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}') \propto \langle G_R^*(\vec{q},\vec{k}) G_R(\vec{q}',\vec{k}') \rangle \times \langle G_R^*(\vec{q}',\vec{k}') G_R(\vec{q},\vec{k}) \rangle - |\langle G_R(\vec{q},\vec{k}) \rangle|^2 |\langle G_R(\vec{q}',\vec{k}') \rangle|^2, \tag{48}$$

and that the angular speckle correlator $C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ satisfies[10–12]

$$C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}') \propto \langle G_R^*(\vec{q},\vec{k}) G_R^*(\vec{q}',\vec{k}') \rangle \times \langle G_R(\vec{q}',\vec{k}') G_R(\vec{q},\vec{k}) \rangle - |\langle G_R(\vec{q},\vec{k}) \rangle|^2 |\langle G_R(\vec{q}',\vec{k}') \rangle|^2. \tag{49}$$

An understanding of the gross geometric properties of $C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ and $C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ can be obtained by using Eqs. (45)–(47) to expand Eqs. (48) and (49) through terms of order $u_{\vec{G}}\sigma^2$. We now turn to a discussion of these terms.

The two-particle Green's functions in Eqs. (48) and (49) can be written as

$$\langle G_R(\vec{q},\vec{k}) G_R(\vec{q}',\vec{k}') \rangle = (2\pi)^3 \delta(\vec{q}-\vec{k}) G(\vec{k}) (2\pi)^3 \times \delta(\vec{q}'-\vec{k}') G(\vec{q}') + G(\vec{q}) G(\vec{q}') \times \Gamma^0(\vec{q},\vec{k}|\vec{q}',\vec{k}') G(\vec{k}) G(\vec{k}'), \tag{50}$$

and

$$\langle G_R^*(\vec{q},\vec{k}) G_R(\vec{q}',\vec{k}') \rangle = (2\pi)^3 \delta(\vec{q}-\vec{k}) G^*(\vec{k}) (2\pi)^3 \times \delta(\vec{q}'-\vec{k}') G(\vec{q}') + G^*(\vec{q}) G(\vec{q}') \times \Gamma^1(\vec{q},\vec{k}|\vec{q}',\vec{k}') G^*(\vec{k}) G(\vec{k}'), \tag{51}$$

where $\Gamma^0(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ and $\Gamma^1(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ are the respective reducible vertex functions arising from the scattering interaction described by $\langle V(\vec{q}|\vec{k}) V(\vec{q}'|\vec{k}') \rangle$ and $\langle V^*(\vec{q}|\vec{k}) V(\vec{q}'|\vec{k}') \rangle$. Computing the reducible vertex function in Eq. (50) to terms of order $u_{\vec{G}}\sigma^2$, we find the general form

$$\Gamma_0^0(\vec{q},\vec{k}|\vec{q}',\vec{k}') = A(\vec{q},\vec{k}|\vec{q}',\vec{k}') \delta(\vec{q}-\vec{k}+\vec{q}'-\vec{k}') + B(\vec{q},\vec{k}|\vec{q}',\vec{k}') \sum_{\vec{G}} \delta(\vec{q}-\vec{k}+\vec{q}'-\vec{k}'+\vec{G}), \tag{52}$$

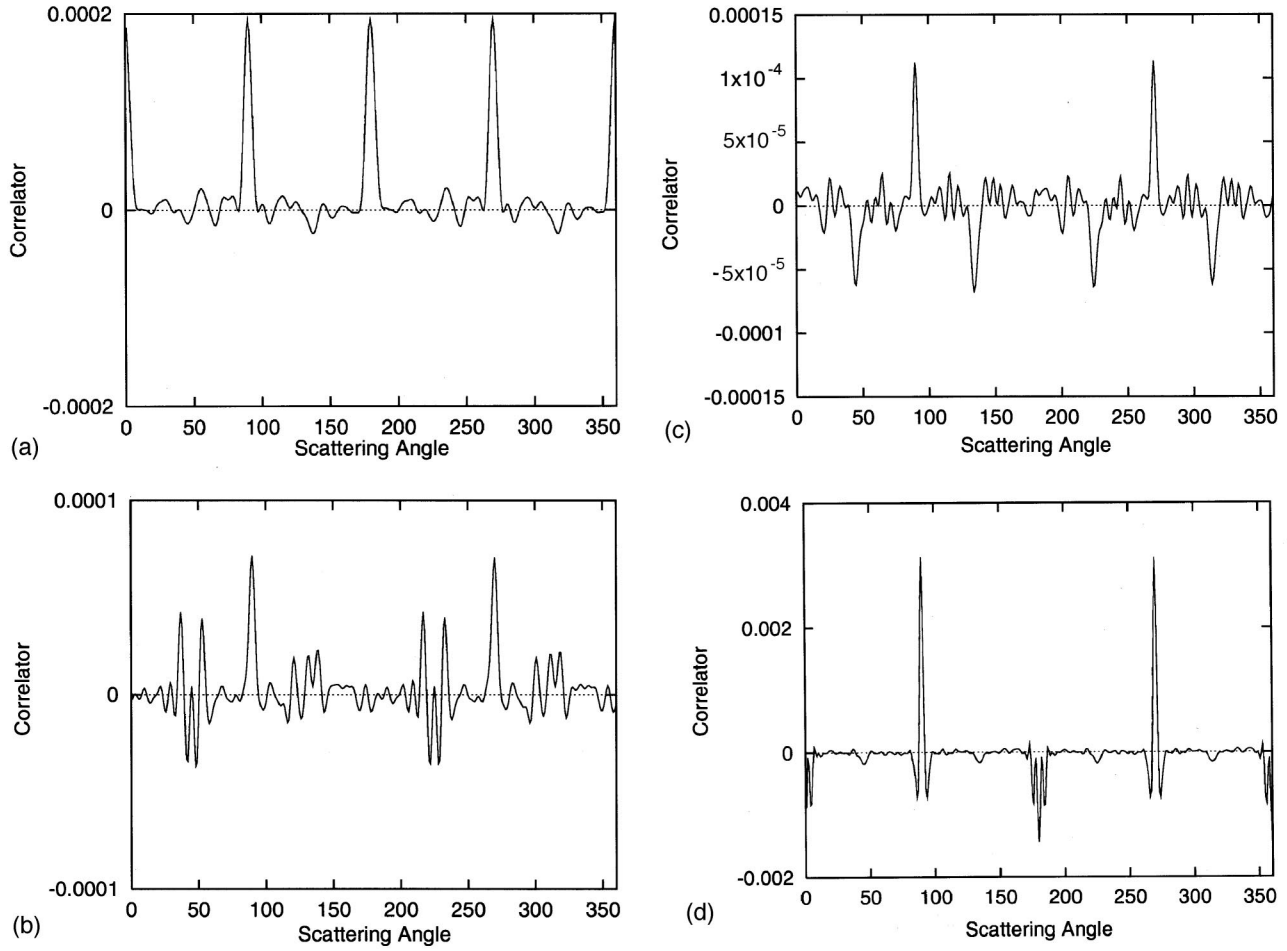

FIG. 10. Results for the angular speckle correlator for the system of Fig. 4(a). Plots are presented for the wave vectors defined in Eqs. (56)–(59) showing the total correlator function as a function of the scattering angle ϕ. Results are shown for: (a) $\theta=30°$; (b) $\theta=45°$; (c) $\theta=50°$; and (d) $\theta=70°$.

where $A(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ and $B(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ are smoothly varying functions of $\vec{q}$, $\vec{k}$, $\vec{q}'$, and $\vec{k}'$. Computing the reducible vertex function in Eq. (51) to terms of order $u_{\vec{G}}\sigma^2$, we find the general form

$$\Gamma_0^1(\vec{q},\vec{k}|\vec{q}',\vec{k}')=D(\vec{q},\vec{k}|\vec{q}',\vec{k}')\,\delta(\vec{q}-\vec{k}-\vec{q}'+\vec{k}') + E(\vec{q},\vec{k}|\vec{q}',\vec{k}')\sum_{\vec{G}}\ \delta(\vec{q}-\vec{k}-\vec{q}'+\vec{k}'+\vec{G}), \tag{53}$$

where $D(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ and $E(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ are smoothly varying functions of $\vec{q}$, $\vec{k}$, $\vec{q}'$, and $\vec{k}'$.

From Eqs. (48)–(53) we see that

$$C^{(1)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')\propto\sum_{\vec{G}}\ \delta(\vec{q}-\vec{k}-\vec{q}'+\vec{k}'+\vec{G}) \tag{54}$$

and

$$C^{(10)}(\vec{q},\vec{k}|\vec{q}',\vec{k}')\propto\sum_{\vec{G}}\ \delta(\vec{q}-\vec{k}+\vec{q}'-\vec{k}'+\vec{G}). \tag{55}$$

The effect on the speckle correlator of the average periodicity in the random system is found in the additional delta-function terms in Eqs. (54) and (55) over the single ($\vec{G}=0$) delta-function terms found in the system that is not periodic on average. In the absence of average periodicity (the $|\vec{G}|\to\infty$ limit), only the $\vec{G}=0$ terms contribute in Eqs. (54) and (55).

To investigate, in the context of our computer simulations discussed in Sec. II, the $\vec{G}\neq 0$ terms in Eqs. (54) and (55), it is useful to take the forms

$$\vec{k}=k_0(\sin\theta,0,\cos\theta), \tag{56}$$

$$\vec{k}'=k_0(\cos\phi\sin\theta,\sin\phi\sin\theta,\cos\theta), \tag{57}$$

$$\vec{q}=k_0(0,\sin\theta,\cos\theta), \tag{58}$$

$$\vec{q}'=k_0(\sin\phi\sin\theta,-\cos\phi\sin\theta,\cos\theta). \tag{59}$$

In the limit that $\theta=90°$ the wave vectors $\vec{k}$, $\vec{k}'$, $\vec{q}$, $\vec{q}'$ in Eqs. (56)–(59) reduce to those given in Eqs. (20)–(23) and reproduce the results in Fig. 4 for the periodic on average system.

For $\theta \neq 90°$ all four vectors have the same z component, and for $\theta = 0°$ all four vectors are parallel to the z axis. Using the wave vectors in Eqs. (56)–(59) we find for general θ and ϕ

$$\vec{q}-\vec{k}-\vec{q}'+\vec{k}' = k_0 \sin\theta(-1-\sin\phi+\cos\phi, 1+\cos\phi + \sin\phi, 0), \tag{60}$$

and

$$\vec{q}-\vec{k}+\vec{q}'-\vec{k}' = k_0 \sin\theta(-1+\sin\phi-\cos\phi, 1-\cos\phi - \sin\phi, 0), \tag{61}$$

which give simple expressions in ϕ and θ for these two important linear combinations of the incident and scattering wave vectors of the angular speckle correlator.

From Eqs. (54) and (55), it is seen that when the right-hand side of either Eq. (60) or (61) becomes equal to a reciprocal-lattice vector, then either $C^{(1)}$ or $C^{(10)}$, respectively, are nonzero. Otherwise, these contributions to the general angular speckle correlator are zero. We shall now use these results to explain the peaks observed in Fig. 4 for the periodic on average system. Following this, we shall present additional results for the periodic on average system considered in Fig. 4 that illustrate the general $\vec{G} \neq 0$ results in Eqs. (54) and (55) for these types of systems.

In Fig. 4 simulation results have been presented for the $\theta = 90°$ case of the wave-vector geometry described by Eqs. (56)–(61). Peaks are observed for $\phi = 0$, 90, 180, and 270°. Upon evaluating Eqs. (60) and (61) for the values of θ and ϕ at which peaks are found in the angular speckle correlator, we find that the four peaks correspond to cases in which the right-hand sides of Eqs. (60) and (61) become equal to reciprocal-lattice vectors. In particular, the peaks observed at $\phi = 0$ and 180° arise from cases in which the right-hand sides of Eqs. (60) and (61) are nonzero reciprocal-lattice vectors, and the peaks at $\phi = 90$ and 270° arise in part from cases in which the right-hand sides of Eqs. (60) and (61) are the zero reciprocal-lattice vector.

In Fig. 10 we extend this analysis to cases in which $\theta \neq 90°$. Scans are presented in ϕ for the angular speckle correlator of the $\epsilon = -9$ periodic on average system treated in Fig. 4(a). Plots are presented of the total correlator $C(\vec{q},\vec{k}|\vec{q}',\vec{k}')$ versus ϕ for $\theta = 30$, 45, 50, 70°, and a variety of peaks and dips are observed as a function of ϕ in these plots. Evaluating Eqs. (60) and (61) we find that the four peaks in the $\theta = 30°$ plot occur when the right-hand sides of Eqs. (60) or (61) equal vectors of the reciprocal lattice. The peaks at $\phi = 0°$ and 180° arise from nonzero reciprocal-lattice vectors, and are consequently absent in the system that is not periodic on average. The peaks at $\phi = 90$ and 270° arise in part from $\vec{G} = 0$ and will be found in the results for the homogeneous random system. From Eqs. (60) and (61) we find that the peaks in the $\theta = 45°$ plot located about $\phi = 0$, 90, 180, 270° occur for cases where Eqs. (60) or (61) become equal to reciprocal-lattice vectors, and the peaks located about $\phi = 45$, 135, 225°, occur for cases in which the right-hand sides of Eqs. (60) or (61) are close to, but not quite equal to, reciprocal-lattice vectors. From Eqs. (60) and (61) we find that the peaks or dips in the $\theta = 50$ and 70° plots for the peaks or dips at $\phi = 45$, 90, 135, 225, 270, 315° occur for cases in which the right-hand side of Eqs. (60) or (61) are close to reciprocal-lattice vectors.

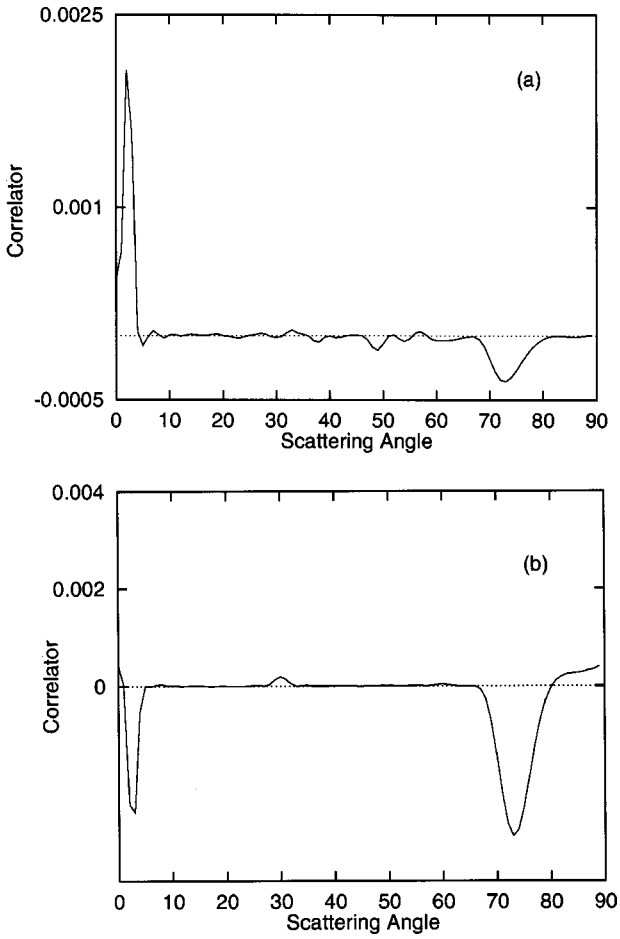

FIG. 11. Results for the angular speckle correlator for the same system as considered in Fig. 10, now plotted as a function of the scattering angle θ for fixed values of: (a) $\phi = 45°$; and (b) $\phi = 180°$.

The delta functions in Eqs. (54) and (55) represent results for an ideal random system that is of infinite extent. For random systems of finite extent we expect that the delta function in Eqs. (54) and (55), $\delta(\vec{q}-\vec{k}-\vec{q}'+\vec{k}'+\vec{G})$ and $\delta(\vec{q}-\vec{k}+\vec{q}'-\vec{k}'+\vec{G})$, are replaced by peaked functions $f(\vec{q}-\vec{k}-\vec{q}'+\vec{k}'+\vec{G})$ and $f(\vec{q}-\vec{k}+\vec{q}'-\vec{k}'+\vec{G})$ with maxima at the zero of their arguments and widths of order $1/l$ where l is a typical linear size dimension (i.e., the radius of the smallest covering sphere) of the finite system. Consequently, the peaks observed near reciprocal-lattice vectors in the plots for $\theta = 45$, 50, 70° should in some cases become less pronounced as the size of the system increases.

To complement the runs made in Figs. 4 and 10, where plots of ϕ for fixed θ are given, it is interesting to fix ϕ at 45 or 180° and plot the angular speckle correlator for θ running from 0 to 90°. These plots are then analogous to the plots made in Figs. 5–8 for the $\vec{G} = 0$ envelopes. In Fig. 11 results

are presented for the $\epsilon=-9$ periodic on average system studied in Fig. 4(a) for $\phi=45$ and 180° for the speckle correlator as a function of θ. In these plots, anomalous behavior is found near sets of angles for which the right-hand sides of Eqs. (60) and (61) are approximately equal to vectors of the reciprocal lattice. It is interesting to note that in both plots negative dips in the correlator are found near $\theta=70°$.

IV. CONCLUSIONS

We have computed the scattering cross section and the speckle correlator for the far-field scattering of light by a finite, random, array of dielectric spheres. Two different types of randomness are considered. In one, the spheres randomly occupy space uniformly subject to the condition that different spheres do not overlap. In a second, the spheres randomly occupy space with a distribution that is periodic on average. The results presented confirm the existence of both the $C^{(1)}$ contribution to the speckle correlator (which has been known to exist from a number of publications[1–12]) and the $C^{(10)}$ contribution to the speckle correlator (which has recently been predicted in three-dimensional scattering[10] but only as yet measured in the light scattered from rough metal surfaces[19]). The simulation has been used to determine the envelope function of these correlators. The envelope functions of $C^{(1)}$ and $C^{(10)}$ are seen to be similar to one another when plotted in the angular variables we have used. Interesting additional features are observed both in $C^{(1)}$ and $C^{(10)}$ when an average periodicity is introduced into the randomness of the system. For the periodic on average systems additional regions of nonzero $C^{(1)}$ and $C^{(10)}$, other that those defined by the conditions $\vec{q}-\vec{k}-\vec{q}'+\vec{k}'=0$ and $\vec{q}-\vec{k}+\vec{q}'-\vec{k}'=0$, are observed, that are caused by scatterings involving one or more Bragg reflections.

In conclusion, we note that recently a contribution to the speckle correlator from volume scattering has been introduced, called the $C^{(0)}$ correlation function.[17,18] The systems studied in Refs. 17 and 18 differ from the systems studied in the present paper in that the speckle in Refs. 17 and 18 arises from a point or extended source of electromagnetic radiation in the medium, respectively, whereas in this paper the scattered electromagnetic waves arise from incident plane waves. It appears, however, that the diagrams contributing to $C^{(0)}$ are part of a set of diagrams entering into the definition of the $C^{(1.5)}$ contribution to the speckle correlator defined for both surface and volume scattering.[10–12] This is clear from the characterization of the $C^{(1.5)}$ diagrams for volume scattering given in the second paragraph of Sec. III B of Ref. 10. This is to say that $C^{(0)}$ is composed from two two-particle Green's functions that are connected by a line of scattering interaction, and this is part of the class of contributions to $C^{(1.5)}$. The reader is also referred to our earlier paper on surface scattering[12] where the $C^{(1.5)}$ contributions are discussed in Sec. 2.2. We note in this regard that Fig. 2 of Ref. 12 does not explicitly list the $C^{(0)}$ diagrams because Fig. 2 of Ref. 12 was a list, taken from the many diagrams contributing to $C^{(1.5)}$, of contributions to $C^{(1.5)}$ that have peaks. These peaks were the focus of the work in Ref. 12 and the $C^{(0)}$ diagrams, though part of of the $C^{(1.5)}$ contributions, do not display peaks. The $C^{(1.5)}$ correlator is not a topic of the results presented in this paper, and we will not pursue this topic further here.

ACKNOWLEDGMENTS

The research of A.A.M. was supported in part by Army Research Office Grant DAAD 19-99-1-0321.

[1] B. Shapiro, Phys. Rev. Lett. **57**, 2168 (1986).
[2] S. Feng, C. Kane, P.A. Lee, and A.D. Stone, Phys. Rev. Lett. **61**, 834 (1988).
[3] I. Freund, M. Rosenbluh, and S. Feng, Phys. Rev. Lett. **61**, 2328 (1988).
[4] R. Berkovits, M. Kaveh, and S. Feng, Phys. Rev. B **40**, 737 (1993).
[5] R. Berkovits and M. Kaveh, Phys. Rev. B **41**, 2635 (1990).
[6] L. Wang and S. Feng, Phys. Rev. B **40**, 8284 (1989).
[7] R. Berkovits, Phys. Rev. B **42**, 10 750 (1990).
[8] N. Garcia and A.Z. Genack, Phys. Rev. Lett. **63**, 1678 (1989).
[9] M.P. van Albada, J.F. De Boer, and A. Lagendijk, Phys. Rev. Lett. **64**, 2787 (1990).
[10] V. Malyshkin, A.R. McGurn, and A.A. Maradudin, Phys. Rev. B **59**, 6167 (1999).
[11] V. Malyshkin, A.R. McGurn, T.A. Leskova, A.A. Maradudin, and M. Nieto-Vesperinas, Opt. Lett. **22**, 946 (1997).
[12] V. Malyshkin, A.R. McGurn, T.A. Leskova, A.A. Maradudin, and M. Nieto-Vesperinas, Waves Random Media **7**, 479 (1997).
[13] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (John Wiley and Sons, Inc., New York, 1991).
[14] G. Brown, V. Celli, M. Haller, and A. Marvin, Surf. Sci. **136**, 136 (1984).
[15] A.R. McGurn, A.A. Maradudin, and V. Celli, Phys. Rev. B **31**, 4866 (1985).
[16] A.R. McGurn and A.A. Maradudin, J. Opt. Soc. Am. B **4**, 910 (1987).
[17] B. Shapiro, Phys. Rev. Lett. **83**, 4733 (1999).
[18] S.E. Skipetrov and R. Maynard, Phys. Rev. B **62**, 886 (2000).
[19] C.S. West and K.A. O'Donnell, Phys. Rev. B **59**, 2393 (1999).

INSTITUTE OF PHYSICS PUBLISHING WAVES IN RANDOM MEDIA

Waves Random Media **12** (2002) 307–319 PII: S0959-7174(02)31152-2

The angular intensity correlation functions $C^{(1)}$ and $C^{(10)}$ for the scattering of light from randomly rough dielectric and metal surfaces

Tamara A Leskova[1], **Ingve Simonsen** [2,3,4] **and Alexei A Maradudin**[4]

[1] Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Russia
[2] NORDITA, Blegdamsvej 17, DK-2100 Copenhagen, Ø, Denmark
[3] Department of Physics, The Norwegian University of Science and Technology, N-7491 Trondheim, Norway
[4] Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine 92697, CA, USA

Received 22 November 2001, in final form 14 February 2002
Published 4 April 2002
Online at stacks.iop.org/WRM/12/307

Abstract

We study the statistical properties of the scattering matrix $S(q|k)$ for the problem of the scattering of light of frequency ω from a randomly rough one-dimensional surface, defined by the equation $x_3 = \zeta(x_1)$, where the surface profile function $\zeta(x_1)$ constitutes a zero-mean, stationary, Gaussian random process. This is done by studying the effects of $S(q|k)$ on the angular intensity correlation function $C(q, k|q', k') = \langle I(q|k)I(q'|k')\rangle - \langle I(q|k)\rangle\langle I(q'|k')\rangle$, where the intensity $I(q|k)$ is defined in terms of $S(q|k)$ by $I(q|k) = L_1^{-1}(\omega/c)|S(q|k)|^2$, with L_1 the length of the x_1 axis covered by the random surface. We focus our attention on the $C^{(1)}$ and $C^{(10)}$ correlation functions, which are the contributions to $C(q, k|q', k')$ proportional to $\delta(q - k - q' + k')$ and $\delta(q - k + q' - k')$, respectively. The existence of both of these correlation functions is consistent with the amplitude of the scattered field obeying complex Gaussian statistics in the limit of a long surface and in the presence of weak surface roughness. We show that the deviation of the statistics of the scattering matrix from complex circular Gaussian statistics and the $C^{(10)}$ correlation function are determined by exactly the same statistical moment of $S(q|k)$. As the random surface becomes rougher, the amplitude of the scattered field no longer obeys complex Gaussian statistics but obeys complex circular Gaussian statistics instead. In this case the $C^{(10)}$ correlation function should therefore vanish. This result is confirmed by numerical simulation calculations.

1. Introduction

The scattering of light from randomly rough surfaces has attracted attention over many years. The majority of the theoretical and experimental studies of such scattering has been devoted to

 © 2002 IOP Publishing Ltd Printed in the UK

coherent interference effects occurring in the multiple scattering of electromagnetic waves from randomly rough surfaces and the related backscattering enhancement phenomenon. These effects are contained in the angular distribution of the intensity of the light scattered incoherently, i.e. in the second moment of the scattered field.

Recently, attention has begun to be directed toward theoretical [1–12] and experimental [2,7,8,12,13] studies of multiple-scattering effects on higher moments of the scattered field, in particular on angular intensity correlation functions. These correlation functions describe how the speckle pattern, formed through the interference of randomly scattered waves, changes when one or more parameters of the scattering system are varied.

The interest in these correlations has been stimulated by the expectation that, just as the inclusion of multiple-scattering processes in the calculation of the angular dependence of the intensity of the light that has been scattered incoherently from, or incoherently through, a randomly rough surface, led to the prediction of enhanced backscattering [14] and enhanced transmission [15], their inclusion in the calculation of higher-order moments of the scattered or transmitted field would also lead to the prediction of new physical effects. This expectation was prompted by the results of earlier theoretical [16,17] and experimental [18–20] investigations of angular intensity correlation functions in the scattering of classical waves from volume disordered media. In a theoretical investigation [9] it was predicted that three types of correlations occur in such scattering, namely short-range correlations, long-range correlations and infinite-range correlations. These were termed the $C^{(1)}$, $C^{(2)}$ and $C^{(3)}$ correlations, respectively. The $C^{(1)}$ correlation function includes both the 'memory effect' and the 'reciprocal memory effect' [9,10], so named because of the wavevector conservation conditions they satisfy. Both of these effects have now been observed in volume scattering experiments [16,17]. The $C^{(2)}$ correlation function has also been observed in volume scattering experiments [18,19], as has the $C^{(3)}$ correlation function [20].

Until recently, only the $C^{(1)}$ correlation function arising in the scattering of light from a randomly rough surface had been studied theoretically and experimentally [1–8]. In a recent series of papers devoted to theoretical studies of angular correlation functions of the intensity of light scattered from one-dimensional [9,10] and two-dimensional [10] randomly rough metal surfaces the long-range $C^{(2)}$ and infinite-range $C^{(3)}$ correlation functions were calculated, and two additional types of correlation functions, a short-range correlation function, named $C^{(10)}$, and a long-range correlation function, named $C^{(1.5)}$, were predicted. In very recent experimental work [12] the envelopes of the $C^{(1)}$ and $C^{(10)}$ correlation functions were measured experimentally for the scattering of p-polarized light from weakly rough, one-dimensional gold surfaces. The $C^{(1.5)}$, $C^{(2)}$ and $C^{(3)}$ correlation functions have yet to be observed experimentally.

The question arises as to whether it possible to determine the relative magnitudes of the different correlation functions from a knowledge of the experimental parameters of the surface roughness and its statistical properties. This question had been raised earlier in [12,16], but not answered definitively. We therefore address it here for the case of a one-dimensional random surface defined by the equation $x_3 = \zeta(x_1)$, on the basis of the single assumption that the surface profile function $\zeta(x_1)$ is a single-valued function of x_1 that constitutes a zero-mean, stationary, Gaussian random process. At the same time we address the question of how the statistical properties of the amplitude of the scattered field are reflected in the symmetry properties of the speckle pattern to which it gives rise.

The outline of this paper is as follows. In section 2 we introduce the angular intensity correlation function and analyse it in terms of the possible statistics of the scattering matrix. In section 3 we illustrate the conclusions of section 2 for the simple example of the scattering of light from the randomly rough surface of a perfect conductor. Finally, in section 4 we present the conclusions drawn from the results obtained in this paper.

2. The angular intensity correlation function

The general angular intensity correlation function $C(q, k|q', k')$ we study in this work is defined by

$$C(q, k|q', k') = \langle I(q|k)I(q'|k')\rangle - \langle I(q|k)\rangle\langle I(q'|k')\rangle, \tag{2.1}$$

where the angle brackets denote an average over the ensemble of realizations of the surface profile function. The intensity $I(q|k)$ entering this expression is defined in terms of the scattering matrix $S(q|k)$ for the scattering of light of frequency ω from a one-dimensional random surface by

$$I(q|k) = \frac{1}{L_1}\left(\frac{\omega}{c}\right)|S(q|k)|^2, \tag{2.2}$$

where L_1 is the length of the x_1 axis covered by the random surface, and the wavenumbers k and q are related to the angles of incidence and scattering, θ_0 and θ_s, measured counterclockwise and clockwise from the normal to the mean scattering surface, respectively, by $k = (\omega/c)\sin\theta_0$ and $q = (\omega/c)\sin\theta_s$.

From equations (2.1) and (2.2) we see that, because the correlation of $I(q|k)$ with itself should generally be stronger than the correlation of $I(q|k)$ with $I(q'|k')$ when $q' \neq q$ and $k' \neq k$, a peak in $C(q, k|q', k')$ is expected when $q' = q$ and $k' = k$. This peak is called the *memory effect peak* for the reason that is explained below. In addition, because $S(q|k)$ is reciprocal, $S(q|k) = S(-k|-q)$, a peak in $C(q, k|q', k')$ is also expected when $q' = -k$ and $k' = -q$. This peak is called the *reciprocal memory effect peak.*

In terms of the scattering matrix $S(q|k)$ the correlation function $C(q, k|q', k')$ becomes

$$\begin{aligned} C(q, k|q', k') = \frac{1}{L_1^2}\frac{\omega^2}{c^2}[&\langle S(q|k)S^*(q|k)S(q'|k')S^*(q'|k')\rangle \\ &- \langle S(q|k)S^*(q|k)\rangle\langle S(q'|k')S^*(q'|k')\rangle]. \end{aligned} \tag{2.3}$$

Since, due to the stationarity of the surface profile function, $\langle S(q|k)\rangle$ is diagonal in q and k, $\langle S(q|k)\rangle = 2\pi\delta(q-k)S(k)$, we introduce the incoherent part of the scattering matrix $\delta S(q|k) = S(q|k) - \langle S(q|k)\rangle$. Then, from the relations between averages of the products of random functions and the corresponding cumulant averages [21, 22] and omitting all terms proportional to $2\pi\delta(q-k)$ and/or $2\pi\delta(q'-k')$ as uninteresting specular effects, equation (2.3) can be rewritten in the form

$$\begin{aligned} C(q, k|q', k') = \frac{1}{L_1^2}\frac{\omega^2}{c^2}[&|\langle\delta S(q|k)\delta S^*(q'|k'))|^2 + |\langle\delta S(q|k)\delta S(q'|k'\rangle|^2 \\ &+ \langle\delta S(q|k)\delta S^*(q|k)\delta S(q'|k')\delta S^*(q'|k')\rangle_c], \end{aligned} \tag{2.4}$$

where $\langle\cdots\rangle_c$ denotes the cumulant average.

Due to the stationarity of the surface profile function $\zeta(x_1)$, $\langle\delta S(q|k)\delta S^*(q'|k')\rangle$ is proportional to $2\pi\delta(q-k-q'+k')$. It gives rise to the contribution to $C(q, k|q', k')$ called $C^{(1)}(q, k|q', k')$ [9, 10] and describes the memory effect and the reciprocal memory effect. The property of a speckle pattern that is characterized by the presence of the factor $2\pi\delta(q-k-q'+k')$ in $C^{(1)}(q, k|q', k')$ is that, if we change the angle of incidence in such a way that k goes into $k' = k + \Delta k$, the entire speckle pattern shifts in such a way that any feature initially at q moves to $q' = q + \Delta k$. This is the reason why the $C^{(1)}$ correlation function was originally named the memory effect. In terms of the angles of incidence and scattering, we have that, if θ_0 is changed into $\theta_0' = \theta_0 + \Delta\theta_0$, any feature in the speckle pattern originally at θ_s is shifted to $\theta_s' = \theta_s + \Delta\theta_s$, where $\Delta\theta_s = \Delta\theta_0\,(\cos\theta_0/\cos\theta_s)$ to first order in $\Delta\theta_0$.

Similarly $\langle \delta S(q|k)\delta S(q'|k')\rangle$ is proportional to $2\pi\delta(q-k+q'-k')$ and contributes to the correlation function $C^{(10)}(q,k|q',k')$ to $C(q,k|q',k')$ [9,10]. The property of a speckle pattern that is characterized by the presence of the factor $2\pi\delta(q-k+q'-k')$ in $C^{(10)}(q,k|q',k')$ is that, if we change the angle of incidence in such a way that k goes into $k'=k+\Delta k$, a feature at $q=k-\Delta q$ will be shifted to $q'=k'+\Delta q$, i.e. to a point as much to one side of the new specular direction as the original point was on the other side of the original specular direction (in wavenumber space). For one and the same incident beam the $C^{(10)}$ correlation function therefore reflects the 'symmetry' of the speckle pattern with respect to the specular direction.

The third term on the right-hand side of equation (2.4), $\langle \delta S(q|k)\delta S^*(q|k)\delta S(q'|k')\delta S^*(q'|k')\rangle_c$, is proportional to $2\pi\delta(0)=L_1$, due to the stationarity of the surface profile function $\zeta(x_1)$ and gives rise to the long-range and infinite-range contributions to $C(q,k|q',k')$ given by the sum $C^{(1.5)}(q,k|q',k')+C^{(2)}(q,k|q',k')+C^{(3)}(q,k|q',k')$ [9,10]. Thus, we have separated out explicitly the contributions to $C(q,k|q',k')$ that have been called $C^{(1)}(q,k|q',k')$ and $C^{(10)}(q,k|q',k')$.

What is more, from equation (2.4) we can easily estimate the relative magnitudes of the different contributions to the general correlation function. Indeed, since $2\pi\delta(0)=L_1$, when the arguments of the δ functions vanish the $C^{(1)}(q,k|q',k')$ and $C^{(10)}(q,k|q',k')$ correlation functions are independent of the length of the surface L_1, because they contain $[2\pi\delta(0)]^2$. At the same time the remaining term in equation (2.4), which yields the sum $C^{(1.5)}(q,k|q',k')+C^{(2)}(q,k|q',k')+C^{(3)}(q,k|q'|k')$, is inversely proportional to the surface length, due to the lack of a second delta function. Therefore, in the limit of a long surface or a large illumination area the long-range and infinite-range correlations are small compared to short-range correlation functions and vanish in the limit of an infinitely long surface. Thus, the experimental observation of the $C^{(1.5)}$, $C^{(2)}$ and $C^{(3)}$ correlation functions requires the use of a short segment of random surface and/or the use of a beam of narrow width for the incident field. A detailed discussion of the conditions under which they may be observed will therefore be deferred to a separate paper.

The preceding results are consistent with the usual assumptions and conclusions encountered in conventional speckle theory [23–25]. Thus, when the surface profile function is assumed to be a stationary random process, and the random surface is assumed to be infinitely long, the scattering matrix $S(q|k)$ becomes the sum of a very large number of independent contributions from different points on the surface. On invoking the central limit theorem, it is found that $S(q|k)$ obeys complex Gaussian statistics. In this case equation (2.4) becomes rigorously [23]

$$C(q,k|q',k')=\frac{1}{L_1^2}\frac{\omega^2}{c^2}[|\langle \delta S(q|k)\delta S^*(q'|k')\rangle|^2+|\langle \delta S(q|k)\delta S(q'|k')\rangle|^2] \quad (2.5)$$

$$=C^{(1)}(q,k|q',k')+C^{(10)}(q,k|q',k'), \quad (2.6)$$

because all cumulant averages of products of more than two Gaussian random processes vanish. The last term on the right-hand side of equation (2.4) therefore gives the correction to the prediction of the central limit theorem due to the finite length of the random surface.

If it is further assumed, as is done in speckle theory, where the disorder is presumed to be strong, that $\delta S(q|k)$ obeys circular complex Gaussian statistics [24, 25], then $\langle \delta S(q|k)\delta S(q'|k')\rangle=0$ and the expression for $C(q,k|q',k')$ simplifies to

$$C(q,k|q',k')=\frac{1}{L_1^2}\frac{\omega^2}{c^2}|\langle \delta S(q|k)\delta S^*(q'|k')|^2 \quad (2.7)$$

$$=C^{(1)}(q,k|q',k'). \quad (2.8)$$

This approximation is often called the factorization approximation to $C(q,k|q'k')$ [17].

We recall that, if the complex random variables F_1 and F_2 are jointly circular complex Gaussian random variables, then the conditions

$$\langle \mathrm{Re}\, F_1 \mathrm{Re}\, F_2 \rangle = \langle \mathrm{Im}\, F_1 \mathrm{Im}\, F_2 \rangle, \tag{2.9}$$

$$\langle \mathrm{Re}\, F_1 \mathrm{Im}\, F_2 \rangle = -\langle \mathrm{Im}\, F_1 \mathrm{Re}\, F_2 \rangle, \tag{2.10}$$

have to be satisfied. To analyse how the scattering matrix transforms from a complex Gaussian random process into a circular complex Gaussian random process we represent the scattering matrix in the form $\delta S(q|k) = \delta S_1(q|k) + \mathrm{i}\delta S_2(q|k)$. The expressions for the averages of the products of the real and imaginary parts of $\delta S(q|k)$ can be written in terms of $\langle \delta S(q|k)\delta S^*(q'|k')\rangle$ and $\langle \delta S(q|k)\delta S(q'|k')\rangle$:

$$\langle \delta S_1(q|k)\delta S_1(q'|k')\rangle = \tfrac{1}{2}\mathrm{Re}\,[\langle \delta S(q|k)\delta S^*(q'|k')\rangle + \langle \delta S(q|k)\delta S(q'|k')\rangle] \tag{2.11}$$

$$\langle \delta S_2(q|k)\delta S_2(q'|k')\rangle = \tfrac{1}{2}\mathrm{Re}\,[\langle \delta S(q|k)\delta S^*(q'|k')\rangle - \langle \delta S(q|k)\delta S(q'|k')\rangle] \tag{2.12}$$

$$\langle \delta S_1(q|k)\delta S_2(q'|k')\rangle = -\tfrac{1}{2}\mathrm{Im}\,[\langle \delta S(q|k)\delta S^*(q'|k')\rangle - \langle \delta S(q|k)\delta S(q'|k')\rangle] \tag{2.13}$$

$$\langle \delta S_2(q|k)\delta S_1(q'|k')\rangle = \tfrac{1}{2}\mathrm{Im}\,[\langle \delta S(q|k)\delta S^*(q'|k')\rangle + \langle \delta S(q|k)\delta S(q'|k')\rangle]. \tag{2.14}$$

When $q = q'$ and $k = k'$, the average $\langle \delta S(q|k)\delta S(q|k)\rangle$, which is proportional to $2\pi\delta(2q-2k)$ due to the stationarity of the surface profile function, is nonzero only in the specular direction $q = k$. Therefore, if the surface is infinitely long, and if we omit the specular direction from equations (2.9), (2.10) and (2.11)–(2.14) we see that the scattering matrix is a circular complex Gaussian random process. Consequently, apart from the specular direction, the speckle contrast $\rho = \sqrt{[\langle(\delta S(q|k)\delta S^*(q|k))^2\rangle/\langle \delta S(q|k)\delta S^*(q|k)\rangle^2] - 1}$ is unity [24–26]. This result contradicts the well known result of [25] and [26] that the statistics of the diffuse component of the scattered field is highly non-circular when the surface is weakly rough, and only in the limit of very rough surfaces is the circularity of the statistics restored. The contradiction stems from the representation of the amplitude of the scattered field as the convolution of a real-valued amplitude weighting function and a random phase factor in [25, 26]. The assumption of a real-valued amplitude weighting function, which represents the finite width of the aperture, is identical to the assumption of a finite length of the randomly rough surface. As a result, the statistics of the scattering amplitude is nonstationary in [25, 26]. In the present work we are interested only in the case where the statistics of the surface profile function, as well as of the scattering matrix, is stationary.

The set of the scattering matrices $\delta S(q|k)$ is a set of jointly circular complex Gaussian random variables when $\langle \delta S(q|k)\delta S(q'|k')\rangle$ vanishes. But when $\langle \delta S(q|k)\delta S(q'|k')\rangle$ vanishes the correlation function $C^{(10)}$ vanishes since, within a coefficient, $C^{(10)}(q, k|q'k') \sim |\langle \delta S(q|k)\delta S(q'|k')\rangle|^2$.

Thus, calculations and measurements of the correlation function $C(q, k|q', k')$ yield important information about the statistical properties of the amplitude of the scattered field. If the random surface is such that only the $C^{(1)}$ and $C^{(10)}$ correlation functions are observed, then $S(q|k)$ obeys complex Gaussian statistics. If the random surface is such that only $C^{(1)}$ is observed, then $S(q|k)$ obeys circular complex Gaussian statistics. Finally, if the random surface is such that $C^{(1.5)}$, $C^{(2)}$ and $C^{(3)}$ are observed in addition to both $C^{(1)}$ and $C^{(10)}$, then $S(q|k)$ is not a Gaussian random process, but the statistics it obeys in this case are not known at the present time.

To conclude this section we introduce the normalized angular intensity correlation functions of interest to us, which in terms of $\delta S(q|k)$ are defined by

$$\Xi^{(1)}(q, k|q', k') = \frac{|\langle \delta S(q|k)\delta S^*(q'|k')\rangle|^2}{\langle \delta S(q|k)\delta S^*(q|k)\rangle\langle \delta S(q'|k')\delta S^*(q'|k')\rangle}, \tag{2.15}$$

and

$$\Xi^{(10)}(q,k|q',k') = \frac{|\langle \delta S(q|k)\delta S(q'|k')\rangle|^2}{\langle \delta S(q|k)\delta S^*(q|k)\rangle \langle \delta S(q'|k')\delta S^*(q'|k')\rangle}. \tag{2.16}$$

We also introduce the envelopes $C_0^{(1)}$ and $C_0^{(10)}$ of the correlation functions $C^{(1)}$ and $C^{(10)}$, which we define by

$$C^{(1)}(q,k|q',k') = 2\pi\delta(q-k-q'+k')C_0^{(1)}(q,k|q',q'-q+k) \tag{2.17}$$

and

$$C^{(10)}(q,k|q',k') = 2\pi\delta(q-k+q'-k')C_0^{(10)}(q,k|q'q'+q-k). \tag{2.18}$$

3. Light scattering from a perfectly conducting randomly rough surface in the framework of phase perturbation theory

In this section we study the statistical properties of the scattering matrix for the problem of the scattering of a scalar plane wave from a randomly rough infinitely long surface defined by the equation $x_3 = \zeta(x_1)$. The region $x_3 > \zeta(x_1)$ is vacuum, while the region $x_3 < \zeta(x_1)$ is a perfectly conducting medium. It is assumed that the Dirichlet boundary condition is satisfied on the surface $x_3 = \zeta(x_1)$.

The surface profile function $\zeta(x_1)$ is assumed to be a single-valued function of x_1 that is differentiable and constitutes a zero-mean, stationary, Gaussian random process defined by the properties

$$\langle \zeta(x_1)\rangle = 0, \qquad \langle \zeta(x_1)\zeta(x_1')\rangle = \delta^2 W(|x_1 - x_1'|). \tag{3.1}$$

In equations (3.1) the angle brackets denote an average over the ensemble of realizations of $\zeta(x_1)$, $\delta = \langle \zeta^2(x_1)\rangle^{1/2}$ is the rms height of the surface and, $W(|x_1|)$ is the surface height autocorrelation function. In numerical examples we will use the Gaussian form for $W(|x_1|)$:

$$W(|x_1|) = \exp(-x_1^2/a^2), \tag{3.2}$$

where a is the transverse correlation length of the surface roughness.

A reciprocal phase-perturbation theory for the scattering matrix $S(q|k)$ was constructed in [27] and [28]. The term of lowest order in the surface profile function was shown to have the form

$$S(q|k) = \int_{-\infty}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{-\mathrm{i}(q-k)x_1} \mathrm{e}^{-2\mathrm{i}\sqrt{\alpha_0(q)\alpha_0(k)}\zeta(x_1)}. \tag{3.3}$$

Since

$$\langle S(q|k)\rangle = 2\pi\delta(q-k)\mathrm{e}^{-2\delta^2\alpha_0(q)\alpha_0(k)}, \tag{3.4}$$

we can write the expression for $\delta S(q|k)$ as

$$\delta S(q|k) = \int_{-\infty}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{-\mathrm{i}(q-k)x_1}\left[\mathrm{e}^{-2\mathrm{i}\sqrt{\alpha_0(q)\alpha_0(k)}\zeta(x_1)} - \mathrm{e}^{-2\delta^2\alpha_0(q)\alpha_0(k)}\right]. \tag{3.5}$$

We calculate the averages $\langle \delta S(q|k)\delta S^*(q'|k')\rangle$ and $\langle \delta S(q|k)\delta S(q'|k')\rangle$ using the expression (3.5) for the scattering matrix. For $\langle \delta S(q|k)\delta S^*(q'|k')\rangle$ we obtain

$$\begin{aligned}\langle \delta S(q|k)\delta S^*(q'|k')\rangle = {} & \int_{-\infty}^{\infty} \mathrm{d}x_1 \int_{-\infty}^{\infty} \mathrm{d}x_1'\, \mathrm{e}^{-\mathrm{i}(q-k)x_1 + \mathrm{i}(q'-k')x_1'} \\ & \times \left\langle \left[\mathrm{e}^{-2\mathrm{i}\sqrt{\alpha_0(q)\alpha_0(k)}\zeta(x_1)} - \mathrm{e}^{-2\delta^2\alpha_0(q)\alpha_0(k)}\right] \right. \\ & \left. \times \left[\mathrm{e}^{2\mathrm{i}\sqrt{\alpha_0(q')\alpha_0(k')}\zeta(x_1')} - \mathrm{e}^{-2\delta^2\alpha_0(q')\alpha_0(k')}\right]\right\rangle \end{aligned} \tag{3.6}$$

$$= \mathrm{e}^{-2\delta^2(\alpha_0(q)\alpha_0(k)+\alpha_0(q')\alpha_0(k'))}\int_{-\infty}^{\infty}\mathrm{d}x_1\int_{-\infty}^{\infty}\mathrm{d}x_1'\,\mathrm{e}^{-\mathrm{i}(q-k)x_1+\mathrm{i}(q'-k')x_1'}$$
$$\times\left[\mathrm{e}^{4\delta^2\sqrt{\alpha_0(q)\alpha_0(q')\alpha_0(k)\alpha_0(k')}W(|x_1-x_1'|)}-1\right] \tag{3.7}$$
$$= 2\pi\delta(q-k-q'+k')\mathrm{e}^{-2\delta^2(\alpha_0(q)\alpha_0(k)+\alpha_0(q')\alpha_0(k'))}$$
$$\times\int_{-\infty}^{\infty}\mathrm{d}u\left[\mathrm{e}^{4\delta^2\sqrt{\alpha_0(q)\alpha_0(q')\alpha_0(k)\alpha_0(k')}W(|u|)}-1\right]\mathrm{e}^{-\mathrm{i}(q'-k')u}, \tag{3.8}$$

while for $\langle\delta S(q|k)\delta S(q'|k')\rangle$ we have

$$\langle\delta S(q|k)\delta S(q'|k')\rangle = \int_{-\infty}^{\infty}\mathrm{d}x_1\int_{-\infty}^{\infty}\mathrm{d}x_1'\,\mathrm{e}^{-\mathrm{i}(q-k)x_1-\mathrm{i}(q'-k')x_1'}$$
$$\times\Big\langle\left[\mathrm{e}^{-2\mathrm{i}\sqrt{\alpha_0(q)\alpha_0(k)}\zeta(x_1)}-\mathrm{e}^{-2\delta^2\alpha_0(q)\alpha_0(k)}\right]$$
$$\times\left[\mathrm{e}^{-2\mathrm{i}\sqrt{\alpha_0(q')\alpha_0(k')}\zeta(x_1')}-\mathrm{e}^{-2\delta^2\alpha_0(q')\alpha_0(k')}\right]\Big\rangle \tag{3.9}$$
$$= \mathrm{e}^{-\delta^2(\alpha_0(q)\alpha_0(k)+\alpha_0(q')\alpha_0(k'))/2}\int_{-\infty}^{\infty}\mathrm{d}x_1\int_{-\infty}^{\infty}\mathrm{d}x_1'\,\mathrm{e}^{-\mathrm{i}(q-k)x_1-\mathrm{i}(q'-k')x_1'}$$
$$\times\left[\mathrm{e}^{-4\delta^2\sqrt{\alpha_0(q)\alpha_0(q')\alpha_0(k)\alpha_0(k')}W(|x_1-x_1'|)}-1\right] \tag{3.10}$$
$$= 2\pi\delta(q-k+q'-k')\mathrm{e}^{-2\delta^2(\alpha_0(q)\alpha_0(k)+\alpha_0(q')\alpha_0(k'))}$$
$$\times\int_{-\infty}^{\infty}\mathrm{d}u\left[\mathrm{e}^{-4\delta^2\sqrt{\alpha_0(q)\alpha_0(q')\alpha_0(k)\alpha_0(k')}W(|u|)}-1\right]\mathrm{e}^{-\mathrm{i}(q'-k')u}. \tag{3.11}$$

It is readily seen that, in contrast to $\langle\delta S(q|k)\delta S^*(q'|k')\rangle$, the average $\langle\delta S(q|k)\delta S(q'|k')\rangle$ vanishes with increasing roughness parameters δ and a, due to the negative exponential under the integral sign in the last line of equation (3.11). Plots of the normalized correlation functions $\Xi^{(1)}(q,k|q',k')$ and $\Xi^{(10)}(q,k|q',k')$ as functions of δ for different values of a are presented in figure 1(a), while plots of the envelopes of the correlation functions $C^{(1)}$ and $C^{(10)}$ as functions of δ for different values of a are presented in figure 1(b), for fixed values of q, k and q', while k' is determined by the constraint of the corresponding δ function. When calculating the results presented in figures 1(a) and (b) the value of q' was chosen to produce the same values of $C^{(1)}$ and $C^{(10)}$ in the limit of a weakly rough surface. From the plots presented in figure 1(a) we see that $\Xi^{(10)}(q,k|q',k')$ vanishes even for quite moderately weakly rough surfaces for which $\Xi^{(1)}(q,k|q',k')$ is still about unity. We note that $C^{(1)}$ also decreases with increasing δ (figure 1(b)). Using equations (2.11)–(2.14), (3.8) and (3.11) we obtain the expressions for $\langle(\delta S_1(q|k))^2\rangle$, $\langle(\delta S_2(q|k))^2\rangle$ and $\langle\delta S_1(q|k)\delta S_2(q|k)\rangle$:

$$\langle(\delta S_1(q|k))^2\rangle = \mathrm{e}^{-4\delta^2\alpha_0(q)\alpha_0(k)}\left[\frac{L_1}{2}\int_{-\infty}^{\infty}\mathrm{d}u\cos(q-k)u(\mathrm{e}^{\delta^2\alpha_0(q)\alpha_0(k)W(|u|)}-1)\right.$$
$$\left.+\tfrac{1}{2}\pi\delta(q-k)\int_{-\infty}^{\infty}\cos(q-k)u(\mathrm{e}^{-\delta^2\alpha_0(q)\alpha_0(k)W(|u|)}-1)\right], \tag{3.12}$$

and

$$\langle(\delta S_2(q|k))^2\rangle = \mathrm{e}^{-4\delta^2\alpha_0(q)\alpha_0(k)}\left[\frac{L_1}{2}\int_{-\infty}^{\infty}\mathrm{d}u\cos(q-k)u(\mathrm{e}^{\delta^2\alpha_0(q)\alpha_0(k)W(|u|)}-1)\right.$$
$$\left.-\tfrac{1}{2}\pi\delta(q-k)\int_{-\infty}^{\infty}\cos(q-k)u(\mathrm{e}^{-\delta^2\alpha_0(q)\alpha_0(k)W(|u|)}-1)\right], \tag{3.13}$$

while

$$\langle\delta S_1(k|k)\delta S_2(k|k)\rangle = 0. \tag{3.14}$$

In figure 2 we present plots of the ratio $\langle(\delta S_2(k|k))^2\rangle/\langle(\delta S_1(k|k))^2\rangle$ as a function of the rms height of the surface roughness δ. Since this ratio is calculated for the specular direction $q=k$,

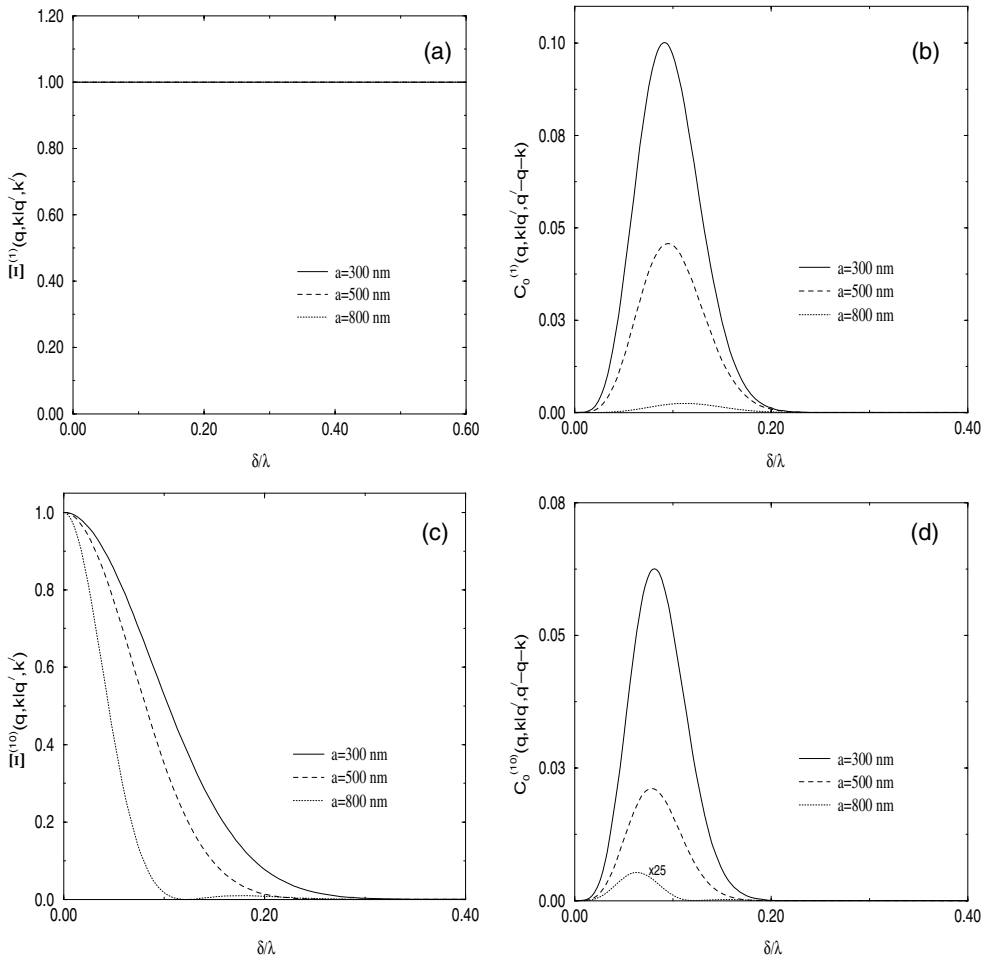

Figure 1. The normalized correlation functions $\Xi^{(1)}$ (*a*) and $\Xi^{(10)}$ (*c*) and the envelopes $C_0^{(10)}$ (*b*) and $C_0^{(10)}$ (*d*) as functions of δ/λ for values of the transverse correlation length $a = 300$ nm, 500 nm and 800 nm. The incident light was s-polarized and of wavelength 632.8 nm. The scattering medium was a randomly rough perfect conductor. Furthermore $\theta_0 = 30^\circ$, $\theta_s = 0^\circ$ and $\theta_s' = 0^\circ$. In figure 1(*a*) the results for the different correlation lengths considered could not be distinguished.

it is independent of the transverse correlation length a. From the plot presented it is easily seen that for large values of the rms height, the incoherent part of the scattering matrix, $\delta S(q|k)$, becomes a circular complex Gaussian variable, even in the specular direction.

4. Light scattering from a randomly rough penetrable surface

The results of the preceding section enable us to make several conclusions when studying the scattering of light from a randomly rough surface of a penetrable medium. For simplicity we consider here the scattering of s-polarized light from a randomly rough surface of a medium characterized by a dielectric function $\epsilon(\omega)$. As is well known (see, e.g., [29–31]), if the surface profile function is such that the conditions for the applicability of the Rayleigh hypothesis are

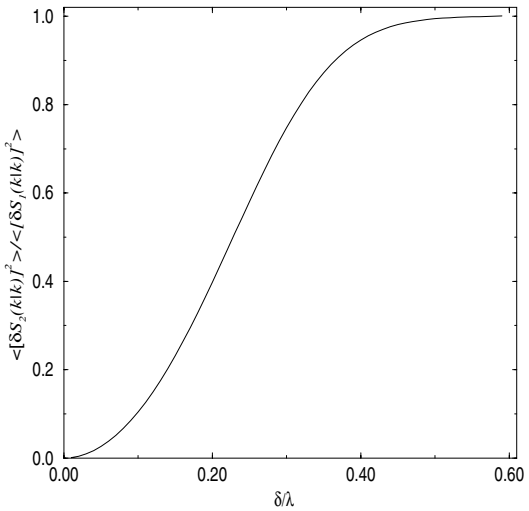

Figure 2. The ratio $\langle(\delta S_2(k|k))^2\rangle/\langle(\delta S_1(k|k))^2\rangle$ as a function of δ/λ.

satisfied the scattering amplitude $R(q|k)$ obeys the reduced Rayleigh equation. Rewritten in terms of the scattering matrix $S(q|k)$ it has the form

$$S(q|k) = 2\pi\delta(q-k)R_0(k) + N(q|k) + \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi}M(q|p)S(p|k), \tag{4.1}$$

where, for the case of the scattering of s-polarized light,

$$R_0(k) = \frac{\alpha_0(k)-\alpha(k)}{\alpha_0(k)+\alpha(k)}, \tag{4.2}$$

$$\alpha_0(k) = \sqrt{\frac{\omega^2}{c^2}-k^2}, \qquad \alpha(k) = \sqrt{\epsilon(\omega)\frac{\omega^2}{c^2}-k^2}, \tag{4.3}$$

$$N(q|k) = -(\epsilon-1)\frac{\omega^2/c^2}{\alpha_0(q)+\alpha(q)}\sqrt{\frac{\alpha_0(q)}{\alpha_0(k)}}\,\frac{J(\alpha(p)+\alpha_0(k)|p-k)}{\alpha(p)+\alpha_0(k)}, \tag{4.4}$$

$$M(q|k) = -(\epsilon-1))\frac{\omega^2/c^2}{\alpha_0(q)+\alpha(q)}\sqrt{\frac{\alpha_0(q)}{\alpha_0(p)}}\,\frac{J(\alpha(p)-\alpha_0(k)|p-k)}{\alpha(p)-\alpha_0(k)}, \tag{4.5}$$

and

$$J(\gamma|Q) = \int_{-\infty}^{\infty}\mathrm{d}x_1\,\mathrm{e}^{-\mathrm{i}Qx_1}(\mathrm{e}^{-\mathrm{i}\gamma\zeta(x_1)}-1). \tag{4.6}$$

We can write the solution of equation (4.1) formally as

$$\begin{aligned}S(q|k) = {} & R_0(k)2\pi\delta(q-k) + F(q|k) + \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi}M(q|p)F(p|k)\\ & + \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi}M(q|p)\int_{-\infty}^{\infty}\frac{\mathrm{d}p'}{2\pi}M(p|p')F(p'|k) + \cdots,\end{aligned} \tag{4.7}$$

where

$$F(q|k) = N(q|k) + M(q|k)R_0(k), \tag{4.8}$$

and we keep all terms in the infinite iterative series. Both $N(q|k)$ and $M(q|p)$ contain the surface disorder only in the functions $J(\gamma|Q)$. Therefore, having in hand the recipe for

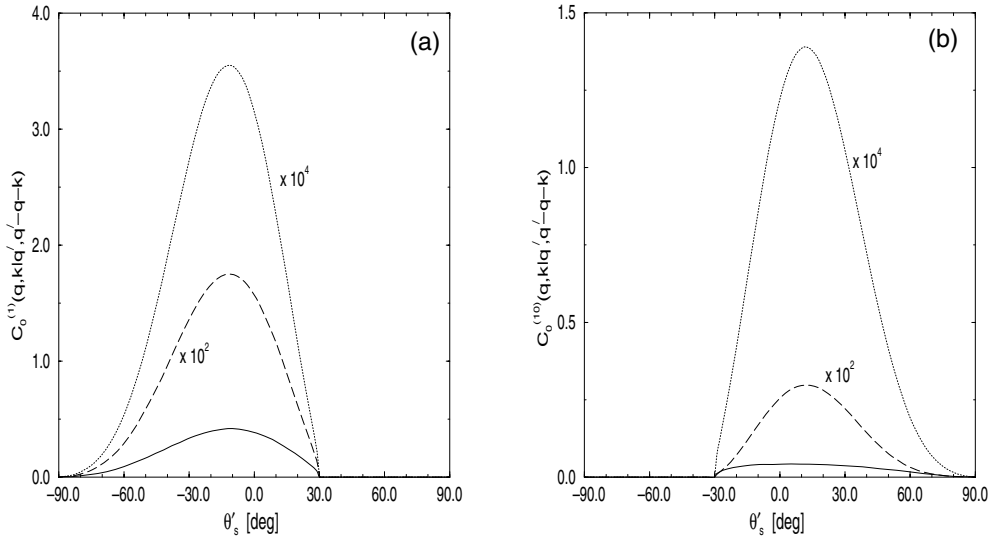

Figure 3. The envelopes of the $C^{(1)}$ (*a*) and $C^{(10)}$ (*b*) correlation functions as functions of θ_s' for $\theta_0 = 30^\circ$ and $\theta_s = 0^\circ$, while θ_0' is determined by the constraints of the corresponding δ functions for the scattering of s-polarized light from a randomly rough silver surface with $a = 500$ nm and $\delta = 20$ nm (full curves), $\delta = 50$ nm (broken curves) and $\delta = 100$ nm (dotted curves).

calculating the average of the product of any number of functions $J(\gamma|Q)$, we can calculate, in principle, both $\langle \delta S(q|k)\delta S(q'|k')\rangle$ and $\langle \delta S(q|k)\delta S^*(q'|k')\rangle$. The basics of such calculations were described in [32].

To calculate the averages $\langle \delta S(q|k)\delta S(q'|k')\rangle$ and $\langle \delta S(q|k)\delta S^*(q'|k')\rangle$ we multiply the series (4.7) for $S(q|k)$ by the corresponding series for $S(q'|k')$ and average the product term-by-term. From the result we subtract the product $\langle S(q|k)\rangle\langle S(q'|k')\rangle$. In a similar fashion we calculate the average $\langle \delta S(q|k)\delta S^*(q'|k')\rangle$ by multiplying the series (4.7) for $S(q|k)$ by the complex conjugate of the corresponding series for $S(q'|k')$, averaging the product term-by-term, and subtracting the product $\langle S(q|k)\rangle\langle S^*(q'|k')\rangle$ from the result. In the product$\langle \delta S(q|k)\delta S^*(q'|k')\rangle$ the contribution of nth order in the functions $J(\gamma|Q)$ and $J^*(\gamma|Q)$ contains $n-1$ terms of the form

$$\sum_{m=1}^{n-1}\left\{\left\langle\prod_{r=1}^{m} J(\gamma_r|Q_r)\prod_{s=1}^{n-m} J^*(\gamma_s'|Q_s')\right\rangle - \left\langle\prod_{r=1}^{m} J(\gamma_r|Q_r)\right\rangle\left\langle\prod_{s=1}^{n-m} J^*(\gamma_s'|Q_s')\right\rangle\right\}. \quad (4.9)$$

To obtain a nonzero contribution, for each value of m at least one $J(\gamma_r|Q_r)$ must be contracted with at least one $J^*(\gamma_s'|Q_s')$. Therefore each term in this sum contains at least one factor with a positive exponential of the form $\exp\{\delta^2\gamma\gamma' W(|u|)\} - 1$. In contrast, when calculating $\langle \delta S(q|k)\delta S(q'|k')\rangle$ the contribution of the nth order in the functions $J(\gamma|Q)$ contains the sum

$$\sum_{m=1}^{n-1}\left\{\left\langle\prod_{r=1}^{m} J(\gamma_r|Q_r)\prod_{s=1}^{n-m} J(\gamma_s'|Q_s')\right\rangle - \left\langle\prod_{r=1}^{m} J(\gamma_r|Q_r)\right\rangle\left\langle\prod_{s=1}^{n-m} J(\gamma_s'|Q_s')\right\rangle\right\}. \quad (4.10)$$

In this case, to obtain a nonzero contribution, for each value of m at least one $J(\gamma_r|Q_s)$ must be contracted with at least one $J(\gamma_s'|Q_s')$. Therefore, each term in this sum contains only negative exponentials of the form $\exp\{-\delta^2\gamma\gamma' W(|u|)\} - 1$. Owing to this lack of the positive exponential, $\langle \delta S(q|k)\delta S(q'|k')\rangle$ vanishes when the roughness parameters increase.

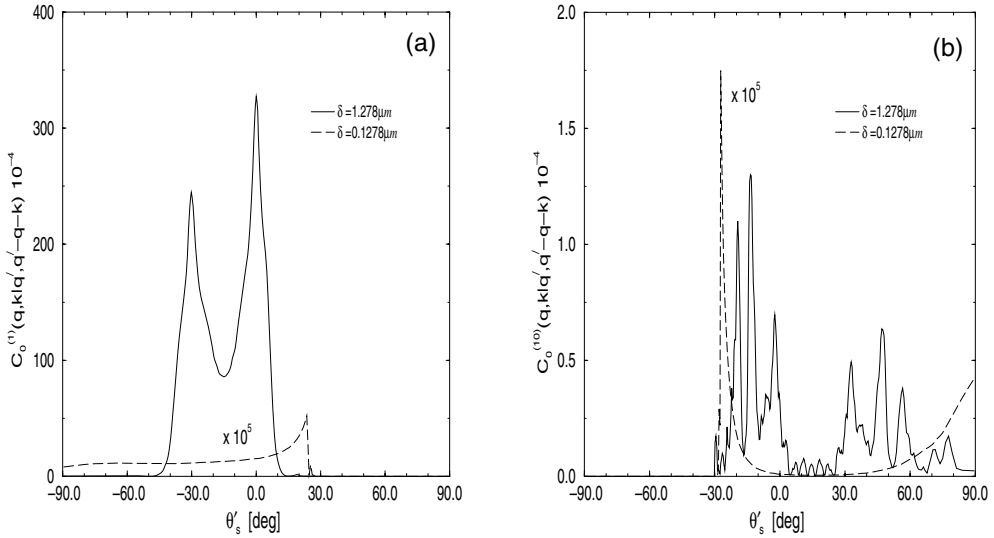

Figure 4. The same as in figure 3, but for $a = 3.85$ μm and $\delta = 1.278$ μm (full curves) and $\delta = 0.1278$ μm (broken curves).

In figure 3 we present plots of the envelopes $C_0^{(1)}$ and $C_0^{(10)}$ of the correlation functions $C^{(1)}$ (figure 3(a)) and $C^{(10)}$ (figure 3(b)) as functions of θ_s' for fixed values of θ_0 and θ_s, while θ_0' is determined by the constraints of the corresponding δ functions. The calculations were carried out for the scattering of s-polarized light, of 612.7 nm wavelength, from a weakly rough random surface of silver characterized by the complex dielectric constant $\epsilon = -17.2 + \mathrm{i}0.479$ for different values of the roughness parameters δ and a. In calculating the results presented in figure 3 we kept all terms in the infinite iterative series equation (4.7) which would give contributions to the averages we calculate through terms of $\mathrm{O}(\delta^8)$ if they were to be expanded in powers of the small parameter $(\omega/c)\delta$.

In figure 4 we present rigorous numerical simulation calculation results [33] for the envelopes of the correlation functions $C^{(1)}$ (figure 4(a)) and $C^{(10)}$ (figure 4(b)). The surface parameters used here were the same as those used in obtaining figure 3, except that the roughness now was $\delta = 1.278$ μm (full curves) and $\delta = 0.1278$ μm (broken curves). It should be pointed out that for the scattering of s-polarized light from a weakly rough random metal surface, there should be no memory or reciprocal memory effect present in $C_0^{(1)}$. This is indeed confirmed by our numerical calculations where the $C_0^{(1)}$ for $\delta = 0.1278$ μm (figure 4(a), broken curve) is a smooth function of its argument, as well as by the results presented in figure 3(a). In particular, there are no peaks at angles $\theta = 0^\circ$ and 30°, which are the positions of the memory and reciprocal memory effects. As the roughness is increased to $\delta = 1.278$ μm one sees from figure 4(a) (full curve) that the overall amplitude of the envelope $C_0^{(1)}$ is increased and, more importantly, that two peaks have developed at the aforementioned angles. These peaks are due, in the large roughness limit, to volume waves scattered multiply at the rough surface. In figure 4(b) the corresponding results for the $C_0^{(10)}$ envelopes are presented. It is observed that in the low roughness limit this envelope is structureless, and that $C_0^{(1)}$ and $C_0^{(10)}$ are roughly of the same order of magnitude. However, as δ is increased, the scattering matrix $S(q|k)$ starts to obey circular complex Gaussian statistics and thus, as discussed earlier, the envelope $C_0^{(10)}$ should in principle vanish. From our numerical results for $\delta = 1.278$ μm (full curve) we

indeed see that $C_0^{(10)}$ is much smaller then the corresponding $C_0^{(1)}$ shown in figure 4(a). In fact, $C_0^{(10)}$ is just noise, consistent with this function vanishing in the large roughness limit.

5. Conclusions

In this paper we calculated the angular intensity correlation functions $C(q, k|q', k')$ by means of an approach that explicitly separates out different contributions to it. We have shown that calculations and measurements of the correlation function $C(q, k|q', k')$ yield important information about the statistical properties of the amplitude of the scattered field. In particular, we have shown that the short-range correlation function $C^{(10)}$ is, in a sense, a measure of the noncircularity of the complex Gaussian statistics of the scattering matrix. Thus, if the random surface is such that only the $C^{(1)}$ and $C^{(10)}$ correlation functions are observed, then $S(q|k)$ obeys complex Gaussian statistics. If the random surface is such that only $C^{(1)}$ is observed, then $S(q|k)$ obeys circular complex Gaussian statistics. Finally, if the random surface is such that $C^{(1.5)}$, $C^{(2)}$ and $C^{(3)}$ are observed, in addition to both $C^{(1)}$ and $C^{(10)}$, then $S(q|k)$ is not a Gaussian random process. In addition, we can conclude that if a surface is sufficiently weakly rough and long enough, its speckle pattern should display the memory and reciprocal memory effects when the angle of incidence is changed and, for a fixed angle of incidence, should be symmetric about the specular direction. However, if the roughness of the surface is sufficiently great and the surface is long enough, its speckle pattern should display only the memory and reciprocal memory effects, when the angle of incidence is changed.

Acknowledgment

The work of TAL and AAM was supported in part by Army Research Office grant no DAAD 19-99-1-0321.

References

[1] Michel T R and O'Donnell K A 1992 *J. Opt. Soc. Am.* A **9** 1374
[2] Knotts M E, Michel T R and O'Donnell K A 1992 *J. Opt. Soc. Am.* A **9** 1822
[3] Nieto-Vesperinas M and Sánchez-Gil J A 1992 *Phys. Rev.* B **46** 3112
[4] Nieto-Vesperinas M and Sánchez-Gil J A 1993 *J. Opt. Soc. Am.* A **10** 150
[5] Arsenieva A D and Feng S 1993 *Phys. Rev.* B **47** 13 047
[6] Nieto-Vesperinas M and Sánchez-Gil J A 1993 *Phys. Rev.* B **48** 4132
[7] Kuga Y, Le C T C, Ishimaru A and Ailes-Sengers L 1996 *IEEE Trans. Geosci. Remote Sens.* **34** 1300
[8] Le C T C, Kuga Y and Ishimaru A 1996 *J. Opt. Soc. Am.* A **13** 1057
[9] Malyshkin V, McGurn A R, Leskova T A, Maradudin A A, and Nieto–Vesperinas M 1997 *Opt. Lett.* **22** 946
[10] Malyshkin V, McGurn A R, Leskova T A, Maradudin A A, and Nieto–Vesperinas M 1997 *Waves Random Media* **7** 479
[11] Nieto-Vesperinas M, Maradudin A A, Shchegrov A V and McGurn A R 1997 *Opt. Commun.* **142** 1
[12] West C S and O'Donnell K A 1999 *Phys. Rev.* B **59** 2393
[13] Lu J Q and Gu Z-H 1997 *Appl. Opt.* **36** 4562
Gu Z-H and Lu J Q 1997 *SPIE* **3141** 269
[14] McGurn A R, Maradudin A A and Celli V 1985 *Phys. Rev.* B **31** 4866
[15] McGurn A R and Maradudin A A 1989 *Opt. Commun.* **72** 279
[16] Feng S, Kane C, Lee P A and Stone A D 1988 *Phys. Rev. Lett.* **61** 834
[17] Shapiro B 1986 *Phys. Rev. Lett.* **57** 2168
[18] Garcia N and Genack A Z 1989 *Phys. Rev. Lett.* **63** 1678
[19] van Albada M P, de Boer J F and Lagendijk A 1990 *Phys. Rev. Lett.* **64** 2787

[20] Freund I and Rosenbluh M 1991 *Opt. Commun.* **82** 362
[21] Kubo R 1962 *J. Phys. Soc. Japan* **17** 1100
[22] Stuart A and Keith Ord J 1987 *Kendall's Advanced Theory of Statistics* vol 1, 5th edn (London: Charles Griffin) p 84ff
[23] Stoffregen B 1979 *Optik* **52** 385
[24] Goodman J W 1985 *Statistical Optics* (New York: Wiley) ch 2
[25] Goodman J W 1975 *Opt. Commun.* **14** 324
[26] Goodman J W 1984 *Laser Speckle and Related Phenomena* ed J C Dainty (Berlin: Springer) ch 2
[27] Shen J and Maradudin A A 1980 *Phys. Rev.* B **22** 4234–40
[28] Fitzgerald R M and Maradudin A A 1994 *Waves Random Media* **4** 275
[29] Lord Rayleigh 1985 *The Theory of Sound* vol 2 (London: Macmillan) pp 89, 297
[30] Petit R and Cadilhac M 1966 *C. R. Acad. Sci., Paris* B **262** 468
[31] Hill N R and Celli V 1978 *Phys. Rev.* B **17** 2478
[32] Leskova T A, Maradudin A A and Novikov I 2000 *J. Opt. Soc. Am.* A **17** 1288
[33] Maradudin A A, Michel T, McGurn A R and Méndez E R 1990 *Ann. Phys., NY* **203** 255

Scattering of light from the random interface between two dielectric media with low contrast

Tamara A. Leskova

Institute of Spectroscopy, Russian Academy of Sciences, Troitsk 142092, Russia

Alexei A. Maradudin and Igor V. Novikov

Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, Irvine, California 92697

Received August 31, 1999; revised manuscript received March 1, 2000; accepted March 28, 2000

We calculate the coherent and incoherent scattering of p- and s-polarized light incident from a dielectric medium characterized by a real, positive, dielectric constant ϵ_0 onto its one-dimensional, randomly rough interface with a dielectric medium characterized by a real, positive, dielectric constant ϵ. We use a perturbation theory with a new small parameter, namely, the dielectric contrast $\eta = \epsilon_0 - \epsilon$ between the medium of incidence and the scattering medium. The proper self-energy entering the expression for the reflectivity is obtained as an expansion in powers of η through the second order in η, and the reducible vertex function in terms of which the scattered intensity is expressed is obtained as an expansion in powers of η through the fourth. The roughness-induced shifts of the Brewster angle (in p polarization) and of the critical angle for total internal reflection ($\epsilon_0 > \epsilon$) are obtained. The angular dependence of the intensity of the incoherent component of the scattered light displays an enhanced backscattering peak, which is due to the coherent interference of multiply scattered lateral waves supported by the interface and their reciprocal partners. Analogs of the Yoneda peaks observed in the scattering of x rays from solid surfaces are also present. The results obtained by our small-contrast perturbation theory are in good agreement with those obtained in computer simulation studies.
© 2000 Optical Society of America [S0740-3232(00)01207-2]
OCIS codes: 240.5770, 290.5880, 290.1350, 290.4210, 260.6970.

1. INTRODUCTION

Each perturbation approach to rough-surface scattering has its own small parameter. Thus in small-amplitude perturbation theory,[1] it is δ/λ, where δ is the rms height of the surface and λ is the wavelength of the incident field; in the small-slope approximation,[2] the small parameter is the rms slope of the surface. In this paper we present a new perturbation theory for rough-surface scattering with a new small parameter, namely, the dielectric contrast between the medium of incidence and the scattering medium.

To illustrate this new perturbation theory, we apply it to calculations of the coherent (specular) and incoherent (diffuse) scattering of light incident from a dielectric medium characterized by an isotropic, real, positive, frequency-independent dielectric constant ϵ_0 onto its one-dimensional, randomly rough interface with a second dielectric medium characterized by an isotropic, real, positive, frequency-independent dielectric constant ϵ. In these calculations we assume that the dielectric contrast $\eta = \epsilon_0 - \epsilon$ is small—of the order of a few tenths. The incident light is either p polarized or s polarized, with the plane of incidence normal to the generators of the random interface. In this geometry there is no cross-polarized scattering, and the plane of scattering coincides with the plane of incidence.

In calculating the reflectivities of this interface for p- and s-polarized light we have chosen to obtain them within the framework of self-energy perturbation theory.[3] However, in contrast with our earlier use of self-energy perturbation theory in calculating the reflectivities of random dielectric surfaces,[4,5] in which the self-energy that enters the expression for the reflectivities of p- and s-polarized light was calculated perturbatively as an expansion in powers of δ/λ, in this paper we calculate it as an expansion in powers of η.

The choice of self-energy perturbation for calculating reflectivities was made for two reasons. It was shown recently that self-energy perturbation theory is well suited to the determination of the roughness-induced shift of the Brewster angle (in p polarization) from its value for a planar interface.[4] In addition, we will show below that self-energy perturbation theory is also well suited to the study of the roughness-induced shift of the critical angle for total internal reflection (when $\epsilon_0 > \epsilon$) from its value for a planar interface in both p and s polarization. The study of these two roughness-induced angular shifts is the first motivation for the present work.

In studying the incoherent scattering of p- and s-polarized light from a low-contrast random dielectric interface, we calculate the reducible vertex function, which determines the intensity of the light that has been scattered incoherently, as an expansion in powers of the dielectric contrast, through terms of the fourth order in η rather than in powers of δ/λ. We show that, as was suggested in Ref. 6, in the scattering of light from a comparatively weakly rough interface between two dielectric media the lateral waves supported by the interface are

excited (see, e.g., Ref. 7), and the coherent interference of each multiply scattered lateral wave path with its reciprocal partner leads to the appearance of an enhanced backscattering peak. Thus in this case it is the weak localization of lateral waves that leads to the backscattering enhancement in our results. The demonstration that lateral waves can give rise to enhanced backscattering of light from weakly rough random surfaces that support neither surface nor guided waves is the second main motivation for the present work. We also show that the angular dependence of the intensity of the light scattered incoherently displays asymmetric peaks when the scattering angle equals the critical angle for total internal reflection (for $\epsilon_0 > \epsilon$), which are analogous to the Yoneda peaks observed in the scattering of x rays from a solid surface.[8]

We note that perturbative expansions in powers of the dielectric contrast have been used recently in a study of the coherent and incoherent scattering of x rays from randomly rough metal surfaces,[9] in which case the value of the contrast is of the order of $\eta = 10^{-6} - 10^{-3}$. However, in that study the intensity of the light scattered coherently—the reflectivity—was calculated directly as an expansion in powers of the dielectric contrast, with terms of the zeroth and second orders in η retained, instead of by the use of self-energy perturbation theory. This result, therefore, lacked the contribution from the higher-order scattering processes that are already present in the latter approach when a low-order expansion of the proper self-energy in powers of η is made. In addition, the intensity of the x rays scattered incoherently was also expanded in powers of the dielectric contrast, with terms of the zeroth and second orders in η obtained. In this approximation no backscattering enhancement can be obtained. The presentation of the present improved version of small-contrast perturbation theory, together with results obtained by its use, is the third main motivation for the present work.

2. CHARACTERIZATION OF A ROUGH INTERFACE

The physical system that we consider in this paper consists of a dielectric medium, characterized by a real, positive, dielectric constant ϵ_0 in the region $x_3 > \zeta(x_1)$ and a second dielectric medium, characterized by a real, positive, dielectric constant ϵ in the region $x_3 < \zeta(x_1)$ (Fig. 1). The profile function $\zeta(x_1)$ is assumed to be a single-valued function of x_1 that is differentiable as many times as is necessary and that constitutes a stationary, zero-mean, Gaussian random process, defined by the properties

$$\langle \zeta(x_1) \rangle = 0, \tag{2.1}$$

$$\langle \zeta(x_1)\zeta(x_1') \rangle = \delta^2 W(|x_1 - x_1'|), \tag{2.2}$$

$$\delta^2 = \langle \zeta^2(x_1) \rangle. \tag{2.3}$$

The angle brackets in Eqs. (2.1)–(2.3) denote an average over the ensemble of realizations of the profile function $\zeta(x_1)$, and δ is the rms height of the surface.

It is also necessary to introduce the Fourier integral representation of $\zeta(x_1)$:

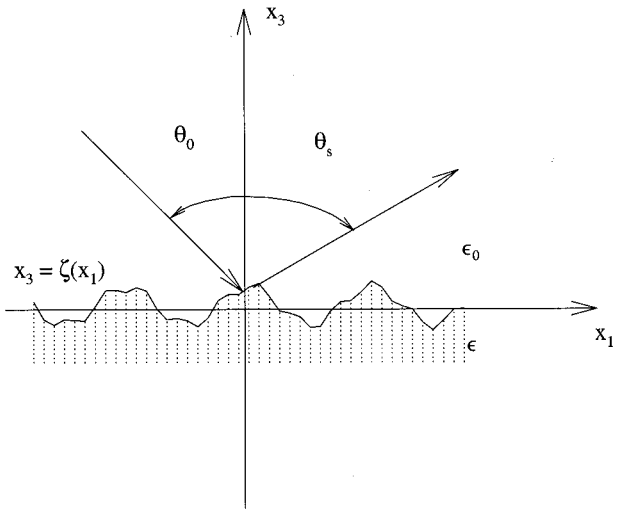

Fig. 1. Physical system studied in this paper.

$$\zeta(x_1) = \int_{-\infty}^{\infty} \frac{dQ}{2\pi} \hat{\zeta}(Q) \exp(iQx_1). \tag{2.4}$$

The Fourier coefficient $\hat{\zeta}(Q)$ is also a zero-mean, Gaussian random process, defined by

$$\langle \hat{\zeta}(Q) \rangle = 0, \tag{2.5}$$

$$\langle \hat{\zeta}(Q)\hat{\zeta}(Q') \rangle = 2\pi\delta(Q + Q')\delta^2 g(|Q|), \tag{2.6}$$

where $g(|Q|)$, the power spectrum of the surface roughness, is related to the surface height autocorrelation function $W(|x_1|)$ by

$$g(|Q|) = \int_{-\infty}^{\infty} dx_1 W(|x_1|) \exp(-iQx_1). \tag{2.7}$$

In this paper we will present results calculated for a random interface characterized by a power spectrum that has the Gaussian form

$$g(|Q|) = \sqrt{\pi} a \exp(-Q^2 a^2/4), \tag{2.8}$$

where a is the transverse correlation length of the interface roughness, as well as for an interface characterized by a power spectrum given by

$$g(|Q|) = (\pi/\Delta k)(\mathrm{rect}\{[Q - (\sqrt{\epsilon}\omega/c)]/\Delta k\} + \mathrm{rect}\{[Q + (\sqrt{\epsilon}\omega/c)]/\Delta k\}), \tag{2.9}$$

where $\Delta k = 2\sqrt{\epsilon}(\omega/c)\sin\theta_{\max}$, and rect denotes the rectangle function. The latter type of surface has been used in recent experimental studies,[10] where its use increased the strength of the enhanced backscattering caused by the surface plasmon polariton mechanism[11] by an order of magnitude compared with that for a surface with the same rms height and rms slope characterized by a Gaussian power spectrum.

3. REDUCED RAYLEIGH EQUATIONS

We will study the scattering of both p- and s-polarized light, incident from the medium whose dielectric constant is ϵ_0 on its interface $x_3 = \zeta(x_1)$ with the medium whose dielectric constant is ϵ. We can treat both polarizations

together by introducing the function $\Phi(x_1, x_3|\omega)$, which is $H_2(x_1, x_3|\omega)$ when the incident light is p polarized, and $E_2(x_1, x_3|\omega)$ when the incident light is s polarized. Then in the region $x_3 > \zeta(x_1)_{\max}$ the function $\Phi(x_1, x_3|\omega)$ is the sum of an incoming incident wave and outgoing scattered waves,

$$\Phi^{>}(x_1, x_3|\omega) = \exp[ikx_1 - i\alpha_0(k)x_3] + \int_{-\infty}^{\infty} \frac{dq}{2\pi} R(q|k) \times \exp[iqx_1 + i\alpha_0(q)x_3], \tag{3.1}$$

while in the region $x_3 < \zeta(x_1)_{\min}$ it consists of outgoing transmitted waves,

$$\Phi^{<}(x_1, x_3|\omega) = \int_{-\infty}^{\infty} \frac{dq}{2\pi} \mathcal{T}(q|k)\exp[iqx_1 - i\alpha(q)x_3], \tag{3.2}$$

where

$$\begin{aligned} \alpha_0(q) &= [\epsilon_0(\omega^2/c^2) - q^2]^{1/2}, & \operatorname{Re}\alpha_0(q) &> 0, \\ & & \operatorname{Im}\alpha_0(q) &> 0, \\ \alpha(q) &= [\epsilon(\omega^2/c^2) - q^2]^{1/2}, & \operatorname{Re}\alpha(q) &> 0, \\ & & \operatorname{Im}\alpha(q) &> 0. \end{aligned} \tag{3.3}$$

In the scattering of electromagnetic waves from a one-dimensional randomly rough surface or interface, what is measured is a mean differential reflection coefficient. This is defined as the fraction of the total energy incident on the surface per unit time that is scattered into an angular interval $d\theta_s$ about the scattering direction defined by the scattering angle θ_s (Fig. 1), averaged over the ensemble of realizations of the profile function $\zeta(x_1)$.

In the present case the contribution to the mean differential reflection coefficient from the coherent component of the scattered light is given by

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle_{\text{coh}} = \frac{1}{L_1}\sqrt{\epsilon_0}\frac{\omega}{2\pi c}\frac{\cos^2\theta_s}{\cos\theta_0}|\langle R(q|k)\rangle|^2, \tag{3.4a}$$

where L_1 is the length of the x_1 axis covered by the random interface, while the contribution from the incoherent component of the scattered light is

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle_{\text{incoh}} = \frac{1}{L_1}\sqrt{\epsilon_0}\frac{\omega}{2\pi c}\frac{\cos^2\theta_s}{\cos\theta_0}[\langle |R(q|k)|^2\rangle - |\langle R(q|k)\rangle|^2]. \tag{3.4b}$$

In Eqs. (3.4), q and k are given by $q = \sqrt{\epsilon_0}(\omega/c)\sin\theta_s$ and $k = \sqrt{\epsilon_0}(\omega/c)\sin\theta_0$. The scattering amplitude $R(q|k)$ satisfies the reduced Rayleigh equation[9]

$$\int_{-\infty}^{\infty} \frac{dq}{2\pi} M(p|q)R(q|k) = -N(p|k), \tag{3.5}$$

where

$$M_p(p|q) = \frac{pq + \alpha(p)\alpha_0(q)}{\alpha(p) - \alpha_0(q)} I[\alpha(p) - \alpha_0(q)|p - q],$$

$$N_p(p|k) = \frac{pk - \alpha(p)\alpha_0(k)}{\alpha(p) + \alpha_0(k)} I[\alpha(p) + \alpha_0(k)|p - k], \tag{3.6a}$$

and

$$M_s(p|q) = \frac{I[\alpha(p) - \alpha_0(q)|p - q]}{\alpha(p) - \alpha_0(q)},$$

$$N_s(p|k) = \frac{I[\alpha(p) + \alpha_0(k)|p - k]}{\alpha(p) + \alpha_0(k)}, \tag{3.6b}$$

while $I(\gamma|Q)$ is defined by

$$I(\gamma|Q) = \int_{-\infty}^{\infty} dx_1 \exp\{-i[Qx_1 + \gamma\zeta(x_1)]\}. \tag{3.7}$$

We will seek the solution of Eq. (3.5) in the form

$$R(q|k) = 2\pi\delta(q - k)R^{(0)}(k) - 2iG^{(0)}(q)T(q|k)G^{(0)}(k)\alpha_0(k), \tag{3.8}$$

where

$$\begin{aligned} R_p^{(0)}(k) &= \frac{\epsilon\alpha_0(k) - \epsilon_0\alpha(k)}{\epsilon\alpha_0(k) + \epsilon_0\alpha(k)} \equiv \frac{\Delta_p(k)}{d_p(k)}, \\ R_s^{(0)}(k) &= \frac{\alpha_0(k) - \alpha(k)}{\alpha_0(k) + a(k)} \equiv \frac{\Delta_s(k)}{d_s(k)}, \end{aligned} \tag{3.9}$$

and

$$\begin{aligned} G_p^{(0)}(k) &= \frac{i\epsilon}{\epsilon\alpha_0(k) + \epsilon_0\alpha(k)}, \\ G_s^{(0)}(k) &= \frac{i}{\alpha_0(k) + \alpha(k)}. \end{aligned} \tag{3.10}$$

The transition matrix $T(q|k)$ is postulated to satisfy the equation

$$\begin{aligned} T(q|k) &= V(q|k) + \int_{-\infty}^{\infty} \frac{dp}{2\pi} V(q|p)G^{(0)}(p)T(p|k) \\ &= V(q|k) + \int_{-\infty}^{\infty} \frac{dp}{2\pi} T(q|p)G^{(0)}(p)V(p|k), \end{aligned} \tag{3.11}$$

and the scattering potential $V(q|k)$ is then found to be the solution of the equation

$$\int_{-\infty}^{\infty} \frac{dq}{2\pi}[N(p|q) - M(p|q)]\frac{V(q|k)}{\alpha_0(q)} = \frac{N(p|k) + M(p|k)R^{(0)}(k)}{G^{(0)}(k)\alpha_0(k)}. \tag{3.12}$$

We will also need the Green's function $G(q|k)$ associated with the randomly rough interface between the medium of incidence and the scattering medium. It is defined as the solution of

$$G(q|k) = 2\pi\delta(q - k)G^{(0)}(k) + G^{(0)}(q) \times \int_{-\infty}^{\infty} \frac{dp}{2\pi} V(q|p)G(p|k) \tag{3.13a}$$

$$= 2\pi\delta(q - k)G^{(0)}(k) + G^{(0)}(q)T(q|k)G^{(0)}(k). \tag{3.13b}$$

From Eqs. (3.8)–(3.10) and (3.13b) we obtain the useful relation

$$R(q|k) = -2\pi\delta(q-k) - 2iG(q|k)\alpha_0(k). \quad (3.14)$$

In what follows we will need to be able to expand $V(q|k)$ in powers of the dielectric contrast $\eta = \epsilon_0 - \epsilon$ instead of in powers of $\zeta(x_1)$. To do this, we must transform Eq. (3.12) into a form in which the dependence on η is explicit. This can be achieved by making the replacement

$$I(\gamma|Q) = 2\pi\delta(Q) + J(\gamma|Q) \quad (3.15)$$

in Eqs. (3.6), where

$$J(\gamma|Q) = \int_{-\infty}^{\infty} dx_1 \exp(-iQx_1)\{\exp[-i\gamma\zeta(x_1)] - 1\}. \quad (3.16)$$

If we introduce the definition

$$V(q|k) = -iv(q|k)/\tilde{\epsilon}, \quad (3.17)$$

where $\tilde{\epsilon} = \epsilon$ is in p polarization and $\tilde{\epsilon} = 1$ is in s polarization, the reduced potential $v(q|k)$ is the solution of the equation

$$v(q|k) = \eta A(q|k) + \eta\int_{-\infty}^{\infty}\frac{dp}{2\pi}F(q|p)v(p|k). \quad (3.18)$$

The matrices $A(q|k)$ and $F(q|k)$ entering this equation are given by

$$A(q|k) = [n(q|k)d(k) + m(q|k)\Delta(k)]\frac{1}{2\tilde{\epsilon}\alpha_0(k)}, \quad (3.19)$$

$$F(q|k) = [m(q|k) - n(q|k)]\frac{1}{2\tilde{\epsilon}\alpha_0(k)}, \quad (3.20)$$

where

$$m_p(q|k) = \frac{qk + \alpha(q)\alpha_0(k)}{\alpha(q) - \alpha_0(k)}J[\alpha(q) - \alpha_0(k)|q-k],$$

$$n_p(q|k) = \frac{qk - \alpha(q)\alpha_0(k)}{\alpha(q) + \alpha_0(k)}J[\alpha(q) + \alpha_0(k)|q-k], \quad (3.21a)$$

and

$$m_s(q|k) = \frac{\omega^2/c^2}{\alpha(q) - \alpha_0(k)}J[\alpha(q) + \alpha_0(k)|q-k],$$

$$n_s(q|k) = \frac{(\omega^2/c^2)}{\alpha(q) + \alpha_0(k)}J[\alpha(q) + \alpha_0(k)|q-k]. \quad (3.21b)$$

The iterative solution of Eq. (3.18) yields the reduced potential $v(q|k)$ formally as an expansion in powers of η:

$$v(q|k) = \eta A(q|k) + \eta^2\int_{-\infty}^{\infty}\frac{dp_1}{2\pi}F(q|p_1)A(p_1|k) + \eta^3\int_{-\infty}^{\infty}\frac{dp_1}{2\pi}\int_{-\infty}^{\infty}\frac{dp_2}{2\pi}F(q|p_1)F(p_1|p_2) \times A(p_2|k) + \cdots. \quad (3.22)$$

4. COHERENT SCATTERING

We see from Eq. (3.4a) that for the calculation of the contribution to the mean differential reflection coefficient from the coherent component of the scattered light we need the ensemble average of $R(q|k)$. In view of Eq. (3.14) this is given by

$$\langle R(q|k)\rangle = -2\pi\delta(q-k) - 2i\langle G(q|k)\rangle\alpha_0(k). \quad (4.1)$$

Owing to the stationarily of $\zeta(x_1)$, the averaged Green's function $\langle G(q|k)\rangle$ appearing in Eq. (4.1) has the form[12]

$$\langle G(q|k)\rangle = 2\pi\delta(q-k)G(k) = 2\pi\delta(q-k)\frac{1}{G^{(0)}(k)^{-1} - M(k)}, \quad (4.2)$$

where $G(k)$ is the averaged Green's function and $M(k)$ is the averaged proper self-energy. The latter is given by

$$\langle M(q|k)\rangle = 2\pi\delta(q-k)M(k), \quad (4.3)$$

where the (unaveraged) proper self-energy $M(q|k)$ is the solution of[12]

$$M(q|k) = V(q|k) + \int_{-\infty}^{\infty}\frac{dp}{2\pi}\int_{-\infty}^{\infty}\frac{dr}{2\pi}M(q|p) \times \langle G(p|r)\rangle w(r|k), \quad (4.4)$$

where we have introduced the notation

$$w(q|k) = V(q|k) - \langle M(q|k)\rangle. \quad (4.5)$$

When the explicit expressions for $G^{(0)}(k)$ given by Eqs. (3.10) are used in Eqs. (4.1) and (4.2), we find that

$$\langle R(q|k)\rangle = 2\pi\delta(q-k)r(\theta_0), \quad (4.6)$$

where

$$r_p(\theta_0) = \frac{\epsilon\sqrt{\epsilon_0}\cos\theta_0 - \epsilon_0(\epsilon_0\cos^2\theta_0 - \eta)^{1/2} + i\epsilon(c/\omega)M_p[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0]}{\epsilon\sqrt{\epsilon_0}\cos\theta_0 + \epsilon_0(\epsilon_0\cos^2\theta_0 - \eta)^{1/2} - i\epsilon(c/\omega)M_p[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0]}, \quad (4.7a)$$

$$r_s(\theta_0) = \frac{\sqrt{\epsilon_0}\cos\theta_0 - (\epsilon_0\cos^2\theta_0 - \eta)^{1/2} + i(c/\omega)M_s[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0]}{\sqrt{\epsilon_0}\cos\theta_0 + (\epsilon_0\cos^2\theta_0 - \eta)^{1/2} - i(c/\omega)M_s[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0]}. \quad (4.7b)$$

It follows from Eq. (3.4a) that the reflectivities for p- and s-polarized light are given by

$$R_{p,s}(\theta_0) = |r_{p,s}(\theta_0)|^2. \tag{4.8}$$

We will calculate the self-energies $M_{p,s}(k)$ as expansions in powers of η, through terms of second order. This is done in two steps. Equation (4.4) is solved iteratively as an expansion in powers of $V(q|k)$, and the result is averaged term by term. The result can be written in the form

$$\langle M(q|k)\rangle = \sum_{n=1}^{\infty} \langle M_n(q|k)\rangle, \tag{4.9}$$

where the subscript denotes the order of the corresponding term in $V(q|k)$. The terms in this expansion that we need are

$$\langle M_1(q|k)\rangle = \{V(q|k)\}, \tag{4.10a}$$

$$\langle M_2(q|k)\rangle = \int_{-\infty}^{\infty} \frac{dp}{2\pi} \frac{\{V(q|p)V(p|k)\}}{G^{(0)}(p)^{-1} - M(p)}, \tag{4.10b}$$

where the bracket symbol $\{...\}$ denotes the cumulant average.[9,13] The expansion of $M(q|k)$ in powers of $V(q|k)$ does not, of course, imply an expansion in powers of η. To obtain $\langle M(q|k)\rangle$ as an expansion in powers of η, we must use the expansion of $V(q|k)$ in powers of η given by Eq. (3.22) and evaluate the cumulant averages appearing in $\langle M_n(q|k)\rangle$ as expansions in powers of η. The contribution to $\langle M_1(q|k)\rangle$ of lowest order in η, i.e., the contribution to $\langle V(q|k)\rangle$ of lowest order, is of $O(1)$, while the contribution to $\langle M_n(q|k)\rangle$, $n = 2,3,...$, of lowest order in η is of $O(\eta^n)$. This will be clearly seen from the following.

From Eqs. (4.10a), (4.10b), and (3.22), we see that to calculate the ensemble average of $V(q|k)$ and products of any number of $V(q|k)$'s we need the average of $J(\gamma|Q)$ and of products of an arbitrary number of $J(\gamma|Q)$'s. With the assumption that $\zeta(x_1)$ is a Gaussian random process, we find that

$$\begin{aligned}
\langle J_1 J_2 \ldots J_n\rangle = {} & \prod_{i=1}^{n} \{J_i\} + \theta(n-2) \sum_{\substack{i,j=1 \\ (i<j)}}^{n} \{J_i J_j\} \prod_{\substack{k=1 \\ (k\neq i,j)}}^{n} \{J_k\} \\
& + \theta(n-3) \sum_{\substack{i,j,k=1 \\ (i<j<k)}}^{n} \{J_i J_j J_k\} \prod_{\substack{l=1 \\ (l\neq i,j,k)}}^{n} \{J_l\} \\
& + \theta(n-4) \sum_{\substack{i,j,k,l=1 \\ (i<j<k<l)}}^{n} \{J_i J_j J_k J_l\} \\
& \times \prod_{\substack{m=1 \\ (m\neq i,j,k,l)}}^{n} \{J_m\} + \theta(n-4) \\
& \times \sum_{\substack{i,j=1 \\ (i<j)}}^{n} \sum_{\substack{k,l=1 \\ (k<l)}}^{n} \{J_i J_j\}\{J_k J_l\} \prod_{\substack{m=1 \\ (m\neq i,j,k,l)}}^{n} \{J_m\} \\
& (i \neq k,l; j \neq k,l) \\
& + \text{terms with fewer than } n-4 \text{ factors of } \{J_p\},
\end{aligned} \tag{4.11}$$

where $\theta(n) = 1$ for $n \geq 0$ and $\theta(n) = 0$ for $n < 0$. The reason for organizing the contributions to $\langle J_1 J_2 \ldots J_n\rangle$ according to the number of factors of $\{J_p\}$ they contain, in decreasing order, is that when we come to average, for example, the series (3.22) term by term, the results given by Eq. (4.11) tell us that the actual order in η of each of the contributions to the average of a term formally proportional to η^n in fact depends on the number of factors of $\langle A(q|k)\rangle$ and $\langle F(q|k)\rangle$ contained in that average when it is evaluated with the use of Eq. (4.11). This is because these averages are diagonal in q and k and, more important, are proportional to $1/\eta$,

$$\langle A(q|k)\rangle = 2\pi\delta(q-k)\frac{a(k)}{\eta}, \tag{4.12a}$$

$$\langle F(q|k)\rangle = 2\pi\delta(q-k)\left[-\frac{f(k)}{\eta}\right], \tag{4.12b}$$

where

$$a(k) = \frac{d(k)\Delta(k)}{2\tilde{\epsilon}\alpha_0(k)}[Y(k) - X(k)], \tag{4.13a}$$

$$f(k) = \frac{d(k)X(k) + \Delta(k)Y(k)}{2\tilde{\epsilon}\alpha_0(k)} - 1, \tag{4.13b}$$

with

$$X(k) = \exp\left\{-\frac{\delta^2}{2}[\alpha(k) - \alpha_0(k)]^2\right\}, \tag{4.14a}$$

$$Y(k) = \exp\left\{-\frac{\delta^2}{2}[\alpha(k) + \alpha_0(k)]^2\right\}. \tag{4.14b}$$

The cumulant averages of products of two or more $F(q|k)$ or $A(q|k)$, or both, do not possess an explicit dependence on η.

Thus the dominant contribution to the expansion of the average of a product of n factors of $F(q|k)$ and $A(q|k)$, e.g., $\langle F(q|p_1)F(p_1|p_2)\ldots A(p_{n-1}|k)\rangle$, in powers of η is obtained by replacing each of the factors by its average. Each of these averages is proportional to $1/\eta$, so that this contribution is of $O(\eta^{-n})$. The next most important term in this expansion is given by the sum of the contributions obtained by pairing two of the factors in all possible ways irrespective of their order [there are $n(n-1)/2$ such pairs], evaluating the cumulant average of their product, and then multiplying the result by the product of the average of each of the remaining unpaired factors. Since there are $n-2$ such factors in an nth-order product, the resulting contribution to $\langle F(q|p_1)F(p_1|p_2)\ldots A(p_{n-1}|k)\rangle$ is of $O(\eta^{2-n})$. Terms of $O(\eta^{3-n})$ and higher are obtained in an analogous way.

The consequences of the preceding results for the calculations of the cumulant averages of products of the scattering potential $V(q|k)$ are two in number. The first is that the dominant contribution to the nth-order cumulant average $\{V(q_1|p_1)\ldots V(q_n|p_n)\}$ for $n \geq 2$ is of $O(\eta^n)$. It is also the case that each cumulant average, e.g., $\{V(q_1|p_1)V(q_2|p_2)\}$, is proportional to a delta function whose argument is the sum of the wave-number transfers in each scattering process occurring within it, viz., to

$2\pi\delta(q_1 - p_1 + q_2 - p_2)$ for the cumulant average indicated. The latter result leads to $\langle M(q|k)\rangle$ having the form given by Eq. (4.3). These results have the consequence that the leading two terms in the expansion of the right-hand side of Eq. (4.10a) in powers of η are of $O(\eta^0)$ and of $O(\eta^2)$; the leading term on the right-hand side of Eq. (4.10b) is of $O(\eta^2)$. Thus if we wish to calculate the self-energy $M(k)$ to second order in η, we need to evaluate the contributions to $\{V(q|k)\}$ of $O(\eta^0)$ and of $O(\eta^2)$ and add the result to that obtained from Eq. (4.10b) by evaluating $\{V(q|p)V(p|k)\}$ in the numerator of the integrand to $O(\eta^2)$ and replacing $M(p)$ in the denominator of the integrand by the contribution of $O(\eta^0)$ obtained from Eq. (4.10a). After the summation of all the contributions is done (see the detailed description in Appendix A), we obtain the expression for the self-energy $M(k)$ in the form

$$M(k) = \frac{1}{i\tilde{\epsilon}}\frac{a(k)}{1 + f(k)} + \eta^2\left(\frac{1}{i\tilde{\epsilon}}\right)^2\frac{1}{[1 + f(k)]^2} \times \int_{-\infty}^{\infty}\frac{dp}{2\pi}\{m(k|p)G^{(0)}(p)X^{-1}(p)A_1(p|k)\}, \tag{4.15}$$

where

$$A_1(q|k) = [n(q|k)d(k)X(k) + m(q|k)\Delta(k)Y(k)]\frac{1}{2\tilde{\epsilon}\alpha_0(k)}. \tag{4.16}$$

The evaluation of the cumulant average in the integrand of the second term in brackets on the right-hand side of Eq. (4.15) utilizes the result that

$$\{J(\gamma_1|Q_1)J(\gamma_2|Q_2)\} = 2\pi\delta(Q_1 + Q_2)\exp\left[-\frac{1}{2}(\gamma_1^2 + \gamma_2^2)\delta^2\right] \times \int_{-\infty}^{\infty}du\,\exp(-iQ_1u)\{\exp[-\gamma_1\gamma_2\delta^2W(|u|)] - 1\}. \tag{4.17}$$

The resulting integrals in $M(k)$ have to be evaluated numerically, recalling that $k = \sqrt{\epsilon_0}(\omega/c)\sin\theta_0$.

5. INCOHERENT SCATTERING

From Eq. (3.14), we find that

$$[\langle|R(q|k)|^2\rangle - |\langle R(q|k)\rangle|^2] = 4\alpha_0^2(k)[\langle|G(q|k)|^2\rangle - |\langle G(q|k)\rangle|^2]. \tag{5.1}$$

The equation for the averaged two-particle Green's function $\langle|G(q|k)|^2\rangle$ can be obtained starting from the equation for the Green's function written in a form[12]

$$G(q|k) = G(q)2\pi\delta(q - k) + G(q)t(q|k)G(k), \tag{5.2}$$

which incorporates into its calculation the averaged Green's function $G(p)$ instead of the unperturbed Green's function $G^{(0)}(p)$. In Eq. (5.2) the operator $t(p|r)$ was introduced to satisfy

$$\int_{-\infty}^{\infty}\frac{dp}{2\pi}w(q|p)G(p|k) = t(q|k)G(k), \tag{5.3}$$

where w has been defined by Eq. (4.5) and is the solution of the equation

$$t(q|k) = w(q|k) + \int_{-\infty}^{\infty}\frac{dp}{2\pi}w(q|p)G(p)t(p|k). \tag{5.4}$$

It follows from Eq. (5.2) that $\langle t(q|k)\rangle = 0$. When we multiply Eq. (5.2) by its complex conjugate and use the procedure outlined in Ref. 12, we obtain

$$\langle G(q|k)G^*(q|k)\rangle = (2\pi)^2\delta^2(q - k)|G(k)|^2 + |G(q)|^2\{\tau(q,k|q,k)\}|G(k)|^2, \tag{5.5}$$

where $\tau(q,k|q',k') = t(q|k)t^*(q'|k')$ is the reducible vertex function, and we have used the result that $\langle G(q|k)\rangle = 2\pi\delta(q - k)G(k)$ and $\langle t(q|k)\rangle = 0$. The use of Eq. (5.5) in Eqs. (5.1) and (3.4b) yields the equation on which our calculation of $\langle\partial R/\partial\theta_s\rangle_{\text{incoh}}$ is based:

$$\left\langle\frac{\partial R}{\partial\theta_s}\right\rangle_{\text{incoh}} = \frac{1}{L_1}\sqrt{\epsilon_0}\frac{2}{\pi}\left(\frac{\omega}{c}\right)^3\cos^2\theta_s\cos\theta_0|G(q)|^2 \times\langle\tau(q,k|q,k)\rangle|G(k)|^2. \tag{5.6}$$

In this paper we will calculate $\langle\tau(q,k|q,k)\rangle$ through terms of fourth order in η, since it is in the fourth-order contribution that the phenomenon of enhanced backscattering first arises. To evaluate $\langle\tau(q,k|q,k)\rangle$ we need to evaluate the cumulant average $\{t(q|k)t^*(q|k)\}$ to fourth order in η. To do this we multiply the iterative solution of Eq. (5.4),

$$t(q|k) = w(q|k) + \int_{-\infty}^{\infty}\frac{dp}{2\pi}w(q|p)G(p)w(p|k) + \int_{-\infty}^{\infty}\frac{dp}{2\pi}\int_{-\infty}^{\infty}\frac{dr}{2\pi}w(q|p)G(p)w(p|r) \times G(r)w(r|k) + \cdots, \tag{5.7}$$

by its complex conjugate and average the resulting expression term by term.

Following the procedure of averaging described in Section 4 and Appendix A, we obtain to fourth order in η,

$$\{\tau(q,k|q,k)\} = \frac{\eta^2}{\tilde{\epsilon}^2}\frac{1}{|1 + f(q)|^2}[\tilde{\tau}_2(q|k) + \eta\tilde{\tau}_3(q|k) + \eta^2\tilde{\tau}_4(q|k)]\frac{1}{|1 + f(k)|^2}, \tag{5.8}$$

where, to make the expressions for $\tilde{\tau}_n(q|k)$ clearer, we introduced a new notation for the cumulant averages, viz.,

$$\{J(\gamma_1|Q_1)\ldots J(\gamma_n|Q_n)\} = 2\pi\delta(Q_1 + Q_2 + \cdots + Q_n)\{J(\gamma_1|Q_1)\ldots J(\gamma_n|Q_n)\}_0,$$

so that after making use of the delta functions in carrying out the integrals, we obtained

$$\tilde{\tau}_2(q|k) = L_1\{A_1(q|k)A_1^*(q|k)\}_0, \tag{5.9a}$$

$$\tilde{\tau}_3(q|k) = 2L_1 \,\mathrm{Im}\left\{ m(q|p)\frac{G(p)}{1+f(p)} \times A_1(p|k)A_1^*(q|k)\right\}_0, \tag{5.9b}$$

$$\begin{aligned}\tilde{\tau}_{4(1)}(q|k) = L_1\Bigg[&\left\{ m(q|p)\frac{G(p)}{1+f(p)}A_1(p|k)m^*(q|r) \times \frac{G^*(r)}{1+f^*(r)}A_1^*(r|k)\right\}_0 \\ &+ 2\,\mathrm{Re}\left\{ m(q|p)\frac{G(p)}{1+f(p)} \times m(p|r)\frac{G(r)}{1+f(r)}A_1(r|k)A_1^*(q|k)\right\}_0\Bigg],\end{aligned} \tag{5.9c}$$

$$\begin{aligned}\tilde{\tau}_{4(2)}(q|k) = L_1\Bigg[&\left\{ m(q|p)\frac{G(p)}{1+f(p)}\{A_1(p|k) \times m^*(q|q+k-p)\}_0\frac{G^*(q+k-p)}{1+f^*(q+k-p)} \times A_1^*(q+k-p|k)\right\}_0 \\ &+ \left\{ m(q|p)\frac{G(p)}{1+f(p)}m^*(q|p)\right\}_0 \times \left\{ A_1(p|k)\frac{G^*(p)}{1+f^*(p)}A_1^*(p|k)\right\}_0 \\ &+ 2\,\mathrm{Re}\left\{ m(q|p)\frac{G(p)}{1+f(p)} \times A_1(p-q+k|k)\right\}_0 \left\{ m(p|p-q+k) \times \frac{G(p-q+k)}{1+f(p-q+k)}A_1^*(q|k)\right\}_0\Bigg],\end{aligned} \tag{5.9d}$$

where $A_1(q|k)$ is given by Eq. (4.16).

Having in hand the explicit expressions for the various cumulant averages $\{\ldots\}_0$ in Eq. (5.8) and the averaged Green's function, obtained in Section 4, we can calculate numerically the angular distribution of the intensity of the scattered light.

6. DISCUSSION

A. Brewster Angle and Critical Angle for Total Internal Reflection

In Figs. 2(a) and 2(b) are plotted the logarithms of the reflectivities $R_p(\theta_0)$ and $R_s(\theta_0)$, respectively, calculated on the basis of Eqs. (4.7a), (4.7b), and (4.8), when the medium of incidence ($\epsilon_0 = 2.4336$) is optically less dense than the scattering medium ($\epsilon = 2.6569$); in Figs. 2(c) and 2(d) the same quantities are plotted for the case where $\epsilon_0 = 2.6569$ and $\epsilon = 2.4336$. The dielectric contrast is therefore $\eta = -0.2233(+0.2233)$ in the former (latter) case. The random roughness of the interface is characterized by the Gaussian height autocorrelation function, Eq. (2.8), and the roughness parameters in both cases are $\delta = 0.2\lambda$ and $a = 5\lambda$, where λ is the vacuum wavelength of the incident light.

In the absence of the interface roughness, the reflectivities of p- and s-polarized light are given by Eqs. (4.8) with $M_{p,s} = 0$. $R_p^{(0)}(\theta_0)$ is seen to vanish at an angle of incidence given by

$$\theta_0 = \theta_B^{(0)} = \tan^{-1}\left(\frac{\epsilon}{\epsilon_0}\right)^{1/2} \tag{6.1}$$

and called the Brewster angle. A corresponding angle does not exist in the case of s polarization.

In the presence of interface roughness the reflectivity $R_p(\theta_0)$ does not vanish at any real value of θ_0 because the term $i\epsilon(c/\omega)M_p[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0]$, appearing in the numerator of the expression (6.13a) for $r_p(\theta_0)$, is complex. In this case we can define the Brewster angle θ_B as the angle of incidence at which $R_p(\theta_0)$ possesses a minimum.[4,14] Note that this definition of the Brewster angle reduces to the usual one, Eq. (6.1), when the random roughness tends to zero, as it should.

The Brewster effect is clearly seen in the plots presented in the insets in Figs. 2(a) and 2(c). For the values of ϵ_0 and ϵ used in obtaining these plots, we find from Eq. (6.1) that the Brewster angle in the case of a planar interface has the values $\theta_B^{(0)} = 46.257°$ and $\theta_B^{(0)} = 43.743°$, respectively. From these figures we see that in the presence of the random roughness the Brewster angle is shifted toward smaller values from its value for a planar

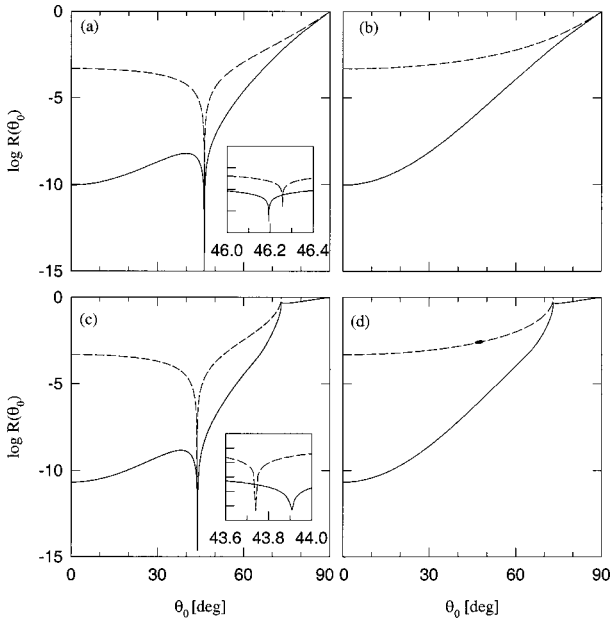

Fig. 2. Logarithm of the reflectivity of a one-dimensional randomly rough dielectric–dielectric interface (solid curves) and of a planar interface (dashed curves) as functions of the angle of incidence, θ_0, when $\epsilon_0 = 2.4336$ and $\epsilon = 2.6569$ for (a) p polarization and (b) s polarization, and when $\epsilon_0 = 2.6569$ and $\epsilon = 2.4336$ for (c) p polarization and (d) s polarization.

interface when $\epsilon_0 < \epsilon$ and toward larger values when $\epsilon_0 > \epsilon$. The former result is in agreement with results obtained earlier for the roughness-induced shift of the Brewster angle obtained by several different approaches in the scattering of p-polarized light, incident from vacuum, on a one-dimensional random dielectric surface.[4,5,14–16] The latter result seems unexpected at first sight, since we kept only the contributions to the averaged self-energy of zeroth and second orders in η. Its explanation lies in the fact that, according to our definition of the Brewster angle in the presence of interface roughness and to Eq. (4.7a), its value is given approximately by the solution of the equation

$$\epsilon\sqrt{\epsilon_0}\cos\theta_0 - \epsilon_0(\epsilon_0\cos^2\theta_0 - \eta)^{1/2} = -\mathrm{Re}\{i\epsilon(c/\omega)M_p[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0]\}, \quad (6.2)$$

where the right-hand side is positive for either sign of η. In the vicinity of the Brewster angle for a planar interface, $\theta_B^{(0)} = \tan^{-1}(\epsilon/\epsilon_0)^{1/2}$, the left-hand side of Eq. (6.2) can be expanded as $\eta[(\sqrt{\epsilon_0}/\epsilon)(\epsilon_0 + \epsilon)\sin\theta_B^{(0)}](\theta_0 - \theta_B^{(0)})$. Thus the roughness-induced shift of the Brewster angle is given approximately by

$$\theta_0 - \theta_B^{(0)} = \frac{\epsilon(c/\omega)\mathrm{Im}\,M_p[\sqrt{\epsilon_0}(\omega/c)\sin\theta_B^{(0)}]}{\eta[(\sqrt{\epsilon_0}/\epsilon)(\epsilon_0 + \epsilon)\sin\theta_B^{(0)}]}. \quad (6.3)$$

Since the sign of the remaining factors is positive irrespective of the sign of η, the presence of the factor $1/\eta$ on the right-hand side of Eq. (6.3) shows that the shift is positive when η is positive and negative when η is negative. From a comparison of the results presented in the insets to Figs. 2(a) and 2(c), we also see that the magnitude of the shift of the Brewster angle from its value for a planar interface is larger when $\epsilon_0 > \epsilon$ than when $\epsilon_0 < \epsilon$. The shift of the Brewster angle depends on the roughness parameters through the imaginary part of the self-energy. In general, this dependence cannot be written out explicitly. However, in the limit of a small correlation length, $a \to 0$, the shift is inversely proportional to the correlation length, $\theta_B - \theta_B^{(0)} \propto \delta^2/(a\lambda)$, as was indicated by Greffet in Ref. 15. It should be pointed out, however, that the use of small-amplitude perturbation theory in Ref. 15 leads to an overestimation of the shift of the Brewster angle.[4,5] In the opposite limit of a large correlation length, the shift becomes independent of it.

Returning to Eqs. (4.7a) and (4.7b), we see that in both polarizations the reflectivity in the absence of surface roughness equals unity at an angle of incidence defined by $\epsilon - \epsilon_0\sin^2\theta_0 = 0$, i.e., for

$$\theta_0 = \theta_c^{(0)} = \sin^{-1}\left(\frac{\epsilon}{\epsilon_0}\right)^{1/2}, \quad (6.4)$$

and remains equal to unity as θ_0 increases beyond $\theta_c^{(0)}$ up to $\pi/2$, because $(\epsilon - \epsilon_0\sin^2\theta_0)^{1/2}$ is pure imaginary for $\theta_c^{(0)} < \theta_0 \leqslant \pi/2$. The angle $\theta_c^{(0)}$ is called the critical angle for total internal reflection. From Eq. (6.4) we see that in order for such an angle to exist the medium of incidence must be the optically more dense medium; i.e., we must have $\epsilon_0 > \epsilon$.

In the case of a randomly rough interface, we see from Eqs. (4.7a) and (4.7b) that the existence of a true critical angle for total internal reflection would require the satisfaction of the condition

$$C_p(\theta_0) \equiv (\epsilon - \epsilon_0\sin^2\theta_0)^{1/2} - i(\epsilon/\epsilon_0)(c/\omega) \times M_p[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0] = 0 \quad (6.5)$$

in the case of p polarization and the satisfaction of the condition

$$C_s(\theta_0) \equiv (\epsilon - \epsilon_0\sin^2\theta_0)^{1/2} - i(c/\omega) \times M_s[\sqrt{\epsilon_0}(\omega/c)\sin\theta_0] = 0 \quad (6.6)$$

in the case of s polarization. However, because $M_{p,s}(k)$ is complex, these conditions can be satisfied only by complex angles of incidence. To obtain a real critical angle in the presence of a randomly rough interface, we define it as the angle $\theta_c^{p,s}$ at which $|C_{p,s}(\theta_0)|^2$ has a minimum in p, s polarization. Again, this definition reduces to the usual one, Eq. (6.4), in the limit as the random roughness tends to zero.

The critical angle θ_c is clearly seen in the results plotted in Figs. 2(c) and 2(d). For the values of ϵ_0 and ϵ used in obtaining these plots, we find that in the case of a planar interface $\theta_c^{(0)} = 73.148°$. To examine the behavior of $R_p(\theta_0)$ and $R_s(\theta_0)$ for θ_0 in the vicinity of $\theta_c^{(0)}$, we have plotted these functions in this range of θ_0 values, using a greatly expanded horizontal scale, in Figs. 3(a) and 3(b), respectively. In the inset to each of these figures we have also plotted $|C_p(\theta_0)|^2$ and $|C_s(\theta_0)|^2$, respectively. The arrows give the position of the minimum of each of the latter functions, viz., the value of the critical angle for total internal reflection in the presence of the interface roughness. We see that, although the reflectivity varies strongly as a function of θ_0 in the vicinity of the critical angle, the shift of the critical angle is small. However, as was noted earlier in Ref. 17, the shift depends strongly on the roughness parameters, the power spectrum, etc., and can have either sign depending on them. This dependence will be discussed in detail in a separate study. In the present study, the value of the critical angle for the roughness parameters characterizing the weakly rough interfaces studied here is shifted by the roughness to values larger than its value for a planar interface for each polarization. The shift of the critical angle for total internal reflection of p-polarized light is larger than that for s-polarized light.

From an experimental standpoint, the measurement of the roughness-induced shift of θ_c could be accomplished by means of a determination of the complex scattering amplitude $r_{p,s}(\theta_0)$. It then follows that

$$C_p(\theta_0) = \frac{\epsilon}{\sqrt{\epsilon_0}}\cos\theta_0\frac{1 - r_p(\theta_0)}{1 + r_p(\theta_0)}, \quad (6.7)$$

$$C_s(\theta_0) = \sqrt{\epsilon_0}\cos\theta_0\frac{1 - r_s(\theta_0)}{1 + r_s(\theta_0)}, \quad (6.8)$$

from which the minima of $|C_p(\theta_0)|^2$ and $|C_s(\theta_0)|^2$ can be determined. The magnitude of the roughness-induced

shift of $\theta_c^{(p,s)}$, however, is smaller than the magnitude of the corresponding shift of the Brewster angle.

B. Enhanced Backscattering Phenomenon

In Fig. 4 we present the contribution to the mean differential reflection coefficient from the incoherent component of p-polarized [Figs. 4(a) and 4(c)] and s-polarized [Figs. 4(b) and 4(d)] scattered light plotted as a function of the scattering angle θ_s for a low-contrast, one-dimensional, randomly rough dielectric–dielectric interface characterized by the Gaussian height autocorrelation function Eq. (2.8). The roughness parameters are $\delta = 0.1\lambda$ and $a = 0.2\lambda$, where λ is the vacuum wavelength of the incident light, when the medium of incidence ($\epsilon_0 = 2.4336$) is optically less dense than the scattering medium ($\epsilon = 2.6569$) [Figs. 4(a) and 4(b)] and when the medium of incidence ($\epsilon_0 = 2.6569$) is optically more dense than the scattering medium ($\epsilon = 2.4336$) [Figs. 4(c) and 4(d)]. The values of η in this case are therefore $\eta = \mp 0.2233$, respectively. For comparison the contribution to the mean differential reflection coefficient from the incoherent component of p- and s-polarized scattered light calculated by means of the formally exact numerical simulation approach[18] is plotted by dashed curves.

As expected, the scattering of light from a rough interface with a low dielectric contrast is extremely weak. As seen from Fig. 4 the s-polarized light is scattered more strongly than the p-polarized light. In the case in which light is incident onto the interface from the optically more dense medium, $\langle \partial R/\partial \theta_s \rangle_{\text{incoh}}$ displays asymmetric peaks at $\theta_s = \theta_c$, analogous to the Yoneda peaks occurring in the scattering of x rays from solid surfaces.[8,9] These peaks come from the factors $\cos\theta_s |G(q)|^2$ and $\cos\theta_0 |G(k)|^2$ in Eq. (5.6). These functions increase with q and k and reach their maxima at the branch points of the Green's functions, that is, when the angles of incidence or scattering reach the critical angle θ_c for total internal reflection; the functions vanish at $\theta_0 = \pi/2$, $\theta_s = \pi/2$, owing to the presence of the factors $\cos\theta_0$ and $\cos\theta_s$. In the case of the scattering of p-polarized light incident from the optically more dense medium [Fig. 4(c)], the minimum analogous to that obtained in Refs. 6 and 19 at an angle of scattering that was called the Brewster angle for diffuse scattering cuts the peaks from the side of the smaller angles of scattering. It is easy to see from the analytical results of the small-amplitude perturbation approach to the scattering of p-polarized light[20] that the contribution to the mean differential reflection coefficient from the term of the lowest order in the surface profile function, which describes the single-scattering processes, vanishes when $\epsilon q k - \epsilon_0 \alpha(q)\alpha(k) = 0$, where

$$q = (\omega/c)\sqrt{\epsilon_0}\sin\theta_s, \quad k = (\omega/c)\sqrt{\epsilon_0}\sin\theta_0,$$

$$\alpha(q) = (\omega/c)(\cos^2\theta_s - \eta)^{1/2},$$

$$\alpha(k) = (\omega/c)(\cos^2\theta_0 - \eta)^{1/2}.$$

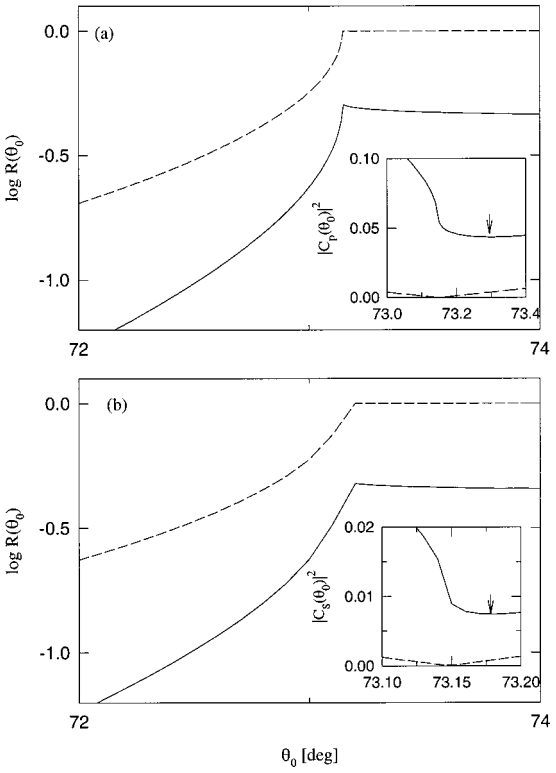

Fig. 3. Reflectivity of a one-dimensional randomly rough dielectric–dielectric interface characterized by the Gaussian power spectrum (solid curves) and of a planar interface (dashed curves) as functions of the angle of incidence, θ_0, for θ_0 in the immediate vicinity of the critical angle for total internal reflection θ_c, when $\epsilon_0 = 2.6569$ and $\epsilon = 2.4336$ for (a) p polarization and (b) s polarization. The insets show plots of $|C_p(\theta_0)|^2$ and $|C_s(\theta_0)|^2$ as functions of θ_0, for θ_0 in the immediate vicinity of θ_c. The arrows indicate the critical angle in the presence of the random roughness.

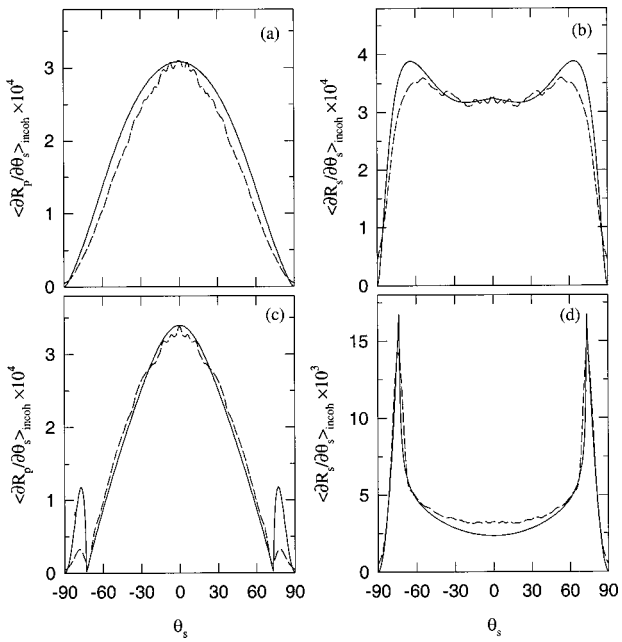

Fig. 4. Contribution to the mean differential reflection coefficient from the incoherent component of the scattered light as a function of the scattering angle θ_s when (a) p-polarized light and (b) s-polarized light are incident from the optically less dense medium and when (c) p-polarized light and (d) s-polarized light are incident from the optically more dense medium. The angle of incidence $\theta_0 = 0°$. The interface roughness is characterized by the Gaussian power spectrum Eq. (2.8) with $\delta = 0.1\lambda$ and $a = 0.2\lambda$. Dashed curves show the results of rigorous numerical simulations.

However, when the interface roughness is increased, the scattering processes of higher order smear out the minimum.

In Fig. 5 we present the results for the contribution to the mean differential reflection coefficient from the incoherent component of the *p*- and *s*-polarized scattered light when the roughness of the interface is characterized by the rectangular power spectrum given by Eq. (2.9) centered at $[\min(\epsilon, \epsilon_0)]^{1/2}(\omega/c)$, when the medium of incidence is optically less dense than the scattering medium [Figs. 5(a) and 5(b), respectively] and when the medium of incidence is optically more dense than the scattering medium [Figs. 5(c) and 5(d), respectively]. The roughness parameters in this case are $\delta = 0.1\lambda$ and $\theta_{\max} = 12°$. As in the case of the scattering from the surface with the Gaussian power spectrum, the *s*-polarized light is scattered more strongly than the *p*-polarized light and produces a stronger enhanced backscattering peak. The situation is similar to that encountered in the scattering of light from a rough dielectric surface with large rms heights and large rms slopes, when the multiple reflection of light from the slopes is important for producing the enhanced backscattering peak.[21] Since the *p*-polarized light is reflected considerably more weakly than the *s*-polarized light, the diffuse scattering of the former is also weaker. The phenomenon of total internal reflection, which amplifies the diffuse scattering and produces the enhanced backscattering peak present in the results of Ref. 22, increases the scattered intensity in the case considered in this paper but is not able to overcome the effects of the low dielectric contrast.

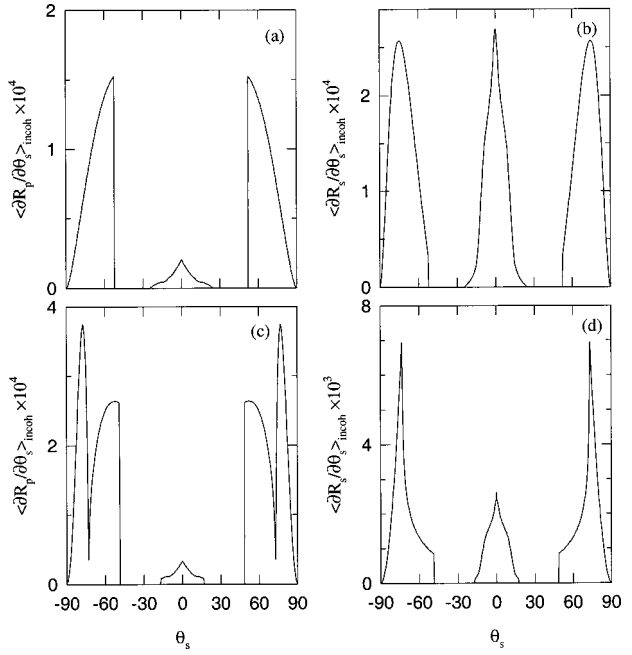

Fig. 5. Contribution to the mean differential reflection coefficient from the incoherent component of the scattered light as a function of the scattering angles θ_s when (a) *p*-polarized light and (b) *s*-polarized light are incident from the optically less dense medium and when (c) *p*-polarized light and (d) *s*-polarized light are incident from the optically more dense medium. The interface roughness is characterized by the rectangular power spectrum Eq. (2.9) centered at $(\min\{\epsilon,\epsilon_0\})^{1/2}\omega/c$ with $\delta = 0.1\lambda$ and $\theta_{\max} = 12°$.

We recall that the expression for the contribution to the mean differential reflection coefficient from the incoherent component of the scattered light [Eqs. (5.6) and (5.8)] contains contributions of second, third, and fourth orders in η. However, each contribution contains all orders in the interface profile function. Therefore there are no restrictions on the degree of the interface roughness except those imposed by the Rayleigh hypothesis, and we can study the scattering of light from fairly rough interfaces. The form of the expression for $\langle \partial R/\partial\theta_s \rangle_{\text{incoh}}$ is similar to that which is usually obtained in the framework of many-body perturbation theory. The first two contributions to $\tilde{\tau}_{4(2)}(q|k)$ are the contributions from the ladder and the maximally crossed diagrams, respectively. In the pole approximation, if we had one, the first term in Eq. (5.9d) would produce an enhanced backscattering peak with a half-width equal to the inverse mean free path of the excitations associated with poles of the Green's function $G(p)$. At the same time, the second term would produce a background of the same magnitude as the magnitude of the peak in the first term. We consider in this paper the interface between two dielectric media, characterized by positive dielectric constants. In this situation there are no surface polaritons, and the Green's function $G(p)$ does not have poles. However, it has branch points at $p = \pm\sqrt{\epsilon_0}\omega/c$ and $p = \pm\sqrt{\epsilon}\omega/c$. In the case of the dielectric–dielectric interface, the branch points lead to peaks in the product of the Green's functions $G(p) \times [G(x-p)]^*$ exactly as the poles do in the case of a metal surface.[20] They are much wider and weaker, of course, than those associated with the surface polaritons, but, nevertheless, they play exactly the same role. It is known (see, e.g., Ref. 7) that lateral waves are associated with the branch point in the Green's function at $p = \pm[\min(\epsilon, \epsilon_0)]^{1/2}\omega/c$. These are waves that propagate along the interface in a wavelike manner, are radiative waves in the optically more dense medium, and are independent of the distance from the interface in the optically less dense medium. In the case in which both media in contact are lossless, the amplitude of the lateral wave field is determined by the difference $(\sqrt{\epsilon_0} - \sqrt{\epsilon})(\omega/c)$ and decays algebraically as $L^{-3/2}$, where L is the distance along the interface. As a result, the width of the enhanced backscattering peak is proportional to the dielectric contrast. The interface roughness induces additional losses, so the enhanced backscattering peak in this case is quite broad. In addition, the peak is non-Lorentzian, since the peaks associated with the branch points are non-Lorentzian. It should also be mentioned here that the power spectrum of the interface roughness should be sufficiently large in the vicinity of the branch points, as it should also be in the vicinity of the poles associated with surface plasmon polaritons, to allow efficient excitation of the lateral waves. For a Gaussian power spectrum of the interface roughness this means that the correlation length a should be small, $a \leqslant \eta\lambda[\min(\epsilon, \epsilon_0)]^{1/2}$, where λ is the vacuum wavelength of the incident light. The use of the rectangular power spectrum $g(|Q|)$ of the interface roughness centered at $Q = [\min(\epsilon, \epsilon_0)]^{1/2}\omega/c$, Eq. (2.9), greatly amplifies the enhanced backscattering in the case in which light is scattered from the random interface between two dielectric media, as it does in the case in which

the enhanced backscattering is caused by the surface plasmon polariton mechanism.[11,20]

7. CONCLUSIONS

In the present paper we have described a perturbation approach to the problem of the scattering of light from a one-dimensional randomly rough interface between two dielectric media with low dielectric contrast. In this approach we have calculated the self-energy and the two-particle Green's function as expansions not in powers of the profile function $\zeta(x_1)$ of the interface but as an expansion in powers of the dielectric contrast $\eta = \epsilon_0 - \epsilon$. Therefore there are no restrictions on the strength of the surface roughness except those required for the validity of the Rayleigh hypothesis. Several conclusions can be drawn from the results described here.

The first is that at a one-dimensional randomly rough dielectric–dielectric interface the Brewster angle (in p polarization) is shifted from the value it has when the interface is planar. The shift is in the direction of smaller angles when the medium of incidence is optically less dense than the scattering medium ($\epsilon_0 < \epsilon$) and is in the direction of larger angles when the medium of incidence is optically more dense than the scattering medium ($\epsilon_0 > \epsilon$). The former result is consistent with results obtained for the roughness-induced shift of the Brewster angle obtained by several different approaches in the scattering of p-polarized light, incident from vacuum, on a one-dimensional random dielectric surface.[4,14–16] It is also consistent with the theoretical results obtained by Baylard *et al.*[5] and Kawanishi et al.,[19] as well as with the experimental results of De Roo and Ulaby,[23] for the scattering of p-polarized electromagnetic waves, incident from vacuum, on a two-dimensional random dielectric surface. For the dielectric contrasts assumed in the present work, $|\eta| = 0.223$, and the roughness parameters adopted, $\delta = 0.2\lambda$ and $a = 5\lambda$, the shifts of the Brewster angle calculated are of the order of 0.1°. This is a small value but one that should be measurable. The magnitude of the shift of the Brewster angle from its value for a planar interface is larger in the presence of total internal reflection than when it is absent.

The second conclusion we can draw is that in the case in which the medium of incidence is optically more dense than the scattering medium ($\epsilon_0 > \epsilon$), so that the phenomenon of total internal reflection occurs, the value of the critical angle for total internal reflection is shifted by the roughness of the interface from its value when the interface is planar. In the case in which the interface roughness is characterized by the Gaussian power spectrum, Eq. (2.8), and for the roughness parameters assumed here, the shift is in the direction of larger angles for incident light of both p and s polarization.

In the scattering of light from a one-dimensional randomly rough dielectric–dielectric interface, the angular dependence of the intensity of the incoherent component of the scattered light displays an enhanced backscattering peak, which is caused by the coherent interference of each multiply scattered lateral wave path with its reciprocal partner. When light is scattered from a rough interface whose roughness is characterized by a rectangular power spectrum centered at $[\min(\epsilon, \epsilon_0)]^{1/2}\omega/c$, the enhanced backscattering peak is amplified considerably. As expected, the enhanced backscattering peak as well as the diffuse background is stronger when light is incident onto the interface from the optically more dense medium. The most favorable situation for observing enhanced backscattering from a dielectric–dielectric interface is when s-polarized light is incident from the optically more dense medium onto a random interface whose roughness is characterized by a rectangular power spectrum. In this case the angular distribution of the intensity of the incoherent component of the scattered light also displays strong asymmetric peaks when the scattering angle equals the critical angle for total internal reflection, which are analogous to the Yoneda bands observed in the scattering of x rays from a solid surface.

Finally, we point out that a variant of small-contrast perturbation theory has been explored recently in connection with the scattering of light from, and its transmission through, a random interface between two different dielectric media.[24–26] It differs from the approach presented here in several respects. In application to a one-dimensional random interface, the expansion parameter is not $\epsilon_0 - \epsilon$ but rather the difference between the dielectric constant of the system, $\epsilon(x_1, x_3) = \epsilon_0\theta[x_3 - \zeta(x_1)] + \epsilon\theta[\zeta(x_1) - x_3]$ in our notation, where $\theta(x)$ is the Heaviside unit step function, and its ensemble average $\epsilon_{\text{eff}}(x_3) = \langle\epsilon(x_1, x_3)\rangle$. In addition, it is an essentially numerical approach. It can yield accurate results for the reflectivity[24] and transmissivity[25] for values of $\epsilon_0 - \epsilon$ up to 2.25 and 5.0, respectively, for rms heights of the roughness up to $\lambda/2$ and transverse correlation lengths of up to 2λ. In its present form, however, it cannot predict enhanced backscattering and enhanced transmission, nor cross-polarized scattering from two-dimensional random surfaces, but this appears to be due more to computational reasons than to any fundamental limitation of the theory.

APPENDIX A

From Eqs. (3.17) and (3.22) we see that to calculate $\{V(q|k)\}$ we have to calculate the average:

$$
\begin{aligned}
&\{v(q|k)\} \\
&\quad = \langle v(q|k)\rangle \\
&\quad = \eta\langle A(q|k)\rangle + \eta^2\int_{-\infty}^{\infty}\frac{dp_1}{2\pi}\langle F(q|p_1)A(p_1|k)\rangle \\
&\qquad + \eta^3\int_{-\infty}^{\infty}\frac{dp_1}{2\pi}\int_{-\infty}^{\infty}\frac{dp_2}{2\pi}\langle F(q|p_1)F(p_1|p_2)A(p_2|k)\rangle \\
&\qquad + \eta^4\int_{-\infty}^{\infty}\frac{dp_1}{2\pi}\int_{-\infty}^{\infty}\frac{dp_2}{2\pi}\int_{-\infty}^{\infty}\frac{dp_3}{2\pi}\langle F(q|p_1) \\
&\qquad \times F(p_1|p_2)F(p_2|p_3)A(p_3|k)\rangle + \cdots . \qquad \text{(A1)}
\end{aligned}
$$

To calculate the contribution to $\langle v(q|k)\rangle$ of zeroth order in η we have to calculate the contribution of this order from

each term on the right-hand side of Eq. (A1) and sum the results. To obtain the contribution of zeroth order in η from any term, it suffices to replace the average of that term by the product of the averages of the individual factors in it. In this way we obtain

$$\{v(q|k)\}_{(0)} = 2\pi\delta(q-k)[a(k) - f(k)a(k) + f^2(k)a(k) - \cdots] = 2\pi\delta(q-k)\frac{a(k)}{1+f(k)}. \tag{A2}$$

The leading contribution to $\{v(q|k)\}$ that is of nonzeroth order in η is of $O(\eta^2)$ and is obtained by summing in each term on the right-hand side of Eq. (A1) the results obtained by pairing two of the factors in all possible ways, irrespective of order, evaluating their cumulant average, and then multiplying the latter by the product of the averages of each of the remaining, unpaired, factors. In the nth-order term ($n \geqslant 2$) two types of contributions arise. The first consists of the $n-1$ terms in which one of the $F(p_i|p_j)$ is paired with $A(p_{n-1}|k)$, while the remaining $n-2$ factors of $F(p_i|p_j)$ are averaged individually; the second consists of the $(n-1)(n-2)/2$ terms in which two of the $F(p_i|p_j)$ are paired, while the remaining $n-2$ factors [including $A(p_{n-1})$] are averaged individually. The contribution to $\{v(q|k)\}$ from the first type of pairing is

$$\{v(q|k)\}_{(21)} = \eta^2[1 - f(q) + f^2(q) - \cdots] \times \int_{-\infty}^{\infty}\frac{dp}{2\pi}\{F(q|p)[1 - f(p) + f^2(p) - \cdots]A(p|k)\} = \eta^2\frac{1}{1+f(q)}\int_{-\infty}^{\infty}\frac{dp}{2\pi}\frac{\{F(q|p)A(p|k)\}}{1+f(p)}. \tag{A3}$$

The contribution from the second type of pairing is

$$\{v(q|k)\}_{(22)} = \eta^2[1 - f(q) + f^2(q) - \cdots] \times \int_{-\infty}^{\infty}\frac{dp}{2\pi}\{F(q|p)[1 - f(p) + f^2(p) - \cdots]F(p|k)\} \times [1 - f(k) + f^2(k) - \cdots]a(k) = \eta^2\frac{1}{1+f(q)}\int_{-\infty}^{\infty}\frac{dp}{2\pi}\frac{\{F(q|p)F(p|k)\}}{1+f(p)} \times \frac{a(k)}{1+f(k)}. \tag{A4}$$

From Eq. (3.17) we also see that to evaluate the cumulant average $\{V(q|p)V(p|k)\}$ appearing in the integrand on the right-hand side of Eq. (4.10b) to second order in η, we have to evaluate the cumulant average,

$$\{v(q|p)v(p|k)\} = \langle v(q|p)v(p|k)\rangle - \langle v(q|p)\rangle\langle v(p|k)\rangle, \tag{A5}$$

to the same order. To do this we multiply two iterative solutions to Eq. (3.18) and average the result term by term. The lowest-order contribution to $\langle v(q|p)v(p|k)\rangle$ as an expansion in powers of η, which is of zeroth order in η, is obtained by replacing each average of the product by the product of the averages of the individual factors. However, since what we need is not $\langle v(q|p)v(p|k)\rangle$ itself, but the difference $\langle v(q|p)v(p|k)\rangle - \langle v(q|p)\rangle\langle v(p|k)\rangle$, the zeroth-order contribution to $\langle v(q|p)v(p|k)\rangle$ can be omitted since it does not contribute to this difference. The contribution to $\langle v(q|p)v(p|k)\rangle$ of second order in η is obtained from the sum of terms in which two factors are paired and the cumulant average of their product is multiplied by the product of the averages of the remaining factors. However, the only terms of this type that contribute to the difference $\langle v(q|p)v(p|k)\rangle - \langle v(q|p)\rangle \times \langle v(p|k)\rangle$ are those in which one of the paired factors whose cumulant average is calculated comes from the expansion of $v(q|p)$ while the other comes from the expansion of $v(p|k)$.

Four classes of terms contribute to $\langle v(q|p)v(p|k)\rangle - \langle v(q|p)\rangle\langle v(p|k)\rangle$ to order η^2. They are defined by the two factors that appear in the pair whose cumulant average is evaluated. They can be written schematically as $\{AA\}$, $\{FA\}$, $\{AF\}$, and $\{FF\}$. These four categories of terms can be summed to yield the result that to $O(\eta^2)$,

$$\{v(q|p)v(p|k)\} = \eta^2\left\{\left[\frac{A(q|p)}{1+f(q)} + \frac{F(q|p)}{1+f(q)}\frac{a(p)}{1+f(p)}\right] \times \left[\frac{A(p|k)}{1+f(p)} + \frac{F(p|k)}{1+f(p)} \times \frac{a(k)}{1+f(k)}\right]\right\}. \tag{A6}$$

ACKNOWLEDGMENT

This research was supported in part by the U.S. Army Research Office grant DAAD 19-99-1-0321.

The corresponding author, A. Maradudin, can be reached by e-mail at aamaradu@uci.edu.

REFERENCES

1. S. O. Rice, "Reflection of electromagnetic waves from slightly rough surfaces," Commun. Pure Appl. Math. **4**, 351–378 (1951).
2. A. G. Voronovich, *Wave Scattering from Rough Surfaces* (Springer-Verlag, New York, 1994).
3. G. C. Brown, V. Celli, M. Coopersmith, and M. Haller, "Unitary and reciprocal expansions in the theory of light scattering from a grating," Surf. Sci. **129**, 507–515 (1983).
4. A. A. Maradudin, R. E. Luna, and E. R. Méndez, "The Brewster effect for a one-dimensional random surface," Waves Random Media **3**, 51–60 (1993).
5. C. Baylard, J.-J. Greffet, and A. A. Maradudin, "Coherent reflection factor of a random rough surface: applications," J. Opt. Soc. Am. A **10**, 2637–2647 (1993).
6. T. Kawanishi, H. Ogura, and Z. L. Wang, "Scattering of electromagnetic wave from a slightly random dielectric surface. Yoneda peak and Brewster angle in incoherent scattering," Waves Random Media **7**, 351–384 (1997).
7. T. Tamir, "The lateral wave," in *Electromagnetic Surface*

Modes, A. D. Boardman, ed. (Wiley, New York, 1982), pp. 521–548.
8. Y. Yoneda, "Anomalous surface reflection of x rays," Phys. Rev. **131**, 2010–2013 (1963).
9. A. A. Maradudin and T. A. Leskova, "X-ray scattering from a randomly rough surface," Waves Random Media **7**, 395–434 (1997).
10. C. S. West and K. A. O'Donnell, "Observations of backscattering enhancement from polaritons on a rough metal surface," J. Opt. Soc. Am. A **12**, 390–397 (1995).
11. A. A. Maradudin, A. R. McGurn, and V. Celli, "Localization effects in the scattering of light from a randomly rough grating," Phys. Rev. B **31**, 4866–4871 (1985).
12. G. C. Brown, V. Celli, M. Haller, A. A. Maradudin, and A. Marvin, "Resonant light scattering from a randomly rough surface," Phys. Rev. B **31**, 4993–5005 (1985).
13. R. Kubo, "Generalized cumulant expansion method," Phys. Soc. Jpn. **17**, 1100–1120 (1962).
14. M. Saillard and D. Maystre, "Scattering from metallic and dielectric rough surfaces," J. Opt. Soc. Am. A **7**, 982–990 (1990).
15. J.-J. Greffet, "Theoretical model of the shift of the Brewster angle on a rough surface," Opt. Lett. **17**, 238–240 (1992).
16. M. Saillard, "A characterization tool for dielectric random rough surfaces: Brewster's phenomenon," Waves Random Media **2**, 67–79 (1992).
17. V. Freilikher, I. Fuks, and M. Pustilnik, "Effect of fluctuations on the reflectivity near the angle of total internal reflection," J. Electromagn. Waves Appl. **9**, 1141–1148 (1995).
18. A. A. Maradudin, T. Michel, A. R. McGurn, and E. R. Méndez, "Enhanced backscattering of light from a random grating," Ann. Phys. (N.Y.) **203**, 255–307 (1990).
19. T. Kawanishi, I. Iwata, M. Kitano, H. Ogura, Z. L. Wang, and M. Izutsu, "Brewster's scattering angle and quasi-anomalous scattering in random scattering from dielectric interfaces," J. Opt. Soc. Am. A **16**, 339–342 (1999).
20. A. A. Maradudin and E. R. Méndez, "Enhanced backscattering of light from weakly rough, random metal surfaces," Appl. Opt. **32**, 3335–3343 (1993).
21. A. A. Maradudin, E. R. Méndez, and T. Michel, "Backscattering effects in the elastic scattering of *p*-polarized light from a large amplitude random grating," in *Scattering in Volumes and Surfaces*, M. Nieto-Vesperinas and J. C. Dainty, eds. (North-Holland, Amsterdam, 1990), pp. 157–174.
22. M. Nieto-Vesperinas and J. A. Sánchez-Gil, "Light scattering from a random rough interface with total internal reflection," J. Opt. Soc. Am. A **9**, 424–436 (1992).
23. D. De Roo and F. T. Ulaby, "Bistatic specular scattering from rough dielectric surfaces," IEEE Trans. Antennas Propag. **42**, 220–227 (1995).
24. A. Sentenac and J.-J. Greffet, "Mean-field theory of light scattering by one-dimensional rough surfaces," J. Opt. Soc. Am. A **15**, 528–532 (1998).
25. A. Sentenac, J. Toso, and M. Saillard, "Study of coherent scattering from one-dimensional rough surfaces with a mean-field theory," J. Opt. Soc. Am. A **15**, 924–931 (1998).
26. O. Calvo-Perez, A. Sentenac, and J.-J. Greffet, "Light scattering by a two-dimensional, rough penetrable medium: a mean-field theory," Radio Sci. **34**, 311–335 (1999).

Amplification of enhanced backscattering from a dye-doped polymer bounded by a rough surface

Zu-Han Gu

Surface Optics Corporation, 11555 Rancho Bernardo Road, San Diego, California 92127

G. D. Peng

School of Electrical Engineering & Telecommunications, University of New South Wales, Kensington, Sydney, NSW 2052, Australia

Received November 8, 1999

We report the experimental study of the enhanced backscattering from a random rough surface through a laser dye-doped polymer. The sample is a slice of pyrromethene-doped polymer coupled with a two-dimensional rough gold layer with a large slope. When the sample is illuminated with an *s*-polarized He–Ne laser and pumped by a cw argon-ion laser, amplified backscattering is observed. The enhanced backscattering peak increases sharply and its width narrows for a sample with low dielectric constant $|\epsilon_2|$. © 2000 Optical Society of America

OCIS codes: 290.5880, 290.1350, 290.5820, 030.0030.

One of the most interesting phenomena associated with the scattering of light from a randomly rough surface is enhanced backscattering. It is characterized by the presence of a well-defined peak in the retroreflection direction, which results primarily from the coherent interference of each multiply reflected optical path with its time-reversed partner.[1–4] Recently a theoretical study of volume and surface scattering showed that novel enhanced backscattering peaks result from an amplifying random medium and surface. An example of an amplifying medium is a laser material. The amplification is more pronounced for long paths that build up the structure of the enhanced backscattering peak and qualitatively change this backscattering. In particular, it is theoretically predicted that the peak will become sharper as the gain of the medium is increased.[5,6] However, we are aware of no previous experiments that would confirm this. In this Letter we investigate enhanced backscattering from a rough surface through a laser dye-doped polymer and report some preliminary experimental results.

Laser dyes are highly efficient gain media either for laser sources with narrow pulse widths and wide tunable ranges or for optical amplifiers with high gain and broad spectral bandwidths. Obviously, it was convenient to use the broadband gain of laser dyes in visible wavelengths for our study. Laser dyes captured in a solid matrix are much easier and safer to handle than their counterparts in liquid form. Moreover, the photostability of some organic laser dyes improves when molecules are trapped in solid matrices.[7,8]

We doped laser dyes in acrylic polymer hosts [e.g., poly(methyl methacrylate) and poly(ethyl methacrylate)] to prepare our gain samples.[9] The laser dyes were dissolved into the mixed monomers before the polymerization. The monomers used included methyl methacrylate and ethyl methacrylate. Specific amounts of initiator (lauroyl peroxide, 0.03 wt. %) and a chain transfer agent (1-butanethiol, 0.35 wt. %) were added to the dye-doped monomers, and the samples were prepared by thermal polymerization. A new class of laser dye, pyrromethene-based laser dyes, has attracted considerable interest recently because these dyes have shown the highest fluorescence quantum efficiency (typically more than 80%) and a triplet extinction coefficient that is merely one fifth that of Rhodamine 560.[10] Hence, in addition to fabricating samples doped with conventional laser dyes such as Rhodamine 6G, Rhodamine B, and fluorescein, we prepared several samples doped with pyrromethene. All the dye-doped gain polymers were polymerized by the same thermal polymerization process, with the process temperature progressively increased from 36 to 98 °C over a period of several days. The process was completed with heat treatment for several days at temperatures from 80 to 98 °C. Finally, the dye-doped polymer samples were cut and polished to various thicknesses for coupling in scattering experiments.

For dye-doped polymers the thermal, chemical, mechanical, and optical properties are of great importance. For the present experiment the most important factor is the gain that scattered light would have through a dye-doped polymer. The gain factor can be found from a simple model based on the well-known singlet–triplet energy-level system of dye molecules.[11] In this model the singlet states (the ground state S_0 and the excited states S_1, S_2, ...) are formed by the two most energetic electrons with their spin antiparallel, resulting in spin $S = 0$. The triplet states (ground state T_0 and excited states T_1, T_2, ...) are formed when the spins of the two electrons are in parallel; thus $S = 1$. They have three possible orientations ($S = -1, 0, 1$) with respect to an external magnetic field. In fact, every excited singlet state (S_1, S_2, etc.) is accompanied by a triplet state (T_1, T_2, etc.) of lower energy. Various optical processes are associated with the electron transitions between these energy levels. For example, the main absorption is normally due to the transition from S_0 to S_1, and the dominant spontaneous (fluroescence) or stimulated emission is due to the transition from S_1 to S_0. These are two processes that are directly related to the gain factor. Other relevant processes are associated with

0146-9592/00/060375-03$15.00/0 © 2000 Optical Society of America

the nonradiative intersystem transitions (e.g., S_1 to T_1 and T_1 to S_0), excited-state absorption (e.g., S_1 to S_2 or S_n, and T_1 to T_2 or T_n), and nonradiative inner-system transitions (e.g., S_n to S_1). In a gain medium the signal intensity, $I = I(\lambda, z)$, where z is the signal propagation direction, is governed by

$$\frac{\mathrm{d}I}{\mathrm{d}z} = \frac{\mathrm{d}I}{\mathrm{d}z}\bigg|_{\mathrm{stim},\,S_1\Rightarrow S_0} - \frac{\mathrm{d}I}{\mathrm{d}z}\bigg|_{\mathrm{ab},\,S_0\Rightarrow S_1} - \frac{\mathrm{d}I}{\mathrm{d}z}\bigg|_{\mathrm{ab},\,T_1\Rightarrow T_2,\,T_3\cdots}, \tag{1}$$

where the total signal intensity increases because of stimulated emission and decreases because of absorption. The signal's optical intensity from stimulated emission is well known[11] and can be expressed as

$$\frac{\mathrm{d}I}{\mathrm{d}z} = \frac{[N_1 - N_0(g_2/g_1)]\lambda^4 E(\lambda)}{8\pi\tau c n^2} I = \gamma I, \tag{2}$$

where N_0, g_1 and N_1, g_2 are the populations (usually atom densities) and degeneracies, respectively, of S_0 and S_1 states. Here c and λ are the velocity and the wavelength of light, respectively; n is the refractive index of the gain medium, τ is the fluorescence decay time for spontaneous emission, $E(\lambda)$ is the spontaneous emission line-shape function that satisfies $\int_0^\infty E(\lambda)\mathrm{d}\lambda = \phi$, where ϕ is the fluorescence quantum yield, and γ is the gain constant. In general, the pump energy E absorbed per unit volume is expressed by $E = \int_0^\infty \tau s(\nu)\alpha(\nu)\mathrm{d}\nu$, where $s(\nu)$ is the optical flux [W/cm^2/Hz], $\alpha(\nu)$ is the absorption coefficient [1/cm], ν is the frequency of light, and τ is the fluorescence decay time [s]. When a single-wavelength pump laser is used, the number of absorbed pump photons can be expressed as $N = (\tau/h\nu)s(\nu)\alpha(\nu)$. If the absorption quantum efficiency is $\eta(\nu)$, the excited molecule number per volume is $N_1 = (\tau/h\nu)\eta(\nu)s(\nu)\alpha(\nu)$. When all the pump power absorbed produces excitation of active dye molecules, e.g., when the excited-state absorption is neglected, we have $\eta(\nu) = 1$. Of course, in most practical cases, $\eta(\nu)$ is less than 1.

In practical applications, the pump wavelength must be sufficiently shorter than the signal wavelength. This ensures an ideal case for optimal gain: The pump excites the molecules efficiently, and the signal is amplified effectively. In these cases, we have simply $\Delta N = [N_1 - N_0(g_2/g_1)] \approx N_1$. Then the gain constant of the dye-doped polymer samples can be evaluated by

$$\gamma \approx \frac{\eta(\nu_p)s(\nu_p)\alpha(\nu_p)\lambda_s{}^4 E(\lambda_s)}{8\pi c n^2 h\nu_p}. \tag{3}$$

As an example, we estimate the gain constant of a Rhodamine 6G–doped polymer sample with an argon laser (514 nm) as the pump and a He–Ne laser (632.8 nm) as the signal source. Taking typical parameters of Rhodamine 6G, e.g., $\eta(\lambda) = 0.9$ and $E(\lambda) \sim 4 \times 10^4$/cm as well as sample parameters $\alpha(\nu) = 10$ dB/mm and $n = 1.48$, we have a gain constant of ~2.3 dB/cm, given a pump power of 10 W with a spot area of 0.2 mm × 0.2 mm. The gain constant defined by Eq. (2) can be related to the imaginary component of dielectric constant ϵ_2 by $I = I_0 e^{\gamma z}$ with a gain constant $\gamma = (2\pi/\lambda)(\epsilon_2/\sqrt{\epsilon_1})$. The gain in the range of a few decibels per centimeter corresponds to small ϵ_2 values.

Figure 1 is a diagram of the optical setup. Usually high pump intensity is needed. For optimal pump efficiency the pump and the scattered beams should be collinearly irradiating the gain layer and rough surface. Hence the scattering main beam is coaxial with the pump beam. Both pump and scattering beams would better be focused, with the pump beam size slightly larger than the scattering beam size. A narrow-band-pass filter is necessary because the fluorescence from laser dyes is typically broadband, and it may easily become mixed with the output signal.

The wavelength of the scattered beam must be sufficiently separated from that of the pump beam to avoid high absorption near the pump wavelength. Additionally, it must be sufficiently close to the fluorescence peak to ensure a significant optical gain. The dye-doped polymer samples have absorption peaks at 470–540 nm wavelengths and fluorescence peaks at 570–670 nm. Hence suitable pump sources for these samples are frequency-doubled YAG lasers (wavelength, 532 nm) and argon lasers (wavelength, 514 or 488 nm). Fluorescence measurement has shown that pyrromethene-doped polymers have their emission peaks near 650 nm. This suggests that pyrromethene-doped polymer would be appropriate for the current experiment, with a He–Ne laser at 633 nm as the signal (scattering) source. The detailed experimental setup is described in Ref. 3.

The main scattering source is a 35-mW, s-polarized He–Ne laser at 633 nm; the pump source is a 10-W argon-ion laser at 514 nm. The laser beams can be focused onto an area as small as 0.1 mm × 0.1 mm. We specially fabricated a sample for our experiment. The mean thickness of the designed photoresist film was ~10 μm, and that of the glass was 2 mm. In view of the need to accommodate one-dimensional or two-dimensional (2-D) scattering theories, a 2-D very random rough surface was fabricated by use of random speckle patterns.[12] Then a layer of gold was deposited onto the rough surface. The surface profile of the 2-D random rough surface was measured with a Dektak 3030 profilometer. Statistical analysis of data from 10 profiles determined that the rms roughness was 1.2 μm and the $1/e$ correlation length was 3.1 μm. Finally, a layer of dye-doped polymer was coupled with the gold rough surface.

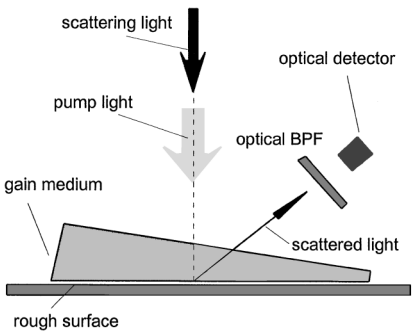

Fig. 1. Sketch of the optical setup.

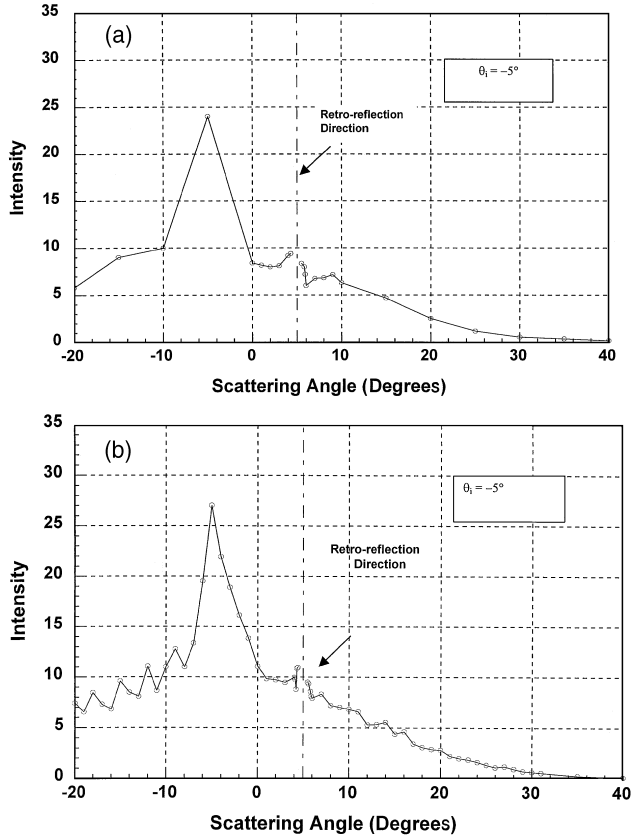

Fig. 2. Backscattering measurement of a dye-doped 2-D gold rough surface (a) without the pump beam and (b) with the pump beam (main beam at 0.6328 μm, pump beam at 0.5145 μm).

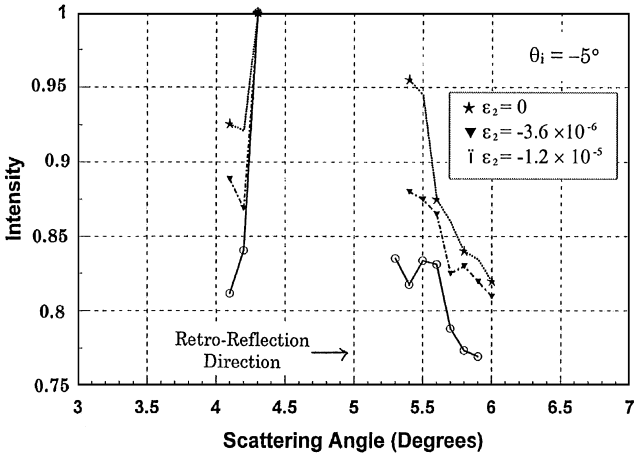

Fig. 3. Comparison of normalized enhanced backscattering peaks.

The amplified scattering for a 2-D random rough gold surface with dye-doped polymer is shown in Fig. 2. Figure 2(a) shows the far-field intensity distribution of the 2-D gold rough surface coupled with pyrromethene-doped polymer when no pump beam is used. The thickness of the polymer is ~3 mm. The *s*-polarized laser beam is applied at an incident angle of −5°. An enhanced backscattering peak at the retroreflection direction of 5° is observed. Figure 2(b) shows the far-field intensity distribution with a pump power of 2 W. Here the incident angle of the main beam remains −5°, so the enhanced backscattering peaks at 5°. Comparing Figs. 2(a) and 2(b), we find that the enhanced backscattering peak has increased from 9.8 to 11.4 and the width has narrowed from 1.9° to 1.5° for amplified scattering from the rough gold surface through the dye-doped polymer. Because of the interference of the source and the detector, the center part of the backscattering peak is missing; however, the general effect should remain the same. Figure 3 depicts the experimental results for the same 2-D gold rough surface with the dye-doped polymer, where the enhanced backscattering peaks obtained from $\epsilon_2 = 0$ (without pumping), $\epsilon_2 = -3.6 \times 10^{-6}$ (with 2-W pumping power), and $\epsilon_2 = -1.2 \times 10^{-5}$ (with 8-W pumping power) have been shifted vertically and normalized such that their maxima coincide. This figure demonstrates clearly that the enhanced backscattering peak is sharper and the width is narrower when the pump power is increased for a low dielectric constant $|\epsilon_2|$. It is also found that the accompanying background noise is the result of nonuniform distribution of the pump power. We believe that the results from an amplifying polymer bounded by a random rough surface here could be explained by the theoretical prediction[6] of the enhanced backscattering for a leaky guided-wave model with a low dielectric constant $|\epsilon_2|$. The enhanced backscattering peak is proportional to $|\Delta|^{-1}$, and its width is proportional to $|\Delta|$, where $|\Delta|$ is the smallest decay/amplification rate of the guided waves supported by the dye-doped polymer in the presence of the surface roughness. $|\Delta|$ decreases with increasing dielectric constant $|\epsilon_2|$. That is why, when dielectric constant $|\epsilon_2|$ increases, the enhanced backscattering peak is sharper and its width is narrower.

The authors are grateful for useful discussions with Zong-Qi Lin. We express our gratitude to the U.S. Army Research Office for its support under contracts DAAG55-98-C-0034 and DAAD19-99-1-0321.

References

1. V. Celli, A. A. Maradudin, A. M. Marvin, and A. R. McGurn, J. Opt. Soc. Am. A **2,** 2225 (1985).
2. E. R. Mendez and K. A. O'Donnell, Opt. Commun. **61,** 91 (1987).
3. Z. H. Gu, J. Q. Lu, A. A. Maradudin, A. Martinez, and E. R. Mendez, Appl. Opt. **32,** 2852 (1993).
4. Z. H. Gu, M. Josse, and M. Ciftan, Opt. Eng. **35,** 370 (1996).
5. D. S. Wiersma, M. P. van Albada, and A. Lagendijk, Phys. Rev. Lett. **75,** 1739 (1995).
6. A. V. Tutov, A. A. Maradudin, T. A. Leskova, A. P. Mayev, and J. A. Sanchez-Gil, Phys. Rev. B **60,** 12692 (1999).
7. A. Dubois, M. Canva, A. Brun, F. Chaput, and J.-P. Boilot, Appl. Opt. **35,** 3193 (1996).
8. F. J. Duarte, Appl. Opt. **33,** 3857 (1994).
9. G. D. Peng, P. L. Chu, Z. Xiong, T. Whitbread, and R. P. Chaplin, J. Lightwave Technol. **14,** 2215 (1996).
10. R. E. Hermes, T. H. Allik, S. Chandra, and J. A. Hutchinson, Appl. Phys. Lett. **63,** 877 (1993).
11. A. Yariv, *Quantum Electronics* (Wiley, New York, 1975), Chap. 8.
12. P. F. Gray, Opt. Acta **25,** 765 (1978).

Light scattering from an amplifying medium bounded by a randomly rough surface: A numerical study

Ingve Simonsen,[1,2] Tamara A. Leskova,[3] and Alexei A. Maradudin[2]
[1]*Department of Physics, The Norwegian University of Science and Technology, Trondheim, Norway*
[2]*Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, California 92697*
[3]*Institute of Spectroscopy, Russian Academy of Sciences, Troitsk 142092, Russia*
(Received 29 August 2000; published 29 June 2001)

By numerical simulations we study the scattering of *s*-polarized light from a rough dielectric film deposited on the planar surface of a semi-infinite perfect conductor. The dielectric film is allowed to be either active or passive, situations that we model by assigning negative and positive values, respectively, to the imaginary part ε_2 of the dielectric constant of the film. We study the reflectance $\mathcal{R}$ and the total scattered energy $\mathcal{U}$ for the system as functions of both ε_2 and the angle of incidence of the light. Furthermore, the positions and widths of the enhanced backscattering and satellite peaks are discussed. It is found that these peaks become narrower and higher when the amplification of the system is increased, and that their widths are linear functions of ε_2. The positions of the backscattering peaks are found to be independent of ε_2, while we find a weak dependence on this quantity in the positions of the satellite peaks.

DOI: 10.1103/PhysRevB.64.035425 PACS number(s): 42.25.Dd, 42.25.Bs

I. INTRODUCTION

In the first half of the 1990s and subsequently, amplifying volume disordered media received a great deal of attention from theorists[1,2] and experimentalists[2,3] alike. This attention was partly motivated by the suggestion of using random volume scattering media to construct a so-called random laser.[4] For scattering systems possessing surface disorder in contrast to volume disorder, the overwhelming majority of theoretical and experimental studies were devoted to scattering from passive (i.e., absorbing) media. Only recently has the surface scattering community begun study surface-disordered amplifying systems. The only literature on the scattering of light from amplifying surface-disordered media known to us is the theoretical study by Tutov *et al.*[5] and the experimental investigation by Gu and Peng.[6] In the theoretical work by Tutov *et al.*,[5] the authors conducted a perturbative study of the scattering of *s*-polarized light from an amplifying film deposited on the planar surface of a perfect conductor, where the vacuum-film interface was a one-dimensional random interface characterized by a Gaussian power spectrum. In this work we consider the same scattering system, but apply a numerical simulation approach for its study. The numerical approach is based on the solution of the reduced Rayleigh equation that the scattering amplitude for the system satisfies. The use of a numerical simulation approach enables us to study possible nonperturbative effects[7] that could not be accounted for by the perturbative technique used in Ref. 5. Furthermore, we also use a different power spectrum of the surface roughness. In particular, a West-O'Donnell (or rectangular) power spectrum[8] is used in this work, in contrast to the Gaussian power spectrum used by Tutov *et al.* Such a power spectrum allows for the suppression of single scattering over a range of scattering angles and, more importantly, it opens the possibility for a strong coupling of the incident light to guided waves supported by the film structure.

In this work we calculate the reflectivity of a surface-random vacuum-dielectric-metal film geometry, illuminated from above by *s*-polarized light of frequency ω. The metal-dielectric interface is assumed to be flat, while the vacuum-dielectric interface is described by the surface profile function $\zeta(x_1)$. The amplifying medium (i.e., the film) is modeled by a dielectric medium whose dielectric constant ε has an imaginary part ε_2 that is negative, while its real part ε_1 is positive. The values of ε_2 are chosen so that they include gains $[g=2\pi|\varepsilon_2|/(\lambda\sqrt{\varepsilon_1})]$ in the medium that are physically realizable. The assumption of a negative imaginary part to ε is the simplest way of modeling stimulated emission in this system. The reflectivity is given by $|R(k)|^2$, where $R(k)$ is defined in terms of the scattering amplitude $R(q|k)$ by $|\langle R(q|k)\rangle|^2 = L_1 2\pi\delta(q-k)|R(k)|^2$. In this relation the wave numbers k and q are related to the angles of incidence and scattering by $k=(\omega/c)\sin\theta_0$ and $q=(\omega/c)\sin\theta_s$, respectively, L_1 is the length of the x_1 axis covered by the random surface, and the angular brackets denote an average over the ensemble of realizations of the surface profile function $\zeta(x_1)$. The scattering amplitude $R(q|k)$ is obtained by numerically solving the reduced Rayleigh equation it satisfies for a large number of realizations of $\zeta(x_1)$, and $\langle R(q|k)\rangle$ is obtained by averaging the results. As expected, the reflectivity of the amplifying medium with a random surface is larger than that of the corresponding absorbing medium, viz. a medium with the same value of $|\varepsilon_2|$ but with ε_2 positive, for all angles of incidence.

II. SCATTERING THEORY

A. Scattering system

The scattering system that will be considered in this paper consists of a dielectric film, with a randomly rough top interface, deposited on the planar surface of a semi-infinite perfect conductor. In particular, it consists of a vacuum in

 ©2001 The American Physical Society

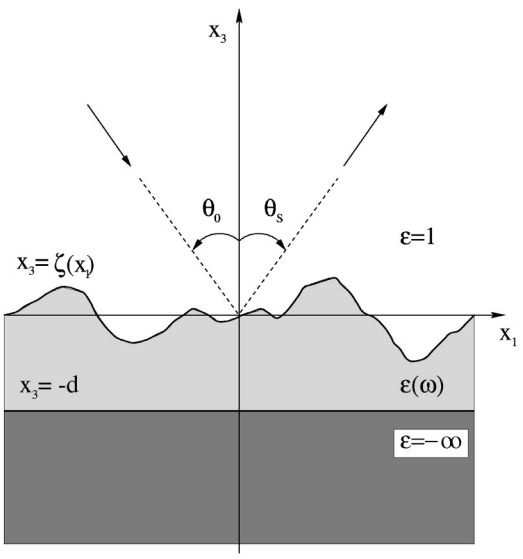

FIG. 1. The scattering geometry considered in the present work.

the region $x_3 > \zeta(x_1)$, an amplifying or absorbing dielectric medium in the region $-d < x_3 < \zeta(x_1)$, and a perfect conductor in the region $x_3 < -d$. This geometry is depicted in Fig. 1. The rough surface profile function, denoted by $\zeta(x_1)$, is assumed to be a single-valued function of its argument, and is differentiable as many times as needed. Furthermore, it is assumed to constitute a zero-mean, stationary, Gaussian random process defined by

$$\langle \zeta(x_1) \rangle = 0, \tag{2.1a}$$

$$\langle \zeta(x_1)\zeta(x_1') \rangle = \delta^2 W(|x_1 - x_1'|), \tag{2.1b}$$

where $\langle \cdot \rangle$ denotes an average over the ensemble of realizations of $\zeta(x_1)$, and δ is the rms height of the rough surface. Moreover, $W(|x_1|)$ denotes the surface height autocorrelation function, and is related to the power spectrum of the surface roughness $g(|k|)$ by

$$g(|k|) = \int_{-\infty}^{\infty} dx_1 W(|x_1|) \; e^{-ikx_1}. \tag{2.2}$$

In the numerical simulation results to be presented later, we will assume a rectangular power spectrum, also known as the West-O'Donnell form,

$$g(|k|) = \frac{\pi}{k_+ - k_-}[\theta(k - k_-)\theta(k_+ - k) + \theta(-k_- - k)\theta(k + k_+)], \tag{2.3}$$

where $\theta(k)$ is the Heaviside unit step function, and $k_\pm$ are parameters to be specified. This power spectrum was recently used in an experimental study of enhanced backscattering from weakly rough surfaces.[8]

B. Scattering equations

If the vacuum-dielectric interface $x_3 = \zeta(x_1)$ is illuminated from the vacuum side by an s-polarized electromagnetic wave of frequency ω, the only nonzero component of the electric field vector in the region $x_3 > \zeta(x_1)_{max}$ is the sum of an incident wave and a scattered field:

$$E_2^>(x_1, x_3|\omega) = e^{ikx_1 - i\alpha_0(k,\omega)x_3} + \int_{-\infty}^{\infty} \frac{dq}{2\pi} R(q|k) e^{iqx_1 + i\alpha_0(q,\omega)x_3}. \tag{2.4}$$

In this equation $R(q|k)$ denotes the scattering amplitude, while we have defined

$$\alpha_0(q,\omega) = \begin{cases} \sqrt{\dfrac{\omega^2}{c^2} - q^2}, & |q| < \omega/c \\ i\sqrt{q^2 - \dfrac{\omega^2}{c^2}}, & |q| > \omega/c. \end{cases} \tag{2.5}$$

From a knowledge of the scattering amplitude one can define the differential reflection coefficient (DRC) $\partial R/\partial \theta_s$. It is defined such that $(\partial R/\partial \theta_s) d\theta_s$ is the fraction of the total time-averaged flux incident on the surface that is scattered into the angular interval $d\theta_s$ about the scattering angle θ_s, in the limit as $d\theta_s \to 0$. The contribution to the mean differential reflection coefficient from the coherent (specular) component of the scattered field is given by[9,10]

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle_{\text{coh}} = \frac{1}{L_1} \frac{\omega}{2\pi c} \frac{\cos^2\theta_s}{\cos\theta_0} |\langle R(q|k) \rangle|^2, \tag{2.6a}$$

and the contribution to the mean differential reflection coefficient from the incoherent (diffuse) component of the scattered field is given by[9,10]

$$\left\langle \frac{\partial R}{\partial \theta_s} \right\rangle_{\text{incoh}} = \frac{1}{L_1} \frac{\omega}{2\pi c} \frac{\cos^2\theta_s}{\cos\theta_0} [\langle |R(q|k)|^2 \rangle - |\langle R(q|k) \rangle|^2]. \tag{2.6b}$$

In Eqs. (2.6), L_1 is the length of the x_1 axis covered by the random surface, and the wave numbers k and q are related to the angles of incidence θ_0 and the angle of scattering θ_s according to

$$k = \frac{\omega}{c} \sin\theta_0, \quad q = \frac{\omega}{c} \sin\theta_s. \tag{2.7}$$

Both these angles are measured from the normal to the mean surface, as indicated in Fig. 1.

From the definition of the mean differential reflection coefficient, we find that the reflectance of the surface is defined according to

$$\mathcal{R} = \int_{-\pi/2}^{\pi/2} d\theta_s \left\langle \frac{\partial R}{\partial \theta_s} \right\rangle_{\text{coh}} = |R(k)|^2, \tag{2.8}$$

where k is given by Eq. (2.7), and $R(k)$ is related to the scattering amplitude $|R(q|k)|^2$ by $|\langle R(q|k) \rangle|^2 = L_1 2\pi\delta(q - k)|R(k)|^2$. Likewise, the total scattered energy (normalized to the incident energy) is defined by

$$\mathcal{U}=\int_{-\pi/2}^{\pi/2} d\theta_s \left\langle \frac{\partial R}{\partial \theta_s} \right\rangle, \tag{2.9}$$

where $\langle \partial R/\partial \theta_s \rangle$ is the total mean DRC, i.e., the sum of the coherent and incoherent contribution as defined in Eqs. (2.6a) and (2.6b), respectively.

So far we have not specified how to obtain the scattering amplitude entering into the above equations. It has previously been shown that $R(q|k)$ is the solution of the so-called reduced Rayleigh equation.[11] This single, inhomogeneous integral equation for $R(q|k)$ for our scattering geometry reads[10]

$$\int_{-\infty}^{\infty} \frac{dq}{2\pi} M(p|q)R(q|k)=N(p|k), \tag{2.10a}$$

where

$$M(p|q)=\frac{e^{i\alpha(p,\omega)d}}{\alpha_0(q,\omega)+\alpha(p,\omega)} I[\alpha_0(q,\omega)+\alpha(p,\omega)|p-q] - \frac{e^{-i\alpha(p,\omega)d}}{\alpha_0(q,\omega)-\alpha(p,\omega)} I[\alpha_0(q,\omega) - \alpha(p,\omega)|p-q], \tag{2.10b}$$

$$N(p|k)=-\frac{e^{i\alpha(p,\omega)d}}{\alpha(p,\omega)-\alpha_0(k,\omega)} I[\alpha(p,\omega)-\alpha_0(k,\omega)|p-k] - \frac{e^{-i\alpha(p,\omega)d}}{\alpha(p,\omega)+\alpha_0(k,\omega)} I[-\alpha(p,\omega) - \alpha_0(k,\omega)|p-k], \tag{2.10c}$$

with

$$I(\gamma|q)=\int_{-\infty}^{\infty} dx_1 e^{i\gamma\zeta(x_1)} e^{-iqx_1}. \tag{2.10d}$$

In writing Eq. (2.10), we have introduced

$$\alpha(q,\omega)=\sqrt{\varepsilon(\omega)\frac{\omega^2}{c^2}-q^2}, \tag{2.11}$$

where the branch of the square root is chosen so that the real part of $\alpha(q,\omega)$ is always positive, while the imaginary part is positive when $\varepsilon_2>0$, but is negative when $\varepsilon_2<0$.

The simulation results to be presented in Sec. III were obtained by directly solving numerically the reduced Rayleigh equation (2.10). This approach can treat much longer rough surfaces as compared to a rigorous numerical simulation approach[9] with the same use of computer power and memory. An additional advantage of a numerical approach based on the reduced Rayleigh equation is that $\mathcal{R}$ and $\mathcal{U}$ can be calculated to high precision, whereas the same quantities calculated by a rigorous approach have been found to be less accurate for the surface lengths typically used in such simulations. We believe this difference in accuracy for $\mathcal{R}$ and $\mathcal{U}$ for these two numerical approaches is related to the difference in the length of the surface that can be handled practically with today's typical computer resources. The numerical solution of the reduced Rayleigh equation is done by converting the integral equation into a set of linear equations obtained by using an appropriate quadrature scheme and solving the resulting system by standard numerical techniques.[12] Due to the increased numerical performance, the calculation of the $I(\gamma|q)$ integrals was based on an expansion of the integrand in powers of the surface profile function. This numerical method was recently applied successfully to a similar scattering geometry[7], and the interested reader is directed to this paper for details of the numerical method.

III. RESULTS AND DISCUSSIONS

For the numerical simulations to be presented below, we have considered the scattering of *s*-polarized incident light of wavelength $\lambda=632.8$ nm. The film was assumed to have a mean thickness $d=500$ nm, and its dielectric constant at the wavelength of the incident light was taken to be $\varepsilon(\omega)=2.6896+i\varepsilon_2$, where ε_2 is allowed to vary over both positive and negative values. The surface profile function was characterized by a power spectrum of the West-O'Donnell type as defined in Eq. (2.3). For the parameters defining the power spectrum, we used $k_-=0.86\,\omega/c$ and $k_+=1.97\,\omega/c$. For these values of $k_\pm$, single scattering should be suppressed for scattering angles in the range $|\theta_s|<55.1°$. The rms height of the surface was taken to be $\delta=30$ nm. Furthermore, the length of the surface was taken to be $L=160\lambda$, and the numerical results were all averaged over $N_\zeta=3000$ realizations of the surface profile function.

In Fig. 2(a) we present the numerical simulation results for the contribution to the mean differential reflection coefficient from the light that has been scattered incoherently, $\langle \partial R/\partial \theta_s \rangle_{\text{incoh}}$, for *s*-polarized light incident normally on the mean surface ($\theta_0=0°$). The values of the imaginary part of the dielectric were (from top to bottom) $\varepsilon_2=-0.0025$, 0, and 0.0025. From this figure we note the enhanced backscattering peaks located at $\theta_s=\theta_0=0°$. Moreover, two satellite peaks, located symmetrically about the position of the enhanced backscattering peak, are easily distinguished from the background. Their positions, as read from Fig. 2(a), fit nicely with their positions, $\theta_\pm=\pm 17.7°$, calculated for the corresponding planar geometry in the limit of vanishing ε_2[7]. The choices made for ε_2 of the film in Fig. 2(a) correspond to an amplifying or active film ($\varepsilon_2=-0.0025$), a neither amplifying nor absorbing film ($\varepsilon_2=0$), and an absorbing or passive film ($\varepsilon_2=0.0025$), respectively. This is reflected in Fig. 2(a), where the contribution to the mean DRC from the light scattered incoherently from the amplifying medium is larger for all scattering angles than for the other two cases, due to the extra energy gained by the scattered light from the amplifying film. Moreover, it is interesting to note that the differences between these curves are largest for small scattering angles, and as one moves to larger scattering angles these differences are reduced. The main reason for this is that for scattering angles $|\theta_s|<55.1°$ the light undergoes multiple

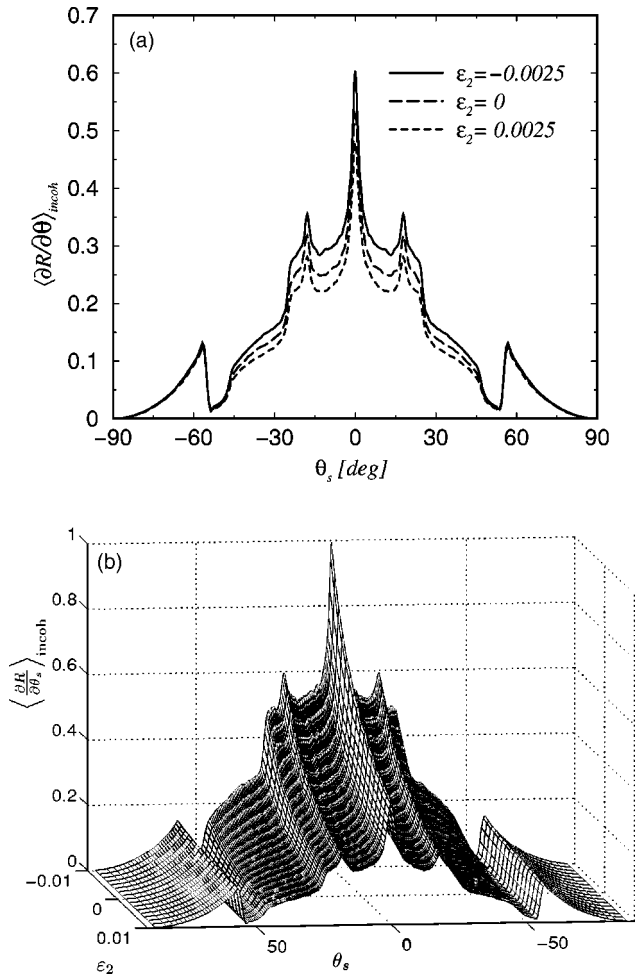

FIG. 2. The mean differential reflection coefficient for the incoherently scattered light for (a) $\varepsilon_2 = 0, \pm 0.0025$, and (b) as a function of the same parameter. For both figures the angle of incidence was $\theta_0 = 0°$, and the wavelength of the incident light was $\lambda = 632.8$ nm. The dielectric constant of a film of mean thickness $d = 500$ nm was $\varepsilon(\omega) = 2.6896 + i\varepsilon_2$, where ε_2 is as indicated in the figure. The randomly rough surface had a rms roughness of $\delta = 30$ nm. The power spectrum was of the West-O'Donnell type defined by the parameters $k_- = 0.86\,\omega/c$ and $k_+ = 1.97\,\omega/c$.

scattering, which will increase the effects of amplification and absorption as compared to a system that is dominated by single-scattering events for such scattering angles. In the wings of the angular dependence of the mean DRC, $|\theta_s| > 55.1°$, where single scattering gives the main contribution, the differences between the curves corresponding to different values of ε_2 is much less pronounced. In order to obtain a more complete picture of how the incoherent component of the mean DRC depends on the imaginary part of the dielectric function, in Fig. 2(b) we present a plot showing $\langle \partial R/\partial\theta_s \rangle_{\rm incoh}$ as a function of ε_2, as well as of the scattering angle θ_s. The angle of incidence here was also chosen to be $\theta_0 = 0°$. As seen from this plot, the positions of the peaks are fixed, or close to fixed, while the overall amplitude of $\langle \partial R/\partial\theta_s \rangle_{\rm incoh}$ increases monotonically with the decreasing imaginary part of the dielectric constant.

To quantify better how the amplification or absorption depends on ε_2, we have studied the total energy scattered by the surface as well as its reflectance. These two quantities, denoted by $\mathcal{U}(\theta_0, \varepsilon_2)$ and $\mathcal{R}(\theta_0, \varepsilon_2)$ respectively, are related to the mean differential reflection coefficient by Eqs. (2.8) and (2.9). The numerical results for these two quantities for normal incidence are given in Figs. 3(a) and 3(b). For small values of ε_2 we find that these quantities are linear in ε_2. However, when the absolute value of the imaginary part of the dielectric constant increases, a deviation from this behavior is observed. The numerical data in both cases are well fitted by a cubic polynomial in ε_2. In Fig. 3(c) we present the numerical results for $\mathcal{U}$ and $\mathcal{R}$ as a function of the angle of incidence θ_0 for $\varepsilon_2 = \pm 0.0025$. It is seen that $\mathcal{U}(\theta_0)$ is a monotonically increasing or decreasing function of the angle of incidence for positive and negative values of ε_2, respectively, and the two curves for $\varepsilon_2 = \pm 0.0025$ are symmetric with respect to the line $\mathcal{U}(\theta_0) = 1$. For negative (positive) ε_2 the total scattered energy is larger (smaller) than unity. However, from the same graph it is observed that $\mathcal{R}(\theta_0)$ is not a monotonic function of ε_2. Instead it has a minimum in the vicinity of 25°. Below this value it is decreasing, while above it is increasing. The reason for this behavior is due to the excitation of a leaky guided wave supported by the scattering geometry.[5] The minimum in $\mathcal{R}(\theta_0)$ occurs for an angle of incidence corresponding to the wave number of the leaky guided wave, and the excitation of this mode will take away scattered energy from the specular direction, resulting in a minimum in $\mathcal{R}(\theta_0, \varepsilon_2)$ for this angle of incidence.

From Fig. 2(a) it can be observed that the widths of both the backscattering and satellite peaks, in contrast to their positions, are sensitive to the value of the imaginary part of the dielectric function. Since the scattering geometry supports (at least) two true guided modes, the widths of these peaks are expected to grow with $\varepsilon_2(\omega)$[5]. This is much more apparent if we shift, but not scale, the tops of the enhanced backscattering peaks to the same height. We have done so by plotting $\langle \partial R/\partial\theta_s \rangle_{\rm incoh} - \langle \partial R/\partial\theta_s \rangle_{\rm incoh}|_{\theta_s = \theta_o}$ as a function of the scattering angle θ_s for various values of ε_2, and the results are shown in Fig. 4 for the backscattering peaks [Fig. 4(a)] and the satellite peaks [Fig. 4(b)]. Figure 4(a) clearly shows that the width of the enhanced backscattering peak increases as the imaginary part of the dielectric constant increases. Or, in other words, the enhanced backscattering peak becomes narrower and taller when the amplification of the medium is increased. This behavior is in qualitative agreement with the experimental results reported recently by Gu and Peng.[6] This finding can theoretically be understood as follows: It can be shown that the enhanced backscattering peak should have a Lorentzian form of total width[5]

$$\Delta_T(\omega) = \Delta_\epsilon(\omega) + \Delta_{\rm sc}(\omega), \tag{3.1}$$

where $\Delta_\epsilon(\omega)$ is the contribution to the width from the attenuation or amplification of the guided waves, while $\Delta_{\rm sc}(\omega)$ is the broadening due to the scattering of such waves by the surface roughness. Moreover, it can be shown that[5]

$$\Delta_\epsilon(\omega) \propto \varepsilon_2. \tag{3.2}$$

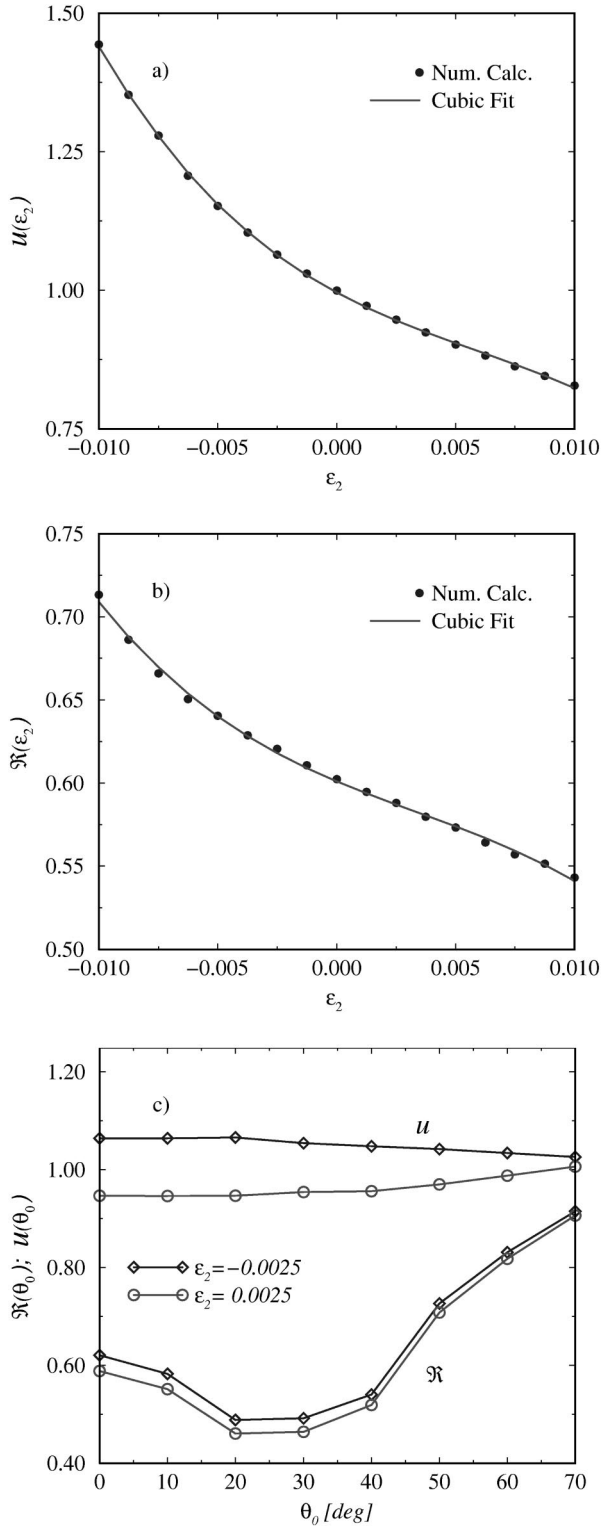

FIG. 3. The total scattered energy $\mathcal{U}$ [Eq. (2.9)] and the reflectance $\mathcal{R}$ [Eq. (2.8)] as functions of the imaginary part of the dielectric constant [(a) and (b)] and of the angle of incidence (c). In (a) and (b) the angle of incidence was $\theta_0 = 0°$, while in (c) the imaginary part of the dielectric constant was $\varepsilon_2 = \pm 0.0025$. The remaining parameters are as given in Fig. 2.

Depending on the geometrical and dielectric parameters of the film, the total width $\Delta_T(\omega)$ can be dominated by either $\Delta_\epsilon(\omega)$ or $\Delta_{sc}(\omega)$. For the parameters considered in this study, however, it is expected[5] that $\Delta_{sc}(\omega) > |\Delta_\epsilon(\omega)| > 0$. Therefore, the width should increase with increasing values of the imaginary part of the dielectric constant.

We will now examine how the full width $W(\varepsilon_2)$ of the backscattering peak depends on ε_2. In Fig. 5 we present $W(\varepsilon_2)$ vs ε_2, as obtained from the numerical simulation results shown in Fig. 2. The width of an enhanced backscatter-

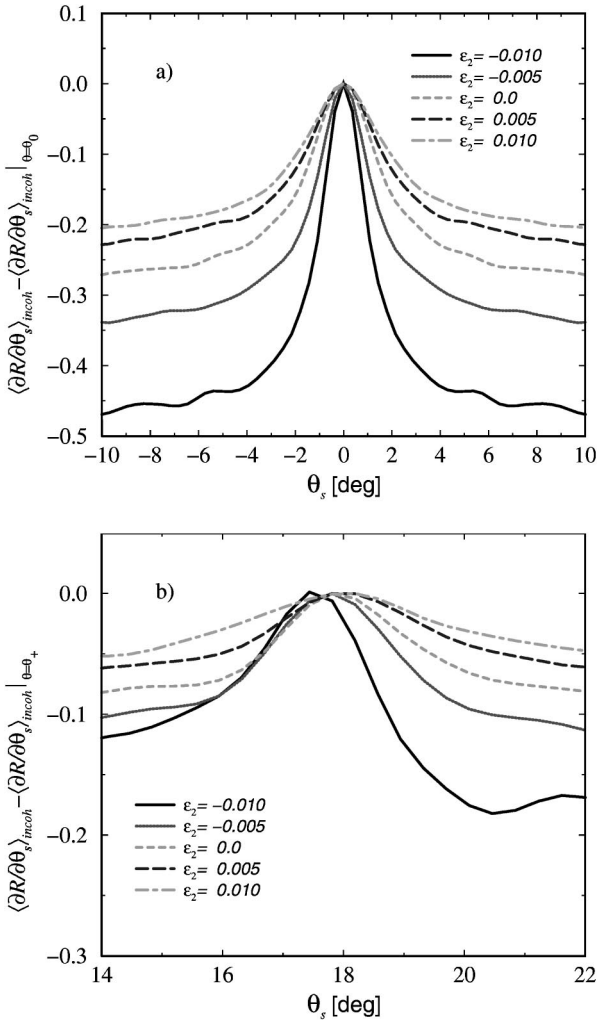

FIG. 4. Shifted plot for the mean DRC around the enhanced backscattering peak (a) and satellite peaks (b). The quantities that are plotted is $\langle \partial R/\partial\theta_s \rangle_{\text{incoh}} - \langle \partial R/\partial\theta_s \rangle_{\text{incoh}}|_{\theta_s=\theta_o}$ for the backscattering peaks and $\langle \partial R/\partial\theta_s \rangle_{\text{incoh}} - \langle \partial R/\partial\theta_s \rangle_{\text{incoh}}|_{\theta_s=\theta_+}$ for the satellite peaks, where θ_+ is the (positive) angular position of the satellite peaks.

ing peak was defined as its full width at half maximum above the background at the position of the peak. Here the background was defined to be located at the minimum value of $\langle \partial R/\partial\theta_s \rangle_{incoh}$ between the backscattering and satellite peaks. Even though the data in Fig. 5 are somewhat noisy, a linear dependence (solid curve) on ε_2, as predicted by Eq. (3.2), is easily seen.

In Fig. 4(b) we present the same kind of plot as in Fig. 4(a), but now for a satellite peak. One sees the that the width of the satellite peaks increases with increasing ε_2, the same behavior found for the enhanced backscattering peak. However, more interesting is the apparent change in the position of the satellite peaks with the value of the imaginary part of the dielectric constant. To the precision of the numerical calculations, the positions of the satellite peaks for an absorbing film ($\varepsilon_2>0$) seem to shift to larger scattering angles (in absolute value) as compared to the position of the satellite

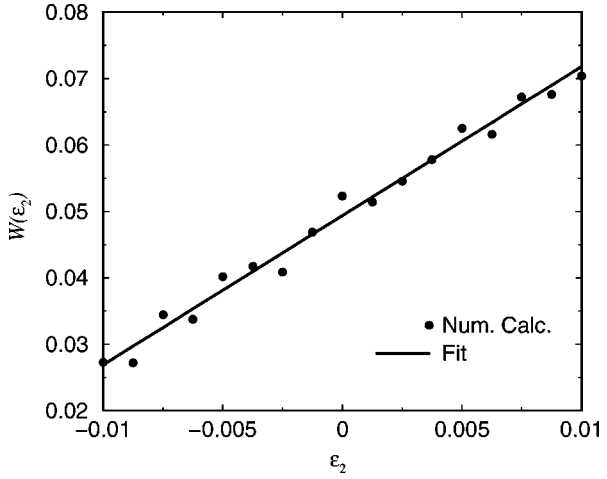

FIG. 5. The full width $W(\varepsilon_2)$ (filled dots) at half maximum above the background at its position of the backscattering peak as a function of the imaginary part ε_2 of the dielectric constant of the film as obtained from the numerical simulation results of Fig. 2. The solid line represents a linear fit in ε_2 to the numerical data.

peaks when $\varepsilon_2=0$. The opposite seems to hold true for an amplifying film ($\varepsilon_2<0$). There are two reasons for this behavior of the satellite peaks. First, the real part of the self-energy has a linear in ε_2 contribution, thus, in the presence of surface roughness the values of the wave numbers of the guided waves acquire a contribution linear in ε_2. The second reason is the strong dependence of the background intensity on the values of ε_2. The increase of the background intensity also shifts the visual positions of the satellite peaks to smaller scattering angles. For the widths of the satellite peaks, the quality of the numerical data, unfortunately, did not allow us to obtain reliable results.

IV. CONCLUSIONS

By numerical simulations we have studied light scattered from an absorbing or amplifying dielectric film deposited on the planar surface of a semi-infinite perfect conductor where the vacuum-dielectric interface is randomly rough. It has been shown that the reflectance $\mathcal{R}(\theta_0,\varepsilon_2)$, as well as the total scattered energy $\mathcal{U}(\theta_0,\varepsilon_2)$, are decreasing functions of the imaginary part of the dielectric function of the film for a fixed angle of incidence. Furthermore, it has been demonstrated that $\mathcal{U}(\theta_0,\varepsilon_2)$ is a monotonically increasing or decreasing function of the angle of incidence for fixed positive and fixed negative values of the imaginary part of the dielectric function, respectively. However, for the reflectance we find that $\mathcal{R}(\theta_0,\varepsilon_2)$ first decreases to a minimum near $\theta_0=25°$ and then increases. This minimum is a result of the leaky guided wave supported by the scattering structure. Moreover, for an amplifying surface both $\mathcal{R}(\theta_0,\varepsilon_2)$ and $\mathcal{U}(\theta_0,\varepsilon_2)$ are smaller then their absorbing equivalents for all angles of incidence.

The width of the enhanced backscattering peaks, as well as the satellite peaks supported by the scattering system, are

found to increase with increasing ε_2. While the location of the enhanced scattering peaks seems to be unaffected by the value of the imaginary part of the dielectric constant of the film, the corresponding positions for the satellite peaks are found to be shifted toward larger (smaller) scattering angles for positive (negative) values of the imaginary part of the dielectric function, respectively. Finally, it is found that the width of the enhanced backscattering peak is a linear function in ε_2.

ACKNOWLEDGMENTS

I. S. would like to thank the Research Council of Norway (Contract No. 32690/213) and Norsk Hydro ASA for financial support. The research of I. S., A. A. M., and T. A. L. was supported in part by Army Research Office Grant No. DAAG 55-98-C-0034. This work has also received support from the Research Council of Norway (Program for Supercomputing) through a grant of computer time.

[1] V. Freilikher, M. Pustilnik, and I. Yurkevich, Phys. Rev. B **56**, 5974 (1997).
[2] D. S. Wiersma and A. Lagendijk, Phys. World **10**, 33 (1997).
[3] D. S. Wiersma, M. P. van Albada, and A. Lagendijk, Phys. Rev. Lett. **75**, 1739 (1995).
[4] M. Kempe, G. A. Berger, and A. Z. Genack, in *Handbook of Optical Properties*, edited by R. E. Hummer and P. Wissman (CRC Press, Boca Raton, FL, 1997), Vol. II, pp. 301–330.
[5] A. V. Tutov, A. A. Maradudin, T. A. Leskova, A. P. Mayer, and J. A. Sánchez-Gil, Phys. Rev. B **60**, 12 692 (1999).
[6] Zu-Han Gu and G. D. Peng, Opt. Lett. **25**, 375 (2000).
[7] I. Simonsen and A. A. Maradudin, Opt. Commun. **162**, 99 (1999).
[8] C. S. West and K. A. O'Donnell, J. Opt. Soc. Am. A **12**, 390 (1995).
[9] A. A. Maradudin, T. Michel, A. R. McGurn, and E. R. Méndez, Ann. Phys. (N.Y.) **203**, 255 (1990).
[10] V. Freilikher, E. Kanzieper, and A. A. Maradudin, Phys. Rep. **288**, 127 (1997).
[11] F. Toigo, A. Marvin, V. Celli, and N. R. Hill, Phys. Rev. B **15**, 5618 (1977).
[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes*, 2nd ed. (Cambridge University Press, Cambridge, 1992).

15 September 2000

ELSEVIER

Optics Communications 183 (2000) 529–545

OPTICS COMMUNICATIONS

www.elsevier.com/locate/optcom

The surface enhanced second harmonic generation of light from a randomly rough metal surface in the Kretschmann geometry

T.A. Leskova [a], M. Leyva-Lucero [b], E.R. Méndez [c], A.A. Maradudin [d,*], I.V. Novikov [d,1]

[a] *Institute of Spectroscopy, Russian Academy of Sciences, Troitsk 142092, Russia*

[b] *Escuela de Ciencias Físico-Matematicas, Universidad Autonoma de Sinaloa, Ciudad Universitaria, C.P. 80000 Culiacan, Sinaloa, Mexico*

[c] *División de Física Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, Apartado Postal 2732, Ensenada, Baja California, Mexico*

[d] *Department of Physics and Astronomy, and Institute for Surface and Interface Science, University of California, Irvine, California, 92697, USA*

Received 11 April 2000; received in revised form 17 July 2000; accepted 18 July 2000

Abstract

We present results of perturbative calculations of the second harmonic light generated in the transmission of *p*-polarized light through a thin metal film with a one-dimensional random surface in the Kretschmann attenuated total reflection (ATR) geometry. The metal film is deposited on the planar surface of a prism through which the light is incident. The back surface of the film is a one-dimensional random surface whose generators are perpendicular to the plane of incidence. It is in contact either with a semi-infinite vacuum or with a semi-infinite nonlinear crystal (quartz). It is shown that when the random surface separates the metal film from vacuum so that the nonlinearity of the film surfaces gives rise to the harmonic light, for a general angle of incidence a dip appears in the angular dependence of the intensity of the transmitted harmonic light in the direction normal to the mean surface. When the second harmonic generation is due to the nonlinearity of the crystal in contact with the metal film, a peak in the angular dependence of the intensity of the transmitted harmonic light occurs in this direction. These dips and peaks are multiple-scattering effects. However, when the angle of incidence is optimal for the excitation of surface plasmon polaritons at the film-vacuum/nonlinear crystal interface the nonlinear mixing of the incident light and the backward propagating surface plasmon polariton leads to an intense peak in the angular dependence of the intensity of the transmitted harmonic light in the direction normal to the mean surface. This peak is already present in the single-scattering approximation. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Experimental and theoretical studies of second harmonic generation (SHG) of light in reflection from a metal surface go back at least three decades, to the first experimental observation [1] and the first theoretical

* Corresponding author. E-mail: aamaradu@uci.edu

[1] Present address: Stanford Research Systems, 1290 ReamWood Ave., Sunnyvale, California, 94089, USA.

0030-4018/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0030-4018(00)00897-X

description of the phenomenon in [2,3]. At this early stage of the studies [2,3] it was established that this phenomenon is a surface effect. It is therefore very sensitive to anything happening on the surface. In the last several years interest in second harmonic generation from a rough metal surface has arisen due to the growing interest in interference effects occurring in the multiple scattering of electromagnetic waves from randomly rough metal surfaces and the related backscattering enhancement phenomenon [4]. It has been expected that the nonlinear optical interactions at a randomly rough metal surface should also produce new features owing to interference effects in the multiple scattering of electromagnetic waves. The results of a perturbative calculation carried out by McGurn et al. [5] predicted that enhanced second harmonic generation of light at a weakly rough, clean, metal surface occurs not only in the retroreflection direction but also in the direction normal to the mean scattering surface. Interference effects in the multiple scattering of surface plasmon polaritons of the fundamental frequency, excited by the incident light through the roughness of the surface, are responsible for the appearance of the peak in the direction normal to the mean surface, while those occurring in the multiple scattering of surface plasmon polaritons at the harmonic frequency are responsible for the appearance of the peak in the retroreflection direction.

This work stimulated several subsequent experimental studies of second-harmonic generation in the scattering of light from random metal surfaces [6–11], and enhanced second harmonic generation peaks in the direction normal to the mean surface and in the retroreflection direction were observed [6–9,11]. In these experiments, however, the scattering system was not a clean random interface between vacuum and a semi-infinite metal. To amplify the second harmonic signal the Kretschmann attenuated total reflection (ATR) geometry [12] was used. Therefore, the scattering system was the random interface with a dielectric or vacuum of a thin metal film deposited on the planar base of a dielectric prism through which the light was incident. In the experiments of Refs. [6,7,9] the scattering system was the random interface between a silver film and a nonlinear quartz crystal, so that the nonlinear interaction occurred in the quartz crystal rather than at the significantly more weakly nonlinear silver surfaces. A well-defined peak of the second harmonic generation in the direction normal to the mean interface was observed in transmission in [6]. When the experiment was carried out with long-range surface plasmon polaritons [13], peaks of the enhanced second harmonic generation were detected both in the retroreflection direction [7,9] and in the direction normal to the surface in transmission [9]. In Refs. [8,10,11] attempts to detect the peaks of the enhanced second harmonic generation at a silver film-vacuum interface were made. A well-defined peak in the direction normal to the mean surface was observed in transmission in [8,11], while only a broad depolarized background, but no peak in the direction normal to the mean surface, was observed in transmission in Ref. [10]. The peak of the intensity of the generated light in the direction normal to the mean surface observed in [6,8,9,11] was interpreted as due to the coherent nonlinear mixing of multiply-scattered contrapropagating surface plasmon polaritons. As is well known [14], the symmetry of the surface nonlinear polarization of a clean metal surface forbids second harmonic generation due to contrapropagating beams of surface plasmon polaritons. In [8,11] it was argued that the surface roughness breaks the symmetry and, as a result, makes the second harmonic generation by contrapropagating beams of surface plasmon polaritons possible. We believe that this explanation is incorrect, and present what we believe to be the correct explanation in the present work.

The main reason for using the Kretschmann ATR geometry in the experiments [6–11] was to excite surface plasmon polaritons associated with the film-vacuum (film-nonlinear crystal) interface through the ATR phenomenon, and thus to enhance the field at the metal film-vacuum (or metal film-nonlinear crystal) interface. For many years this experimental configuration was used to enhance the second harmonic generation in reflection from a metal surface [15].

If, however, the metal surface is rough the surface plasmon polaritons can be excited through the surface roughness without a prism coupler. The nonlinear interaction of the excited surface plasmon polaritons with the incident light also leads to a strong enhancement of the second harmonic generation [5,16,17]. Such processes lead to strong peaks, up to several orders of magnitude above the background, in the angular distribution of the intensity of light of frequency 2ω in the directions determined by the conditions $q = k \pm k_{sp}(\omega)$, where

$q = (2\omega/c)\sin\theta_s$, $k = (\omega/c)\sin\theta_0$, θ_s and θ_0 are the angles of scattering and incidence, and $k_{sp}(\omega)$ is the wavenumber of the surface plasmon polaritons of frequency ω that are excited through the surface roughness. These peaks arise already in single-scattering processes, and their positions depend on the angle of incidence.

In the Kretschmann geometry the angle of incidence is tuned so that the surface polaritons on the rough metal film-vacuum interface or the rough metal film-nonlinear crystal interface are excited efficiently. This means that $k = k_{sp}(\omega)$, and the single-scattering peaks move to $q = 0$, i.e. to the direction normal to the mean surface, and to $q = 2k_{sp}(\omega)$, i.e. to the nonradiative region. Due to the single-scattering origin of the peak in the direction normal to the mean surface, it could be expected that its intensity is much higher than that of the peak predicted in [5], which is associated with multiple-scattering processes. As a result, the weak features associated with the coherent effects in the multiple scattering of surface plasmon polaritons are masked by the strong single-scattering contribution. The presence of roughness leads also to the excitation of the surface plasmon polaritons associated with the prism–metal film interface. As a result, two additional strong peaks appear in the angular distribution of the intensity of the generated light. Their positions are determined by the conditions $q = k \pm k_{sp}^{(p)}(\omega)$, where $k_{sp}^{(p)}(\omega)$ is the wavenumber of the surface plasmon polaritons associated with the prism–metal film interface.

In this paper we present results of perturbative calculations of the second harmonic generation of light in transmission in the Kretschmann ATR geometry when the interface of a thin metal film with vacuum or a nonlinear crystal is weakly rough. The emphasis will be on the existence of the peak of the enhanced second harmonic generation in the direction normal to the mean surface, which is due to the interference of multiply-scattered surface plasmon polaritons of the frequency of the incident light. The results of rigorous numerical simulations, applicable to more strongly corrugated surfaces, will be presented elsewhere.

2. Formulation of the scattering problem

The physical system we consider consists of a dielectric prism characterized by a positive dielectric constant ϵ_0 in the region $x_3 > D$, a metal film characterized by an isotropic, complex, frequency-dependent dielectric function $\epsilon_1(\omega) = \epsilon'(\omega) + i\epsilon''(\omega)$ in the region $D > x_3 > \zeta(x_1)$ and, generally, a nonlinear crystal characterized by a linear dielectric function $\epsilon_2(\omega)$ in the region $x_3 < \zeta(x_1)$ (Fig. 1).

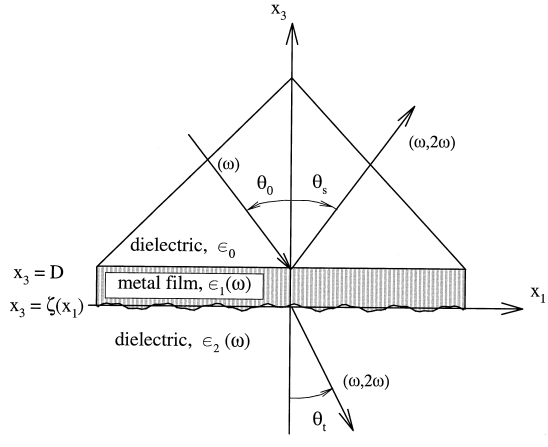

Fig. 1. The system studied in this paper.

The surface profile function $\zeta(x_1)$ is assumed to be a single-valued function of x_1 that is differentiable as many times as is necessary, and to constitute a zero-mean, stationary, Gaussian random process defined by

$$\langle \zeta(x_1)\rangle = 0, \tag{2.1a}$$

$$\langle \zeta(x_1)\zeta(x_1')\rangle = \delta^2 W(|x_1 - x_1'|). \tag{2.1b}$$

The angle brackets in Eqs. (2.1) denote an average over the ensemble of realizations of the surface profile function, and $\delta = \langle \zeta^2(x_1)\rangle^{1/2}$ is the rms height of the surface.

We also introduce the Fourier integral representation of the surface profile function,

$$\zeta(x_1) = \int_{-\infty}^{\infty} \frac{dQ}{2\pi} \hat{\zeta}(Q) e^{iQx_1}. \tag{2.2}$$

The Fourier coefficient $\hat{\zeta}(Q)$ is a zero-mean Gaussian random process defined by

$$\langle \hat{\zeta}(Q)\rangle = 0, \tag{2.3a}$$

$$\langle \hat{\zeta}(Q)\hat{\zeta}(Q')\rangle = 2\pi\delta(Q+Q')\delta^2 g(|Q|), \tag{2.3b}$$

where $g(|Q|)$, the power spectrum of the surface roughness, is given by

$$g(|Q|) = \int_{-\infty}^{\infty} dx_1 W(|x_1|) e^{-iQx_1}. \tag{2.4}$$

In this paper we will present results calculated for a random surface characterized by a Gaussian power spectrum given by

$$g(|Q|) = \sqrt{\pi}\, a e^{-Q^2a^2/4}. \tag{2.5}$$

In our treatment of second harmonic generation we neglect the influence of the nonlinearity on the fundamental fields (undepleted pump approximation). We treat the nonlinearity of the metal film through effective nonlinear boundary conditions, while the bulk nonlinearity of the crystal is taken into account through the solution of the Maxwell equations. Since we study the Kretchsmann ATR geometry, and the surface roughness is one-dimensional, only *p*-polarized incident waves are of interest to us. As is known, the nonlinear polarization of a metal surface has only *p*-polarized components independent of whether the incident light is *p*- or *s*-polarized. Therefore, in what follows we assume that a *p*-polarized plane wave of frequency ω is incident from the prism onto the planar prism–metal interface. The angle of incidence, measured counterclockwise from the normal to the mean surface is θ_0, and the plane of incidence is the x_1x_3-plane. We first solve the linear scattering problem, and with its solution in hand we determine the surface and bulk nonlinear polarizations at the harmonic frequency 2ω. For the sake of simplicity we assume that the orientation of the crystal is chosen so that the bulk nonlinear polarization in it has a nonzero tangential component [18]

$$P_1^{(\mathrm{NL})}(x_1, x_3|2\omega) = d_{11}\left(E_1^2(x_1, x_3|\omega) - E_3^2(x_1, x_3|\omega)\right), \tag{2.6}$$

where d_{11} is the second order nonlinear susceptibility. In solving the scattering problem for the harmonic fields we will use the nonlinear boundary conditions [19], whose general form is

$$\begin{aligned} &H_2^{(f)}(x_1, x_3|2\omega)|_{x_3=D,\zeta(x_1)} - H_2^{(p,v)}(x_1, x_3|2\omega)|_{x_3=D,\zeta(x_1)} \\ &\quad = \frac{2ic}{\omega}\mu_3^{(p,v)}\frac{1}{\phi^2(x_1)}\frac{1}{\epsilon_1(\omega)}\frac{\partial}{\partial N}H_2^{(f)}(x_1, x_3|\omega)|_{x_3=D,\zeta(x_1)}\frac{d}{dx_1}H_2^{(f)}(x_1, x_3|\omega)|_{x_3=D,\zeta(x_1)} \\ &\quad \equiv A^{(p,v)}(x_1|2\omega), \end{aligned} \tag{2.7a}$$

$$\frac{1}{\epsilon_1(2\omega)}\left[\frac{\partial}{\partial N}H_2^{(f)}(x_1,x_3|2\omega)\right]\Bigg|_{x_3=D,\zeta(x_1)} - \frac{1}{\epsilon_{0,2}(2\omega)}\left[\frac{\partial}{\partial N}H_2^{(p,\upsilon)}(x_1,x_3|2\omega)\right]\Bigg|_{x_3=D,\zeta(x_1)}$$
$$= \frac{2ic}{\omega}\frac{d}{dx_1}\left\{\frac{1}{\phi^2(x_1)}\left[\mu_1^{(p,\upsilon)}\left(\frac{d}{dx_1}H_2^{(f)}(x_1,x_3|\omega)\Big|_{x_3=D,\zeta(x_1)}\right)^2\right.\right.$$
$$\left.\left.+\frac{\mu_2^{(p,\upsilon)}}{\epsilon_1^2(\omega)}\left(\frac{\partial}{\partial N}H_2^{(f)}(x_1,x_3|\omega)\Big|_{x_3=D,\zeta(x_1)}\right)^2\right]\right\} \equiv B^{(p,\upsilon)}(x_1|2\omega), \tag{2.7b}$$

where $\partial/\partial N=(-\zeta'(x_1)\partial/\partial x_1,0,\partial/\partial x_3)$ is the unnormalized derivative along the normal to the surface, $\phi(x_1)=[1+(\zeta'(x_1))^2]^{1/2}$ and the indices p, f, and υ denote the fields in the prism, metal film, and vacuum, respectively. In Eqs. (2.7) $\mu_{1,2,3}^{(p,\upsilon)}$ are phenomenological nonlinear coefficients at the prism–film (p) and film-vacuum (υ) interfaces, and are to be determined from experimental data or by using some particular model for the surface nonlinear polarization.

We assume that the metal film-crystal interface is weakly rough and satisfies the condition for the validity of the Rayleigh hypothesis, $|\zeta'(x_1)|\ll 1$ [20–22]. In this case we can seek the x_2-component of the magnetic fields of frequencies ω and 2ω in the form

$$H_2^{(p)}(x_1,x_3|\Omega)=H_0e^{ikx_1}e^{-i\alpha_0(k,\Omega)(x_3-D)}+\int_{-\infty}^{\infty}\frac{dq}{2\pi}R(q,\Omega)e^{iqx_1}e^{i\alpha_0(q,\Omega)(x_3-D)} \tag{2.8a}$$

in the prism,

$$H_2^{(f)}(x_1,x_3|\Omega)=\int_{-\infty}^{\infty}\frac{dq}{2\pi}e^{iqx_1}\left[T_1(q,\Omega)e^{i\alpha_1(q,\Omega)x_3}+T_2(q,\Omega)e^{-i\alpha_1(q,\Omega)x_3}\right] \tag{2.8b}$$

in the metal film, and

$$H_2^{(\upsilon)}(x_1,x_3|\Omega)=\int_{-\infty}^{\infty}\frac{dq}{2\pi}e^{iqx_1}\left[T(q,\Omega)e^{-i\alpha_2(q,\Omega)x_3}+\mathscr{P}(q,\Omega|x_3)\right] \tag{2.8c}$$

in the crystal. In Eqs. (2.8) Ω stands for either ω or 2ω, $R(q,\Omega)$ and $T(q,\Omega)$ are the scattering and transmission amplitudes, respectively, $\alpha_0(q,\Omega)=\sqrt{\epsilon_0(\Omega^2/c^2)-q^2}$, $\mathrm{Re}(\alpha_0(q,\Omega))>0$, $\mathrm{Im}(\alpha_0(q,\Omega))>0$, $\alpha_1(q,\Omega)=\sqrt{\epsilon_1(\Omega)(\Omega^2/c^2)-q^2}$, $\mathrm{Re}(\alpha_1(q,\Omega))>0$, $\mathrm{Im}(\alpha_1(q,\Omega))>0$, and $\alpha_2(q,\Omega)=\sqrt{\epsilon_2(\Omega)(\Omega^2/c^2)-q^2}$, $\mathrm{Re}(\alpha_2(q,\Omega))>0$, $\mathrm{Im}(\alpha_2(q,\Omega)))>0$.

When we consider fields of the fundamental frequency ω, H_0 is the amplitude of the incident field, $k=\sqrt{\epsilon_0}(\omega/c)\sin\theta_0$ and $\alpha_0(k,\omega)=\sqrt{\epsilon_0}(\omega/c)\cos\theta_0$ are the tangential and normal components of the wavector of the incident light, and $\mathscr{P}(q,\omega|x_3)=0$. For the second harmonic fields $H_0=0$ and $\mathscr{P}(q,2\omega|x_3)$ describes the particular solution of the Maxwell equations in the nonlinear crystal:

$$\mathscr{P}(q,2\omega|x_3,)=-\frac{2cd_{11}}{\omega\epsilon_2^2(\omega)}\int_{-\infty}^{\infty}\frac{dp}{2\pi}e^{-i(\alpha_2(p,\omega)+\alpha_2(q-p,\omega))x_3}T(q-p,\omega)T(p,\omega)F(q,p), \tag{2.9}$$

where

$$F(q,p)=\frac{(\alpha_2(p,\omega)+\alpha_2(q-p,\omega))(p(q-p)-\alpha_2(p,\omega)\alpha_2(q-p,\omega))}{(\alpha_2(p,\omega)+\alpha_2(q-p,\omega))^2-\alpha_2^2(q,2\omega)}. \tag{2.10}$$

Since in the experiments [6,8–11] it was the transmitted signal that was measured, we solve the scattering problem for the transmission amplitude $T(q,\Omega)$. Using the boundary conditions for the tangential component of the magnetic field and the tangential component of the electric field across the interfaces at the fundamental and

harmonic frequencies, we can obtain a single integral equation for the transmission amplitude $T(q,\Omega)$ that is analogous to the reduced Rayleigh equation of the scattering problem,

$$T(q,\Omega) = G_0(q,\Omega)\int_{-\infty}^{\infty}\frac{dp}{2\pi}V(q|p,\Omega)T(p,\Omega) + G_0(q,\Omega)Q(q,\Omega), \tag{2.11}$$

where the scattering potential $V(q|k,\Omega)$ is

$$V(q|k,\Omega) = i\frac{\epsilon_1(\Omega)-\epsilon_2(\Omega)}{\epsilon_2(\Omega)\epsilon_1(\Omega)}\left[m(q|k,\Omega) + r_{01}(q,\Omega)^{2i\alpha_{1(q,\Omega)D}}n(q|k,\Omega)\right], \tag{2.12}$$

with

$$m(q|k,\Omega) = \frac{qk+\alpha_1(q,\Omega)\alpha_2(k,\Omega)}{\alpha_1(q,\Omega)-\alpha_2(k,\Omega)}J(-(\alpha_1(q,\Omega)-\alpha_2(k,\Omega))|q-k), \tag{2.13a}$$

$$n(q|k,\Omega) = \frac{qk-\alpha_1(q,\Omega)\alpha_2(k,\Omega)}{\alpha_1(q,\Omega)+\alpha_2(k,\Omega)}J(\alpha_1(q,\Omega)+\alpha_2(k,\Omega)|q-k), \tag{2.13b}$$

$$J(\gamma|Q) = \int_{-\infty}^{\infty}dx_1 e^{-iQx_1}(e^{-i\gamma\zeta(x_1)}-1), \tag{2.13c}$$

and where $G_0(q,\Omega)$ is the Green's function associated with a three-layer system with planar interfaces

$$G_0(k,\Omega) = \frac{i}{D(k,\Omega)}, \tag{2.14a}$$

with

$$D(k) = \left(\frac{\alpha_1(k,\Omega)}{\epsilon_1(\Omega)}+\frac{\alpha_2(k,\Omega)}{\epsilon_2(\Omega)}\right)\left[1+r_{01}(k)r_{12}(k)e^{2i\alpha_1(k,\Omega)D}\right], \tag{2.14b}$$

and

$$r_{ij}(q,\Omega) = \frac{(\alpha_i(q,\Omega)\epsilon_j(\Omega)-\alpha_j(q,\Omega)\epsilon_i(\Omega))}{(\alpha_i(q,\Omega)\epsilon_j(\Omega)+\alpha_j(q,\Omega)\epsilon_i(\Omega))}, \quad i,j=0,1,2. \tag{2.15}$$

The driving term $Q(q,\Omega)$ in (2.11) has different forms for the fundamental and harmonic fields. For the fundamental field it is related to the incident field,

$$Q(q,\omega) = -2i\frac{\alpha_1(k,\omega)}{\epsilon_1(\omega)}t_{01}(k,\omega)e^{i\alpha_1(k,\omega)D}H_0 2\pi\delta(q-k), \tag{2.16}$$

where

$$t_{ij}(k,\Omega) = \frac{2\alpha_i(k,\Omega)\epsilon_j(\Omega)}{\alpha_i(k,\Omega)\epsilon_j(\Omega)+\alpha_j(k,\Omega)\epsilon_i(\Omega)}, \quad i,j=0,1,2, \tag{2.17}$$

so that $G_0(q,\omega)Q(q,w) = T_0(k,\omega)2\pi\delta(q-k)$, where $T_0(k,\omega)$ is the Fresnel transmission coefficient for the layer system

$$T_0(k,\omega) = \frac{t_{12}(k,\omega)t_{01}(k,\omega)e^{i\alpha_1(k,\omega)D}}{1+r_{12}(k,\omega)r_{01}(k,\omega)e^{2i\alpha_1(k,\omega)D}}. \tag{2.18}$$

For the second harmonic field the nonlinear driving source $Q(q,2\omega)$ has three contributions,

$$Q(q,2\omega) = Q^{(p)}(q,2\omega) + Q^{(v)}(q,2\omega) + Q^{(b)}(q,2\omega), \tag{2.19}$$

where $Q^{(p)}(q,2\omega)$ is associated with the prism–metal film interface

$$Q^{(p)}(q,2\omega) = it_{10}(q,2\omega)\, e^{i\alpha_1(q,2\omega)D}\left(\frac{\alpha_0(q,2\omega)}{\epsilon_0}A^{(p)}(q) + B^{(p)}(q)\right), \tag{2.20a}$$

$Q^{(v)}(q,2\omega)$ is associated with the metal film-crystal interface

$$\begin{aligned} Q^{(v)}(q,2\omega) = i\int_{-\infty}^{\infty}\frac{dp}{2\pi}\Bigg\{\frac{1}{\epsilon_1(2\omega)}\left[\alpha_1(q,2\omega) + \frac{q(q-p)}{\alpha_1(q,2\omega)}\right]A^{(v)}(p)\big[I(-\alpha_1(q,2\omega)|q-p) \\ + r_{01}(q,2\omega)\, e^{2i\alpha_1(q,2\omega)D} I(\alpha_1(q,2\omega)|q-p)\big] + iB^{(v)}(p)\big[I(-\alpha_1(q,2\omega)|q-p) \\ - r_{01}(q,2\omega)\, e^{2i\alpha_1(q,2\omega)D} I(\alpha_1(q,2\omega)|q-p)\big]\Bigg\}, \end{aligned} \tag{2.20b}$$

and $Q^{(b)}(q,2\omega)$ arises from the nonlinearity of the crystal

$$\begin{aligned} Q^{(b)}(q,2\omega) = -i\int_{-\infty}^{\infty}\frac{dp}{2\pi}\Bigg\{\frac{1}{\epsilon_1(2\omega)}\left(\alpha_1(q,2\omega) + \frac{q(q-p)}{\alpha_1(q,2\omega)}\right)\mathscr{P}(p)\big[I(-\alpha_1(q,2\omega)|q-p) \\ + r_{01}(q,2\omega)\, I(\alpha_1(q,2\omega)|q-p)\big] + \frac{1}{\epsilon_2(2\omega)}\mathscr{P}_1(p)\big[I(-\alpha_1(q,2\omega)|q-p) \\ - r_{01}(q,2\omega)\, I(\alpha_1(q,2\omega)|q-p)\big]\Bigg\}, \end{aligned} \tag{2.21}$$

where $\mathscr{P}(p)$ and $\mathscr{P}_1(p)$ are the Fourier components of the driving term in the nonlinear medium and its normal derivative evaluated at the rough interface, and are given by

$$\mathscr{P}(p) = -\frac{2cd_{11}}{\omega\epsilon_2^2(\omega)}\int_{-\infty}^{\infty}\frac{dr}{2\pi}\int_{-\infty}^{\infty}\frac{dr'}{2\pi} I(\alpha_1(r,\omega) + \alpha_1(r-r',\omega)|p-r)F(r,r')T(r-r',\omega)T(r',\omega), \tag{2.22}$$

and

$$\begin{aligned} \mathscr{P}_1(p) = i\frac{2cd_{11}}{\omega\epsilon_2^2(\omega)}\int_{-\infty}^{\infty}\frac{dr}{2\pi}\int_{-\infty}^{\infty}\frac{dr'}{2\pi} I(\alpha_1(r,\omega) + \alpha_1(r-r',\omega)|p-r)\Bigg[\alpha_1(r,\omega) + \alpha_1(r-r',\omega) \\ - \frac{r(p-r)}{\alpha_1(r,\omega) + \alpha_1(r-r',\omega)}\Bigg]F(r,r')T(r-r',\omega)T(r',\omega). \end{aligned} \tag{2.23}$$

The solution of the reduced Rayleigh equation (2.11) and explicit expressions for the nonlinear driving term are presented in Appendices A and B, respectively.

The quantity we are interested in is the intensity of the second harmonic light, which we define as the total power of the harmonic light generated diffusely, normalized by the square of the power of the incident light and multiplied by the illuminated area S:

$$p_{2\omega} = \frac{P_{2\omega}}{(P_{\text{in}})^2}S, \tag{2.24}$$

where $S = L_1 L_2$, and L_1 and L_2 are the lengths of the surface along the x_1- and x_2-axes. If we define the angle of transmission θ_t explicitly by setting $q = (\sqrt{\epsilon_2(2\omega)}\, 2\omega/c)\sin\theta_t$, so that $\alpha_2(q,2\omega) = (\sqrt{\epsilon_2(2\omega)}\, 2\omega/c)\cos\theta_t$, the mean normalized intensity of second harmonic light takes the form

$$\langle I(\theta_t|2\omega)\rangle_{\text{incoh}} = \frac{8\omega\epsilon_0}{L_1 c^2 |H_0|^4} \frac{\cos^2\theta_t}{\cos^2\theta_0} \left[\langle |T(q,2\omega)|^2\rangle - |\langle T(q,2\omega)\rangle|^2\right]. \tag{2.25}$$

3. Results and discussion

In Fig. 2(a), (b), and (c) we present the mean intensity of the second harmonic light, calculated by the use of the small-amplitude perturbation approach, when *p*-polarized light is incident through a prism on a one-dimen-

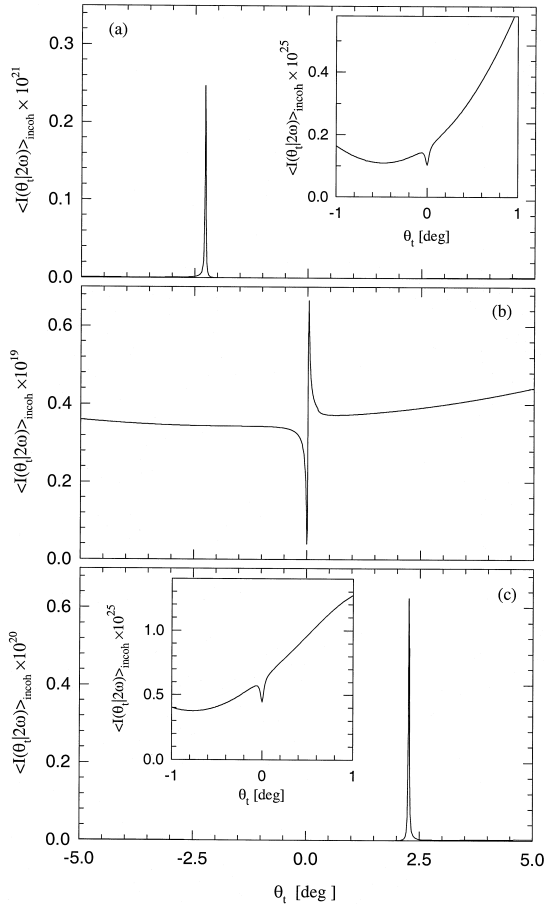

Fig. 2. The mean differential intensity of the second harmonic light as a function of the angle of transmission θ_t for the scattering of *p*-polarized light through a randomly rough silver film-vacuum interface. The interface roughness is characterized by the Gaussian power spectrum Eq. (2.5) with $\delta = 5$ nm and $a = 113.5$ nm. The nonlinear coefficients are given by the free-electron model. The angles of incidence are (a) $\theta_0 = 22°$, (b) $\theta_0 = 24°$, and (c) $\theta_0 = 26°$. The optimal angle for excitation of surface plasmon polaritons at the film-vacuum interface is 24°.

sional, random silver film-vacuum interface, characterized by the Gaussian power spectrum (2.5) with an rms height $\delta = 5$ nm and a transverse correlation length $a = 113.5$ nm. The mean thickness of the film is $D = 55$ nm, and the angles of incidence are $\theta_0 = 22°$ (a), $\theta_0 = 24°$ (b), and $\theta_0 = 26°$ (c). The refractive index of the prism is $n_0 = 2.479$, so that the optimal angle for the excitation of surface plasmon polaritons is $\theta_0 = 24°$. The nonlinear coefficients $\mu_{1,2,3}^{(p,v)}$ were calculated on the basis of the free-electron model [19].

In Fig. 3(a), (b), and (c) we present the mean intensity of the second harmonic light, calculated by the perturbative approach, when *p*-polarized light is incident through a prism on a one-dimensional, random silver film-quartz interface, characterized by the Gaussian power spectrum (2.5) with an rms height $\delta = 5$ nm and a transverse correlation length $a = 113.5$ nm. The mean thickness of the film is $D = 55$ nm, and the angles of incidence are $\theta_0 = 37.2°$ (a), $\theta_0 = 39.2°$ (b), and $\theta_0 = 41.2°$ (c). The refractive indices of the quartz are $n_2(\omega) = \sqrt{\epsilon_2(\omega)} = 1.536$ and $n_2(2\omega) = \sqrt{\epsilon_2(2\omega)} = 1.542$, and the optimal angle for the excitation of surface plasmon polaritons is $\theta_0 = 39.2°$.

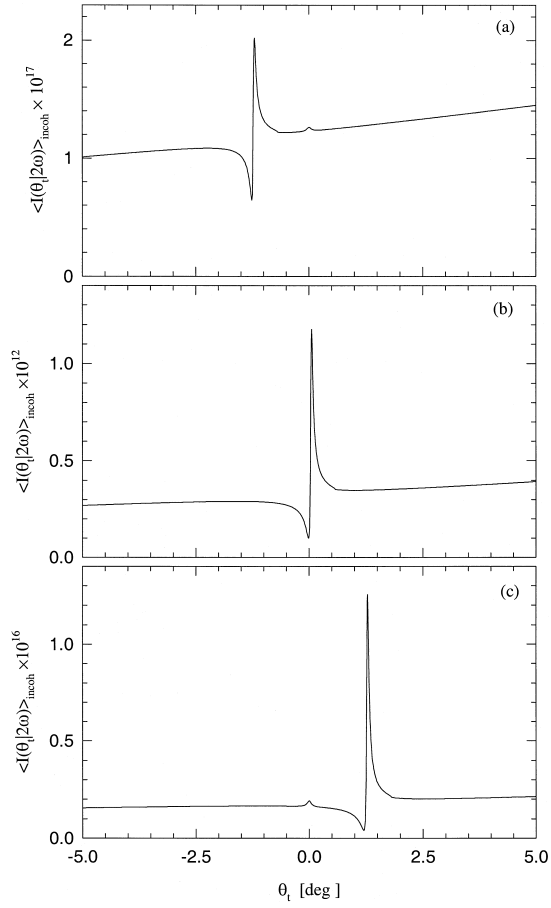

Fig. 3. The mean differential intensity of the second harmonic light as a function of the angle of transmission θ_t for the scattering of *p*-polarized light from a randomly rough silver film-nonlinear quartz interface, whose roughness is characterized by the Gaussian power spectrum Eq. (2.5) with $\delta = 5$ nm and $a = 113.5$ nm. The nonlinear susceptibility of the crystal is $d_{11} = 0.93 \times 10^{-9}$ esu. The angles of incidence are (a) $\theta_0 = 37.2°$, (b) $\theta_0 = 39.2°$, and (c) $\theta_0 = 41.2°$. The optimal angle for excitation of surface plasmon polaritons at the film-quartz interface is 39.2°.

To analyze the main features of the angular dependence of the intensity of the generated light we need explicit expressions for the different contributions to the mean differential intensity of the harmonic light. The effects of the multiple scattering of surface plasmon polaritons of frequency ω, which are of interest to us here, are contained in the the nonlinear driving term $Q(q,\Omega)$ on the right hand side of Eq. (2.11) evaluated at $\Omega = 2\omega$. The integral term in Eq. (A.13) of Appendix A describes the multiple scattering of the waves of frequency 2ω and, in particular, contains the effects of the multiple scattering of surface plasmon polaritons of frequency 2ω. Since these effects do not include the peak in the direction normal to the mean surface, which is of primary interest to us here, we omit their contribution to the mean differential intensity $\langle I(\theta_t|2\omega)\rangle_{\rm incoh}$. Since all three contributions to $Q(q,2\omega)$, Eq. (2.19), contain the product of the transmission amplitudes $T(p,\omega)T(p',\omega)$, by using the expression for the transmission amplitude $T(q,\omega) = \mathscr{T}_0(k,\omega)H_0[2\pi\delta(q-k) + G(q,\omega)t(q|k)]$ (see, Eq. (A.9)) we can subdivide each term in $Q(q,2\omega)$ into three contributions with different physical meanings. The first contributions to $Q^{(p,v,b)}(q,2\omega)$, which contain only the product of δ-functions, $(2\pi)^2\delta(p-k)\delta(p'-k)$, describe the nonlinear mixing of the fields of frequency ω which would be specular if the film-crystal interface were planar. The parts of $Q^{(p,v,b)}(q,2\omega)$ that contain the product of a δ-function and a term with the Green's function, $2\pi\delta(p-k)G(p',2\omega)t(p'|k) + 2\pi\delta(p'-k)G(p,2\omega)t(p|k)$, describe the interaction of the 'specular' and scattered fields, including the nonlinear mixing of the excited surface plasmon polaritons with the incident light. Finally, the parts of $Q^{(p,v,b)}(q,2\omega)$ which contain the product of two Green's functions, $G(p,2\omega)t(p|k)G(p',2\omega)t(p'|k)$, describe the nonlinear mixing of the scattered fields, and include the mixing of co- and contrapropagating surface plasmon polaritons. According to the classification we have just given we separate explicitly these three contributions to the nonlinear driving source, and include all the roughness induced corrections to them into the additional source function $\mathscr{E}_{\rm sc}(q,2\omega)$, so that we obtain

$$Q(q,2\omega) = \frac{2ic}{\omega}\mathscr{T}_0^2(k,\omega)H_0^2\Bigg[\Gamma_0(k)2\pi\delta(q-2k) + \Gamma_1(q,k)G(q-k)t(q-k|k) + \int_{-\infty}^{\infty}\frac{dp}{4\pi}\Gamma_2(q,p)G(q-p,\omega)G(p,\omega)t(q-p|k)t(p|k) + \mathscr{E}_{\rm sc}(q,2\omega)\Bigg], \tag{3.1}$$

where the nonlinear coefficients Γ_i are given by $\Gamma_i = \gamma_i^{(p)} + \gamma_i^{(v)} + \gamma_i^{(b)}$, and the explicit expressions for $\gamma_i^{(p,v,b)}$ are presented in Appendix B. The most resonant contributions to the intensity of the harmonic light have the form

$$\langle I(\theta_s|2\omega)\rangle_{\rm res} = \frac{512\omega^3}{\epsilon_0 c^4}\left|t_{10}(k,\omega)e^{2i\alpha_1(k,\omega)D}\right|^2\cos^2\theta_s\cos^2\theta_0|G(q,2\omega)|^2|G(k,\omega)|^4| \times\Bigg[|\Gamma_1(q,k)|^2|G(q-k)|^2\tau(q-k|k) + \int_{-\infty}^{\infty}\frac{dp}{2\pi}|\Gamma_2(q,p)|^2|G(q-p,\omega)G(p,\omega)|^2\tau(q-p|k)\tau(p|k)\Bigg], \tag{3.2}$$

where we have defined $\langle t(q|k)t^*(p|k)\rangle = 2\pi\delta(q-p)\tau(q|k)$, and $\tau(q|k)$ is the averaged reducible vertex function that can be calculated by using, for example, the pole approximation for the Green's function [4].

The presence of the Green's functions $|G(q,2\omega)|^2$ and $|G(k,\omega)|^4$ in Eq. (3.2) leads to the amplification of nonlinear processes at the metal film-crystal interface in the Kretschmann geometry [23], because the Green's function $G(q,2\omega)$ has poles at the wavenumbers of the surface plasmon polaritons, $k_{sp}(2\omega)$, and $G(k,\omega)$ has poles at the wavenumbers of the surface plasmon polaritons, $k_{sp}(\omega)$. Thus, when the angle of incidence is the optimal one for the excitation of surface plasmon polaritons the enhancement of the intensity of the second harmonic light can reach several orders of magnitude compared with that for scattering from a single surface. This is true for both the coherent and incoherent parts of the intensity. When the angle of scattering is the optimal one for the detection of surface plasmon polaritons of frequency 2ω resonance enhancement of the intensity can also occur [24]. However, this enhancement can be observed only in the intensity of the reflected light.

The second interesting feature of Eq. (3.2) is the presence of the term that contains the factor $|G(q-k,\omega)|^2$. This term describes the second harmonic generation due to the nonlinear mixing of the incident light with the surface plasmon polaritons which are excited due to the surface roughness. Its presence leads to the appearance of strong resonant peaks at $q = k \pm k_{sp}(\omega)$ and $q = k \pm k_{sp}^{(p)}(\omega)$. This is the origin of the peak of the strong enhancement of second harmonic generation in reflection from a randomly rough metal surface that was first observed in [16], and analyzed by Deck and Grygier [17] in the framework of the first order perturbation theory in the small roughness. These peaks of the enhanced second harmonic generation are due to the resonant interaction of the incident light and forward/backward propagating surface plasmon polaritons excited in a single scattering event. No interference effects are involved in the formation of the peaks. The main reason for using the Kretschmann geometry in the experiments [6,8–11] was to excite surface plasmon polaritons associated with the metal–vacuum (metal–nonlinear crystal) interface through the ATR phenomenon. Thus, the angle of incidence was tuned so that $k = k_{sp}(\omega)$. In this case the peaks due to the poles of $G(q-k,\omega)$ move so that the peak due to the mixing of the incident light with the forward propagating surface plasmon polariton moves into the nonradiative region, while that due to the mixing of the incident light with the backward propagating surface plasmon polariton moves to the direction of the normal to the mean surface, $q = k - k_{sp}(\omega) = 0$. The strength and the shape of the peak is determined by the effective nonlinear coefficient $\Gamma_1(q,k)$. In the case when the nonlinearity of the system is due to the film interfaces only, this effective nonlinear coefficient is linear in q for small q. This is the result of the symmetry of the surface nonlinear polarization that forbids second harmonic generation by contrapropagating surface plasmon polaritons [14]. Since in the case where the angle of incidence is optimal for exciting surface plasmon polaritons the incident light is the surface plasmon polariton propagating in the forward direction, its nonlinear mixing with the scattered surface plasmon polaritons propagating in the backward direction is forbidden by this symmetry. As a result, the resonant peak has an antiresonant shape. The width of the peak is determined by the decay rate of the surface plasmon polaritons on the rough interface $\mathrm{Im}(k_{sp}(\omega)) = \Delta_{tot}(\omega) = \Delta_{\epsilon}(\omega) + \Delta_{sc}(\omega)$, where $\Delta_{\epsilon}(\omega)$ is the decay rate of the surface plasmon polaritons of frequency ω due to ohmic losses, while $\Delta_{sc}(\omega)$ is their decay rate due to their roughness-induced scattering into other surface plasmon polaritons.

The effects of the multiple scattering of surface plasmon polaritons are contained in the function $\tau(q|k)$. As was shown in [4,25,26], because our scattering system supports two surface plasmon polaritons at the frequency ω, the function $\tau(q|k)$ contains a superposition of two Lorentzian peaks centered at $q + k = 0$ (enhanced backscattering) with halfwidths $\Delta_{tot}(\omega)$and $\Delta_{tot}^{(p)}(\omega)$. It also contains Lorentzian peaks centered at $q + k \pm (k_{sp}(\omega) - k_{sp}^{(p)}(\omega)) = 0$ (satellite peaks) with a half-width $\Delta_{tot}(\omega) + \Delta_{tot}^{(p)}(\omega)$. In these results $\Delta_{tot}^{(p)}(\omega)$ is the decay rate of the surface plasmon polariton associated with the prism–metal interface. Therefore, $\tau(q-k|k)$ entering the second term in Eq. (3.2) has a peak at $q = 0$. Since this contribution arises due to the nonlinear mixing of the incident and scattered waves, it describes the coherent generation of the second harmonic light by the incident and backscattered radiation, enhanced due to the multiple scattering of surface plasmon polaritons by the roughness.

The strongest contribution associated with the multiple scattering of surface plasmon polaritons of frequency ω comes, however, from the second term in Eq. (3.2), which describes the nonlinear mixing of the

multiply-scattered surface plasmon polaritons of frequency ω propagating in opposite directions. In the pole approximation for the Green's function [4,25] this contribution has the form

$$\langle I(\theta_s|2\omega)\rangle_{\rm res} = \frac{512\omega^3}{\epsilon_0 c^4}\left|t_{10}(k,\omega)e^{2i\alpha_1(k,\omega)D}\right|^2|G(q,2\omega)|^2|G(k,\omega)|^2\left\{\frac{C^4(\omega)}{\Delta_{\rm tot}(\omega)}\frac{1}{q^2+4\Delta_{\rm tot}^2(\omega)}\right.$$
$$\times\left[|\Gamma_2(q,k_{\rm sp}(\omega))|^2\tau(k_{\rm sp}(\omega)|k)\tau(q-k_{\rm sp}(\omega)|k)+|\Gamma_2(q,-k_{\rm sp}(\omega))|^2\tau(-k_{\rm sp}(\omega)|k)\right.$$
$$\left.\times\tau(q+k_{\rm sp}(\omega)|k)\right]+C_p^4(\omega)\Delta_{\rm tot}^{(p)}(\omega)\frac{1}{q^2+4(\Delta_{\rm tot}^{(p)}(\omega))^2}\left[|\Gamma_2(q,k_{\rm sp}^{(p)}(\omega))|^2\right.$$
$$\times\tau(k_{\rm sp}^{(p)}(\omega)|k)\tau(q-k_{\rm sp}^{(p)}(\omega)|k)+|\Gamma_2(q,-k_{\rm sp}^{(p)}(\omega))|^2$$
$$\left.\left.\times\tau(-k_{\rm sp}^{(p)}(\omega)|k)\tau(q+k_{\rm sp}^{(p)}(\omega)|k)\right]\right\}, \tag{3.3}$$

where $C(\omega)$ and $C_p(\omega)$ are the residues of the Green's function at the poles $\pm k_{\rm sp}(\omega)$ and $\pm k_{\rm sp}^{(p)}(\omega)$, respectively. We have not included in Eq. (3.3) the contribution from the possible satellite peaks [26], since their positions are far from the direction normal to the mean surface. As expected, the result given by Eq. (3.3) contains Lorentzian factors centered at $q=0$. However, the efficiency of the nonlinear mixing of the contrapropagating surface plasmon polaritons is determined by the effective nonlinear coefficients $\Gamma_2(q,\pm k_{\rm sp}(\omega))$ and $\Gamma_2(q,\pm k_{\rm sp}^{(p)}(\omega))$. In the case where the nonlinearity of the system is due to the film interfaces only, these effective nonlinear coefficients are linear in q for small q. This is the manifestation of the well known fact that the symmetry of the nonlinear polarization of a metal surface forbids such processes [14]. As a result, the contribution given by Eq. (3.3) displays a dip rather than a peak in the direction normal to the mean surface. The depth of this dip depends strongly on the values of the material parameters and the angle of incidence of the fundamental light. However, when the metal film is in contact with a nonlinear quartz crystal, a suitable choice of the orientation of the crystal [18] makes possible second harmonic generation by contrapropagating surface plasmon polaritons, and a peak of the enhanced second harmonic generation occurs in the direction normal to the mean surface.

Recently the first experimental studies of multiple-scattering effects in the second harmonic generation of light scattered from a clean one-dimensional vacuum-metal interface were carried out in a series of papers by O'Donnell and his colleagues [27–29]. It was found that, for weakly rough silver surfaces a dip is present in the retroreflection direction in the angular dependence of the intensity of the scattered second harmonic light rather than the peak that occurs in scattering at the fundamental frequency [27,29], while no well-defined peak or dip in the direction normal to the mean surface was observed [28]. In the latter experiment the random surfaces were fabricated in a special way to produce a strong excitation of surface plasmon polaritons of the fundamental frequency, propagating in both the forward and backward directions. The results of Ref. [28] suggest a surface nonlinear polarization of a form different from those usually used in studies of second harmonic generation in scattering from an isotropic metal surface. The possible anisotropy of the surface could be the reason for the absence of the dip in the direction normal to the mean surface. If, for some reason, the nonlinear surface polarization is such that no peak or dip in the direction normal to the mean surface, which is due to the multiple scattering of surface plasmon polaritons by the surface roughness, can be formed, in this case in experiments in the Kretschmann geometry the resonant peak at $q=0$ will also be absent. Finally, the different qualitative results of [8,11] and [10] suggest that the mechanism for the nonlinear polarization might be different for the strongly rough surfaces studied in [8,11] and for the weakly rough surfaces studied in [10]. As an example, in Fig. 4 we present the mean intensity of the second harmonic light, calculated by the small-amplitude perturbation approach, assuming that the surface nonlinearity arises from a thin subsurface layer of the metal

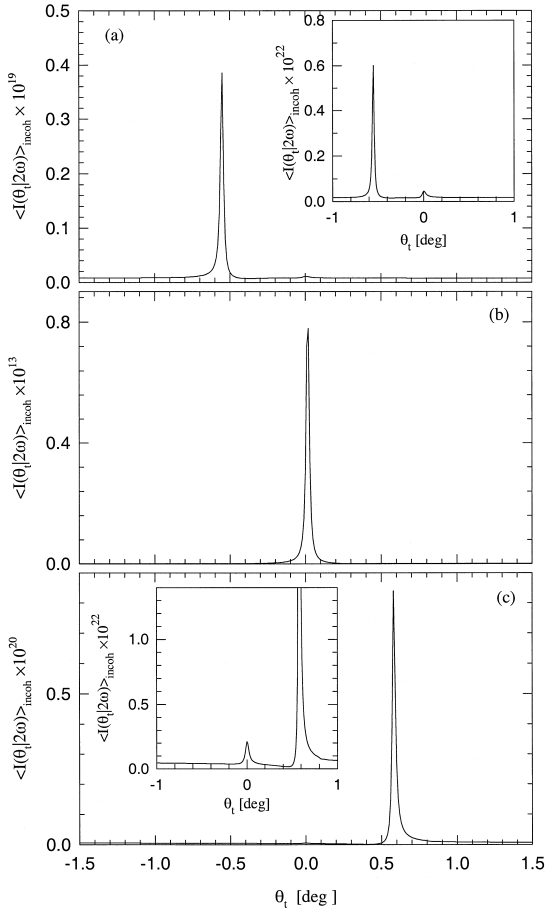

Fig. 4. The mean differential intensity of the second harmonic light as a function of the angle of transmission θ_t for the transmission of *p*-polarized light through the same film as in Fig. 2, but for the case where the surface nonlinear polarization has an artificial nature that is described in the text. The inset presents the contribution to $\langle I(\theta_t|2\omega)\rangle_{\text{incoh}}$ from double-scattering processes only. The angles of incidence are (a) $\theta_0 = 23.5°$, (b) $\theta_0 = 24°$, and (c) $\theta_0 = 24.5°$. The optimal angle for excitation of surface plasmon polaritons at the film-vacuum interface is 24°.

film, or from a monolayer of nonlinear molecules covering the metal film. In either case the surface nonlinear polarization is associated not with the surface itself but is given by the 'bulk' nonlinear polarization of the layer, which is assumed to be given by Eq. (2.6). In this case both the nonlinear mixing of the incident light with the backward propagating surface plasmon polariton and the nonlinear mixing of the multiply-scattered surface plasmon polaritons lead to the appearance of the peak in transmission in the direction normal to the mean surface.

4. Conclusions

In this work we have discussed the second harmonic generation of light in transmission in the Kretschmann ATR geometry, in the case where a metal film-vacuum or a metal film-nonlinear quartz crystal interface is weakly rough. We have analyzed the origin of the peak of the enhanced second harmonic generation observed in

the direction normal to the mean surface in [6,8,11]. We have shown that in the case where the metal film is in contact with vacuum, so that the nonlinearity of the film surfaces gives rise to the harmonic light, the interference effects in the multiple scattering of surface plasmon polaritons lead, due to the symmetry of the nonlinear polarization, to the appearance of a dip rather than a peak in the direction normal to the mean surface. When the second harmonic generation is due to the nonlinearity of the crystal adjacent to the metal film a peak of the enhanced second harmonic generation occurs in this direction.

However, when the angle of incidence is chosen to be optimal for the excitation of surface plasmon polaritons localized at the film-vacuum/nonlinear crystal interface, a much stronger mechanism leads to the appearance of a peak of the enhanced second harmonic generation in the direction normal to the mean interface. It is the nonlinear mixing of the incident light, which in this case is the resonantly excited surface plasmon polariton, and the backward-propagating surface plasmon polariton excited through the roughness that forms the peak. No interference effects are involved in the formation of the peak. Since this peak appears already in the single-scattering processes, it is considerably stronger than the weak features associated with multiple scattering. When the angle of incidence is shifted from the optimal one, the peak moves away from the direction normal to the mean interface, so that the weak peak/dip which is due to the multiple-scattering effects, can be observed. However, the efficiency of second harmonic generation decreases considerably as the angle of incidence shifts from the optimal one.

Acknowledgements

The work of A.A.M., I.V. N., and T.A.L. was supported by Army Research Office Grant DAAG 55-98-C-0034. The work of E.R.M., and M. L.-L. was supported by CONACYT Grant 32254–E.

Appendix A.

To solve the linear scattering problem we seek the solution of Eq. (2.11) in the form

$$T(q,\omega) = -2iH_0 G_\omega(q|k)\frac{\alpha_1(k,\omega)}{\epsilon_1(\omega)}t_{01}(k,\omega)e^{i\alpha_1(k,\omega)D}, \tag{A.1}$$

where we have introduced the Green's function $G_\omega(q|k)$ associated with the randomly rough interface between the vacuum and the metal that is the solution of the equation

$$G_\omega(q|k) = G(q,\omega)2\pi\delta(q-k) + G(q,\omega)t(q|k)G(k,\omega). \tag{A.2}$$

In Eq. (A.2) $G(q,\omega)$, defined by $\langle G_\omega(q|k)\rangle = 2\pi\delta(q-k)G(q,\omega)$, is the averaged Green's function and is given by

$$G(q,\omega) = 2\pi\delta(q-k)\frac{1}{G_0^{-1}(k,\omega) - M(k,\omega)}, \tag{A.3}$$

where $M(k,\omega)$ is the averaged proper self-energy. The latter is given by

$$\langle M(q|k)\rangle = 2\pi\delta(q-k)M(k,\omega), \tag{A.4}$$

where the (unaveraged) proper self-energy $M(q|k)$ is the solution of [25]

$$M(q|k) = V(q,k|\omega) + \int_{-\infty}^{\infty}\frac{dp}{2\pi}M(q|p)G(p,\omega)W(p,k|\omega), \tag{A.5}$$

and we have introduced the notation

$$W(q,k|\omega) = V(q,k|\omega) - \langle M(q|k)\rangle. \tag{A.6}$$

The operator $t(p|r)$ in Eq. (A.2) was introduced to satisfy

$$\int_{-\infty}^{\infty} \frac{dp}{2\pi} W(q,p|\omega) G(p|k,\omega) = t(q|k) G(k,\omega), \tag{A.7}$$

and is the solution of the equation

$$t(q|k) = W(q,k|\omega) + \int_{-\infty}^{\infty} \frac{dp}{2\pi} \int_{-\infty}^{\infty} \frac{dr}{2\pi} W(q,p|\omega) G(p,\omega) t(p|k). \tag{A.8}$$

From Eq. (A.2) it follows that $\langle t(q|k)\rangle = 0$. The transmission amplitude $T(q,\omega)$ in terms of the averaged Green's function $G(p,\omega)$ and the operator $t_\omega(q|k)$ is then given by

$$T(q,\omega) = \mathscr{T}_0(k,\omega) H_0\left[2\pi\delta(q-k) + G(q,\omega) t(q|k)\right], \tag{A.9}$$

where

$$\mathscr{T}_0(k,\omega) = -2i \frac{\alpha_0(k,w}{\epsilon_0} t_{10}(k,\omega) e^{i\alpha_1(k,\omega)D} G(k,\omega). \tag{A.10}$$

Having in hand the solution for $T(q,\omega)$ we can calculate the fields of frequency ω in the prism and in the metal film and, consequently, the nonlinear driving term of frequency 2ω.

In the same manner we can seek the solution of the nonlinear scattering problem. If we now introduce a new unknown function $S(q|2k)$ by the equation

$$T(q,2\omega) = iG_0(p,2\omega) Q(q,2\omega) + iG_0(q,2\omega) \int_{-\infty}^{\infty} \frac{dq}{2\pi} S(p|q) G_0(q,2\omega) Q(q,2\omega), \tag{A.11}$$

it follows from Eq. (2.11) that the function $S(q|p)$ satisfies the equation

$$S(q|p) = V(q|p,2\omega) + \int_{-\infty}^{\infty} \frac{dr}{2\pi} V(q|r,2\omega) G_0(r,2\omega) S(r|p). \tag{A.12}$$

The amplitude $S(q|p)$ describes the multiple scattering of electromagnetic waves of frequency 2ω, including the scattering of surface plasmon polaritons of frequency 2ω. It can be rewritten in terms of the Green's function associated with the rough interface of the metal film exactly as this was done for the fundamental fields. As a result we obtain the following expression for the transmission amplitude

$$T(p,2\omega) = iG(p,2\omega) Q(p,2\omega) + iG(p,2\omega) \int_{-\infty}^{\infty} \frac{dq}{2\pi} t_{2\omega}(p|q) G(q,2\omega) Q(q,2\omega). \tag{A.13}$$

Appendix B.

In this appendix we present the explicit expression for the effective nonlinear coefficients $\gamma_i^{(p,v,b)}$.

$$\gamma_0^{(p)}(k) = t_{10}(2k,2\omega) e^{i\alpha_1(2k,2\omega)D} \left\{ \mu_3^{(p)} k \frac{\alpha_0(2k,2\omega)}{\epsilon_0} \frac{\alpha_1(k,\omega)}{\epsilon_1(w)} f_+(k) f_-(k) \right.$$
$$\left. -2k\left[\mu_1^{(p)} k^2 f_+^2(k) + \mu_2^{(p)} \left(\frac{\alpha_1(k,\omega)}{\epsilon_1(\omega)} \right)^2 f_-^2(k) \right] \right\} \tag{B.1}$$

$$\gamma_1^{(p)}(q,k) = t_{10}(q,2\omega)\, e^{i\alpha_1(q,2\omega)D}\Bigg\{ \mu_3^{(p)} \frac{\alpha_0(q,2\omega)}{\epsilon_0}\Bigg[k\frac{\alpha_1(q-k,\omega)}{\epsilon_1(w)} f_+(k) f_-(q-k) + (q-k)\frac{\alpha_1(k,\omega)}{\epsilon_1(w)} f_+(q-k) f_-(k)\Bigg] - 2q\Bigg[\mu_1^{(p)}(q-k)\, k f_+(q-k) f_+(k) + \mu_2^{(p)} \frac{\alpha_1(k,\omega)}{\epsilon_1(\omega)} \frac{\alpha_1(q-k,\omega)}{\epsilon_1(\omega)} f_-(q-k) f_+(k)\Bigg]\Bigg\} \tag{B.2}$$

$$\gamma_2^{(p)}(q,p) = t_{10}(q,2\omega)\, e^{i\alpha_1(q,2\omega)D}\Bigg\{ \mu_3^{(p)} \frac{\alpha_0(q,2\omega)}{\epsilon_0}\Bigg[p\frac{\alpha_1(q-p,\omega)}{\epsilon_1(w)} f_+(p) f_-(q-p) + (q-p)\frac{\alpha_1(p,\omega)}{\epsilon_1(w)} f_+(q-p) f_-(p)\Bigg] - 2q\Bigg[\mu_1^{(p)}(q-p)\, p f_+(q-p) f_+(p) + \mu_2^{(p)} \frac{\alpha_1(q-p,\omega)}{\epsilon_1(\omega)} \frac{\alpha_1(p,\omega)}{\epsilon_1(\omega)} f_-(q-p) f_+(p)\Bigg]\Bigg\}, \tag{B.3}$$

with

$$f_\pm(q) = \frac{1 \pm r_{12}(q,\omega)\, e^{2i\alpha_1(q,\omega)D}}{t_{12}(q,\omega)\, e^{i\alpha_1(q,\omega)D}} \tag{B.4}$$

and

$$\gamma_0^{(v)}(k) = \mu_3^{(v)} k \frac{\alpha_1(2k,2\omega)}{\epsilon_1(2\omega)} \frac{\alpha_2(k,\omega)}{\epsilon_2(w)} g_+(2k) - 2k g_-(2k)\Bigg[\mu_1^{(v)} k^2 + \mu_2^{(v)} \left(\frac{\alpha_2(k,\omega)}{\epsilon_2(\omega)} \right)^2 \Bigg], \tag{B.5}$$

$$\gamma_1^{(v)}(q,k) = \mu_3^{(v)} \frac{\alpha_1(q,2\omega)}{\epsilon_1(2\omega)} g_+(q)\Bigg[k\frac{\alpha_2(q-k,\omega)}{\epsilon_2(w)} + (q-k)\frac{\alpha_2(k,\omega)}{\epsilon_2(w)}\Bigg] - 2qg_-(q)\Bigg[\mu_1^{(v)}(q-k)\, k + \mu_2^{(v)} \frac{\alpha_2(k,\omega)}{\epsilon_2(\omega)} \frac{\alpha_2(q-k,\omega)}{\epsilon_2(\omega)}\Bigg] \tag{B.6}$$

$$\gamma_2^{(v)}(q,p) = \mu_3^{(v)} \frac{\alpha_1(q,2\omega)}{\epsilon_1(2\omega)} g_+(q)\Bigg[p\frac{\alpha_2(q-p,\omega)}{\epsilon_2(w)} + (q-p)\frac{\alpha_2(p,\omega)}{\epsilon_2(w)}\Bigg] - 2qg_-(q)\Bigg[\mu_1^{(v)}(q-p)\, p + \mu_2^{(v)} \frac{\alpha_2(q-p,\omega)}{\epsilon_2(\omega)} \frac{\alpha_2(p,\omega)}{\epsilon_2(\omega)}\Bigg], \tag{B.7}$$

with

$$g_\pm(q) = 1 \pm r_{01}(q,2\omega)\, e^{2i\alpha_1(q,2\omega)D}. \tag{B.8}$$

$$\gamma_0^{(b)}(k) = \frac{id_{11}}{\epsilon_2^2(\omega)}\Bigg\{ \frac{\alpha_1(2k,2\omega)}{\epsilon_1(2\omega)}\left(1 + r_{01}(2k,2\omega)\, e^{2i\alpha_1(2k,2\omega)D}\right) + \frac{2\alpha_2(k,\omega)}{\epsilon_2(2\omega)}\left(1 - r_{01}(2k,2\omega)\, e^{2i\alpha_1(2k,2\omega)D}\right)\Bigg\} F(2k,k), \tag{B.9}$$

$$\gamma_1^{(b)}(q,k) = \frac{id_{11}}{\epsilon_2^2(\omega)}\left\{\frac{\alpha_1(q,2\omega)}{\epsilon_1(2\omega)}\left(1 + r_{01}(q,2\omega)\,e^{2i\alpha_1(q,2\omega)D}\right) + \frac{\alpha_2(k,\omega) + \alpha_1(q-k,\omega)}{\epsilon_2(2\omega)}\left(1 - r_{01}(q,2\omega)\,e^{2i\alpha_1(q,2\omega)D}\right)\right\}\left[F(q,k) + F(q,q-k)\right], \tag{B.10}$$

$$\gamma_2^{(b)}(q,p) = \frac{id_{11}}{\epsilon_2^2(\omega)}\left\{\frac{\alpha_1(q,2\omega)}{\epsilon_1(2\omega)}\left(1 + r_{01}(q,2\omega)\,e^{2i\alpha_1(q,2\omega)D}\right) + \frac{\alpha_2(p,\omega) + \alpha_1(q-p,\omega)}{\epsilon_2(2\omega)}\left(1 - r_{01}(q,2\omega)\,e^{2i\alpha_1(q,2\omega)D}\right)\right\}F(q,p). \tag{B.11}$$

References

[1] F. Brown, R.E. Parks, A.M. Sleeper, Phys. Rev. Lett. 14 (1965) 1029.
[2] S.S. Jha, Phys. Rev. Lett. 15 (1965) 412.
[3] N. Bloembergen, R.K. Chang, S.S. Jha, C.H. Lee, Phys. Rev. 74 (1968) 813.
[4] A.R. McGurn, A.A. Maradudin, V. Celli, Phys. Rev. B 31 (1985) 4866.
[5] A.R. McGurn, T.A. Leskova, V.M. Agranovich, Phys. Rev. B 44 (1991) 11441.
[6] X. Wang, H.J. Simon, Opt. Lett. 16 (1991) 1475.
[7] H.J. Simon, Y. Wang, L.B. Zhou, Z. Chan, Opt. Lett. 17 (1992) 1268.
[8] O.A. Aktsipetrov, V.N. Golovkina, O.I. Kapusta, T.A. Leskova, N.N. Novikova, Phys. Lett. A 170 (1992) 231.
[9] Y. Wang, H.J. Simon, Phys. Rev. B 47 (1993) 13695.
[10] L. Kuang, H.J. Simon, Phys. Lett. A 197 (1995) 257.
[11] S.I. Bozhevolnyi, K. Pedersen, Surface Science 377–379 (1997) 384.
[12] E. Kretschmann, Z. Physik 241 (1971) 313.
[13] D. Sarid, Phys. Rev. Lett. 47 (1981) 1927.
[14] M. Fukui, J.E. Sipe, V.C.Y. So, G.I. Stegeman, Solid State Commun. 27 (1978) 1265.
[15] H.J. Simon, D.E. Mitchell, J.G. Watson, Phys. Rev. Lett. 33 (1974) 1531.
[16] C.K. Chen, A.R.B. de Castro, Y.R. Shen, Phys. Rev. Lett. 46 (1981) 145.
[17] R.T. Deck, R.K. Grygier, Appl. Opt. 23 (1984) 3202.
[18] C.K. Chen, A.R.B. de Castro, Y.R. Shen, Opt. Lett. 4 (1979) 393.
[19] I.V. Novikov, A.A. Maradudin, T.A. Leskova, E.R. Méndez, M. Leyva-Lucero (unpublished work).
[20] Lord Rayleigh, The Theory of Sound, 2nd ed, vol. II, MacMillan, London, 1895, pp. 89 and 297–311.
[21] R. Petit, M. Cadilhac, C.R. Acad. Sci. Paris B 262 (1966) 468.
[22] N.R. Hill, V. Celli, Phys. Rev. B 17 (1978) 2478.
[23] Y.J. Chen, E. Burstein, Nuovo Cimento 39B (1977) 807.
[24] H.J. Simon, R.E. Benner, J.G. Rako, Opt. Commun. 23 (1977) 245.
[25] G.C. Brown, V. Celli, M. Haller, A.A. Maradudin, A. Marvin, Phys. Rev. B 31 (1985) 4993.
[26] See, for example, V. Freilikher, E. Kanzieper, A.A. Maradudin, Phys. Rep. 288 (1997) 127.
[27] K.A. O'Donnell, R. Torre, C.S. West, Opt. Lett. 21 (1996) 1738.
[28] K.A. O'Donnell, R. Torre, C.S. West, Phys. Rev. B 55 (1997) 7985.
[29] K.A. O'Donnell, R. Torre, Opt. Commun. 138 (1997) 341.

INSTITUTE OF PHYSICS PUBLISHING WAVES IN RANDOM MEDIA

Waves Random Media **11** (2001) 183–231 www.iop.org/Journals/wr PII: S0959-7174(01)18799-9

Second harmonic generation of light in the Kretschmann attenuated total reflection geometry in the presence of surface roughness

I V Novikov[1,4], **A A Maradudin**[1], **T A Leskova**[2], **E R Méndez**[3] **and M Leyva-Lucero**[3,5]

[1] Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California Irvine, CA 92697, USA
[2] Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, 142092, Russia
[3] División de Física Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, Baja California 22800, Mexico

Received 9 November 2000, in final form 12 February 2001

Abstract

Using a computer simulation approach we study the generation of second harmonic light in reflection and in transmission in the Kretschmann attenuated total reflection geometry. In this geometry a thin metal film is deposited on the planar base of a dielectric prism, through which p-polarized light is incident on the film. The back surface of the film, which separates the film from vacuum, is a one-dimensional, randomly rough surface, whose generators are normal to the plane of incidence of the light. The nonlinearity responsible for the second harmonic generation is assumed to arise at the prism–metal and metal–vacuum interfaces, and thus enters the problem only through the boundary conditions at these interfaces at the harmonic frequency. The source terms entering these boundary conditions are obtained from the solutions of the corresponding scattering and transmission problems at the fundamental frequency. It is found that a peak in the angular dependence of the intensity of both the transmitted and reflected second harmonic light occurs in the directions normal to the mean scattering surface, in addition to an enhanced backscattering peak in the retroreflection direction. The enhanced transmission peak occurs in the non-radiative region, and therefore cannot be observed in the far field.

1. Introduction

The coherent interference of multiply scattered electromagnetic waves propagating in volume disordered media, or reflected from or transmitted through random surfaces, has attracted a great deal of attention in recent years due to their analogy with the interference effects

[4] Current address: Stanford Research Systems, 1290 Ream Wood Ave, Sunnyvale, CA 94089, USA.
[5] Current address: Escuela de Ciencias Físico-Matematicas, Universidad Autonoma de Sinaloa, Ciudad Universitaria, CP 80000 Culiacan, Sinaloa, Mexico.

0959-7174/01/030183+49$30.00 © 2001 IOP Publishing Ltd Printed in the UK

caused by the multiple scattering of electrons propagating in strongly disordered media that are responsible for Anderson localization. Even if the volume disorder or surface roughness is weak, this interference is important because it can give rise to weak localization effects, including enhanced backscattering, enhanced transmission and satellite peaks.

For randomly rough metallic surfaces, the multiple scattering that gives rise to enhanced backscattering involves the excitation of surface plasmon polaritons [1–3] or the multiple scattering of volume electromagnetic waves [4–7]. The former is the main mechanism involved in the case of weakly rough surfaces, while the latter appears to be the dominant mechanism in the case of strongly rough surfaces.

Recently, the main ideas of weak localization theory have been applied to nonlinear optical phenomena in disordered media (see, e.g., [8]). In particular, a perturbative calculation of the second harmonic generation of light in the scattering of p-polarized light from a clean, weakly rough, one-dimensional, random metal surface predicted that a peak in the angular dependence of the intensity of the second harmonic light should occur not only in the retroreflection direction but also in the direction normal to the mean surface [9]. The second harmonic radiation in this study was assumed to be generated in a metal–vacuum interface layer that has a finite thickness on the microscopic scale, within which the isotropy of the bulk metal is broken. Within this layer the electromagnetic fields and material constants vary rapidly, and their gradients give rise to the optical nonlinearity of the surface. Therefore, the nonlinear polarization enters the scattering problem only through the boundary conditions on the second harmonic fields at the metal–vacuum interface.

A subsequent computer simulation and perturbation theory study of second harmonic generation from weakly rough one-dimensional random metal surfaces showed that either a peak or a dip can occur in the retroreflection direction, depending on the values of the phenomenological constants entering the nonlinear boundary conditions, and on the angle of incidence, and that there is no peak in the direction normal to the mean surface [10]. The experimental observation that only a dip occurs in the retroreflection direction for all angles of incidence [11, 12], yielded valuable information about the phenomenological constants in the nonlinear boundary conditions [13]. Not surprisingly, surface plasmon polaritons play a prominent role in the formation of the features observed in the retroreflection direction, which have been shown to be due to the nonlinear excitation of surface plasmon polaritons at the harmonic frequency.

Experimental [14] and numerical simulation [15] results for second harmonic generation in the scattering of p-polarized light from clean, one-dimensional, random metal surfaces with large RMS heights and large RMS slopes agree in showing well defined dips in the retroreflection direction. In this case the multiple-scattering processes that give rise to these dips are of a different nature from those occurring at weakly rough metal surfaces, and involve the multiple scattering of volume electromagnetic waves by the surface roughness.

Thus, the experimental results for second harmonic generation in the scattering of light from both weakly rough and strongly rough, clean, one-dimensional, random surfaces, are explained well by the numerical simulation and perturbation theory studies of this generation.

The earliest experimental investigations of the second harmonic generation of light in the scattering of light from randomly rough metal surfaces, however, were not carried out on clean, one-dimensional, random metal surfaces. Instead, to amplify the second harmonic signal the Kretschmann attenuated total reflection (ATR) geometry [16] was used. In the experiments of [17–19] the scattering system consisted of a silver film deposited on the planar base of a dielectric prism, through which p-polarized light was incident on the film. The unilluminated (back) surface of the film was a randomly rough surface that separated the

film from a nonlinear quartz crystal. In this case the nonlinear interaction occurred in the quartz crystal rather than at the more weakly nonlinear silver surfaces. A well defined peak of second harmonic generation in the direction normal to the mean silver–quartz interface was observed in transmission in [17]. When the experiment was carried out with long-range surface plasmon polaritons [20], peaks of the second harmonic generation were present both in the retroreflection direction [18, 19], and in the direction normal to the mean silver–quartz interface in transmission [19].

In [21–23] efforts were made to observe peaks of enhanced second harmonic generation in the geometry used in [17–19] but with the quartz crystal replaced by vacuum, i.e. at a silver–film vacuum interface. A well defined peak in the direction normal to the interface was observed in transmission in [21, 23]. However, in [23] the peak was observed when the film was a gold film, but not when it was a silver film. In contrast, in [22] only a broad depolarized background, but no peak in the direction of the normal to the mean silver–vacuum interface was observed.

The Kretschmann ATR geometry, especially in the case where the metal film is in contact with vacuum, has attracted a good deal of attention from both theorists and experimentalists, because it is designed to yield experimentally the dispersion relation of surface plasmon polaritons. This structure supports two surface plasmon polaritons, of which the one bound to the metal–vacuum interface leaks into the prism. At a given wavelength of the incident light one can find the optimal condition for exciting the surface plasmon polaritons of this wavelength localized to the metal–vacuum interface by varying the angle of incidence. At the optimal angle of incidence the transmissivity of the system displays a maximum, while the reflectivity simultaneously displays a minimum. From a determination of this optimal angle of incidence the wavenumber of the surface plasmon polariton at the wavelength of the incident light can be obtained. The strong excitation of surface plasmon polaritons at this angle of incidence in the Kretschmann ATR geometry make it an attractive one to use in studies of second harmonic generation of light at weakly rough metal surfaces, where the multiple scattering of these surface waves is expected to play a dominant role in the formation of weak localization features in the angular distribution of the second harmonic light scattered from or transmitted through the randomly rough surfaces.

A perturbative theory of the second harmonic generation of light in transmission in the Kretschmann ATR geometry, both in the presence of a nonlinear crystal in contact with the metal film and in its absence, has been developed very recently [24], for the case where the metal–vacuum/nonlinear crystal interface is a one-dimensional random interface whose generators are perpendicular to the plane of incidence of p-polarized light. This theory predicts that when the angle of incidence is optimal for the excitation of surface plasmon polaritons at the film–vacuum/nonlinear crystal interface an intense peak should be present in the angular dependence of the intensity of the transmitted harmonic light in the direction normal to the mean interface. This peak is already present in the single-scattering approximation, and is therefore not a multiple-scattering phenomenon. For an angle of incidence that differs from the optimal one, a dip appears in the angular dependence of the intensity of the transmitted harmonic light in the direction normal to the mean interface when the metal is in contact with vacuum. When the metal is in contact with a nonlinear crystal a peak in the angular dependence of the transmitted light occurs in this direction. These dips and peaks are multiple-scattering effects. The explanations for the peaks in the direction normal to the mean metal–vacuum/nonlinear crystal interface correct the one presented in [9, 21, 23].

In this paper we extend the results of [24], which are valid only for a weakly rough metal–vacuum/nonlinear crystal interface, to more strongly rough interfaces, by carrying out a numerical simulation of the scattering of p-polarized light from, and its transmission through,

a metal film in the Kretschmann ATR geometry, in the case studied experimentally in [21–23], where the metal film is in contact with vacuum. In our analysis the metal–vacuum interface is assumed to be a one-dimensional, random interface. We assume that the second harmonic generation occurs in the metal–prism and metal–vacuum interface layers that have a finite thickness on the microscopic scale and neglect the small contribution to the nonlinearity coming from the bulk of the film. In this case the nonlinear polarization of the interfaces can be taken into account through the boundary conditions for the second harmonic fields. Therefore, both the fundamental and harmonic fields satisfy Helmholtz equations in the prism, the metal film and in the vacuum. We first solve the linear problem of the scattering of light of the fundamental frequency ω, and use its solution to determine the surface nonlinear polarization at the harmonic frequency 2ω. Having in hand the boundary conditions for the harmonic fields across the prism–metal film and metal film–vacuum interfaces we then solve the linear problem of the scattering of light of second harmonic frequency. Our results predict the occurrence of a peak in the direction normal to the mean metal–vacuum interface in the angular dependence of the intensity of the second harmonic light generated both in reflection from and in transmission through the rough metal film. The explanations for these features are in agreement with those presented in [24].

However, it will not be possible for us to make a quantitative comparison between the theoretical predictions of this work and the experimental results presented in [21–23]. In these papers information such as the dielectric constant of the prism and the mean thickness of the metal film was not provided. Moreover, the metal–vacuum interfaces in these studies were two dimensional rather than one dimensional, and their roughness was not characterized in a quantitative fashion. We have therefore chosen values for these characteristics of the scattering system that are compatible with those used in earlier studies of the scattering of light from, and its transmission through, metal films [7, 25], and in studies of the second harmonic generation of light from clean, randomly rough metal surfaces [10, 15, 24]. It is our hope that the results presented in this paper will stimulate experimental studies of the second harmonic generation of light in the Kretschmann geometry on well characterized systems.

The outline of this paper is as follows. In section 2 we define the scattering system studied in this work, and present the two models of the random surface roughness that will be used in this study. The two surface plasmon polaritons supported by the scattering system, and their consequences for the angular distribution of the intensity of the scattered and transmitted light at the fundamental frequency, are described in section 3. The scattering problem is formulated in section 4, and the scattering equations at the fundamental frequency are derived in section 5. The boundary conditions at the harmonic frequency are derived in section 6. The scattering equations at the harmonic frequency are obtained in section 7. The results obtained from the solution of these equations are presented in section 8, and the conclusions reached on their basis are summarized in section 9. Two appendices containing material used in the text complete the paper. Readers who do not need the detailed description of the numerical methods used in this work can skip sections 4–7, as well as both appendices. The quantities we are discussing in section 8, are introduced in the final paragraphs of sections 5 and 7.

2. Scattering geometry

The scattering system we consider in this paper is depicted in figure 1. It consists of a dielectric in the region $x_3 > D$, a metal in the region $\zeta(x_1) < x_3 < D$ and vacuum in the region $x_3 < \zeta(x_1)$. The dielectric medium is characterized by a real, positive dielectric constant ϵ_0, while the metal is characterized by an isotropic, frequency-dependent, complex dielectric function $\epsilon(\omega)$. We will be interested in the frequency range in which the real part of $\epsilon(\omega)$ is

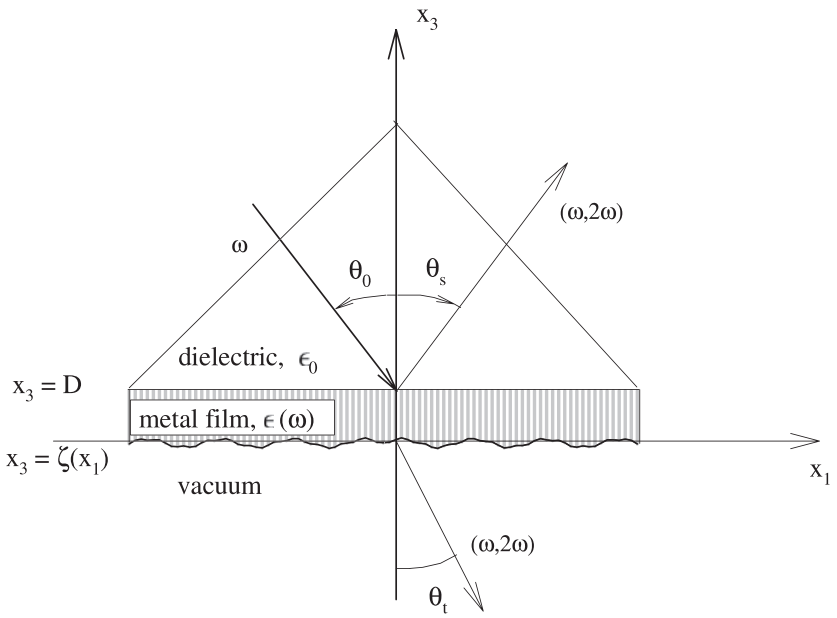

Figure 1. A schematic diagram of the scattering geometry.

negative, since it is in this range that surface plasmon polaritons can exist in this scattering system. The surface profile function $\zeta(x_1)$ is assumed to be a single-valued function of x_1 that is at least twice differentiable, and constitutes a stationary, zero-mean, Gaussian random process defined by the properties

$$\langle \zeta(x_1) \rangle = 0 \tag{2.1a}$$

$$\langle \zeta(x_1)\zeta(x_1') \rangle = \delta^2 W(|x_1 - x_1'|) \tag{2.1b}$$

$$\langle \zeta^2(x_1) \rangle = \delta^2. \tag{2.1c}$$

The angle brackets in equations (2.1) denote an average over the ensemble of realizations of $\zeta(x_1)$, and δ is the RMS height of the surface.

The Fourier integral representation of $\zeta(x_1)$ is

$$\zeta(x_1) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \hat{\zeta}(k)\, \mathrm{e}^{\mathrm{i}kx_1}. \tag{2.2}$$

The Fourier coefficient $\zeta(k)$ is also a zero-mean Gaussian random process defined by

$$\langle \hat{\zeta}(k) \rangle = 0 \tag{2.3a}$$

$$\langle \hat{\zeta}(k)\hat{\zeta}(k') \rangle = 2\pi\delta(k + k')\delta^2 g(|k|) \tag{2.3b}$$

where $g(|k|)$ is the power spectrum of the surface roughness,

$$g(|k|) = \int_{-\infty}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{-\mathrm{i}kx_1} W(|x_1|). \tag{2.4}$$

In the numerical calculations carried out in this paper two forms for the power spectrum $g(|k|)$ will be used. The first is the Gaussian power spectrum,

$$g(|k|) = \sqrt{\pi} a \exp(-a^2k^2/4) \tag{2.5}$$

where the characteristic length a is called the transverse correlation length of the surface roughness. The second power spectrum we use is given by

$$g(|k|) = \frac{\pi h_1}{k_{max}^{(1)} - k_{min}^{(1)}}[\theta(k - k_{min}^{(1)})\theta(k_{max}^{(1)} - k) + \theta(-k - k_{min}^{(1)})\theta(k_{max}^{(1)} + k)]$$
$$+ \frac{\pi h_2}{k_{max}^{(2)} - k_{min}^{(2)}}[\theta(k - k_{min}^{(2)})\theta(k_{max}^{(2)} - k) + \theta(-k - k_{min}^{(2)})\theta(k_{max}^{(2)} + k)] \tag{2.6}$$

where $\theta(z)$ is the Heaviside unit step function, and $h_1 + h_2 = 1$. The manner in which $k_{min}^{(1)}$, $k_{max}^{(1)}$, $k_{min}^{(2)}$, $k_{max}^{(2)}$ are chosen will be described in section 8. Surfaces characterized by a power spectrum of this form have been used in recent experimental [11, 12] and theoretical [10, 13] studies of second harmonic generation in reflection from clean, one-dimensional, random metal surfaces.

3. Surface plasmon polaritons and leaky waves

At each frequency ω in the interval $0 < \omega < \omega_{max}$, where ω_{max} is the solution of $\epsilon_0 + \mathrm{Re}\,\epsilon(\omega) = 0$, the structure depicted in figure 1 supports two p-polarized surface plasmon polaritons, whose wavenumbers are denoted by $q_1(\omega)$ and $q_2(\omega)$, with $q_1(\omega) < q_2(\omega)$. In the absence of surface roughness ($\zeta(x_1) \equiv 0$), these wavenumbers are obtained as the solutions of the dispersion relation

$$[\epsilon_0\beta(q,\omega) + \epsilon(\omega)\beta_p(q,\omega)][\epsilon(\omega)\beta_0(q,\omega) + \beta(q,\omega)]$$
$$= -[\epsilon_0\beta(q,\omega) - \epsilon(\omega)\beta_p(q,\omega)][\epsilon(\omega)\beta_0(q,\omega) - \beta(q,\omega)]\exp[-2\beta(q,\omega)D] \tag{3.1}$$

where $\beta_p(q,\omega) = [q^2 - \epsilon_0(\omega/c)^2]^{1/2}$, with $\mathrm{Re}\,\beta_p(q,\omega) > 0$, $\mathrm{Im}\,\beta_p(q,\omega) < 0$; $\beta(q,\omega) = [q^2 - \epsilon(\omega)(\omega/c)^2]^{1/2}$, with $\mathrm{Re}\,\beta(q,\omega) > 0$, $\mathrm{Im}\,\beta(q,\omega) < 0$; and $\beta_0(q,\omega) = [q^2 - (\omega/c)^2]^{1/2}$, with $\mathrm{Re}\,\beta_0(q,\omega) > 0$, $\mathrm{Im}\,\beta_0(q,\omega) < 0$.

If, for the time being, we neglect the imaginary part of $\epsilon(\omega)$, we find that true surface plasmon polaritons exist only in the region $q > \sqrt{\epsilon_0}(\omega/c)$, where both $\beta_p(q,\omega)$ and $\beta_0(q,\omega)$ are real and positive, so that the electromagnetic field in both the prism and the vacuum decays exponentially with increasing distance from the metal film ($\beta(q,\omega)$ is real and positive because we have assumed that $\epsilon(\omega)$ is real and negative). In the limit $D \to \infty$ the vanishing of the first factor on the left-hand side of equation (3.1) is the dispersion relation for a surface plasmon polariton localized to the prism–metal interface; the vanishing of the second factor is the dispersion relation for a surface plasmon polariton localized to the interface between the metal film and the vacuum.

The solution of equation (3.1) for finite values of D in the frequency range of interest to us yields two real values of q, $q_1(\omega)$ and $q_2(\omega)$. In this frequency range $q_2(\omega) > \sqrt{\epsilon_0(\omega/c)}$, so that both $\beta_0(q_2(\omega))$ and $\beta_p(q_2(\omega))$ are real and positive. The field associated with this mode has its maximum at the prism–metal interface and decays exponentially both into the prism and into the vacuum. This is a true surface plasmon polariton mode, which in the limit of a thick film becomes the surface plasmon polariton at a prism–metal interface.

The solution $q_1(\omega)$, although it is greater than $(\omega/c, q_1(\omega) > (\omega/c)$, is smaller than $\sqrt{\epsilon_0}(\omega/c)$, $q_1(\omega) < \sqrt{\epsilon_0}(\omega/c)$, so that $\beta_0(q_1(\omega))$ is real, while $\beta_p(q_1(\omega))$ is imaginary. The field associated with this mode has its maximum at the metal–vacuum interface and decays exponentially into the vacuum; however, it radiates into the prism. Thus, this mode is a leaky

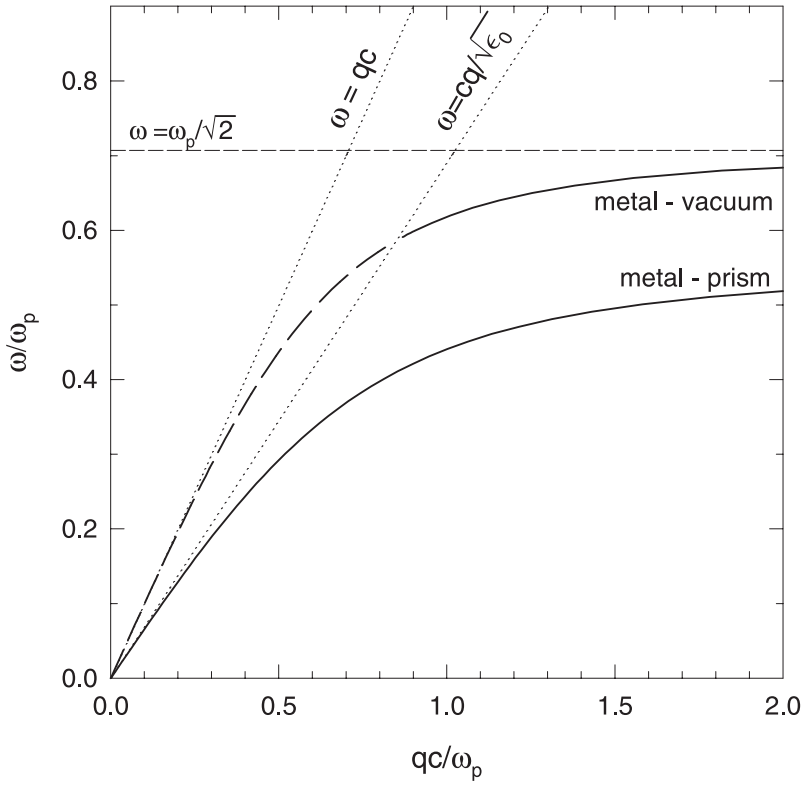

Figure 2. The dispersion curves for the surface polaritons supported by a planar metal film characterized by the dielectric function $\epsilon(\omega) = 1 - \omega_p^2/\omega^2$. The medium above the film is a dielectric ($\epsilon_0 = 2.25$). The medium below the film is vacuum. The leaky mode is depicted by the broken curve.

wave. In the limit of a thick film this mode becomes the surface plasmon polariton at a metal–vacuum interface. The fact that for finite values of D this mode is a leaky wave makes it possible to be excited in the Kretschmann ATR configuration.

The dispersion curves resulting from the solution of equation (3.1) are plotted in figure 2 for the case where $\epsilon_0 = 2.25$ and $\epsilon(\omega)$ has the simple free electron form $\epsilon(\omega) = 1 - (\omega_p/\omega)^2$, where ω_p is the plasma frequency of the conduction electron in the metal. The portion of one branch of the dispersion curve that corresponds to the leaky wave is depicted by the broken curve, and represents a plot of ω as a function of $\mathrm{Re}\, q$.

When the imaginary part of the dielectric function is restored, i.e. when $\epsilon(\omega)$ is now taken to be complex, both wavenumbers $q_1(\omega)$ and $q_2(\omega)$ become complex. Their imaginary parts are generally small enough compared with their real parts that the dispersion curves plotted in figure 2 are almost unchanged, except that the entire horizontal axis must now be understood to be labelled as $\mathrm{Re}\, q$.

The fact that the scattering system depicted in figure 1 supports two surface plasmon polaritons has the consequence that a pair of satellite peaks can occur in the angular distribution of the intensity of the light scattered incoherently from, or transmitted incoherently through, this system [26–28]. If we denote by $q_1(\omega)$ and $q_2(\omega)$ the real parts of the wavenumbers of the two surface plasmon polaritons of frequency ω supported by the scattering system, an analysis of the kind carried out in [26–28] predicts that in scattering the two satellite peaks occur at the

scattering angle θ_s given by

$$\sin\theta_s = -\sin\theta_0 \pm \frac{c}{\sqrt{\epsilon_0}\omega}[q_1(\omega) - q_2(\omega)] \tag{3.2}$$

where θ_0 is the angle of incidence (figure 1), while in transmission the two satellite peaks occur at angles of transmission θ_t (figure 1) given by

$$\sin\theta_t = -\sqrt{\epsilon_0}\sin\theta_0 \pm \frac{c}{\omega}[q_1(\omega) - q_1(\omega)]. \tag{3.3}$$

These satellite peaks in scattering and transmission are multiple-scattering phenomena. It should be noted that among the satellite peaks in scattering and transmission that should occur in the corresponding radiative region of the optical spectrum, not all may be intense enough to be observable. This is especially the case in scattering where the light has to pass through the metal film twice, and is attenuated during each pass.

The fact that one of the branches of the dispersion curves obtained from the solution of equation (3.1) is that of a leaky wave for a certain range of frequencies (figure 2), gives rise to the possibility of additional peaks in the angular distribution of the intensity of the incoherent component of the reflected light. In contrast with the satellite peaks discussed above, these leaky wave peaks are single-scattering phenomena, and hence are generally more intense than the former. To estimate the scattering angles at which the leaky wave peaks occur we use the simple argument that the component of the wavevector of the scattered light parallel to the mean scattering surface is equal in magnitude to the wavenumber of the leaky surface plasmon polariton,

$$\mathrm{Re}[q_1(\omega)] = \pm\sqrt{\epsilon_0}\frac{\omega}{c}\sin\theta_{leaky}. \tag{3.4}$$

The angles at which the leaky wave peaks occur are therefore independent of the angle of incidence θ_0. We note that because $\mathrm{Re}[q_1(\omega)]$ is smaller than $\sqrt{\epsilon_0}(\omega/c)$ but larger than (ω/c), it follows that there are no leaky peaks in transmission, because in the transmission region the dielectric constant is unity, and the corresponding peaks lie in the non-radiative region of the optical spectrum. For the same reason, leaky wave peaks are absent in the scattering of light from a free-standing metal film ($\epsilon_0 = 1$). Because the leaky wave peaks give a significant contribution to the angular distribution of the intensity of the scattered light, we may expect that multiple-scattering effects will manifest themselves more clearly under conditions where the leaky wave peaks are absent.

4. Formulation of the scattering problem

Monochromatic p-polarized light of frequency ω, whose plane of incidence is the x_1x_3-plane, illuminates the prism–metal interface $x_3 = D$ from the side of the prism. In this case it is convenient to work with the single non-zero component of the magnetic vector at the fundamental frequency ω. It is assumed that the nonlinearities that give rise to the second harmonic field arise in a thin region containing the prism–metal interface, and in a thin region containing the metal–vacuum interface. The nonlinearity of the metal therefore enters the problem only through the boundary conditions at these interfaces at the harmonic frequency. It is also implicit in our formulation of the problem that the harmonic field, being small, has no influence on the fields or the scattering problem at the fundamental frequency. Then the x_2-components of the magnetic fields at ω and 2ω satisfy Helmholtz equations in each of the regions $x_3 > D$ (region I), $\zeta(x_1) < x_3 < D$ (region II) and $x_3 < \zeta(x_1)$ (region III). By

applying Green's second integral identity in the plane [29] to each of these three regions in turn, we obtain the following equations:

$$\theta(x_3 - D)H_2^{(I)}(x_1, x_3|\Omega) = H_0(x_1, x_3|\Omega) + \frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \times \left\{ \left[\frac{\partial}{\partial x_3'} G_{\epsilon_0}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{x_3'=D} H^{(I)}(x_1'|\Omega) - [G_{\epsilon_0}^{(\Omega)}(x_1, x_3|x_1', x_3')]_{x_3'=D} L^{(I)}(x_1'|\Omega) \right\} \tag{4.1}$$

$$\theta(x_3 - \zeta(x_1))\theta(D - x_3)H_2^{(II)}(x_1, x_3|\Omega) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial x_3'} G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right] H_2^{(II)}(x_1', x_3'|\Omega) - G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \frac{\partial}{\partial x_3'} H_2^{(II)}(x_1', x_3'|\Omega) \right\}_{x_3'=D} + \frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial N'} G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right] H_2^{(II)}(x_1', x_3'|\Omega) - G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \frac{\partial}{\partial N'} H_2^{(II)}(x_1', x_3'|\Omega) \right\}_{x_3'=\zeta(x_1')} \tag{4.2}$$

$$\theta(\zeta(x_1) - x_3)H_2^{(III)}(x_1, x_3|\Omega) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial N'} G_0^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{x_3'=\zeta(x_1')} H^{(III)}(x_1'|\Omega) - \left[G_0^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{x_3'=\zeta(x_1')} L^{(III)}(x_1'|\Omega) \right\} \tag{4.3}$$

respectively. In these equations Ω denotes either ω or 2ω, $\theta(z)$ is the Heaviside unit step function,

$$\frac{\partial}{\partial N} = -\zeta'(x_1)\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \tag{4.4}$$

and

$$H_0(x_1, x_3|\Omega) = H_2^{(I)}(x_1, x_3|\omega)_{inc} \qquad \Omega = \omega \tag{4.5a}$$

$$= 0 \qquad \Omega = 2\omega. \tag{4.5b}$$

The source functions $H^{(I)}(x_1|\Omega)$, $L^{(I)}(x_1|\Omega)$, $H^{(III)}(x_1|\Omega)$, and $L^{(III)}(x_1|\Omega)$ are defined by

$$H^{(I)}(x_1|\Omega) = H_2^{(I)}(x_1, x_3|\Omega)\Big|_{x_3=D} \tag{4.6a}$$

$$L^{(I)}(x_1|\Omega) = \frac{\partial}{\partial x_3} H_2^{(I)}(x_1, x_3|\Omega)\Big|_{x_3=D} \tag{4.6b}$$

$$H^{(III)}(x_1|\Omega) = H_2^{(III)}(x_1, x_3|\Omega)\Big|_{x_3=\zeta(x_1)} \tag{4.6c}$$

$$L^{(III)}(x_1|\Omega) = \frac{\partial}{\partial N} H_2^{(III)}(x_1, x_3|\Omega)|_{x_3=\zeta(x_1)}. \tag{4.6d}$$

The functions $G_{\epsilon_0}^{(\Omega)}(x_1, x_3|x_1', x_3')$, $G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3')$ and $G_{0}^{(\Omega)}(x_1, x_3|x_1', x_3')$ are, respectively, the Green functions for a dielectric medium characterized by a dielectric constant ϵ_0, for a metal characterized by a dielectric function $\epsilon(\Omega)$, and for vacuum, at frequencies ω and 2ω. They are given explicitly by

$$G_{\epsilon_0}^{(\Omega)}(x_1, x_3|x_1', x_3') = \int_{-\infty}^{\infty} \frac{\mathrm{d}q}{2\pi} \frac{2\pi \mathrm{i}}{\alpha_p(q, \Omega)} \exp[\mathrm{i}q(x_1 - x_1') + \mathrm{i}\alpha_p(q, \Omega)|x_3 - x_3'|] \tag{4.7a}$$

$$= \mathrm{i}\pi H_0^{(1)}\left(\epsilon_0^{1/2} \frac{\Omega}{c}[(x_1 - x_1')^2 + (x_3 - x_3')^2]^{1/2}\right) \tag{4.7b}$$

$$G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') = \int_{-\infty}^{\infty} \frac{\mathrm{d}q}{2\pi} \frac{2\pi}{\beta(q, \Omega)} \exp[\mathrm{i}q(x_1 - x_1') - \beta(q, \Omega)|x_3 - x_3'|] \tag{4.8a}$$

$$= \mathrm{i}\pi H_0^{(1)}(n_\epsilon(\Omega) \frac{\Omega}{c}[(x_1 - x_1')^2 + (x_3 - x_3')^2]^{1/2}) \tag{4.8b}$$

$$G_{0}^{(\Omega)}(x_1, x_3|x_1', x_3') = \int_{-\infty}^{\infty} \frac{\mathrm{d}q}{2\pi} \frac{2\pi \mathrm{i}}{\alpha_0(q, \Omega)} \exp[\mathrm{i}q(x_1 - x_1') + \mathrm{i}\alpha_0(q, \Omega)|x_3 - x_3'|] \tag{4.9a}$$

$$= \mathrm{i}\pi H_0^{(1)}\left(\frac{\Omega}{c}[(x_1 - x_1')^2 + (x_3 - x_3')^2]^{1/2}\right) \tag{4.9b}$$

where

$$\alpha_p(q, \Omega) = \left(\epsilon_0 \frac{\Omega^2}{c^2} - q^2\right)^{1/2} \qquad q^2 < \epsilon_0 \frac{\Omega^2}{c^2} \tag{4.10a}$$

$$= \mathrm{i}\left(q^2 - \epsilon_0 \frac{\Omega^2}{c^2}\right)^{1/2} \qquad q^2 > \epsilon_0 \frac{\Omega^2}{c^2} \tag{4.10b}$$

$$\alpha_0(q, \Omega) = \left(\frac{\Omega^2}{c^2} - q^2\right)^{1/2} \qquad q^2 < \frac{\Omega^2}{c^2} \tag{4.11a}$$

$$= \mathrm{i}\left(q^2 - \frac{\Omega^2}{c^2}\right)^{1/2} \qquad q^2 > \frac{\Omega^2}{c^2} \tag{4.11b}$$

and

$$\beta(q, \Omega) = \left(q^2 - \epsilon(\Omega) \frac{\Omega^2}{c^2}\right)^{1/2} \qquad \mathrm{Re}\,\beta(q, \Omega) > 0 \qquad \mathrm{Im}\,\beta(q, \Omega) < 0 \tag{4.12}$$

while

$$n_\epsilon(\Omega) = (\epsilon(\Omega))^{1/2} \qquad \mathrm{Re}\, n_\epsilon(\Omega) > 0 \qquad \mathrm{Im}\, n_\epsilon(\Omega) > 0 \tag{4.13}$$

and $H_0^{(1)}(z)$ is a Hankel function of the first kind.

The scattered field in the prism is given by the second term on the right-hand side of equation (4.1):

$$H_2^{(I)}(x_1, x_2|\Omega)_{sc} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial x_3'} G_{\epsilon_0}^{(\Omega)}(x_1, x_3|x_1', x_3')\right]_{x_3'=D} H^{(I)}(x_1'|\Omega) \right.$$

$$\left. - \left[G_{\epsilon_0}^{(\Omega)}(x_1, x_3|x_1', x_3')\right]_{x_3'=D} L^{(I)}(x_1'|\Omega) \right\}. \tag{4.14}$$

The transmitted field in the vacuum is given by the right-hand side of equation (4.3):

$$H_2^{(III)}(x_1, x_3|\Omega)_{tr} = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial N'} G_0^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{x_3'=\zeta(x_1')} H^{(III)}(x_1'|\Omega) \right.$$
$$\left. - \left[G_0^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{x_3'=\zeta(x_1')} L^{(III)}(x_1'|\Omega) \right\}. \tag{4.15}$$

To obtain the equations satisfied by the source functions $H^{(I,III)}(x_1|\Omega)$ and $L^{(I,III)}(x_1|\Omega)$ we proceed as follows. By setting $x_3 = D + \eta$, where η is a positive infinitesimal, in equations (4.1) and (4.2), and using equations (4.6), we obtain the following pair of equations:

$$H^{(I)}(x_1|\Omega) = H_0(x_1|\Omega) + \frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial x_3'} G_{e_0}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=D+\eta \\ x_3'=D}} \right.$$
$$\left. \times H^{(I)}(x_1'|\Omega) - \left[G_{\epsilon_0}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=D+\eta \\ x_3'=D}} L^{(I)}(x_1'|\Omega) \right\} \tag{4.16}$$

$$0 = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial x_3'} G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=D+\eta \\ x_3'=D}} \left[H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=D} \right.$$
$$- \left[G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=D+\eta \\ x_3'=D}} \left[\frac{\partial}{\partial x_3'} H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=D} \Bigg\}$$
$$+ \frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial N'} G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=D \\ x_3'=\zeta(x_1')}} \left[H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=\zeta(x_1')} \right.$$
$$\left. - \left[G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=D \\ x_3'=\zeta(x_1')}} \left[\frac{\partial}{\partial N'} H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=\zeta(x_1')} \right\} \tag{4.17}$$

where

$$H_0(x_1|\Omega) = H_0(x_1, x_3|\Omega)|_{x_3=D}. \tag{4.18}$$

We obtain two more equations by setting $x_3 = \zeta(x_1) + \eta$ in equations (4.2) and (4.3):

$$H_2^{(II)}(x_1, x_3|\Omega)\Big|_{x_3=\zeta(x_1)} = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial x_3'} G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=\zeta(x_1) \\ x_3'=D}} \right.$$
$$\times \left[H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=D} - \left[G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=\zeta(x_1) \\ x_3'=D}}$$
$$\left. \times \left[\frac{\partial}{\partial x_3'} H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=D} \right\}$$
$$+ \frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial N'} G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}} \right.$$
$$\times \left[H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=\zeta(x_1')} - \left[G_{\epsilon}^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}}$$
$$\left. \times \left[\frac{\partial}{\partial N'} H_2^{(II)}(x_1', x_3'|\Omega) \right]_{x_3'=\zeta(x_1')} \right\} \tag{4.19}$$

$$0 = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}x_1' \left\{ \left[\frac{\partial}{\partial N'} G_0^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}} H^{(III)}(x_1'|\Omega) \right.$$
$$\left. - \left[G_0^{(\Omega)}(x_1, x_3|x_1', x_3') \right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}} L^{(III)}(x_1'|\Omega) \right\}. \tag{4.20}$$

The four equations (4.16), (4.17) and (4.19), (4.20) provide the starting point for the calculations that yield the source functions. We now consider the fields at ω and 2ω in turn.

5. The linear scattering equations

The continuity of the tangential components of the electric and magnetic fields across the surfaces $x_3 = D$ and $x_3 = \zeta(x_1)$ at frequency ω can be expressed as

$$H_2^{(I)}(x_1, x_3|\omega)|_{x_3=D} = H_2^{(II)}(x_1, x_3|\omega)|_{x_3=D} \tag{5.1a}$$

$$\frac{1}{\epsilon_0}\frac{\partial}{\partial x_3} H_2^{(I)}(x_1, x_3|\omega)|_{x_3=D} = \frac{1}{\epsilon(\omega)}\frac{\partial}{\partial x_3} H_2^{(II)}(x_1, x_3|\omega)|_{x_3=D} \tag{5.1b}$$

and

$$H_2^{(II)}(x_1, x_3|\omega)|_{x_3=\zeta(x_1)} = H_2^{(III)}(x_1, x_3|\omega)|_{x_3=\zeta(x_1)} \tag{5.2a}$$

$$\frac{1}{\epsilon(\omega)}\frac{\partial}{\partial N} H_2^{(II)}(x_1, x_3|\omega)|_{x_3=\zeta(x_1)} = \frac{\partial}{\partial N} H_2^{(III)}(x_1, x_3|\omega)|_{x_3=\zeta(x_1)} \tag{5.2b}$$

respectively. When these boundary conditions are used in the integral equations (4.16), (4.17) and (4.19), (4.20) with $\Omega = \omega$, and the definitions of the source functions, equations (4.6), for $\Omega = \omega$ are used, we obtain the following set of integral equations for the four source functions $H^{(I)}(x_1|\omega)$, $L^{(I)}(x_1|\omega)$, $H^{(III)}(x_1|\omega)$, and $L^{(III)}(x_1|\omega)$:

$$H^{(I)}(x_1|\omega) = H(x_1|\omega)_{inc} + \int_{-\infty}^{\infty} \mathrm{d}x_1' \left[H_p^{DD}(x_1|x_1') H^{(I)}(x_1'|\omega) \; - L_p^{DD}(x_1|x_1') L^{(I)}(x_1'|\omega) \right] \tag{5.3}$$

$$0 = \int_{-\infty}^{\infty} \mathrm{d}x_1' \left[-H_\epsilon^{DD}(x_1|x_1') H^{(I)}(x_1'|\omega) + \frac{\epsilon(\omega)}{\epsilon_0} L_\epsilon^{DD}(x_1|x_1') L^{(I)}(x_1'|\omega) \right.$$
$$\left. + H_\epsilon^{DO}(x_1|x_1') H^{(III)}(x_1'|\omega) - \epsilon(\omega) L_\epsilon^{DO}(x_1|x_1') L^{(III)}(x_1'|\omega) \right] \tag{5.4}$$

$$H^{(III)}(x_1|\omega) = \int_{-\infty}^{\infty} \mathrm{d}x_1' \left[-H_\epsilon^{OD}(x_1|x_1') H^{(I)}(x_1'|\omega) + \frac{\epsilon(\omega)}{\epsilon_0} L_\epsilon^{OD}(x_1|x_1') L^{(I)}(x_1'|\omega) \right.$$
$$\left. + H_\epsilon^{OO}(x_1|x_1') H^{(III)}(x_1'|\omega) - \epsilon(\omega) L_\epsilon^{OO}(x_1|x_1') L^{(III)}(x_1'|\omega) \right] \tag{5.5}$$

$$0 = \int_{-\infty}^{\infty} \mathrm{d}x_1' \left[-H_0^{OO}(x_1|x_1') H^{(III)}(x_1'|\omega) + L_0^{OO}(x_1|x_1') L^{(III)}(x_1'|\omega) \right]. \tag{5.6}$$

The kernels appearing in these equations are written out explicitly in appendix A. Equations (5.3)–(5.6) are solved by converting them into matrix equations. The way in which this has been done in this paper is outlined in appendix B.

For the incident field we use the superposition of an infinite number of incoming plane waves,

$$H_2^{(I)}(x_1, x_3|\omega)_{inc} = \sqrt{\epsilon_0}\frac{\omega}{c}\frac{w}{2\sqrt{\pi}}\int_{-\pi/2}^{\pi/2} \mathrm{d}\theta \exp\left[-\epsilon_0\frac{\omega^2 w^2}{4c^2}(\theta-\theta_0)^2\right] \times \exp\left[\mathrm{i}\sqrt{\epsilon_0}\frac{\omega}{c}(x_1\sin\theta - (x_3 - D)\cos\theta)\right]. \tag{5.7}$$

In the limit $\sqrt{\epsilon_0}(\omega w/2c) \gg 1$ this expression defines a Gaussian beam of width $2w$, whose angle of incidence is θ_0. The total time-averaged flux incident on the surface $x_3 = D$ is

$$P_{inc} = \left|\int_{-\infty}^{\infty}\mathrm{d}x_1\int_{-L_2/2}^{L_2/2}\mathrm{d}x_2\,\mathrm{Re}[S_3^c]_{inc}\right| = L_2\frac{cw}{16\sqrt{2\pi\epsilon_0}}\left\{\mathrm{Erf}\left(\sqrt{\epsilon_0}\frac{\omega w}{\sqrt{2}c}\left(\frac{\pi}{2}-\theta_0\right)\right)+\mathrm{Erf}\left(\sqrt{\epsilon_0}\frac{\omega w}{\sqrt{2}c}\left(\frac{\pi}{2}+\theta_0\right)\right)\right\} \tag{5.8}$$

where S_3^c is the three-component of the complex Poynting vector, L_2 is the length of the surface along the x_2-axis, and $\mathrm{Erf}(z)$ is the error function.

With the use of the integral representation (4.7a) of the Green function $G_{\epsilon_0}^{(\omega)}(x_1, x_3|x_1', x_3')$ we can write the amplitude of the scattered field (4.14) in the far field in the form

$$H_2^{(I)}(x_1, x_3|\omega)_{sc} = \int_{-\infty}^{\infty}\frac{\mathrm{d}q}{2\pi}R(q,\omega)\exp[\mathrm{i}qx_1 + \mathrm{i}\alpha_p(q,\omega)x_3] \tag{5.9}$$

where

$$R(q,\omega) = \frac{1}{2\mathrm{i}\alpha_p(q,\omega)}\int_{-\infty}^{\infty}\mathrm{d}x_1\exp[-\mathrm{i}qx_1 - \mathrm{i}\alpha_p(q,\omega)D] \times[\mathrm{i}\alpha_p(q,\omega)H^{(I)}(x_1|\omega) + L^{(I)}(x_1|\omega)]. \tag{5.10}$$

The total time-averaged scattered flux crossing the plane $x_3 = \text{constant} > D$ is

$$P_{sc}^{(\omega)} = \int_{-\infty}^{\infty}\mathrm{d}x_1\int_{-L_2/2}^{L_2/2}\mathrm{d}x_2\,\mathrm{Re}[S_3^c]_{sc} = \int_{-\pi/2}^{\pi/2}\mathrm{d}\theta_s P_{sc}(\theta_s|\omega) \tag{5.11}$$

where

$$P_{sc}(\theta_s|\omega) = L_2\frac{c^2}{64\pi^2\epsilon_0\omega}|r(\theta_s|\omega)|^2 \tag{5.12}$$

and

$$r(\theta_s|\omega) = 2\mathrm{i}\sqrt{\epsilon_0}\frac{\omega}{c}\cos\theta_s R\left(\sqrt{\epsilon_0}\frac{\omega}{c}\sin\theta_s, \omega\right). \tag{5.13}$$

In a similar fashion, with the use of the integral representation (4.9a) of the Green function $G_0^{(\omega)}(x_1, x_3|x_1', x_3')$ we can write the amplitude of the transmitted field (4.15) as

$$H_2^{(III)}(x_1, x_3|\omega)_{tr} = \int_{-\infty}^{\infty}\frac{\mathrm{d}q}{2\pi}T(q,\omega)\exp[\mathrm{i}qx_1 - \mathrm{i}\alpha_0(q,\omega)x_3] \tag{5.14}$$

where

$$T(q,\omega) = \frac{1}{2\mathrm{i}\alpha_0(q,\omega)}\int_{-\infty}^{\infty}\mathrm{d}x_1\exp[-\mathrm{i}qx_1 + \mathrm{i}\alpha_0(q,\omega)\zeta(x_1)] \times[\mathrm{i}(q\zeta'(x_1) + \alpha_0(q,\omega))H^{(III)}(x_1|\omega) - L^{(III)}(x_1|\omega)]. \tag{5.15}$$

The total time-averaged flux crossing the plane $x_3 = \text{constant} < \zeta(x_1)_{\min}$ is

$$P_{tr}^{(\omega)} = \int_{-\infty}^{\infty} \mathrm{d}x_1 \int_{-L_2/2}^{L_2/2} \mathrm{d}x_2 \, \mathrm{Re}[S_3^c]_{tr}$$
$$= \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta_t \, P_{tr}(\theta_t|\omega) \tag{5.16}$$

where

$$P_{tr}(\theta_t|\omega) = L_2 \frac{c^2}{64\pi^2\omega} |t(\theta_t|\omega)|^2 \tag{5.17}$$

and

$$t(\theta_t|\omega) = 2\mathrm{i}\frac{\omega}{c} \cos\theta_t \, T\left(\frac{\omega}{c}\sin\theta_t, \omega\right). \tag{5.18}$$

The differential reflection coefficient, defined as the fraction of the total incident power that is scattered per unit angle, is

$$\frac{\partial R(\theta_s|\omega)}{\partial \theta_s} = \frac{P_{\mathrm{sc}}(\theta_s|\omega)}{P_{inc}} = L_2 \frac{c^2}{64\pi^2\epsilon_0\omega} \frac{|r(\theta_s|\omega)|^2}{P_{inc}}. \tag{5.19}$$

Similarly, the differential transmission coefficient, defined as the fraction of the total incident power that is transmitted per unit angle, is

$$\frac{\partial T(\theta_t|\omega)}{\partial \theta_t} = \frac{P_{tr}(\theta_t|\omega)}{P_{inc}} = L_2 \frac{c^2}{64\pi^2\omega} \frac{|t(\theta_t|\omega)|^2}{P_{inc}}. \tag{5.20}$$

From these results we find that the contributions to the mean differential reflection and transmission coefficients from the incoherent component of the scattered and transmitted light are given by

$$\left\langle \frac{\partial R(\theta_s|\omega)}{\partial \theta_s} \right\rangle_{incoh} = L_2 \frac{c^2}{64\pi^2\epsilon_0\omega} \frac{\langle |r(\theta_s|\omega)|^2\rangle - |\langle r(\theta_s|\omega)\rangle|^2}{P_{inc}} \tag{5.21}$$

$$\left\langle \frac{\partial T(\theta_t|\omega)}{\partial \theta_t} \right\rangle_{\mathrm{incoh}} = L_2 \frac{c^2}{64\pi^2\omega} \frac{\langle |t(\theta_t|\omega)|^2\rangle - |\langle t(\theta_t|\omega)\rangle|^2}{P_{inc}}. \tag{5.22}$$

6. Nonlinear boundary conditions

Homogeneous and isotropic metals possess inversion symmetry. Therefore, the dipolar contribution to the bulk nonlinear polarization is absent ($\chi^{(2)} = 0$). The presence of a surface breaks the inversion symmetry and, since both the electromagnetic fields and material constants vary rapidly at the interfaces, their gradients give rise to the optical nonlinearity of each of the interfaces. Therefore, the second harmonic radiation we are interested in is generated in a prism–metal interface layer and in a metal–vacuum interface layer that have finite thicknesses on the microscopic scale. Consequently, the resulting nonlinear polarization can be taken into account through the boundary conditions for the second harmonic fields. In what follows we will neglect the small contribution to the nonlinearity coming from the bulk and the possible anisotropy of the material constants. We also assume that the metal film has a macroscopic thickness, so that the interface layers are well separated. Several models of the surface nonlinear polarization have been discussed in the literature [30–36]. In this paper we derive the nonlinear

boundary conditions for the general form of the nonlinear polarization in a centrosymmetric metal proposed in [33], which in the case of a metal film is

$$\boldsymbol{P}^{NL}_{I,III} = \frac{1}{4\pi}[\alpha_{I,III}\boldsymbol{E}(\nabla\cdot\boldsymbol{E}) + \beta_{I,III}(\boldsymbol{E}\cdot\nabla)\boldsymbol{E} + \gamma_{I,III}\boldsymbol{E}\times(\nabla\times\boldsymbol{E}) + \boldsymbol{E}(\boldsymbol{E}\cdot\nabla\rho_{I,III}) + (\boldsymbol{E}\cdot\boldsymbol{E})\nabla\kappa_{I,III}] \tag{6.1}$$

where the subscripts I and III denote the prism–metal and metal–vacuum interfaces, respectively, the coefficients, in general, depend on the distance from the corresponding interface and $\boldsymbol{E}$ is the macroscopic electric field. The coefficients entering the expression (6.1) can be related to the nonlinear parameters of the other models.

Both material constants and electromagnetic fields vary strongly across the interface layers. However, outside the interface layers the fields are finite. This implies that in spite of the singular behaviour of the nonlinear polarization, the tangential components of the electric and magnetic fields and the normal components of the displacement vector must be finite. The latter, in turn, implies a singular behaviour of the normal component of the electric field across the interface layer. The nonlinear boundary conditions for the 2ω fields are obtained by integrating the Maxwell equations for them across the interface layers, and then passing to the limit of vanishing thicknesses of the layers.

We first derive the nonlinear boundary conditions for the rough metal–vacuum interface, $x_3 = \zeta(x_1)$. We introduce a local coordinate system with unit vectors $\{\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}\}$, where $\hat{\boldsymbol{x}} = (1, 0, \zeta'(x_1))/\phi(x_1)$ and $\hat{\boldsymbol{z}} = (-\zeta'(x_1), 0, 1)/\phi(x_1)$ are unit vectors tangent and normal to the interface in the plane perpendicular to its generators, $\phi(x_1) = [1 + (\zeta'(x_1))^2]^{1/2}$, and $\hat{\boldsymbol{y}}$ is the unit vector along the x_2-axis. In this coordinate system all material parameters, $\epsilon(\omega)$, $\epsilon(2\omega)$, α, β, γ, ρ and κ, depend only on the distance from the interface along the normal to it, z. Integrating the tangential component of the equation $\nabla\times\boldsymbol{H} = -(2\mathrm{i}\omega/c)\boldsymbol{D}$, namely $\partial\boldsymbol{H}_t/\partial z = (2\mathrm{i}\omega/c)\hat{\boldsymbol{z}}\times\boldsymbol{D}_t + \nabla_t H_z$, along a contour around the metal–vacuum interface, we obtain with the use of the relation $\boldsymbol{D} = \epsilon(2\omega, z)\boldsymbol{E} + 4\pi\boldsymbol{P}^{NL}$

$$\lim_{\eta\to 0}[\boldsymbol{H}_t(x,\eta|2\omega) - \boldsymbol{H}_t(x,-\eta|2\omega)] = \lim_{\eta\to 0}\int_{-\eta}^{+\eta}\mathrm{d}z\left\{\frac{2\mathrm{i}\omega}{c}\hat{\boldsymbol{z}}\times[\epsilon(2\omega,z)\boldsymbol{E}_t(x,z|2\omega) + 4\pi\boldsymbol{P}^{NL}_t(x,z|2\omega)] + \frac{\partial}{\partial\boldsymbol{t}}H_z(x,z|2\omega)\right\} \tag{6.2}$$

where the subscript t denotes the components tangential to the local interface and $\partial/\partial\boldsymbol{t} = (\partial/\partial x, \partial/\partial y, 0)$.

The non-vanishing contribution to the integral on the right-hand side of equation (6.2) in the limit $\eta\to 0$ comes from the most singular part of $\boldsymbol{P}^{NL}_t$. Since the normal component of the electric field and the nonlinear material parameters are discontinuous across the interface, their derivatives with respect to the normal to the interface determine the unbounded, most singular, part of the nonlinear polarization. Therefore, from equation (6.1) it follows that the unbounded part of the tangential component $\boldsymbol{P}^{NL}_t$ has the form

$$\boldsymbol{P}^{NL}_t = \frac{1}{4\pi}\boldsymbol{E}_t(x,z|\omega)D_z(x,z|\omega)\left(\alpha_{III}(z)\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{\epsilon(\omega,z)} + \frac{1}{\epsilon(\omega,z)}\frac{\mathrm{d}\rho_{III}(z)}{\mathrm{d}z}\right) \tag{6.3}$$

where $\epsilon(\omega, z>\eta) = \epsilon(\omega)$ in the metal and $\epsilon(\omega, z<-\eta) = 1$ in vacuum. Since $\boldsymbol{E}_t$ and D_z are finite they can be evaluated at the surface and removed from under the integral sign. As a result we obtain

$$\boldsymbol{H}^{(II)}_t(x,z|2\omega)_{|z=0} - \boldsymbol{H}^{(III)}_t(x,z|2\omega)_{|z=0} = \frac{2\mathrm{i}\omega}{c}\mu_3^{(III)}\hat{\boldsymbol{z}}\times\boldsymbol{E}_t(x,z|\omega)D_z(x,z|\omega)_{|z=0} \tag{6.4}$$

where the superscripts (III) and (II) denote the fields in vacuum and in the metal film, respectively, and the phenomenological nonlinear constant $\mu_3^{(III)}$ is given by

$$\mu_3^{(III)} = \lim_{\eta\to 0}\int_{-\eta}^{\eta} \mathrm{d}z \left[\alpha_{III}(z)\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{\epsilon(\omega, z)} + \frac{1}{\epsilon(\omega, z)}\frac{\mathrm{d}\rho_{III}(z)}{\mathrm{d}z}\right]. \tag{6.5}$$

The field components entering the nonlinear source on the right-hand side of equation (6.4) are continuous across the local interface $z = 0$ and can be evaluated either in the film or in the vacuum.

The analogous boundary condition for the tangential component of the magnetic field across the planar metal–prism interface, $x_3 = D$, is obtained in the same manner, with the result that

$$\begin{aligned}\boldsymbol{H}_t^{(I)}(x_1, x_3|2\omega)_{|x_3=D} - \boldsymbol{H}_t^{(II)}(x_1, x_3|2\omega)_{|x_3=D}\\ = \frac{2\mathrm{i}\omega}{c}\mu_3^{(I)}\hat{\boldsymbol{x}}_3 \times \boldsymbol{E}_t(x_1, x_3|\omega)D_3(x_1, x_3|\omega)_{|x_3=D}\end{aligned} \tag{6.6}$$

where $\hat{\boldsymbol{x}}_3$ is the unit vector normal to the interface directed into the prism and the superscript (I) denotes the fields in the prism. The phenomenological nonlinearity constant $\mu_3^{(I)}$ is given by

$$\mu_3^{(I)} = \lim_{\eta\to 0}\int_{D-\eta}^{D+\eta} \mathrm{d}x_3 \left[\alpha_I(x_3)\frac{\mathrm{d}}{\mathrm{d}x_3}\frac{1}{\epsilon(\omega, x_3)} + \frac{1}{\epsilon(\omega, x_3)}\frac{\mathrm{d}\rho_I(x_3)}{\mathrm{d}x_3}\right] \tag{6.7}$$

and $\epsilon(\omega, x_3 < D - \eta) = \epsilon(\omega)$ in the metal film and $\epsilon(\omega, z > D + \eta) = \epsilon_0$ in the prism.

The second boundary condition at the metal–vacuum interface $x_3 = \zeta(x_1)$ is obtained by integrating the tangential component of the equation $\nabla \times \boldsymbol{E} = (2\mathrm{i}\omega/c)\boldsymbol{H}$, namely $\partial \boldsymbol{E}_t/\partial z = -(2\mathrm{i}\omega/c)\hat{\boldsymbol{z}} \times \boldsymbol{H}_t + \nabla_t E_z$, around the metal–vacuum interface:

$$\begin{aligned}\lim_{\eta\to 0}[\boldsymbol{E}_t(x, \eta|2\omega) - \boldsymbol{E}_t(x, -\eta|2\omega)]\\ = \lim_{\eta\to 0}\int_{-\eta}^{\eta} \mathrm{d}z \left(-\frac{2\mathrm{i}\omega}{c}\hat{\boldsymbol{z}} \times \boldsymbol{H}_t(x, z|2\omega) + \frac{\partial}{\partial \boldsymbol{t}}E_z(x, z|2\omega)\right).\end{aligned} \tag{6.8}$$

The normal component of the electric field can be singular at the interface. To exclude it we use the relation $E_z(x, z|2\omega) = [D_z(x, z|2\omega) - 4\pi P_z^{NL}(x, z|2\omega)]/\epsilon(2\omega, z)$, where $\epsilon(2\omega, z < -\eta) = 1$ in the vacuum and $\epsilon(2\omega, z > \eta) = \epsilon(2\omega)$ in the metal. Since the tangential component of the magnetic field and the normal component of the displacement are finite across the interface layer, the non-vanishing contribution to the integral on the right-hand side of equation (6.8) as $\eta \to 0$ comes from the most strongly varying part of the normal component of the surface nonlinear polarization. From equation (6.1) it follows that this is given by

$$\begin{aligned}P_z^{NL}(x, z|2\omega) = \frac{1}{4\pi}\left[D_z^2(x, z|\omega)\left(\frac{1}{2}(\alpha_{III}(z) + \beta_{III}(z))\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{\epsilon^2(\omega, z)}\right.\right.\\ \left.\left. + \frac{1}{\epsilon^2(\omega, z)}\frac{\mathrm{d}}{\mathrm{d}z}(\rho_{III}(z) + \kappa_{III}(z))\right) + \boldsymbol{E}_t^2(x, z|\omega)\frac{\mathrm{d}\kappa_{III}(z)}{\mathrm{d}z}\right].\end{aligned} \tag{6.9}$$

On evaluating the continuous fields at the interface and removing them from under the integral sign we obtain

$$\boldsymbol{E}_t^{(II)}(x, z|2\omega)_{|z=0} - \boldsymbol{E}_t^{(III)}(x, z|2\omega)_{|z=0} = -\frac{\partial}{\partial \boldsymbol{t}}[\mu_1^{(III)}D_z^2(x, z|\omega) + \mu_2^{(III)}\boldsymbol{E}_t^2(x, z|\omega)]_{|z=0}. \tag{6.10}$$

The phenomenological constants $\mu_1^{(III)}$ and $\mu_2^{(III)}$ appearing in the boundary condition, equation (6.10), are given by

$$\mu_1^{(III)} = \lim_{\eta\to 0}\int_{-\eta}^{\eta} \mathrm{d}z\, \frac{1}{\epsilon(2\omega, z)}\left[\frac{1}{2}(\alpha_{III}(z)+\beta_{III}(z))\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{\epsilon^2(\omega, z)} + \frac{1}{\epsilon^2(\omega, z)}\frac{\mathrm{d}}{\mathrm{d}z}(\rho_{III}(z)+\kappa_{III}(z))\right] \tag{6.11}$$

and

$$\mu_2^{(III)} = \lim_{\eta\to 0}\int_{-\eta}^{\eta} \mathrm{d}z\, \frac{1}{\epsilon(2\omega, z)}\frac{\mathrm{d}\kappa_{III}(z)}{\mathrm{d}z}. \tag{6.12}$$

In the same manner we obtain the second boundary condition for the prism–metal interface $x_3 = D$ in the form

$$\boldsymbol{E}_t^{(I)}(x_1, x_3|2\omega)_{|x_3=D} - \boldsymbol{E}_t^{(II)}(x_1, x_3|2\omega)_{|x_3=D} = -\frac{\partial}{\partial \boldsymbol{t}}\left[\mu_1^{(I)} D_z^2(x_1, x_3|\omega) + \mu_2^{(I)}\boldsymbol{E}_t^2(x_1, x_3|\omega)\right]_{|x_3=D}. \tag{6.13}$$

The phenomenological constants $\mu_1^{(I)}$ and $\mu_2^{(I)}$ appearing in the boundary condition (6.13) are given by

$$\mu_1^{(I)} = \lim_{\eta\to 0}\int_{D-\eta}^{D+\eta} \mathrm{d}x_3 \frac{1}{\epsilon(2\omega, x_3)}\left[\frac{1}{2}(\alpha_I(x_3)+\beta_I(x_3))\frac{\mathrm{d}}{\mathrm{d}x_3}\frac{1}{\epsilon^2(\omega, x_3)} + \frac{1}{\epsilon^2(\omega, x_3)}\frac{\mathrm{d}}{\mathrm{d}x_3}(\rho_I(x_3)+\kappa_I(x_3))\right] \tag{6.14}$$

and

$$\mu_2^{(I)} = \lim_{\eta\to 0}\int_{D-\eta}^{D+\eta} \mathrm{d}x_3 \frac{1}{\epsilon(2\omega, x_3)}\frac{\mathrm{d}\kappa_I(x_3)}{\mathrm{d}x_3}. \tag{6.15}$$

From the boundary conditions, equations (6.4), (6.6), (6.10) and (6.13), it follows that when the incident field is p-polarized the nonlinear sources on the right-hand sides of the equations are non-zero only for p-polarized fields of frequency 2ω. In this case it is convenient to work with the single non-zero component of the magnetic field $H_2(x_1, x_3|\Omega)$, and in the laboratory coordinate system (x_1, x_2, x_3) the nonlinear boundary conditions (6.4) take the form

$$H_2^{(II)}(x_1, x_3|2\omega)_{|x_3=\zeta(x_1)} - H^{(III)}(x_1|2\omega) = \frac{2\mathrm{i}c}{\omega}\mu_3^{(III)}\frac{L^{(III)}(x_1|\omega)}{\phi^2(x_1)}\frac{\mathrm{d}}{\mathrm{d}x_1}H^{(III)}(x_1|\omega) \equiv A^{(III)}(x_1|2\omega) \tag{6.16a}$$

and

$$\frac{1}{\epsilon(2\omega)}\left[\frac{\partial}{\partial N}H_2^{(II)}(x_1, x_3|2\omega)\right]_{|x_3=\zeta(x_1)} - L^{(III)}(x_1|2\omega) = \frac{2\mathrm{i}c}{\omega}\frac{\mathrm{d}}{\mathrm{d}x_1}\left\{\frac{1}{\phi^2(x_1)}\left[\mu_1^{(III)}\left(\frac{\mathrm{d}}{\mathrm{d}x_1}H^{(III)}(x_1|\omega)\right)^2 + \mu_2^{(III)}L^{(III)}(x_1|\omega)^2\right]\right\} \equiv B^{(III)}(x_1|2\omega) \tag{6.16b}$$

where $H^{(I,III)}(x_1|\Omega)$ and $L^{(I,III)}(x_1|\Omega)$ are given by equations (4.6), while $\partial/\partial N$ is the unnormalized derivative along the normal to the surface and is given by equation (4.4).

The nonlinear boundary conditions (6.2) and (6.8) in terms of the source functions take the form

$$H^{(I)}(x_1|2\omega) - H_2^{(II)}(x_1, x_3|2\omega)_{|x_3=D} = \frac{2ic}{\omega}\mu_3^{(I)}\frac{1}{\epsilon_0}L^{(I)}(x_1|\omega)\frac{d}{dx_1}H^{(I)}(x_1|\omega)$$
$$\equiv A^{(I)}(x_1|2\omega) \tag{6.17a}$$

and

$$\frac{1}{\epsilon_0}L^{(I)}(x_1|2\omega) - \frac{1}{\epsilon(2\omega)}\frac{\partial}{\partial x_3}H_2^{(II)}(x_1, x_3|2\omega)_{|x_3=D}$$
$$= \frac{2ic}{\omega}\frac{d}{dx_1}\left[\mu_1^{(I)}\left(\frac{d}{dx_1}H^{(I)}(x_1|\omega)\right)^2 + \mu_2^{(I)}\frac{1}{\epsilon_0^2}L^{(I)}(x_1|\omega)^2\right]$$
$$\equiv B^{(I)}(x_1|2\omega). \tag{6.17b}$$

The nonlinear boundary conditions (6.16) and (6.17) were derived for the general form of the surface nonlinear polarization given by equation (6.1). The requirement of energy conservation leads to the following relations among the coefficients in equation (6.1): $\alpha_{I,III} = -\beta_{I,III} = -\gamma_{I,III}$, and $\nabla\rho_{I,III} = \nabla(\alpha_{I,III}+2\kappa_{I,II})$. In this case the nonlinear coefficients $\mu_j^{(I,III)}$ takes the form

$$\mu_3^{(III)} = \lim_{\eta\to 0}\int_{-\eta}^{\eta} dz\left[\frac{d}{dz}\frac{\alpha_{III}(z)}{\epsilon(\omega, z)} + \frac{2}{\epsilon(\omega, z)}\frac{d\kappa_{III}(z)}{dz}\right] \tag{6.18a}$$

$$\mu_1^{(III)} = \lim_{\eta\to 0}\int_{-\eta}^{\eta} dz\,\frac{1}{\epsilon^2(\omega, z)\epsilon(2\omega, z)}\frac{d}{dz}(\alpha_{III}(z) + 3\kappa_{III}(z)) \tag{6.18b}$$

$$\mu_2^{(III)} = \lim_{\eta\to 0}\int_{-\eta}^{\eta} dz\,\frac{1}{\epsilon(2\omega, z)}\frac{d\kappa_{III}(z)}{dz} \tag{6.18c}$$

and

$$\mu_3^{(I)} = \lim_{\eta\to 0}\int_{D-\eta}^{D+\eta} dz\left[\frac{d}{dz}\frac{\alpha_I(z)}{\epsilon(\omega, z)} + \frac{2}{\epsilon(\omega, z)}\frac{d\kappa_I(z)}{dz}\right] \tag{6.19a}$$

$$\mu_1^{(I)} = \lim_{\eta\to 0}\int_{D-\eta}^{D+\eta} dz\,\frac{1}{\epsilon^2(\omega, z)}\frac{d}{dz}(\alpha_I(z) + 3\kappa_I(z)) \tag{6.19b}$$

$$\mu_2^{(I)} = \lim_{\eta\to 0}\int_{D-\eta}^{D+\eta} dz\,\frac{1}{\epsilon(2\omega, z)}\frac{d\kappa_I(z)}{dz}. \tag{6.19c}$$

If we assume that the coefficients $\rho_{I,III}$ and $\kappa_{I,III}$ vanish, while $\alpha_{I,III}$, $\beta_{I,III}$ and $\gamma_{I,III}$ are independent of the distance from the interface, the nonlinear polarization reduces to the form introduced by Bloembergen *et al* [30]. In this case $\mu_2^{(I,III)}$ vanish, while the nonlinear coefficients $\mu_1^{(I,III)}$ and $\mu_3^{(I,III)}$ take the forms

$$\mu_3^{(III)} = \lim_{\eta\to 0}\alpha_{III}\int_{-\eta}^{\eta} dz\,\frac{d}{dz}\frac{1}{\epsilon(\omega, z)} \tag{6.20a}$$

$$\mu_1^{(III)} = \lim_{\eta\to 0}\frac{1}{2}(\alpha_{III} + \beta_{III})\int_{-\eta}^{\eta} dz\,\frac{1}{\epsilon(2\omega, z)}\frac{d}{dz}\frac{1}{\epsilon^2(\omega, z)} \tag{6.20b}$$

and

$$\mu_3^{(I)} = \lim_{\eta\to 0} \alpha_I \int_{D-\eta}^{D+\eta} \mathrm{d}z \, \frac{\mathrm{d}}{\mathrm{d}z} \frac{1}{\epsilon(\omega, z)} \tag{6.21a}$$

$$\mu_1^{(I)} = \lim_{\eta\to 0} \frac{1}{2}(\alpha_I + \beta_I) \int_{D-\eta}^{D+\eta} \mathrm{d}z \, \frac{1}{\epsilon(2\omega, z)} \frac{\mathrm{d}}{\mathrm{d}z} \frac{1}{\epsilon^2(\omega, z)}. \tag{6.21b}$$

Finally, if we assume that $\rho_{I,III} = \kappa_{I,III} = 0$, while $\alpha_{I,III} = e/(8\pi m\omega^2) \equiv \beta$, $\gamma_{I,III} = \beta_{I,III} = e^3 n_0(z)/(4m^2\omega^4) \equiv 2\gamma(z)$, where e is the magnitude of the electron charge, m is the electric mass and $n_0(z)$ is the electron number density, $n_0(z) = n_0\theta(z)\theta(D-z)$, where $\theta(z)$ is the Heaviside unit step function, the surface nonlinear function given by equation (6.1) reduces to that of the free-electron model [32, 36–38]. In our numerical calculations we will use this model of the surface nonlinear polarization because its great advantage is that all the nonlinear coefficients can be expressed in terms of the dielectric functions of the adjacent media. The resulting nonlinear coefficients are known to yield a good fit to the experimental data on second harmonic generation in reflection from a planar silver surface [38].

Since in the free-electron model the nonlinear coefficients β and γ are known and we can use the Drude model to relate the dielectric function to the electron number density $n_0(z)$, the nonlinear coefficients μ_1, μ_2 and μ_3 can be expressed in terms of $\epsilon(\omega, z)$, whose behaviour at the surface is known.

In the Drude model the dielectric function $\epsilon(\omega)$ is $\epsilon(\omega) = 1 - (4\pi e^2 n_0)/(m\omega^2)$, so we have $\epsilon(2\omega) = (3 + \epsilon(\omega))/4$ and $\gamma = [\beta(1 - \epsilon(\omega))/4]\theta(z)\theta(D - z)$. To calculate the phenomenological nonlinear coefficients μ_1, μ_2 and μ_3 we can represent $\epsilon(2\omega, z)$ and $\gamma(z)$ in the forms

$$\epsilon(2\omega, z) = \tilde{\epsilon} - (\tilde{\epsilon} - \epsilon(\omega, z))\left(1 - \tfrac{3}{4}\alpha\right) \tag{6.22}$$

and

$$\gamma(z) = \tfrac{1}{4}\beta(\tilde{\epsilon} - \epsilon(\omega, z))\alpha \tag{6.23}$$

where

$$\alpha = \frac{1 - \epsilon(\omega)}{\tilde{\epsilon} - \epsilon(\omega)} \tag{6.24}$$

with $\tilde{\epsilon} = 1$ for the metal film–vacuum interface and $\tilde{\epsilon} = \epsilon_0$ for the prism–metal film interface. Therefore, the phenomenological nonlinearity constants $\mu_3^{(III)}$ and $\mu_3^{(I)}$ entering the boundary conditions equations (6.4) and (6.6) are now given by

$$\begin{aligned} \mu_3^{(III)} &= \beta \lim_{\eta\to 0} \int_{-\eta}^{\eta} \mathrm{d}z \, \frac{\mathrm{d}}{\mathrm{d}z} \frac{1}{\epsilon(\omega, z)} \\ &= \beta \frac{1 - \epsilon(\omega)}{\epsilon(\omega)} \end{aligned} \tag{6.25}$$

and

$$\begin{aligned} \mu_3^{(I)} &= \beta \lim_{\eta\to 0} \int_{D-\eta}^{D+\eta} \mathrm{d}z \, \frac{\mathrm{d}}{\mathrm{d}z} \frac{1}{\epsilon(\omega, z)} \\ &= -\beta \left(\frac{1}{\epsilon(\omega)} - \frac{1}{\epsilon_0} \right). \end{aligned} \tag{6.26}$$

Since $\kappa_{I,III}$ vanish in the free-electron model, the phenomenological nonlinear constant $\mu_2^{(I,III)}$ also vanishes, while for $\mu_1^{(III)}$ we obtain the expression

$$\mu_1^{(III)} = \lim_{\eta\to 0}\int_{-\eta}^{\eta} \mathrm{d}z\, \frac{1}{\epsilon(2\omega,z)}\left(\gamma(z)+\frac{\beta}{2}\right)\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{\epsilon^2(\omega,z)}$$
$$= -2\beta\int_1^{\epsilon(\omega)} \mathrm{d}\epsilon(\omega,z)\,\frac{3-\epsilon(\omega,z)}{\epsilon(\omega,z)+3}\frac{1}{\epsilon^3(\omega,z)}. \tag{6.27}$$

As a result of the integration we obtain that

$$\mu_1^{(III)} = \frac{1}{3}\beta\left[\frac{(\epsilon(\omega)-1)(\epsilon(\omega)-3)}{\epsilon^2(\omega)} - \frac{4}{3}\ln\left(\frac{\epsilon(\omega)}{\epsilon(2\omega)}\right)\right]. \tag{6.28}$$

For the prism–metal film interface we have

$$\mu_1^{(I)} = \lim_{\eta\to 0}\int_{D-\eta}^{D+\eta} \mathrm{d}z\, \frac{1}{\epsilon(2\omega,z)}\left(\gamma(z)+\frac{\beta}{2}\right)\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{\epsilon^2(\omega,z)}$$
$$= -2\beta\int_{\epsilon(\omega)}^{\epsilon_0} \frac{2+\epsilon_0\alpha-\alpha\epsilon(\omega,z)}{\epsilon(\omega,z)(4-3\alpha)+3\epsilon_0\alpha}\frac{1}{\epsilon^3(\omega,z)}\frac{\mathrm{d}\epsilon(\omega,z)}{\mathrm{d}z}. \tag{6.29}$$

On evaluating the integral in the second line of equation (6.29) we obtain

$$\mu_1^{(I)} = -2\beta\left\{A\,\ln\left[\frac{\epsilon(\omega)}{\epsilon(2\omega)}\right] - B\left[\frac{1}{\epsilon_0}-\frac{1}{\epsilon(\omega)}\right] - \frac{C}{2}\left[\frac{1}{\epsilon_0^2}-\frac{1}{\epsilon^2(\omega)}\right]\right\} \tag{6.30}$$

where

$$C = \frac{2}{3}\frac{2+\epsilon_0\alpha}{\epsilon_0\alpha} \tag{6.31a}$$

$$B = -\frac{2}{3\epsilon_0} + \frac{4-3\alpha}{3\epsilon_0\alpha}C \tag{6.31b}$$

and

$$A = -\frac{4-3\alpha}{3\epsilon_0\alpha}B. \tag{6.31c}$$

The values of the effective nonlinear susceptibilities of the metal–vacuum interface $\mu_1^{(III)}$ and $\mu_3^{(III)}$, given by equations (6.25) and (6.28) coincide with those proposed by Mendoza and Mochan [35] for semiconductors possessing a centre of inversion.

7. The scattering equations at 2ω

With the boundary conditions (6.16) and (6.17) in hand, we can now return to equations (4.16)–(4.20) to obtain the integral equations satisfied by the source functions $H^{(I,III)}(x_1|2\omega)$ and $L^{(I,III)}(x_1|2\omega)$. These can be written as

$$H^{(I)}(x_1|2\omega) = \int_{-\infty}^{\infty}\mathrm{d}x_1'[\tilde{H}_p^{DD}(x_1|x_1')H^{(I)}(x_1'|2\omega) - \tilde{L}_p^{DD}(x_1|x_1')L^{(I)}(x_1'|2\omega)] \tag{7.1}$$

$$Q^{(I)}(x_1|2\omega) = \int_{-\infty}^{\infty} \mathrm{d}x_1' \Big[-\tilde{H}_\epsilon^{DD}(x_1|x_1') H^{(I)}(x_1'|2\omega) + \frac{\epsilon(2\omega)}{\epsilon_0} \tilde{L}_\epsilon^{DD}(x_1|x_1') L^{(I)}(x_1'|2\omega) + \tilde{H}^{DO} \epsilon(x_1|x_1') H^{(III)}(x_1'|2\omega) - \epsilon(2\omega) \tilde{L}_\epsilon^{DO}(x_1|x_1') L^{(III)}(x_1'|2\omega) \Big] \tag{7.2}$$

$$H^{(III)}(x_1|2\omega) = Q^{(III)}(x_1|2\omega) + \int_{-\infty}^{\infty} \mathrm{d}x_1' \Big[-\tilde{H}_\epsilon^{OD}(x_1|x_1') H^{(I)}(x_1'|2\omega) + \frac{\epsilon(2\omega)}{\epsilon_0} \tilde{L}_\epsilon^{OD}(x_1|x_1') L^{(I)}(x_1'|2\omega) + \tilde{H}_\epsilon^{OO}(x_1|x_1') H^{(III)}(x_1'|2\omega) - \epsilon(2\omega) \tilde{L}_\epsilon^{OO}(x_1|x_1') L^{(III)}(x_1'|2\omega) \Big] \tag{7.3}$$

$$0 = \int_{-\infty}^{\infty} \mathrm{d}x_1' \Big[-\tilde{H}_0^{OO}(x_1|x_1') H^{(III)}(x_1'|2\omega) + \tilde{L}_0^{OO}(x_1|x_1') L^{(III)}(x_1'|2\omega) \Big] \tag{7.4}$$

where

$$Q^{(I)}(x_1|2\omega) = \int_{-\infty}^{\infty} \mathrm{d}x_1' \Big[-\tilde{H}_\epsilon^{DD}(x_1|x_1') A^{(I)}(x_1'|2\omega) + \epsilon(2\omega) \tilde{L}_\epsilon^{DD}(x_1|x_1') B^{(I)}(x_1'|2\omega) - \tilde{H}_\epsilon^{DO}(x_1|x_1') A^{(III)}(x_1'|2\omega) + \epsilon(2\omega) \tilde{L}_\epsilon^{DO}(x_1|x_1') B^{(III)}(x_1'|2\omega) \Big] \tag{7.5}$$

$$Q^{(III)}(x_1|2\omega) = -A^{(III)}(x_1|2\omega) + \int_{-\infty}^{\infty} \mathrm{d}x_1' \Big[\tilde{H}_\epsilon^{OD}(x_1|x_1') A^{(I)}(x_1'|2\omega) - \epsilon(2\omega) \tilde{L}_\epsilon^{OD}(x_1|x_1') B^{(I)}(x_1'|2\omega) + \tilde{H}_\epsilon^{OO}(x_1|x_1') A^{(III)}(x_1'|2\omega) - \epsilon(2\omega) \tilde{L}_\epsilon^{OO}(x_1|x_1') B^{(III)}(x_1'|2\omega) \Big]. \tag{7.6}$$

The matrix elements entering these equations are obtained from those tabulated in appendix A through the replacement of ω everywhere by 2ω.

We note that the functions $Q^{(I)}(x_1|2\omega)$ and $Q^{(III)}(x_1|2\omega)$ play the role of sources in equations (7.2) and (7.3).

The integral equations (7.1)–(7.4) are converted into matrix equations following the procedure outlined in appendix B, which are then solved numerically to yield $H^{(I,III)}(x_1|2\omega)$ and $L^{(I,III)}(x_1|2\omega)$. The results are then used to calculate the reflection and transmission efficiencies of the second harmonic fields as follows.

With the use of the integral representation (4.7*a*) of the Green function $G_{\epsilon_0}^{(2\omega)}(x_1, x_3|x_1', x_3')$ we can write the two-component of the scattered field in the form

$$H_2^{(I)}(x_1, x_3|2\omega)_{sc} = \int_{-2\sqrt{\epsilon_0}\omega/c}^{2\sqrt{\epsilon_0}\omega/c} \frac{\mathrm{d}q}{2\pi} \exp(\mathrm{i}qx_1 + \mathrm{i}\alpha_p(q, 2\omega)x_3) \tag{7.7}$$

where

$$R(q, 2\omega) = \frac{1}{2\mathrm{i}\alpha_p(q, 2\omega)} \int_{-\infty}^{\infty} \mathrm{d}x_1 \exp(-\mathrm{i}qx_1 - \mathrm{i}\alpha_p(q, 2\omega)D) \times [\mathrm{i}\alpha_p(q, 2\omega) H^{(I)}(x_1|2\omega) + L^{(I)}(x_1|2\omega)]. \tag{7.8}$$

The total, time-averaged flux crossing the place $x_3 = \text{constant} > D$ is

$$P_{2\omega}^{sc} = \frac{c^2}{128\pi^2\omega\epsilon_0}\int_{-\pi/2}^{\pi/2} \mathrm{d}\theta_s \, |r(\theta_s|2\omega)|^2 \tag{7.9}$$

where

$$r(\theta_s|2\omega) = 4\mathrm{i}\sqrt{\epsilon_0}\,\frac{\omega}{c}\cos\theta_s R\left(2\sqrt{\epsilon_0}\,\frac{\omega}{c}\sin\theta_s, 2\omega\right). \tag{7.10}$$

We rewrite $P_{2\omega}^{sc}$ in the form

$$P_{2\omega}^{sc} = \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta_s \; P_s(\theta_s|2\omega) \tag{7.11}$$

where $P_s(\theta_s|2\omega)$ is given by

$$P_s(\theta_s|2\omega) = \frac{c^2}{128\pi^2\omega\epsilon_0}|r(\theta_s|2\omega)|^2. \tag{7.12}$$

Similarly, for the two-component of the transmitted field we have

$$H_2^{(III)}(x_1, x_3|2\omega)_{tr} = \int_{-2\omega/c}^{2\omega/c} \frac{\mathrm{d}q}{2\pi}\, T(q, 2\omega)\exp(\mathrm{i}qx_1 - \mathrm{i}\alpha_0(q, 2\omega)x_3) \tag{7.13}$$

where

$$\begin{aligned} T(q, 2\omega) = {} & \frac{1}{2\mathrm{i}\alpha_0(q, 2\omega)}\int_{-\infty}^{\infty} \mathrm{d}x_1 \exp(-\mathrm{i}qx_1 + \mathrm{i}\alpha_0(q, 2\omega)\zeta(x_1)) \\ & \times[\mathrm{i}(q\zeta'(x_1) + \alpha_0(q, 2\omega))H^{(III)}(x_1|2\omega) - L^{(III)}(x_1|2\omega)]. \end{aligned} \tag{7.14}$$

The total, time-averaged transmitted flux crossing the plane $x_3 = \text{constant} < \zeta(x_1)_{min}$ is

$$P_{2\omega}^{tr} = \frac{c^2}{128\pi^2\omega}\int_{-\pi/2}^{\pi/2} \mathrm{d}\theta_t \, |t(\theta_t|2\omega)|^2 \tag{7.15}$$

where

$$t(\theta_t|\omega) = 4\mathrm{i}\frac{\omega}{c}\cos\theta_t \, T\left(2\frac{\omega}{c}\sin\theta_t, 2\omega\right). \tag{7.16}$$

In a similar fashion we rewrite $P_{2\omega}^{tr}$ in the form

$$P_{2\omega}^{tr} = \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta_t \; P_t(\theta_t|2\omega) \tag{7.17}$$

where $P_t(\theta_t|2\omega)$ is given by

$$P_t(\theta_t|2\omega) = \frac{c^2}{128\pi^2\omega}|t(\theta_t|2\omega)|^2. \tag{7.18}$$

We define the efficiency of second harmonic generation in reflection as the total power of the scattered harmonic light, normalized by the square of the power of the incident field and multiplied by the illuminated area S:

$$I_{2\omega}^{sc} = \frac{P_{2\omega}^{sc}}{P_{inc}^2}S. \tag{7.19}$$

We take the value of S to be equal to $2gL_2$, where g is the half-width of the intercept of the incident beam with the upper plane ($x_3 = D$) of the film, L_2 is a large arbitrary length along the x_2-axis, and P_{inc} is given by equation (5.8).

Similarly, the efficiency of second harmonic generation in transmission is given by

$$I_{2\omega}^{tr} = \frac{P_{2\omega}^{tr}}{P_{inc}^2} S. \tag{7.20}$$

The efficiencies defined in this way do not depend on the amplitude of the incident field, and are therefore convenient to use in experimental measurements of the second harmonic generation.

We then define the differential reflection and transmission efficiencies of second harmonic light by

$$\frac{\partial}{\partial\theta_s} I_{2w}^{sc}(\theta_s) = \frac{P_s(\theta_s|2\omega)}{P_{inc}^2} S = \frac{c^2 S}{128\pi^2\omega\epsilon_0 P_{inc}^2} |r(\theta_s|2\omega)|^2 \tag{7.21a}$$

$$\frac{\partial}{\partial\theta_t} I_{2w}^{tr}(\theta_t) = \frac{P_t(\theta_s|2\omega)}{P_{inc}^2} S = \frac{c^2 S}{128\pi^2\omega P_{inc}^2} |t(\theta_t|2\omega)|^2. \tag{7.21b}$$

To obtain the mean differential reflection and transmission efficiencies in the Kretschmann geometry with a random surface we must average the expressions given by equation (7.21) over the ensemble of realizations of the surface profile function,

$$\left\langle \frac{\partial}{\partial\theta_s} I_{2w}^{sc}(\theta_s) \right\rangle = \frac{c^2 S}{128\pi^2\omega\epsilon_0 P_{inc}^2} \langle |r(\theta_s|2\omega)|^2 \rangle \tag{7.22a}$$

$$\left\langle \frac{\partial}{\partial\theta_t} I_{2w}^{tr}(\theta_t) \right\rangle = \frac{c^2 S}{128\pi^2\omega P_{inc}^2} \langle |t(\theta_s|2\omega)|^2 \rangle. \tag{7.22b}$$

The contributions to the mean differential reflection and transmission efficiencies from the incoherent components of the scattered and transmitted light, respectively, are then given by

$$\left\langle \frac{\partial}{\partial\theta_s} I_{2\omega}^{sc}(\theta_s) \right\rangle_{incoh} = \frac{c^2 S}{128\pi^2\omega\epsilon_0 P_{inc}^2} [\langle |r(\theta_s|2\omega)|^2 \rangle - |\langle r(\theta_s|2\omega) \rangle|^2] \tag{7.23a}$$

$$\left\langle \frac{\partial}{\partial\theta_t} I_{2\omega}^{tr}(\theta_t) \right\rangle_{incoh} = \frac{c^2 S}{128\pi^2\omega P_{inc}^2} [\langle |t(\theta_t|2\omega)|^2 \rangle - |\langle t(\theta_t|2\omega) \rangle|^2]. \tag{7.23b}$$

8. Results and discussion

In this section we present results for the scattering of p-polarized light from, and its transmission through, the structure depicted in figure 1, at both the fundamental and harmonic frequencies. The numerical results illustrated in this section represent the averages of the corresponding results obtained from $N_p \geqslant 1000$ realizations of the random surface. In each case the number of realizations N_p is given in the caption to the corresponding figure.

The results obtained at the fundamental frequency, in particular the source functions $H^{(I)}(x_1|\omega)$, $L^{(I)}(x_1|\omega)$, $H^{(III)}(x_1|\omega)$ and $L^{(III)}(x_1|\omega)$, are needed for solving the problems of scattering and transmission at the harmonic frequency. However, the angular distributions of the intensity of the light scattered and transmitted incoherently at this frequency are also of interest in themselves, because they can display features associated both with the fact that the scattering structure supports two surface plasmon polaritons at each frequency ω below $\omega = \omega_p/\sqrt{1+\epsilon_0}$, where ω_p is the plasma frequency of the conduction electrons in the metal film, and with the presence of the prism on whose base the metal film is deposited. To the best

of our knowledge the scattering and transmission of p-polarized light from a rough metal film in the Kretschmann attenuated total reflection geometry has not been studied theoretically up to the present time.

8.1. Results at the fundamental frequency

To illustrate the variety of single-scattering and multiple-scattering phenomena to which the system depicted in figure 1 can give rise, we consider a silver film of mean thickness 48 nm with a randomly rough surface characterized by a Gaussian or a West–O'Donnell power spectrum. We have chosen two wavelengths of the incident beam, namely 394.7 and 612.7 nm. At these frequencies the complex dielectric function of silver has the values $\epsilon(\omega) = -4.28 + \mathrm{i}0.21$ and $\epsilon(\omega) = -17.2 + \mathrm{i}0.498$, respectively. The medium above the film is assumed to be either a dielectric or vacuum. In the latter case (a free-standing film) there are no leaky wave peaks and, as explained earlier, we therefore expect that multiple scattering effects will manifest themselves more clearly in the scattering spectrum in this case than in the case of the Kretschmann geometry.

In figure 3 we plot the contributions to the mean differential reflection (DRC) and transmission (DTC) coefficients from the incoherent components of the scattered and transmitted fields for p-polarized light of wavelength $\lambda = 394.7$ nm incident on a free-standing film with a random back surface characterized by a Gaussian power spectrum (2.5) and roughness parameters $\delta = 8.8$, $a = 250$ nm. The angle of incidence is $\theta_0 = 0^\circ$. The mean thickness of the film is $D = 48$ nm. The numerical solution of the dispersion equation (3.1) for the film with planar surfaces in the case where $\epsilon_0 = 1$, yields two surface plasmon polaritons, the real parts of whose wavenumbers are $q_1(\omega) = 1.07(\omega/c)$, $q_2(\omega) = 1.26(\omega/c)$. The imaginary parts of these wavenumbers are three orders of magnitude smaller than their real parts, so that we will neglect them here and in what follows. The broken lines in figure 3 indicate the angles at which the satellite peaks occur according to these values of $q_1(\omega)$ and $q_2(\omega)$ and equations (3.2) and (3.3). These peaks are weak but well defined, and do not disappear as the number of sampling points used increases. There are no leaky peaks in scattering from a free-standing film. This result is consistent with the arguments presented above.

The situation is different when the upper medium is not vacuum but a dielectric. In figure 4 we plot the contributions to the mean DRC and DTC from the incoherent components of the scattered and transmitted fields in the scattering of p-polarized light in the Kretschmann geometry. In obtaining these results it was assumed that $\theta_0 = 0^\circ$, $\lambda = 394.7$ nm, $\epsilon_0 = 2.25$, $D = 48$ nm, $\epsilon(\omega) = -4.28 + \mathrm{i}0.21$, $\delta = 8.8$ nm, $a = 101.7$ nm. The vertical broken lines mark the positions of the leaky wave peaks as predicted by equation (3.4) with the use of $q_1(\omega) = 1.13(\omega/c)$ $(q_2(\omega)) = 2.23(\omega/c))$. Satellite peaks in reflection are not observed; in transmission they are clearly visible, but are not well pronounced.

In figure 5 we plot the results of calculations of $\langle \partial R/\partial\theta_s \rangle_{incoh}$ and $\langle \partial T/\partial\theta_t \rangle_{incoh}$ for the scattering of a p-polarized beam in the Kretschmann geometry with a random surface characterized by the West–O'Donnell power spectrum (2.6) with $k_{min}^{(1)} = 0.86(\omega/c)$, $k_{max}^{(1)} = 1.80(\omega/c)$, $h_1 = 1$, $h_2 = 0$ and $\delta = 8.8$ nm. In obtaining these results it was assumed that $\theta_0 = 0^\circ$, $\lambda = 394.7$ nm, $\epsilon_0 = 2.25$, $D = 48$ nm, $\epsilon(\omega) = -4.28 + \mathrm{i}0.21$. The numerical solution of the dispersion equation (3.1) for the film with planar surfaces yields two surface plasmon polaritons with wavenumbers $q_1(\omega) = 1.13\omega/c$, $q_2(\omega) = 2.23\omega/c$. Thus, with our choices of $k_{min}^{(1)}$ and $k_{max}^{(1)}$ the incident field can excite only the surface plasmon polariton with wavenumber $q_1(\omega)$. Therefore, satellite peaks will not be observed. Furthermore, it follows from equation (3.3) that the satellite peaks in transmission should appear in the non-

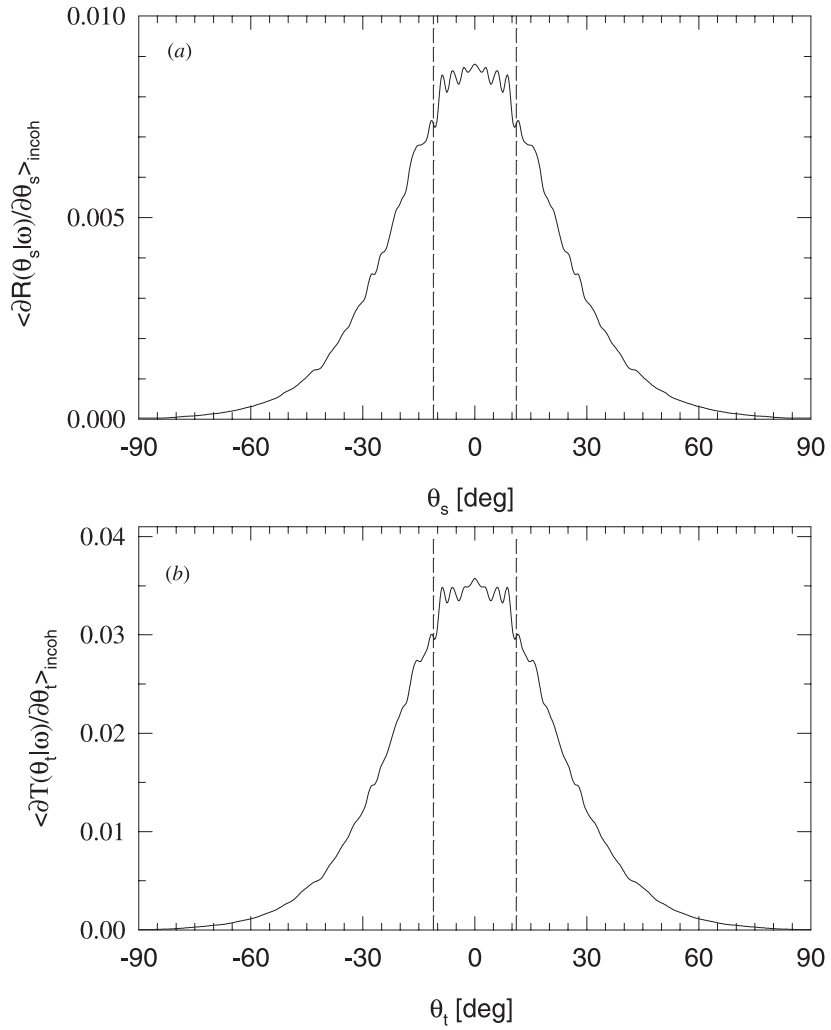

Figure 3. The differential reflection and transmission coefficients for the scattering of a p-polarized beam of light from a one-dimensional free-standing silver film with a random back surface characterized by a Gaussian power spectrum. $L = 2.52\ \mu$m, $\lambda = 394.7$ nm, $\theta_0 = 0^\circ$, $g = L/5$, $\delta = 8.8$ nm, $a = 250$, $D = 0.048$ nm, $\epsilon(\omega) = -4.28 + \mathrm{i}0.28$, $N_p = 1100$, $q_1(\omega) = 1.14\omega/c$, $q_2(\omega) = 1.26\omega/c$. The vertical dotted lines indicate the expected positions of the satellite peaks.

radiative region of the optical spectrum. A prominent peak is observed in transmission. This is the enhanced transmission peak. However, in reflection the amplitude of the enhanced backscattering peak is much smaller than the amplitude of the leaky peaks.

Now we consider a scattering system in which both wavenumbers are inside the window $k_{min}^{(1)} < q_1(\omega), q_2(\omega) < k_{max}^{(1)}$. Thus, the power spectrum is non-zero for a range of wavenumbers that includes $q_1(\omega)$ and $q_2(\omega)$. In figure 6 we plot the DRC for the scattering of a p-polarized beam in the Kretschmann geometry with a random surface characterized by the West–O'Donnell power spectrum with $k_{min}^{(1)} = 0.86\,(\omega/c)$, $k_{max}^{(1)} = 1.80(\omega/c)$, $h_1 = 1$, $h_2 = 0$ and $\delta = 8.8$ nm. It is further assumed that $\theta_0 = 0^\circ$, $\lambda = 612.7$ nm, $\epsilon_0 = 2.25$, $D = 48$ nm, $\epsilon(\omega) = -17.2 + \mathrm{i}0.498$. From the solution of the dispersion equation (3.1) we obtain for these

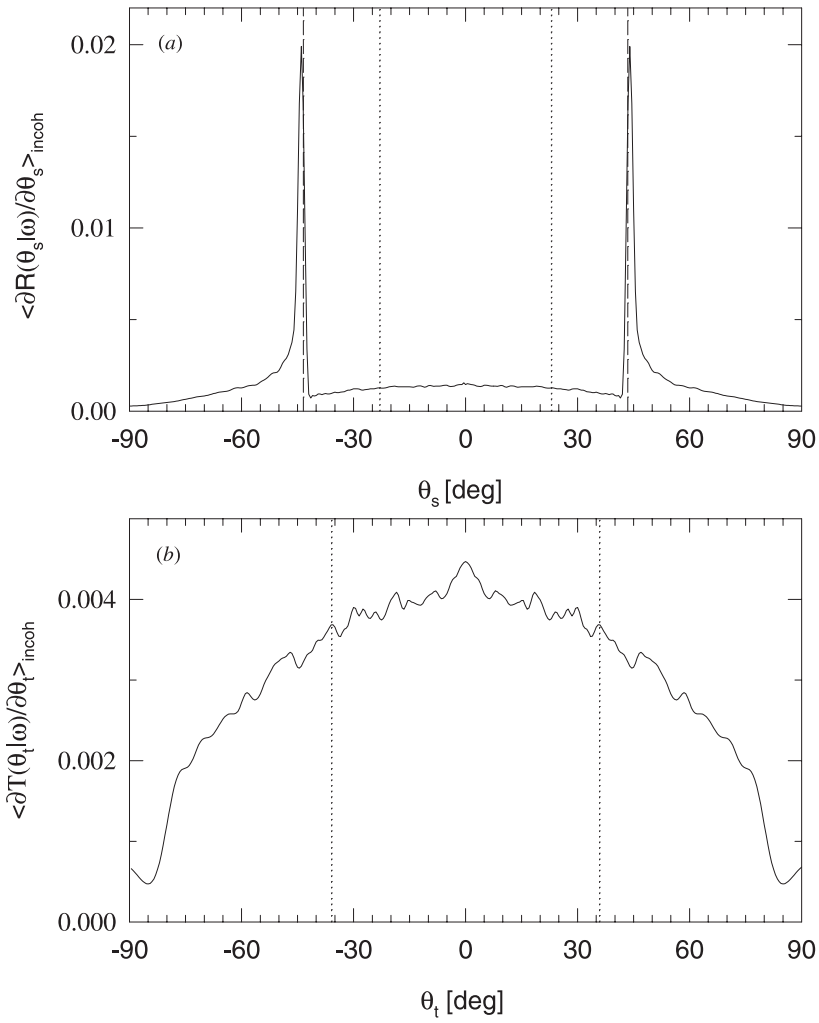

Figure 4. The differential reflection coefficient for the scattering of a p-polarized beam of light in the Kretschmann geometry with a randomly rough silver film the roughness of whose back surface is characterized by a Gaussian power spectrum. $L = 25.2\ \mu$m, $\lambda = 394.7$ nm, $\theta_0 = 0^\circ$, $g = L/5, \delta = 8.8$ nm, $a = 250$ nm, $D = 48$ nm, $\epsilon_0 = 2.25, \epsilon(\omega) = -4.28 + \mathrm{i}0.21$, $N_p = 1100, q_1(\omega) = 1.13\omega/c, q_2(\omega) = 2.23\omega/c$. The vertical dotted and broken lines indicate the expected positions of the satellite and leaky peaks, respectively.

parameters the wavenumbers of the two surface plasmon polaritons given by $q_1(\omega) = 1.03\ (\omega/c), q_2(\omega) = 1.62\,(\omega/c)$. There are two well defined satellite peaks (dotted lines) in transmission and reflection (see the inset) at scattering angles that agree with the expected positions defined by equations (3.2) and (3.3). The angles at which satellite peaks occur are given by $\theta_s^{sat} = \pm 22.99^\circ$ and $\theta_t^{sat} = \pm 35.87^\circ$, in reflection and transmission, respectively. The leaky wave peaks occur at $\theta_{leaky} = \pm 43.43^\circ$ (dotted lines).

We should note that the results plotted in figure 6 show large wings in the DTC. This is a side effect that occurs due to the short segment of integration, i.e. the length of the rough surface in this case is shorter than the propagation length of the surface plasmon polaritons.

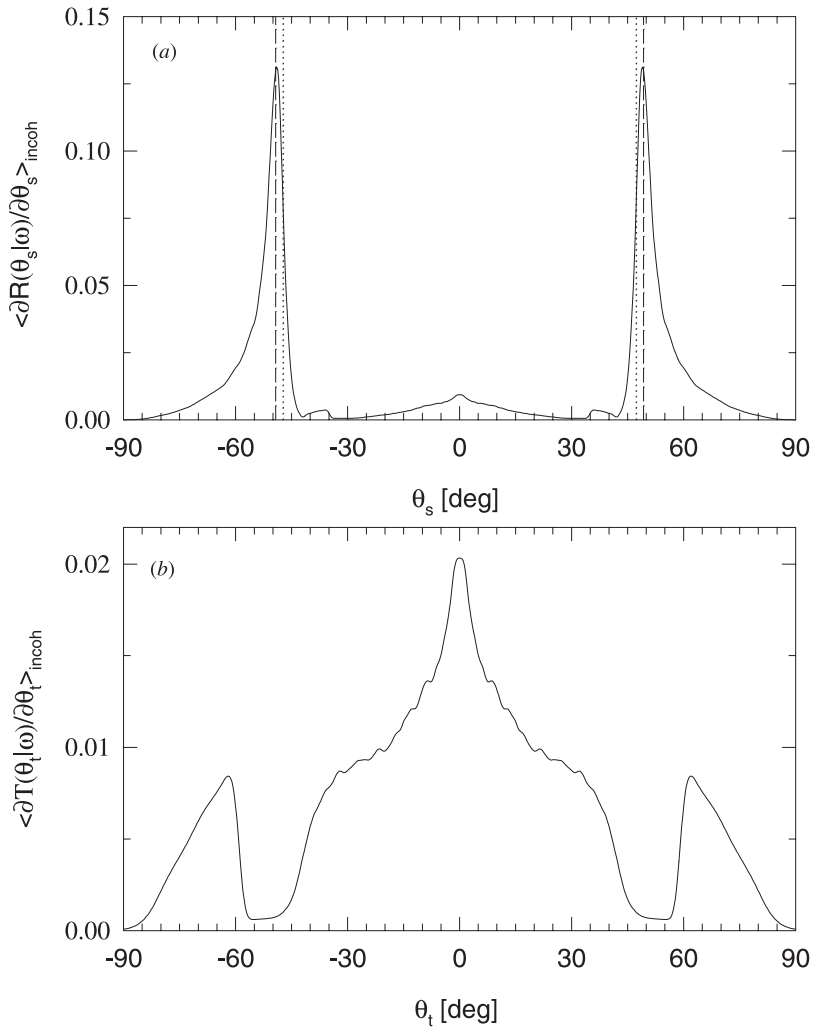

Figure 5. The differential reflection coefficient for the scattering of a p-polarized beam in the Kretschmann geometry with a random back surface characterized by the West–O'Donnell power spectrum. $k_{min}^{(1)} = 0.86\,\omega/c$, $k_{max}^{(1)} = 1.80\,\omega/c$, $h_1 = 1$, $h_2 = 0$, $L = 25.2$ μm, $q_1(\omega) = 1.14\omega/c$, $q_2(\omega) = 1.26\omega/c$, $\lambda = 394.7$ nm, $\theta_0 = 0^\circ$, $g = L/5$, $\delta = 8.8$ nm, $a = 250$ nm, $D = 48$ nm, $\epsilon_0 = 2.25$, $\epsilon(\omega) = -4.28 + \mathrm{i}0.21$, $N = 400$, $N_p = 2000$. The vertical dotted and broken lines indicate the expected positions of the satellite and leaky peaks, respectively.

We intentionally include this plot to compare it with figure 7, which was obtained for the same parameters except that $L = 25.2$ μm (twice as large), $g = L/5$, $N = 500$. A comparison of figures 6 and 7 demonstrates that a short integration segment gives rise to computational errors that can be eliminated by making the integration path sufficiently long. To ensure that the numerical values of the DRC and DTC as well as the differential reflection and transmission efficiencies of second harmonic light are reliable, the length of integration was chosen so that the wings at large angles of scattering or transmission are absent and a further increase of the length does not influence the results. However, in any case, the positions of the satellite and leaky peaks are not greatly affected by the length of the integration interval.

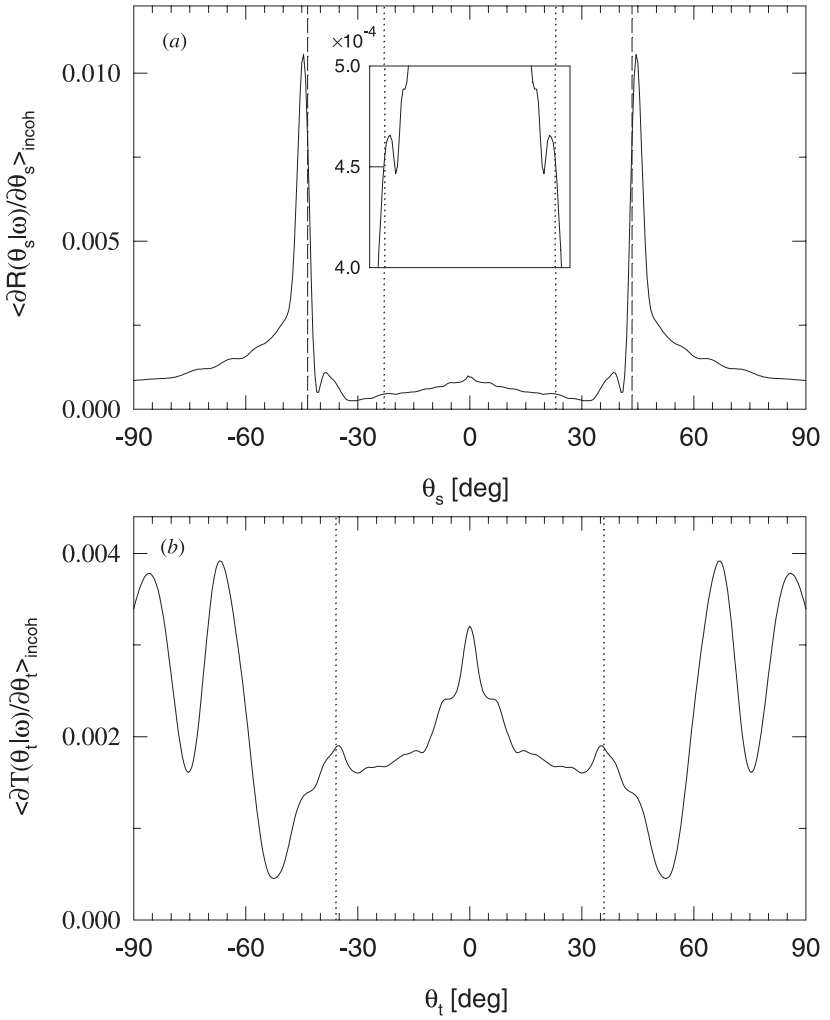

Figure 6. The differential reflection coefficient for the scattering of a p-polarized beam in the Kretschmann geometry with a random back surface characterized by the West–O'Donnell power spectrum. $k_{min}^{(1)} = 0.86\,\omega/c$, $k_{max}^{(1)} = 1.80\,\omega/c$, $h_1 = 1$, $h_2 = 0$, $q_1(\omega) = 1.03\,\omega/c$, $q_2(\omega) = 1.62\,\omega/c$, $L = 12.6\,\mu$m, $\lambda = 612.7$ nm, $\theta_0 = 0^\circ$, $g = L/4$, $\delta = 8.8$ nm, $a = 250$ nm, $D = 48$ nm, $\epsilon_0 = 2.25$, $\epsilon(\omega) = -17.2 + \mathrm{i}0.498$, $N = 400$, $N_p = 2000$. The vertical dotted and broken lines indicate the predicted positions of the satellite and leaky peaks, respectively.

In figure 8 we present the results of calculations for the same parameters as in figure 7, but with the angle of incidence $\theta_0 = 10^\circ$. The positions of the satellite peaks should shift according to equations (3.2) and (3.3) ($\theta_s^{sat} = -34.35^\circ, 12.53^\circ$, $\theta_t^{sat} = -57.82^\circ, 18.99^\circ$), but only the right peak is well manifested, while the left peak disappears. The reverse pattern is observed for the leaky wave peaks: the left peak is much more intense than the right peak. We note that the positions of the leaky wave peaks do not depend on the angle of the incidence, in agreement with equation (3.4).

At first glance, the preceding results seem to be counterintuitive. It might be thought that it is easier to excite the surface polariton in the forward direction, so that the amplitude of

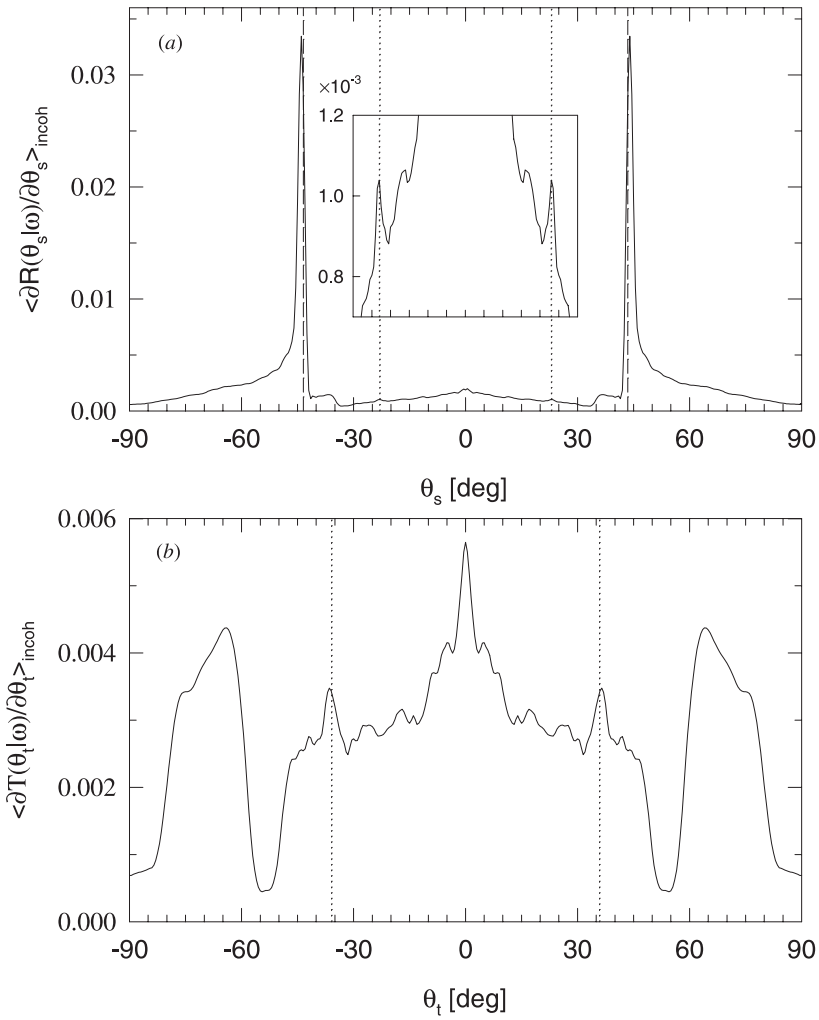

Figure 7. The same as in figure 6 but with $L = 25.2\ \mu$m, $g = L/5$, $N = 500$.

the right leaky peak should be larger than the amplitude of the left leaky peak. In fact, this observation is correct only if the power spectrum of a random surface has a Gaussian form. However, in the case of the West–O'Donnell spectrum the incident light couples strongly to the surface polaritons over a limited range of angles of incidence. Consequently, for some angles of incidence it is more probable to excite a surface polariton in the backward direction.

Let us analyse the results obtained in figure 8. As we mentioned earlier, if the angle of incidence is zero, both wavenumbers satisfy the following set of inequalities:

$$k_{min}^{(1)} < q_1(\omega) < q_2(\omega) < k_{max}^{(1)} \tag{8.1a}$$

$$-k_{max}^{(1)} < -q_2(\omega) < -q_1(\omega) < -k_{min}^{(1)} \tag{8.1b}$$

and both polaritons can propagate in the forward and backward directions with the same

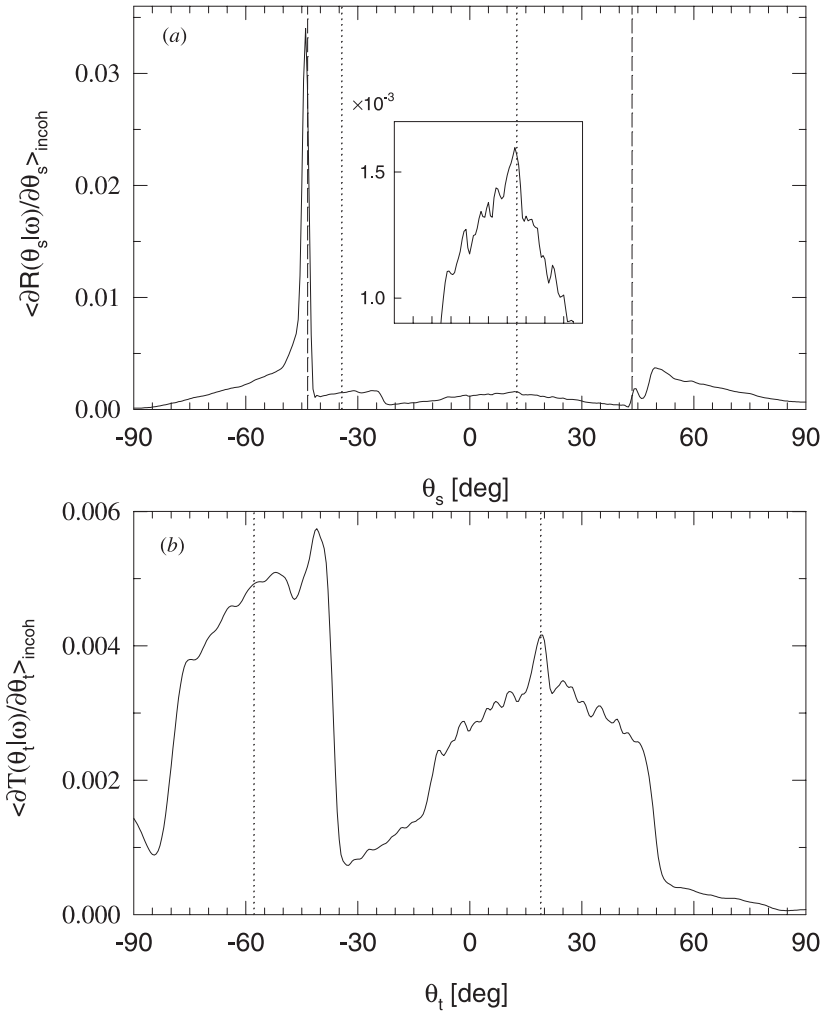

Figure 8. The same as in figure 7, but with $\theta_0 = 10^\circ$.

probability. If the angle of incidence $\theta_0 = 10^\circ$, we have the following set of inequalities:

$$-q_2(\omega) < -k^{(1)}_{max} + \sqrt{\epsilon_0}\,\frac{\omega}{c}\sin\theta_0 < -q_1(\omega) < -k^{(1)}_{min} + \sqrt{\epsilon_0}\,\frac{\omega}{c}\sin\theta_0 \qquad (8.2a)$$

$$q_1(\omega) < k^{(1)}_{min} + \sqrt{\epsilon_0}\,\frac{\omega}{c}\sin\theta_0 < q_2(\omega) < k^{(1)}_{max} + \sqrt{\epsilon_0}\,\frac{\omega}{c}\sin\theta_0. \qquad (8.2b)$$

Thus the wavenumbers $-q_1(\omega)$ and $q_2(\omega)$ are inside the effective coupling window. This means that a surface polariton with wavenumber $q_1(\omega)$ can be excited more effectively in the backward direction, while a surface polariton with wavenumber $q_2(\omega)$ can be excited more effectively in the forward direction. This leads us to the conclusion that the leaky peaks should be more intense in the backward direction than in the forward direction, because the wavenumber $q_1(\omega)$ corresponds to the leaky mode.

A similar argument can be applied to the analysis of the satellite peaks. The more intense satellite peaks in reflection and transmission should occur at the angles given by equations (3.2) and (3.3), where we choose the minus sign before the second term on the

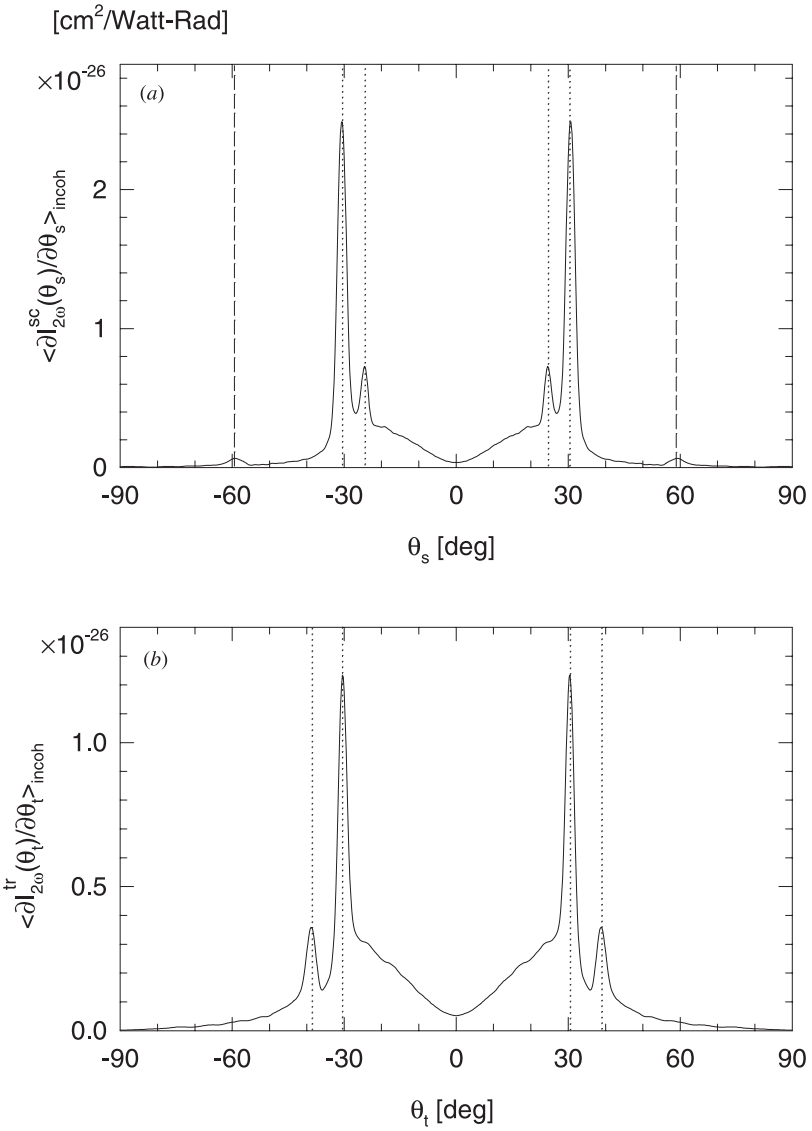

Figure 9. The differential reflection and transmission efficiencies of the second harmonic light in the Kretschmann geometry with a randomly rough silver film characterized by a Gaussian power spectrum. $L = 25.0\ \mu$m, $\lambda = 1.063\ \mu$m, $\theta_0 = 0^\circ$, $g = L/4$, $\delta = 7$ nm, $a = 600$ nm, $D = 50$ nm, $\epsilon_0 = 1.5$, $\epsilon(\omega) = -56.9 + i0.6$, $\epsilon(2\omega) = -11.56 + i0.37$, $N_p = 1000$. The vertical dotted lines indicate the predicted positions of the resonance peaks, the vertical broken lines indicate the positions of the leaky peaks.

right-hand sides of these equations. We believe that this mechanism is also the reason that there are no strong enhanced backscattering and transmission peaks in this case, because for any scattering event the excitation of a surface polariton in the forward direction has a significantly different probability to occur than the probability of the excitation of the surface polariton in the backward direction.

8.2. Results at the harmonic frequency

In figure 9 we present the contributions to the mean differential reflection and transmission efficiencies from the incoherent component of the light scattered from and transmitted through a rough silver film at frequency 2ω, when the lower, random, surface is characterized by the Gaussian power spectrum (2.5) and the roughness parameters $\delta = 7$, $a = 600$ nm. The mean film thickness is $D = 50$ nm, and the angle of incidence $\theta_0 = 0^\circ$. The dielectric constant of the prism is $\epsilon_0 = 1.5$, the wavelength of the incident light is $\lambda = 1.063$ μm, and the dielectric constants of silver at λ and $\lambda/2$ are $\epsilon(\omega) = -56.9 + i0.6$ and $\epsilon(2\omega) = -11.56 + i0.27$, respectively. Four prominent peaks are observed in both reflection and transmission. These peaks occur due to the resonant nonlinear mixing of surface plasmon polaritons of frequency ω with the incident light. A schematic diagram of this nonlinear interaction is shown in figure 10. This scattering geometry supports two surface plasmon polaritons, and they can propagate both in the forward and backward directions. The incident light can excite both surface waves at the fundamental frequency, which then interact nonlinearly with the incident light (this interaction is denoted schematically by the broken circles in figure 10), and are converted into volume electromagnetic waves at the harmonic frequency. This process gives rise to four peaks in the spectrum of the incoherently generated second harmonic light, both in reflection and transmission. The positions of these peaks can be obtained from the equations

$$k \pm q_{1,2}(\omega) = q_{sc}^{2\omega} \tag{8.3a}$$

$$k \pm q_{1,2}(\omega) = q_{tr}^{2w} \tag{8.3b}$$

where $k = \sqrt{\epsilon_0}(\omega/c)\sin\theta_0$ is the x_1-component of the wavevector of the incident light, $q_{1,2}(\omega)$ are the wavenumbers of the surface plasmon polaritons of frequency ω, $q_{sc}^{2\omega} = (2\sqrt{\epsilon_0}\omega/c)\sin\theta_s$ is the x_1-component of the wavevector of the scattered second harmonic light, and $q_{tr}^{2\omega} = (2\omega/c)\sin\theta_t$ is the x_1-component of the wavevector of the transmitted second harmonic light. For the parameters chosen the real parts of the wavenumbers of the surface polaritons are given by

$$q_1(\omega) = 1.009\frac{\omega}{c} \qquad q_2(\omega) = 1.242\frac{\omega}{c} \tag{8.4a}$$

$$q_1(2\omega) = 1.046\frac{2\omega}{c} \qquad q_2(2\omega) = 1.323\frac{2\omega}{c} \tag{8.4b}$$

where $q_1(\omega)$ and $q_1(2\omega)$ are the wavenumbers of leaky modes.

We note that this process is not mediated by surface plasmon polaritons at the second harmonic frequency, since the particular Gaussian power spectrum we used ensures the excitation of the surface polaritons of frequency ω, while the excitation of surface polaritons of frequency 2ω is effectively suppressed. This is because in the latter case the wavenumber of the surface polariton lies in the wings of the power spectrum (especially at normal incidence). However, we can see from figure 9 that even the weak excitation of surface plasmon polaritons of frequency 2ω gives rise to the small leaky peaks in reflection marked by the broken lines. The positions of the leaky peaks are given by

$$q_{leaky}(2\omega) = \pm 2\sqrt{\epsilon_0}\frac{\omega}{c}\sin\theta_{leaky} \tag{8.5}$$

where $q_{leaky}(\omega)$ is the wavenumber of the leaky surface polariton, and θ_{leaky} is the angle that defines the positions of the leaky peaks. The physical mechanism for this process is the same as that discussed in section 3.

In figure 11 we plot the contributions to the mean differential reflection and transmission efficiencies from the incoherent components of the light scattered and transmitted at frequency

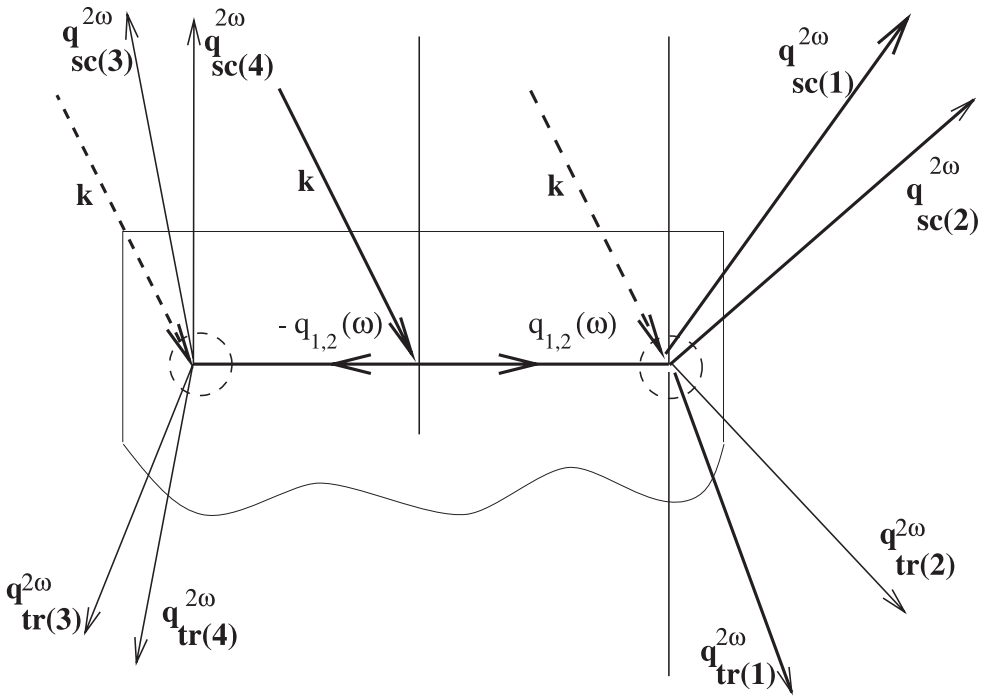

Figure 10. A schematic diagram of the nonlinear interaction of the incident light with a surface plasmon polariton of frequency ω which gives rise to four peaks in the spectrum of the second harmonic light in both reflection and transmission.

2ω for the same parameters as in figure 9, but with $\theta_0 = 5^\circ$. The positions of the resonant peaks are shifted according to equation (8.3), and the right-hand peaks are more intense than the left-hand peaks. The leaky peaks retain the same positions, and the right leaky peak is more intense than the left-hand one. This leads us to the conclusion that the surface polaritons of frequency ω and 2ω are easier to excite in the forward direction than in the backward direction.

In figure 12 we use the same parameters as in figure 9 but with $a = 500$, $\delta = 10$ nm, so that the random surface is rougher. The efficiencies of diffuse scattering and transmission are more than ten times larger than those plotted in figure 9, and the leaky peaks are more pronounced. This is a consequence of the fact that the Gaussian peak in the power spectrum is higher and broader than the one used in obtaining figure 9, which makes it easier to couple the incident light into the surface plasmon polaritons. We should also point out that if $\theta_0 = 0^\circ$ only a dip occurs in the direction normal to the mean surface, both in reflection and transmission. This is due to the symmetry of the nonlinear polarization, namely the second harmonic generation in the direction normal to the mean surface is produced by the nonlinear mixing of contrapropagating surface polaritons, which is forbidden by this symmetry.

In figure 13 we plot the contributions to the mean differential reflection and transmission efficiencies from the incoherent components of the light scattered and transmitted at frequency 2ω when the lower, random, surface of the film is characterized by the West–O'Donnell power spectrum with $k_{min}^{(1)} = 0.86\omega/c$, $k_{max}^{1)} = 1.58\omega/c$, $h_1 = 1$, $h_2 = 0$, $\delta = 7$ nm. The mean film thickness is $D = 50$ nm, and the angle of incidence is $\theta_0 = 0^\circ$. The dielectric constant of the prism is $\epsilon_0 = 1.5$, the wavelength of the incident light is $\lambda = 1.063$ μm, $\epsilon(\omega) = -56.9 + \mathrm{i}0.6$ and $\epsilon(2\omega) = -11.56 + \mathrm{i}0.27$. The wavenumbers of the surface plasmon polaritons for these

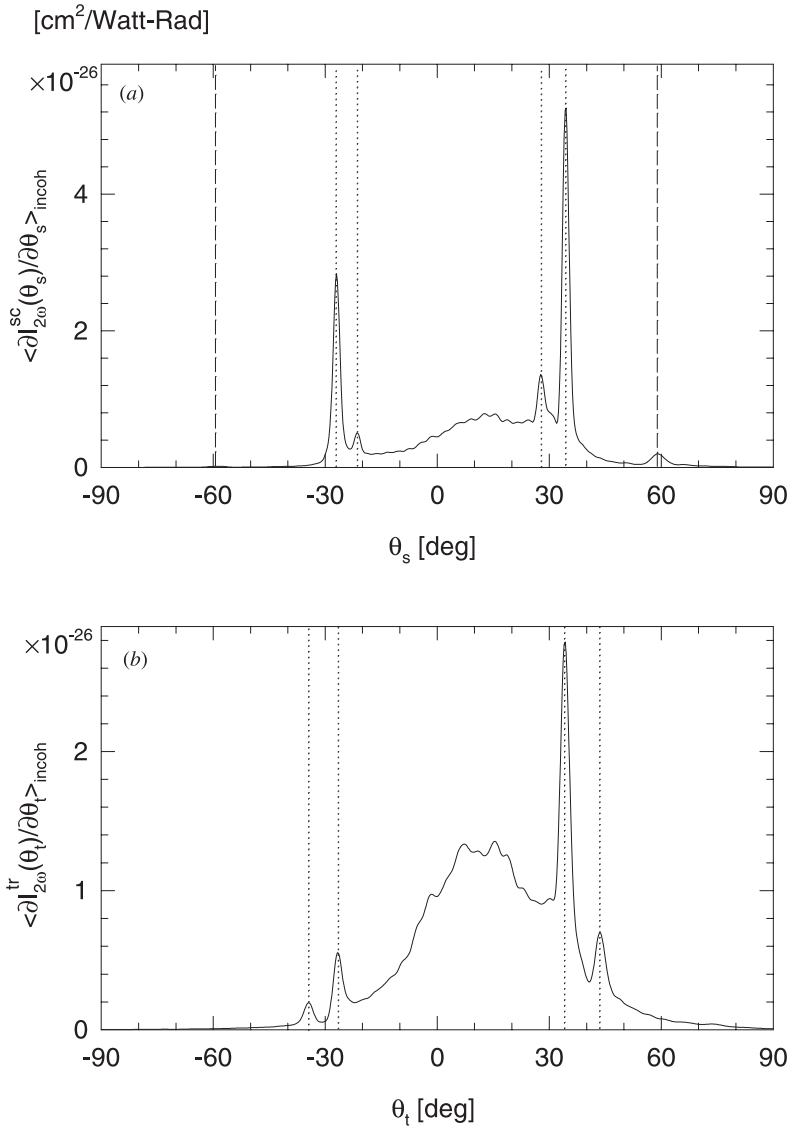

Figure 11. The same as in figure 9, but with $\theta_0 = 5^\circ$.

parameters are those given by equations (8.4). The inset in figure 13 magnifies the region in the vicinity of $\theta_s = 0^\circ$ to show a dip.

Up to now we have considered only small angles of incidence. However, it is worthwhile to study the case in which the angle of incidence is optimal for exciting a surface plasmon polariton. If the angle of incidence is optimal the incident light couples directly into a surface plasmon polariton even if the film has planar interfaces. We can find the optimal angle by matching the x_1-component of the wavevector of the incident light with the wavenumber of the surface plasmon polariton,

$$\sqrt{\epsilon_0}\,\frac{\omega}{c}\sin\theta_{opt} = q_{1,2}(\omega). \tag{8.6}$$

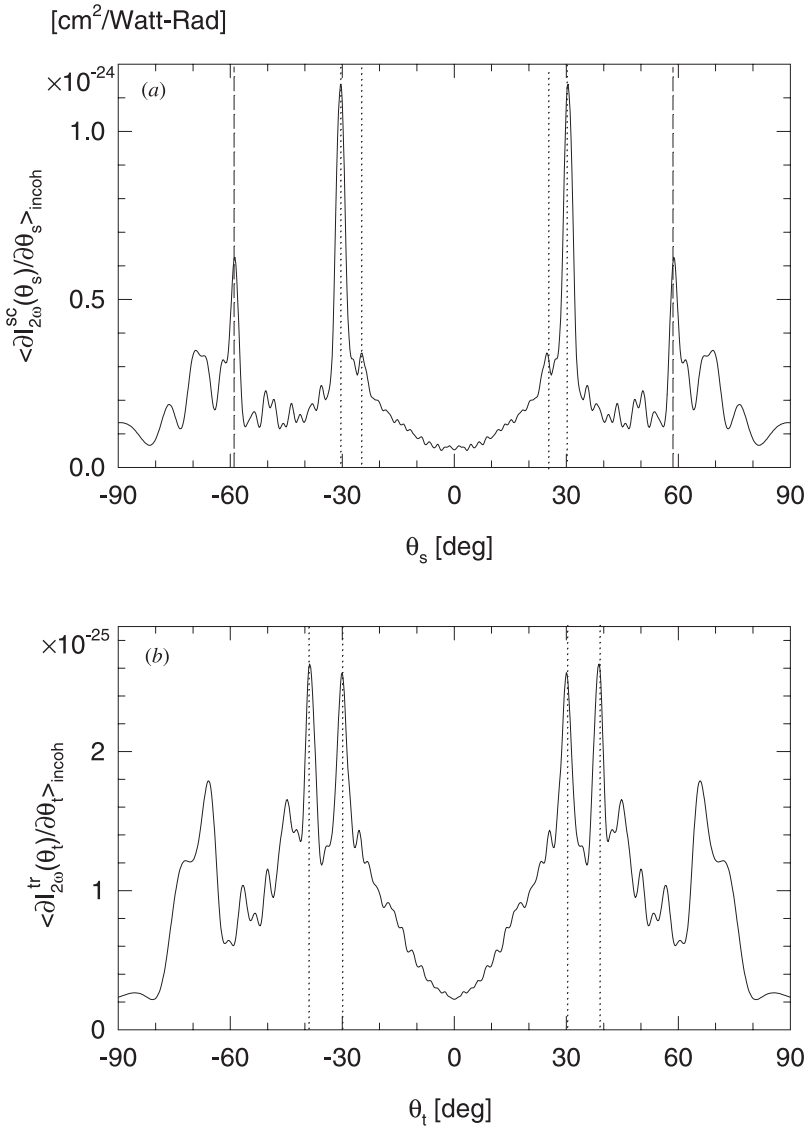

Figure 12. The differential reflection and transmission efficiencies of the second harmonic light in the Kretschmann geometry with a randomly rough silver film characterized by a Gaussian power spectrum. $L = 25.0\ \mu\text{m}, \lambda = 1.063\ \mu\text{m}, \theta_0 = 0^\circ, g = L/4, \delta = 10, a = 500, D = 50$ nm, $\epsilon_0 = 1.5, \epsilon(\omega) = -56.9 + \text{i}0.6, \epsilon(2\omega) = 11.56 + \text{i}0.37, N_p = 1000$. The vertical dotted lines indicate the predicted positions of the resonance peaks, the vertical broken lines indicate the positions of the leaky peaks.

We assume the following parameters for the scattering system: $\epsilon_0 = 16, \epsilon(\omega) = -56.9 + \text{i}0.6$, $D = 50$ nm. For these parameters we find $q_1(\omega) = 1.009\omega/c$ and $q_2(\omega) = 4.73\omega/c$. Therefore, there is only one optimal angle, $\theta_{opt} = 14.611^\circ$. To verify this result we calculate the reflectivity and transmittivity of this system as functions of the angle of incidence, using equations (B.25)–(B.29). The position of the dip in reflectivity or the position of the peak in transmittivity defines the optimal angle. We note that the peak in the transmissivity in the

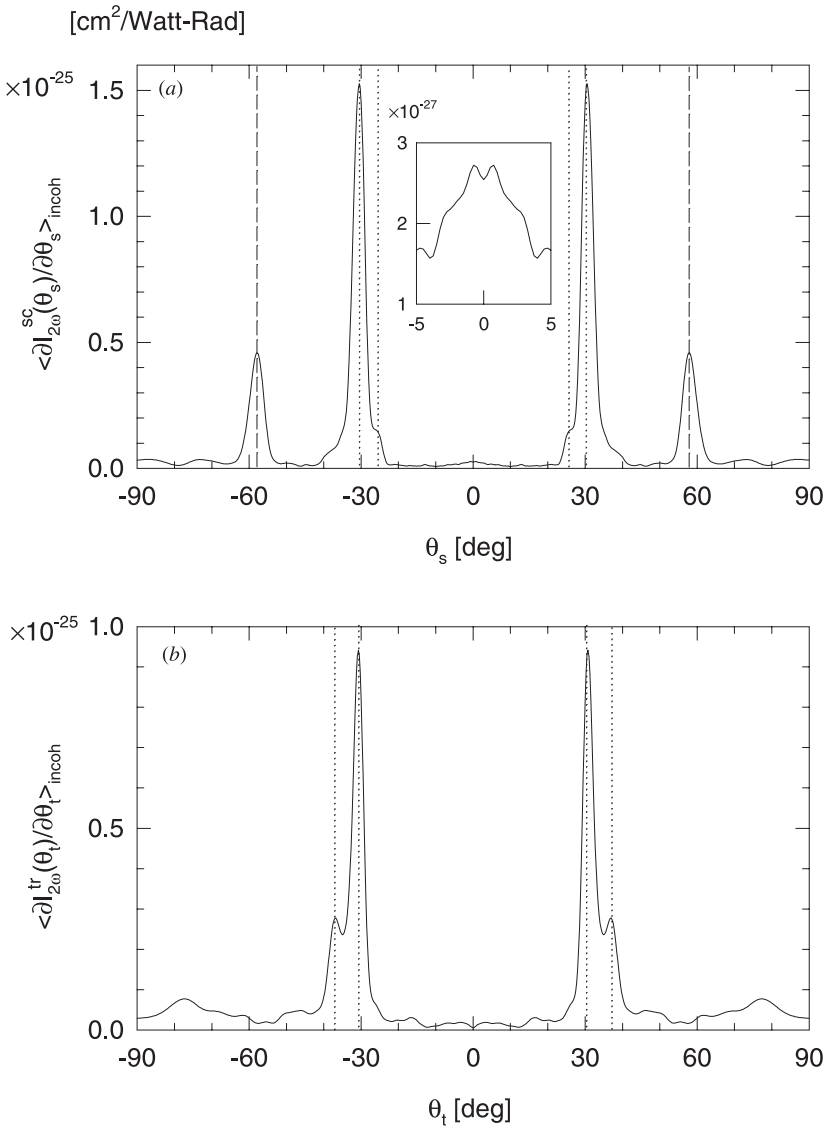

Figure 13. The differential reflection and transmission efficiencies of second harmonic light in the Kretschmann geometry with a randomly rough silver film characterized by the West–O'Donnell power spectrum. $k_{min} = 0.8605\omega/c$, $k_{max} = 1.58\omega/c$, $L = 25.0$ μm, $\lambda = 1.063$ μm, $\theta_0 = 5^\circ$, $g = L/4, \delta = 0.01$ μm, $d = 0.05$ μm, $\epsilon_0 = 1.5, \epsilon(\omega) = -56.9 + \mathrm{i}0.6$, $\epsilon(2\omega) = -11.56 + \mathrm{i}0.37$, $N_p = 1000$. The vertical dotted lines indicate the predicted positions of the resonance peaks, the vertical broken lines indicate the positions of the leaky peaks.

vacuum falls into the non-radiative region since $\sin\theta_t = \sqrt{\epsilon_0}\sin\theta_{opt} = 1.009 > 1$. From figure 14 we find that the optimal angle is equal to 14.613°, which is in good agreement with our previous estimate. We can see from equation (8.3) by using the condition (8.6) that there must be a resonance peak in the direction normal to the mean surface in both reflection and transmission. We pointed out earlier that for small angles of incidence the positions of the resonance peaks shift according to equation (8.3), and thus if the condition for the optimal

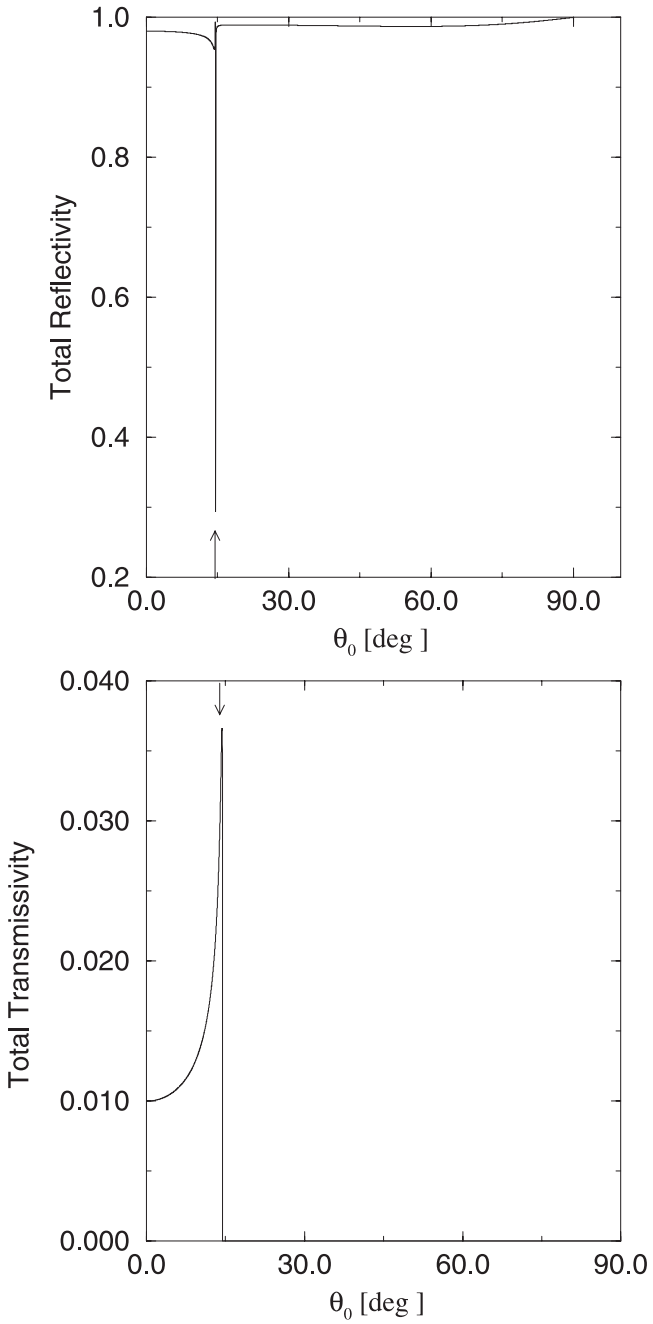

Figure 14. Plot of the reflectivity and transmissivity of p-polarized light in the Kretschmann geometry for a silver film with planar interfaces versus the angle of incidence. $D = 0.05\ \mu$m, $\epsilon_0 = 16$, $\epsilon(\omega) = -56.9 + \mathrm{i}0.6$. The vertical arrows indicate the predicted positions of the optimal angle.

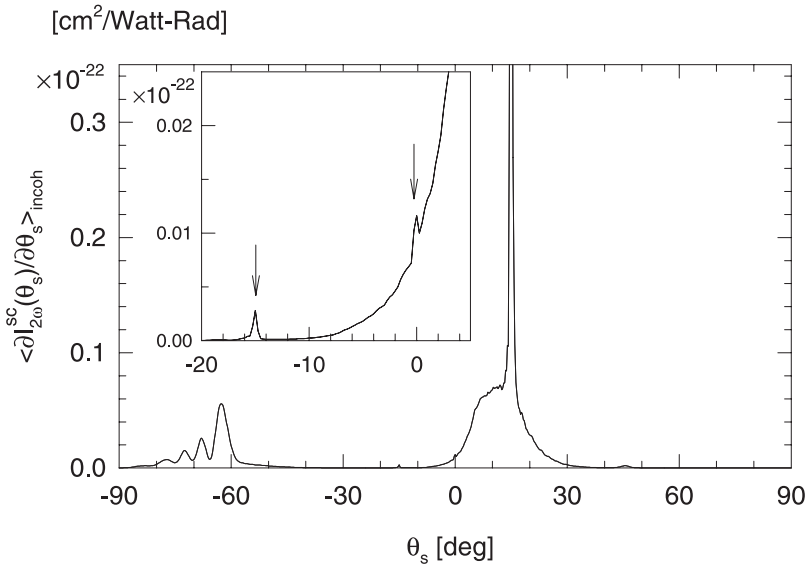

Figure 15. The differential reflection efficiencies of the second harmonic light in the Kretschmann geometry for a silver film with a randomly rough surface characterized by a Gaussian power spectrum. The angle of incidence is optimal for exciting the surface polariton. $L = 30.0\ \mu$m, $N = 400$, $\lambda = 1.063\ \mu$m, $\theta_0 = 14.61^\circ$, $g = L/5$, $\delta = 8$ nm, $a = 300$ nm, $D = 50$ nm, $\epsilon_0 = 15$, $\epsilon(\omega) = -56.9 + \mathrm{i}0.6$, $\epsilon(2\omega) = -11.56 + \mathrm{i}0.37$, $N_p = 1000$. The vertical arrows indicate the predicted positions of the resonance and backscattering peaks.

excitation of surface polaritons is satisfied, then one of the right-hand peaks in reflection or transmission occurs at a scattering angle equal to zero.

To verify these qualitative arguments we calculate the contributions to the mean differential reflection and transmission efficiencies of the second harmonic light when the angle of incidence is optimal, $\theta = 14.61^\circ$. The roughness of the lower, random surface of the metal film is characterized by the Gaussian power spectrum (2.5) and the roughness parameters $\delta = 8$ nm, $a = 300$ nm. We further assume that $\lambda = 1.063\ \mu$m, $\epsilon_0 = 16$, $D = 50$ nm, $\epsilon(\omega) = -56.9 + \mathrm{i}0.6$, $\epsilon(2\omega) = -11.56 + \mathrm{i}0.37$. For these values of the parameters we find from equation (3.1) that $q_1(\omega) = 1.009(\omega/c)$ and $q_2(\omega) = 4.73(\omega/c)$, while $q_1(2\omega)$ and $q_2(2\omega)$ do not exist. The results plotted in figures 15 and 16 show distinct peaks in the direction normal to the mean surface both in reflection and transmission. A well defined peak is present in the retroreflection direction, but there is no enhanced transmission peak. This is expected, because for these parameters the enhanced transmission peak should occur in the non-radiative region. Thus these calculations confirm the existence of peaks in the direction normal to the mean surface. However, the amplitudes of the peaks can be very small, especially in reflection.

9. Conclusions

In this work we have studied the scattering of p-polarized light from, and its transmission through, a thin metal film at both the frequency ω of the incident light and at the second harmonic frequency 2ω. The illuminated surface of the film was planar, while the back surface was a one-dimensional, randomly rough surface in contact with vacuum. In the calculations carried out at the fundamental frequency ω, two cases were considered: either the metal film was deposited on the planar base of a dielectric prism through which the light was incident

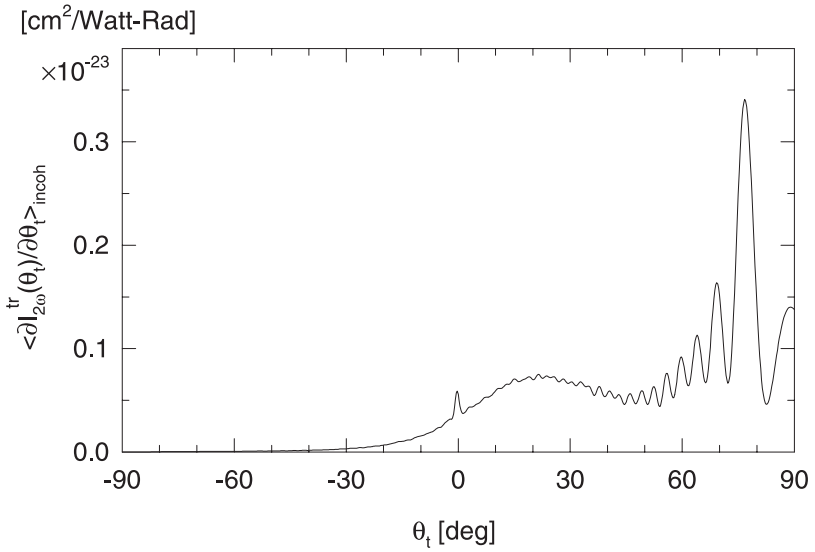

Figure 16. The differential transmission efficiency of the second harmonic light in the Kretschmann geometry with the same parameters as in figure 15.

(the Kretschmann attenuated total reflection geometry), or it was a free-standing film. In the calculations carried out at the harmonic frequency 2ω only the Kretschmann geometry was considered.

In the case of the linear scattering and transmission problem we have found that multiple-scattering effects in reflection are not as well pronounced as in transmission. This is due mainly to the fact that in order for constructive interference of multiply scattered optical paths to occur in reflection, each interfering optical path must cross a strongly attenuating film at least twice, while in transmission it must do so only once. Another reason is that single-scattering processes (the leaky peaks) contribute a large portion of the reflected intensity. Thus, the multiple-scattering effects are not well manifested in reflection.

We illustrated the latter argument by comparing the contributions to the mean differential reflection and transmission coefficients from the incoherent components of the scattered and transmitted light, for the Kretschmann geometry and for a free-standing film. In both cases the back interface was a one-dimensional random surface characterized by a Gaussian power spectrum. The only difference between these two geometries was that in the former case the medium of incidence was a dielectric prism, while in the latter case it was vacuum. However, this difference is essential, because the Kretschmann geometry supports a leaky surface mode, while the free-standing film does not. In transmission, weak, but well defined, satellite peaks were observed for both geometries, while in reflection they were observed only for the free-standing film. The enhanced backscattering and transmission peaks were not well pronounced in both geometries.

To enhance the multiple-scattering effects we utilized the West–O'Donnell power spectrum to characterize the roughness of the random surface of the film. Strong enhanced backscattering peaks were observed in the Kretschmann geometry. The intensity of the enhanced backscattering peak was approximately twice that of the background at its position. If the non-zero region of the West–O'Donnell power spectrum included the wavenumber of only one of the surface plasmon polaritons supported by the Kretschmann geometry, no satellite

peaks were observed. If the wavenumbers of both surface plasmon polaritons fall within this region, well defined satellite peaks were observed. These results confirm the physical mechanism for the origin of the satellite peaks as the constructive interference of two multiple-scattering sequences mediated by two degenerate surface plasmon polaritons with different wavenumbers.

Our study of the generation of second harmonic light in the Kretschmann geometry was stimulated by recent experimental studies of this effect, in which a peak in the angular dependence of the intensity of the transmitted second harmonic light was observed in the direction normal to the mean scattering surface [21, 23], at the same time that no peak in this direction was observed in another such study [22]. In our work it was assumed that the second harmonic radiation is generated in the prism–metal and metal–vacuum interfaces, where the components of the electromagnetic fields can vary rapidly in the directions normal to these interfaces. The results presented in section 8.2 indicate that if the angle of incidence is very small, then four resonance peaks occur in the angular dependence of the intensity of the second harmonic light, both in reflection and in transmission. These peaks arise from resonant processes in which surface plasmon polaritons of frequency ω mix nonlinearly with the incident light. Through the surface roughness the incident light excites two surface plasmon polaritons of frequency ω, which propagate in the forward and backward directions, and then interact nonlinearly with the incident light to produce four peaks at angles of scattering and transmission defined by equations (8.3*a*) and (8.3*b*), respectively. When the angle of incidence is optimal for exciting the surface plasmon polaritons, as was the case in the experimental studies, one of these resonance peaks is in the direction normal to the mean scattering surface. These peaks are well defined. However, their amplitudes are small. We also observe a peak in the retroreflection direction, whose amplitude, however, is small. No enhanced transmission peak is observed, because it should occur in the non-radiative region. Thus, our results confirm the existence of peaks in the retroreflection and normal directions.

We emphasize that the work presented here covers only several aspects of multiple-scattering phenomena in second harmonic generation. The analytical and numerical study of the second harmonic generation of light in three-layer geometries is far from over. The analytical theories are limited by the necessary assumption of weak roughness. On the other hand, the numerical approach needs to be improved to handle the long mean free paths of surface polaritons that are often the case for weakly rough surfaces.

Finally, we have demonstrated the efficacy of a new numerical approach to the solution of the integral equations for the source functions arising in a rigorous approach to the scattering of light from a one-dimensional random surface. Based on the Gauss–Legendre numerical integration scheme, this approach gives accurate results for large values of ϵ_0 and $\epsilon(\omega)$, and provides an efficient set of abscissas in coordinate space for a particular scattering problem, whose number can be significantly smaller than when a set of uniformly spaced abscissas is used to obtain the same level of accuracy. One of the possible future applications of this method is to the non-uniform rescaling of a long segment of the scattering surface into a shorter segment, which can be useful for scattering problems with grazing angles of incidence, or with the excitation of surface or guided waves with long mean free paths.

Acknowledgments

The research of IVN, AAM and TAL was supported in part by Army Research Office Grants DAAH 0-96-1-0187 and DAAD 19-99-1-0321. The research of ERM and ML-L was supported in part by CONACYT grant 32254-E.

Appendix A. Kernels

The kernels appearing in equations (5.3)–(5.6) are defined by

$$H_p^{DD}(x_1|x_1') = \frac{1}{4\pi}\left[\frac{\partial}{\partial x_3'}G_{\epsilon_0}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=D+\eta \\ x_3'=D}} \tag{A.1a}$$

$$L_p^{DD}(x_1|x_1') = \frac{1}{4\pi}\left[G_{\epsilon_0}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=D+\eta \\ x_3'=D}} \tag{A.1b}$$

$$H_\epsilon^{DD}(x_1|x_1') = \frac{1}{4\pi}\left[\frac{\partial}{\partial x_3'}G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=D+\eta \\ x_3'=D}} \tag{A.1c}$$

$$L_\epsilon^{DD}(x_1|x_1') = \frac{1}{4\pi}\left[G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=D+\eta \\ x_3'=D}} \tag{A.1d}$$

$$H_\epsilon^{DO}(x_1|x_1') = \frac{1}{4\pi}\left[\frac{\partial}{\partial N'}G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=D \\ x_3'=\zeta(x_1')}} \tag{A.1e}$$

$$L_\epsilon^{DO}(x_1|x_1') = \frac{1}{4\pi}\left[G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=D \\ x_3'=\zeta(x_1')}} \tag{A.1f}$$

$$H_\epsilon^{OD}(x_1|x_1') = \frac{1}{4\pi}\left[\frac{\partial}{\partial x_3'}G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=\zeta(x_1) \\ x_3'=D}} \tag{A.1g}$$

$$L_\epsilon^{OD}(x_1|x_1') = \frac{1}{4\pi}\left[G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=\zeta(x_1) \\ x_3'=D}} \tag{A.1h}$$

$$H_\epsilon^{OO}(x_1|x_1') = \frac{1}{4\pi}\left[\frac{\partial}{\partial N'}G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}} \tag{A.1i}$$

$$L_\epsilon^{OO}(x_1|x_1') = \frac{1}{4\pi}\left[G_{\epsilon}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}} \tag{A.1j}$$

$$H_0^{OO}(x_1|x_1') = \frac{1}{4\pi}\left[\frac{\partial}{\partial N'}G_{0}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}} \tag{A.1k}$$

$$L_0^{OO}(x_1|x_1') = \frac{1}{4\pi}\left[G_{0}^{(\omega)}(x_1, x_3|x_1', x_3')\right]_{\substack{x_3=\zeta(x_1)+\eta \\ x_3'=\zeta(x_1')}}. \tag{A.1l}$$

Appendix B. The non-uniform grid method

The systems of equations (5.3)–(5.6) and (7.1)–(7.4) for the source functions at the fundamental and harmonic frequencies, respectively, are solved by converting them into matrix equations. This is done in two steps [7]. In the first, the infinite range of integration is replaced by the finite range $(-L/2, L/2)$. In the second, integration over this interval is replaced by summation through the use of a numerical quadrature scheme. In this paper we have carried out the second step in a somewhat non-standard way, which we describe in this appendix and compare with

the more usual approach [7]. To simplify the presentation, without loss of generality we apply it to an inhomogeneous Fredholm equation of the second kind

$$H(x_1) = G(x_1) + \int_a^b \mathrm{d}x_1'\, K(x_1|x_1')H(x_1') \qquad a \leqslant x_1 \leqslant b. \tag{B.1}$$

The kernel in this equation has one of two forms. It is either

$$\begin{aligned} K_1(x_1|x_1') &= \frac{1}{4\pi} G(x_1, x_3|x_1', x_3')\Big|_{\substack{x_3'=\zeta(x_1') \\ x_3=\zeta(x_1)+\eta}} \\ &= \frac{\mathrm{i}}{4} H_0^{(1)}\left(\sqrt{\epsilon}\frac{\Omega}{c}[(x_1 - x_1')^2 + (\zeta(x_1) - \zeta(x_1') + \eta)^2]^{1/2}\right) \end{aligned} \tag{B.2}$$

or

$$\begin{aligned} K_2(x_1|x_1') &= \frac{1}{4\pi}\left[\left(-\zeta'(x_1')\frac{\partial}{\partial x_1'} + \frac{\partial}{\partial x_3'}\right) G(x_1, x_3|x_1', x_3')\right]_{\substack{x_3'=\zeta(x_1') \\ x_3=\zeta(x_1)+\eta}} \\ &= \frac{\mathrm{i}}{4}\sqrt{\epsilon}\frac{\Omega}{c}\frac{H_1^{(1)}(\sqrt{\epsilon}\frac{\Omega}{c}[(x_1 - x_1')^2 + (\zeta(x_1) - \zeta(x_1') + \eta)^2]^{1/2})}{[(x_1 - x_1')^2 + (\zeta(x_1) - \zeta(x_1') + \eta)^2]^{1/2}} \\ &\quad \times[-(x_1 - x_1')\zeta'(x_1') + (\zeta(x_1) - \zeta(x_1') + \eta)]. \end{aligned} \tag{B.3}$$

In these expressions ϵ is either ϵ_0, $\epsilon(\Omega)$, or 1, η is a positive infinitesimal, and we are interested in the limit as $\eta \to 0+$. We note that both $K_1(x_1|x_1')$ and $K_2(x_1|x_1')$ are singular along the diagonal line $x_1' = x_1$.

In a common approach to the solution of equation (B.1) [7] the interval (a, b) is divided into N intervals of equal length, and the integral in it is written as

$$I(x_1) = \int_a^b \mathrm{d}x_1'\, K(x_1|x_1')H(x_1') = \sum_{n=1}^{N} \int_{x_n - \frac{1}{2}\Delta x}^{x_n + \frac{1}{2}\Delta x} \mathrm{d}x_1'\, K(x_1|x_1')H(x_1') \tag{B.4}$$

where $x_n = a + (n - \frac{1}{2})\Delta x$ and $\Delta x = (b - a)/N$. Then, on the assumption that $H(x_1')$ is a slowly varying function of x_1' in each of the N intervals $(x_n - \frac{1}{2}\Delta x, x_n + \frac{1}{2}\Delta x)$, we evaluate it at the midpoint of each interval, and remove it from the integral:

$$\begin{aligned} I(x_1) &\cong \sum_{n=1}^{N} \left\{ \int_{x_n - \frac{1}{2}\Delta x}^{x_n + \frac{1}{2}\Delta x} \mathrm{d}x_1'\, K(x_1|x_1') \right\} H(x_n) \\ &= \sum_{n=1}^{N} \left\{ \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathrm{d}u\, K(x_1|x_n + u) \right\} H(x_n). \end{aligned} \tag{B.5}$$

In general, the integral over u in this result is evaluated to the lowest non-zeroth order in Δx. When equation (B.5) is substituted into equation (B.1), and x_1 is replaced by x_m $(m = 1, 2, \ldots, N)$, we obtain the $N \times N$ matrix equation

$$H(x_m) = G(x_m) + \sum_{n=1}^{N} \left\{ \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathrm{d}u\, K(x_m|x_n + u) \right\} H(x_n) \qquad m = 1, 2, \ldots, N \tag{B.6}$$

for determining the values of $H(x)$ at the points $\{x_m\}$.

One of the advantages of the method of solution described is that its implementation is straightforward. Another advantage is that the singularity of the kernel can be treated analytically. However, in the evaluation of the singular diagonal elements analytically, care

must be exercised. To see this consider the evaluation of the diagonal elements for the kernel $K_1(x|x')$. We have that

$$\int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathrm{d}u\, K_1(x_m|x_m+u) = \frac{\mathrm{i}}{4}\int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathrm{d}u\, H_0^{(1)}(a_m|u|) + \mathrm{O}((\Delta x)^3) \tag{B.7}$$

where $a_m = \sqrt{\epsilon}(\Omega/c)[1 + (\zeta'(x_m))^2]^{1/2}$. If we use the result that for small z $H_0^{(1)}(z) = (2\mathrm{i}/\pi)[\ln(z/2) + \gamma] + 1 + \mathrm{O}(z^2 \ln z)$, and rewrite the right-hand side of this equation as

$$\frac{\mathrm{i}}{2a_m}\int_0^{\frac{1}{2}a_m\Delta x} \mathrm{d}u\, H_0^{(1)}(u) = \frac{\mathrm{i}}{4}\Delta x H_0^{(1)}\left(\frac{a_m\Delta x}{2e}\right) + \mathrm{O}\left(\left(\frac{a_m\Delta x}{2}\right)^2 \ln\left(\frac{a_m\Delta x}{2}\right)\right) \tag{B.8}$$

we see that to obtain a diagonal matrix element when $|\sqrt{\epsilon}| > 1$ that is comparable in accuracy to its value when $|\sqrt{\epsilon}| \equiv 1$, we either have to keep higher-order terms in Δx on the right-hand side of equation (B.8), which erodes the simplicity of the implementation of the approach described, or we have to reduce Δx by a factor of $|\sqrt{\epsilon}|$. If the limits of integration are kept fixed at $x_1 = \pm L/2$, the latter option means that the number of abscissas $\{x_m\}$ must be increased from N to $|\sqrt{\epsilon}|N$. Although the difficulties arising in calculations in which a large value of $|\sqrt{\epsilon}|$ is present can be avoided by the use of an impedance boundary condition [39,40], there are limitations on the use of such a boundary condition [39,40]. A final drawback of the method based on a set of uniformly spaced abscissas is its slow rate of convergence as the number of abscissas N is increased. Similar problems arise in the evaluation of the diagonal matrix elements of the kernel $K_2(x_1|x_1')$. Therefore, it is useful to have a computational method that is free from these difficulties.

It is known [41] that it is often possible to obtain more accurate results in the numerical evaluation of an integral by the use of irregularly spaced abscissas, as in Gauss–Legendre integration, than when the same number of equally spaced abscissas is used. Thus, in what follows we consider the use of Gauss–Legendre integration for converting equation (B.1) into a matrix equation. However, this approach easily allows changing of the integration rule without substantial changes in the algorithm.

A straightforward application of an n-point numerical integration rule to equation (B.1) results in the matrix equation

$$H_m = G_m + \sum_{n=1}^{N} K_{mn} w_n H_n \tag{B.9}$$

where $H_m = H(x_m)$, $G_m = G(x_m)$, $K_{mn} = K(x_m|x_n)$, $\{x_m\}$ are the abscissas of the integration scheme adopted and $\{w_m\}$ are the corresponding weights. In the present case such a straightforward application of the Gauss–Legendre method fails, because the kernel $K(x_1|x_1')$ is evaluated at the singular points $x_m = x_n$. To overcome this problem we use the method of subtraction of singularities [42]. In this method we write the integral term on the right-hand side of equation (B.1) in the form

$$\int_a^b \mathrm{d}x_1'\, K(x_1|x_1')H(x_1') = \int_a^b \mathrm{d}x_1'\, K(x_1|x_1')[H(x_1') - H(x_1)] + H(x_1)R(x_1) \tag{B.10}$$

where

$$R(x_1) = \int_a^b \mathrm{d}x_1'\, K(x_1|x_1'). \tag{B.11}$$

We first consider the evaluation of $R(x_1)$, which we rewrite as

$$R(x_1) = \int_a^{x_1-\delta/2} \mathrm{d}x_1' \, K(x_1|x_1') + \int_{x_1+\delta/2}^{b} \mathrm{d}x_1' \, K(x_1|x_1') + \int_{x_1-\delta/2}^{x_1+\delta/2} \mathrm{d}x_1' \, K(x_1|x_1'). \tag{B.12}$$

The contribution to $R(x_1)$ from the first two terms on the right-hand side of this equation can be calculated numerically by applying the standard Gauss–Legendre rule in the open intervals $(a, x_1 - \delta/2)$ and $(x_1 + \delta/2, b)$. For these calculations we use $\delta = 10^{-6}L$. The last term on the right-hand side of this equation requires an analytic evaluation for each of the two types of kernels appearing in the integral equations for the source functions, $K_1(x_1|x_1')$ and $K_2(x_1|x_1')$. In the former case we have

$$\begin{aligned}\int_{x_1-\delta/2}^{x_1+\delta/2} \mathrm{d}x_1' \, K_1(x_1|x_1') &= \tfrac{1}{4}\mathrm{i} \int_{-\delta/2}^{\delta/2} \mathrm{d}u \, H_0^{(1)}(a|u|) + \mathrm{O}(\delta^3) \\ &= \tfrac{1}{4}\mathrm{i}\delta H_0^{(1)}\left(\frac{a\delta}{2e}\right) + \mathrm{O}(\delta^3 \ln \delta).\end{aligned} \tag{B.13}$$

Consequently, we can neglect this contribution for the small value of δ we have assumed. Thus, we obtain the result that

$$\int_a^b \mathrm{d}x_1' \, K_1(x_1|x_1') = P \int_a^b \mathrm{d}x_1' \, K_1(x_1|x_1') \tag{B.14}$$

where P denotes the Cauchy principal value, and the integral is evaluated numerically in the manner indicated above. For the second type of kernel we find that

$$\begin{aligned}\int_{x_1-\delta/2}^{x_1+\delta/2} \mathrm{d}x_1' \, K_2(x_1|x_1') &= \lim_{\eta\to 0+} \frac{1}{2\pi} \int_{-\delta/2}^{\delta/2} \mathrm{d}u \, \frac{\eta + \frac{1}{2}\zeta''(x_1)u^2}{(\phi^2(x_1) - \eta\zeta''(x_1))u^2 - 2\eta\zeta'(x_1)u + \eta^2} \\ &\quad + \mathrm{O}(\delta^3 \ln \delta) \\ &= \lim_{\eta\to 0+} \frac{1}{2\pi} \int_{-\delta/2\eta}^{\delta/2\eta} \mathrm{d}u \, \frac{1}{\phi^2(x_1) - 2\zeta'(x_1)u + 1} + \frac{1}{4\pi}\frac{\zeta''(x_1)}{\phi^2(x_1)} \int_{-\delta/2}^{\delta/2} \mathrm{d}u + \mathrm{O}(\delta^3 \ln \delta) \\ &= \frac{1}{2} + \frac{\zeta''(x_1)}{4\pi\phi^2(x_1)}\delta + \mathrm{O}(\delta^3 \ln \delta).\end{aligned} \tag{B.15}$$

Thus, we have the result that

$$\int_a^b \mathrm{d}x_1' \, K_2(x_1|x_1') = \tfrac{1}{2} + P \int_a^b \mathrm{d}x_1' \, K_2(x_1|x_1'). \tag{B.16}$$

We next turn to the evaluation of the first term on the right-hand side of equation (B.10). We assume that $H(x_1)$ is continuous and has a continuous derivative for x_1 in the interval (a, b). We then write this term as

$$\begin{aligned}&\int_a^b \mathrm{d}x_1' \, K(x_1|x_1')[H(x_1') - H(x_1)] \\ &\quad = \left[\int_a^{x_1-\delta/2} + \int_{x_1-\delta/2}^{x_1+\delta/2} + \int_{x_1+\delta/2}^{b}\right] \mathrm{d}x_1' \, K(x_1|x_1')[H(x_1') - H(x_1)] \\ &\quad = P\int_a^b \mathrm{d}x_1' \, K(x_1|x_1')[H(x_1') - H(x_1)] \\ &\qquad + \int_{x_1-\delta/2}^{x_1+\delta/2} \mathrm{d}x_1' \, K(x_1|x_1')[H(x_1') - H(x_1)].\end{aligned} \tag{B.17}$$

Under our assumptions the product $K(x_1|x_1')[H(x_1') - H(x_1)]$ has a weaker singularity at $x_1' = x_1$ than does $K(x_1|x_1')$ itself. Thus we have that

$$\begin{aligned} J(x_1) &= \int_{x_1-\delta/2}^{x_1+\delta/2} \mathrm{d}x_1'\, K(x_1|x_1')[H(x_1') - (H(x_1)] \\ &= \int_{-\delta/2}^{\delta/2} \mathrm{d}u\, K(x_1|x_1+u)\big[uH'(x_1) + \tfrac{1}{2}u^2H''(x_1) + \cdots\big]. \end{aligned} \tag{B.18}$$

For the first type of kernel we find that

$$\begin{aligned} J_1(x_1) &\cong H'(x_1)\int_{-\delta/2}^{\delta/2} \mathrm{d}u\, uK_1(x_1|x_1+u) + \tfrac{1}{2}H''(x_1)\int_{-\delta/2}^{\delta/2} \mathrm{d}u\, u^2K_1(x_1|x_1+u) \\ &= -H'(x_1)\frac{\mathrm{i}}{4}\sqrt{\epsilon}\frac{\Omega}{c}\frac{\zeta'(x_1)\zeta''(x_1)}{2\phi(x_1)}\int_{-\delta/2}^{\delta/2} \mathrm{d}u\, |u|u^2H_1^{(1)}\left(\sqrt{\epsilon}\frac{\Omega}{c}\phi(x_1)|u|\right) \\ &\quad + \frac{1}{2}H''(x_1)\frac{\mathrm{i}}{4}\int_{-\delta/2}^{\delta/2} \mathrm{d}u\, u^2H_0^{(1)}\left(\sqrt{\epsilon}\frac{\Omega}{c}\phi(x_1)|u|\right) + \cdots \\ &= 0(\delta^3\ln\delta). \end{aligned} \tag{B.19}$$

For the second type of kernel we find that

$$\begin{aligned} J_2(x_1) &\cong H'(x_1)\int_{-\delta/2}^{\delta/2} \mathrm{d}u\, uK_2(x_1|x_1+u) + \tfrac{1}{2}H''(x_1)\int_{-\delta/2}^{\delta/2} \mathrm{d}u\, u^2K_2(x_1|x_1+u) + \cdots \\ &= \lim_{\eta\to 0+}\frac{H'(x_1)\eta}{2\pi}\int_{-\delta/2\eta}^{\delta/2\eta} \mathrm{d}u\, \frac{u}{\phi^2(x_1)u^2 - 2\zeta'(x_1)u + 1} \\ &\quad + \frac{\zeta''(x_1)H'(x_1)}{4\pi\phi^2(x_1)}\int_{-\delta/2}^{\delta/2} \mathrm{d}u\, u + \mathrm{O}(\delta^3). \end{aligned} \tag{B.20}$$

Thus $J_2(x_1) = \mathrm{O}(\delta^3)$ because the first integral vanishes as $\eta \to 0^+$, and the second term is exactly zero. As a result, we can set $J_{1,2}(x_1)$ equal to zero for the small value of δ we have assumed, and we obtain

$$\int_a^b \mathrm{d}x_1'\, K(x_1|x_1')[H(x_1') - H(x_1)] = P\int_a^b \mathrm{d}x_1'\, K(x_1|x_1')[H(x_1') - H(x_1)]. \tag{B.21}$$

When we put together the preceding results we find that the matrix equation equivalent to equation (B.1) is

$$H_m = G_m + \sum_{n=1}^{N}\left\{K_{mn}^{(1)}w_n + [R_m - \sum_{\ell=1}^{N}K_{m\ell}^{(1)}w_\ell]\delta_{mn}\right\}H_n \tag{B.22}$$

in the case that the kernel is $K_1(x_1|x_1')$, where $H_m = H(x_m)$, $G_m = G(x_m)$, $K_{mn}^{(1)} = K_1(x_m|x_n)$, with $K_{mm}^{(1)} = 0$, and

$$R_m = \int_a^{x_m-\delta/2} \mathrm{d}x_1'\, K_1(x_m|x_1') + \int_{x_m+\delta/2}^{b} \mathrm{d}x_1'\, K_1(x_m|x_1'). \tag{B.23}$$

In the case that the kernel is $K_2(x_1|x_1')$ we obtain the matrix equation

$$H_m = 2G_m + 2\sum_{n=1}^{N}\left\{K_{mn}^{(2)}w_n + [R_m - \sum_{\ell=1}^{N}K_{m\ell}^{(2)}w_\ell]\delta_{mn}\right\}H_n \tag{B.24}$$

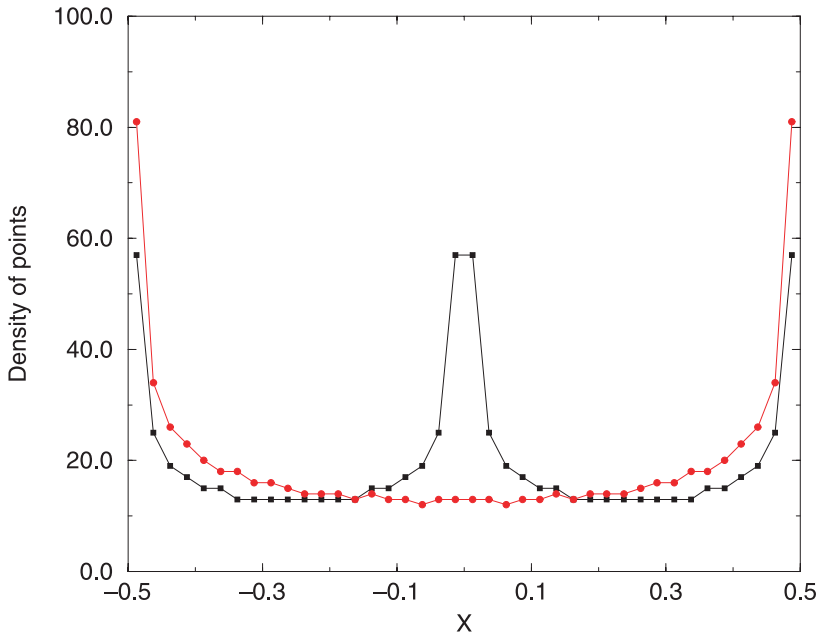

Figure B1. The density of points on the surface. The total number of sampling points $N = 800$. The surface is divided into 40 equally spaced intervals. The curve with the circles represents the density of points for grid 1. The curves with the squares represent the density of points for grid 2. (This figure is in colour only in the electronic version, see `www.iop.org`)

where $K^{(2)}_{mn} = K_2(x_m|x_n)$, with $K^{(2)}_{mm} = 0$, and H_m, G_m, and R_m are defined as in the case of the kernel $K_1(x_1|x_1')$. The difference between equations (B.22) and (B.24) arises from the presence of the first term on the right-hand side of equation (B.16).

With the original integral equation (B.1) transformed into an equivalent matrix equation, we turn to the question of the optimal choice of the abscissas to be used in defining the elements H_m, G_m, R_m and K_{mn}. We require that they be highly dense in the vicinity of the central point $x_1 = 0$. This is a necessary condition because the incident beam has a Gaussian form. Hence the surface is most strongly illuminated in the centre. In figure B1 we plot the density of abscissas for the Gauss–Legendre rule for the interval $-0.5 < x_1 < 0.5$ (the curve with circles). We use the rescaled interval $-0.5 < x_1 < 0.5$ instead of the usual interval $-1 < x_1 < 1$ because it is convenient to have the total length of the surface equal unity in dimensionless units. In obtaining this figure we have assumed that the total number of abscissas is $N = 800$, and have divided the surface into 40 equally spaced intervals (grid 1). The density of abscissas is therefore the number of abscissas in a given one of these intervals. We can see that this set of abscissas has a high density at the edges (80 points per interval), but the most important region, in the centre, is not covered effectively (13 points per interval). To increase the number of abscissas in the centre of the surface we rescale the Gauss–Legendre rule into the interval $-0.5 < x_1 < 0$, using $N/2 = 400$ abscissas (grid 2). Then in the interval $0 < x_1 < 0.5$ we use the same set of abscissas reflected about the origin of coordinates. In figure B1 we plot the density of the abscissas obtained in this way (the curve with squares). The central region now has the same density of abscissas as the edges do (57 points per interval). After the abscissas $\{x_m\}$ have been determined as the zeros of $P_{N/2}(x_1)$ in the rescaled interval $-0.5 < x_1 < 0$ and then, by reflection, in the interval

$0 < x_1 < 0.5$, the corresponding weights $\{w_n\}$ are obtained by the use of standard recursive algorithms [37].

It is now necessary to generate the surface profile function $\zeta(x_1)$ at the values of the abscissas that were generated in the manner just described. We do this by using a numerical technique [7] that generates $\zeta(x_1)$ and its derivatives at a set of N_0 equally spaced points within the finite interval of integration $(-L/2, L/2)$. We assume that N_0 is at least twice as large as N. Then by interpolation we obtain $\zeta(x_1)$ and its derivatives $\zeta'(x_1)$ and $\zeta''(x_1)$ at any intermediate point between those at which they are generated. The matrix elements for each of the kernels given by equations (A.1) can now be evaluated readily.

To check the accuracy of the non-uniform grid method relative to that of the uniform grid method, we apply both methods to a problem for which an analytic solution is known, namely the reflectivity (R) and transmissivity (T) of a metal film with planar surfaces in the Kretschmann geometry. We define R and T as the ratios of the total time-averaged scattered and transmitted fluxes to the total time-averaged incident flux. Assuming the incident field to be a plane wave, we have the results that

$$R = \frac{|r(k)|^2}{|\Delta(k)|^2} \qquad T = \frac{\epsilon_0 \alpha_0(k)}{\alpha_p(k)} \frac{|t(k)|^2}{|\Delta(k)|^2} \tag{B.25}$$

where

$$r(k) = [\epsilon(\omega)\alpha_p(k)\alpha(k) - \epsilon_0\epsilon(\omega)\alpha(k)\alpha_0(k)]\cos[\alpha(k)D] + \mathrm{i}[\epsilon_0\alpha^2(k) - \epsilon^2(\omega)\alpha_p(k)\alpha_0(k)]\sin[\alpha(k)D] \tag{B.26}$$

$$t(k) = 2\epsilon(\omega)\alpha_p(k)\alpha(k) \tag{B.27}$$

$$\Delta(k) = [\epsilon(\omega)\alpha_p(k)\alpha(k) + \epsilon_0\epsilon(\omega)\alpha(k)\alpha_0(k)]\cos[\alpha(k)D] - \mathrm{i}[\epsilon_0\alpha^2(k) + \epsilon^2(\omega)\alpha_p(k)\alpha_0(k)]\sin[\alpha(k)D] \tag{B.28}$$

and

$$\alpha_p(k) = [\epsilon_0(w/c)^2 - k^2]^{1/2} \qquad \mathrm{Re}\,\alpha_p(k) > 0 \qquad \mathrm{Im}\,\alpha_p(k) > 0 \tag{B.29a}$$

$$\alpha(k) = [\epsilon(\omega)(\omega/c)^2 - k^2]^{1/2} \qquad \mathrm{Re}\,\alpha(k) > 0 \qquad \mathrm{Im}\,\alpha(k) > 0 \tag{B.29b}$$

$$\alpha_0(k) = [(\omega/c)^2 - k^2]^{1/2} \qquad \mathrm{Re}\,\alpha_0(k) > 0 \qquad \mathrm{Im}\,\alpha_0(k) > 0. \tag{B.29c}$$

The wavenumber k here is related to the angle of incidence θ_0 in the prism by $k = \sqrt{\epsilon_0}(\omega/c)\sin\theta_0$. If we assume for the experimental parameters the values $\lambda = 612.7$, $D = 48$ nm, $\epsilon_0 = 9$, $\epsilon(\omega) = -17.2 + \mathrm{i}0.498$ and $L = 12.6$ μm, it follows from equation (B.25) that at normal incidence ($k = 0$) $R = 0.941$ and $T = 0.002\,83$. The results based on the uniform and non-uniform grid methods with $N = 500$ are $R = 0.899$, $T = 0.002\,00$ and $R = 9.939$, $T = 0.002\,85$, respectively, when a plane wave is used for the incident field. The results obtained by the non-uniform grid method are much closer to their analytical counterparts.

We have tested the non-uniform grid method for different parameters, and have found that in many cases it gives more accurate results than the uniform grid method does. In particular, the solution is robust in the range of large values of ϵ_0 and $\epsilon(\omega)$, where the utility of uniform grid methods can be limited.

References

[1] McGurn A R, Maradudin A A and Celli V 1985 Localization effects in the scattering of light from a randomly rough grating *Phys. Rev.* B **31** 4866–71

[2] Maradudin A A and Méndez E R 1994 Enhanced backscattering of light from weakly rough, random metal surfaces *Appl. Opt.* **32** 3335–43

[3] West C S and O'Donnell K A 1995 Observations of backscattering enhancement from polaritons on a rough metal surface *J. Opt. Soc. Am.* A **12** 390–7

[4] Méndez E R and O'Donnell K A 1987 Observation of depolarization and backscattering enhancement in light scattering from Gaussian random surfaces *Opt. Commun.* **61** 91–5

[5] O'Donnell K A and Méndez E R 1987 An experimental study of scattering from characterized random surfaces *J. Opt. Soc. Am.* A **4** 1194–205

[6] Nieto-Vesperinas M and Soto-Crespo J M 1987 Monte Carlo simulations for scattering of electromagnetic waves from perfectly conductive random rough surfaces *Opt. Lett.* **12** 979–81

[7] Maradudin A A, Michel T, McGurn A R and Méndez E R 1990 Enhanced backscattering of light from a random grating *Ann. Phys., NY* **203** 255–307

[8] Agranovich V M and Grigorishin K I 1993 Effects of weak localization of photons in nonlinear optics of disordered media *Nonlinear Opt.* **5** 3–11

[9] McGurn A R, Leskova T A and Agranovich V M 1991 Weak-localization effects in the generation of second harmonics of light at a randomly rough vacuum–metal grating *Phys. Rev.* B **44** 11441–56

[10] Leyva-Lucero M, Méndez E R, Leskova T A, Maradudin A A and Lu Jun Q 1996 Multiple-scattering effects in the second-harmonic generation of light in reflection from a randomly rough surface *Opt. Lett.* **21** 1809–11

[11] O'Donnell K A, Torre R and West C S 1996 Observations of backscattering effects in second-harmonic generation from a weakly rough metal surface *Opt. Lett.* **21** 1738–40

[12] O'Donnell K A, Torre R and West C S 1997 Observations of second harmonic generation from randomly rough metal surfaces *Phys. Rev.* B **55** 7985–92

[13] Leskova T A, Maradudin A A, Leyva-Lucero M and Méndez E R (unpublished)

[14] O'Donnell K A and Torre R 1997 Second-harmonic generation from a strongly rough metal surface *Opt. Commun.* **138** 341–4

[15] Leyva-Lucero M, Méndez E R, Leskova T A and Maradudin A A 1999 Destructive interference effects in the second harmonic light generated at randomly rough surfaces *Opt. Commun.* **161** 79–94

[16] Kretschmann E 1971 The determination of the optical constants of metals by excitation of surface plasmons *Z. Phys.* **241** 313–24

[17] Wang X and Simon H J 1991 Directionally scattered optical second-harmonic generation with surface plasmons *Opt. Lett.* **16** 1475–7

[18] Simon H J, Wang Y, Zhou L-B and Chan Z 1992 Coherent backscattering of optical second-harmonic generation with long-range surface plasmons *Opt. Lett.* **17** 1268–70

[19] Wang Y and Simon H J 1993 Coherent backscattering of optical second-harmonic generation in silver films *Phys. Rev.* B **47** 13 695–9

[20] Sarid D 1981 Long-range surface-plasma waves on very thin metal films *Phys. Rev. Lett.* **47** 1927–30

[21] Aktsipetrov O K, Golovkina V N, Kapusta O I, Leskova T A and Novikova N N 1992 Anderson localization effects in the second harmonic generation at a weakly rough metal surface *Phys. Lett.* A **170** 231–4

[22] Kuang L and Simon H J 1995 Diffusely scattered second harmonic generation from a silver film due to surface plasmons *Phys. Lett.* A **197** 257–61

[23] Bozhevolnyi S I and Pedersen K 1997 Second harmonic generation due to surface plasmon localization *Surf. Sci.* **377–9** 384–7

[24] Leskova T A, Leyva-Lucero M, Méndez E R, Maradudin A A and Novikov I V 2000 The surface enhanced second harmonic generation of light from a randomly rough metal surface in the Kretschmann geometry *Opt. Commun.* **183** 529–45

[25] Lu Jun Q and Maradudin A A 1997 Surface plasmon polaritons in light scattering from free-standing randomly rough thin metal films *SPIE* **3141** 186–94

[26] Sánchez-Gil J A, Maradudin A A, Lu Jun Q, Freilikher V D, Pustilnik M and Yurkevich I 1994 Scattering of electromagnetic waves from a bounded medium with a random surface *Phys. Rev.* B **50** 15 353–68

[27] Sánchez-Gil A J, Maradudin A A, Lu Jun Q and Freilikher V D 1995 Transmission of electromagnetic waves through thin metal films with randomly rough surfaces *Phys. Rev.* B **51** 17 100–15

[28] Sánchez-Gil J A, Maradudin A A, Lu Jun Q, Freilikher V D, Pustilnik M and Yurkevich I 1996 Satellite peaks in the scattering of p-polarized light from a randomly rough film on a perfectly conducting substrate *J. Mod. Opt.* **43** 435–52

[29] Danese A E 1965 *Advanced Calculus* vol I (Boston, MA: Allyn and Bacon) p 123
[30] Bloembergen N, Chang R K, Jha S S and Lee C H 1968 Optical second-harmonic generation in reflection from media with inversion symmetry *Phys. Rev.* **174** 813–22
[31] Rudnick J and Stern E A 1971 Second-harmonic radiation from metal surfaces *Phys. Rev.* B **4** 4274–90
[32] Sipe J E, So V C Y, Fukui M and Stegeman G 1980 Analysis of second-harmonic generation at metal surfaces *Phys. Rev.* B **21** 4389–402
[33] Agranovich V M and Darmanyan S A 1982 Theory of second-harmonic generation upon reflection of light from a medium with a center of inversion *JETP Lett.* **35** 80–2
[34] Maytorena J A, Mochan W L and Mendoza B S 1995 Hydrodynamic model for second-harmonic generation at conductor surfaces with continuous profiles *Phys. Rev.* B **51** 2556–62
[35] Mendoza B S and Mochan W L 1996 Exactly solvable model of surface second-harmonic generation *Phys. Rev.* B **53** 4999–5006
[36] Maystre D, Neviére M and Reinisch R 1986 Nonlinear polarization inside metals: a mathematical study of the free-electron model *Appl. Phys.* A **39** 115–21
[37] Coutaz J L, Maystre D, Neviére M and Reinisch R 1987 Optical second harmonic generation from silver at 1.064 μm pump wavelength *J. Appl. Phys.* **62** 1529–31
[38] Neviére M, Vincent P, Maystre D, Reinisch R and Coutaz J L 1988 Differential theory for metallic gratings in nonlinear optics: second harmonic generation *J. Opt. Soc. Am.* B **5** 330–6
[39] Garcia-Molina R, Maradudin A A and Leskova T A 1990 The impedance boundary condition for a curved surface *Phys. Rep.* **194** 351–9
[40] Maradudin A A and Méndez E R 1994 Theoretical studies of the enhanced backscattering of light from one-dimensional randomly rough metal surfaces by the use of a nonlocal impedance boundary condition *Physica* A **207** 302–14
[41] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes* 2nd edn (Cambridge: Cambridge University Press) section 4.2
[42] Delves L M and Mohamed J L 1985 *Computational Methods for Integral Equations* (Cambridge: Cambridge University Press) p 268

ELSEVIER

Physica B 296 (2001) 85–97

PHYSICA B

www.elsevier.com/locate/physb

Localization of surface plasmon polaritons on a random surface

A.A. Maradudin[a,*], I. Simonsen[a,b], T.A. Leskova[c], E.R. Méndez[d]

[a]*Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, CA 92697, USA*
[b]*Department of Physics, The Norwegian University of Science and Technology, N-7491 Trondheim, Norway*
[c]*Institute of Spectroscopy, Russian Academy of Sciences, Troitsk 142092, Russia*
[d]*División de Física Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, Apartado Postal 2732, Ensenada, Baja California 22800, Mexico*

Abstract

We study the possibility of the strong localization of surface plasmon polaritons propagating along a metal surface a finite part of which is randomly rough. The surface roughness is such that the roughness-induced conversion of a surface plasmon polariton propagating on it into volume electromagnetic waves in the vacuum above the surface is suppressed. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Localization; Surface plasmons; Polaritons; Random surfaces

1. Introduction

Since the prediction of localization of electrons in a disordered random system [1], there has been a great interest in localization phenomenon in the physics community. Although this phenomenon was predicted for "quantum" waves, it is not restricted to these kinds of waves, and should in particular also apply to classical waves in random media. For example, the experimental observation of the localization of light was reported recently in a bulk disordered semiconductor [2,3]. In the present work we discuss the Anderson localization of another type of classical wave by disorder of a different nature, namely the localization of surface plasmon polaritons on a randomly rough metal surface in contact with vacuum. This effect has been believed to be difficult to observe due to its being masked by competing effects such as roughness-induced conversion of the surface plasmon polariton into volume waves in the vacuum above the surface (leakage), and ohmic losses due to the non-vanishing imaginary part of the dielectric function of the metal [4–6]. In this work we show how to circumvent this problem by using a specially designed randomly rough surface that suppresses leakage.

The Anderson localization length of a surface plasmon polariton of frequency ω propagating along a one-dimensional randomly rough surface of a metal in contact with vacuum can be determined by calculating the amplitude $t(\omega, L)$ of the surface plasmon polariton transmitted through a finite length L of the random surface. The surface plasmon polariton transmission coefficient is then given by $T(\omega, L) = |t(\omega, L)|^2$. For large L the average of the self-averaging quantity $\ln T(\omega, L)$ over

* Corresponding author.
E-mail address: aamaradu@uci.edu (A.A. Maradudin).

0921-4526/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0921-4526(00)00784-5

the ensemble of realizations of the random surface, $\langle \ln T(\omega, L)\rangle$, is expected to display a linear dependence on L,

$$\langle \ln T(\omega, L)\rangle = \text{const.} - L/\ell_T(\omega), \tag{1.1}$$

where the characteristic length $\ell_T(\omega)$ is called the Lyapunov exponent. It is not the localization length of the surface plasmon polariton $\ell(\omega)$, but can instead be related to the latter by

$$\ell_T^{-1}(\omega) = \ell^{-1}(\omega) + \ell_\varepsilon^{-1}(\omega) + \ell_{\text{rad}}^{-1}(\omega), \tag{1.2}$$

where $\ell_\varepsilon(\omega)$ and $\ell_{\text{rad}}(\omega)$ are, respectively, the characteristic decay lengths associated with the leakage and ohmic losses [7].

In the present paper we analyze analytically and numerically the propagation of surface plasmon polaritons along a metal surface a finite part of which is randomly rough. This random roughness is chosen to constitute a random process that suppresses leakage, i.e. the roughness-induced conversion of a surface plasmon polariton propagating on it into volume electromagnetic waves in the vacuum above the surface. The use of a random surface that suppresses leakage facilitates the investigation of the strong localization of surface plasmon polaritons by random surface roughness by removing the contribution $\ell_{\text{rad}}^{-1}(\omega)$ from the expression for $\ell_T^{-1}(\omega)$ (Eq. (1.2)).

In the approach to the suppression of leakage taken by Sornette and his colleagues [4,5], it was assumed that the random surface was not planar on average, but periodic, so that the dispersion curve of the surface plasmon polaritons supported by the mean surface displays a gap at the boundary of the one-dimensional first Brillouin zone defined by the period of the mean surface. Leakage should then either vanish or decrease significantly for the surface plasmon frequency at the band edge. However, this was not observed in the numerical simulation calculations of leakage carried out in Refs. [6,8].

In this work we first present an approach to designing a one-dimensional random surface that suppresses the leakage of a surface plasmon polariton as it propagates across it that differs from that proposed by Sornette et al. [4,5]. Although the power spectrum of the resulting surface is nonzero in a narrow range of wave numbers, that surface is not periodic on average. However, as with the surface proposed by Sornette et al. our surface is specific to the frequency of the surface plasmon polariton propagating across it: if that frequency is changed, a new surface has to be designed.

For a weakly rough random surface of this nature we analyze the possibility of the localization of surface plasmon polaritons by an analytic approach. In the case of a strongly rough surface we solve the problem of surface polariton propagation numerically.

2. The transmitted field

We study the scattering of a p-polarized surface plasmon polariton of frequency ω propagating in the x_1-direction that is incident on a segment of a one-dimensional randomly rough surface defined by the equation $x_3 = \zeta(x_1)$. The surface profile function $\zeta(x_1)$ is assumed to be a single-valued function of x_1 that is nonzero only in the interval $-L/2 < x_1 < L/2$ (Fig. 1).

The region $x_3 > \zeta(x_1)$ is vacuum and the region $x_3 < \zeta(x_1)$ is a metal characterized by an isotropic, frequency-dependent, complex dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$. We are interested in the frequency range in which $\varepsilon_1(\omega) < -1$, $\varepsilon_2(\omega) > 0$, within which surface plasmon polaritons exist. We write the surface profile function $\zeta(x_1)$ in the form

$$\zeta(x_1) = \Gamma(x_1)s(x_1), \tag{2.1}$$

where $s(x_1)$ is a single-valued function of x_1 that is differentiable and constitutes a stationary,

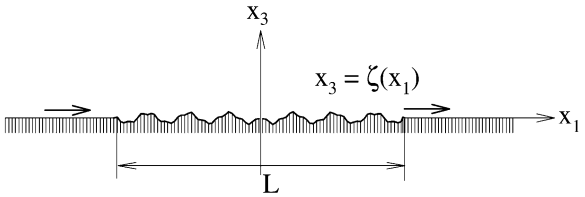

Fig. 1. The scattering system considered in the study of Anderson localization of surface plasmon polaritons on a random surface.

zero-mean, Gaussian random process defined by

$$\langle s(x_1)\rangle = 0, \tag{2.2a}$$

$$\langle s(x_1)s(x'_1)\rangle = \delta^2 W(|x_1 - x'_1|), \tag{2.2b}$$

$$\langle s^2(x_1)\rangle = \delta^2. \tag{2.2c}$$

The angle brackets in Eqs. (2.2) denote an average over the ensemble of realizations of $s(x_1)$, δ is the rms height of the roughness, and $W(|x_1 - x'_1|)$ is the surface height autocorrelation function. The form of the power spectrum of the surface roughness, which is defined by

$$g(|Q|) = \int_{-\infty}^{\infty} dx_1\, e^{-iQx_1} W(|x_1|), \tag{2.3}$$

appropriate for our purposes will be specified below. The function $\Gamma(x_1)$ serves to restrict the nonzero values of $\zeta(x_1)$ to the interval $-L/2 < x_1 < L/2$. One form $\Gamma(x_1)$ can have is

$$\Gamma(x_1) = \theta\left(\frac{L}{2} + x_1\right)\theta\left(\frac{L}{2} - x_1\right), \tag{2.4a}$$

where $\theta(x_1)$ is the Heaviside unit step function. A smoother, differentiable version of $\Gamma(x_1)$ is provided by

$$\Gamma(x_1) = \frac{1 + \cosh(1/2)\kappa L}{\cosh \kappa x_1 + \cosh(1/2)\kappa L}, \tag{2.4b}$$

where the parameter κ controls the range of x_1 values over which $\Gamma(x_1)$ decreases from 1 to 0. In view of the factor $\Gamma(x_1)$ in Eq. (2.1), the surface profile function $\zeta(x_1)$ is not a stationary random process even though $s(x_1)$ is.

We assume that the surface roughness is sufficiently weak that the surface profile function $\zeta(x_1)$ satisfies the conditions for the validity of the Rayleigh hypothesis [9]. In this case the single nonzero component of the magnetic field in the vacuum region $x_3 > \zeta(x_1)_{max}$ can be written as the sum of the fields of the incident and scattered waves

$$H_2^{>}(x_1, x_3|\omega) = \exp[ikx_1 - \beta_0(k,\omega)x_3] + \int_{-\infty}^{\infty} \frac{dq}{2\pi} R^{>}(q,\omega)\exp[iqx_1 - \beta_0(q,\omega)x_3], \tag{2.5a}$$

while in the region of the metal, $x_3 < \zeta(x_1)_{min}$,

$$H_2^{<}(x_1, x_3|\omega) = \exp[ikx_1 + \beta(k,\omega)x_3] + \int_{-\infty}^{\infty} \frac{dq}{2\pi} R^{<}(q,\omega)\exp[iqx_1 + \beta(q,\omega)x_3]. \tag{2.5b}$$

In Eqs. (2.5) k is the wave number of the incident surface plasmon polariton,

$$k = \frac{\omega}{c}\left[\frac{\varepsilon(\omega)}{\varepsilon(\omega)+1}\right]^{1/2} = k_1(\omega) + ik_2(\omega), \tag{2.6}$$

while the functions $R^{>}(q,\omega)$ and $R^{<}(q,\omega)$ are the scattering amplitudes of the surface plasmon polariton in the vacuum and in the metal, respectively, and

$$\beta_0(q,\omega) = \left(q^2 - \frac{\omega^2}{c^2}\right)^{1/2}, \quad \mathrm{Re}\,\beta_0(q,\omega) > 0, \quad \mathrm{Im}\,\beta_0(q,\omega) < 0, \tag{2.7a}$$

$$\beta(q,\omega) = \left(q^2 - \varepsilon(\omega)\frac{\omega^2}{c^2}\right)^{1/2} \mathrm{Re}\,\beta(q,\omega) > 0, \quad \mathrm{Im}\,\beta(q,\omega) < 0. \tag{2.7b}$$

The scattering amplitude $R^{>}(q,\omega)$ satisfies the reduced Rayleigh equation [10]

$$R^{>}(p,\omega) = G_0(p)\left\{v(p|k)J(\beta(p,\omega) - \beta_0(k,\omega)|p-k) + \int_{-\infty}^{\infty} \frac{dq}{2\pi} v(p|q)J(\beta(p,\omega) - \beta_0(q,\omega)|p-q)R^{>}(q,\omega)\right\}, \tag{2.8}$$

where

$$J(\gamma|Q) = \int_{-\infty}^{\infty} dx_1\, e^{-iQx_1}\frac{e^{\gamma\zeta(x_1)} - 1}{\gamma} \tag{2.9}$$

and

$$v(p|q) = \frac{1 - \varepsilon(\omega)}{\varepsilon(\omega)}[pq - \beta(p,\omega)\beta_0(q,\omega)], \tag{2.10}$$

while

$$G_0(p) = \frac{\varepsilon(\omega)}{\varepsilon(\omega)\beta_0(p,\omega) + \beta(p,\omega)} \tag{2.11}$$

is the Green's function of surface plasmon polaritons associated with the planar surface. We note that $G_0(p)$ has simple poles at $p = \pm k$.

If we introduce the transition matrix $T(p|k)$ by the relation

$$R^{>}(p,\omega) = G_0(p)T(p|k), \tag{2.12}$$

then Eq. (2.8) takes the form

$$T(p|k) = V(p|k) + \int_{-\infty}^{\infty} \frac{dq}{2\pi} V(p|q)G_0(q)T(q|k), \tag{2.13}$$

where the scattering potential $V(p|q)$ is given by

$$V(p|q) = v(p|q)J(\beta(p,\omega) - \beta_0(q,\omega)|p-q). \tag{2.14}$$

From Eqs. (2.5a) and (2.12) we see that the scattered field in the vacuum region can be written in the form

$$H_2^{>}(x_1, x_3|\omega)_{sc} = \int_{-\infty}^{\infty} \frac{dq}{2\pi} G_0(q)T(q|k)e^{iqx_1 - \beta_0(q,\omega)x_3}. \tag{2.15}$$

The field of the scattered surface plasmon polariton in the region $x_1 > L/2$ is given by the contribution from the pole in the integrand of the integral in Eq. (2.15), and has the form

$$H_2^{>}(x_1 > L/2, x_3|\omega)_{sc} = iCT(k|k)e^{ikx_1}e^{-\beta_0(k,\omega)x_3}, \tag{2.16}$$

where

$$C = \frac{\sqrt{-\varepsilon^3(\omega)}}{\varepsilon^2(\omega) - 1} \tag{2.17}$$

is the residue of the Green's function $G_0(q)$ at $q = \pm k$. The amplitude of the transmitted surface polariton is therefore

$$t(\omega, L) = 1 + iCT(k|k). \tag{2.18}$$

The surface plasmon polariton transmission coefficient $T(\omega, L)$ is then defined by

$$T(\omega, L) = \frac{P_{tr}(L/2)}{P_{inc}(-L/2)} = |t(\omega, L)|^2 \exp\left(-\frac{L}{\ell_\varepsilon(\omega)}\right), \tag{2.19}$$

where

$$\ell_\varepsilon(\omega) = \frac{1}{2k_2(\omega)} \tag{2.20}$$

is the propagation length of the surface plasmon polaritons due to the ohmic losses in the metal, and gives the fraction of the flux entering the random segment of the metal surface at $x_1 = -L/2$, $P_{inc}(-L/2)$, that leaves it at $x_1 = L/2$, $P_{tr}(L/2)$.

From Eq. (2.19) we find that

$$\langle \ln T(\omega, L)\rangle = \langle \ln|t(\omega, L)|^2\rangle - \frac{L}{\ell_\varepsilon(\omega)}. \tag{2.21}$$

We can rewrite Eq. (2.21) in the form

$$\langle \ln T(\omega, L)\rangle = 2\mathrm{Re}\,\langle \ln[t(\omega, L)]\rangle - \frac{L}{\ell_\varepsilon(\omega)}. \tag{2.22}$$

Therefore, in view of Eqs. (1.1) and (1.2), and in the absence of leakage, we expect that

$$2\mathrm{Re}\,\langle \ln[t(\omega, L)]\rangle = \text{const.} - \frac{L}{\ell(\omega)}. \tag{2.23}$$

3. The random surface

Before proceeding, several remarks on the properties of the rough surface have to be made. From Eq. (2.5a) it follows that the total power scattered into the vacuum above the surface is

$$P_{sc} = L_2 \frac{\omega}{16\pi^2} \int_{-\pi/2}^{\pi/2} d\theta_s \cos^2\theta_s \left| R^{>}\left(\frac{\omega}{c}\sin\theta_s, \omega\right)\right|^2, \tag{3.1}$$

where L_2 the length of the surface along the x_2-axis. The scattering angle θ_s, measured clockwise from the x_3-axis, is related to the wavenumber q by $q = (\omega/c)\sin\theta_s$. Since the integrand in Eq. (3.1) is non-negative, we see that the only way in which leakage can be suppressed, i.e. the only way in which P_{sc} can be made to vanish, is to design a one-dimensional random surface for which the amplitude $R^{>}(q,\omega)$ is identically zero for $-(\omega/c) < q < (\omega/c)$. Several ways to design surfaces that give rise

to specified angular distributions of the scattered intensity have been proposed recently [11,12]. To suppress the leakage we will use a surface characterized by the power spectrum [13]

$$g(|Q|) = \frac{\pi}{2\Delta k}[\theta(Q - k_{\min})\theta(k_{\max} - Q) + \theta(-Q - k_{\min})\theta(k_{\max} + Q)], \tag{3.2}$$

where

$$k_{\min} = 2k_1(\omega) - \Delta k, \tag{3.3a}$$

$$k_{\max} = 2k_1(\omega) + \Delta k, \tag{3.3b}$$

and Δk must satisfy the inequality

$$\Delta k \ll k_1(\omega) - (\omega/c). \tag{3.4b}$$

That a surface characterized by the power spectrum (3.2) suppresses leakage can be seen from the following argument: The incident surface plasmon polariton has a wave number whose real part is $k_1(\omega)$. After its first interaction with the surface roughness it will be scattered into waves the real parts of whose wave numbers lie in the two intervals $(3k_1(\omega) - \Delta k, 3k_1(\omega) + \Delta k)$ and $(-k_1(\omega) - \Delta k, -k_1(\omega) + \Delta k)$. This is because the wave numbers in the spectrum of the surface roughness with which $k_1(\omega)$ can combine lie in the intervals $(2k_1(\omega) - \Delta k, 2k_1(\omega) + \Delta k)$ and $(-2k_1(\omega) - \Delta k, -2k_1(\omega) + \Delta k)$. For the same reason, after the scattered waves interact again with the surface roughness the real parts of the wave numbers of the doubly scattered waves will lie in the three intervals $(5k_1(\omega) - 2\Delta k, 5k_1(\omega) + 2\Delta k)$, $(k_1(\omega) - 2\Delta k, k_1(\omega) + 2\Delta k)$, and $(-3k_1(\omega) - 2\Delta k, -3k_1(\omega) + 2\Delta k)$. After the third interaction with the surface roughness the real parts of the wave number of the scattered waves will lie in the four intervals $(7k_1(\omega) - 3\Delta k, 7k_1(\omega) + 3\Delta k)$, $(3k_1(\omega) - 3\Delta k, 3k_1(\omega) + 3\Delta k)$, $(-k_1(\omega) - 3\Delta k, -k_1(\omega) + 3\Delta k)$, and $(-5k_1(\omega) - 3\Delta k, -5k_1(\omega) + 3\Delta k)$, and so on. Thus, for example, if $-k_1(\omega) + 3\Delta k < -(\omega/c)$, so that $\Delta k < \frac{1}{3}(k_1(\omega) - (\omega/c))$, after triple scattering the surface plasmon polaritons will not be converted into volume electromagnetic waves. In general, if we wish the surface plasmon polariton to scatter n times from the surface roughness without being converted into volume electromagnetic waves, we must require that $\Delta k < (1/n)(k_1(\omega) - (\omega/c))$. It is clear that the strongest scattering processes are those whose final states are backward or forward propagating surface plasmon polaritons, since they result in propagating excitations while all others final states are strongly decaying electromagnetic waves.

A randomly rough surface with a power spectrum of the form of two Gaussian peaks centered at $q = \pm 2k_1(\omega)$ has been assumed in Ref. [14] in a search for the localization-induced enhancement of the surface plasmon polariton field. However, such a power spectrum is nonzero in the range $|q| < (\omega/c)$ and, therefore, such a surface does not suppress leakage.

4. Analytical arguments

Let us consider Eq. (2.13) for the transition matrix $T(q|k)$. In deriving this equation we have assumed that the conditions for the validity of Rayleigh hypothesis are satisfied. The scattering potential $V(q|p)$, given by Eq. (2.14), does not have any poles in the complex plane of the variables q and p, and the inequality $(e^{(\beta(k,\omega)-\beta_0(k,\omega))\zeta(x_1)} - 1) \ll 1$ is satisfied. In view of the power spectrum of the surface roughness assumed, the main contribution to the integral term in Eq. (2.13) comes from the poles of the Green's function $G_0(p)$. In the pole approximation for the Green's function [15] we can rewrite Eq. (2.13) as

$$T(q|k) = V(q|k) + iCV(q|k)T(k|k) + iCV(q|-k)T(-k|k), \tag{4.1a}$$

$$T(k|k) = iCV(k|-k)T(-k|k), \tag{4.1b}$$

$$T(-k|k) = V(-k|k) + iCV(-k|k)T(k|k). \tag{4.1c}$$

From Eqs. (4.1) we obtain

$$T(q|k) = \frac{V(q|k) + iCV(q|-k)V(-k|k)}{1 + C^2V(k|-k)V(-k|k)} \tag{4.2a}$$

and

$$T(k|k) = \frac{iCV(k|-k)V(-k|k)}{1 + C^2V(k|-k)V(-k|k)}. \tag{4.2b}$$

Therefore, from Eq. (2.18) we find that the amplitude of the transmitted field is given by

$$t(\omega,L)=\frac{1}{1+C^2V(k|-k)V(-k|k)}. \tag{4.3}$$

Using Eq. (4.3) we can calculate the desired quantity $\langle \ln T(\omega,L)\rangle$,

$$\begin{aligned}\langle \ln T(\omega,L)\rangle &= 2\mathrm{Re}\langle \ln t(\omega,L)\rangle \\ &= -2\mathrm{Re}\langle \ln[1+C^2V(k|-k)V(-k|k)]\rangle. \end{aligned}\tag{4.4}$$

In order to calculate the average in the second line of Eq. (4.4) we shall use the Taylor expansion of the logarithm

$$\begin{aligned}&\langle \ln[1+C^2V(k|-k)V(-k|k)]\rangle \\ &\quad=\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}C^{2n}\langle [V(k|-k)V(-k|k)]^n\rangle, \end{aligned}\tag{4.5}$$

and calculate the moments $\langle [V(k|-k)V(-k|k)]^n\rangle$. From Eq. (2.14) it follows that

$$\begin{aligned}\langle [V(k|-k)V(-k|k)]^n\rangle &\equiv [v(k|-k)v(-k|k)]^n \\ &\times\langle J^n(\beta(k,\omega)-\beta_0(k,\omega)|2k)J^n(\beta(k,\omega) \\ &-\beta_0(k,\omega)|-2k)\rangle. \end{aligned}\tag{4.6}$$

To calculate these moments we use their representation in terms of cumulant averages [16,17]. In doing so we will use the fact that due to the nonstationarity of the surface the cumulant average of $J(\gamma|Q)$ is given by

$$\langle J(\gamma|Q)\rangle_c = L\,\mathrm{sinc}\left[Q\frac{L}{2}\right]\frac{1}{\gamma}(e^{\gamma^2\delta^2/2}-1), \tag{4.7}$$

the cumulant average of the product $J(\gamma|Q_1)J(\gamma|Q_2)$ is given by

$$\begin{aligned}\langle J(\gamma|Q_1)J(\gamma|Q_2)\rangle_c &= j(Q_1,Q_2) \\ &\equiv L^2\int_{-\infty}^{\infty}\frac{dr}{2\pi}\mathrm{sinc}\left[\frac{(Q_1-r)L}{2}\right] \\ &\times\mathrm{sinc}\left[\frac{(Q_2+r)L}{2}\right]\hat{g}(r), \end{aligned}\tag{4.8}$$

where

$$\hat{g}(r)=\frac{1}{\gamma^2}e^{\gamma^2\delta^2}\int_{-\infty}^{\infty}du\,e^{-iru}[e^{\delta^2\gamma^2W(|u|)}-1] \tag{4.9}$$

and so on. As a result, the cumulants of any odd numbers of J functions from Eq. (4.6) are found to be proportional to at least one factor $\mathrm{sinc}(k_1L)$, which is small when $k_1L\gg 1$. Only the cumulant averages of products of even numbers of J functions, which contain equal numbers of $J(\gamma|2k)$ and $J(\gamma|-2k)$, do not contain this small factor.

When the length of the rough part of the surface is not very large, so that $L\delta^2\ll|\varepsilon(\omega)|\lambda^3$, where $\lambda=(2\pi c/\omega)$, the main contribution to the average of the product of J functions comes from the products of pair cumulant averages, so that the moments are found to be given by

$$\begin{aligned}&\langle J^n(\beta(k,\omega)-\beta_0(k,\omega)|2k)J^n(\beta(k,\omega)-\beta_0(k,\omega)|-2k)\rangle \\ &\quad\approx n![\langle J(\beta(k,\omega)-\beta_0(k,\omega)|2k) \\ &\qquad\times J(\beta(k,\omega)-\beta_0(k,\omega)|-2k)\rangle_c]^n \\ &\quad= n!j^n(2k,-2k). \end{aligned}\tag{4.10}$$

With the use of Eq. (4.10) we can rewrite the infinite series in Eq. (4.5) as

$$\begin{aligned}&\langle \ln[1+C^2V(k|-k)V(-k|k)]\rangle \\ &\quad= C^2v(k|-k)v(-k|k)\sum_{n=0}^{\infty}(-1)^n n!C^{2n} \\ &\qquad\times[v(k|-k)v(-k|k)j(2k,-2k)]^n \\ &\quad= C^2v(k|-k)v(-k|k)j(2k,-2k) \\ &\qquad\times\int_0^{\infty}du\frac{e^{-u}}{1+C^2v(k|-k)v(-k|k)j(2k,-2k)u}. \end{aligned}\tag{4.11}$$

In the limit of a weakly rough surface, $\delta\ll\lambda\sqrt{|\varepsilon(\omega)|}$, the function $\hat{g}(r)$ can be approximated by $\hat{g}(2k)=\delta^2g(|r|)$, so that $j(2k,-2k)$ takes the form

$$\begin{aligned}j(2k,-2k) &= L^2\delta^2\int_{-\infty}^{\infty}\frac{dr}{2\pi}g(|r|) \\ &\times\mathrm{sinc}^2\left[(2k-r)\frac{L}{2}\right]. \end{aligned}\tag{4.12}$$

When the length of the rough part of the surface is small, so that the conditions $L\Delta k\ll 1$ and $k_2L\ll 1$ are satisfied, the function $j(2k,-2k)$ becomes $j(2k,-2k)=L^2\delta^2/2$. Since in this case

$$C^2v(k|-k)v(-k|k)j(2k,-2k)\ll 1, \tag{4.13}$$

the integral in Eq. (4.11) can be replaced by unity, and therefore, we obtain

$$\langle \ln T(\omega, L)\rangle \approx -\mathrm{Re}[C^2L^2\delta^2 v(k|-k)v(-k|k)]. \tag{4.14}$$

To illustrate these results we assume that surface polaritons of frequency ω corresponding to a wavelength $\lambda = (2\pi c/\omega) = 457.9$ nm propagate along the silver surface, whose dielectric function at this wavelength is $\varepsilon(\omega) = -7.5 + \mathrm{i}0.24$. Let the surface roughness be characterized by the parameters $\Delta k = 0.3(k_1(\omega) - (\omega/c)) \propto 0.15\omega/(c|\varepsilon(\omega)|)$ and the rms height $\delta = 3$ nm. For these parameters the two main characteristic lengths of the problem are $\ell_{\mathrm{corr}} \equiv 1/(\Delta k) = 4\lambda$ and $\ell_{\varepsilon} \equiv 1/2k_2(\omega) = 30.1\lambda$. Therefore, to satisfy the condition $L\Delta k \ll 1$ the length of the surface should be of the order of a few wavelengths. Then, since $L < \ell_{\mathrm{eff}} = 5.3\lambda$, where $\ell_{\mathrm{eff}}(\omega) = \{\mathrm{Re}[C^2 v(k|-k)v(-k|k)]\}^{-1/2}$, the condition at which Eq. (4.14) have been obtained is satisfied, and the average logarithm of the transmission coefficient of surface plasmon polaritons has the form $\langle \ln T(L)\rangle = \text{const.} - (L/\ell_{\mathrm{eff}}(\omega))^2$.

However, when studying the localization of classical waves we are interested in the limit $L \to \infty$. Therefore, in this limit $L\Delta k \gg 1$, and the function $j(2k|-2k)$ has the form

$$j(2k|-2k) = L\frac{2\pi}{\Delta k}\delta^2. \tag{4.15}$$

If, in addition, the condition (4.13) is satisfied, we obtain the expression for the averaged logarithm of the transmission coefficient in the form

$$\langle \ln T(\omega, L)\rangle \approx -\mathrm{Re}\left[C^2L\frac{2\pi}{\Delta k}\delta^2 v(k|-k)v(-k|k)\right], \tag{4.16}$$

i.e.

$$\langle \ln T(\omega, L)\rangle \approx -\frac{L}{\ell(\omega)}, \tag{4.17}$$

where

$$\ell^{-1}(\omega) = \mathrm{Re}\left[C^2\frac{2\pi}{\Delta k}\delta^2 v(k|-k)v(-k|k)\right] \tag{4.18}$$

and coincides with the scattering length

$$\ell_{\mathrm{sc}}^{-1}(\omega) = 2\Delta_{\mathrm{sc}}(\omega), \tag{4.19}$$

where $\Delta_{\mathrm{sc}}(\omega)$ is the roughness-induced decay rate of surface plasmon polaritons. Indeed, let us introduce the exact Green's function $G(q|p)$ in accordance with Ref. [18] by the relation

$$G_0(p)T(p|k) = \int_{-\infty}^{\infty}\frac{\mathrm{d}q}{2\pi}G(p|q)V(q|k). \tag{4.20}$$

In the limit $L\Delta k \gg 1$ the stationarity of the surface is almost restored and the average Green's function $G(q)$, which is then defined by $\langle G(q|p)\rangle = G(q)2\pi\delta(q-p)$, has the form

$$G(q) = \frac{1}{G_0^{-1}(q) - M(q)}, \tag{4.21}$$

where $M(q)$ is the averaged self-energy defined by $\langle M(q|p)\rangle = M(q)2\pi\delta(q-p)$. The self-energy $M(q|p)$ satisfies the equation [18]

$$M(q|k) = V(q|k) + \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi}V(q|p)G_0(p)[M(p|k) - \langle M(p|k)\rangle]. \tag{4.22}$$

In the pole approximation the averaged self-energy can be obtained in the same manner in which we calculated the transition matrix $T(q|k)$. The result is

$$\langle M(k|k)\rangle = \frac{\langle \mathrm{i}CV(k|-k)V(-k|k)/[1 + C^2V(k|-k)V(-k|k)]\rangle}{\langle [1 + \mathrm{i}CV(k|-k)]/[1 + C^2V(k|-k)V(-k|k)]\rangle}. \tag{4.23}$$

In the limit in which Eq. (4.16) was obtained the averaged self-energy is given by

$$\langle M(k|k)\rangle \approx \mathrm{i}Cv(k|-k)v(-k|k)j(2k|-2k). \tag{4.24}$$

Therefore, $\Delta_{\mathrm{sc}}(\omega) = \mathrm{Im}\, CM(k)$, and is indeed the roughness–induced decay rate of surface plasmon polaritons.

For example, for the case where the rms height of the surface roughness is $\delta = 3$ nm while the length of the rough part of the surface is small $L < 20\lambda$, the length $\ell_{\mathrm{sc}}(\omega)$ turns out to be $\ell \approx 8\lambda$, and is smaller

than the propagation length of surface plasmon polaritons associated with the ohmic losses. However, the exponential decay of the transmission coefficient of surface plasmon polaritons does not necessarily imply that they are localized. For such a weakly rough surface, although the scattering by surface roughness leads to a strong damping of the surface waves, nevertheless they remain propagating electromagnetic waves, since $k_2(\omega) + \Delta_{sp}(\omega) \ll k_1(\omega)$. And if the rough surface is not very long, the surface polaritons can escape the rough part of the surface.

The situation is different when the surface is moderately rough, or the length of the rough part of the surface increases, so that $\delta^2 L \gg |\varepsilon(\omega)|\lambda^3$. In this limit the main contribution to the average of the product of nJ functions comes not from the products of pair cumulants but from the cumulant average of largest order that is

$$\langle J^n(\beta(k,\omega) - \beta_0(k,\omega)|2k) J^n(\beta(k,\omega) - \beta_0(k,\omega)| - 2k)\rangle$$

$$\approx \langle J^n(\beta(k,\omega) - \beta_0(k,\omega)|2k) J^n(\beta(k,\omega) - \beta_0(k,\omega)| - 2k)\rangle_c = Ln\hat{g}^{2n-1}(2k), \tag{4.25}$$

where $\hat{g}(q)$ is given by Eq. (4.9). In this case we can sum the infinite series in Eq. (4.5) with the result

$$\langle \ln[1 + C^2 V(|k| - k) V(-k|k)]\rangle = L \sum_{n=0}^{\infty} (-1)^n C^{2n+2} \times L[v(|k| - k) v(-k|k)]^{n+1} \hat{g}^{2n+1}(2k)$$

$$= 2L \frac{C^2 v(|k| - k) v(-k|k)\hat{g}(2k)}{1 + C^2 v(|k| - k) v(-k|k)\hat{g}^2(2k)}. \tag{4.26}$$

The average of the logarithm of the transmission coefficient in this case takes the form

$$\langle \ln T(\omega, L)\rangle = -2L \,\mathrm{Re} \times \left\{ \frac{C^2 v(|k| - k) v(-k|k)\hat{g}(2k)}{1 + C^2 v(|k| - k) v(-k|k)\hat{g}^2(2k)} \right\}. \tag{4.27}$$

Therefore, we obtained the linear dependence

$$\langle \ln T(\omega, L)\rangle \approx -\frac{L}{\ell(\omega)}, \tag{4.28}$$

where

$$\ell^{-1}(\omega) = 2\mathrm{Re}\left\{ \frac{C^2 v(|k| - k) v(-k|k)\hat{g}(2k)}{1 + C^2 v(|k| - k) v(-k|k)\hat{g}^2(2k)} \right\}. \tag{4.29}$$

For a rough surface with an rms height $\delta = 10$ nm, we find that $C^2 v(|k| - k) v(-k|k)\hat{g}^2(2k) \gg 1$. Therefore, instead of Eq. (4.27) we can write

$$\langle \ln T(\omega, L)\rangle \approx -2L \,\mathrm{Re}\left[\frac{1}{\hat{g}(2k)} \right], \tag{4.30}$$

so that

$$\ell^{-1}(\omega) = 2\mathrm{Re}\left[\frac{1}{\hat{g}(2k)} \right]. \tag{4.31}$$

Thus, in this case the localization length is $\ell(\omega) \approx 0.1\lambda$. In this case, the surface polariton field is overdamped and the waves cease to propagate. As in the case of a weakly rough surface the length $\ell(\omega)$ coincides with the roughness induced decay length of surface plasmon polaritons, $\ell(\omega) = 1/(2\Delta_{sc}(\omega))$, where $\Delta_{sc}(\omega) = \mathrm{Im}\, CM(k)$, since in this limit the average self energy, Eq. (4.23) is

$$M(k) = -\frac{i}{C\hat{g}(2k)}. \tag{4.32}$$

Several remarks have to be made concerning the pole approximation we have used here. When using the pole approximation and, thus, reducing the integral equations (2.13) to a system of algebraic equation, we lose the contributions from the nonsingular part of the integrand. This might be significant if the transition matrix $T(q|k)$ has strong peaks. However, the heights of the peaks can be expected to be of the order of $1/\Delta_{sc}(\omega)$ and $1/\Delta k$. Since $\Delta_{sc}(\omega) \gg k_2(\omega)$ and $\Delta k \gg k_2(\omega)$, the contributions of these possible peaks to the integral part of Eq. (2.13) are much smaller than the contribution from the poles of the Green's function $G_0(q)$ and can be neglected.

In the analytical approach described above we assumed that the surface is weakly rough, so that $\delta\sqrt{\varepsilon(\omega)} \ll \lambda$. In the case when the surface roughness is quite strong, so that $\delta\sqrt{\varepsilon(\omega)} \geqslant \lambda$, but the conditions for the validity of the Rayleigh hypothesis are, nevertheless, satisfied, to study the localization of surface plasmon polaritons we solved Eq. (2.13) numerically.

5. Numerical solution

In order to solve the equation for the transition matrix care has to be taken. We recall that our ultimate goal is calculate the transmission amplitude $t(\omega, L)$ defined in Eq. (2.18), i.e. essentially to obtain $T(\omega, L)$ from a numerical solution of the inhomogeneous Fredholm integral equation of the second kind satisfied by the transition matrix. In doing so one is facing at least two major challenges: (i) How to calculate the transition matrix $T(p|k)$ at the wave number of the surface plasmon polariton $p = k$, which is complex due to the non-vanishing imaginary part of the dielectric function of the metal, and (ii) how to handle the poles of $G_0(q)$ at $q = \pm k$. The numerical technique used to calculate $T(k|k)$ numerically is a two step process. It is started by observing that for a real (absorbing) metal there will be no poles located directly on the real axis. Therefore, at least in principle, one can calculate $T(p|k)$ for all real arguments p. Since the kernel is well defined for all real q's, this is done by converting the integral equation into a set of linear equation that can be solved by standard techniques [19]. For step 2 we notice that the integral equation (2.13) is valid for all momenta p and q, both real and complex. Thus, one can calculate the desired transition matrix at $p = k$ by integrating along the real q-axis because here $T(q|k)$ is already known from the preceding step, i.e. one calculates

$$T(k|k) = \int_{-\infty}^{\infty} \frac{\mathrm{d}q}{2\pi}\, V(k|q) G_0(q) T(q|k). \tag{5.1}$$

Note that the scattering potential vanishes at $q = k$, $V(q|k) = 0$, since $\beta(k, \omega)\beta_0(k, \omega) \equiv k^2$.

Although the poles are not located on the real axis, they are still rather close to it. Therefore, in the vicinity of $q = \pm k_1$, where the poles are closest to the real axis, the integrand in Eq. (5.1) changes rapidly. Furthermore, $T(q|k)$ might have weaker peaks, due to multiply scattered surface plasmon polaritons, located at $q = \pm (2n + 1)k_1(\omega)$ $n = 1, 2, 3 \ldots$. It is therefore beneficial not to use a uniform discretization grid, so that a higher density of points can be used around these wave numbers. This was done by first replacing the upper and lower limits in the integral in Eq. (5.1) by finite values and then subdividing this resulting region of integration into subintervals where different densities of points were allowed for. The integration range was divided into a total of 27 subintervals and, in particular, small intervals with high densities of points were chosen around $\pm k_1(\omega)$, $\pm 3k_1(\omega)$, and $\pm 5k_1(\omega)$. Within each subinterval the grid points corresponding to different densities were obtained by the classic Gauss–Legendre method [19]. The total number of points used in the integration was $N = 1850$.

In this way we solved the integral equation (5.1) satisfied by the transmission matrix $T(k|k)$ needed to calculate e.g. the transmission coefficient of surface plasmon polaritons as a function of the length L of the rough part of the surface, for each particular realization of the surface. In numerical calculations the function $\Gamma(x_1)$ which serves to restrict the nonzero values of $s(x_1)$ to the interval $-L/2 < x_1 < L/2$, was taken in the form given by Eq. (2.4b) with $\kappa = 100L$ so that $s(x_1)$ was cut off smoothly.

The traditional way of generating randomly rough surfaces with a well-defined power spectrum and Gaussian height distribution is to use the so-called Fourier filtering method [20, Appendix A]. This method consists of generating Gaussian uncorrelated random variables that are filtered with the desired (decaying) power spectrum. By Fourier transforming this filtered sequence back into real space one obtains a randomly rough surface with the desired statistical properties. In most implementations of this algorithm, it is beneficial to take advantage of the fast Fourier transform for performing the inverse transform needed. However, to generate numerically surfaces that suppress leakage as defined in the preceding sections, the use of the fast Fourier transform (FFT) is not necessarily the

best option. The reason for this is that the power spectrum, according to its definition, is nonvanishing only in a very narrow interval of width $2\Delta k$ about $\pm 2k_1(\omega)$, where $k_1(\omega)$ is the real part of the wave number of the surface plasmon polaritons supported by the planar surface. This has the consequence that the number of points needed in order to resolve the nonvanishing part of the power spectrum in a satisfactory matter is very large. Since it is the widths of the rectangles contained in the power spectrum that makes the surface randomly rough, we want a rather good resolution here. For example for the numerical results for silver to be shown later, $\Delta k \sim 10^{-2}\,\omega/c$ and the number of points needed in order to generate surfaces in a satisfactory manner by using the FFT was $N \geqslant 10^4$. The FFT for this number of points is a computationally costly algorithm, and we therefore calculated the Fourier transform by straightforward numerical integration for which a high-density discretization in momentum space is possible at lower computational costs. Another advantage of this numerical integration approach is, as we will see below, that the surface now may be generated directly on a non-uniform grid without any need for any interpolation. In Fig. 2 an example of a surface realization generated by the method just described is presented.

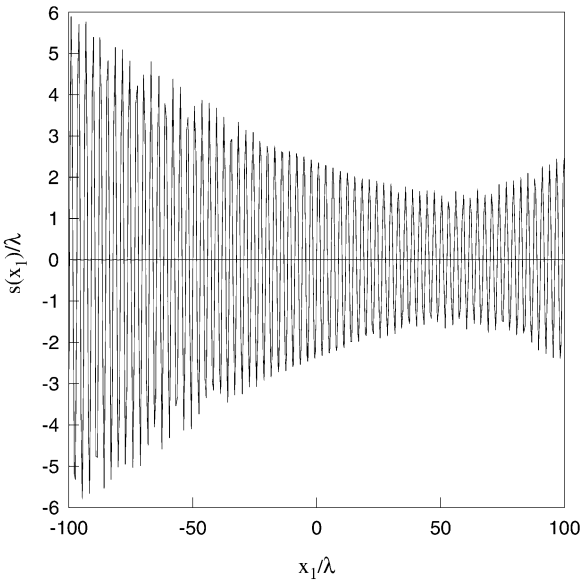

Fig. 2. Example of a random surface.

6. Numerical results

We start our discussion of the numerical results by presenting a result that explicitly shows that the surfaces generated in the way described above suppress leakage, i.e. that the scattering amplitude vanishes in the radiative region $|q| < \omega/c$. In Fig. 3 we present a plot of $(\omega/c)^2\langle|R^>(q,\omega)|^2\rangle$ as a function of cq/ω for a silver surface where the rough portion had a length $L = 20\lambda$. The result plotted in Fig. 3 was calculated from the analytical expression for the transition matrix $T(q|k)$, Eq. (4.2a), obtained in the pole approximation for the Green's function. No expansion in powers of the surface profile function $s(x_1)$ was used when averaging $|R^>(q,\omega)|^2$. The vacuum wavelength of the surface plasmon

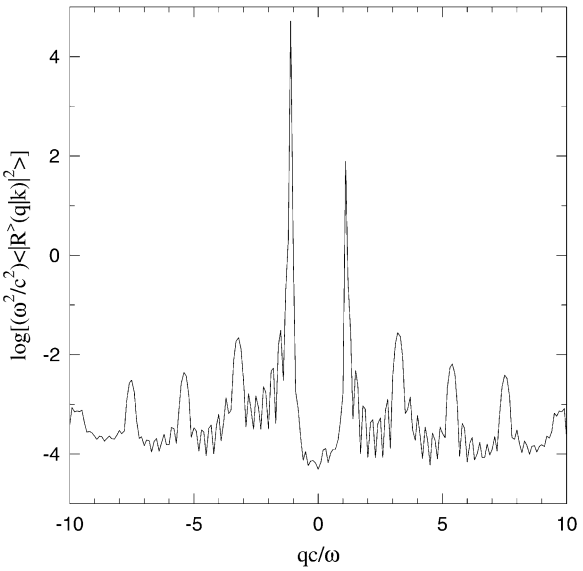

Fig. 3 . Plot of $(\omega/c)^2\langle|R^>(q,\omega)|^2\rangle$ as a function of cq/ω calculated by averaging analytically the expression (5.1) for the transition matrix $T(q|k)$ for a silver surface characterized by the parameters $\Delta k = 0.3(k_1(\omega) - (\omega/c))$ and $\delta = 3$ nm. The rough portion of the surface has length $L = 20\lambda$. The wave number of the surface plasmon polariton, $k(\omega) = k_1(\omega) + ik_2(\omega) = (1.0741 + i0.0026)\omega/c$, corresponds to a vacuum wavelength of $\lambda = 457.9$ nm, and the dielectric function of silver at this frequency is $\varepsilon(\omega) = -7.5 + i0.24$.

polaritons was taken to be $\lambda = (2\pi c/\omega) = 457.9$ nm, so that the dielectric function of silver at this frequency is $\varepsilon(\omega) = -7.5 + i0.24$. The corresponding wave number of the surface plasmon polariton is $k(\omega) = k_1(\omega) + ik_2(\omega) = (1.0741 + i0.0026)\omega/c$. The surface roughness was characterized by the parameters $\Delta k = 0.3(k_1(\omega) - (\omega/c))$ and $\delta = 3$ nm. With this value of Δk the surface should suppress leakage due to scattering processes of up to, and including, third order. We observe from Fig. 3 that $\langle|R^{>}(q,\omega)|^2\rangle$ is indeed suppressed in the radiative region. The analogous results obtained by means of a numerical solution of Eq. (2.13) is presented in Fig. 4 for the case where the rms height of the surface roughness was taken to be $\delta = 30$ nm. From this figure we see that although $\langle|R^{>}(q,\omega)|^2\rangle$ is heavily suppressed in the region of small values of $q \ll (\omega/c)$, it is far from zero for almost grazing directions of radiation $q \leqslant (\omega/c)$. This is due to the strong higher-order scattering processes which are possible for such a strongly rough surface. Only six peaks corresponding to $\pm k_1(\omega)$, $\pm 3k_1(\omega)$, and $\pm 5k_1(\omega)$ are easily seen in this figure. It should also be noted that when $|q| \geqslant 7\omega/c$, $(\omega/c)^2\langle|R^{>}(q|k)|^2\rangle$ becomes flatter. This flatness is due to leakage setting in for such scattering wave numbers, and they are no longer restricted to well-defined values as is the case for $|q| \leqslant 7\omega/c$. The reason for the rapid dip at $q \approx k_1(\omega)$ is caused by the vanishing of the scattering potential $V(k|q)$ at $q = k$.

By numerically solving Eq. (2.13), and calculating $T(k|k)$ by numerical integration in Eq. (5.1), the transmission amplitude $t(\omega, L)$ defined in Eq. (2.18), and the transmission coefficient $T(\omega, L)$ defined by Eq. (2.19) could be calculated for different values of the length L of the rough portion of the surface. From this equation we also recall that the decay due to ohmic losses in the metal could be factored out leaving only possible leakage or Anderson localization in the quantity $|t(\omega, L)|^2$. In Fig. 5 we show numerical simulation results (filled circles) for $\langle \ln T(\omega, L)\rangle$ as a function of the length of the rough portion of the surface. The remaining parameters are those used to obtain the results of Fig. 4. The error bars indicate errors due to the use of a finite number of samples. These errors tend to increase with increasing system size L because of numerical difficulties related to the peaks that can be seen in Fig. 4 becoming narrower and higher.

We observe from Fig. 5 that the behavior of $\langle \ln T(\omega, L)\rangle$ within the error bars is consistent with the behavior predicted in Eq. (2.23), i.e. with an exponential decay of the surface plasmon polariton

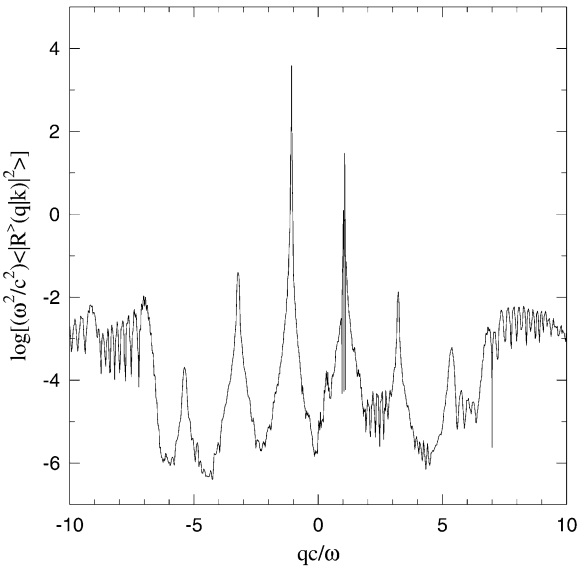

Fig. 4. The same as in Fig. 3 but calculated by means of a numerical solution of Eq. (2.13). The rms height of the surface roughness is $\delta = 30$ nm. The results for 50 realizations of the surface profile function were averaged numerically to obtain the results plotted in this figure.

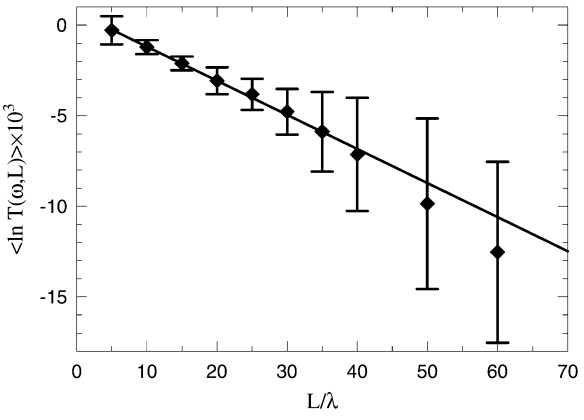

Fig. 5. Numerical simulation results for $\langle \ln T(\omega, L)\rangle$ versus the length of the rough portion of the surface L.

transmission coefficient. The solid line in Fig. 5 represents a χ^2-fit [19] to the simulation data. We recall that in the absence of leakage the slope of this straight line gives according to Eq. (2.23) the inverse of the Anderson localization length, $\ell(\omega)$. The numerical value that we obtain in this way is

$$\ell(\omega) = (5319 \pm 905)\lambda. \tag{6.2}$$

Thus, by large-scale numerical simulations we have shown that for such a strongly rough random surface the average logarithm of the transmission coefficient is a linear function of the length of the rough part of the surface. However, in this case the characteristic length of the decay of the transmission coefficient $\ell(\omega)$ is of the order of many thousands of wavelengths. As we have seen in Fig. 4 in this case the bulk electromagnetic waves are quite efficiently radiated into the vacuum in the directions almost parallel to the surface. These scattered bulk waves are, in their turn, scattered by the surface roughness and excite surface plasmon polaritons. Just these processes of reexcitation of surface plasmon polaritons lead to such a long decay length $\ell(\omega)$. Thus, in this case, although the scattering length $\ell_{sc}(\omega)$ might be quite small, any possibility of localization is destroyed by the strong reexcitation of surface plasmon polaritons.

7. Conclusions

In this paper we have presented an approach to generating a one-dimensional random surface that suppresses leakage. The suppression of leakage is essential for being able to observe the localization of surface plasmon polaritons on a randomly rough surface. We have shown that in the case of a weakly and moderately rough surface the transmission coefficient $T(\omega, L)$ of surface plasmon polaritons decays exponentially with the length L of the rough part of the surface. The inverse of the characteristic length of the decay $\ell_{sc}^{-1}(\omega)$ is determined by the roughness-induced decay rate of the surface plasmon polaritons $\ell_{sc}^{-1}(\omega) = 2\Delta_{sc}(\omega) = 2\,\mathrm{Im}\,CM(k)$. In the case of a weakly rough surface and when the length L of the rough part of the surface is small, although the localization length is smaller than the propagation length $\ell_{\varepsilon}(\omega)$, which is due to the ohmic losses in the metal, it is large enough to ensure the propagative nature of the surface plasmon polariton field. With the increase of the strength of the surface roughness, the localization length becomes considerably smaller than the vacuum wavelength of the surface plasmon polariton. In this case the surface plasmon polaritons lose their wave-like nature and their field is localized. With a further increase of the rms height, the scattering processes of higher order (higher than the third) become efficient. These processes lead to the appearance of leakage and, more important, to the processes of reexcitation of surface plasmon polaritons. By large-scale numerical simulations we showed that the localization length in this case is of the order of many thousands of wavelengths. This is because in this case the surface plasmon polaritons which propagate along the planar surface away from the rough part of the surface ($x_1 > L/2$) are actually not the transmitted surface plasmon polaritons but can be regarded as the surface plasmon polaritons excited by the effective modes of the rough surface: coupled multiply scattered surface plasmon polaritons and bulk electromagnetic waves.

Acknowledgements

It is a pleasure to dedicate the paper to Eleftherios Economou on the occasion of his 60th birthday, in recognition of his many contributions to the theory of strong localization. The research of A.A.M. and T.A.L. was supported in part by Army Research Office Grant DAAD 19-99-1-0321. I.S. would like to thank the Research Council of Norway (Contract No. 32690/213) and Norsk Hydro ASA for financial support. The work of E.R.M. was supported by CONACYT Grant 3804P-A. This work has also received support from the Research Council of Norway (Program for Supercomputing) through a grant of computer time.

References

[1] P.W. Anderson, Phys. Rev. 109 (1958) 1492.
[2] D.S. Wiersma, P. Bartolini, A. Lagendijk, R. Righini, Nature 390 (1977) 671.

[3] F. Scheffold, R. Lenke, R. Tweer, G. Maret, Nature 398 (1997) 206.
[4] D. Sornette, L. Macon, J. Coste, J. Phys. (France) 49 (1988) 1683.
[5] J.-P. Desideri, L. Macon, D. Sornette, Phys. Rev. Lett. 63 (1989) 390.
[6] J.A. Sánchez-Gil, A.A. Maradudin, Phys. Rev. B 56 (1997) 1103.
[7] V.D. Freilikher, S.A. Gredeskul, in: E. Wolf (Ed.), Progress in Optics, Vol. XXX, North-Holland, Amsterdam, 1992, p. 137.
[8] F. Pincemin, J.-J. Greffet, J. Opt. Soc. Am. B 13 (1996) 1499.
[9] Lord Rayleigh, The Theory of Sound, Vol. II, 2nd Edition, MacMillan, London, 1896, pp. 89–96, 297–311.
[10] M.F. Pascual, W. Zierau, T.A. Leskova, A.A. Maradudin, Opt. Commun. 155 (1998) 351.
[11] T.A. Leskova, A.A. Maradudin, I.V. Novikov, A.V. Shchegrov, E.R. Méndez, Appl. Phys. Lett. 73 (1998) 1943.
[12] A.A. Maradudin, I. Simonsen, T.A. Leskova, E. R. Méndez, Opt. Lett. 24 (1999) 1257.
[13] C.S. West, K.A. O'Donnell, J. Opt. Soc. Am. A 12 (1995) 390.
[14] H. Ogura, Z.L. Wang, Y. Sasakura, V. Freilikher, Opt. Commun. 134 (1997) 1.
[15] A.A. Maradudin, A.R. McGurn, V. Celli, Phys. Rev. B 31 (1985) 4866.
[16] R. Kubo, J. Phys. Soc. Japan 17 (1962) 1100.
[17] T.A. Leskova, A.A. Maradudin, I.V. Novikov, J. Opt. Soc. Am. A 17 (2000) 1288.
[18] G. Brown, V. Celli, M. Haller, A.A. Maradudin, A. Marvin, Phys. Rev. B 31 (1985) 4993.
[19] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes, 2nd Edition, Cambridge University Press, Cambridge, 1992.
[20] A.A. Maradudin, T. Michel, A.R. McGurn, E.R. Méndez, Ann. Phys. 203 (1990) 255.

Channel polaritons

I. V. Novikov* and A. A. Maradudin
Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, California 92697
(Received 5 October 2001; revised manuscript received 13 February 2002; published 28 June 2002)

By means of Green's second integral identity in the plane and a parametrization of the surface profile function, the dispersion relation for surface electromagnetic waves guided by a straight channel cut into the otherwise planar surface of a solid in contact with vacuum has been determined numerically. The solid can be either a metal or a dielectric medium and is characterized by an isotropic, real, frequency-dependent dielectric function $\epsilon(\omega)$ that is negative in some range of frequencies. The resulting propagating bound electromagnetic modes are called *channel polaritons.*

DOI: 10.1103/PhysRevB.66.035403 PACS number(s): 78.20.Bh, 42.82.Et, 42.79.Gn

I. INTRODUCTION

The ability to confine surface electromagnetic waves in the plane perpendicular to their direction of propagation could make these waves useful in device applications such as signal transmission and routing and in the construction of couplers and power splitters.[1] This possibility has stimulated recent experimental and theoretical investigations of structures that give rise to such confinement. Thus, Bozhevolni *et al*.[2] have studied experimentally the guiding of surface plasmon polaritons by line defects in 400-nm-period triangular lattice structures formed from ~200-nm-wide and ~45-nm-high gold scatterers on a gold film deposited on a glass substrate. Subsequently, Berini[1] has studied theoretically the bound electromagnetic modes of propagation supported by a structure consisting of a thin lossy metal film of finite width on a dielectric substrate and covered by a different dielectric superstrate. More recently Weeber *et al*. have studied experimentally[3] and theoretically[4] the guiding of surface plasmon polaritons along thin metal films of finite width (metal stripes) deposited on a glass substrate.

In this paper we obtain the dispersion relation for the bound electromagnetic modes of propagation supported by a much simpler structure: namely, a straight channel cut into an otherwise planar surface of a solid in contact with vacuum. The solid, which can be either a metal or a dielectric medium, is characterized by a real, isotropic, frequency-dependent dielectric function $\epsilon(\omega)$ that is negative in a certain frequency range. The restriction to a real dielectric function is not essential to the calculations reported here, but is justified by the fact that the mean free path of surface plasmon polaritons on a planar silver surface, due to Ohmic losses, exceeds 100 μm for wavelengths of light in the near infrared and is approximately 21 μm at the wavelength of a He-Ne laser. These are long enough that such electromagnetic modes can be useful in integrated optical circuits operating at telecommunications wavelengths.[2] We have called these electromagnetic modes *channel polaritons*. The present work is thus an extension of an earlier study[5] in which the same problem was studied in the electrostatic limit.

The outline of this paper is the following. In Sec. II the problem of electromagnetic waves guided by a channel is formulated as a waveguide problem, and the equations satisfied by the components of the electric and magnetic fields in the system parallel to the axis of the channel are obtained, together with the boundary conditions satisfied by these field components at the vacuum-solid interface. In Sec. III we describe a procedure for the parametrization of the surface profile function. This procedure was first introduced in Ref. 6, in which light scattering by a reentrant one-dimensional surface was studied. These results, together with Green's second integral identity in the plane,[7] are used in Sec. III to obtain a set of four coupled, homogeneous integral equations for the values of the components of the electric and magnetic fields parallel to the channel and of their normal derivatives evaluated at the vacuum-solid interface. The solvability condition for this system of equations is the dispersion relation for channel polaritons. The numerical solution of the dispersion relation and the calculation of the fields for a single mode are outlined in Sec. IV, and results obtained by this approach are presented in Sec. V for three different forms of the cross section of the channel. In Sec. VI a discussion of the results obtained concludes the paper.

II. MAXWELL'S EQUATIONS AND THE BOUNDARY CONDITIONS

The physical system studied in this paper consists of a vacuum in region I $[x_3 > \zeta(x_1)]$ and a solid characterized by a real, isotropic, frequency-dependent, dielectric function $\epsilon(\omega)$ in region II $[x_3 < \zeta(x_1)]$, as shown in Fig. 1. The surface profile function $\zeta(x_1)$ is assumed to be either a single-

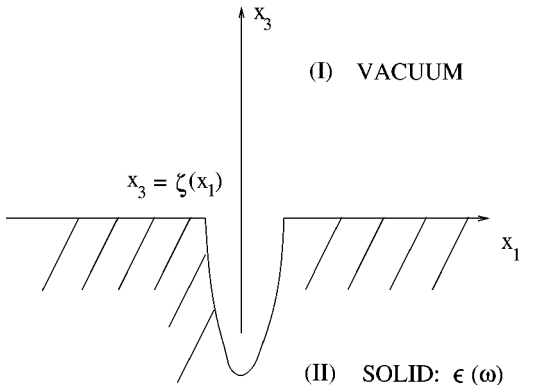

FIG. 1. A schematic diagram of the geometry studied in this paper.

 ©2002 The American Physical Society

valued or a multivalued function of x_1, and we further assume that it is an even function of x_1, $\zeta(-x_1)=\zeta(x_1)$, although this is an unessential simplification.

We seek solutions of the Maxwell curl equations

$$\nabla\times\mathbf{E}(\mathbf{x};t)=-\frac{1}{c}\frac{\partial}{\partial t}\mathbf{H}(\mathbf{x};t), \tag{2.1a}$$

$$\nabla\times\mathbf{H}(\mathbf{x};t)=\frac{1}{c}\frac{\partial}{\partial t}\mathbf{D}(\mathbf{x};t), \tag{2.1b}$$

in this system in the form

$$E_\alpha(\mathbf{x};t)=\hat{E}_\alpha(x_1,x_3|k,\omega)e^{ikx_2-i\omega t}, \quad \alpha=1,2,3, \tag{2.2a}$$

$$H_\alpha(\mathbf{x};t)=\hat{H}_\alpha(x_1,x_3|k,\omega)e^{ikx_2-i\omega t}, \quad \alpha=1,2,3, \tag{2.2b}$$

which reflects the invariance of our system against displacements along the x_2 axis. From Eqs. (2.1) and (2.2), together with the Maxwell equation $\nabla\cdot\mathbf{H}=0$, we obtain the equations satisfied by the components of the magnetic and electric field amplitudes parallel to the generators of the surface, $\hat{H}_2(x_1,x_3|k,\omega)$ and $\hat{E}_2(x_1,x_3|k,\omega)$, respectively. In region I these components satisfy the equations

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}-\beta_0^2(k,\omega)\right)\begin{Bmatrix}\hat{H}_2^>(x_1,x_3|k,\omega)\\ \hat{E}_2^>(x_1,x_3|k,\omega)\end{Bmatrix}=0. \tag{2.3a}$$

In region II they satisfy the equations

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}-\beta^2(k,\omega)\right)\begin{Bmatrix}\hat{H}_2^<(x_1,x_3|k,\omega)\\ \hat{E}_2^<(x_1,x_3|k,\omega)\end{Bmatrix}=0, \tag{2.3b}$$

where

$$\beta_0(k,\omega)=\left(k^2-\frac{\omega^2}{c^2}\right)^{1/2} \quad \left(|k|>\frac{\omega}{c}\right)$$

$$=-i\left(\frac{\omega^2}{c^2}-k^2\right)^{1/2} \quad \left(|k|<\frac{\omega}{c}\right) \tag{2.4a}$$

and

$$\beta(k,\omega)=\left[k^2-\epsilon(\omega)\frac{\omega^2}{c^2}\right]^{1/2}, \quad \mathrm{Re}\,\beta(k,\omega)>0,$$

$$\mathrm{Im}\,\beta(k,\omega)<0. \tag{2.4b}$$

We denote components of the electromagnetic field in region I by a superscript ">" and those in the region II by a superscript "<." We seek solutions of Eqs. (2.3a) that vanish as $x_3\to\infty$, since we seek solutions in the nonradiative region $|k|>\omega/c$, where β_0^2 is positive. Similarly we seek solutions of Eqs. (2.3b) that vanish as $x_3\to-\infty$ [since $\epsilon(\omega)$ is negative, β^2 is always positive].

The remaining, transverse, components of the magnetic and electric fields can be obtained in terms of $\hat{H}_2(x_1,x_3|k,\omega)$ and $\hat{E}_2(x_1,x_3|k,\omega)$[8]

$$\hat{\mathbf{H}}_\perp=\frac{i}{k^2-\epsilon\omega^2/c^2}\left(-k\nabla_\perp\hat{H}_2-\epsilon\frac{\omega}{c}[\hat{\mathbf{x}}_2\times\nabla_\perp\hat{E}_2]\right), \tag{2.5a}$$

$$\hat{\mathbf{E}}_\perp=\frac{i}{k^2-\epsilon\omega^2/c^2}\left(k\nabla_\perp\hat{E}_2+\frac{\omega}{c}[\hat{\mathbf{x}}_2\times\nabla_\perp\hat{H}_2]\right), \tag{2.5b}$$

where ϵ is the dielectric constant of the medium in which the fields are being calculated, $\nabla_\perp=(\partial/\partial x_1,0,\partial/\partial x_3)$ and the subscript $\perp$ denotes the field components transverse to the x_2 axis.

Thus, in what follows we will regard $\hat{H}_2(x_1,x_3|k,\omega)$ and $\hat{E}_2(x_1,x_3|k,\omega)$ as the primary field components and $\hat{E}_{1,3}(x_1,x_3|k,\omega)$ and $\hat{H}_{1,3}(x_1,x_3|k,\omega)$ as the derived field components. Therefore, we next seek the boundary conditions at the solid-vacuum interface in terms of $\hat{H}_2(x_1,x_3|k,\omega)$ and $\hat{E}_2(x_1,x_3|k,\omega)$ alone. These boundary conditions (not all independent) can be written in the forms

$$\hat{\mathbf{n}}\times\mathbf{E}^>=\hat{\mathbf{n}}\times\mathbf{E}^<, \quad \hat{\mathbf{n}}\cdot\mathbf{E}^>=\epsilon(\omega)\hat{\mathbf{n}}\cdot\mathbf{E}^<, \tag{2.6a}$$

$$\hat{\mathbf{n}}\times\mathbf{H}^>=\hat{\mathbf{n}}\times\mathbf{H}^<, \quad \hat{\mathbf{n}}\cdot\mathbf{H}^>=\hat{\mathbf{n}}\cdot\mathbf{H}^<, \tag{2.6b}$$

where $\hat{\mathbf{n}}$ is the unit vector normal to the interface at each point, directed from the solid into the vacuum. From Eqs. (2.5) and (2.6) we obtain the following boundary conditions:

$$\hat{E}_2^>=\hat{E}_2^<, \tag{2.7a}$$

$$\hat{H}_2^>=\hat{H}_2^<, \tag{2.7b}$$

$$\frac{\partial}{\partial n}\hat{E}_2^<=\frac{\beta^2(k,\omega)}{\epsilon(\omega)\beta_0^2(k,\omega)}\frac{\partial}{\partial n}\hat{E}_2^>-\left(\frac{\epsilon(\omega)-1}{\epsilon(\omega)}\right)\frac{\omega k}{c\beta_0^2(k,\omega)}\frac{\partial}{\partial t}\hat{H}_2^>, \tag{2.7c}$$

$$\frac{\partial}{\partial n}\hat{H}_2^<=\frac{\beta^2(k,\omega)}{\beta_0^2(k,\omega)}\frac{\partial}{\partial n}\hat{H}_2^>+[\epsilon(\omega)-1]\frac{\omega k}{c\beta_0^2(k,\omega)}\frac{\partial}{\partial t}\hat{E}_2^>, \tag{2.7d}$$

where $\partial/\partial n$ and $\partial/\partial t$ are the normalized derivatives along the normal and the tangent to the interface at each point. Explicit expressions for them will be given in Sec. III.

III. DISPERSION RELATION

We begin the derivation of the dispersion relation for channel polaritons by introducing two Green's functions $G_0(x_1,x_3|x_1',x_3')$ and $G_\epsilon(x_1,x_3|x_1',x_3')$, which satisfy the equations

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}-\beta_0^2(k,\omega)\right)G_0(x_1,x_3|x_1',x_3')$$

$$=-4\pi\delta(x_1-x_1')\delta(x_3-x_3') \tag{3.1a}$$

and

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}-\beta^2(k,\omega)\right)G_\epsilon(x_1,x_3|x_1',x_3') = -4\pi\delta(x_1-x_1')\delta(x_3-x_3'), \tag{3.1b}$$

respectively. $G_0(x_1,x_3|x_1',x_3')$ satisfies vanishing ($\beta_0^2>0$) or outgoing ($\beta_0^2<0$) wave boundary conditions at infinity, while $G_\epsilon(x_1,x_3|x_1',x_3')$ satisfies vanishing boundary conditions at infinity [recall that $\epsilon(\omega)$ is negative so that β^2 is always positive]. Explicit expressions for $G_0(x_1,x_3|x_1',x_3')$ and $G_\epsilon(x_1,x_3|x_1',x_3')$ are

$$G_0(x_1,x_3|x_1',x_3') = 2K_0\{\beta_0(k,\omega)[(x_1-x_1')^2 + (x_3-x_3')^2]^{1/2}\} \tag{3.2a}$$

$$= G_0(x_1',x_3'|x_1,x_3), \tag{3.2b}$$

$$G_\epsilon(x_1,x_3|x_1',x_3') = 2K_0\{\beta(k,\omega)[(x_1-x_1')^2+(x_3-x_3')^2]^{1/2}\} \tag{3.3a}$$

$$= G_\epsilon(x_1',x_3'|x_1,x_3), \tag{3.3b}$$

where $K_0(z)$ is a modified Bessel function.

The derivation of the dispersion relation is based on Green's second integral identity in the plane:[7]

$$\int dx_1\int dx_3\left[u(x_1,x_3)\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}\right)v(x_1,x_3)-v(x_1,x_3)\times\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}\right)u(x_1,x_3)\right] = \int_C ds\left[u(x_1,x_3)\frac{\partial}{\partial\nu}v(x_1,x_3)-v(x_1,x_3)\frac{\partial}{\partial\nu}u(x_1,x_3)\right], \tag{3.4}$$

where $u(x_1,x_3)$ and $v(x_1,x_3)$ are two arbitrary functions of x_1 and x_3 that are continuous and have continuous derivatives in an area A bounded by a closed curve C. The element of arc length along this curve is ds, and the derivative $\partial/\partial\nu$ is taken along the normal to the curve C at each point, directed away from the area A.

We first apply Eq. (3.4) to region I. The curve C is then the union of the curve $x_3=\zeta(x_1)$, which we denote by $\hat{s}$, and a semicircle of infinite radius in the upper half-plane, which we denote by $C^{+\infty}$. If we choose for $u(x_1,x_3)$ the function $\hat{\Phi}^>(x_1,x_3|k,\omega)$, which is either $\hat{H}_2^>(x_1,x_3|k,\omega)$ or $\hat{E}_2^>(x_1,x_3|k,\omega)$, and assume $v(x_1,x_3)$ to be $G_0(x_1,x_3|x_1',x_3')$, then in view of Eqs. (2.3a) and (3.1a), we obtain from Eq. (3.4) the result that

$$\theta^I(\mathbf{x})\hat{\Phi}^>(x_1,x_3|k,\omega) = \frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_0(x_1,x_3|x_1',x_3')\right]\times\hat{\Phi}^>(x_1',x_3'|k,\omega)-G_0(x_1,x_3|x_1',x_3')\times\left[\frac{\partial}{\partial n'}\hat{\Phi}^>(x_1',x_3'|k,\omega)\right]\right\}, \tag{3.5}$$

where $\mathbf{x}=(x_1,0,x_3)$ and

$$\theta^I(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x}\in \text{I},\\ 0 & \text{otherwise.}\end{cases} \tag{3.6}$$

Because of the vanishing boundary conditions at infinity satisfied by $\hat{\Phi}^>(x_1,x_3|k,\omega)$ and $G_0(x_1,x_3|x_1',x_3')$, there is no contribution to the right-hand side of Eq. (3.5) from the integral along the curve $C^{+\infty}$. We have also used the symmetry of $G_0(x_1,x_3|x_1',x_3')$, Eq. (3.2b), in writing Eq. (3.5).

We next apply Eq. (3.4) to region II. The curve C in this case is the union of the curve $\hat{s}$ and a semicircle of infinite radius in the lower half-plane, which we denote by $C^{-\infty}$. We choose for $u(x_1,x_3)$ the function $\hat{\Phi}^<(x_1,x_3|k,\omega)$, which is either $\hat{H}_2^<(x_1,x_3|k,\omega)$ or $\hat{E}_2^<(x_1,x_3|k,\omega)$, and assume $v(x_1,x_3)$ to be $G_\epsilon(x_1,x_3|x_1',x_3')$. Then in view of Eqs. (2.3b) and (3.1b), we obtain from Eq. (3.4) the result that

$$\theta^{II}(\mathbf{x})\hat{\Phi}^<(x_1,x_3|k,\omega) = -\frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_\epsilon(x_1,x_3|x_1',x_3')\right]\hat{\Phi}^<(x_1',x_3'|k,\omega) - G_\epsilon(x_1,x_3|x_1',x_3')\left[\frac{\partial}{\partial n'}\hat{\Phi}^<(x_1',x_3'|k,\omega)\right]\right\}, \tag{3.7}$$

where

$$\theta^{II}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x}\in \text{II},\\ 0 & \text{otherwise.}\end{cases} \tag{3.8}$$

Again, because of the vanishing boundary conditions at infinity satisfied by $\hat{\Phi}^<(x_1,x_3|k,\omega)$ and $G_\epsilon(x_1,x_3|x_1',x_3')$, there is no contribution to the right-hand side of Eq. (3.7) from the integral along the curve $C^{-\infty}$. We have also used the symmetry of $G_\epsilon(x_1,x_3|x_1',x_3')$, Eq. (3.3b), in writing this equation.

Since we have assumed that the surface profile function $\zeta(x_1)$ can be a multivalued function of x_1, we next describe the method of the parametrization of the surface profile with the notation introduced in Ref. 6. We consider that the surface profile is a curve with a finite number of singular points. This curve can be characterized in terms of a parameter s that is chosen as the arc length along the curve. The coordinates of each point of this curve are defined by a single-valued vector function $\mathbf{R}(s)$ which has at least second derivatives at each regular point. The vector $\mathbf{R}(s)$ is given by

$$\mathbf{R}(s)=(\xi(s),0,\eta(s)), \tag{3.9}$$

where $\xi(s)$ and $\eta(s)$ are parametric functions of s. The normal and tangential derivatives to the interface at each regular point in terms of s are defined by

$$\frac{\partial}{\partial n}=-\eta'(s)\frac{\partial}{\partial \xi}+\xi'(s)\frac{\partial}{\partial \eta}, \tag{3.10a}$$

$$\frac{\partial}{\partial t}=\xi'(s)\frac{\partial}{\partial \xi}+\eta'(s)\frac{\partial}{\partial \eta}. \tag{3.10b}$$

The prime in Eqs. (3.10) denotes differentiation with respect to s. The two equations represented by Eq. (3.5) can then be written explicitly in the forms

$$\theta^{\mathrm{I}}(\mathbf{x})\hat{H}_2^{>}(x_1,x_3|k,\omega)=\frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_0(x_1,x_3|\xi(s'),\eta(s'))\right]\times H(s'|k,\omega)-G_0(x_1,x_3|\xi(s'),\eta(s'))L(s'|k,\omega)\right\}, \tag{3.11a}$$

$$\theta^{\mathrm{I}}(\mathbf{x})\hat{E}_2^{>}(x_1,x_3|k,\omega)=\frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_0(x_1,x_3|\xi(s'),\eta(s'))\right]\times E(s'|k,\omega)-G_0(x_1,x_3|\xi(s'),\eta(s'))F(s'|k,\omega)\right\}, \tag{3.11b}$$

where we have defined the four source functions

$$H(s|k,\omega)=\hat{H}_2^{>}(\xi(s),\eta(s)|k,\omega), \tag{3.12a}$$

$$L(s|k,\omega)=\frac{\partial}{\partial n}\hat{H}_2^{>}(\xi(s),\eta(s)|k,\omega), \tag{3.12b}$$

$$E(s|k,\omega)=\hat{E}_2^{>}(\xi(s),\eta(s)|k,\omega), \tag{3.12c}$$

$$F(s|k,\omega)=\frac{\partial}{\partial n}\hat{E}_2^{>}(\xi(s),\eta(s)|k,\omega). \tag{3.12d}$$

Turning now to the two equations represented by Eq. (3.7), we write them out explicitly, using the boundary conditions (2.7) with the results that

$$\theta^{II}(\mathbf{x})\hat{H}_2^{<}(x_1,x_3|k,\omega)=-\frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right]H(s'|k,\omega)-\frac{\beta^2(k,\omega)}{\beta_0^2(k,\omega)}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))L(s'|k,\omega)-[\epsilon(\omega)-1]\frac{\omega k}{c\beta_0^2(k,\omega)}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\frac{\partial}{\partial t'}\hat{E}_2^{>}(\xi(s'),\eta(s')|k,\omega)\right\} \tag{3.13a}$$

and

$$\theta^{II}(\mathbf{x})\hat{E}_2^{<}(x_1,x_3|k,\omega)=-\frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right]E(s'|k,\omega)-\frac{\beta^2(k,\omega)}{\epsilon(\omega)\beta_0^2(k,\omega)}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))F(s'|k,\omega)+\frac{\epsilon(\omega)-1}{\epsilon(\omega)}\frac{\omega k}{c\beta_0^2(k,\omega)}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\frac{\partial}{\partial t'}\hat{H}_2^{>}(\xi(s'),\eta(s')|k,\omega)\right\}. \tag{3.13b}$$

At this point we note the results that

$$\left[\frac{\partial}{\partial t}\hat{E}_2^{>}(\xi(s),\eta(s)|k,\omega)\right]=\frac{d}{ds}\hat{E}_2^{>}(\xi(s),\eta(s)|k,\omega) \tag{3.14a}$$

$$=\frac{d}{ds}E(s|k,\omega), \tag{3.14b}$$

$$\left[\frac{\partial}{\partial t}\hat{H}_2^{>}(\xi(s),\eta(s)|k,\omega)\right]=\frac{d}{ds}\hat{H}_2^{>}(\xi(s),\eta(s)|k,\omega) \tag{3.14c}$$

$$=\frac{d}{ds}H(s|k,\omega). \tag{3.14d}$$

When the results given by Eqs. (3.14) are substituted into Eqs. (3.13), after an integration by parts we obtain a second pair of equations containing the source functions

$$\theta^{II}(\mathbf{x})\hat{H}_2^{<}(x_1,x_3|k,\omega)=-\frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right]H(s'|k,\omega)-\frac{\beta^2(k,\omega)}{\beta_0^2(k,\omega)}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))L(s'|k,\omega)\right.$$
$$\left.+[\epsilon(\omega)-1]\frac{\omega k}{c\beta_0^2(k,\omega)}\left[\frac{\partial}{\partial t'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right]E(s'|k,\omega)\right\} \tag{3.15}$$

and

$$\theta^{II}(\mathbf{x})\hat{E}_2^{<}(x_1,x_3|k,\omega)=-\frac{1}{4\pi}\int_{\hat{s}}ds'\left\{\left[\frac{\partial}{\partial n'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right]E(s'|k,\omega)\right.$$
$$-\frac{\beta^2(k,\omega)}{\epsilon(\omega)\beta_0^2(k,\omega)}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))F(s'|k,\omega)$$
$$\left.-\frac{\epsilon(\omega)-1}{\epsilon(\omega)}\frac{\omega k}{c\beta_0^2(k,\omega)}\left[\frac{\partial}{\partial t'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right]H(s'|k,w)\right\}. \tag{3.16}$$

To complete our derivation of the dispersion relation for channel polaritons we set $\mathbf{x}=\mathbf{R}(s)+v\hat{\mathbf{n}}$, where v is a positive infinitesimal, in Eqs. (3.11), (3.15), and (3.16), to obtain

$$H(s|k,\omega)=\int_{\hat{s}}ds'[H_0(s|s')H(s'|k,\omega)]-L_0(s|s')L(s'|k,\omega)], \tag{3.17a}$$

$$E(s|k,\omega)=\int_{\hat{s}}ds'[H_0(s|s')H(s'|k,\omega)]-L_0(s|s')F(s'|k,\omega)], \tag{3.17b}$$

$$0=-\int_{\hat{s}}ds'\left[H_\epsilon(s|s')H(s'|k,\omega)-\frac{\beta^2(k,\omega)}{\beta_0^2(k,\omega)}L_\epsilon(s|s')L(s'|k,\omega)+[\epsilon(\omega)-1]\frac{\omega k}{c\beta_0^2(k,\omega)}N_\epsilon(s|s')E(s'|k,\omega)\right], \tag{3.17c}$$

$$0=-\int_{\hat{s}}ds'\left[H_\epsilon(s|s')E(s'|k,\omega)-\frac{\beta^2(k,\omega)}{\epsilon(\omega)\beta_0^2(k,\omega)}L_\epsilon(s|s')F(s'|k,\omega)-\left(\frac{\epsilon(\omega)-1}{\epsilon(\omega)}\right)\frac{\omega k}{c\beta_0^2(k,\omega)}N_\epsilon(s|s')H(s'|k,\omega)\right], \tag{3.17d}$$

where the kernels are given by

$$H_0(s|s')=\frac{1}{4\pi}\left[\frac{\partial}{\partial n'}G_0(x_1,x_3|\xi(s'),\eta(s'))\right],\quad x_1=\xi(s)-v\,\eta'(s),\quad x_3=\eta(s)+v\,\xi'(s), \tag{3.18a}$$

$$L_0(s|s')=\frac{1}{4\pi}[G_0(x_1,x_3|\xi(s'),\eta(s'))],\quad x_1=\xi(s)-v\,\eta'(s),\quad x_3=\eta(s)+v\,\xi'(s), \tag{3.18b}$$

$$H_\epsilon(s|s')=\frac{1}{4\pi}\left[\frac{\partial}{\partial n'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right],\quad x_1=\xi(s)-v\,\eta'(s),\quad x_3=\eta(s)+v\,\xi'(s), \tag{3.18c}$$

$$L_\epsilon(s|s')=\frac{1}{4\pi}[G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))],\quad x_1=\xi(s)-v\,\eta'(s),\quad x_3=\eta(s)+v\,\xi'(s), \tag{3.18d}$$

$$N_\epsilon(s|s')=\frac{1}{4\pi}\left[\frac{\partial}{\partial t'}G_\epsilon(x_1,x_3|\xi(s'),\eta(s'))\right],\quad x_1=\xi(s)-v\,\eta'(s),\quad x_3=\eta(s)+v\,\xi'(s). \tag{3.18e}$$

Equations (3.17) constitute a set of four coupled, homogeneous integral equations for the four source functions $H(s|k,\omega)$, $L(s|k,\omega)$, $E(s|k,\omega)$, and $F(s|k,\omega)$. The solvability condition for this system of equations is the dispersion relation for channel polaritons.

We now turn to the numerical solution of this system of equations.

IV. NUMERICAL SOLUTION OF THE DISPERSION RELATION

To solve the system of equations (3.17) we begin by replacing the infinite range of integration by the finite range $(0,2L+\Delta s)$, where $L=N\Delta s$ is a half-length of the surface curve and N is a large positive integer. We next introduce the set of $2N+1$ equally spaced points $\{s_n\}$, where

$$s_n=\Delta s\left(\frac{1}{2}+n+N\right), \quad n=-N,-N+1,\ldots,N-1,N, \tag{4.1}$$

and rewrite Eq. (3.17a) as

$$H(s|k,\omega)=\sum_{n=-N}^{N}\int_{s_n-\Delta s/2}^{s_n+\Delta s/2}ds'[H_0(s|s')H(s'|k,\omega) - L_0(s|s')L(s'|k,\omega)]. \tag{4.2}$$

On the assumption that $H(s|k,\omega)$ and $L(s|k,\omega)$ are slowly varying functions of s in each interval $(s_n-\frac{1}{2}\Delta s, s_n+\frac{1}{2}\Delta s)$, we replace Eq. (4.2) by

$$H(s|k,\omega)=\sum_{n=-N}^{N}\int_{s_n-\Delta s/2}^{s_n+\Delta s/2}ds'[H_0(s|s')H(s_n|k,\omega) - L_0(s|s')L(s_n|k,\omega)]. \tag{4.3}$$

By replacing s with s_m and s' with s_n+u, we obtain the matrix equation

$$H(s_m|k,\omega)=\sum_{n=-N}^{N}[H_{mn}^{(0)}H(s_n|k,\omega)-L_{mn}^{(0)}L(s_n|k,\omega)], \tag{4.4}$$

where

$$H_{mn}^{(0)}=\int_{-\Delta s/2}^{\Delta s/2}duH_0(s_m|s_n+u), \tag{4.5}$$

$$L_{mn}^{(0)}=\int_{-\Delta s/2}^{\Delta s/2}duL_0(s_m|s_n+u). \tag{4.6}$$

In this way we convert the four equations (3.17) into the following matrix equations:

$$H(s_m|k,\omega)=\sum_{n=-N}^{N}[H_{mn}^{(0)}H(s_n|k,\omega)-L_{mn}^{(0)}L(s_n|k,\omega)], \tag{4.7a}$$

$$E(s_m|k,\omega)=\sum_{n=-N}^{N}[H_{mn}^{(0)}E(s_n|k,\omega)-L_{mn}^{(0)}F(s_n|k,\omega)], \tag{4.7b}$$

$$0=-\sum_{m=-N}^{N}\left[H_{mn}^{(\epsilon)}H(s_n|k,\omega)-\frac{\beta^2(k,\omega)}{\beta_0^2(k,\omega)}L_{mn}^{(\epsilon)}L(s_n|k,\omega) + [\epsilon(\omega)-1]\frac{\omega k}{c\beta_0^2(k,\omega)}N_{mn}^{(\epsilon)}E(s_n|k,\omega)\right], \tag{4.7c}$$

$$0=-\sum_{m=-N}^{N}\left[E_{mn}^{(\epsilon)}H(s_n|k,\omega)-\frac{\beta^2(k,\omega)}{\epsilon(\omega)\beta_0^2(k,\omega)}L_{mn}^{(\epsilon)}F(s_n|k,\omega) - \left(\frac{\epsilon(\omega)-1}{\epsilon(\omega)}\right)\frac{\omega k}{c\beta_0^2(k,\omega)}N_{mn}^{(\epsilon)}H(s_n|k,\omega)\right], \tag{4.7d}$$

where

$$H_{mn}^{(\epsilon)}=\int_{-\Delta s/2}^{\Delta s/2}duH_\epsilon(s_m|s_n+u), \tag{4.8a}$$

$$L_{mn}^{(\epsilon)}=\int_{-\Delta s/2}^{\Delta s/2}duL_\epsilon(s_m|s_n+u), \tag{4.8b}$$

$$N_{mn}^{(\epsilon)}=\int_{-\Delta s/2}^{\Delta s/2}duN_\epsilon(s_m|s_n+u). \tag{4.8c}$$

The matrix elements $H_{mn}^{(\epsilon)}$ and $L_{mn}^{(\epsilon)}$ have been evaluated by a standard method[9] and to the lowest nonzero order in Δs are given by

$$H_{mn}^{(\epsilon)}=\frac{\Delta s}{2\pi}\beta^2(k,\omega)\frac{K_1(\beta(k,\omega)\{[\xi(s_m)-\xi(s_n)]^2+[\eta(s_m)-\eta(s_n)]^2\}^{1/2})}{\beta(k,\omega)\{[\xi(s_m)-\xi(s_n)]^2+[\eta(s_m)-\eta(s_n)]^2\}^{1/2}}\{\xi'(s_n)[\eta(s_m)-\eta(s_n)] - \eta'(s_n)[\xi(s_m)-\xi(s_n)]\} \quad (m\neq n) \tag{4.9a}$$

$$=\frac{1}{2}+\Delta s\frac{\eta''(s_m)\xi'(s_m)-\xi''(s_m)n'(s_m)}{4\pi} \quad (m=n), \tag{4.9b}$$

$$L_{mn}^{(\epsilon)}=\frac{\Delta s}{2\pi}K_0(\beta(k,\omega)\{[\xi(s_m)-\xi(s_n)]^2+[\eta(s_m)-\eta(s_n)]^2\}^{1/2}) \quad (m\neq n) \tag{4.10a}$$

$$=\frac{\Delta s}{2\pi}K_0\left(\frac{\beta(k,\omega)\Delta s}{2e}\right) \quad (m=n). \tag{4.10b}$$

The matrix elements $H^{(0)}_{mn}$ and $L^{(0)}_{mn}$ are obtained from $H^{(\epsilon)}_{mn}$ and $L^{(\epsilon)}_{mn}$ by setting $\beta(k,\omega)=\beta_0(k,\omega)$ in them. The matrix element $N^{(\epsilon)}_{mn}$ is given to first order in Δs by

$$N^{(\epsilon)}_{mn}=\frac{\Delta s}{2\pi}\beta^2(k,\omega)\frac{K_1(\beta(k,\omega)\{[\xi(s_m)-\xi(s_n)]^2+[\eta(s_m)-\eta(s_n)]^2\}^{1/2})}{\beta(k,\omega)\{[\xi(s_m)-\xi(s_n)]^2+[\eta(s_m)-\eta(s_n)]^2\}^{1/2}}\{\xi'(s_n)[\xi(s_m)-\xi(s_n)]$$
$$+\eta'(s_n)[\eta(s_m)-\eta(s_n)]\} \quad (m\neq n) \tag{4.11a}$$
$$=0 \quad (m=n). \tag{4.11b}$$

We can simplify Eqs. (4.7) by using the symmetry of the problem. Because we have assumed that the surface profile function is an even function of x_1 or $(s-L)$, then $\xi(s_{-m})=-\xi(s_m)$ and $\eta(s_{-m})=\eta(s_m)$. We can see from Eqs. (4.9) that the matrix elements have the following properties:

$$L^{(0)}_{mn}=L^{(0)}_{-m,-n}, \quad L^{(\epsilon)}_{mn}=L^{(\epsilon)}_{-m,-n},$$
$$H^{(0)}_{mn}=H^{(0)}_{-m,-n}, \quad H^{(\epsilon)}_{mn}=H^{(\epsilon)}_{-m,-n}, \tag{4.12}$$
$$N^{(\epsilon)}_{mn}=-N^{(\epsilon)}_{-m,-n}.$$

By introducing the even and odd functions

$$H^{(e)}(s_m|\omega,k)=H(s_m|\omega,k)+H(s_{-m}|\omega,k),$$
$$H^{(o)}(s_m|\omega,k)=H(s_m|\omega,k)-H(s_{-m}|\omega,k),$$
$$E^{(e)}(s_m|\omega,k)=E(s_m|\omega,k)+E(s_{-m}|\omega,k), \tag{4.13}$$
$$E^{(o)}(s_m|\omega,k)=E(s_m|\omega,k)-E(s_{-m}|\omega,k)$$

and

$$F^{(e)}(s_m|\omega,k)=F(s_m|\omega,k)+F(s_{-m}|\omega,k),$$
$$F^{(o)}(s_m|\omega,k)=F(s_m|\omega,k)-F(s_{-m}|\omega,k),$$
$$L^{(e)}(s_m|\omega,k)=L(s_m|\omega,k)+L(s_{-m}|\omega,k), \tag{4.14}$$
$$L^{(o)}(s_m|\omega,k)=L(s_m|\omega,k)-L(s_{-m}|\omega,k)$$

for the mode in which the electric field is an even function of $(s-L)$ and the magnetic field is an odd function of $(s-L)$, we can rewrite Eqs. (4.4) and (4.7) as

$$\sum_{n=0}^{N}[\mathcal{H}^{even(0)}_{mn}E^{(e)}(s_n|\omega,k)-\mathcal{L}^{even(0)}_{mn}F^{(e)}(s_n|\omega,k)]=0,$$
$$\sum_{n=1}^{N}[\mathcal{H}^{odd(0)}_{mn}H^{(o)}(s_n|\omega,k)-\mathcal{L}^{odd(0)}_{mn}L^{(o)}(s_n|\omega,k)]=0, \tag{4.15}$$

$$\sum_{n=0}^{N}\left[\mathcal{H}^{even(\epsilon)}_{mn}E^{(\epsilon)}(s_n|\omega,k)-\frac{\beta^2(k,\omega)}{\epsilon(\omega)\beta_0^2(k,\omega)}\mathcal{L}^{even(\epsilon)}_{mn}F^{(e)}(s_n|\omega,k)\right]$$
$$+\frac{[\epsilon(\omega)-1]\omega k}{\epsilon(\omega)c\beta_0^2(k,\omega)}\sum_{n=1}^{N}[\mathcal{N}^{odd(\epsilon)}_{mn}H^{(o)}(s_n|\omega,k)]=0,$$
$$\sum_{n=1}^{N}\left[\mathcal{H}^{odd(\epsilon)}_{mn}H^{(o)}(s_n|\omega,k)-\frac{\beta^2(k,\omega)}{\beta_0^2(k,\omega)}\mathcal{L}^{odd(\epsilon)}_{mn}L^{(o)}(s_n|\omega,k)\right]$$
$$-\frac{[\epsilon(\omega)-1]\omega k}{c\beta_0^2(k,\omega)}\sum_{n=1}^{N}[\mathcal{N}^{even(\epsilon)}_{mn}E^{(e)}(s_n|\omega,k)]=0.$$

For brevity we will call this mode an $E^{(e)}H^{(o)}$ mode. The matrix elements are defined by

$$\mathcal{H}^{even(\epsilon,0)}_{m0}=H^{(\epsilon,0)}_{m0}-\delta\delta_{m0},$$
$$\mathcal{H}^{even(\epsilon,0)}_{mn}=H^{(\epsilon,0)+}_{mn}H^{(\epsilon,0)}_{m,-n}-\delta(\delta_{mn}+\delta_{m,-n}), \quad 1\leqslant n\leqslant N,$$
$$\mathcal{H}^{odd(\epsilon,0)}_{mn}=H^{(\epsilon,0)-}_{mn}H^{(\epsilon,0)}_{m,-n}-\delta(\delta_{mn}-\delta_{m,-n}), \quad 1\leqslant n\leqslant N,$$
$$\mathcal{L}^{even(\epsilon,0)}_{m0}=H^{(\epsilon,0)}_{m0},$$
$$\mathcal{L}^{even(\epsilon,0)}_{mn}=L^{(\epsilon,0)}_{mn}+L^{(\epsilon,0)}_{m,-n}, \quad 1\leqslant n\leqslant N, \tag{4.16}$$
$$\mathcal{L}^{odd(\epsilon,0)}_{mn}=L^{(\epsilon,0)}_{mn}-L^{(\epsilon,0)}_{m,-n}, \quad 1\leqslant n\leqslant N,$$
$$\mathcal{N}^{even(\epsilon,0)}_{m0}=N^{(\epsilon,0)}_{m0},$$
$$\mathcal{N}^{even(\epsilon,0)}_{mn}=N^{(\epsilon,0)}_{mn}+N^{(\epsilon,0)}_{m,-n}, \quad 1\leqslant n\leqslant N,$$
$$\mathcal{N}^{odd(\epsilon,0)}_{mn}=N^{(\epsilon,0)}_{mn}-N^{(\epsilon,0)}_{m,-n}, \quad 1\leqslant n\leqslant N,$$

where the superscript $(\epsilon,0)$ can be either ϵ or 0; $\delta=1$ if the superscript is 0; and $\delta=0$ if the superscript is ϵ. The size of the supermatrix in Eq. (4.15) is $(4N-2)\times(4N-2)$, essentially half the size of the supermatrix defined by Eq. (4.7). The solvability condition for the matrix equation (4.15) yields the dispersion curve for the $E^{(e)}H^{(o)}$ mode. For the

$E^{(o)}H^{(e)}$ mode, in which the magnetic field is an even function of $(s-L)$, while the electric field is an odd function of $(s-L)$, the supermatrix is defined by

$$\sum_{n=1}^{N}[\mathcal{H}_{mn}^{odd(0)}E^{(o)}(s_n|\omega,k)-\mathcal{L}_{mn}^{odd(0)}F^{(o)}(s_n|\omega,k)]=0,$$

$$\sum_{n=0}^{N}[\mathcal{H}_{mn}^{even(0)}H^{(e)}(s_n|\omega,k)-\mathcal{L}_{mn}^{even(0)}L^{(e)}(s_n|\omega,k)]=0,$$

$$\sum_{n=0}^{N}\left[\mathcal{H}_{mn}^{odd(\epsilon)}E^{(o)}(s_n|\omega,k)-\frac{\beta^2(k,\omega)}{\epsilon(\omega)\beta_0^2(k,\omega)}\mathcal{L}_{mn}^{odd(\epsilon)}F^{(o)}(s_n|\omega,k)\right]+\frac{(\epsilon(\omega)-1)\omega k}{\epsilon(\omega)c\beta_0^2(k,\omega)}\sum_{n=0}^{N}[\mathcal{N}_{mn}^{even(\epsilon)}H^{(o)}(s_n|\omega,k)]=0,$$

$$\sum_{n=0}^{N}\left[\mathcal{H}_{mn}^{even(\epsilon)}H^{(e)}(s_n|\omega,k)-\frac{\beta^2(k,\omega)}{\beta_0^2(k,\omega)}\mathcal{L}_{mn}^{even(\epsilon)}L^{(e)}(s_n|\omega,k)\right]-\frac{(\epsilon(\omega)-1)\omega k}{\epsilon(\omega)c\beta_0^2(k,\omega)}\sum_{n=1}^{N}[\mathcal{N}_{mn}^{odd(\epsilon)}E^{(o)}(s_n|\omega,k)]=0. \tag{4.17}$$

The solvability condition for this matrix equation gives the dispersion curves for the $E^{(o)}H^{(e)}$ channel polaritons. The determinant of the matrix is a function of k and ω. The roots of this function define the branches of the dispersion curve.

The numerical approach to finding of the roots is the following. First, we assume a value of k and increase ω systematically until the determinant of the supermatrix changes its sign. Second, by the method of dichotomy we limit the interval in which the determinant changes its sign to obtain the required accuracy. Then we continue increasing ω to get all roots of the determinant for a given value of k. By repeating this procedure for different values of k we find all branches of the dispersion relation.

Now we turn to the calculation of the coordinate dependences of the fields of a channel polariton. Based on the preceding results we can choose any point on a branch of the dispersion relation. To illustrate the field patterns of a channel polariton we take points on the distinct branches. Every point on a branch defined by a frequency and wave number (ω,k) of a channel polariton characterizes a particular single mode. First, we consider an $E^{(e)}H^{(o)}$ mode. The fields everywhere in our system can be obtained from Eqs. (4.7a) and (4.7b) and Eqs. (2.5), where the source functions $E^{(e)}(s_m|\omega,k)$ and $H^{(o)}(s_m|\omega,k)$ are the solutions of the matrix equations (4.15). Here ω and k define a point on the branch, and therefore the determinant of this matrix is equal to zero. In fact, these source functions are the components of the eigenvector of this matrix with the corresponding eigenvalue $\lambda=0$, which can be calculated by using a standard numerical routine. We normalize the source functions assuming that the norm of this eigenvector is unity. The same procedure applies to the $E^{(o)}H^{(e)}$ mode.

V. RESULTS

The dispersion relation for the $E^{(e)}H^{(o)}$ modes obtained from Eq. (4.15) and the dispersion relation for the $E^{(o)}H^{(e)}$ modes obtained from Eqs. (4.17), each has an infinite number of branches. This can be understood most simply by considering the limit as $k\to\infty$, where the electrostatic approximation is valid. In this limit, as was shown in Ref. 5, each branch is related to an eigenvalue of an $N\times N$ matrix, in the limit as $N\to\infty$. Thus the number of branches tends to infinity in this limit. In our numerical calculations, to simplify the resulting figures, we considered only six branches for every case considered, which for a given value of the wave number k were the three lowest-frequency branches and the three highest-frequency branches.

In Fig. 2 we plot the branches of the dispersion curve for channel polaritons guided by a channel defined by a Gaussian surface profile function $\zeta(x_1)=-A\exp(-x_1^2/R^2)$, with $A/R=8$. We assume that the substrate is a metal and take for its dielectric function the simple free electron form $\epsilon(\omega)=1-(\omega_p^2/\omega^2)$, where ω_p is the plasma frequency of the electrons in the bulk of the metal. We can vary one independent dimensionless parameter of the system, $\alpha=\omega_pR/c$. Here $\alpha=0$ is the electrostatic case, which was considered in Ref. 5. The results for $\alpha=0.1,0.5,1.0$ are presented in Figs. 2(a)–2(c), corresponding to the $E^{(e)}H^{(o)}$ mode. For these values of α there are branches with frequencies above and below $\omega_p/\sqrt{2}$, which is the limiting frequency as $k\to\infty$ of surface plasmon polaritons at a planar vacuum-metal interface. In Figs. 2(d)–2(f) the results are plotted for the same α but for electric fields that are odd functions of x_1 and magnetic fields that are even functions of x_1. For $\alpha=0.1$ there are some branches that lie above $\omega_p/\sqrt{2}$, but the increase of α shifts all dispersion curves well below $\omega_p/\sqrt{2}$. The frequencies of all branches lie inside the range $0<\omega<\omega_p$, and all of them tend to $\omega_p/\sqrt{2}$ as $k\to\infty$. This is because as k becomes infinite and the wavelength of the channel polariton becomes smaller than any linear dimension characterizing the channel, the channel polariton ''sees'' a locally planar surface, and the frequency of each branch tends to the $k\to\infty$ limiting frequency of surface polaritons at a planar vacuum-metal interface. This limiting behavior occurs for all of the channel profiles considered in this paper.

We also note that although we have plotted the dispersion curves for six branches, in Figs. 2(a)–2(c) the second and third branches, counted from the lowest-frequency branch, are degenerate on the scale of the figure, while in Fig. 2(d) the fourth and fifth branches are degenerate on the same scale. Consequently only five branches appear to be plotted in these figures.

In each figure we have also plotted for comparison the

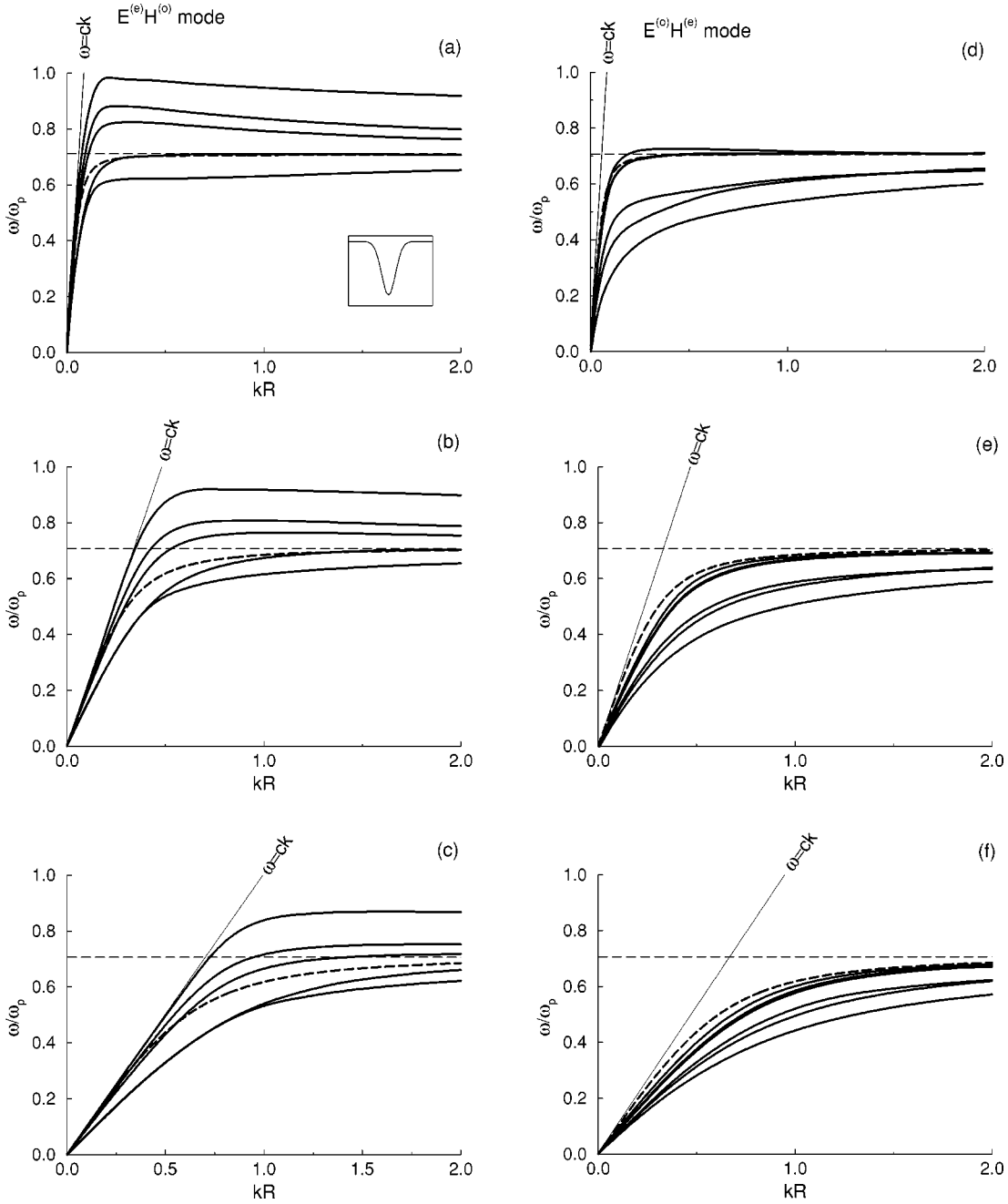

FIG. 2. The branches of the dispersion curve for channel polaritons guided by a channel on a metallic substrate defined by the profile function $\zeta(x_1) = -A \exp(-x_1^2/R^2)$, $A/R=8$. The left side of the plot corresponds to the $E^{(e)}H^{(o)}$ modes while the right side corresponds to the $E^{(o)}H^{(e)}$ modes. (a) and (d) $\alpha=0.1$; (b) and (e) $\alpha=0.5$; (c) and (f) $\alpha=1.0$, where $\alpha=\omega_p R/c$. The dark dashed curve in each figure is the dispersion curve for surface plasmon polaritons at a planar vacuum-metal interface.

dispersion curve for surface plasmon polaritons on a planar vacuum-metal interface obtained from the dispersion relation $kR=\alpha\tilde{\omega}[(\tilde{\omega}^2-1)/(2\tilde{\omega}^2-1)]^{1/2}$, where $\tilde{\omega}=\omega/\omega_p$.

In Fig. 3 we plot the amplitude of the second component of the electric field of a channel polariton along the x_3 axis for the four highest $E^{(e)}H^{(o)}$ modes in the system with the same parameters as in Fig. 2(c), with wave number $q=1/R$ and frequencies given by $\omega/\omega_p=0.883$, 0.774, 0.732, and 0.707. The field for the lower-frequency modes has a larger number of nodes across the channel. The field decreases so rapidly with increasing x_3 that it is practically zero at the mouth of the channel and in the vacuum above it.

In Figs. 4–7 we present contour plots of the squared modulus of the second component of the electric field inside the channel. The lighter regions correspond to the larger values of the electric field. The most intense field is concentrated in the vicinity of the tip of the groove. From these figures we also see that the electric field is strongly confined to the channel. For all practical purposes it vanishes outside the channel.

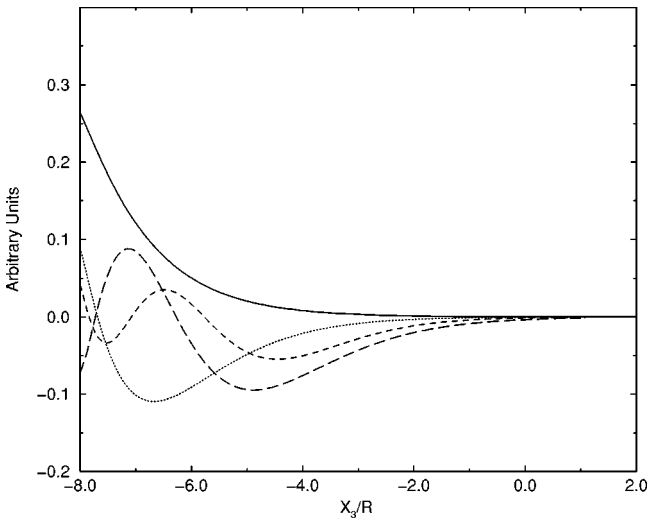

FIG. 3. The amplitude of the electric field of a channel polariton for four $E^{(e)}H^{(o)}$ modes in the system with the same parameters as in Fig. 2(c), with the wave number $q=1/R$ and the frequencies given by $\omega/\omega_p=0.883$ (solid line), 0.774 (dotted line), 0.732 (dashed line), and 0.707 (long dashed line).

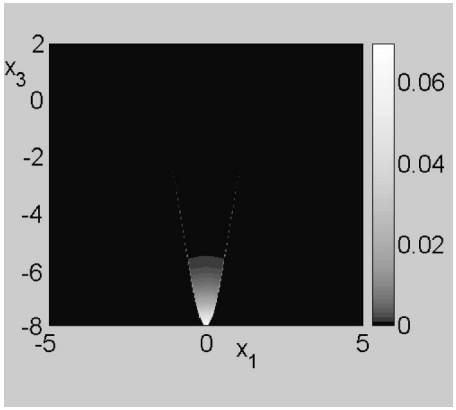

FIG. 4. A contour plot of the squared modulus of the second component of the electric field in a channel defined by a Gaussian profile. The wave number of the electric field is $q=1/R$, and the frequency is given by $\omega/\omega_p=0.883$. x_1 and x_3 are expressed in units of R.

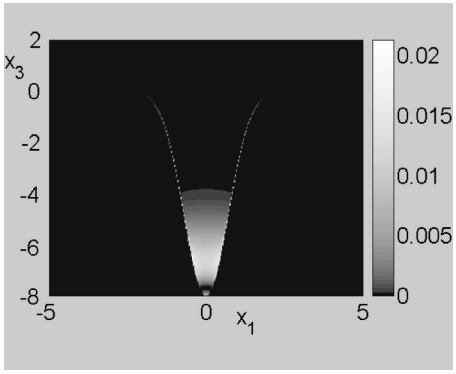

FIG. 5. The same as in Fig. 4, but with the frequency of the electric field given by $\omega/\omega_p=0.774$.

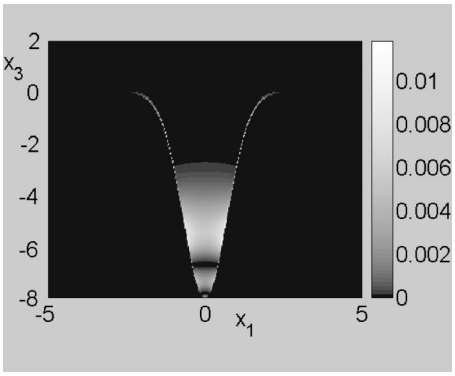

FIG. 6. The same as in Fig. 4, but with the frequency of the electric field given by $\omega/\omega_p=0.732$.

In Fig. 8 we present the branches of the dispersion curve for channel polaritons guided by a channel defined by the surface profile function

$$\zeta(x_1)=-d\,\frac{2\cosh^2(\beta w/4)}{\cosh(\beta w/2)+\cosh\beta x_1}. \tag{5.1}$$

In the limit as $\beta\to\infty$ this profile defines the rectangular channel of width w and depth d whose surface profile function is given by

$$\zeta(x_1)=-d \quad \left(|x_1|<\frac{w}{2}\right) \tag{5.2a}$$

$$=0 \quad \left(|x_1|>\frac{w}{2}\right). \tag{5.2b}$$

For this profile we vary a dimensionless parameter $\alpha=\omega_p w/c$. Figures 8(a)–8(c) present dispersion curves with $\alpha=(0.1,0.5,1.0)$, corresponding to electric fields that are even functions of x_1 and magnetic fields that are odd functions of x_1. In Figs. 8(d)–8(f) we plot the dispersion curves with the same α corresponding to electric fields that are odd functions of x_1 and magnetic fields that are even functions of x_1. As in the case of the Gaussian profile function the modes supported by this channel are dispersive, and for large k their frequencies tend to $\omega_p/\sqrt{2}$.

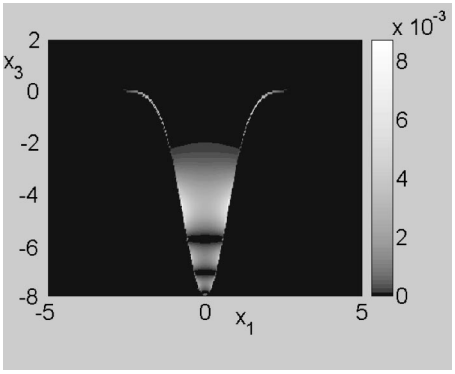

FIG. 7. The same as in Fig. 4, but with the frequency of the electric field given by $\omega/\omega_p=0.707$.

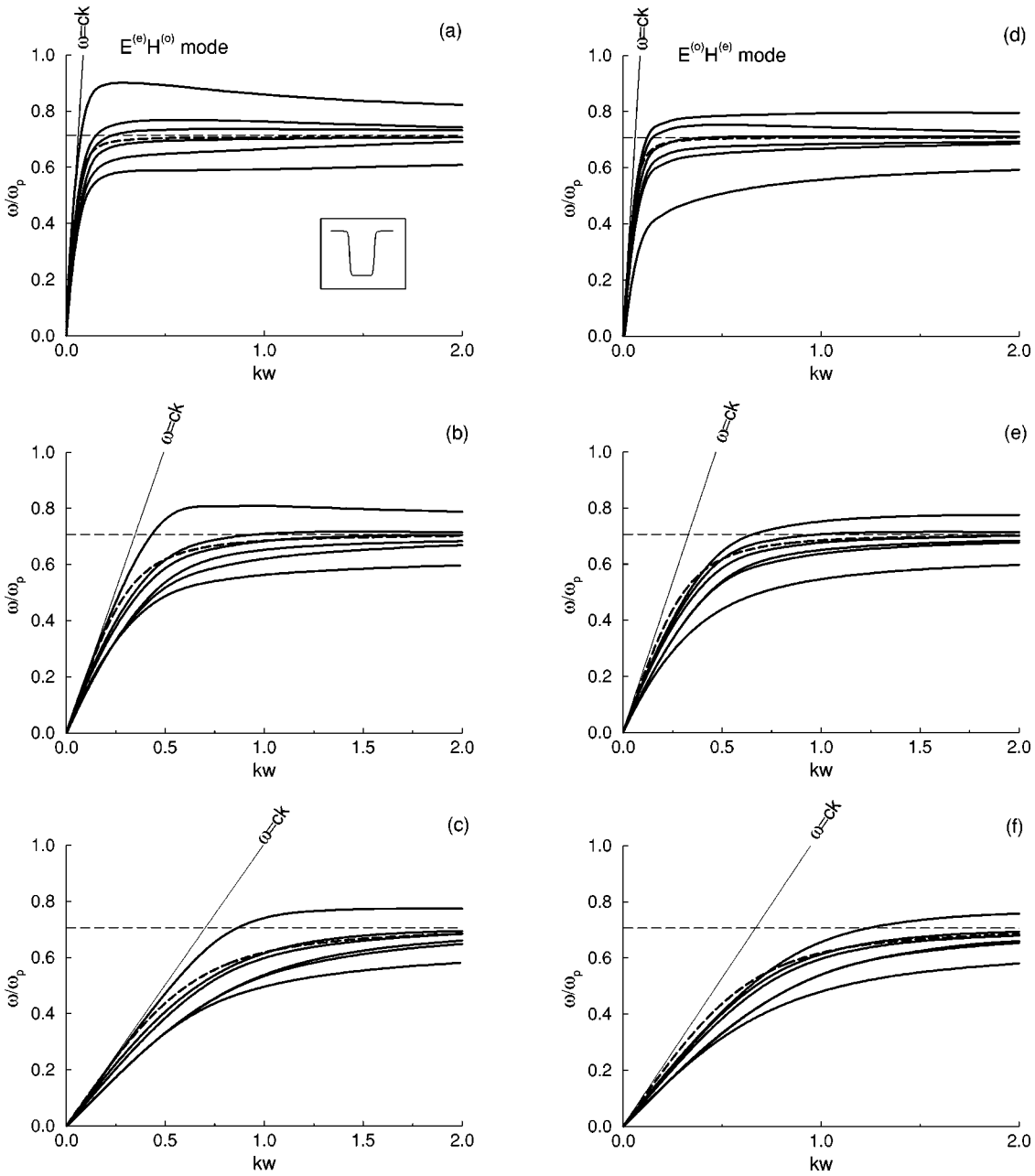

FIG. 8. The branches of the dispersion curve for the surface polaritons guided by a channel on a metallic substrate defined by the profile function $\zeta(x_1) = -2d\cosh^2(\beta w/4)[\cosh(\beta w/2) + \cosh(\beta x_1)]$, $d/w = 1$, $\beta w = 24$. The left side of the plot corresponds to the $E^{(e)}H^{(o)}$ modes, while the right side corresponds to the $E^{(o)}H^{(e)}$ modes. (a) and (d) $\alpha = 0.1$; (b) and (e) $\alpha = 0.5$; (c) and (f) $\alpha = 1.0$, where $\alpha = \omega_p w/c$. The dark dashed curve in each figure is the dispersion curve for surface plasmon polaritons at a planar vacuum-metal interface.

To illustrate the power of the method of the parametric representation of the surface profile we have also calculated the dispersion curves for channel polaritons guided by the truncated cylindrical channel depicted in Fig. 9. The radius of the cylindrical channel is R_0, and the angle subtended by its throat is $\theta_0 = 30°$. The resulting dispersion curves are plotted in Fig. 10 for the case where the dimensionless parameter $\alpha = \omega_p R_0/c$ is equal to unity. In Fig. 10(a) the second and third branches are degenerate on the scale of the figure, while in Fig. 10(b) the fourth and fifth branches are degenerate on the same scale.

VI. CONCLUSIONS

In this paper we have extended the definition of channel plasmons[5] to the electromagnetic case. We have called the resulting guided waves channel polaritons and have obtained their dispersion relation. Using the symmetry of the surface profile functions assumed, we have distinguished two different types of modes. The first type corresponds to electric fields that are even functions of x_1 and magnetic fields that are odd functions of x_1. The second type corresponds to electric fields that are odd functions of x_1 and magnetic fields that are even functions of x_1. We have made calculations for

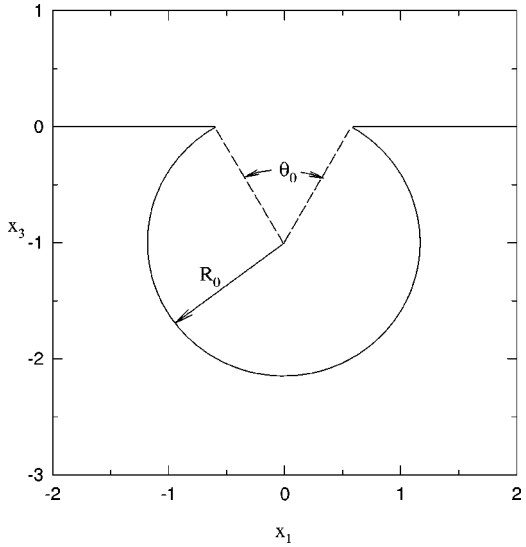

FIG. 9. A truncated cylindrical channel.

channels defined by a deep Gaussian profile function, a rectangular channel, and a truncated cylindrical channel. The frequencies of the modes supported by these channels are dispersive and are well separated from the frequency of a surface plasmon polariton on a planar vacuum-metal interface.

We have calculated the coordinate dependences of the fields of a channel polariton and have plotted two-dimensional charts of the squared modulus of the second component of the electric field inside the channel defined by a Gaussian profile. The results show that channel polaritons are well localized to the immediate vicinity of the channel.

Finally, we note that the present treatment has been based on the assumption that the dielectric function $\epsilon(\omega)$ of the material into which the channel is cut is real. This means that the mean free path of the channel polaritons is infinitely long. If the dielectric function is assumed to be complex, their mean free path will become finite due to their attenuation by Ohmic losses. To estimate what the mean free path of a typical channel polariton on a lossy metal surface might be, we have studied the case of a channel polariton on a silver surface whose wavelength is that of a He-Ne laser, $\lambda = 632.8$ nm. The dielectric constant of silver at this wavelength is $\epsilon(\omega) = -15.8836 + 1.075\,73i$, obtained by interpolation from the data of Palik.[10] The propagation constant of a surface plasmon polariton at a planar vacuum-silver interface at this frequency is obtained from the relation $\mathrm{Re}\,k(\omega) = (\omega/c)\mathrm{Re}\{\epsilon(\omega)/[\epsilon(\omega)+1]\}^{1/2}$ and has the value $\mathrm{Re}\,k(\omega) = 10.26\ \mu\mathrm{m}^{-1}$. The mean free path of this surface plasmon polariton is given by $l(\omega) = [2\,\mathrm{Im}\,k(\omega)]^{-1}$, where $\mathrm{Im}\,k(\omega) = (\omega/c)\mathrm{Im}\{\epsilon(\omega)/[\epsilon(\omega)+1]\}^{1/2}$, and has the value $l(\omega) = 21.54\ \mu$m. For the same value of the frequency, $\omega = 2.979\times 10^{15}\ \mathrm{s}^{-1}$, we searched for the complex roots of the dispersion relation for $E^{(e)}H^{(o)}$ modes obtained from Eqs. (4.15). The surface profile function was given by $\zeta(x_1) = -A\exp(-x_1^2/R^2)$, with $A/R = 8$. These calculations are difficult, and we cannot claim great accuracy for the imaginary parts of these roots, but they should have the correct order of magnitude. Two modes at this frequency had the complex wave numbers $kR = 0.3028 + 0.000\,495i$ and $kR = 0.3368 + 0.000\,776i$. If we require that the propagation constant of a channel polariton be larger than that of light in vacuum—i.e., that it should be in the nonradiative region $\mathrm{Re}\,k(\omega) > (\omega/c)$—the value of R must be smaller than $R \approx 0.03\ \mu$m. For a value of $R = 0.025\ \mu$m the propagation constants of these two channel polaritons are 12.11 and 13.47 $\mu\mathrm{m}^{-1}$, respectively, while their mean free paths are 25.25 and 16.11 μm, respectively. The latter are comparable to the mean free path of the surface plasmon polariton of the same frequency at a planar vacuum-silver interface. The reason that the mean free path of a channel polariton can be longer than that of a surface plasmon polariton on a planar surface would seem to be the strong localization of its electromagnetic field to the vicinity of the channel, which reduces the volume within which the dielectric losses occur, relative to that for a planar surface.

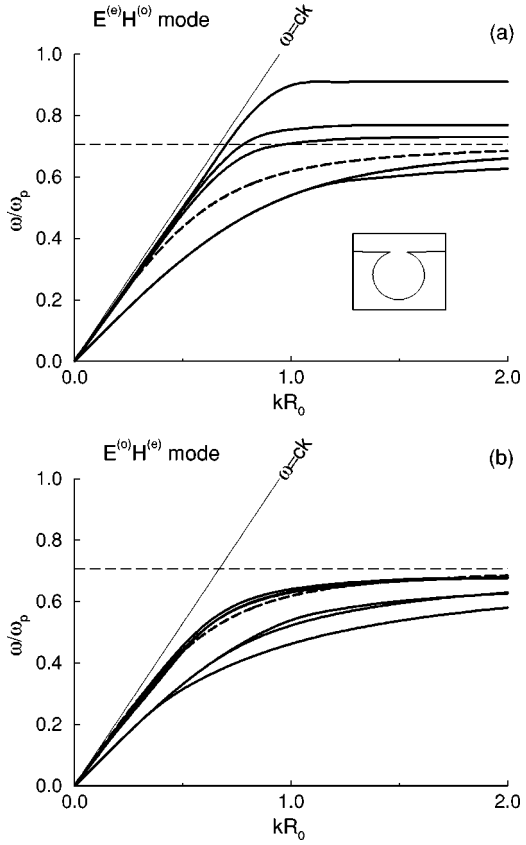

FIG. 10. The branches of the dispersion curve for the surface polaritons guided by a truncated cyclindrical channel. $\alpha = \omega_p R_0/c = 1$. The dark dashed curve in each figure is the dispersion curve for surface plasmon polaritons at a planar vacuum-metal interface.

ACKNOWLEDGMENT

This research was supported in part by Army Research Office Grant No. DAAD 19-99-1-0321.

*Stanford Research Systems, 1290 Ream Wood Ave., Sunnyvale, CA 94089.

[1] P. Berini, Phys. Rev. B **63**, 125417 (2001).

[2] S. I. Bozhevolnyi, J. Erland, K. Leosson, P. M. W. Skovgaard, and J. M. Hvam, Phys. Rev. Lett. **86**, 3008 (2001); S. I. Bozhevolnyi, V. S. Volkov, K. Leosson, and J. Erland, Opt. Lett. **26**, 734 (2001).

[3] J.-C. Weeber, J. R. Krenn, A. Dereux, B. Lamprecht, Y. Lacroute, and J. P. Goudonnet, Phys. Rev. B **64**, 045411 (2001).

[4] J.-C. Weeber, A. Dereux, C. Girard, J. R. Krenn, and J. P. Goudonnet, Phys. Rev. B **60**, 9061 (1999).

[5] Jun Q. Lu and A. A. Maradudin, Phys. Rev. B **42**, 11 159 (1990).

[6] A. Mendoza-Suárez and E. R. Méndez, Appl. Opt. **36**, 214 (1997).

[7] A. E. Danese, *Advanced Calculus* (Allyn and Bacon, Boston, 1965), Vol. I, p. 123.

[8] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1975), p. 341.

[9] A. A. Maradudin, T. R. Michel, A. R. McGurn, and E. R. Méndez, Ann. Phys. (N.Y.) **203**, 255 (1990).

[10] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, New York, 1985).

Comment on "Radiative transfer over small distances from a heated metal"

A. A. Maradudin

Department of Physics and Astronomy and Institute for Surface and Interface Science, University of California, Irvine, Irvine, California 92697

Received April 24, 2000

The chief result obtained in Ref. 1, namely, that, when the width l of the gap between a metal and a lossless dielectric shrinks to zero, the coefficient of heat transfer from the heated metal to the colder dielectric is finite, is correct. However, this result was already obtained many years ago by Levin *et al.*[2] They also showed that, if the dielectric is lossy, the heat transfer diverges as l^{-2} as $l \to 0$. Consequently, the criticism in Ref. 1 of earlier work, in particular, that of Loomis and Maris,[3] for not yielding a finite heat transfer in this limit is unjustified, because in Ref. 3 the dielectric medium was assumed to be lossy. In fact, the results reported in Ref. 3 are consistent with those of Levin *et al.*

References

1. J. L. Pan, Opt. Lett. **25,** 369 (2000).
2. M. L. Levin, V. G. Polevoi, and S. M. Rytov, Sov. Phys. JETP **52,** 1054 (1980).
3. J. J. Loomis and H. J. Maris, Phys. Rev. B **50,** 18517 (1994).

OCIS codes: 030.5620, 260.3060, 130.5990.

0146-9592/01/080479-01$15.00/0 © 2001 Optical Society of America