Optical response of microscopically rough surfaces

D. E. Aspnes
Bell Communications Research, Inc. (Bellcore), P.O. Box 7040, Red Bank, New Jersey 07701-7040
(Received 7 August 1989; revised manuscript received 27 December 1989)

The effect of microscopic roughness on the complex normal-incidence reflectance $\tilde{r}$ at an interface between a transparent ambient and a microscopically rough substrate is treated analytically in the classical limit to lowest order in d/λ, where d is the effective thickness of the microscopically rough layer and λ is the wavelength of light. The ambient and substrate dielectric functions, ϵ_a and ϵ_s, respectively, are assumed to be local, scalar, and isotropic. The mathematical definition of microscopic roughness is shown to include the distance to the observer as well as λ and the spatial Fourier components of the surface, which is characteristic of coherent (Fraunhofer) diffraction. The effect on $\tilde{r}$, expressed as a dielectric response of an equivalent thin film, reduces to standard effective-medium form, thereby providing a theoretical justification of the highly successful empirical approach. This expression further reduces to a surface integral of the local-field potential weighted by the component of the local normal surface vector parallel to the mean surface plane, a form that allows the microscopic-roughness-induced change $\Delta\tilde{r}/\tilde{r}$ in $\tilde{r}$ to be calculated directly if the local potential is known by symmetry, numerical analysis, conformal mapping, or other means. If $|\epsilon_s|$ is large, as for semiconductors and metals, and if self-consistency effects can be ignored, then the problem becomes isomorphic to that of fluid flow and can be solved analytically for simple geometries by conformal mapping. In this limit, I obtain analytic solutions for the weak sinusoidal grating, the low ridge, and the isolated low step, all of which are incipiently rough surfaces that cannot be treated by effective-medium theory but which represent morphologies commonly encountered in crystal growth. The expressions for the ridge and step contain singularities that are logarithmic in the ratio of step height to separation, although these singularities are expected to be limited by self-consistency effects that remain to be established numerically. The present results also provide a theoretical framework for more general treatments.

I. INTRODUCTION

The mathematically flat surfaces assumed in the derivation of the Fresnel equations[1] are fictitious, as surfaces are always rough to some degree. From an optical perspective, the most obvious manifestation of surface roughness is the nonspecular scattering of light that occurs when the length scale of the irregularities is comparable to or larger than λ, the wavelength of light.[2–4] Because nonspecularly scattered light arises entirely from deviations from ideality, it is easy to recognize and to investigate quantitatively. As a result, the theory of optical scattering from macroscopically rough surfaces has been developed extensively.[2–28]

Yet irregularities also exist on length scales much less than λ. But because microscopic roughness does not scatter appreciable light out of the specular beam, its existence is not self-evident. Instead, microscopic roughness acts primarily as an antireflection coating, reducing the amount of specularly reflected light by softening the dielectric discontinuity between sample and ambient.[29] Because this affects only our perception of the dielectric properties of the sample, which may not be accurately known in the first place, it is not surprising that the recognition[30] of microscopic roughness and the realization[30] that it could be described by effective-medium theory[31,32] is relatively recent. The empirical verification of these ideas had to await the development of spectroellipsometry and is even more recent.[29,33–43] It is now well established that microscopic roughness is endemic, and that its optical effect can be described accurately by the so-called three-phase (substrate–thin-film–ambient) model where the microscopically rough surface is represented as a thin film whose dielectric response is an effective-medium blend of the dielectric responses of substrate and ambient.

Despite its ubiquity and importance, microscopic roughness has received almost no first-principles theoretical attention. To my knowledge, the single exception is the recent discussion of Mochán and Barrera[44] of nearly flat surfaces with random irregularities, treated as an extension of their theory[45] of the electromagnetic response of systems with spatial fluctuations. The present work was motivated by our real-time studies[46–50] of semiconductor crystal growth, which use a normal-incidence optical-probe, reflectance-difference spectroscopy[51,52] (RDS), to follow surface conditions to better than 0.01 monolayer. Because growth surfaces are typically covered by large densities of biatomic steps, we needed to establish, analytically if possible, the approximate dielectric responses of these morphological features. The present paper addresses this problem.

Beginning with Green's-function solutions to Maxwell's equations[1,53] for a classical system consisting of a substrate and a transparent ambient with local, isotropic, and scalar dielectric responses, and using vector calculus to eliminate detail that does not contribute to the result, I obtain new expressions that describe the effect of microscopic roughness on the complex normal-incidence reflectance $\tilde{r}$. These expressions provide insights that were previously lacking, and with some approximations can be solved analytically for simple morphologies. For example, I show that the mathematical definition of microscopic roughness must include the distance between the surface and the observer as well as λ and the spatial Fourier coefficients of the rough surface. This is a general characteristic of Fraunhofer diffraction,[1] and illustrates the coherent (cooperative, self-consistent, "many-body") nature of optical scattering from microscopically rough surfaces. The effective-medium formalism is also obtained, thereby providing the theoretical justification of this highly successful empirical approach.

The effective-medium expression is transformed into a surface integral of the local-field potential weighted by the component of the local surface-normal vector parallel to the mean surface plane. The derivation establishes the direct link between the relative roughness-induced change $\Delta\tilde{r}/\tilde{r}$ in $\tilde{r}$ and the screening or depolarization charge that develops as a result of microscopic roughness, and also establishes a natural reference plane (the equivalent smooth surface) against which to measure these changes. For example, with respect to this reference plane $\Delta\tilde{r}/\tilde{r}=0$ for any geometry for which the applied field is everywhere locally parallel to the surface, as, for example, for normally incident light polarized parallel to unidirectional steps on a surface. This symmetry has not been recognized previously. This transformation also allows $\Delta\tilde{r}/\tilde{r}$ to be calculated to first order if the local potential has already been determined by symmetry, conformal mapping,[54,55] or numerical analysis. It also provides a means of expanding $\Delta\tilde{r}/\tilde{r}$ as a power series in $\epsilon_a/|\epsilon_s|$, where ϵ_a and ϵ_s are the dielectric functions of ambient and substrate, respectively. This expansion is appropriate to metals and semiconductors in the visible and near ultraviolet where $|\epsilon_s| \gg \epsilon_a$.

If $\epsilon_a/|\epsilon_s| \ll 1$ and self-consistency effects are negligible, then $\Delta\tilde{r}/\tilde{r}$ can be expressed simply in terms of properties of the conformal map that generates a given roughness geometry. With this approach I obtain approximate analytic expressions for $\Delta\tilde{r}/\tilde{r}$ for the low grating, the symmetric low ridge, and the isolated low step. None of these configurations can be treated by the effective-medium approach. All yield a bilinear dependence of $\Delta\tilde{r}/\tilde{r}$ on the peak-to-valley amplitude d and the mean slope d/L of the particular morphological features, where L is the distance between features. The symmetric low ridge and isolated low step show an additional weak logarithmic divergence $\ln(L/d)$ as a result of field distortions near the step, although this divergence must saturate with increasing L owing to screening-charge self-consistency effects not included in the calculation.

The initial part of the present approach most closely parallels the effective-surface-current treatments of Rice,[56] Kröger and Kretschmann,[57,58] and Juranek.[59] I take advantage of the formal solutions of Maxwell's equations,[1,53] which embody the general properties shared by all configurations, to establish general properties of a specific class of configurations, here interfaces that are nonideal due to microscopic roughness, as a perturbation of the ideal two-phase substrate-ambient system. While Refs. 56–59 also deal with nonideal interfaces, neglect of local-field effects makes their results inapplicable to the present problem. I also mention the work of Bagchi *et al.*,[60] who explored the physics of the two-phase system with a nonlocal ϵ_s. The work of Mochán and Barrera[44,45] is qualitatively different, being directed toward the calculation of the macroscopic dielectric response of a spatially inhomogeneous system as a function of local fluctuations of the dielectric response. Possibly the most comprehensive treatment of surface roughness within the framework of formal solutions is that of Arya and Zeyher,[23] although their one- and two-photon Green's-function treatment is designed for numerical analysis. The random-diffraction-grating or Rayleigh-Fano approach,[5,6] where roughness is treated by expanding the height function $z(x,y)$ of the surface as a two-dimensional Fourier transform whose components are then treated independently, is adequate for describing the incoherent scattering from macroscopic irregularities[5,6,15,16,19–21,23,26] but is not suited for describing the coherent scattering from microscopic roughness. In any case Rayleigh-Fano calculations are arcane and provide little or no insight into ranges of validity or even into such elementary properties as spectral or polarization[61] dependences of the scattered light.

II. GENERAL FORMALISM

Figure 1 illustrates the configuration, a continuous rough surface $S(\mathbf{r})=S(x,y,z)=0$ separating the substrate s from the ambient a near the plane $z=0$, and a normally incident vector-potential wave $\mathbf{A}_i(\mathbf{r})$ generating reflected and transmitted waves $\mathbf{A}_r(\mathbf{r})$ and $\mathbf{A}_t(\mathbf{r})$, respec-

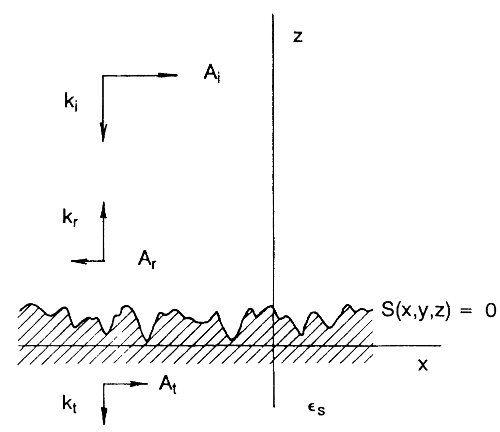

FIG. 1. Schematic diagram of a cross section of a microscopically rough surface separating an ambient a and a substrate s. The surface is defined by $S(x,y,z)=0$. As drawn, the reference plane $z=0$ lies below the real surface and that of the equivalent smooth sample, which is obtained by smoothing out the asperities. The incident, transmitted, and reflected vector potentials are marked.

tively. The substrate and ambient are assumed to be homogeneous, isotropic, and local, the ambient to be transparent $[\mathrm{Im}(\epsilon_a)=0]$, and the substrate to be at least infinitesimally absorbing $[\mathrm{Im}(\epsilon_s)\neq 0]$. The location of the reference plane $z=0$ is arbitrary, although a correction term vanishes if it is chosen to coincide with the equivalent smooth surface formed by averaging out all irregularities. The zeroth-order solution, which describes normal-incidence reflectance at an ideal boundary, is

$$\mathbf{A}_{i0}(\mathbf{r})=\hat{\mathbf{x}}A_{i0}e^{-ik_a z-i\omega t}, \quad z>0 \tag{1a}$$

$$\mathbf{A}_{r0}(\mathbf{r})=\hat{\mathbf{x}}A_{r0}e^{ik_a z-i\omega t}, \quad z>0 \tag{1b}$$

$$\mathbf{A}_{t0}(\mathbf{r})=\hat{\mathbf{x}}A_{t0}e^{-ik_s z-i\omega t}, \quad z<0 \tag{1c}$$

where A_{i0}, A_{r0}, and A_{t0} are complex amplitude coefficients, the complex wave-vector amplitudes k_a and k_s are defined by $\epsilon_a=c^2k_a^2/\omega^2$ and $\epsilon_s=c^2k_s^2/\omega^2$, respectively, and $\omega=2\pi c/\lambda$. The boundary conditions on continuous tangential electric and magnetic fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{H}(\mathbf{r})$, respectively, lead to the Fresnel relations[1,53]

$$A_{r0}=\frac{k_a-k_s}{k_a+k_s}A_{i0}\ , \tag{2a}$$

$$A_{t0}=\frac{2k_a}{k_a+k_s}A_{i0}\ . \tag{2b}$$

I formulate the rough-surface problem with the constitutive equation $\mathbf{D}(\mathbf{r})=\epsilon(\mathbf{r})\mathbf{E}(\mathbf{r})$, considering ϵ to be spatially dependent. Even though the spatial dependence of ϵ is essentially trivial, it is advantageous for analytic reasons to retain it. In the Coulomb gauge $\nabla\cdot\mathbf{A}(\mathbf{r})=0$, the electric and magnetic fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{H}(\mathbf{r})$ of the differential form of Maxwell's equations are replaced by the vector and scalar potentials $\mathbf{A}(\mathbf{r})$ and $\phi(\mathbf{r})$ according to

$$\mathbf{H}(\mathbf{r})=\nabla\times\mathbf{A}(\mathbf{r})\ , \tag{3a}$$

$$\mathbf{E}(\mathbf{r})=(i\omega/c)\mathbf{A}(\mathbf{r})-\nabla\phi(\mathbf{r})\ , \tag{3b}$$

in which case Maxwell's equations become

$$0=[(i\omega/c)\mathbf{A}(\mathbf{r})-\nabla\phi(\mathbf{r})]\cdot\nabla\epsilon(\mathbf{r})-\epsilon(\mathbf{r})\nabla^2\phi(\mathbf{r})\ , \tag{4a}$$

$$\left[-\nabla^2-\frac{\omega^2}{c^2}\epsilon(\mathbf{r})\right]\mathbf{A}(\mathbf{r})=\frac{i\omega}{c}\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})\ . \tag{4b}$$

Although, in general, (4) must be solved self-consistently, both (4a) and (4b) are readily expanded in powers of d/λ, where d is a characteristic thickness of the microscopically rough region. To lowest order (4a) describes a local-field response $-\nabla\phi(\mathbf{r})$ to a driving term $(i\omega/c)\mathbf{A}(\mathbf{r})$ that can be substituted into (4b) to obtain the first-order correction to $\mathbf{A}_{r0}(\mathbf{r})$.

I first consider the determination of $-\nabla\phi(\mathbf{r})$. Because $\nabla\cdot\mathbf{A}(\mathbf{r})=0$ and $\epsilon(\mathbf{r})$ are spatially invariant within substrate and ambient, (4a) reduces to

$$\nabla^2\phi(\mathbf{r})=0, \quad \mathbf{r}\text{ not on }S \tag{5a}$$

$$\text{normal component of }\epsilon(\mathbf{r})\frac{i\omega}{c}\mathbf{A}(\mathbf{r})-\nabla\phi(\mathbf{r})\text{ is conserved across }S\ . \tag{5b}$$

Equation (5a) implies that

$$\phi(\mathbf{r})\approx\phi_l(\mathbf{r})=\int_S d^2r_s'\frac{\sigma(\mathbf{r}_s')}{|\mathbf{r}-\mathbf{r}_s'|}-\mathbf{E}_0\cdot\mathbf{r}+\phi_0\ , \tag{6a}$$

where $\sigma(\mathbf{r}_s)$ is the surface charge density on S at $\mathbf{r}=\mathbf{r}_s$. The integration constants $\mathbf{E}_0$ and ϕ_0 are eliminated immediately because $\phi(\mathbf{r})$ must vanish for $z\to-\infty$, as s is assumed to be weakly absorbing. The subscript l denotes that $\phi_l(\mathbf{r})$ is a *local* potential arising from the depolarization charge $\sigma(\mathbf{r})$ that develops as a result of microscopic roughness; justification that $\phi_l(\mathbf{r})$ is indeed local follows in a later paragraph. The local and total electric fields are therefore

$$\mathbf{E}_l(\mathbf{r})=-\nabla\phi_l(\mathbf{r})=\int_S d^2r_s'\frac{\sigma(\mathbf{r}_s')(\mathbf{r}-\mathbf{r}_s')}{|\mathbf{r}-\mathbf{r}'|^3}\ , \tag{6b}$$

$$\mathbf{E}(\mathbf{r})=\frac{i\omega}{c}\mathbf{A}(\mathbf{r})+\mathbf{E}_l(\mathbf{r})\ . \tag{6c}$$

Because (6b) is general, it must also be valid even for points infinitesimally close to S. Taking $\mathbf{r}=\mathbf{r}_s+\delta\hat{\mathbf{n}}(\mathbf{r}_s)$, where the surface-normal vector $\hat{\mathbf{n}}(\mathbf{r}_s)=\nabla S(\mathbf{r}_s)/|\nabla S(\mathbf{r}_s)|$ points from s to a, expanding to second order in δ, invoking (5b), and evaluating the result for $\delta\to 0$ leads to

$$\sigma(\mathbf{r}_s)=\frac{\epsilon_s-\epsilon_a}{2\pi(\epsilon_s+\epsilon_a)}\hat{\mathbf{n}}(\mathbf{r}_s)\cdot\left[\frac{i\omega}{c}\mathbf{A}(\mathbf{r}_s)+\int_{\bar{S}} d^2r_s'\frac{\sigma(\mathbf{r}_s')(\mathbf{r}_s-\mathbf{r}_s')}{|\mathbf{r}_s-\mathbf{r}_s'|^3}\right]\ , \tag{7}$$

where the integral spans S except for an infinitesimally small circle about $\mathbf{r}_s$. The locality of $\sigma(\mathbf{r}_s)$, expressed as the vanishing of the average of $\sigma(\mathbf{r}_s)$ over a large enough area, is easily verified by applying the divergence theorem to the integral of (4a) over a volume of substrate bounded on one side by S. This vanishing removes the inherent ambiguity of (7), which specifies how $\sigma(\mathbf{r}_s)$ must change at the medium-ambient discontinuity, but does not place a constraint on its cumulative value. A similar ambiguity in $\mathbf{E}_l(\mathbf{r})$ is removed because the line integral of $\phi_l(\mathbf{r})$ must also vanish on the average.

Equation (7) establishes $\sigma(\mathbf{r}_s)$ as a screening or surface polarization charge whose existence depends on the existences of a dielectric mismatch between substrate and ambient and of components of $\mathbf{A}(\mathbf{r}_s)$ locally perpendicular to S. Thus $\sigma(\mathbf{r}_s)$ vanishes even for microscopically rough surfaces if $\epsilon_s=\epsilon_a$, which may happen at certain wavelengths, or if $\mathbf{A}(\mathbf{r})$ is everywhere parallel to the boundary, as for a unidirectionally stepped surface illuminated by light polarized along the steps.

I now consider the determination of $\mathbf{A}(\mathbf{r})$. The left-hand side of (4b) reduces to its zeroth-order form if the term involving $\epsilon(\mathbf{r})$ is moved to the right-hand side and if $(\omega^2/c^2)\epsilon_a$ or $(\omega^2/c^2)\epsilon_s$ is subtracted from both sides according to whether z is greater or less than 0:

$$\left[-\nabla^2-\frac{\omega^2}{c^2}\epsilon_a\right]\mathbf{A}(\mathbf{r})=\frac{\omega^2}{c^2}[\epsilon(\mathbf{r})-\epsilon_a]\mathbf{A}(\mathbf{r})+\frac{i\omega}{c}\epsilon(\mathbf{r})\nabla\phi_l(\mathbf{r}) \quad (8a)$$

$$=\frac{4\pi}{c}\mathbf{j}_+(\mathbf{r}),\quad z>0 \quad (8b)$$

$$\left[-\nabla^2-\frac{\omega^2}{c^2}\epsilon_s\right]\mathbf{A}(\mathbf{r})=-\frac{\omega^2}{c^2}[\epsilon_s-\epsilon(\mathbf{r})]\mathbf{A}(\mathbf{r})+\frac{i\omega}{c}\epsilon(\mathbf{r})\nabla\phi_l(\mathbf{r}) \quad (8c)$$

$$=\frac{4\pi}{c}\mathbf{j}_-(\mathbf{r}),\quad z<0\ . \quad (8d)$$

Clearly, the extra polarizability that results from substrate material protruding beyond $z=0$ and the loss of bulk polarizability that results from ambient material penetrating below $z=0$ are completely equivalent to the local screening charge $\nabla\phi_l(\mathbf{r})$ from the perspective of the effective source currents $\mathbf{j}_+(\mathbf{r})$ and $\mathbf{j}_-(\mathbf{r})$.

If the other half-space did not exist, the solution of (8) could be written formally in each half-space using the Green's function of the Helmholtz equation:[53]

$$\mathbf{A}(\mathbf{r})=\mathbf{A}_0(\mathbf{r})+\frac{1}{c}\int d^3r'\frac{e^{ik_a|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\mathbf{j}_+(\mathbf{r}'),\quad z,z'>0 \quad (9a)$$

$$\mathbf{A}(\mathbf{r})=\mathbf{A}_0(\mathbf{r})+\frac{1}{c}\int d^3r'\frac{e^{ik_s|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\mathbf{j}_-(\mathbf{r}'),\quad z,z'<0 \quad (9b)$$

where $\mathbf{A}_0(\mathbf{r})$ is a solution of the homogeneous equation. When reflectance at the ambient-substrate interface is taken into account, Eqs. (9) become self-consistent expressions for $\mathbf{A}(\mathbf{r})$. But if $\mathbf{j}_+(\mathbf{r}')$ and $\mathbf{j}_-(\mathbf{r}')$ are nonzero only in a narrow region near $z=0$, then $\mathbf{A}(\mathbf{r})$ in (7) and on the right-hand sides of (8) can be replaced by $\mathbf{A}_0(\mathbf{r})$, yielding the leading term of a perturbation expansion that describes the effect of the rough surface on $\tilde{r}$. In this first Born approximation, (1) and (2) also simplify to yield

$$\mathbf{A}_0(\mathbf{r})\approx\hat{\mathbf{x}}A_{i0}\frac{2k_a}{k_a+k_s},\quad z\approx 0\ . \quad (10)$$

Retardation effects (exponential dependences involving z) are neglected in this approximation.

I next calculate the effect of the perturbation term of (9a) on the reflected component of $\mathbf{A}_r(\mathbf{r})$; (9b) can be treated similarly. With the observer at large positive z, (9a) becomes approximately

$$\lim_{z\to\infty}\mathbf{A}(\mathbf{r})\approx\mathbf{A}_0(\mathbf{r})+\frac{2k_a}{zc(k_a+k_s)}e^{ik_az}\int_0^\infty dz'\int dx'\int dy'\,e^{ik_a(x'^2+y'^2)/2z}\mathbf{j}_+(\mathbf{r}')\ , \quad (11)$$

where retardation effects are again neglected, except in the prefactor. Equation (11) includes both the directly backscattered and the backreflected forward-scattered contributions from the kernel of (9a). The upper limit of the z' integration is removed to infinity because $\mathbf{j}_+(\mathbf{r})$ vanishes rapidly with increasing z. With (1) and (2) I obtain the relative first-order correction directly:

$$\lim_{z\to\infty}A_r(\mathbf{r})=A_{r0}(z)\left[\hat{\mathbf{x}}+\frac{2k_a}{A_{i0}zc(k_a-k_s)}\int_0^\infty dz'\int dx'\int dy'\,e^{ik_a(x'^2+y'^2)/2z}\mathbf{j}_+(\mathbf{r}')\right]\ , \quad (12a)$$

where

$$A_{r0}(z)=\frac{k_a-k_s}{k_a+k_s}A_{i0}e^{ik_az}\ . \quad (12b)$$

To proceed further, I express the x and y dependences of $\mathbf{j}_+(\mathbf{r})$ as the two-dimensional Fourier transform

$$\mathbf{j}_+(\mathbf{r})=\int dq_x\int dq_y\,C_+(q_x,q_y;z)e^{iq_xx+iq_yy}\ , \quad (13)$$

which when substituted into (12) yields an integral over x and y that can be expressed in closed form:

$$\lim_{z\to\infty}\mathbf{A}_r(\mathbf{r})=A_{r0}(z)\left[\hat{\mathbf{x}}+\frac{4\pi i}{A_{i0}c(k_a-k_s)}\int_0^\infty dz'\int_{-\infty}^\infty dq_x\int_{-\infty}^\infty dq_y C_+(q_x,q_y;z')e^{-iz(q_x^2+q_y^2)/2k_a}\right]\ . \quad (14)$$

The definition of a microscopically rough surface and its connection to Fraunhofer diffraction are now clear. If all coefficients $C_+(q_x,q_y)$ are zero unless $|\mathbf{q}|=0$ or $|\mathbf{q}|$ is large enough to scatter light out the field of the observer [$|\mathbf{q}|\geq k_aD/(2z)$, where D is the diameter of the illuminated spot and z is the distance between surface and observer], then the backscattered wave at z will be a coherent superposition of contributions arising from a small region of approximate radius $(2z/k_a)^{1/2}=(z\lambda/\pi)^{1/2}$ about the point of stationary phase $(x',y')=(0,0)$. In this case I can represent $\mathbf{j}_+(\mathbf{r})$ by its zero-$|\mathbf{q}|$ component:

$$\mathbf{j}_+(\mathbf{r})=\langle \mathbf{j}_+(z)\rangle=\frac{1}{\Omega}\int_\Omega dx\,dy\,\mathbf{j}_+(\mathbf{r})\ , \tag{15}$$

where Ω is the projection of the illuminated part of S onto the x-y plane. Then (12a) becomes

$$\lim_{z\to\infty} A_r(\mathbf{r})=A_{r0}(z)\left[\hat{\mathbf{x}}+\frac{4\pi i}{A_{i0}c(k_a-k_s)\Omega}\int_\Omega dx'\,dy'\int_0^\infty dz'\,\mathbf{j}_+(\mathbf{r}')\right]\ . \tag{16}$$

A parallel development for $\mathbf{j}_-(\mathbf{r})$ leads to an analogous expression, which when combined with (16) yields

$$\lim_{z\to\infty} \mathbf{A}_r(\mathbf{r})=A_{r0}(z)\left[\hat{\mathbf{x}}+\frac{1}{(n_a-n_s)A_{i0}\Omega}\int_\Omega dx'\,dy'\int_{-\infty}^\infty dz'\left[\frac{i\omega}{c}\mathbf{A}_0(\mathbf{r}')[\epsilon(\mathbf{r}')-\epsilon_a u(z')-\epsilon_s u(-z')]-\epsilon(\mathbf{r}')\nabla\phi_l(\mathbf{r}')\right]\right]\ , \tag{17}$$

where $n_a=ck_a/\omega$ and $n_s=ck_s/\omega$. The integral over z' converges because the expression in large square brackets rapidly approaches zero away from the reference plane.

Equation (17) can be transformed into effective-medium form by noting that the multiplier of $\epsilon(\mathbf{r})$ is just the total electric field (6c). Since, for $z\approx 0$, $\langle \mathbf{E}(\mathbf{r})\rangle=\hat{\mathbf{x}}E_0$, where $\langle f(\mathbf{r})\rangle$ denotes the volume average of $f(\mathbf{r})$, and since

$$\mathbf{E}_0(\mathbf{r})\approx\hat{\mathbf{x}}E_0=\frac{i\omega}{c}\mathbf{A}_0(\mathbf{r})=\hat{\mathbf{x}}A_{i0}\frac{2i\omega k_a}{c(k_a+k_s)}\ , \tag{18}$$

it follows that

$$\lim_{z\to\infty} \mathbf{A}_r(\mathbf{r})=A_{r0}(z)\left[\hat{\mathbf{x}}+\frac{4\pi i n_a}{\lambda(\epsilon_s-\epsilon_a)}\int_{-\infty}^\infty dz\left[\hat{\mathbf{x}}[\epsilon_s u(-z)+\epsilon_a u(z)]-\frac{1}{\Omega E_0}\int_\Omega dx\,dy\,\epsilon(\mathbf{r})\mathbf{E}(\mathbf{r})\right]\right]\ . \tag{19}$$

Equation (19) is precisely of effective-medium form, describing the microscopic-roughness correction as a ratio of a macroscopic displacement field $\mathbf{D}=\langle\epsilon(\mathbf{r})\mathbf{E}(\mathbf{r})\rangle$ to the macroscopic electric field $\langle\mathbf{E}(\mathbf{r})\rangle$. The volume integral of (19) can be transformed into a simple surface integral. Writing the volume integral as

$$I=\int_{-D}^D dz\left[\hat{\mathbf{x}}[\epsilon_s u(-z)+\epsilon_a u(z)]-\frac{1}{\Omega E_0}\int_\Omega dx\,dy\,\epsilon(\mathbf{r})\mathbf{E}(\mathbf{r})\right] \tag{20a}$$

$$=-\hat{\mathbf{x}}D(\epsilon_s-\epsilon_a)(f_s-f_a)+\frac{1}{\Omega E_0}\int_V d^3r\,\epsilon(\mathbf{r})\nabla\phi_l(\mathbf{r})\ , \tag{20b}$$

where D is arbitrarily large, I take advantage of the spatial invariance of $\epsilon(\mathbf{r})$ within substrate and ambient to break (20) into two parts and to convert each part into a surface integral over the respective volumes. The only surviving contribution comes from the interface, leading to

$$I=(\epsilon_s-\epsilon_a)\left[-\hat{\mathbf{x}}D(f_s-f_a)+\frac{1}{\Omega E_0}\int_S d^2r_s'\,\hat{\mathbf{n}}(\mathbf{r}_s)\phi_l(\mathbf{r}_s')\right]\ , \tag{20c}$$

where $\hat{\mathbf{n}}(\mathbf{r}_s)$ is the unit outward normal at $\mathbf{r}_s$ that points from s to a. With (20c), (19) becomes

$$\lim_{z\to\infty} \mathbf{A}_r(\mathbf{r})=A_{r0}(z)\left[\hat{\mathbf{x}}-\frac{4\pi i n_a}{\lambda}\left[\hat{\mathbf{x}}\,\Delta z-\frac{1}{E_0\Omega}\int_S d^2r_s\,\hat{\mathbf{n}}(\mathbf{r}_s)\phi(\mathbf{r}_s)\right]\right]\ , \tag{21}$$

where the "correction term" $\Delta z=D(f_s-f_a)$ describes the phase shift accumulated by propagation between the equivalent smooth surface and the selected $z=0$ reference plane, in effect moving the arbitrarily chosen laboratory reference plane to the equivalent smooth surface. If the reference plane were already chosen to be the equivalent smooth surface, and if the x, or polarization direction, were a plane of symmetry (the usual effective-medium assumption), then $\Delta z=0$ and the x projection can be taken throughout, leading to the simple result

$$\lim_{z\to\infty} \mathbf{A}_r(\mathbf{r})=\hat{\mathbf{x}}A_{r0}(z)\left[1+\frac{4\pi i n_a}{\lambda E_0\Omega}\int_S d^2r_s\,\hat{\mathbf{x}}\cdot\hat{\mathbf{n}}(\mathbf{r}_s)\phi_l(\mathbf{r}_s)\right]\ , \tag{22a}$$

where Ω is the projection of S onto the x-y plane. In a more common notation, the correction term can be written alone as

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{4\pi i n_a}{\lambda E_0\Omega}\int_S d^2r_s\,\hat{\mathbf{x}}\cdot\hat{\mathbf{n}}(\mathbf{r}_s)\phi_l(\mathbf{r}_s)\ . \tag{22b}$$

Thus, to first order the effect of microscopic roughness is described as a surface integral of the local potential weighted by the component of the local surface-normal vector parallel to the mean surface plane.

To my knowledge, this formulation of the problem is new. Several symmetries are clearly incorporated, such as that discussed above concerning orthogonal polarization and local surface-normal vectors. Two other general properties follow from (6a) and (7), which show that $\phi_l(\mathbf{r})$ is real if either ϵ_s is real or $|\epsilon_s|\to\infty$. In either case, $\Delta\tilde{r}/\tilde{r}$ is purely imaginary. Since the relative change $\Delta R/R$ in the power reflectance R is given by $\Delta R/R=2\,\mathrm{Re}(\Delta\tilde{r}/\tilde{r})$, it follows that microscopic roughness has no first-order effect on R if either ϵ_s is real or $\epsilon_s\to\infty$. The former reproduces a well-known property of transparent films on transparent substrates. The latter shows that the effect of microscopic roughness on R reduces as ϵ_s increases. Equations (22) also allow the effect of microscopic roughness to be calculated if the local potential is already determined by symmetry, conformal mapping, numerical analysis, or other means. In addition, (6a), (7), and (22) provide a natural basis for analyzing microscopic roughness as a power series in $\epsilon_a/|\epsilon_s|$, appropriate for metals and semiconductors in the visible and near-ultraviolet spectral regions.

Before discussing applications, it is useful to make contact with previous treatments of surface dielectric responses where the dielectric function $\epsilon(\mathbf{r})$ is taken to be z dependent and where the $z=0$ reference plane is assigned to various locations. For $\epsilon(\mathbf{r})=\epsilon(z)$, (17) can be rewritten as

$$\frac{\Delta\tilde{r}}{\tilde{r}}=-\frac{4\pi i n_a}{\lambda(\epsilon_s-\epsilon_a)}\int_{-\infty}^{\infty}dz'[\epsilon(z')-\epsilon_a u(z')-\epsilon_s u(-z)]\ , \tag{23}$$

taking advantage of (18) and (6b) and (7) to eliminate $\phi_l(\mathbf{r})$. The generic form (23) has been previously derived, although by different aproaches, by Plieth and Naegele[62] as a generalization of the three-phase model,[63] and by Bagchi *et al.*[60] and Feibelman[64] for systems where the dielectric response is nonlocal. For a laminar film of dielectric constant ϵ_o and thickness $d\ll\lambda$ deposited on s, (23) becomes

$$\frac{\Delta\tilde{r}}{\tilde{r}}=-\frac{4\pi i d n_a}{\lambda}\frac{\epsilon_o-\epsilon_a}{\epsilon_s-\epsilon_a}\ , \tag{24a}$$

if the reference plane is located at the substrate-film boundary and

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{4\pi i d n_a}{\lambda}\frac{\epsilon_s-\epsilon_o}{\epsilon_s-\epsilon_a} \tag{24b}$$

if the reference plane is located at the film-ambient boundary. Since the two configurations differ only by a rigid shift of the substrate by $\Delta z=d$, it is not surprising that (24a) and (24b) differ only by the extra phase retardation $4\pi i d n_a/\lambda$ accumulated by the zeroth-order wave propagating this extra distance in the latter case. As absolute phase differences are of no moment whatsoever in photometry, these two "three-phase-model" expressions have been used interchangeably in the literature. However, in comparative measurements, such as reflectance-difference spectroscopy, this distinction is important.

By comparing (19) with (24), the phenomenological response ϵ_o and thickness d of the three-phase model of microroughness can be related to the microscopic parameters of the system:

$$d(\epsilon_a-\epsilon_o)=\int_0^{\infty}\left[\epsilon_a-\frac{1}{E_0\Omega}\int_\Omega dx\,dy\,\epsilon(\mathbf{r})\hat{\mathbf{x}}\cdot\mathbf{E}(\mathbf{r})\right]\ ,\quad z>0 \tag{25a}$$

$$d(\epsilon_s-\epsilon_o)=\int_{-\infty}^{0}\left[\epsilon_s-\frac{1}{E_0\Omega}\int_\Omega dx\,dy\,\epsilon(\mathbf{r})\hat{\mathbf{x}}\cdot\mathbf{E}(\mathbf{r})\right]\ ,\quad z<0\ . \tag{25b}$$

Equation (25) show that ϵ_o and d cannot be obtained independently, a well-known characteristic of the thin-film limit, where retardation effects are ignored.

III. APPLICATIONS

A. Lamellar grating

As a simple example, I consider the lamellar grating of Fig. 2, making the standard effective-medium assumption that the grating period is small compared to the depth d of the grooves, so that fringing fields can be ignored. If t is the thickness of the substrate protrusions, then the surface of the equivalent smooth sample is located a distance d_r above the bottom of the grooves, where $d_r=(t/L)d$. To evaluate (22b), $\phi_l(\mathbf{r})=\phi_l(x)$ must be obtained. Equations (6a) and (7) cannot be used if fringing fields are neglected because the integral reduces to an indeterminate infinite sum. But $\phi_l(\mathbf{r})$ can be determined by taking advantage of periodicity, the continuity boundary condition on normal **D**, and the fact that the line integral of **E** per period is the increment of the total potential. Taking $\phi(0)=0$, it follows that

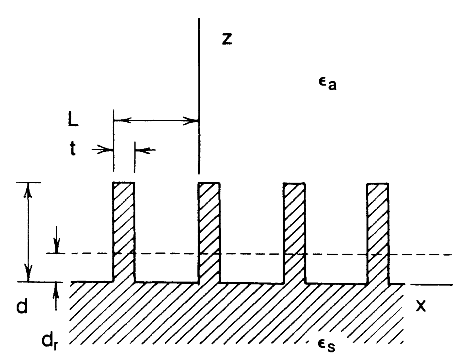

FIG. 2. Schematic diagram of the cross section of microscopic roughness as a thin laminar film. The reference plane $z=0$ is taken to coincide with the bottom of the laminae.

$$\phi(x)=\begin{cases}-xE_s, & 0\le x\le t \\ -E_s t - E_a(x-t), & t\le x\le L\end{cases} \tag{26a, 26b}$$

over the single grating period $0\le x\le L$, where

$$E_s = E_0\frac{\epsilon_a L}{\epsilon_s L-(\epsilon_s-\epsilon_a)t}\ , \tag{26c}$$

$$E_a = E_0\frac{\epsilon_s L}{\epsilon_s L-(\epsilon_s-\epsilon_a)t} \tag{26d}$$

are the fields in s and a, respectively. By symmetry, I need to evaluate (22b) only over one period, from x just less than 0 to x just less than L, and to integrate y over some arbitrary fixed length Y. The choice $\phi(0)=0$ eliminates the contribution from the vertical interface at $x=0$, so I need only the contribution from the vertical interface at $x=t$. Since $\phi_l(t)=(E_0-E_s)t$, it follows immediately that, for light linearly polarized along $\hat{\mathbf{x}}$,

$$\frac{\Delta\tilde r}{\tilde r}=\frac{4\pi i d n_a}{\lambda L}\frac{(\epsilon_s-\epsilon_a)(L-t)t}{\epsilon_s L-(\epsilon_s-\epsilon_a)t}\ . \tag{27}$$

With this choice of reference plane, $\Delta\tilde r/\tilde r=0$ for light linearly polarized along y.

For this configuration, $\Delta\tilde r/\tilde r$ can be calculated from effective-medium theory and the three-phase model. From (26) the effective film dielectric function for polarization perpendicular to the laminations is

$$\epsilon_o=\frac{\epsilon_a\epsilon_s L}{\epsilon_a t+\epsilon_s(L-t)}\ . \tag{28}$$

If the reference plane is assumed to lie at the base of the laminations, then (24a) yields

$$\frac{\Delta\tilde r}{\tilde r}=\frac{4\pi i d n_a}{\lambda}\frac{\epsilon_s(L-t)}{\epsilon_s(L-t)+\epsilon_a t}\ , \tag{29}$$

which differs from (27) only by the absolute phase shift $(-4\pi i d n_a/\lambda)[(L-t)/L]$ corresponding to the phase accumulated in propagation between the two different reference planes. Thus the two approaches yield the same result.

B. Expansion for large $|\epsilon_s|$

For large $|\epsilon_s|$ the interface-potential problem simplifies enormously by factoring into independent parts for substrate and ambient. Here, the substrate field can have no component perpendicular to the interface, so the substrate solution is isomorphic to that of fluid flow past an impenetrable boundary. For simple two-dimensional geometries this can be treated analytically by conformal mapping. Moreover, because the relative energy density in s is of the order of $|\epsilon_s|/\epsilon_a$, in the limit of large $|\epsilon_s|$ the interface potential is determined entirely by the substrate, so the surface charge need not be calculated at all. This limit is relevant because $|\epsilon_s|$ is much greater than ϵ_a for semiconductors and metals in the infrared and visible–near-ultraviolet spectral ranges.

Consequently, I investigate the possibility of obtaining solutions of (22) in powers of ϵ_a/ϵ_s. The connection can be made through (6a) and (7). Equating like orders in a series expansion in $\epsilon_a/|\epsilon_s|$ yields the zeroth-order relation

$$\sigma_0(\mathbf{r}_s)=\frac{1}{2\pi}\hat{\mathbf{n}}(\mathbf{r}_s)\cdot\left[\hat{\mathbf{x}}E_0+\int_{\bar S}d^2r_s'\,\frac{\sigma_0(\mathbf{r}_s')(\mathbf{r}_s-\mathbf{r}_s')}{|\mathbf{r}_s-\mathbf{r}_s'|^3}\right]\ , \tag{30a}$$

and the first-order correction, $\Delta\sigma_0$, to σ_0,

$$\Delta\sigma(\mathbf{r}_s)=-\frac{2\epsilon_a}{\epsilon_s}\sigma_0+\frac{1}{2\pi}\hat{\mathbf{n}}(\mathbf{r}_s)\cdot\int_{\bar S}d^2r_s'\frac{\Delta\sigma_0(\mathbf{r}_s')(\mathbf{r}_s-\mathbf{r}_s')}{|\mathbf{r}_s-\mathbf{r}_s'|^3}\ . \tag{30b}$$

In general, (30b) must be solved numerically. But where self-consistency effects are small, I have approximately

$$\frac{\Delta\tilde r}{\tilde r}\approx\frac{4\pi i n_a}{\lambda E_0\Omega}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\int_S d^2r_s\,\hat{\mathbf{x}}\cdot\hat{\mathbf{n}}(\mathbf{r}_s)\phi_l^0(\mathbf{r}_s)\ . \tag{30c}$$

One such case is the symmetric lamellar grating discussed above. If $t=L/2$ the self-consistency term cancels by symmetry, and

$$\frac{\Delta\tilde r}{\tilde r}\approx\frac{4\pi i d n_a}{\lambda}\frac{1}{2}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\ . \tag{31}$$

The leading term is a phase shift that effectively moves the reference plane from the equivalent smooth surface to the top of the laminations, showing that, for $\epsilon_s\to\infty$, a microstructured layer is as good a reflector as a smooth substrate. The correction term describes field-penetration losses for finite ϵ_s. Since $\mathrm{Im}(\epsilon_s)>0$, it follows that $\Delta R/R<0$, i.e., that R is reduced by the impedance-matching effect of microstructuring.

C. Conformal mapping solutions

If $|\epsilon_s|$ is large and the microstructure varies spatially in only two dimensions, $\phi_l(\mathbf{r})$ can be obtained by conformal mapping. As a conformal map preserves angles from one plane to another, a function $\phi(x,y)$ that is harmonic in $z=x+iy$ transforms into a function $\phi(u,v)=\phi(x(u,v),y(u,v))$ that is also harmonic in $w=f(z)=u+iv$ at every point z, where $f(z)$ is analytic and $f'(z)\neq 0$.[55] Let $w(z)$ be the map that takes the solution $\phi(x,y)=-E_0x$ for the ideal smooth sample into something more complicated, for example, a weakly sinusoidal surface, a symmetric low ridge, or an isolated low step, as will be discussed in following subsections. Maps of interest become identity transformations $w=z$ deep within the substrate, and consequently preserve projected lengths (e.g., repeat distances) along the surface. The problem then reduces to evaluating (30c) in the coordinate system (u,v) for the local-potential part of the function $\phi(u,v)=-E_0x(u,v)$ defined by the mapping $w=w(z)$.

In w, (30c) can be written

$$\frac{\Delta\tilde{r}}{\tilde{r}} \approx \frac{4\pi i n_a}{\lambda E_0 L}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\int_C ds\,\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}(s)\phi_l^0(s)\ , \qquad (32)$$

where L is the width of a full cycle of the map in w, $C=w(x,0)$ is the interface, i.e., the mapping of the ideal interface $y=0$ onto w, and s and ds relate to the position variable along C. The integration perpendicular to the xy plane cancels a common contribution from Ω, as already expressed in (32). As the local potential is equal to $\phi(u,v)=-E_0x(u,v)$, less the background or average potential, and as the mapping is assumed to be an identity at the endpoints of the segment of interest, it follows that $\phi_l(u,v)=-E_0(x(u,v)-u)$, in which case

$$\frac{\Delta\tilde{r}}{\tilde{r}} \approx \frac{4\pi i n_a}{\lambda L}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\int_C ds\,\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}(s)[u-x(u,v)]\ . \qquad (33)$$

However, (33) is written in terms of the inverse map $x+iy=z(u,v)$ rather than the direct map $u+iv=w(x,y)$. For convenience I recast (33) into variables defined on z, in which case $u\rightarrow u(x,0)$ and $x(u,v)\rightarrow x$. s can be replaced parametrically by x since C is defined as $C=w(x,0)$. The product $\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}\,ds$ reduces simply to $-(\partial v/\partial x)dx$, whence (33) becomes

$$\frac{\Delta\tilde{r}}{\tilde{r}} \approx \frac{4\pi i n_a}{\lambda L}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\int_0^L dx\frac{\partial v}{\partial x}[x-u(x,0)]\ . \qquad (34)$$

Not only is (34) simpler than (33), all quantities can be obtained from the direct map $w=w(z)$. This is the reverse of the usual procedure, where the goal is the field distribution and $z=z(w)$ is the function of interest.

D. The low sinusoidal grating

The formalism developed above permits the investigation of geometries to which effective-medium theory cannot be applied. One is the low sinusoidal grating, for which $\hat{\mathbf{n}}(\mathbf{r}_s')\cdot(\mathbf{r}_s-\mathbf{r}_s')\approx 0$ for all $\mathbf{r}_s,\mathbf{r}_s'$. Here, the self-consistent term in (30b) should be small and (34) should be a good approximation as long as the peak-to-valley height d is much less than the grating period L. The configuration is defined in Fig. 3. The transformation

$$w=u(x,y)+v(x,y)=z+\tfrac{1}{2}ie^{-2\pi i z/L} \qquad (35)$$

evaluated for $y=0$ yields the boundary variation shown in Fig. 3. The quantities needed in (34) are readily determined, yielding the result

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\left[\frac{4\pi i d n_a}{\lambda}\frac{1}{2}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\right]\frac{\pi d}{2L}\ . \qquad (36)$$

The square-bracketed part of (36) is the result for the symmetric lamellar grating, which provides a useful benchmark. For this protypical incipiently microscopically rough surface, $\Delta\tilde{r}/\tilde{r}$ is bilinar in the peak-to-peak height d and the slope d/L, consistent with the results of Mochán and Barrera for weak random roughness.[45] The apparent divergence of (40) as $1/L$ for vanishing L is an artifact of the neglect of self-consistency, an invalid assumption for slopes of order unity or greater.

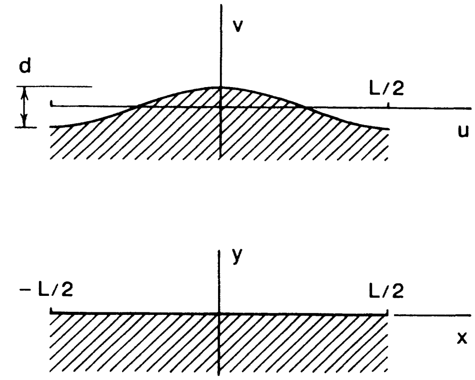

FIG. 3. Top: schematic diagram of the cross section of a low sinusoidal grating. Bottom: schematic diagram of the ideal surface that is conformally mapped to generate the low sinusoidal grating.

E. The low symmetric ridge and the isolated low step

Two cases that can be treated similarly are the low symmetric ridge and the isolated low step, both of interest with respect to crystal growth. The configurations are shown in Figs. 4 and 5. As the ridge can be considered a symmetric pair of steps, I take L to be half of the repeat distance of the ridge. For simplicity I consider only the symmetric case where the ridge and valley are equally wide, or, equivalently, where the step is situated near the midpoint of the range of interest. To achieve an analytic result, I assume that the self-consistent contribution term is small, but this assumption should be checked by numerical methods.

The symmetric ridge is generated by the Schwartz-Christoffel transformation[55]

$$\frac{dw}{dz}=\left[\frac{\cos(kx_1)-\cos(kz)}{\cos(kx_2)-\cos(kz)}\right]^{1/2}, \qquad (37)$$

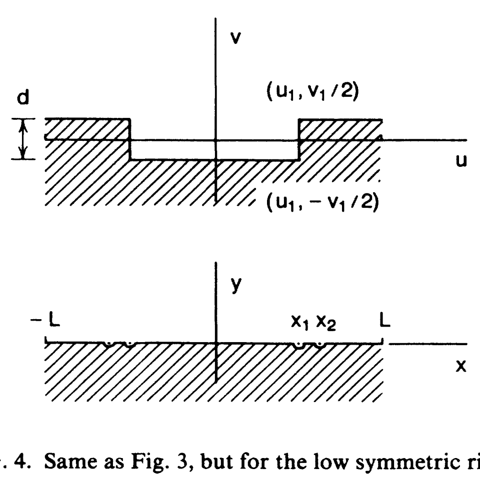

FIG. 4. Same as Fig. 3, but for the low symmetric ridge.

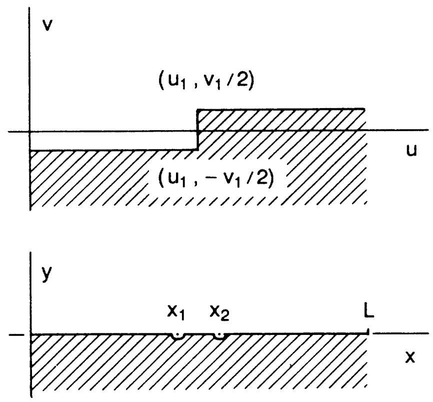

FIG. 5. Same as Fig. 3, but for the isolated low step.

where $k=\pi/L$. Clearly, $x=0$ and $\pm L/2$ are mirror planes of the transformation, which has a repeat period $\Delta x=2L$. By symmetry I consider only the segment $0\le x\le L$. The transformation maps $z=(0,0)$, $(x_1,0)$, $(x_2,0)$, and $(L,0)$ onto $w=(0,-v_1/2)$, $(u_1,-v_1/2)$, $(u_1,v_1/2)$, and $(L,v_1/2)$, respectively. Let $x_1=L/2-\Delta x/2$ and $x_2=L/2+\Delta x/2$, where Δx will be related to the step height d. From (37) and (33),

$$u_1=\int_0^{x_1} dx\left[\frac{\cos(kx)-\cos(kx_1)}{\cos(kx)-\cos(kx_2)}\right]^{1/2}, \tag{38a}$$

$$v_1=d=\int_{x_1}^{x_2} dx\left[\frac{\cos(kx_1)-\cos(kx)}{\cos(kx)-\cos(kx_2)}\right]^{1/2}, \tag{38b}$$

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{4\pi i n_a}{\lambda L}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\times\int_{x_1}^{x_2} dx(x-u_1)\left[\frac{\cos(kx_1)-\cos(kx)}{\cos(kx)-\cos(kx_2)}\right]^{1/2}. \tag{38c}$$

Equation (38a) can be converted into standard elliptic-integral form by substituting $t=\cos(kx)$. With

$$a=\cos(kx_1)\approx\frac{\pi\Delta x}{2L}, \tag{39a}$$

$$b=\cos(kx_2)\approx -a, \tag{39b}$$

(38a) reduces to[65]

$$u_1=\frac{4La}{\pi(1+a)}[\Pi(-\kappa,\kappa)-K(\kappa)], \tag{40}$$

where $\kappa=(1-a)/(1+a)$. $\Pi(-\kappa,\kappa)$ can be evaluated with the identity $2\Pi(-\kappa,\kappa)=K(\kappa)+\pi/[2(1-\kappa)]$, in which case

$$u_1\approx\frac{L}{2}-\frac{\Delta x}{2}\ln\left[\frac{8L}{\pi\Delta x}\right], \tag{41a}$$

where I have also used $K(\kappa)\approx\ln[4/(1-\kappa)]$ as $\kappa\to 1$. Equation (38b) can be treated similarly, but since v_1 is already of order of the step height, the integration can be simplified substantially, yielding

$$v_1=d\approx\frac{\pi}{2}\Delta x. \tag{41b}$$

Likewise, the integral $I(x_1,x_2,u_1,k)$ in (38c) becomes

$$I(x_1,x_2,u_1,k)\approx\frac{d^2}{2\pi}\left[1+2\ln\left[\frac{4L}{d}\right]\right]. \tag{41c}$$

Combining Eqs. (41) yields

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{4\pi i n_a}{\lambda L}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\frac{d^2}{\pi}\left[\tfrac{1}{2}+\ln 4+\ln\left[\frac{L}{d}\right]\right]. \tag{42}$$

Thus, in addition to the bilinear dependence on slope and height, $\Delta\tilde{r}/\tilde{r}$ exhibits a logarithmic singularity, which arises from the long-range distortion of the field map away from the immediate vicinity of the step. This singularity is expected to saturate with increasing L owing to self-consistency effects that are not included in the above, but which must be calculated numerically.

A similar but simpler calculation can be performed for the isolated step of Fig. 5. Here,

$$\frac{dw}{dz}=\left[\frac{z-x_1}{z-x_2}\right]^{1/2}, \tag{43}$$

from which it follows that

$$u_1=\int_0^{x_1} dx\frac{(x_1-x)^{1/2}}{(x_2-x)^{1/2}}, \tag{44a}$$

$$v_1=d=\int_{x_1}^{x_2} dx\frac{(x-x_1)^{1/2}}{(x_2-x)^{1/2}}, \tag{44b}$$

$$\frac{\Delta\tilde{r}}{\tilde{r}}=\frac{4\pi i n_a}{\lambda L}\left[1-\frac{2\epsilon_a}{\epsilon_s}\right]\frac{d^2}{\pi}\left[\tfrac{1}{2}+\ln\pi+\ln\left[\frac{L}{d}\right]\right]. \tag{44c}$$

The only difference between (42) and (44c) is the replacement of the term $\ln 4=1.886$ by $\ln\pi=1.645$, a consequence of the different constraints imposed on the long-range properties of the map of the isolated step by not incorporating periodicity.

IV. CONCLUSIONS

In this work, I establish a general approach to calculate the effect of microscopic roughness on the normal-incidence complex reflectance when empirical effective-medium models cannot be applied, and use it to obtain simple analytic expressions for microscopic-roughness effects for several geometries of current interest. In addition to providing the first estimates of the effects of incipient microroughness and low steps on $\tilde{r}$, the work also establishes the mathematical definition of microscopic roughness, justifies the empirical effective-medium approach for normal-incidence illumination, and illustrates some of the essential physics of the problem of the microscopically rough surface. The framework established here should also be generalizable to non-normal incidence and to thicker films where retardation cannot be ignored. Such elaborations are necessary for spectroellipsometric applications.

The normal-incidence results given here are accurate enough for the present purposes, but numerical computations should yield improvements in the form of estimates of the magnitudes of the self-consistency effects that are neglected in the present treatment and of the ranges of validity of the current expressions. In addition, such calculations will allow atomic polarizabilities to be incorporated.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge useful conversations with J. M. Worlock about Schwartz-Christoffel transformations.

[1]See, e.g., M. Born and E. Wolf, *Principles of Optics,* 5th ed. (Pergamon, New York, 1975).

[2]P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (Pergamon, Oxford, 1963).

[3]P. Beckmann, *The Depolarization of Electromagnetic Waves* (Golem, Boulder, CO, 1968).

[4]J. M. Bennett and L. Mattsson, *Introduction to Surface Roughness and Scattering* (Optical Society of America, New York, 1990), Vol. 1. Bennett and Mattsson comprehensively review both experimental and theoretical work in this field.

[5]Lord Rayleigh, Philos. Mag. **14**, 70 (1907).

[6]U. Fano, J. Opt. Soc. Am. **31**, 213 (1941).

[7]B. A. Lippmann, J. Opt. Soc. Am. **43**, 408 (1953).

[8]H. Davies, Proc. IEEE **101**, 209 (1954).

[9]H. E. Bennett and J. O. Porteus, J. Opt. Soc. Am. **51**, 123 (1961).

[10]P. A. Fedders, Phys. Rev. **165**, 580 (1968).

[11]J. Crowell and R. H. Ritchie, J. Opt. Soc. Am. **60**, 795 (1970).

[12]J. M. Elson and R. H. Ritchie, Phys. Rev. B **4**, 4129 (1971).

[13]I. Ohlidal and F. Lukes, Opt. Acta **19**, 817 (1972).

[14]J. M. Elson and R. H. Ritchie, Phys. Status Solidi B **62**, 461 (1974).

[15]A. A. Maradudin and D. L. Mills, Phys. Rev. B **11**, 1392 (1975).

[16]V. Celli, A. Marvin, and F. Toigo, Phys. Rev. B **11**, 1779 (1975).

[17]E. N. Economou and K. L. Ngai, Adv. Chem. Phys. **23**, 1989 (1976).

[18]I. Ohlidal, F. Lukes, and K. Navratil, J. Phys. (Paris) Colloq. **38**, 77 (1977).

[19]F. Toigo, A. Marvin, V. Celli, and N. R. Hill, Phys. Rev. B **15**, 5618 (1977).

[20]B. Laks and D. L. Mills, Phys. Rev. B **20**, 4962 (1979).

[21]B. Laks and D. L. Mills, Phys. Rev. B **21**, 5175 (1980).

[22]A. M. Brodsky and M. I. Urbakh, Surf. Sci. **115**, 417 (1982).

[23]K. Arya and R. Zeyher, Phys. Rev. B **28**, 4090 (1983).

[24]M. M. Wind and J. Vlieger, Physica **125A**, 75 (1984).

[25]M. M. Wind and J. Vlieger, Physica **131A**, 1 (1985).

[26]A. Wirgin and A. A. Maradudin, Phys. Rev. **31**, 5573 (1985).

[27]O. Mata-Méndez, Phys. Rev. B **37**, 8182 (1988).

[28]M. Nieto-Vesperinas and J. M. Soto-Crespo, Phys. Rev. B **38**, 7250 (1988).

[29]D. E. Aspnes, J. Vac. Sci. Technol. **17**, 1057 (1980).

[30]C. A. Fenstermaker and F. L. McCrackin, Surf. Sci. **16**, 85 (1969).

[31]J. C. M. Garnett, Philos. Trans. R. Soc. London **203**, 385 (1904); **205**, 237 (1906).

[32]D. A. G. Bruggeman, Ann. Phys. (Leipzig) **24**, 636 (1935).

[33]D. E. Aspnes, J. B. Theeten, and F. Hottier, Phys. Rev. B **20**, 3292 (1979).

[34]F. Hottier and J. B. Theeten, J. Cryst. Growth **48**, 644 (1980).

[35]D. E. Aspnes, Physica B+C (Amsterdam) **117&118B**, 359 (1983).

[36]J. R. Blanco, P. J. McMarr, and K. Vedam, Appl. Opt. **24**, 3773 (1985).

[37]P. G. Snyder, M. C. Rost, G. H. Bu-Abbud, J. Oh, J. A. Woollam, D. Poker, D. E. Aspnes, D. Ingram, and P. Pronko, J. Appl. Phys. **60**, 770 (1986).

[38]S. F. Nee and H. E. Bennett, Proc. SPIE **675**, 260 (1986).

[39]A. M. Antoine, B. Drevillon, and P. Roca i Cabarrocas, J. Appl. Phys. **61**, 2501 (1987).

[40]R. W. Collins and J. M. Cavese, J. Appl. Phys. **61**, 1662 (1987).

[41]B. T. Sullivan and R. R. Parsons, J. Vac. Sci. Technol. A **5**, 3399 (1987).

[42]P. J. McMarr and J. R. Blanco, Appl. Opt. **27**, 4265 (1988).

[43]S.-M. F. Nee, Appl. Opt. **27**, 2819 (1988).

[44]W. L. Mochán and R. G. Barrera, Phys. Rev. B **32**, 4989 (1985).

[45]W. L. Mochán and R. G. Barrera, Phys. Rev. B **32**, 4984 (1985).

[46]D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, Phys. Rev. Lett. **59**, 1687 (1987).

[47]D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, J. Vac. Sci. Technol. A **6**, 1327 (1988).

[48]D. E. Aspnes, E. Colas, A. A. Studna, R. Bhat, M. A. Koza, and V. G. Keramidas, Phys. Rev. Lett. **61**, 2782 (1988).

[49]D. E. Aspnes, R. Bhat, E. Colas, V. G. Keramidas, M. A. Koza, and A. A. Studna, J. Vac. Sci. Technol. A **7**, 711 (1989).

[50]D. E. Aspnes, A. A. Studna, L. T. Florez, Y.-C. Chang, J. P. Harbison, M. K. Kelly, and H. H. Farrell, J. Vac. Sci. Technol. B **7**, 901 (1989).

[51]D. E. Aspnes and A. A. Studna, Phys. Rev. Lett. **54**, 1956 (1985).

[52]D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).

[53]J. D. Jackson, *Classical Electrodynamics,* 2nd ed. (Wiley, New York, 1975).

[54]F. Hildebrand, *Advanced Calculus for Engineers* (Prentice-Hall, Englewood Cliffs, NJ, 1949).

[55]R. V. Churchhill, *Introduction to Complex Variables and Applications* (McGraw-Hill, New York, 1948).

[56]S. O. Rice, Commun. Pure Appl. Math. **4**, 351 (1951).

[57]E. Kretschmann, Z. Phys. **227**, 412 (1969).

[58]E. Kröger and E. Kretschmann, Z. Phys. **237**, 1 (1970).

[59]H. J. Juranek, Z. Phys. **233**, 324 (1970).

[60]A. Bagchi, R. G. Barrera, and A. K. Rajagopal, Phys. Rev. B **20**, 4824 (1979).

[61]M. D. Williams and D. E. Aspnes, Phys. Rev. Lett. **41**, 1667 (1978).

[62]W. J. Plieth and K. Naegele, Surf. Sci. **64**, 484 (1977).

[63]J. D. E. McIntyre and D. E. Aspnes, Surf. Sci. **24**, 417 (1971).

[64]P. J. Feibelman, Phys. Rev. B **14**, 762 (1976).

[65]W. Gröbner and N. Hofreiter, *Integraltafel,* 3rd ed. (Springer, Wien, 1961).