Convergent scheme for light scattering from an arbitrary deep metallic grating

Dan Agassi
Naval Surface Weapons Center, White Oak, Silver Spring, Maryland 20910

Thomas F. George
Departments of Chemistry and Physics, State University of New York at Buffalo, Buffalo, New York 14260
(Received 5 August 1985)

The justification for continuing the Rayleigh expansion to the grating's surface (the Rayleigh hypothesis) and its convergence properties are considered. A class of gratings for which the Rayleigh hypothesis is exact is identified, a prime example of which is the sinusoidal grating. Based on the exposure of the origin for the Rayleigh expansion limited convergence, a modified expansion is introduced, dubbed the dressed Rayleigh expansion. This new expansion has presumably excellent convergence properties as explicitly demonstrated for the sinusoidal grating. The dimensionality N of the matrix which must be inverted for a sinusoidal grating of arbitrary depth g and periodicity d is found to be $N \sim 8\pi g/d$.

I. INTRODUCTION

Light scattering from a grating is hardly a new phenomenon, and consequently the literature on the subject is enormous.[1,2] However, it is important to realize that, barring a few recent exceptions, the majority of the literature deals with very shallow gratings, i.e., when the ratio between the height g and periodicity d is $g/d \ll 1$. Shallow gratings are realized in many important physical situations, such as optical-refraction gratings, holograms, gratings induced by surface phonons,[3] and a large class of volume gratings.[4]

Studies of light scattering from deep gratings, i.e., when $g/d \approx 1$ or larger, are relatively recent.[5–11] The motivation to consider this regime stems in part from the advent of new fabrication capabilities. By using holographic techniques,[12] or laser-induced deposition from a volatile organometallic gas,[13] it is possible to form gratings with $g/d \leqq 1$. Another reason for the interest in deep gratings is the expected new qualitative features. When $\beta \ll 1$, where hereafter we use $\beta = 2\pi g/d$ as the natural parameter, the surface can exchange only to a limited degree surface-parallel momentum quanta $k_G = 2\pi/d$ with the incident light. Consequently, for shallow gratings it is sufficient to consider a few Bragg reflections in addition to the dominating specularly reflected and transmitted waves. For deep gratings ($\beta \gg 1$), the situation is quite different. Here the grating can efficiently exchange a large number of quanta k_G with the incident light. Therefore, the reflected wave, for example, is made up of many Bragg reflections, strongly interfering with each other and with the specularly reflected wave. Thus, unlike the $\beta \ll 1$ regime, when $\beta \gg 1$ the identity of the specular components and the individual Bragg reflections is blurred and the reflected light acquires a new qualitative character.

The qualitative distinction between the $\beta \ll 1$ and $\beta \gg 1$ regimes is reminiscent of the situation with regard to band structure in condensed-matter physics.[14] Umklapp scatterings play the role of Bragg reflections, and the regimes of weak and strong periodic ionic potentials are the counterparts of the shallow- and deep-grating regimes, respectively. In keeping with this analogy, the qualitative difference between light scattering from a grating and from a rough surface corresponds to the qualitative difference between electronic states in an ordered and disorderd system. At the origin of this distinction is the underlying symmetry: a discrete translational symmetry in gratings (crystals) and none at all in rough surfaces (disordered systems). Consequently the character of the ensuing interference among the many reflected components is fundamentally different.

Attempts to study theoretically the deep-grating regime have encountered numerical convergence problems[1,15] when $g/d \gtrsim 0.1$. The associated literature[1,2,16,17] often makes reference to two related issues which are at the focus of this work: the validity of the Rayleigh hypothesis and the convergence of the Rayleigh expansion. The Rayleigh expansion[18] of the electromagnetic fields above the selvedge domain [region a in Fig. 1(a)] embodies the underlying symmetry of the grating as expressed by the Floquet-Bloch theorem[14] and the boundary conditions for $z \rightarrow \infty$ [in the notation of Fig. 1(a)]. Thus the Rayleigh expansion results from a general principle. The first issue is whether, or when, it can be continued into the selvedge domain all the way to the grating's surface [domain b in Fig. 1(a)]. The assumption that this is possible is referred to as the Rayleigh hypothesis. The second issue is what are the convergence properties of the Rayleigh expansion, in light of the numerical difficulties encountered in applying it for deep gratings.

Notwithstanding the many attempts to settle these issues for a general grating,[1,2,16,17,19] there is no clear answer yet. This state of affairs is particularly puzzling in view of the generality of the Rayleigh expansion (which is actually a Fourier series over a unit cell of length d). In an effort to understand these issues, we focus on the analysis of light scattering from a class of simple gratings, with the sinusoidal grating (SG) serving as a prototype. The two results of this work are as follows.

(i) We identify a class of gratings for which the Rayleigh hypothesis is *exact*. Gratings in this class, such as the SG, have the property that the domain above the selvedge and the adjacent domain in the selvedge [e.g., domains a and b in Fig. 1(a)], share the whole x axis except for a set of isolated nonsingular points.

(ii) With regard to the convergence properties of the Rayleigh-expansion, we identify the cause for divergence when $\beta \gg 1$. This analysis, in turn, leads naturally to a simple alternative expansion—the dressed Rayleigh expansion of Eqs. (3.5) and (3.6)—which has presumably excellent convergence properties. This proposition is explicitly checked for the SG. We find in this case that the dressed expansion converges for an arbitrary β, and the estimated order of the matrix needed to be inverted is 4β. Hence, for example, for a SG with $g/d=1$, inverting a matrix of order ≤ 50 should be adequate.

The paper is organized as follows. In Sec. II we introduce the Rayleigh expansion and discuss the class of gratings for which the Rayleigh hypothesis is exact. Section III is devoted to examining the convergence properties of the Rayleigh expansion and the introduction of the dressed expansion. In Sec. IV the latter is analyzed for the SG. Discussion and concluding remarks are given in Sec. V.

II. RAYLEIGH EXPANSION AND HYPOTHESIS

To introduce the Rayleigh expansion, we start by considering the symmetry of the system: Since the grating is invariant under translations of the type $x \rightarrow x + Ld$, where L is an integer and d is the periodicity [all notations are defined in Fig. 1(a)], so are the solutions of Maxwell's equations. Consequently, according to the Floquet-Bloch theorem,[6,8,14,17] the electric or magnetic fields $\Psi(x,z)$ satisfy

$$\Psi(x+d\,z)=e^{ik_{\parallel}d}\Psi(x,z)\ , \tag{2.1}$$

where $k_{\parallel}$ is a surface-parallel (in the x direction) momentum label. For a scattering wave mode, $k_{\parallel}$ is the x component of the incident wave. For a surface wave mode (surface plasmons in the case of metallic gratings), $k_{\parallel}$ is a continuous label restricted to the first Brillouin zone $-k_G/2 \leq k_{\parallel} \leq k_G/2$. To simplify the subsequent analysis, we consider only the configuration of a p-polarized plane wave incident perpendicular to the direction of the grating's grooves (the y axis). In this case considering the xz plane is sufficient.

The symmetry property (2.1) determines the functional form of the fields up to constants, to be determined by matching the proper boundary conditions. Assuming an $e^{-i\omega_0 t}$ time dependence of the electromagnetic fields, where ω_0 is the field frequency, we can always expand the solutions of Maxwell's equations (at least when they are not singular) in terms of the following complete set of functions in the $0<x\leq d$ interval:[8,16,18]

$$\mathbf{E}_\alpha(x,z)=\sum_{l=-\infty}^{\infty}\left(C_\alpha(l)\hat{\mathbf{p}}_{\alpha,-}(l)\exp\{i[k_l x - W_\alpha(l)z]\}+A_\alpha(l)\hat{\mathbf{p}}_{\alpha,+}(l)\exp\{i[k_l x + W_\alpha(l)z]\}\right) \tag{2.2a}$$

and

$$\mathbf{B}_\alpha(x,z)=\frac{k(\alpha)}{k}\hat{\mathbf{s}}\sum_{l=-\infty}^{\infty}\left(C_\alpha(l)\exp\{i[k_l x - W_\alpha(l)z]\}+A_\alpha(l)\exp\{i[k_l x + W_\alpha(l)z]\}\right)\ , \tag{2.2b}$$

where[20]

$$\begin{aligned} &k(\alpha)=(\epsilon_\alpha)^{1/2}k,\ \ k=\omega_0/c,\ \ k_G=2\pi/d\ ,\\ &W_\alpha(l)=[k^2(\alpha)-k_l^2]^{1/2},\ \ k_l=k_{\parallel}+lk_G\ ,\\ &\hat{\mathbf{p}}_{\alpha,\pm}(l)=\frac{1}{k(\alpha)}[k_l\hat{\mathbf{z}}\mp W_\alpha(l)\hat{\mathbf{x}}],\ \ \hat{\mathbf{s}}=\hat{\mathbf{x}}\times\hat{\mathbf{z}}\ . \end{aligned} \tag{2.2c}$$

In (2.2) we choose always $\mathrm{Im}[W_\alpha(l)]$ or $\mathrm{Re}[W_\alpha(l)]$ as positive, $\hat{\mathbf{x}}$, $\hat{\mathbf{z}}$, and $\hat{\mathbf{s}}$ are unit vectors in the x, z, and y directions [Fig. 1(a)], and α denotes a domain in the xz plane with a constant dielectric function ϵ_α. (ϵ_α can depend on ω_0 in the following analysis since ω_0 is kept fixed throughout.) Once a convenient partition of the xz plane into totally filling domains has been chosen, the coefficients $C_\alpha(l), A_\alpha(l)$ are determined by matching the boundary conditions across the boundaries of the domains.

To demonstrate the choice of domains α, consider in particular a SG-type example [Fig. 1(a)]. It is obvious that in this case there are four relevant domains and hence four expansions of the type (2.2) to determine. In these terms, the issue of whether the Rayleigh expansion can, or cannot, be continued to the grating's surface is tantamount to whether the expansions in domains a and b (and similarly c and d) are identical. This issue must be decided by matching the boundary conditions between domains a and b along the separating plane $z=g$.

Consider for example the tangential components of $\mathbf{E}$ and $\mathbf{B}$ along $\hat{\mathbf{x}}$ and $\hat{\mathbf{s}}$ [Eq. (2.2)]. By comparing coefficients of the complete set $\{e^{ik_l x}\}$ we obtain

$$\begin{aligned} &C_a(l)e^{-R}-A_a(l)e^{+R}=C_b(l)e^{-R}-A_b(l)e^{+R}\ ,\\ &C_a(l)e^{-R}+A_a(l)e^{+R}=C_b(l)e^{-R}+A_b(l)e^{+R}\ , \end{aligned} \tag{2.3}$$

where

$$R=iW_a(l)g=iW_b(l)g\ . \tag{2.4}$$

Equations (2.3) yield $C_a(l)=C_b(l)$ and $A_a(l)=A_b(l)$, and hence the Rayleigh hypothesis is exact. In establishing (2.3) we used the fact that domains a and b share the entire x axis (or unit cell). Thus it is possible to equate coef-

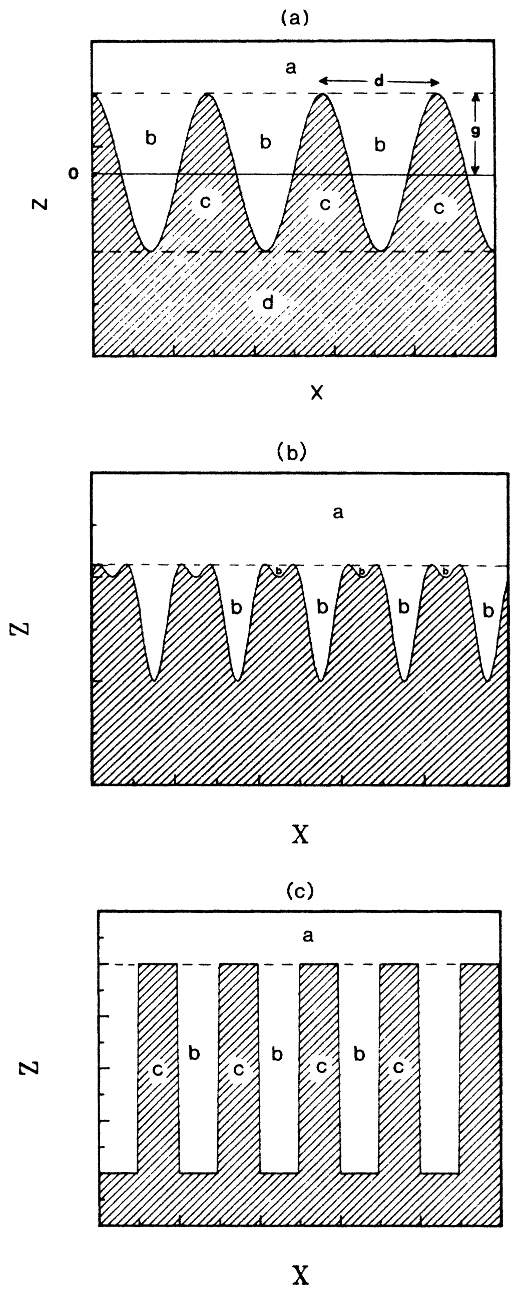

FIG. 1. (a) Schematic display of a grating for which there are four obvious domains with constant dielectric constant. The domains are denoted by a, b, c, and d. The hatched and plain areas indicate the dielectric and "air," respectively. The maximum $|z|$ extension and periodicity of the grating are denoted by g and d, respectively. (b) A more complex grating for which the Rayleigh hypothesis is exact, i.e., field expansions in domains a and b are identical (see Sec. II). (c) The square-well grating for which the Rayleigh expansion is not exact.

ficients of each member of the complete set $\{e^{ik_l x}\}$. It has been tacitly assumed at this junction that the set of isolated nonsingular points ($x = nd$, $z = g$), jointly shared by domains a, b, and c, do not break the orthogonality of the set $\{e^{ik_l x}\}$ over a unit cell. The above argument remains valid for *any* grating such that the selvedge domain and the domain above it [e.g., domains a and b in Fig. 1(a)] share the entire x axis except for a set of isolated nonsingular points. Figure 1(b) depicts a slightly more complicated example where the Rayleigh expansions in domains a and b are identical.

An example for which the above argument fails is the square-well grating, Fig. 1(c). Here again we start by dividing the xz plane into four domains and try to match domains a and b. It is obvious that since the boundary between domains a and b is not the entire x axis, Eq. (2.3) is not valid; hence the Rayleigh hypothesis is *not* exact. For a shallow square-well grating, however, the Rayleigh hypothesis may provide a good approximation.

The foregoing discussion is in keeping with the accepted point of view that in general the Rayleigh expansion above the grating cannot be continued to the grating's surface. We have demonstrated, nevertheless, that for a certain class of gratings, and the SG in particular, the Rayleigh hypothesis is exact. Note also that we restricted ourselves to nonsingular fields (and first derivatives) only. When singularities are present, the implication is of surface charges, an infinite amount of charge at particular points, surface currents, etc. Consequently, the above conclusions, which are based on the boundary conditions of no surface charges or currents and the very existence of expansion (2.2), are expected to be modified.

III. DRESSED RAYLEIGH EXPANSION

Even when the Rayleigh hypothesis is exact, the usefulness of the Rayleigh expansion must be substantiated by good convergence properties. As difficulties in applying the expansion to deep gratings indicate,[15] this is probably not the case. It is for this reason that alternative schemes have been suggested[1,2,6–8,16] with presumably a wider range of convergence. Our strategy, on the other hand, is first to expose the deficiency of the Rayleigh expansion rather than to abandon it. Once this is achieved, we are naturally led to a modified expansion which we believe (and demonstrate for the SG in the next section) to have excellent convergence properties for deep gratings.

When the Rayleigh hypothesis is exact, it is sufficient to consider only two domains in the xz plane (see Fig. 2). Consequently there are only two expansions to deal with [Eq. (2.2)], indexed as 0 and 1. The electric field expansions are therefore

$$\mathbf{E}_0(x,z)=\sum_{l=-\infty}^{\infty}\left(C_0(l)\hat{\mathbf{p}}_{0,-}(l)\exp\{i[k_l x - W_0(l)z]\}+A_0(l)\hat{\mathbf{p}}_{0,+}(l)\exp\{i[k_l x + W_0(l)z]\}\right),$$

$$\mathbf{E}_1(x,z)=\sum_{l=-\infty}^{\infty}C_1(l)\hat{\mathbf{p}}_{1,-}(l)\exp\{i[k_l x - W_1(l)z]\}, \tag{3.1a}$$

and the magnetic field expansions are

$$\mathbf{B}_0(x,z)=\frac{k(0)}{k}\hat{\mathbf{s}}\sum_{l=-\infty}^{\infty}(C_0(l)\exp\{i[k_lx-W_0(l)z]\}+A_0(l)\exp\{i[k_lx+W_0(l)z]\})\ ,$$
$$\mathbf{B}_1(x,z)=\frac{k(1)}{k}\hat{\mathbf{s}}\sum_{l=-\infty}^{\infty}C_1(l)\exp\{i[k_lx-W_1(l)z]\}\ . \tag{3.1b}$$

In Eq. (3.1)

$$C_0(l)=\delta_{l,0}=E_{0p}^{-}\ , \tag{3.2}$$

where E_{0p}^{-} is the arbitrary amplitude of the incident p-polarized light. Equation (3.2) embodies the boundary condition of the downward-propagating incident wave in region 0 (Fig. 2), and (3.1) incorporates the outgoing-wave boundary conditions at $|z|\to\infty$.

The deficiency of (3.1) becomes apparent by considering, for example, the $A_0(l)$ term in (3.1b). For $|l|\to\infty$, the x-independent factor tends to

$$A_0(l)e^{iW_0(l)z}\underset{|l|\to\infty}{\longrightarrow}A_0(l)e^{-k_G|l|z}\ , \tag{3.3}$$

where (2.2c) was used. For the $z>0$ portion of domain 0, the z-dependent factor in (3.3) converges exponentially with $|l|$. However, (3.1) is valid throughout region 0. In particular, at the bottom of the troughs where $-g\leqq z\leqq 0$, the z-dependent factor in (3.3) diverges exponentially with $|l|$. Consequently, in order to keep the total field $\mathbf{B}_0(x,z)$ finite for $z\leqq 0$, it follows that the *exact* $A_0(l)$ must converge for $|l|\to\infty$ at least exponentially. (This argument fails for gratings which give rise to singularities in the fields or in their derivatives.) Therefore, for $z<0$ the Rayleigh expansion of $\mathbf{E}_0$ and $\mathbf{B}_0$ is a sum of many terms, most of which (large $|l|$) are products of exponentially large numbers (the z-dependent factors) times exponentially small numbers [the $A_0(l)$]. However, there are always errors in the *calculated* $A_0(l)$. These, in turn, will lead to large variations in the total fields, since the $A_0(l)$ are multiplied in (3.1) by very large exponents. This type of numerical instability, demonstrated here for the $z\leqq 0$ portion of the 0 domain, has an adverse effect on other portions of the xz plane by virtue of the extinction theorem.[1,2,16,21] Therefore, for cases when many Bragg reflections ($|l|$) contribute to (3.1), i.e., for deep gratings, the Rayleigh expansion is intrinsically unsuitable. For shallow gratings, where only a few Bragg reflections contribute, the instability just described does not arise.

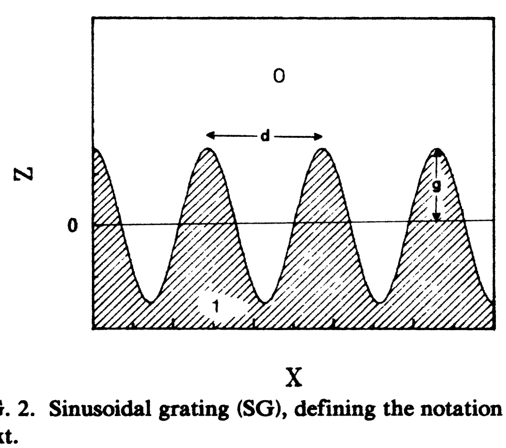

FIG. 2. Sinusoidal grating (SG), defining the notation used in the text.

The above deficiency of (3.1) can be easily remedied in the following manner. Consider again, for example, the $A_0(l)$ term in (3.1b). It can be rewritten as

$$A_0(l)\exp[iW_0(l)z]=\alpha_0(l)\exp[iW_0(l)(z+g)]\ , \tag{3.4}$$

where g is the (positive) minimum of the grating in region 0 (see Fig. 2) and $\alpha_0(l)=A_0(l)\exp[-iW_0(l)g]$. Since by construction $z+g\geqq 0$ throughout region 0, the exponential factor on the right-hand side of (3.4) never diverges; in fact, it always converges exponentially (and is unity at the grating's profile). Furthermore, since (3.1) is valid throughout region 0, including at the grating surface, the exact $\alpha_0(l)$ converge with very high l to render a finite total field. Consequently, the *relevant* exact $\alpha_0(l)$ are not necessarily exponentially small, and hence small errors in the calculated $\alpha_0(l)$ will not lead to instabilities in the total field. These considerations suggest that we transcribe the Rayleigh expansion (3.1) into a "dressed Rayleigh expansion" by writing

$$A_0(l)\exp[iW_0(l)z]=\alpha_0(l)\exp[iW_0(l)(z+g)]\ ,$$
$$C_1(l)\exp[-iW_1(l)z]=\gamma_1(l)\exp[-iW_1(l)(z-g)]\ , \tag{3.5}$$

where the dressed-expansion coefficients are given by

$$\alpha_0(l)=A_0(l)\exp[-iW_0(l)g]\ ,$$
$$\gamma_1(l)=C_1(l)\exp[-iW_1(l)g]\ . \tag{3.6}$$

The central assertion of this work is that the dressed Rayleigh expansion, defined by (3.5) and (3.6), has excellent convergence properties for a wide range of values of β and is a suitable framework for deep-grating calculations. This proposition is explicitly demonstrated for the SG in the next section. Furthermore, the procedure used in Sec. IV for the SG can be generalized to other gratings (see Sec. V), indicating the wide applicability of the dressed Rayleigh expansion.

IV. SINUSOIDAL GRATING (SG)

The proposition stated as a conjecture in Sec. III with regard to the dressed Rayleigh expansion, (3.5) and (3.6), is now explicitly checked for the SG. Our starting point is the exact infinite set of coupled linear equations[16] satisfied by the expansion coefficients of (3.1) (see Appendix A for the derivation),

$$\sum_{l=-\infty}^{\infty} M^B_{m,l} A_0(l) = \mu^B(m)\ ,$$
$$\sum_{l=-\infty}^{\infty} N^B_{m,l} C_1(l) = \nu^B(m)\ , \tag{4.1}$$

where the superscript B is a remainder that (4.1) pertains to the expansion coefficients of (3.1). These coefficients are hereafter referred to as bare, in distinction from the dressed expansion coefficients defined in (3.6). The entries to (4.1) are explicitly given by (see Appendix):

$$M^B_{m,l} = \frac{W_0(l)W_1(m) + k_l k_m}{W_0(l) - W_1(m)} \times i^{m-l} J_{m-l}(g[W_0(l) - W_1(m)])\ ,$$
$$N^B_{m,l} = \frac{W_1(l)W_0(m) + k_l k_m}{W_l(l) - W_0(m)} \times i^{m-l} J_{m-l}(g[-W_l(l) + W_0(m)])\ ,$$
$$\mu^B(m) = \frac{-W_0(0)W_1(m) + k_0 k_m}{W_0(0) + W_1(m)} \times i^m J_m(-g[W_0(0) + W_1(m)])E^-_{0p}\ , \tag{4.2}$$
$$\nu^B(m) = \frac{2\epsilon_0\epsilon_1}{\epsilon_1 - \epsilon_0} W_0(m)\delta_{m,0} E^-_{0p}\ ,$$

where the symbol J_m denotes the Bessel function of order m, and all other symbols have been defined in (2.2). The Bessel function factor in $\underline{M}^B$ and $\underline{N}^B$ determines the mixing between the Bragg reflections of order m and l, and hence involves the g parameter. Thus, for instance, when $g=0$ (flat surface), it follows that $m=l$, i.e., no mixing. The role of the other factors in (4.2) is elaborated elsewhere.[22]

The exact equations (4.1) and (4.2) expose from yet another vantage point the deficiency of the Rayleigh expansion (3.1). Consider for example $\underline{M}^B$ and the corresponding equations in (4.1). When m is fixed and $|l| \to \infty$, the fastest changing factor in (4.2) is the Bessel function. For $|l| \to \infty$ it follows that $W_1(m) \ll W_0(l) \sim ik_G|l|$, and therefore, up to uninteresting factors (phases and powers of $|l|$), the matrix elements of $\underline{M}^B$ behave as[23]

$$\lim_{\substack{|l|\to\infty \\ m \text{ fixed}}} M^B_{m,l} \sim \lim_{|l|\to\infty} J_l(i\beta|l|) \sim e^{|l|\eta(\beta)}\ , \tag{4.3}$$

where again $\beta = k_G g$ and

$$\eta(\beta) = (1+\beta^2)^{1/2} + \ln\{\beta/[1+(1+\beta^2)^{1/2}]\}\ . \tag{4.4}$$

Hence, for $\beta \gg 1$ the matrix elements of $\underline{M}^B$ diverge exponentially with $|l|$. On the other hand, for fixed m the right-hand side of (4.1) is a constant. Consequently, the *exact* $A_0(l)$ must converge at least as $e^{-|l|\eta(\beta)}$, which is in keeping with the analysis of Sec. III. Furthermore, since $\underline{M}^B$ diverges with $|l|$ as indicated in (4.3), there is no natural point at which to truncate (in $|l|$) the matrix and to control the corrections in an actual calculation.

The situation is dramatically changed once (4.1) are transcribed to equations for the dressed coefficients of Eq. (3.6). These are

$$\sum_{l=-\infty}^{\infty} M^D_{m,l} \alpha_0(l) = \mu^B(m)\ ,$$
$$\sum_{l=-\infty}^{\infty} N^D_{m,l} \gamma_1(l) = \nu^B(m)\ , \tag{4.5}$$

where the l-dressed matrices $\underline{M}^D$ and $\underline{N}^D$ are

$$M^D_{m,l} = M^B_{m,l} \exp[iW_0(l)g]\ ,$$
$$N^D_{m,l} = M^B_{m,l} \exp[iW_1(l)g]\ . \tag{4.6}$$

By the same procedure which led to (4.3) and (4.4), we obtain

$$\lim_{\substack{|l|\to\infty \\ m \text{ fixed}}} M^D_{m,l} \sim \lim_{|l|\to\infty} J_l(i\beta|l|)e^{-\beta|l|} \sim e^{|l|\phi(\beta)}\ , \tag{4.7}$$

where

$$\phi(\beta) = \eta(\beta) - \beta\ . \tag{4.8}$$

The central point of this section is to recognize that $\phi(\beta)$ is negative for *all* values of β (see Fig. 3). Moreover, since

$$\phi(\beta) \sim \begin{cases} 1 + \ln\beta & \text{for } \beta \ll 1 \\ -\dfrac{1}{2\beta} & \text{for } \beta \gg 1, \end{cases} \tag{4.9}$$

we have an estimate for the dimension N of the effective $\underline{M}^D$ matrix (comprised of the significant matrix elements in $\underline{M}^D$). For $\beta \gg 1$, Eq. (4.9) yields

$$N \approx 2(2\beta) = 8\pi g/d\ , \tag{4.10}$$

where the extra factor of 2 in (4.10) is added to account for positive as well as negative l values. The estimated N in (4.10) is roughly the number of significant terms in the

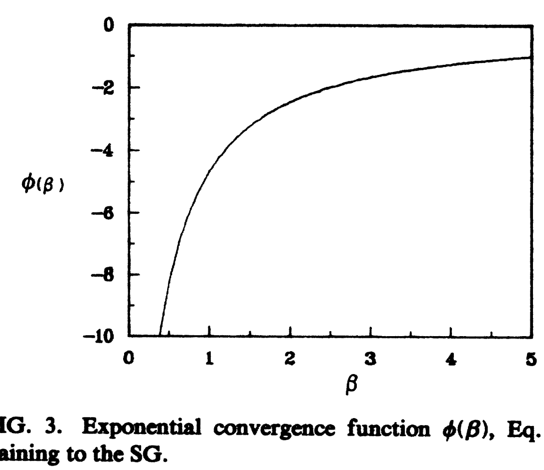

FIG. 3. Exponential convergence function $\phi(\beta)$, Eq. (4.8), pertaining to the SG.

dressed Rayleigh expansion, and provides the natural truncation point of $\underline{M}^D$ in an actual calculation. Thus, for instance, when $g/d=1$ the inversion of a 50×50 matrix should be quite adequate.

To complete the analysis, we now consider the convergence of (4.5) for the case $|m|\to\infty$ and l is fixed. Obviously, the m-convergence factor to be used is arbitrary. A convenient choice, symmetric to (4.6), is

$$\begin{aligned} M_{m,l}&=\exp\{i[W_1(m)+W_0(l)]g\}M^B_{m,l}\ ,\\ \mu(m)&=\exp[iW_1(m)g]\mu^B(m)\ ,\\ N_{m,l}&=\exp\{i[W_0(m)+W_1(l)]g\}N^B_{m,l}\ ,\\ \nu(m)&=\exp[iW_0(m)g]\nu^B(m)\ , \end{aligned} \tag{4.11}$$

and the fully dressed equations are

$$\begin{aligned} \sum_{l=-\infty}^{\infty} M_{m,l}\alpha_0(l)&=\mu(m)\ ,\\ \sum_{l=-\infty}^{\infty} N_{m,l}\gamma_1(l)&=\nu(m)\ . \end{aligned} \tag{4.12}$$

The set (4.12) is exponentially convergent to both the m and l "directions" for an *arbitrary* β. It is therefore a perfectly stable framework for calculations of deep SG. The size of the matrix needed to be inverted is of the order given by (4.10), which is well within the limits of a reasonable effort. The dressed coefficients $\alpha_0(l)$ and $\gamma_1(l)$ are then to be used in the dressed Rayleigh expansion (3.5) to obtain the total fields.

V. DISCUSSION AND SUMMARY

The dressed Rayleigh expansion, i.e., (3.1) with the substitution (3.5), takes the form (e.g., for the electrical field with the notation of Fig. 2)

$$\begin{aligned} \mathbf{E}_0(x,z)&=\sum_{l=-\infty}^{\infty}(\gamma_0(l)\hat{\mathbf{p}}_{0,-}(l)\exp\{i[k_lx-W_0(l)(z-g)]\}+\alpha_0(l)\hat{\mathbf{p}}_{0,+}(l)\exp\{i[k_lx+W_0(l)(z+g)]\})\ ,\\ \mathbf{E}_1(x,z)&=\sum_{l=-\infty}^{\infty}\gamma_1(l)\hat{\mathbf{p}}_{1,-}(l)\exp\{i[k_lx-W_1(l)(z-g)]\}\ , \end{aligned} \tag{5.1}$$

where

$$\gamma_0(l)=\delta_{l,0}\exp[-iW_0(l)g]E^-_{0p}\ ,$$

and the coefficients $\alpha_0(l)$ and $\gamma_1(l)$ satisfy (4.12). The convergence properties of (5.1) can be analyzed for a general grating shape along the lines applied for the SG in Sec. IV. All that is necessary are the exact expression of the corresponding $M^B_{m,l}$ and an examination of its asymptotic behavior. The explicit expression for $M^B_{m,l}$ pertaining to a grating of a general shape is known[16] (see also the appendix).

The comparison between the convergence properties of the dressed and bare Rayleigh expansion can be explicitly discussed in the context of the SG. These are determined by the asymptotic behavior of $M^B_{m,l}$ and $M^D_{m,l}$, respectively. With regard to the bare Rayleigh expansion (3.1), the asymptotic behavior is given by (4.3) and (4.4). Since $\eta(\beta)$ is monotonically increasing and negative for $\beta\ll 1$ (see Fig. 3), it follows that the maximum value of β for which $M^B_{m,l}$ is l convergent (exponentially) satisfies

$$\eta(\beta)=(1+\beta^2)^{1/2}+\ln\{\beta/[1+(1+\beta^2)^{1/2}]\}=0\ , \tag{5.2}$$

or equivalently, using the notation of Petit and Cadilhac,[24]

$$\frac{\theta+1}{\theta-1}=\exp\left[\frac{\theta^2+1}{2\theta}\right],\quad \beta=\frac{\theta^2-1}{2\theta}\ . \tag{5.3}$$

The solution to (5.2) or (5.3) is $\beta\approx 0.66$ ($\theta\approx 1.86$), or $g/d\approx 0.1$. This value is in agreement with the reported[15] maximum g/d ratio for which the bare Rayleigh expansion is found applicable. By contrast, the asymptotic behavior of $M^D_{m,l}$, Eqs. (4.7), yields convergence for all β! Note also that alternative schemes for solving the light-grating scattering problem[1,7,16] involve the bare coefficients $A_0(l)$ and $C_1(l)$. By dressing the coefficients, as articulated in (3.6), the convergence properties of these schemes are expected to improve. In this regard, the dressed Rayleigh expansion provides the simplest possible convergent scheme.

The value $\theta\approx 1.86$ deduced from (5.2) for a SG is somewhat larger than $\theta\approx 1.54$,[21,24] which is the rigorously proven threshold for a *perfect* metallic reflection SG ($\epsilon_1=-\infty$). This difference highlights the distinction between singular and nonsingular grating (i.e., smooth profile, finite ϵ_1). We can expect on physical grounds that when there is an infinite amount of charge at the surface (with a δ-function distribution), the point ($x=d$, $z=g$) separating the two adjacent selvedge domains $0\leq x\leq d$ and $d\leq x\leq 2d$ is singular (Fig. 2). Therefore, two such adjacent selvedge domains become "disconnected," the Rayleigh hypothesis is no longer exact and the convergence criterion is modified. The same conclusion follows also mathematically: when $\epsilon_1=-\infty$, the analysis leading to (4.3) and (4.4) is totally changed.

We comment now on a related convergence criterion[21] for the bare Rayleigh expansion, which is presumably valid for any dielectric constant ϵ_1. This criterion is based on the application of the steepest descent method[25] to an exact expression of $A_0(l)$ or $C_1(l)$. When applied to a SG, the maximum value of β for which there is convergence satisfies[24]

$$\frac{\theta+1}{\theta-1}=e^{\theta}\ , \tag{5.4}$$

where θ is defined in (5.3). We understood the difference between our result (5.3) and (5.4) in terms of the limited applicability of the steepest-descent method[25] for the SG case. Phrased differently, the corrections to the steepest-descent result for the SG case are not negligible, and hence it is not clear how stringent condition (5.4) really is. [The solution to (5.4) is $\theta\approx 1.54$. Numerically,[15] however, convergent results are obtained for higher values of θ up to about our result of $\theta\approx 1.86$.] This statement can be simply verified by examining the stationary-phase function $\psi(x)$ derivatives around the stationary points of z_s. We find that although $\psi''(z_s)<0$, the fourth derivative is positive and has a contribution larger than $\psi''(z_s)$.

In summary, we have analyzed the two questions pertaining to the Rayleigh expansion: the possibility to continue it to the grating's surface (the Rayleigh hypothesis) and its convergence properties. We identified a class of gratings, the SG being one example, for which the Rayleigh hypothesis is exact. We have also exposed a cause for the instability of the Rayleigh expansion when applied to a deep grating calculation. To remedy this instability, we propose a modified expansion, Eqs. (3.5), (3.6), and (5.1), dubbed the dressed Rayleigh expansion. The dressed expansion has presumingly excellent convergence properties for deep gratings. For the case of the SG, we explicitly show that the dressed expansion converges for an arbitrarily large value of β (or g/d ratio), and the associated matrix to be inverted is of the order $8\pi g/d$ [Eq. (4.10)]. Our analysis is valid for a frequency-dependent dielectric function since the whole discussion pertains to a fixed incoming-wave frequency (all Bragg scatterings are elastic). The dressed Rayleigh expansion is applied elsewhere[22] to study the very deep ($\beta\rightarrow\infty$) SG grating limit.

ACKNOWLEDGMENTS

We would like to thank M. Hutchinson and W. C. Murphy for stimulating discussions and criticism. We are grateful to A. C. Beri for his assistance in preparing the figures. This research was supported for the U.S. Office of Naval Research and the U.S. Air Force Office of Scientific Research, United States Air Force, under Contract No. F49620-86-C-009.

APPENDIX: EXACT EQUATIONS FOR THE (REFLECTIVITY) $A_0(l)$ AND (TRANSMITTIVITY) $C_1(l)$ COEFFICIENTS

The exact infinite set of coupled linear equations for the coefficients of expansion (3.1) has been derived by Toigo, Marvin, Celli, and Hill for a grating of an arbitrary shape.[16] We recap here their derivation for the SG for the sake of completeness and to unify notations.

We choose to match the tangential components of **B** and **E**. For this purpose we introduce the following triad of unit vectors which are normal and tangential to the SG profile given by $z=g\cos(k_Gx)$ (the notation of Fig. 2 is used throughout):

$$\mathbf{n}=-\hat{\mathbf{x}}\beta\sin(k_Gx)-\hat{\mathbf{z}},\quad \mathbf{t}=-\hat{\mathbf{x}}+\hat{\mathbf{z}}\beta\sin(k_Gx),\quad \mathbf{b}=\hat{\mathbf{s}}\ , \tag{A1}$$

where $\beta=k_Gg$, and the carats denote the unit vectors. The normal vector **n** and the tangential vector **t** need not be normalized for our purposes. By projecting the field (3.1) along **t** and **b** at the grating surface, the following exact equations are obtained:

$$\sum_{l=-\infty}^{\infty} e^{ik_lx}\left[A_0(l)\exp[iW_0(l)g\cos(k_Gx)]-\frac{k(1)}{k(0)}C_1(l)\exp[-iW_1(l)g\cos(k_Gx)]\right]=Q_b\ , \tag{A2a}$$

$$\begin{aligned}\sum_{l=-\infty}^{\infty} e^{ik_lx}\Bigg[&A_0(l)[W_0(l)+\beta k_l\sin(k_Gx)]\exp[iW_0(l)g\cos(k_Gx)]\\ &-\frac{k(0)}{k(1)}C_1(l)[-W_1(l)+\beta k_l\sin(k_Gx)]\exp[-iW_1(l)g\cos(k_Gx)\Bigg]=Q_t\ ,\end{aligned} \tag{A2b}$$

where

$$\begin{aligned}Q_b&=-\sum_{l=-\infty}^{\infty} C_1(l)\exp(ik_lx)\exp[-iW_0(l)g\cos(k_Gx)]\ ,\\ Q_t&=-\sum_{l=-\infty}^{\infty} C_1(l)[-W_0(l)+\beta k_l\sin(k_Gx)]\exp(ik_lx)\exp[-iW_0(l)g\cos(k_Gx)]\ .\end{aligned} \tag{A3}$$

Equations (A2) can be cast into a set of coupled linear equations for $A_0(l)$ and $C_1(l)$ by expanding all terms in the complete set $\{e^{ik_lx}\}$. It is remarkable that it is in fact possible to further eliminate exactly $A_0(l)$ or $C_1(l)$, thus breaking (A2) into two (infinite) sets of equations. This is achieved by first multiplying (A2a) with

$$M_b=[-W_0(m)+\beta k_m\sin(k_Gx)]\exp[-ik_mx+iW_0(m)g\cos(k_Gx)] \tag{A4a}$$

and multiplying (A2b) with

$$M_t = \exp[-ik_m x + iW_0(m) g \cos(k_G x)] \ , \tag{A4b}$$

adding the two equations and integrating over the interval $0 \leqq x \leqq d$. This manipulation yields the equations for $C_1(l)$. By repeating the same procedure with

$$M_b = \frac{k(0)}{k(1)}[-W_1(m) - \beta k_m \sin(k_G x)]\exp[-ik_m x - iW_1(m) g \cos(k_G x)] \tag{A5a}$$

and

$$M_t = -\frac{k(1)}{k(0)}\exp[-ik_m x - iW_1(m) g \cos(k_G x)] \ , \tag{A5b}$$

an equation for the $A_0(l)$ is obtained. In these manipulations for arbitrary a, b, d, and q, we have used[26]

$$\int_0^d dx\,[a + b\sin(k_G x)]\exp[i(k_l - k_m)x + iq\cos(k_G x)] = \left[a + \frac{b(k_l - k_m)}{qk_G}\right] d i^{m-l} J_{m-l}(q) \ , \tag{A6}$$

where J_n denotes the Bessel function of order n.

[1] *Electromagnetic Theory of Diffraction Gratings,* edited by R. Petit (Springer, New York, 1980).

[2] *Surface Polaritons,* edited by V. M. Agranovich and D. L. Mills (North-Holland, New York, 1982).

[3] D. L. Mills, Phys. Rev. B **15**, 3097 (1977).

[4] P. St. J. Russell, Phys. Rep. **71**, 209 (1980).

[5] B. Laks, D. L. Mills, and A. A. Maradudin, Phys. Rev. B **23**, 4965 (1981).

[6] N. E. Glass and A. A. Maradudin, Phys. Rev. B **24**, 595 (1981).

[7] P. Sheng, R. S. Stepleman, and P. N. Sanda, Phys. Rev. B **26**, 2907 (1982); K. T. Lee and T. F. George, *ibid.* **31**, 5106 (1985).

[8] M. Weber and D. L. Mills, Phys. Rev. B **27**, 2698 (1983).

[9] R. Reinisch and M. Neviere, Phys. Rev. B **28**, 1870 (1983).

[10] A. Wirgin and T. Lopez-Rios, Opt. Commun. **48**, 416 (1984); Solid State Commun. **52**, 197 (1984).

[11] N. Garcia, G. Diaz, J. J. Saenz, and C. Ocal, Surf. Sci. **143**, 338 (1984).

[12] See, for example, J. K. Powers and D. A. Bryan, Opt. Lett. **5**, 513 (1980).

[13] S. R. J. Brueck and D. J. Ehrlich, Phys. Rev. Lett. **48**, 1678 (1982).

[14] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Reinehart and Winston, New York, 1976), Chaps. 8—10.

[15] N. Garcia, V. Celli, N. R. Hill, and N. Cabrera, Phys. Rev. B **18**, 5184 (1978).

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

[17] J. L. Uretsky, Ann. Phys. (N.Y.) **33**, 400 (1965).

[18] Lord Rayleigh, *Theory of Sound,* 2nd ed. (Dover, New York, 1945), Vol. II, p. 89.

[19] W. C. Meecham, J. Appl. Phys. **27**, 361 (1956).

[20] J. E. Sipe, Surf. Sci. **105**, 489 (1981).

[21] N. R. Hill and V. Celli, Phys. Rev. B **17**, 2478 (1978).

[22] D. Agassi and T. F. George, Surf. Sci. (to be published).

[23] *Handbook of Mathematical Functions,* edited by M. Abramovitz and I. E. Stegun (National Bureau of Standards, Washington, D.C., 1964), p. 366.

[24] R. Petit and M. Cadilhac, C. R. Acad. Sci. Ser. B **262**, 468 (1966).

[25] M. Born and E. Wolf, *Principles of Optics,* 6th ed. (Pergamon, Oxford, 1984), p. 747.

[26] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965), p. 973.