Launching surface plasmon waves via vanishingly small periodic gratings

DAVID P. NICHOLLS,1* SANG-HYUN OH,2 TIMOTHY W. JOHNSON,2 AND FERNANDO REITICH3

1Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, Illinois 60607, USA
2Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota 55455, USA
3CAP S.A., Santiago, Chile
*Corresponding author: davidn@uic.edu

Received 20 October 2015; revised 30 December 2015; accepted 30 December 2015; posted 5 January 2016 (Doc. ID 251574); published 5 February 2016

The scattering of electromagnetic waves by periodic layered media plays a crucial role in many applications in optics and photonics, in particular in nanoplasmonics for topics as diverse as extraordinary optical transmission, photonic crystals, metamaterials, and surface plasmon resonance biosensing. With these applications in mind, we focus on surface plasmon resonances excited in the context of insulator–metal structures with a periodic, corrugated interface. The object of this contribution is to study the geometric limits required to generate these fundamentally important phenomena. For this we use the robust, rapid, and highly accurate field expansions method to investigate these delicate phenomena and demonstrate how very small perturbations (e.g., a 5 nm deviation on a 530 nm period grating) can generate strong (in this instance 20%) plasmonic absorption, and vanishingly small perturbations (e.g., a 1 nm deviation on a 530 nm period grating) can generate nontrivial (in this instance 1%) plasmonic absorption. © 2016 Optical Society of America

OCIS codes: (050.1755) Computational electromagnetic methods; (240.6680) Surface plasmons.

http://dx.doi.org/10.1364/JOSAA.33.000276

1. INTRODUCTION

The scattering of linear electromagnetic waves by periodic layered media plays a crucial role in many applications of scientific and engineering interest. In optics and photonics, for instance, this can be seen in nanoplasmonics [1–3], where one can investigate topics as diverse as extraordinary optical transmission [4,5], photonic crystals [6,7], metamaterials [8], surface acousto-optic systems [9,10], and surface plasmon resonance (SPR) biosensing [11–17].

The objects of our current investigation are the SPRs at the heart of a wide range of highly accurate and robust biosensing devices of current interest (see, e.g., [12]). These SPRs are generated when a surface plasmon polariton (an exponentially confined time-harmonic electromagnetic wave propagating at the interface between an insulator and a metal) is excited and coupled to the illuminating radiation.

A well-known property of these SPRs is that they cannot be excited at the boundary of an insulator–metal structure with a flat interface (though they can be launched in the Kretschmann configuration) [1–3]. It is not difficult to show that there is insufficient “momentum” to generate an SPR. With the addition of periodic corrugations at the interface, such momentum can be provided and SPRs can be excited. (Periodic gratings are among several methods of producing SPRs, and we refer the interested reader to [1–3] for other approaches.)

It is also known that these SPRs can be generated by quite small deformations. The object of this contribution is to study the limits of the size of the deformation required to generate surface plasmon waves. In concert with this, a remarkable feature of these resonances is that for a well-designed (i.e., the metal should be a superior conductor, e.g., having permittivity with a very negative real part) but fixed configuration, they are excited only for a very narrow band of illumination frequencies. Clearly, the ability to simulate such configurations numerically in a rapid, robust, and high-order fashion is of paramount importance.

While all of the classical numerical algorithms have been brought to bear upon this problem, each has shortcomings. Typically one considers methods based on finite elements (see, e.g., [18–21]) and finite differences (see, e.g., [22–24]). However, these volumetric approaches are clearly disadvantaged by an unnecessarily large number of unknowns for the problem at hand, which features piecewise-constant dielectric constants. Methods based on traditional integral equation (IE) formulations [25] are a natural candidate but face several challenges. While most have been adequately addressed with
appropriate quadrature rules and iteration procedures (see [26] and [27] for different strategies), the authors recently argued [17,28] that these methods are noncompetitive for the periodic, parameterized problems we consider as compared with "high-order perturbation of surfaces" (HOPS) algorithms, which we advocate here.

More specifically, in [17,28] we pointed out that standard IE methods have the following properties:

1. For periodic problems, the relevant Green's function must be periodized if one is to restrict the domain of integration to a single period cell. This is a well-known problem (see, e.g., the introduction of [29] for a full description), and the slow convergence of the periodization must be accelerated (e.g., with techniques such as Ewald summation). However, even with such technology, these IE methods demand an additional discretization parameter: the number of terms retained in the approximation of the periodized Green's function. (We note the recent results [30] and [31], which significantly ameliorate this concern.)

2. For configurations parameterized by the real value \( h \) (for us the height/slope of the crossed interface), an IE solver will return the scattering returns only for a particular value of \( h \). If this value is changed, then the solver must be run again.

3. The dense, nonsymmetric positive definite systems of linear equations, which must be inverted with each simulation.

In this contribution we utilize a particular HOPS approach: the method of field expansions (FE), which traces its roots to the low-order calculations of Rayleigh [32] and Rice [33]. Its high-order incarnation for doubly layered media was first introduced by Bruno and Reitich for the two-dimensional scalar case in [34,35] and for the fully three-dimensional vector Maxwell case in [36]. It was further enhanced and stabilized by Nicholls and Reitich [37–39], and expanded to multiple layers in the two-dimensional scalar case by Malcolm and Nicholls [40] and the fully three-dimensional vector electromagnetic case by Nicholls [28].

As we pointed out in [17,28], this formulation is particularly compelling, as it maintains the advantageous properties of classical IE formulations (e.g., surface formulation and exact enforcement of far-field conditions) while avoiding the shortcomings listed above:

1. As this HOPS scheme utilizes the eigenfunctions of the Laplacian (suitable complex exponentials) on a periodic domain, the quasiperiodicity of solutions is built in and does not need to be further approximated.

2. Since the method is built upon expansions in the boundary parameter, \( b \), once the Taylor coefficients are known for the scattering quantities, it is simply a matter of summarizing these (rather than beginning a new simulation) for any given choice of \( b \) to recover the returns.

3. Due to the perturbative nature of the scheme, at every perturbation order one need only invert a single sparse operator corresponding to the flat-interface approximation of the problem.

With this method we can show that vanishingly small-amplitude (e.g., a 5 nm deviation on a 530 nm period grating) perturbations can give rise to substantial plasmonic absorption on the order of a 20% dip from the base, flat-interface return. This is of particular relevance to experimentalists, since the roughness of as-deposited metal films can be on the order of 5 nm, unless ultraflat metal films are produced using techniques such as template stripping [41].

From here the paper is organized as follows: In Section 2 we briefly recall the equations that govern the propagation of electromagnetic waves in a periodic structure that is invariant in one direction, rendering this problem two-dimensional. In Section 3 we specify the method of FE for numerically approximating solutions to these governing equations [28,34–36,40]. In Section 4 we make a simple explanation of the SPR phenomena in terms of the FE framework introduced in Section 3. In Section 5 we discuss our numerical simulations, including the implementation (Section 5.A) and results for vanishingly small perturbations (Section 5.B).

2. GOVERNING EQUATIONS

In Fig. 1 we show the geometry of the configuration we consider: A \( y \)-invariant doubly layered insulator–metal structure. The insulator (vacuum with refractive index \( n^{(i)} = 1 \)) occupies the domain \( \{ z > g(x) \} \), and the metal (with index of refraction \( n^{(m)} \)) fills \( \{ z < g(x) \} \).

The grating is \( d \)-periodic so that \( g(x + d) = g(x) \). The structure is illuminated from above by monochromatic plane-wave incidence of frequency \( \omega \), aligned with the grooves

\[
\mathbf{E}(x, z, t) = \mathbf{A}e^{i \omega t - j k z}, \quad \mathbf{H}(x, z, t) = \mathbf{B}e^{i \omega t - j k z}.
\]

We consider the reduced electric and magnetic fields

\[
\mathbf{E}(x, z) = e^{i \alpha z} \mathbf{E}, \quad \mathbf{H}(x, z) = e^{i \alpha z} \mathbf{H},
\]

which, like the reduced scattered fields, are \( \alpha \)-quasiperiodic due to the incident radiation. Finally, the scattered radiation must be "outgoing" (upward propagating in \( S^{(a)} \) and downward propagating in \( S^{(d)} \)).

As shown in Petit [42], in this two-dimensional setting, the time-harmonic Maxwell equations decouple into two scalar...
Helmholtz problems that govern the transverse electric (TE) and transverse magnetic (TM) polarizations. We denote the invariant (\(y\)) directions of the scattered electric and magnetic fields by

\[
u = u(x, z), \quad w = w(x, z),
\]

in \(S^{(u)}\) and \(S^{(w)}\), respectively. The incident radiation in the upper layer is specified by \(u'\).

In light of all of this, we are led to seek outgoing, \(\alpha\)-quasiperiodic solutions of

\[
\Delta u + \left( k^{(u)} \right)^2 u = 0, \quad z > g(x), \quad (2.1a)
\]

\[
\Delta w + \left( k^{(w)} \right)^2 w = 0, \quad z < g(x), \quad (2.1b)
\]

\[
\frac{\partial}{\partial z} u - \tau^2 \frac{\partial}{\partial z} w = \psi, \quad z = g(x), \quad (2.1c)
\]

where \(k^{(m)} = n^{(m)} \alpha / c\), \(N = (-\partial_{g}, 1)^T\), the Dirichlet and Neumann data are

\[
\zeta(x) := -u'(x, g(x)) = -e^{i\pi \gamma \alpha} \rho_g(x),
\]

\[
\psi(x) := - (\frac{\partial}{\partial z} u')(x, g(x)) = (i\gamma^{(u)} + i\alpha(\partial_{g})) e^{i\pi \gamma \alpha} \rho_g(x),
\]

and

\[
\tau^2 = \left\{ \begin{array}{ll} 1 & \text{TE} \\ \left( \frac{k^{(u)}}{k^{(w)}} \right)^2 = \left( \frac{n^{(u)}}{n^{(w)}} \right)^2 & \text{TM}. \end{array} \right.
\]

Appealing to the classical study of SPRs [1], we restrict our attention to TM polarization from here.

**A. Rayleigh Expansions**

Separation of variables gives the Rayleigh expansions [42], which are quasiperiodic, outgoing solutions of Eqs. (2.1a) and (2.1b). The electric fields can be written

\[
u(x, z) = \sum_{p = -\infty}^{\infty} \hat{\alpha}_p e^{i\phi^{(z)}_p} e^{i\gamma^{(z)}_p \rho_g(x)} z, \quad (2.2a)
\]

\[
w(x, z) = \sum_{p = -\infty}^{\infty} \hat{\gamma}_p e^{i\phi^{(z)}_p} e^{-i\gamma^{(z)}_p \rho_g(x)} z, \quad (2.2b)
\]

where, for \(p \in \mathbb{Z} \) and \(m = u, w\),

\[
a_p := a + \left( \frac{2\pi}{\alpha} \right) p, \quad \gamma^{(m)}_p := \begin{cases} \sqrt{\left( k^{(m)} \right)^2 - \alpha^2} & p \in \mathcal{U}^{(m)} \\ \frac{\gamma^{(m)}_p}{\left( k^{(m)} \right)^2} - \alpha^2 & p \notin \mathcal{U}^{(m)} \end{cases},
\]

and

\[
\mathcal{U}^{(m)} = \{ p \in \mathbb{Z} | \alpha^2_p < \left( k^{(m)} \right)^2 \},
\]

which are the “propagating modes” in the upper and lower layers. We point out that \(\hat{\alpha}_p\) and \(\hat{\gamma}_p\) are the upward and downward propagating Rayleigh amplitudes. Quantities of great interest are the efficiencies

\[
e_p^{(m)} = \left( \frac{\gamma^{(m)}_p}{\gamma^{(m)}_p} \right) |\hat{\alpha}_p|^2, \quad e_w^{(m)} = \left( \frac{\gamma^{(m)}_p}{\gamma^{(m)}_p} \right) |\hat{\gamma}_p|^2,
\]

and the objects of fundamental importance to the design of SPR biosensors [12–17] are the “reflectivity map” and “normalized reflectivity,”

\[
R := \sum_{p \in \mathcal{E}^{(m)}} \varepsilon_p^{(m)}, \quad B(\lambda, \hat{b}) := \frac{\varepsilon_p^{(m)}(\lambda, \hat{b})}{\varepsilon_0^{(m)}(\lambda, 0)}, \quad (2.3)
\]

respectively. If the lower layer is filled with a perfect electric conductor, then if \(S^{(w)}\) contains an insulator (such as the vacuum), conservation of energy requires that \(R = 1\). This is not the case for a metal (such as gold or silver) in the lower domain, and drops in its value to a tenth or even a hundredth are the fundamental phenomena behind the utility of these sensors.

**3. FIELD EXPANSIONS**

The method of FE [34–36] is a perturbative approach to enforcing the boundary conditions in Eqs. (2.1c) and (2.1d) with the \(\{\hat{\alpha}_p, \hat{\gamma}_p\}\) from the Rayleigh expansions in Eq. (2.2) as unknowns. Here we take the point of view advocated by one of the authors in [28], which we believe, simplifies the presentation. First, we define the functions

\[
a(x) := u(x, 0) = \sum_{p = -\infty}^{\infty} \hat{\alpha}_p e^{i\phi^{(x)}_p} \exp(-i\gamma^{(x)}_p x), \quad (3.1a)
\]

\[
d(x) := w(x, 0) = \sum_{p = -\infty}^{\infty} \hat{\gamma}_p e^{i\phi^{(x)}_p} \exp(-i\gamma^{(x)}_p x), \quad (3.1b)
\]

which are the “flat interface” field traces.

We further define the Dirichlet trace operators

\[
\mathcal{D}^{(u)}: a \rightarrow u(x, g(x)), \quad \mathcal{D}^{(w)}: d \rightarrow w(x, g(x)),
\]

and their Neumann counterparts

\[
\mathcal{N}^{(u)}: a \rightarrow (\partial_{z} u - (\partial_{g}) u)(x, g(x)), \quad \mathcal{N}^{(w)}: d \rightarrow (\partial_{z} w - (\partial_{g}) w)(x, g(x)).
\]

The idea behind these operators \(\mathcal{D}\) and \(\mathcal{N}\) is that they map, respectively, the function pair \((a, d)\) to the upper and lower Dirichlet and Neumann traces. It can be shown that

\[
\mathcal{D}^{(u)} = \exp(g(\hat{\gamma}^{(u)}_D)), \quad \mathcal{D}^{(w)} = \exp(-i\hat{\gamma}^{(w)}_D)),
\]

and

\[
\mathcal{N}^{(u)} = \exp(g(\hat{\gamma}^{(u)}_D)) (i\hat{\gamma}^{(u)}_D) - (\partial_{g}) \exp(g(\hat{\gamma}^{(u)}_D)) \partial_{g}, \quad \mathcal{N}^{(w)} = \exp(-i\hat{\gamma}^{(w)}_D)) (i\hat{\gamma}^{(w)}_D) - (\partial_{g}) \exp(-i\hat{\gamma}^{(w)}_D)) \partial_{g},
\]

where we have used Fourier multiplier notation, e.g.,

\[
m(D) \xi(x) := \sum_{p = -\infty}^{\infty} m(p) \hat{\xi}_p \exp(i\phi^{(x)}_p),
\]

where \(\hat{\xi}_p\) is the generalized \(p\)th Fourier coefficient of \(\xi(x)\).

In these terms, the Dirichlet boundary condition, Eq. (2.1c), becomes

\[
\mathcal{D}^{(u)}[d] - \mathcal{D}^{(w)}[d] = \zeta, \quad (3.2)
\]

while the Neumann condition, Eq. (2.1d), becomes

\[
\mathcal{N}^{(u)}[d] - \tau^2 \mathcal{N}^{(w)}[d] = \psi. \quad (3.3)
\]

We state the boundary conditions [Eqs. (3.2) and (3.3)] abstractly as

\[
\mathbf{Mv} = \mathbf{b}, \quad (3.4)
\]

where
\[ M = \begin{pmatrix} D_0^{(a)} & D_0^{(w)} \\ \Lambda^{(a)} & -\tau^2 \Lambda^{(w)} \end{pmatrix}, \quad v = \begin{pmatrix} a \\ d \end{pmatrix}, \quad b = \begin{pmatrix} \zeta \\ \psi \end{pmatrix}. \]

A. Taylor Expansions

The FE approach to this problem is to consider deformations of the form \( g(x) = hf(x) \) where \( f = O(1) \) and note that for a sufficiently smooth \( f \) (Lipschitz) and sufficiently small \( h \), the linear operator \( M \) and inhomogeneity \( b \) are both analytic in \( h \) [43,44]. Furthermore, an analytic solution \( v \) can be shown to exist. More specifically, the following expansions can be demonstrated to be strongly convergent:

\[ \{M, v, b\}(h^p) = \sum_{n=0}^{\infty} \{M_n(f), v_n(f), b_n(f)\}h^n. \]

Crucially, an algorithm for recovering \( v_n \) can be devised based on regular perturbation theory. In short, we write Eq. (3.4) as

\[ \left( \sum_{n=0}^{\infty} M_n h^n \right) \left( \sum_{n=0}^{\infty} v_n h^n \right) = \sum_{n=0}^{\infty} b_n h^n, \]

and, equating at each perturbation order, we find

\[ M_0 v_n = b_n - \sum_{\ell=1}^{n-1} M_{n-\ell} v_{\ell}. \quad (3.5) \]

At order zero we recover the flat-interface solution, giving the Fresnel coefficients, while higher-order corrections, \( v_n \), can be computed by appealing to Eq. (3.5). Of great importance is the fact that one only need invert the same linear operator, \( M_0 \), at every perturbation order. All that remains is a specification of the terms \( \{M_{\ell}, b_{\ell}\} \).

Regarding the Dirichlet trace operators, upon defining

\[ F_a(x) := f(x) e^{i\tau x}/m, \]

one can show that

\[ D_0^{(a)} = F_a(i\tau_D^{(a)})^n, \quad D_0^{(w)} = F_a(-i\tau_D^{(w)})^n. \]

For their Neumann counterparts, we have

\[ N_0^{(a)} = F_a(i\tau_D^{(a)})^{-n+1} - (\partial_x f) F_{a-1} \partial_x (i\tau_D^{(a)})^{-n-1}, \quad N_0^{(w)} = F_a(-i\tau_D^{(w)})^{-n+1} - (\partial_x f) F_{a-1} \partial_x (-i\tau_D^{(w)})^{-n-1}. \]

Finally, for the surface data, \( b_n \), it is easy to show that

\[ \zeta_n = -F_a(-i\tau_D^{(a)})^n e^{i\tau x}, \quad \varphi_n = F_a(i\tau_D^{(a)})^{-n} e^{i\tau x} + (\partial_x f) F_{a-1} (i\alpha (i\tau_D^{(a)})^{-n-1} e^{i\tau x}, \]

where \( F_{-1}(x) \equiv 0 \) and \( F_0(x) \equiv 1 \).

4. PERSISTENCE OF SPRS AT VANISHINGLY SMALL AMPLITUDE

The SPR phenomena have been observed since the time of Wood [45], and explanations for the onset of these surface waves for extremely narrow bands of illumination wavelengths can be traced to the foundational work of Rayleigh [32] (see the fascinating article of Maystre in Chapter 1 of [3]). This is all discussed in a number of sources including the books of Raether [1], Maier [2], and Novotny and Hecht [3]. Here we recover these results in the language of the boundary formulation [Eq. (3.4)] and the FE approach.

The onset of an SPR is indicated by a precipitous drop in the normalized reflectivity, \( B \) [Eq. (2.3)], as a function of \( \lambda \). In the current formulation,

\[ B(\lambda, h) = 1 + B_1(\lambda) h + B_2(\lambda) h^2 + O(h^3), \]

where

\[ B_1 = \frac{\hat{a}_{0,1}(\alpha_0 + \hat{a}_{0,0} \alpha_1)}{|\hat{a}_{0,0}|^2}, \]

and

\[ B_2 = \frac{\hat{a}_{0,2}(\alpha_0 + \hat{a}_{0,0} \alpha_1) + \hat{a}_{0,0} \alpha_2}{|\hat{a}_{0,0}|^2}. \]

We will now demonstrate that \( B_1 \equiv 0 \), while \( B_2 \) takes on very large negative values at an SPR.

To simplify the presentation, we use the notation

\[ U := (i\tau_D^{(a)}), \quad W := (-i\tau_D^{(w)}), \quad Z := (-i\tau_D^{(w)}), \]

and further fix upon normally incident illumination so that \( \alpha = 0 \). With these, we have

\[ \text{D}_n^{(a)} = F_a U^n, \quad \text{D}_n^{(w)} = F_a W^n, \]

\[ \text{N}_n^{(a)} = F_a U^{n-1} - (\partial_x f) F_{a-1} \partial_x U^{n-1}, \]

\[ \text{N}_n^{(w)} = F_a W^{n-1} - (\partial_x f) F_{a-1} \partial_x W^{n-1}, \]

\[ \zeta_n = -F_n Z^n, \quad \varphi_n = -F_n Z^{n+1}, \]

which gives

\[ M_0 = \begin{pmatrix} I \\ -I \end{pmatrix}, \quad b_0 = -\begin{pmatrix} 1 \\ Z \end{pmatrix}, \]

and

\[ M_n = \begin{pmatrix} U^n \\ -W^n \end{pmatrix}, \quad b_n = -F_n Z^n \begin{pmatrix} 1 \\ Z \end{pmatrix}. \]

For future reference, it can be shown that

\[ M_0 = \Delta^{-1} \begin{pmatrix} -\tau^2 W & I \\ -U & I \end{pmatrix}, \]

where \( \Delta := U - \tau^2 W \), and

\[ \Delta^{-1}[e^{i\tilde{p}}] = \frac{1}{\Delta} e^{i\tilde{p}}, \]

where \( \tilde{p} := i\tau_D^{(a)} + i\tau_D^{(w)} \) and \( \tilde{p} := (2\pi/d)p \).

A. Order Zero: Fresnel Coefficients

At order zero, Eq. (3.4) delivers

\[ \begin{pmatrix} a_0 \\ d_0 \end{pmatrix} = v_0 = M_0 b_0 = -\Delta^{-1} \begin{pmatrix} -\tau^2 W & I \\ -U & I \end{pmatrix} \begin{pmatrix} 1 \\ Z \end{pmatrix}, \]

\[ = -\Delta^{-1} \begin{pmatrix} -\tau^2 W[1] + Z \\ -U[1] + Z \end{pmatrix}. \]
Recalling that
\[ U[1] = (i\gamma^{(a)}), \quad W[1] = (-i\gamma^{(a)}), \]
we find the solutions
\[ a_0(x, z) = R e^{i\gamma^{(a)}z}, \quad d_0(x, z) = T e^{-i\gamma^{(a)}z}, \]  
with the Fresnel (reflection and transmission) coefficients
\[ R = \frac{\gamma^{(a)} - i\gamma^{(a)}}{\Delta_0}, \quad T = \frac{2i\gamma^{(a)}}{\Delta_0}. \]
There is no appreciable variation in \( R \) as \( \lambda \) is varied; one requires \( f \neq 0 \) to find an SPR.

**B. Order One: \( B_1 \)**

At order one, Eq. (3.4) becomes
\[ M_0 v_1 = b_1 - M_1 v_0 = S_1, \]
where
\[ b_1 = -f Z \left( \frac{1}{Z} \right), \]
and
\[ M_1 = f \begin{pmatrix} U & -W \\ U^2 & -\tau^2 W^2 \end{pmatrix} - (\partial_x f) \partial_x \begin{pmatrix} 0 \\ I \end{pmatrix}, \]
so that
\[ S_1 = f \begin{pmatrix} (1 - R) (i\gamma^{(a)}) - T (i\gamma^{(a)}) \\ -(1 + R) (i\gamma^{(a)})^2 + T^2 \end{pmatrix}. \]
We seek solutions of the form
\[ a_1(x) = \sum_{p=-\infty}^{\infty} \hat{a}_{1,p} e^{ipx}, \quad d_1(x) = \sum_{p=-\infty}^{\infty} \hat{d}_{1,p} e^{ipx}, \]
which has the solution
\[ \left( \begin{array}{c} \hat{a}_{1,p} \\ \hat{d}_{1,p} \end{array} \right) = M_0^{-1} (S_1) = \hat{f}_{p} \Delta_{2,p}. \]
It is not difficult to show that
\[ \hat{R}_{p} = -\tau^2 (i\gamma^{(a)})^2 (Z) Z - (\tau^2 (i\gamma^{(a)})^2 + i\gamma^{(a)}) (i\gamma^{(a)}) R + \hat{T}_{p} = (i\gamma^{(a)})^2 (Z) Z + (i\gamma^{(a)}) - i\gamma^{(a)}) (i\gamma^{(a)}) R \]
and
\[ \hat{T}_{p} = (i\gamma^{(a)})^2 (Z) Z + (i\gamma^{(a)}) - i\gamma^{(a)}) (i\gamma^{(a)}) R \]

Importantly, since the mean of \( f \) is zero, we have \( \hat{f}_{0} = 0 \), so that \( \hat{a}_{1,0} = \hat{d}_{1,0} = 0 \), which implies that
\[ B_1 \equiv 0. \]
We must move to order two to see the SPR.

**C. Order Two: \( B_2 \)**

At order two, Eq. (3.4) becomes
\[ M_0 v_2 = b_2 - M_1 v_1 - M_2 v_0 = S_2, \]
where
\[ b_2 = -\frac{1}{2} f^2 Z \left( \frac{1}{Z} \right), \]
and
\[ M_2 = \frac{1}{2} f^2 \begin{pmatrix} U^2 & -W^2 \\ U^3 & -\tau^2 W^2 \end{pmatrix} - (\partial_x f) \partial_x \begin{pmatrix} 0 \\ U \end{pmatrix} \]
Again, we seek solutions of the form
\[ a_2(x) = \sum_{p=-\infty}^{\infty} \hat{a}_{2,p} e^{ipx}, \quad d_2(x) = \sum_{p=-\infty}^{\infty} \hat{d}_{2,p} e^{ipx}, \]
which can be solved to yield
\[ \left( \begin{array}{c} \hat{a}_{2,p} \\ \hat{d}_{2,p} \end{array} \right) = M_0^{-1} (S_2) = \sum_{q=-\infty}^{\infty} \hat{f}_{p-q} \hat{f}_{q} \left( \begin{array}{c} \hat{R}_{p-q} \\ \hat{L}_{p-q} \end{array} \right). \]
Unfortunately, the forms for \( \{\hat{R}_{p,q}, \hat{L}_{p,q}\} \) are rather unwieldy and we do not report them here. However, there are a few things one can say.

Our goal is to investigate the nature of the SPR and the role played by the classic condition [1-3]
\[ \tilde{\Delta}_{1} \approx 0. \]
For this we focus on \( S_2 \), and since the determinant of \( M_0 \) at wavenumber \( p = 0 \) (\( \tilde{\Delta}_{0} \)) is of order one, we seek near-singularities in this right-hand side to generate the dramatic SPR response. We note that \( b_2 \) and \( M_2 v_0 \) are also of order one throughout the range of \( \lambda \), and thus we focus on \( M_1 v_1 \). In Fourier space it can be shown that
\[ M_1^{-1} (S_1) \]
where
\[ \hat{\xi}_{p} = (i\gamma^{(a)}) \hat{R}_{p} + (i\gamma^{(a)}) \hat{T}_{p}, \]
\[ \hat{\nu}_{p} = (i\gamma^{(a)})^2 \hat{R}_{p} + \hat{T}_{p}, \]
\[ \hat{\kappa}_{p} := -(i\gamma^{(a)}) \hat{R}_{p} + (i\gamma^{(a)}) \hat{T}_{p}, \]
Importantly, while \( \hat{f}_{0} = 0 \), \( \hat{f}_{1} \neq 0 \) generically, so that contributions to \( \{\hat{M}_1 v_1\}_1 \) will come from \( \hat{\xi}_{1}, \hat{\nu}_{1}, \hat{\kappa}_{1}, \) and \( \hat{\xi}_{1}, \hat{\nu}_{1}, \hat{\kappa}_{1} \).

We summarize our findings:

1. In Fig. 2 we see that the denominator of \( (M_1^{-1} v_1) \), \( \tilde{\Delta}_{1} \), approaches zero very rapidly as \( \lambda \) tends towards the SPR value \( \lambda_{SPR} \approx 557.4 \) nm, while being far from zero near the “passing-off value” (which we denote as the Rayleigh value \( \lambda_{Ray} = 650 \) nm).
2. In Fig. 3 we plot three components of the numerator of \( (M_1^{-1} v_1) \),
\[ \{ |\hat{\xi}_{1}|^2, |\hat{\nu}_{1}|^2, |\hat{\kappa}_{1}|^2 \}, \]
versus \( \lambda \). Two of these tend to zero as \( \lambda \) approaches \( \lambda_{Ray} \), while all three vary quite continuously (and none become particularly large) as \( \lambda \) changes.
3. By contrast, in Fig. 4 we display three components of \( (M_1^{-1} v_1) \),
\[ \{ |\hat{\xi}_{1}|^2, |\hat{\nu}_{1}|^2, |\hat{\kappa}_{1}|^2 \}, \]
versus \( \lambda \). Two of these tend to zero as \( \lambda \) approaches \( \lambda_{SPR} \), but all three become anomalously large as \( \lambda \) approaches \( \lambda_{SPR} \). From this it becomes
clear that the only reason for the SPR in terms of these equations is the near-zero value of $\hat{\Delta}_1$ at $\lambda_{\text{SPR}}$.

While it is difficult to explain all of the constituent parts of $\hat{a}_{2,0}$, and therefore $B_2$, we can say that it depends linearly upon 
\[ \{\hat{\xi}/\hat{\Delta}_1, \hat{\nu}_1/\hat{\Delta}_1, \hat{k}_1/\hat{\Delta}_1\}. \]

In this way we see quite explicitly how at very low (second) order, the SPR condition $\hat{\Delta}_1 \approx 0$ can generate its remarkably strong and specific effect.

5. NUMERICAL RESULTS

We are now in a position to explore the conclusions above for the case of a very small perturbation of an ultraflat silver interface. In these we investigate deformations with a size on the order of the roughness of as-deposited metal films (approximately 5 nm on a grating of period 530 nm) in the absence of specialized deposition techniques [41].

A. Numerical Implementation

The method described in Section 3 is essentially a Fourier collocation [46]/Taylor method [37,47] enhanced by Padé approximation [44,48]. More specifically, we approximate the fields $\{u, w\}$ by

\begin{equation}
\begin{aligned}
u_{N,0}^N &= ∑_{p=0}^{N/2} a_{n,p} e^{i\gamma x} e^{i\Delta_1 x} \mu^p, \\
u_{N,0}^N &= ∑_{p=0}^{N/2} a_{n,p} e^{i\gamma x} e^{i\Delta_1 x} \mu^p, \\
u_{N,0}^N &= ∑_{p=0}^{N/2} a_{n,p} e^{i\gamma x} e^{i\Delta_1 x} \mu^p.
\end{aligned}
\end{equation}

[cf. Eq. (2.2)]. We insert these into Eq. (3.5) and determine $\{\nu_n\}$.

A crucial consideration is how the Taylor series in $h$ are summed. To be specific, to approximate $u$ we consider the truncation $u_{N,0}^N$, which amounts to the approximation $\hat{u}_p(b) := ∑_{n=0}^{\infty} a_{n,p} b^n$ by $\hat{u}_N(b) := ∑_{n=0}^{N} a_{n,p} b^n$. The classical numerical analytic continuation technique of Padé approximation [48] has been successfully brought to bear upon HOPS methods in the past (see, e.g., [28,35,44]), and we advocate its use here. Padé approximation seeks to simulate the truncated Taylor series $\hat{u}_N(b)$ by the rational function

\begin{equation}
Λ(b) := A^L(b) B^M(b) := ∑_{m=0}^{L} A_m b^m \Big/ ∑_{m=1}^{M} B_m b^m, \tag{5.2}
\end{equation}

where $L + M = N$ and
\( [L/M](b) = \tilde{A}_F^N(h) + \mathcal{O}(b^{L+M+1}) \);

well-known formulas for the coefficients \( \{A_F, B_m\} \) can be found in [48]. This approximant has remarkable properties of enhanced convergence, and we refer the interested reader to Section 2.2 of Baker and Graves-Morris [48] and the insightful calculations of Section 8.3 of Bender and Orszag [49] for a thorough discussion of the capabilities and limitations of Padé approximants.

**B. Small Perturbations**

We now display results that show that SPRs can be launched for surface profiles with amplitudes on the order of the roughness of as-deposited metal films. To illustrate this, we consider a doubly layered structure of vacuum (an insulator) above silver (a metal), separated by the profile \( g(x) = h f(x) \).

To begin, we consider \( f(x) = -(1/2) \cos(2\pi x/d) \); see Fig. 5. By definition, the refractive index of vacuum is \( n^v = 1 \), while the refractive index of silver is still the subject of current research. For this we use a Lorenz model [50],

\[
e^{i\Delta g} = e^{i\Delta g^v} + \sum_{j=1}^{6} a_j e^{i\Delta g^s} \alpha_j^s + b_j e^{i\Delta g^s} \sigma_j^s + c_j e^{i\Delta g^s},
\]

where \( \omega = 2\pi/\lambda, e^{i\Delta g} = 1 \), and \( \Delta g^v, \Delta g^s, b J^s, \) and \( c_j \) can be found in [50]. For physical and numerical parameters we choose

\[
\alpha = 0, \quad \gamma = (\gamma^s, \gamma^s)^T, \quad (5.3a)
\]

\[
b = 0, \ldots, 5, \quad d = 530, \quad (5.3b)
\]

\[
N_x = 32, \quad N = 2. \quad (5.3c)
\]

In Fig. 6 we display the normalized reflectivity, \( B(\lambda, h) \) [cf. Eq. (2.3)], for this configuration, which shows a strong 13% plasmonic response for a perturbation of size 5 nm (<1% of \( d \)).

We also consider the “lamellar” profile (see Fig. 1),

\[
f(x) = \frac{1}{2} \left\{ \tanh \left[ \sigma \left( x - \frac{d}{2} \right) + \frac{W}{2} \right] - \tanh \left[ \sigma \left( x - \frac{d}{2} \right) - \frac{W}{2} \right] \right\};
\]

in this, \( W \) is the linewidth and \( \sigma \) measures the “steepness” of the line. For instance, in Fig. 1 we chose \( \epsilon = 50, d = 530 \) nm, and \( W = d/2 \). For physical and numerical parameters we choose

\[
\alpha = 0, \quad \gamma = (\gamma^s, \gamma^s)^T, \quad (5.4a)
\]

\[
b = 0, \ldots, 5, \quad d = 530, \quad W = d/2, \quad (5.4b)
\]

\[
N_x = 32, \quad N = 2. \quad (5.4c)
\]
In Fig. 8 we display the squared norm of the field $|u|^2$ in the lamellar configuration [Eq. (5.4)] at $\lambda = 557.4 \approx \lambda_{\text{SPR}}$, with $N_x = 32, N = 2$.

In Table 1 we report the numerical values of $B(\lambda, h)$ for $\lambda = 557.4 \approx \lambda_{\text{SPR}}$ for the cosine and lamellar profiles of period $d = 530$ nm.

Table 1. Numerical Values of the Normalized Reflectivity $B$ [Eq. (2.3)] for $\lambda = 557.4$ nm $\approx \lambda_{\text{SPR}}$ for the Cosine and Lamellar Profiles of Period $d = 530$ nm

<table>
<thead>
<tr>
<th>$b$ (nm)</th>
<th>Cosine Profile</th>
<th>Lamellar Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0.994839</td>
<td>0.992012</td>
</tr>
<tr>
<td>2</td>
<td>0.979436</td>
<td>0.968234</td>
</tr>
<tr>
<td>3</td>
<td>0.954031</td>
<td>0.929244</td>
</tr>
<tr>
<td>4</td>
<td>0.919026</td>
<td>0.876019</td>
</tr>
<tr>
<td>5</td>
<td>0.874981</td>
<td>0.809930</td>
</tr>
</tbody>
</table>

In Fig. 9 we display the normalized reflectivity $R(\lambda, h)$ [cf. Eq. (2.3)] versus incident wavelength $\lambda$ and deformation height $h$. Results for the lamellar configuration [Eq. (5.4)] with grating period $d = 785$ nm and $N_x = 32, N = 2$.

In Table 2 we report the numerical values of $B(\lambda, h)$ for $\lambda = 682.8$ nm $\approx \lambda_{\text{SPR}}$ for the cosine and lamellar profiles of period $d = 633$ nm.

Table 2. Numerical Values of the Normalized Reflectivity $B$ [Eq. (2.3)] for $\lambda = 682.8$ nm $\approx \lambda_{\text{SPR}}$ for the Cosine and Lamellar Profiles of Period $d = 633$ nm

<table>
<thead>
<tr>
<th>$b$ (nm)</th>
<th>Cosine Profile</th>
<th>Lamellar Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0.994977</td>
<td>0.992139</td>
</tr>
<tr>
<td>2</td>
<td>0.979984</td>
<td>0.968749</td>
</tr>
<tr>
<td>3</td>
<td>0.955248</td>
<td>0.930404</td>
</tr>
<tr>
<td>4</td>
<td>0.921149</td>
<td>0.878058</td>
</tr>
<tr>
<td>5</td>
<td>0.878219</td>
<td>0.813045</td>
</tr>
</tbody>
</table>

In Fig. 7 we display the normalized reflectivity, $B(\lambda, h)$ [cf. Eq. (2.3)], for this configuration, which shows a strong 20% plasmonic response for a perturbation of size 5 nm (<1% of $d$).

In Table 3 we report the numerical values of $B(\lambda, h)$ for $\lambda = 801.0$ nm $\approx \lambda_{\text{SPR}}$ for the cosine and lamellar profiles of period $d = 785$ nm.

Table 3. Numerical Values of the Normalized Reflectivity $B$ [Eq. (2.3)] for $\lambda = 801.0$ nm $\approx \lambda_{\text{SPR}}$ for the Cosine and Lamellar Profiles of Period $d = 785$ nm

<table>
<thead>
<tr>
<th>$b$ (nm)</th>
<th>Cosine Profile</th>
<th>Lamellar Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0.995426</td>
<td>0.992743</td>
</tr>
<tr>
<td>2</td>
<td>0.981766</td>
<td>0.971126</td>
</tr>
<tr>
<td>3</td>
<td>0.959209</td>
<td>0.935622</td>
</tr>
<tr>
<td>4</td>
<td>0.92807</td>
<td>0.887027</td>
</tr>
<tr>
<td>5</td>
<td>0.88791</td>
<td>0.82646</td>
</tr>
</tbody>
</table>

6. CONCLUSION

In this contribution we studied the limits of the surface deformation required to generate SPRs. We explicitly identified the crucial role that the SPR condition, $\Delta_{\lambda} \approx 0$, plays in this phenomena. Additionally, we used the robust, rapid, and highly accurate FE method to investigate this delicate phenomenon, which is difficult to perform using conventional finite difference time domain or finite element methods due to the low accuracy and exorbitant computational cost of these algorithms. We demonstrated how very small perturbations (e.g., a 5 nm deviation on a 530 nm period grating) can generate strong (in this instance 20%) plasmonic absorption, and vanishingly small perturbations (e.g., a 1 nm deviation on a 530 nm period grating) can generate nontrivial (in this instance 1%) plasmonic absorption. Our findings support the contention that ultrasmooth metal surfaces (roughness on the order of 1 nm or below) can improve the performance of plasmonic...
resonators, and provide the means to quantitatively explore such experiments.

**Funding.** National Science Foundation (NSF) (DMS-1115333, DMS-1522548, CMMI-1363334).

**Acknowledgment.** D.P.N. gratefully acknowledges support from the National Science Foundation through grants DMS-1115333 and DMS-1522548. S.-H.O. and T.W.J. acknowledge support from the National Science Foundation through an NSF CAREER award and grant CMMI-1363334.

**REFERENCES**