Effects of spatial dispersion on electromagnetic surface modes and on modes associated with a gap between two half spaces

Bo E. Sernelius*
Department of Physics and Measurement Technology, Linköping University, SE-581 83 Linköping, Sweden
(Received 20 December 2004; revised manuscript received 24 March 2005; published 27 June 2005)

We study the effects of spatial dispersion on the conditions for having electromagnetic normal modes at a single interface and in the presence of two interfaces. The first result has bearing on the dispersion of surface modes, like surface plasmons, and the second result on the van der Waals and Casimir force between two half-spaces.

DOI: 10.1103/PhysRevB.71.235114 PACS number(s): 71.10.−w, 68.90.+g, 71.36.+c, 41.20.−q

I. INTRODUCTION

In the most elementary textbooks on electromagnetism the dielectric function is treated as a constant. This works well in many situations. The Fresnel coefficients describing the reflection and refraction of light in a glass prism can be calculated quite accurately. It works less well in other situations. With such a treatment the group and phase velocities would be equal, there would be no so-called surface modes[1] and the van der Waals attraction between objects would not exist. To incorporate these effects the frequency dependence of the dielectric function has to be taken into account. Expressed in another way, temporal dispersion has to be included. Now, the dielectric function depends on the momentum, also—there is spatial dispersion. When an electromagnetic wave impinges on an interface, frequency is conserved. This means that the Fresnel coefficients can be used to determine the resulting waves even if the dielectric functions on the two sides of the interface are frequency dependent. The momentum on the other hand is not conserved at the interface. This means that the Fresnel coefficients can no longer be used for systems where the spatial dispersion is important. Fortunately, spatial dispersion is not important in most situations.

The first mention of spatial dispersion was already in 1811 when Arago[2] discovered the rotatory power of quartz. Spatial dispersion has since become a whole research field of its own. The interested reader is referred to a rather recent book: *Spatial Dispersion in Solids and Plasmas*,[3] edited by P. Halevi, which is fully devoted to this topic. There are also two chapters in the book: *Surface Polaritons*,[4] edited by V. M. Agronovich and D. L. Mills which are of interest here: one by Lagois and Fischer[5] and one by Agronovich.[6] The spatial dispersion problem for metal surfaces has been widely studied in the past[7,8] as have the nonlocal effects on the surface-normal modes, the surface-plasmon polaritons.[9,10]

In the present work we are concerned with metals and, in particular, the matching of fields at metal-vacuum interfaces. For this problem spatial dispersion may be important in some frequency regions. These regions are regions where the dielectric function starts to have an important momentum dependence already at small momentum. The anomalous skin effect in the microwave region is such a manifestation of spatial dispersion. Here the single-particle continuum comes close to the frequency axis in the ωq plane. Another region where spatial dispersion may be important is near the plasma frequency. A third region is near the threshold for interband transitions. This region will not be covered in the numerical calculations performed in this work. We use model dielectric functions that do not include contributions from interband transitions.

Our main purpose in this work is to find out how spatial dispersion affects the Casimir force between real metal plates. There have been very resent publications[11,12] on the effects of spatial dispersion on the Casimir force. The second publication studies the nonlocal effects in much detail, but unfortunately the authors did not consider finite temperatures.

There are two quite different predictions in the literature regarding the room-temperature Casimir force for large separations. In one version[13] the result is equal to the perfect-metal result; in the other version[14–16] the result is one-half of that result. The reason is that one type of mode does not contribute in this particular limit. The result is very sensitive to the behavior of the transverse dielectric function in the zero-frequency limit;[1] thus, the matching conditions of the fields in the low-frequency limit are crucial, and this is one of the regions where spatial dispersion can have important effects.

The material is outlined in the following way. In Sec. II we introduce the formalism and in Sec. III we study electromagnetic normal modes, surface modes, at a single interface separating two materials. Section IV is devoted to modes in the vacuum gap between two half-spaces and in Sec. V we calculate the force between two gold plates. Finally, a summary and conclusions are found in Sec. VI.

II. THE FORMALISM

In general, the sources for electromagnetic fields are charge and current densities. One method[1] to find the electromagnetic normal modes of a system is the following. Start with a reasonable guess for a set of external charge and current densities that obey the equation of continuity; find the resulting electromagnetic fields; find the proper combination of ω and $\mathbf{q}$ for which these fields can survive, even if the amplitudes of the external densities go toward zero. If one

succeeds one has found the dispersion curve of a normal mode. Here we start from valid (obeying the equation of continuity) charge and current densities at the interface, find the resulting fields, use boundary conditions that force the densities to be induced and not external. This leads to two solutions: one is the trivial solution when the densities and fields are all zero; the other solution is one where the densities and fields are in harmony with each other—we have self-sustained fields or equivalently self-sustained, induced charge and current densities.

Let us study the interface between two materials *1* and *2*; medium *1* to the left and *2* to the right. We use, consistently, the idealization that the interface is perfectly sharp; the dielectric function on each side is represented by the bulk function all the way up to the interface; all potentials on either side of the interface are screened by the corresponding dielectric function. This means that the charge and current densities that produce the self-sustained fields, the key element of an electromagnetic normal mode, are two-dimensional charge and current densities at the plane of the interface. Keeping strictly to this idealization, we do not need to introduce any so-called ABCs (*Additional Boundary Conditions*).[17] The ABC problem, i.e., that there are discrepancies between the results obtained from using different ABCs, has been eliminated before with the proper choice of surface potential.[18]

The charge and current densities can then be divided into two classes: the strict two-dimensional charge and current densities located at the interface, ρ and $\mathbf{J}$, respectively; the induced charge and current densities in the bulk of the materials on either side of the interface. Each of these separately obeys the equation of continuity. We let the xy plane coincide with the interface and let the z direction point to the right. To get the fields in material *1*, i.e., to the left of the interface, we assume that these are generated by charge and current densities at the position of the interface, and as if the whole space were filled with medium *1*. To get the fields in material *2*, i.e., to the right of the interface we again assume that these are generated by charge and current densities at the position of the interface, and as if the whole space were filled with medium *2*. These two sets of charge and current densities need not be the same. This is in analogy with the mirror-charge formalism, but we address the real surface-charge densities instead of the artificial equivalent mirror charges. The equation of continuity demands the presence of accompanying surface-current densities.

The modes at an interface are characterized by the 2D (two-dimensional) wave vector, $\mathbf{k}$, in the plane of the interface. We let a general wave vector be denoted by $\mathbf{q}$, and $\mathbf{k}$ is then the in-plane component. We neglect any effects from imperfections of the interface. This means that the in-plane momentum is conserved. Thus $\mathbf{k}$ is a good quantum number for the modes. For isotropic materials there is no preferred direction in the xy plane. We arbitrarily choose the propagation of the mode to be in the x direction.

In the Coulomb gauge, the surface charge and current densities give rise to the following potentials for nonmagnetic materials:[1]

$$\Phi^{(i)}(\mathbf{q},\omega)=\frac{4\pi\rho^{(i)}(\mathbf{q},\omega)}{\varepsilon_L^{(i)}(\mathbf{q},\omega)q^2}; \tag{1}$$

$$\mathbf{A}^{(i)}(\mathbf{q},\omega)=\frac{4\pi\mathbf{J}_\perp^{(i)}(\mathbf{q},\omega)/c}{q^2-(\omega/c)^2\varepsilon_\perp^{(i)}(\mathbf{q},\omega)}, \tag{2}$$

where the index $i=1,2$ specifies the medium. In the Coulomb gauge the vector potential, $\mathbf{A}$, is transverse and depends only on the transverse part of the current density. Note that the transverse part of the current is orthogonal to $\mathbf{q}$ and not necessarily to $\mathbf{k}$.

These charge and current densities are strictly 2D. Let us study a 2D Fourier component of the surface-charge density propagating in the x direction:

$$\begin{aligned}\rho(\mathbf{r},t)&=\rho(k,0,-;\omega)\cdot e^{i(kx-\omega t)}\delta(z)\\&=\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\rho(k,0,-;\omega)\cdot e^{i(kx+0y+q_zz-\omega t)}.\end{aligned} \tag{3}$$

It has a spatial variation in the x direction, no variation in the y direction, and is a delta function in the z direction. A Fourier transform of a general 2D charge density at the interface depends on the momentum components in the x and y directions but not on the momentum component in the z direction.

A time-varying charge density is associated with a current density via the equation of continuity:

$$\omega\rho(\mathbf{q},\omega)-\mathbf{q}\mathbf{J}(\mathbf{q},\omega)=0. \tag{4}$$

In our case we have

$$\begin{aligned}&\omega\rho(k,0,-;\omega)-(k\hat{x}+q_z\hat{z})[J_x(k,0,q_z;\omega)\hat{x}+J_y(k,0,q_z;\omega)\hat{y}]\\&\quad=0,\end{aligned} \tag{5}$$

which is reduced into

$$\omega\rho(k,0,-;\omega)-kJ_x(k,0,q_z,\omega)=0. \tag{6}$$

Thus the charge density couples to a current density in the x direction only. So a surface-normal mode described by our chosen surface-charge density has the following surface-current density:

$$\mathbf{J}(k,q_y,q_z;\omega)=\hat{x}\rho(k,0,-;\omega)\frac{\omega}{k}\delta_{q_y,0}. \tag{7}$$

We will find, below that the electric fields associated with this mode are in the plane of incidence (xz plane) and the magnetic fields are perpendicular to this plane. The fields are p polarized or TM fields (transverse magnetic). One could be led to believe that there are also modes associated with a $\mathbf{J}_y$ component and with no accompanying induced charge density. However, there are no such solutions to the Maxwell equations whose resulting fields obey the boundary conditions. We will show later that in the geometry with two parallel interfaces there are modes generated by coupled $\mathbf{J}_y$ components at the two interfaces. These modes are TE modes with s-polarized fields.

Let us now return to our mode. The charge density gives rise to a scalar potential and the current density to a vector potential. Since we are working in the Coulomb gauge, only the transverse part of the current contributes to the vector potential. To get the transverse part we subtract the longitudinal part from the current density,

$$\mathbf{J}_\perp = \mathbf{J} - \left(\frac{\mathbf{q}}{q}\mathbf{J}\right)\frac{\mathbf{q}}{q} = J[(q^2-k^2)\hat{x} - kq_z\hat{z}]\frac{1}{q^2} = J(q_z^2\hat{x} - kq_z\hat{z})\frac{1}{q^2}. \tag{8}$$

Let us, in what follows, drop the superscript representing the medium, on the potentials and fields. The result is valid on each side of the interface, we just add the proper superscript at the end. We have

$$\Phi(q_x,q_y,q_z;\omega) = \frac{4\pi\rho(k,0,-;\omega)}{\varepsilon_L(q,\omega)q^2}\delta_{q_y,0}; \tag{9}$$

$$\begin{aligned}\mathbf{A}(q_x,q_y,q_z;\omega) &= \frac{4\pi J(q_z^2\hat{x} - kq_z\hat{z})}{cq^2[q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2]} \\ &= \frac{4\pi\rho(k,0,-;\omega)\omega(q_z^2\hat{x} - kq_z\hat{z})}{cq^2k[q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2]}\delta_{q_y,0}.\end{aligned} \tag{10}$$

We are interested in the electric and magnetic fields associated with the charge and current densities. The electric field is

$$\mathbf{E} = -\nabla\Phi - \frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}, \tag{11}$$

and its Fourier transform is

$$\begin{aligned}\mathbf{E}(k,q_y,q_z;\omega) &= -i\mathbf{q}\Phi(k,q_y,q_z;\omega) - \frac{-i\omega}{c}\mathbf{A}(k,q_y,q_z;\omega) \\ &= -i\delta_{q_y,0}4\pi\rho(k,0,-;\omega)\left\{\hat{x}k\left[\frac{1}{q^2\varepsilon_L(q,\omega)}\right.\right. \\ &\quad - \frac{(\omega/c)^2}{k^2[q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2]} \\ &\quad \left.\left. + \frac{(\omega/c)^2}{q^2[q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2]}\right] + \hat{z}\left[\frac{q_z}{q^2\varepsilon_L(q,\omega)}\right.\right. \\ &\quad \left.\left. + \frac{q_z(\omega/c)^2}{q^2[q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2]}\right]\right\}.\end{aligned} \tag{12}$$

Similarly, the magnetic induction is

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{13}$$

and its Fourier transform is

$$\begin{aligned}\mathbf{B}(q_x,q_y,q_z;\omega) &= i\mathbf{q}\times\mathbf{A}(q_x,q_y,q_z;\omega) = i(k\hat{x} + q_z\hat{z}) \\ &\quad \times \frac{4\pi\rho(k,0,-;\omega)\omega(q_z^2\hat{x} - kq_z\hat{z})}{cq^2k[q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2]}\delta_{q_y,0} \\ &= \hat{y}i\delta_{q_y,0}\frac{4\pi\rho(k,0,-;\omega)\omega}{cq^2k[q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2]}[k^2q_z \\ &\quad + q_zq_z^2] \\ &= \hat{y}i\delta_{q_y,0}\frac{4\pi\rho(k,0,-;\omega)\omega}{ck}\frac{q_z}{q^2 - \varepsilon_\perp(q,\omega)(\omega/c)^2}.\end{aligned} \tag{14}$$

Thus, each of the two components of the **E** field has a longitudinal and a transverse part. The **B** field has only a transverse part.

Many functions will appear repeatedly when we apply the boundary conditions and we give them names to make the expressions simpler. We need the following functions:

$$g_a^{(i)}(k,\omega) = 2k\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{q^2\varepsilon_L^{(i)}(q,\omega)}; \tag{15}$$

$$g_b^{(i)}(k,\omega) = 2\gamma\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{[q^2 - \varepsilon_\perp^{(i)}(q,\omega)(\omega/c)^2]}; \tag{16}$$

$$g_c^{(i)}(k,\omega) = \frac{2(\omega/c)^2k\gamma}{k-\gamma}\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{q^2[q^2 - \varepsilon_\perp^{(i)}(q,\omega)(\omega/c)^2]}; \tag{17}$$

$$g_d^{(i)}(k,\omega) = -2i\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{q_ze^{iq_z0^+}}{[q^2 - \varepsilon_\perp^{(i)}(q,\omega)(\omega/c)^2]}, \tag{18}$$

where

$$\gamma = \gamma(k,\omega) = \sqrt{k^2 - (\omega/c)^2}. \tag{19}$$

The prefactors have been chosen such that all g functions are unity in vacuum:

$$g_a^0(k,\omega) = 2k\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{q^2} = 1; \tag{20}$$

$$g_b^0(k,\omega) = 2\gamma\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{[q^2 - (\omega/c)^2]} = 1; \tag{21}$$

$$g_c^0(k,\omega) = \frac{2(\omega/c)^2k\gamma}{k-\gamma}\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{q^2[q^2 - (\omega/c)^2]} = 1; \tag{22}$$

$$g_d^0(k,\omega) = -2i\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{q_ze^{iq_z0^+}}{[q^2 - (\omega/c)^2]} = 1. \tag{23}$$

These results are easily found from the standard integrals:

$$\int_{-\infty}^{\infty}\frac{dr}{2\pi}e^{irs}\frac{1}{r^2+a^2} = \frac{1}{2a}e^{-a|s|}; \tag{24}$$

$$\int_{-\infty}^{\infty}\frac{dr}{2\pi}e^{irs}\frac{r}{r^2+a^2} = \frac{i}{2}\,\mathrm{sign}(s)e^{-a|s|}. \tag{25}$$

We also need the g functions with the neglect of spatial dispersion. If we neglect spatial dispersion, indicated by an extra subscript 1, we have

$$g_{a,1}^{(i)}(k,\omega) = 2k\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{q^2\varepsilon_L^{(i)}(\omega)} = \frac{1}{\varepsilon_L^{(i)}(\omega)}; \tag{26}$$

$$g_{b,1}^{(i)}(k,\omega) = 2\gamma\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{[q^2 - \varepsilon_\perp^{(i)}(\omega)(\omega/c)^2]} = \frac{\gamma}{\gamma^{(i)}}; \tag{27}$$

$$g_{c,1}^{(i)}(k,\omega) = \frac{2(\omega/c)^2 k\gamma}{k-\gamma}\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{1}{q^2[q^2-\varepsilon_\perp^{(i)}(\omega)(\omega/c)^2]} = \frac{\gamma(\omega/c)^2}{\gamma^{(i)}(k-\gamma)(k+\gamma^{(i)})}; \quad (28)$$

$$g_{d,1}^{(i)}(k,\omega) = -2i\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{q_z e^{iq_z 0^+}}{[q^2-\varepsilon_\perp^{(i)}(\omega)(\omega/c)^2]} = 1, \quad (29)$$

where

$$\gamma^{(i)} = \gamma^{(i)}(k,\omega) = \sqrt{k^2+\varepsilon_\perp^{(i)}(\omega/c)^2}. \quad (30)$$

The g_d functions are calculated just to the right of the interface and have all the value unity. The corresponding value to the left of the interface is minus unity.

We will furthermore need the following combinations of g functions:

$$G^{(i),TM}(k,\omega) = \frac{k}{\gamma}g_a^{(i)}(k,\omega) - \frac{(\omega/c)^2}{\gamma^2}g_b^{(i)}(k,\omega) + \frac{k(k-\gamma)}{\gamma^2}g_c^{(i)}(k,\omega), \quad (31)$$

and

$$G^{(i),TE}(k,\omega) = g_b^{(i)}(k,\omega). \quad (32)$$

The overall scaling of these functions has been chosen such that the functions are unity in vacuum, i.e.,

$$G^{(0),TM}(k,\omega) = 1; G^{(0),TE}(k,\omega) = 1. \quad (33)$$

If we neglect spatial dispersion, these functions become

$$G_1^{(i),TM}(k,\omega) = \frac{k}{\gamma}g_{a,1}^{(i)}(k,\omega) + \frac{k(k-\gamma)}{\gamma^2}g_{c,1}^{(i)}(k,\omega) - \frac{(\omega/c)^2}{\gamma^2}g_{b,1}^{(i)}(k,\omega) = \frac{k}{\gamma}\frac{1}{\varepsilon_L^{(i)}(\omega)} + \frac{k(\omega/c)^2}{\gamma\gamma^{(i)}(k+\gamma^{(i)})} - \frac{(\omega/c)^2}{\gamma\gamma^{(i)}} = \frac{\gamma^{(i)}}{\gamma}\frac{1}{\varepsilon^{(i)}(\omega)}, \quad (34)$$

and

$$G_1^{(i),TE}(k,\omega) = g_{b,1}^{(i)}(k,\omega) = \frac{\gamma}{\gamma^{(i)}}, \quad (35)$$

respectively. In the first function we have used the fact that the longitudinal and transverse dielectric functions are equal in the limit of vanishing momentum.

We will use three different model-dielectric functions. Sometimes it is useful to express the functions in terms of the following dimensionless variables: $Q=q/2k_F$, $K=k/2k_F$, $W=\hbar\omega/4E_F$, $W_{pl}=\hbar\omega_{pl}/4E_F$, and $y=m_e e^2/\hbar^2 k_F$.

The first and simplest of the dielectric functions is the so-called hydrodynamic dielectric function,[19]

$$\varepsilon(q,\omega) = 1-\omega_{pl}^2/(\omega^2-a^2q^4), \quad (36)$$

or with $a=\hbar/2m_e$,

$$\varepsilon(Q,W) = 1-W_{pl}^2/(W^2-Q^4). \quad (37)$$

Here we have neglected any dissipation.

The second dielectric function is in the so-called plasmon-pole approximation,[20]

$$\varepsilon(q,\omega) = 1-\omega_{pl}^2/\{\omega^2-(4E_F\hbar)^2[(q2k_F)^2 3+(q2k_F)^4]\}, \quad (38)$$

and in terms of the dimensionless variables,

$$\varepsilon(Q,W) = 1-W_{pl}^2/\{W^2-[Q^2 3+Q^4]\}. \quad (39)$$

The RPA dielectric functions, on the real frequency axis, expressed in the dimensionless variables are

$$\varepsilon_{L,\perp}(Q,W) = 1+\alpha_{L,\perp}(Q,W), \quad (40)$$

where the longitudinal and transverse polarizabilities are[21]

$$\alpha_L(Q,W) = \alpha_L^0(Q,W) = \frac{y}{2\pi}\frac{1}{Q^2}\left\{1+\frac{Q^2-(W-Q^2)^2}{4Q^3}\times\ln\left[\frac{Q-(W-Q^2)}{-Q-(W-Q^2)}\right]-\frac{Q^2-(W+Q^2)^2}{4Q^3}\times\ln\left[\frac{-Q+(W+Q^2)}{Q+(W+Q^2)}\right]\right\}, \quad (41)$$

and

$$\alpha_\perp(Q,W) = \alpha_\perp^0(Q,W) = \frac{y}{2\pi}\frac{1}{Q^2}\frac{1}{4W^2}\left\{-(Q^4+Q^2+3W^2)+\frac{[Q^2-(W-Q^2)^2]^2}{4Q^3}\ln\left[\frac{Q-(W-Q^2)}{-Q-(W-Q^2)}\right]-\frac{[Q^2-(W+Q^2)^2]^2}{4Q^3}\ln\left[\frac{-Q+(W+Q^2)}{Q+(W+Q^2)}\right]\right\}, \quad (42)$$

respectively. The logarithm in these expressions is taken from the branch for which $|\arg[\ln(z)]|<\pi$.

These results are in neglect of dissipation. Including dissipation, or damping, in a simple way leads to the following modifications:[22]

$$\alpha_L(Q,W,\Delta) = \frac{(W+i\Delta)\alpha_L^0(Q,W+i\Delta)}{W+i\Delta[\alpha_L^0(Q,W+i\Delta)/\alpha_L^0(Q,0)]}, \quad (43)$$

and[23]

$$\alpha_\perp(Q,W,\Delta) = \frac{W+i\Delta}{W}\alpha_\perp^0(Q,W+i\Delta), \quad (44)$$

respectively.

For the force calculations we need the polarizabilities on the imaginary frequency axis. There they are

$$\alpha_L'^0(Q,W)=\frac{y}{2\pi Q^2}\left\{1+\frac{(W^2+Q^2-Q^4)}{4Q^3}\times\ln\left[\frac{(Q+Q^2)^2+W^2}{(Q-Q^2)^2+W^2}\right]-\frac{W}{Q}\left[\tan^{-1}\left(\frac{Q+Q^2}{W}\right)+\tan^{-1}\left(\frac{Q-Q^2}{W}\right)\right]\right\},\tag{45}$$

and

$$\alpha_\perp'^0(Q,W)=\frac{1}{8}\frac{y}{\pi Q^2W^2}\left\{-(3W^2-Q^2-Q^4)+\frac{(2WQ^2)^2-(W^2+Q^2-Q^4)^2}{4Q^3}\times\ln\left[\frac{(Q+Q^2)^2+W^2}{(Q-Q^2)^2+W^2}\right]+\frac{2W(W^2+Q^2-Q^4)}{Q}\times\left[\tan^{-1}\left(\frac{Q+Q^2}{W}\right)+\tan^{-1}\left(\frac{Q-Q^2}{W}\right)\right]\right\},\tag{46}$$

respectively. The inverse tangent functions are taken from the branch where their absolute values are less than $\pi/2$.

Including dissipation we now have

$$\alpha_L'(Q,W,\Delta)=\frac{(W+\Delta)\alpha_L'^0(Q,W+\Delta)}{W+\Delta[\alpha_L'^0(Q,W+\Delta)/\alpha_L'^0(Q,0)]},\tag{47}$$

and

$$\alpha_\perp'(Q,W,\Delta)=\frac{W+\Delta}{W}\alpha_\perp'^0(Q,W+\Delta),\tag{48}$$

respectively.

We are now done with the introduction of the formalism and will in what follows use this material to find the electromagnetic normal modes. We have two momentum scales in the present problem; one is defined by the Fermi momentum, k_F; one by the surface mode frequency divided by the speed of light, ω_s/c. These scales are quite different and $\omega_s/c \ll k_F$. Different effects are visible on the two scales and one has to view them separately. The first scale is important if we are interested in the surface energy or the interaction between objects at small separations, the van der Waals force. The second scale is important for the interaction at larger separations, in the Casimir range. Retardation effects, effects from the finite speed of light, enter for very small in-plane momenta and should be viewed on the second scale.

III. MODES AT A SINGLE INTERFACE

We now derive the condition for having a surface-normal mode from the standard boundary conditions for the fields at the interface: the continuity of the in-plane components of the **E** and **H** fields and the normal components of the **D** and **B** fields; we only need the first two. These boundary conditions are valid if there are no external charge or current densities at the interface, only induced densities from the self-sustained fields.

We need the $\mathbf{E}_x$ and $\mathbf{H}_y$ components on the two sides of the interface. We have

$$\mathbf{E}_x(\mathbf{r},t)=\mathbf{E}_x(x,z,t)=-i4\pi\rho(k,0,-;\omega)ke^{i(kx-\omega t)}\times\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{iq_zz}\left[\frac{1}{q^2\varepsilon_L(q,\omega)}-\frac{(\omega/c)^2}{k^2[q^2-\varepsilon_\perp(q,\omega)(\omega/c)^2]}+\frac{(\omega/c)^2}{q^2[q^2-\varepsilon_\perp(q,\omega)(\omega/c)^2]}\right].\tag{49}$$

Just to the left of the interface we have

$$\mathbf{E}_x(x,0^-,t)=-i2\pi\rho^{(1)}(k,0,-;\omega)e^{i(kx-\omega t)}\times\left[g_a^{(1)}(k,\omega)-\frac{(\omega/c)^2}{\gamma k}g_b^{(1)}(k,\omega)+\frac{(k-\gamma)}{\gamma}g_c^{(1)}(k,\omega)\right]=-2\pi\rho^{(1)}(k,0,-;\omega)e^{i(kx-\omega t)}\frac{\gamma}{k}G^{(1),TM}(k,\omega),\tag{50}$$

and just to the right we have

$$\mathbf{E}_x(x,0^+,t)=-i2\pi\rho^{(2)}(k,0,-;\omega)e^{i(kx-\omega t)}\frac{\gamma}{k}G^{(2),TM}(k,\omega).\tag{51}$$

Furthermore, we need the $\mathbf{H}_y$ component on the two sides of the interface. The general expression is

$$\mathbf{H}_y(\mathbf{r},t)=\mathbf{B}_y(\mathbf{r},t)=B_y(x,z,t)=i\frac{4\pi\rho(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)}\times\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{iq_zz}\frac{q_z}{q^2-\varepsilon_\perp(q,\omega)(\omega/c)^2}.\tag{52}$$

Just to the left of the interface it is

$$\mathbf{H}_y(x,0^-,t)=\frac{2\pi\rho^{(1)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)}g_d^{(1)}(k,\omega),\tag{53}$$

and to the right it is

$$\mathbf{H}_y(x,0^+,t)=-\frac{2\pi\rho^{(2)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)}g_d^{(2)}(k,\omega).\tag{54}$$

To make the expressions for the fields more compact we introduce the common factors A and B, not to be confused with the vector potential and magnetic induction,

$$A=\frac{2\pi\rho^{(1)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)};\qquad B=\frac{2\pi\rho^{(2)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)}.\tag{55}$$

Then the continuity of the $\mathbf{H}_y$ component gives

$$g_d^{(1)}(k,\omega)A = -g_d^{(2)}(k,\omega)B \rightarrow A + B = 0, \tag{56}$$

since the g_d functions are unity. The continuity of the $\mathbf{E}_x$ component gives

$$-\frac{ic\gamma}{\omega}G^{(1),TM}(k,\omega)A + \frac{ic\gamma}{\omega}G^{(2),TM}(k,\omega)B = 0, \tag{57}$$

or

$$G^{(1),TM}(k,\omega)A - G^{(2),TM}(k,\omega)B = 0. \tag{58}$$

Thus, we have obtained the following system of equations:

$$\begin{pmatrix} 1 & 1 \\ G^{(1),TM}(k,\omega) & -G^{(2),TM}(k,\omega) \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} = 0. \tag{59}$$

This system has the trivial solution that $A=B=0$. It has also nontrivial solutions which are the modes we are looking for. The condition for normal modes is found as

$$\begin{vmatrix} 1 & 1 \\ G^{(1),TM}(k,\omega) & -G^{(2),TM}(k,\omega) \end{vmatrix} = 0, \tag{60}$$

or

$$G^{(1),TM}(k,\omega) + G^{(2),TM}(k,\omega) = 0. \tag{61}$$

Neglecting spatial dispersion, we arrive at

$$\gamma^{(1)}\varepsilon^{(2)}(\omega) + \gamma^{(2)}\varepsilon^{(1)}(\omega) = 0, \tag{62}$$

which is the standard condition for having a normal mode in the neglect of spatial dispersion.[1]

We have now derived the general conditions for having surface-normal-modes at an interface between two media with and without spatial dispersion taken into account. Here we will study the effect from spatial dispersion on the surface-plasmon dispersion, i.e., the overall effect on the k_F scale, and the effect on the surface-plasmon-polariton dispersion, i.e., on the ω_s/c scale.

A. Overall effect of spatial dispersion

In finding the overall effect, which we do first, we may neglect the retardation effects. We let medium *1* be a metal and medium *2* vacuum. In neglect of retardation effects the condition for modes is

$$g_a^{(1)}(k,\omega) + g_a^0(k,\omega) = 0, \tag{63}$$

or

$$g_a^{(1)}(k,\omega) + 1 = 0, \tag{64}$$

or

$$\tilde{g}_a^{(1)}(k,\omega) + 2 = 0. \tag{65}$$

Similar relations have been obtained earlier.[24]

The tilde over a g function means the g function calculated inside the medium minus the corresponding function calculated in vacuum. The tilde versions of the g functions were introduced since they represent faster converging integrals. To benefit from this the two integrals in each tilde version of a g function should be combined into one.

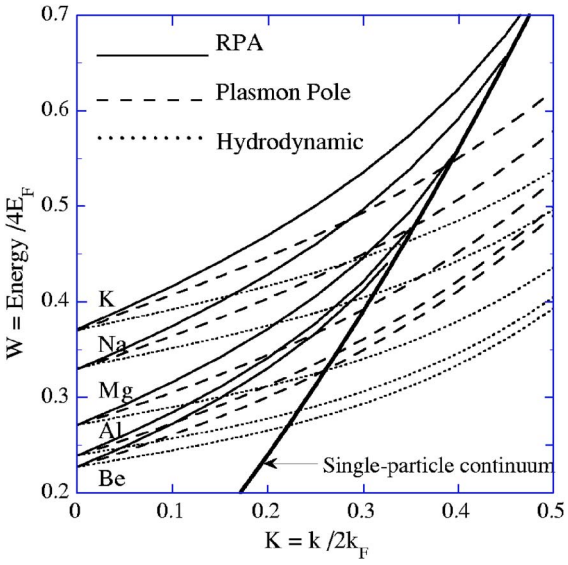

FIG. 1. The surface-plasmon dispersion curves for the metals K, Na, Mg, Al, and Be as a result from taking spatial dispersion into account. The dotted curves are for the hydrodynamic approximation, the dashed curves are for the plasmon-pole approximation and the solid curves for RPA. Neglecting spatial dispersion would produce dispersionless horizontal curves at the values where the present curves start out from the frequency axis. The thick solid curve is the boundary of the single-particle continuum in the in-plane direction of the surface plasmon.

We find the surface-plasmon modes for simple metals for the three different dielectric functions. For the two first approximations, the g_a function can be obtained analytically. In the hydrodynamic approximation the g_a function has the form

$$g_a(k,\omega) = \frac{1}{\omega^2 - \omega_{pl}^2}\left[\omega^2 - \frac{1}{2}\omega_{pl}^2\sqrt{ak^2}\left(\frac{1}{\sqrt{ak^2 + \sqrt{\omega^2 - \omega_{pl}^2}}} + \frac{1}{\sqrt{ak^2 - \sqrt{\omega^2 - \omega_{pl}^2}}}\right)\right], \tag{66}$$

and in the plasmon-pole approximation it is

$$g_a(K,W) = \left[1 - \frac{W_{pl}^2}{W^2}\right]\left[1 + K\frac{W_{pl}^2}{2W^2}\left(\frac{1}{6rt} - \frac{1}{t} - \frac{1}{6rs} - \frac{1}{s}\right)\right], \tag{67}$$

where

$$r = \sqrt{W^2 - W_{pl}^2 + 1/36}; \quad s = \sqrt{K^2 + 1/6 - r}; \quad t = \sqrt{K^2 + 1/6 + r}. \tag{68}$$

The RPA g_a function has to be found numerically.

The result for the simple metals K, Na, Mg, Al, and Be are shown in Fig. 1. In all three approximations we find the surface-plasmon dispersion curves start out at the same frequency, the frequency also obtained at the neglect of spatial dispersion, and then varies linearly with momentum. The

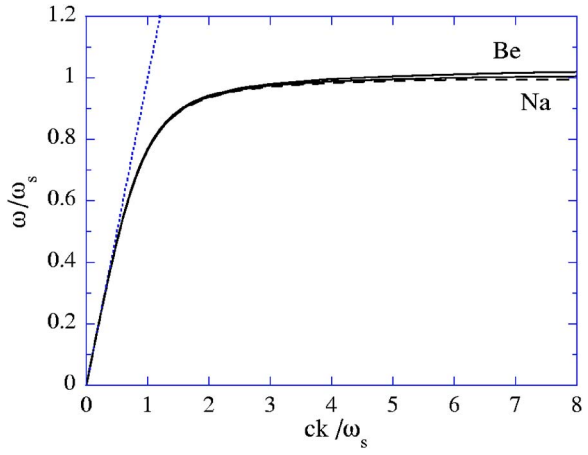

FIG. 2. The surface-plasmon dispersion curves for Be (upper solid curve) and Na (lower solid curve) in the range where retardation effects are visible. The dashed curve is the universal result when spatial dispersion is neglected. The retardation effects push the surface-plasmon dispersion curves down below the dispersion curve for light in vacuum (dotted straight line).

slope is different in the three approximations. Experimentally it turns out that band-structure effects like those from interband transitions, neglected here, are very important, and may actually lead to a negative slope of the dispersion.[25–28] Thus the treatment here is not good enough for obtaining quantitative dispersion curves. However, we find that spatial dispersion is one very important component, although not the only one, influencing the dispersion of the surface modes.

B. Retardation effects at long wavelengths

Retardation effects will modify the surface plasmon dispersion at the long wavelength range. The modes are given as solutions to the equation

$$G^{(1),TM}(k,\omega)+1=0, \tag{69}$$

or expressed in modified g functions, the "tilde functions," as

$$\frac{k}{\gamma}\tilde{g}_a^{(1)}(k,\omega)-\frac{(\omega/c)^2}{\gamma^2}\tilde{g}_b^{(1)}(k,\omega)+\frac{k(k-\gamma)}{\gamma^2}\tilde{g}_c^{(1)}(k,\omega)+2=0. \tag{70}$$

Remember that the tilde over a g function means the g function calculated inside the medium minus the corresponding function calculated in vacuum.

Here, we need both the longitudinal and transverse versions of the dielectric function so we restrict the actual calculation to the RPA approximation. In Fig. 2 we show the results in the range where the retardation effects are important. We find that spatial dispersion has a very small effect on the results here. The figure gives the result for the two of our metals with lowest and highest electron densities, respectively, in comparison with the result without spatial dispersion, dashed curve.

IV. MODES ASSOCIATED WITH A GAP BETWEEN TWO HALF-SPACES

We start by studying the modes generated by the same type of charge and accompanying current densities on the two interfaces. Then we treat the case where we instead have coupled current densities on the two interfaces. In the more complicated geometry, treated here, we need to define two new factors:

$$C=\frac{2\pi\rho^{(0,left)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)};$$
$$D=\frac{2\pi\rho^{(0,right)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)}. \tag{71}$$

The first superscript represents medium 0, which is a vacuum. The second represents which of the two interfaces the charge distribution is located at; the one to the left or the one to the right. The fields inside the left (right) metal halfspace are generated by the charge and current densities at the left (right) interface. The fields in the middle region, in vacuum, are generated by both sets of densities. We start by studying the boundary conditions for the magnetic fields.

The continuity of the $\mathbf{H}_y$ component gives, for the left interface,

$$\mathbf{H}_y(x,0^-,t)=\mathbf{H}_y(x,0^+,t), \tag{72}$$

which means that

$$i2A\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{iq_z0^-}\frac{q_z}{q^2-\varepsilon_\perp^{(1)}(q,\omega)(\omega/c)^2}$$
$$=i2C\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{iq_z0^+}\frac{q_z}{q^2-(\omega/c)^2}$$
$$+i2D\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{-iq_zd}\frac{q_z}{q^2-(\omega/c)^2}, \tag{73}$$

which reduces into

$$A=-C+e^{-\gamma d}D. \tag{74}$$

Similarly, the continuity at the second interface gives

$$i2C\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{iq_zd}\frac{q_z}{q^2-(\omega/c)^2}+i2D\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{-iq_z0^-}\frac{q_z}{q^2-(\omega/c)^2}$$
$$=i2B\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}e^{-iq_z0^+}\frac{q_z}{q^2-\varepsilon_\perp^{(2)}(q,\omega)(\omega/c)^2}, \tag{75}$$

or

$$-e^{-\gamma d}C+D=-B. \tag{76}$$

These two equations may be written as

$$\begin{pmatrix}A\\B\end{pmatrix}=\begin{pmatrix}-1 & e^{-\gamma d}\\ e^{-\gamma d} & -1\end{pmatrix}\begin{pmatrix}C\\D\end{pmatrix}. \tag{77}$$

The continuity of the x component of the $\mathbf{E}$ field at the left interface gives

$$-Ack\left[g_a^{(1)}(k,\omega)+\frac{(k-\gamma)}{\gamma}g_c^{(1)}(k,\omega)-\frac{(\omega/c)^2}{k\gamma}g_b^{(1)}(k,\omega)\right]$$
$$=-Cck\left[g_a^{(0)}(k,\omega)+\frac{(k-\gamma)}{\gamma}g_c^{(0)}(k,\omega)-\frac{(\omega/c)^2}{k\gamma}g_b^{(0)}(k,\omega)\right]$$
$$-Dck\left[g_a^{(0)}(k,\omega)+\frac{(k-\gamma)}{\gamma}g_c^{(0)}(k,\omega)-\frac{(\omega/c)^2}{k\gamma}g_b^{(0)}(k,\omega)\right]e^{-\gamma d}, \quad (78)$$

or

$$-\frac{ic\gamma}{\omega}G^{(1),TM}(k,\omega)A=-\frac{ic\gamma}{\omega}C-\frac{ic\gamma}{\omega}e^{-\gamma d}D, \quad (79)$$

and after common factors have been removed,

$$G^{(1),TM}(k,\omega)A=C+e^{-\gamma d}D. \quad (80)$$

The corresponding equation from the right interface gives

$$G^{(2),TM}(k,\omega)B=e^{-\gamma d}C+D. \quad (81)$$

Here we will limit ourselves to having the same material on both sides of the gap. Then we may combine the results to

$$G^{(1),TM}(k,\omega)\begin{pmatrix}A\\B\end{pmatrix}=\begin{pmatrix}1 & e^{-\gamma d}\\ e^{-\gamma d} & 1\end{pmatrix}\begin{pmatrix}C\\D\end{pmatrix}. \quad (82)$$

All the boundary conditions have resulted in the following system of equations:

$$\begin{pmatrix}A\\B\end{pmatrix}=\begin{pmatrix}-1 & e^{-\gamma d}\\ e^{-\gamma d} & -1\end{pmatrix}\begin{pmatrix}C\\D\end{pmatrix};$$
$$G^{TM}(k,\omega)\begin{pmatrix}A\\B\end{pmatrix}=\begin{pmatrix}1 & e^{-\gamma d}\\ e^{-\gamma d} & 1\end{pmatrix}\begin{pmatrix}C\\D\end{pmatrix}, \quad (83)$$

where we have dropped superscripts indicating the medium. We have a system of equations consisting of four equations and four unknowns. We could write this in the form of one matrix equation with a 4×4 matrix and end up with the condition for the mode being that the determinant of this matrix is zero. However, it is easier to first eliminate the two unknowns A and B in favor of C and D and end up with a 2×2 matrix. The elimination leads to

$$\begin{pmatrix}G^{TM}(k,\omega)+1 & e^{-\gamma d}[1-G^{TM}(k,\omega)]\\ e^{-\gamma d}[1-G^{TM}(k,\omega)] & G^{TM}(k,\omega)+1\end{pmatrix}\begin{pmatrix}C\\D\end{pmatrix}=0. \quad (84)$$

The condition for self-sustained fields is that the determinant of the matrix vanishes, i.e.,

$$\begin{vmatrix}G^{TM}(k,\omega)+1 & e^{-\gamma d}[1-G^{TM}(k,\omega)]\\ e^{-\gamma d}[1-G^{TM}(k,\omega)] & G^{TM}(k,\omega)+1\end{vmatrix}=0. \quad (85)$$

On the dispersion curves, defined by this equation, fields may appear spontaneously without any external charge and current densities at the interfaces. The equation results in the following condition for normal modes:

$$[G^{TM}(k,\omega)+1]^2-e^{-2\gamma d}[G^{TM}(k,\omega)-1]^2=0. \quad (86)$$

This may be generalized to the case of two different materials and results in

$$[G^{(1),TM}(k,\omega)+1][G^{(2),TM}(k,\omega)+1]-e^{-2\gamma d}[G^{(1),TM}(k,\omega)-1]\times[G^{(2),TM}(k,\omega)-1]=0. \quad (87)$$

Let us now find out the corresponding result when spatial dispersion is neglected. Then, since

$$G_1^{(i),TM}(k,\omega)=\frac{\gamma^{(i)}}{\gamma}\frac{1}{\varepsilon^{(i)}(\omega)}, \quad (88)$$

our condition is reduced into

$$\left[\frac{\varepsilon^{(1)}(\omega)}{1}+\frac{\gamma^{(1)}(k,\omega)}{\gamma(k,\omega)}\right]\left[\frac{\varepsilon^{(2)}(\omega)}{1}+\frac{\gamma^{(2)}(k,\omega)}{\gamma(k,\omega)}\right]-e^{-2\gamma(k,\omega)d}\left[\frac{\varepsilon^{(1)}(\omega)}{1}-\frac{\gamma^{(1)}(k,\omega)}{\gamma(k,\omega)}\right]\left[\frac{\varepsilon^{(2)}(\omega)}{1}-\frac{\gamma^{(2)}(k,\omega)}{\gamma(k,\omega)}\right]=0. \quad (89)$$

This is the well-established condition for TM modes when spatial dispersion is neglected.[1]

We are now done with the TM modes. In the present configuration the current component in the y direction at one interface may couple to the corresponding component at the other interface. The modes we have treated so far are all TM modes. When we solved the equation of continuity only the component of the current in the x direction coupled to the charge density. We also found a component of the current pointing in the y direction, but this did not couple to the charge density and did not help in the formation of self-sustained fields. Now, if we have two interfaces these current components at the two interfaces may couple and give rise to a mode; this is the TE mode. The scalar potential does not contribute to the fields in this case. The fields are genuinely transverse in character and only the transverse dielectric function enters the relations. The TM modes on the other hand have both longitudinal and transverse character and both types of dielectric function enter the formalism. Let us study the fields from a current component in the y direction.

The transverse current is

$$\mathbf{J}_\perp=\mathbf{J}-\left(\frac{\mathbf{q}}{q}\mathbf{J}\right)\frac{\mathbf{q}}{q}=J\hat{y}-\left(\frac{\mathbf{q}}{q}J\hat{y}\right)\frac{\mathbf{q}}{q}=J\hat{y}, \quad (90)$$

and the vector potential is

$$\mathbf{A}^{(i)}=\frac{4\pi J^{(i)}\hat{y}}{c[q^2-\varepsilon_\perp^{(i)}(\mathbf{q},\omega)(\omega/c)^2]}. \quad (91)$$

This vector potential gives rise to the electric field,

$$\mathbf{E}_y^{(i)}=\mathbf{E}^{(i)}=\frac{i\omega}{c}\mathbf{A}^{(i)}=\frac{i\omega 4\pi J^{(i)}\hat{y}}{c^2[q^2-\varepsilon_\perp^{(i)}(\mathbf{q},\omega)(\omega/c)^2]}, \quad (92)$$

and to the magnetic field,

$$\mathbf{H}^{(i)} = \mathbf{B}^{(i)} = i\mathbf{q} \times \mathbf{A}^{(i)} = i\frac{4\pi J^{(i)}}{c[q^2 - \varepsilon_{\perp}^{(i)}(\mathbf{q},\omega)(\omega/c)^2]}(k\hat{x} + q_z\hat{z}) \times \hat{y}. \quad (93)$$

This last relation means that

$$\mathbf{B}_x^{(i)} = -iq_z\frac{4\pi J^{(i)}}{c[q^2 - \varepsilon_{\perp}^{(i)}(\mathbf{q},\omega)(\omega/c)^2]}\hat{x}; \quad (94)$$

$$\mathbf{B}_z^{(i)} = ik\frac{4\pi J^{(i)}}{c[q^2 - \varepsilon_{\perp}^{(i)}(\mathbf{q},\omega)(\omega/c)^2]}\hat{z}. \quad (95)$$

We treat the current at one of the interfaces in the same way as we did with the charge density before,

$$J(\mathbf{r},t) = J(k,0,-;\omega)\cdot e^{i(kx-\omega t)}\delta(z) = \int_{-\infty}^{\infty}\frac{dq_z}{2\pi}J(k,0,-;\omega)\cdot e^{i(kx+0y+q_z z-\omega t)}. \quad (96)$$

We let

$$A_J = \frac{2\pi J^{(1)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)};$$

$$B_J = \frac{2\pi J^{(2)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)};$$

$$C_J = \frac{2\pi J^{(0,left)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)}; \quad (97)$$

$$D_J = \frac{2\pi J^{(0,right)}(k,0,-;\omega)\omega}{ck}e^{i(kx-\omega t)}.$$

The continuity of the $\mathbf{B}_x$ component at the left interface gives

$$-\frac{k}{\omega}A_J g_d^{(1)}(k,\omega) = \frac{k}{\omega}C_J g_d^{(0)}(k,\omega) - \frac{k}{\omega}D_J g_d^{(0)}(k,\omega)e^{-\gamma d}, \quad (98)$$

or

$$A_J = -C_J + e^{-\gamma d}D_J. \quad (99)$$

The continuity at the right interface gives

$$\frac{k}{\omega}C_J g_d^{(0)}(k,\omega)e^{-\gamma d} - \frac{k}{\omega}D_J g_d^{(0)}(k,\omega) = \frac{k}{\omega}B_J g_d^{(2)}(k,\omega), \quad (100)$$

or

$$B_J = e^{-\gamma d}C_J - D_J. \quad (101)$$

Thus, we have so far found that

$$\begin{pmatrix} A_J \\ B_J \end{pmatrix} = \begin{pmatrix} -1 & e^{-\gamma d} \\ e^{-\gamma d} & -1 \end{pmatrix}\begin{pmatrix} C_J \\ D_J \end{pmatrix}. \quad (102)$$

The continuity of the $\mathbf{E}_y$ component or $\mathbf{B}_z$ component at the left interface gives

$$A_J\frac{ik}{c\gamma}g_b^{(1)}(k,\omega) = C_J\frac{ik}{c\gamma}g_b^{(0)}(k,\omega) + D_J\frac{ik}{c\gamma}e^{-\gamma d}, \quad (103)$$

or

$$G^{(1),TE}(k,\omega)(k,\omega)A_J = C_J + e^{-\gamma d}D_J, \quad (104)$$

and at the right interface we get

$$e^{-\gamma d}\frac{ik}{c\gamma}C_J + \frac{ik}{c\gamma}D_J = g_b^{(2)}(k,\omega)\frac{ik}{c\gamma}B_J, \quad (105)$$

or

$$e^{-\gamma d}C_J + D_J = G^{(2),TE}(k,\omega)B_J. \quad (106)$$

We have for two equal materials,

$$G^{(1),TE}(k,\omega)\begin{pmatrix} A_J \\ B_J \end{pmatrix} = \begin{pmatrix} 1 & e^{-\gamma d} \\ e^{-\gamma d} & 1 \end{pmatrix}\begin{pmatrix} C_J \\ D_J \end{pmatrix}. \quad (107)$$

Eliminating the unknowns A_J and B_J from the two systems of equation leads to the condition for modes as

$$\begin{vmatrix} G^{(1),TE}(k,\omega)+1 & e^{-\gamma d}[1-G^{(1),TE}(k,\omega)] \\ e^{-\gamma d}[1-G^{(1),TE}(k,\omega)] & G^{(1),TE}(k,\omega)+1 \end{vmatrix} = 0, \quad (108)$$

or

$$[G^{(1),TE}(k,\omega)+1]^2 - e^{-2\gamma(k,\omega)d}[G^{(1),TE}(k,\omega)-1]^2 = 0. \quad (109)$$

In the more general case of two different materials, we have

$$[G^{(1),TE}(k,\omega)+1][G^{(2),TE}(k,\omega)+1] - e^{-2\gamma(k,\omega)d}[G^{(1),TE}(k,\omega) - 1][G^{(2),TE}(k,\omega)-1] = 0. \quad (110)$$

Neglecting spatial dispersion leads to

$$[\gamma(k,\omega)+\gamma^{(1)}(k,\omega)][\gamma(k,\omega)+\gamma^{(2)}(k,\omega)] - e^{-2\gamma(k,\omega)d}[\gamma(k,\omega) - \gamma^{(1)}(k,\omega)][\gamma(k,\omega)-\gamma^{(2)}(k,\omega)] = 0, \quad (111)$$

which is the correct expression.[1] So in summary we have the two mode types from the relations

$$[G^{(1),TM,TE}(k,\omega)+1][G^{(2),TM,TE}(k,\omega)+1] - e^{-2\gamma(k,\omega)d}[G^{(1),TM,TE}(k,\omega)-1][G^{(2),TM,TE}(k,\omega)-1] = 0. \quad (112)$$

We here just note in passing that the relation between our G functions and the so-called surface impedance can be found from the relations above. The surface impedance for p-polarized and s-polarized waves are

$$Z^p = E_x/H_y = \left[-\frac{ic\gamma}{\omega}G^{(i),TM}(k,\omega)A\right] \Big/ [g_d^{(i)}(k,\omega)A] = -\frac{ic\gamma}{\omega}G^{(i),TM}(k,\omega), \quad (113)$$

and

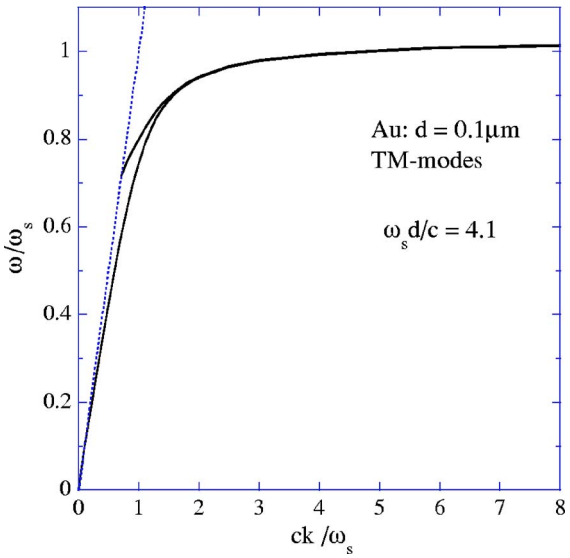

FIG. 3. The two TM-surface-mode branches for two half-spaces of gold separated by a gap of 1 μm. The upper solid curve is just that part of a branch where the mode is a true surface mode, in that the fields decay exponentially away from the two interfaces on both sides. The dotted straight line is the light dispersion curve in vacuum. See the text for more details.

$$Z^s = E_y/H_x = \left[A_J \frac{ik}{c\gamma} g_b^{(i)}(k,\omega)\right] \Big/ \left[-\frac{k}{\omega} A_J g_d^{(i)}(k,\omega)\right]$$
$$= -\frac{i\omega}{c\gamma} G^{(i),TE}(k,\omega), \tag{114}$$

respectively. One should be a little careful with the sign. It is understood that the z direction is into the interface. So to find the relation here we have used the fields at the left interface. If we had used the fields at the interface to the right we would have had to change sign.

In Fig. 3 we show the two TM modes for two gold half-spaces separated by a gap of 1 μm. In the region where retardation is important, the spatial dispersion is once again negligible. This is also the region that is important for the force in the Casimir range. From these findings one would guess that spatial dispersion has a negligible effect on the Casimir force. We will show later that this is actually what we find for the contribution from TM modes, but not from the TE modes at room temperature. At the high momentum side of the figure we note that the frequency of the mode has increased a little beyond the value unity. This very modest effect is caused by the spatial dispersion. Now, one of the two TM modes seems to end at the light-dispersion curve in the figure. It actually continues on the other side, but is not a true surface mode there. There the fields do not decay exponentially away from the metal. Instead they form standing waves between the two plates. This mode has one node between the plates. There are also modes with 2,3,…, nodes. All the TE modes come in the region to the left of and above the light dispersion curve. They have 0,1,…, number of nodes. There is a larger number of modes present the larger the separation between the plates.

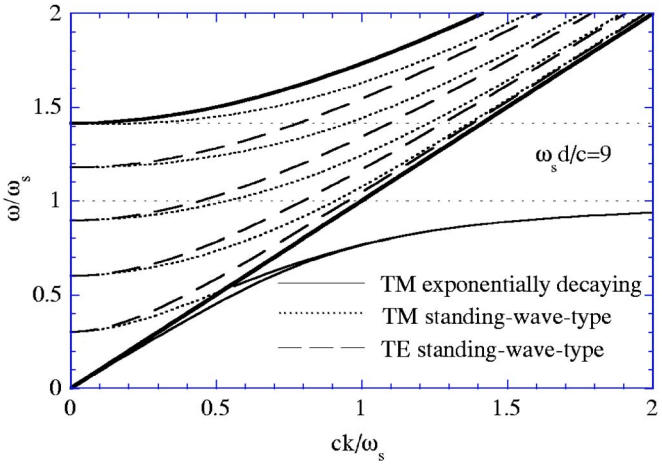

FIG. 4. The complete set of normal modes between gold plates (obtained in neglect of spatial dispersion) for roughly twice the separation of the setup in Fig. 3. The upper thick solid curve is the boundary for bulk-polariton-modes and the lower thick solid straight line is the light-dispersion curve. The thin solid curves are the true-surface-mode-type-modes, the dotted curves are the standing-wave-type TM modes and the dashed curves the standing-wave-type TE modes. Note that one of TM-mode branches changes character when crossing the light-dispersion curve.

In Fig. 4 we show the results for a little larger separation and now in neglect of spatial dispersion. In this case there are six TM modes and four TE modes. One of the TM modes is a true surface mode; one is, for a part of the dispersion curve, a true surface mode and, for the rest of the dispersion curve, a standing-wave type or wave-guide type of mode. The four remaining TM modes and the four TE modes are wave-guide modes. The upper solid curve is the boundary of the continuum of bulk-polariton modes. Above this curve the metals can no longer keep the modes in the gap; they can propagate freely through the geometry and do not contribute to the dispersion forces. The larger the separation between the plates the larger the number of modes between the light dispersion curve and the bulk-polariton continuum.

V. DISPERSION FORCES BETWEEN TWO GOLD PLATES IN VACUUM

In the previous section we found that the conditions for the two mode types between two nonmagnetic metal half-spaces separated by a vacuum gap of thickness d are

$$[G^{TM,TE}(k,\omega)+1]^2 - e^{-2\gamma d}[G^{TM,TE}(k,\omega)-1]^2 = 0, \tag{115}$$

where

$$G^{TM}(k,\omega) = \frac{k}{\gamma}\tilde{g}_a(k,\omega) + \frac{k(k-\gamma)}{\gamma^2}\tilde{g}_c(k,\omega) - \frac{(\omega c)^2}{\gamma^2}\tilde{g}_b(k,\omega) + 1;$$
$$G^{TE}(k,\omega) = \tilde{g}_b(k,\omega) + 1. \tag{116}$$

From this the interaction energy is found to be[1,14]

$$\Delta E(d) = \frac{1}{2\pi\beta}\sum_{\omega_n}{}' \int_0^\infty dk\, k \times \ln\left\{1 - e^{-2\gamma'(k,\omega_n)d}\frac{[G'^{TM}(k,\omega_n)-1]^2}{[G'^{TM}(k,\omega_n)+1]^2}\right\} + \frac{1}{2\pi\beta}\sum_{\omega_n}{}' \int_0^\infty dk\, k \times \ln\left\{1 - e^{-2\gamma'(k,\omega_n)d}\frac{[G'^{TE}(k,\omega_n)-1]^2}{[G'^{TE}(k,\omega_n)+1]^2}\right\}, \tag{117}$$

where

$$\omega_n = \frac{2\pi n}{\hbar\beta}; \quad n = 0,1,2,\ldots. \tag{118}$$

The prime on the summation sign indicates that the $n=0$ term is multiplied by the factor $1/2$. The prime on a function indicates that the frequency argument is on the imaginary frequency axis. For zero temperature the summation is replaced by an integration:

$$\frac{1}{\beta}\sum_{\omega_n}{}' \to \hbar\int_0^\infty \frac{d\omega}{2\pi}. \tag{119}$$

The results are presented in Fig. 5. All results in the figure are the interaction energy divided by the zero-temperature (Casimir) result for a perfect metal ($\hbar c\pi^2/720d^3$). The filled circles with error bars is the experimental result by Lamoreaux.[29] The solid curve is our previous result,[14] where we used the experimental dielectric properties. This means that dissipation was included, but no spatial dispersion. Very similar results are obtained from using the Drude expression, including dissipation, for the dielectric function.[1,14] The momentum dependence of the dielectric function cannot be obtained experimentally, so one is forced to use theoretical expressions for the dielectric functions. The triangles show our present result taking spatial dispersion into account. We have also performed calculations using the versions of the dielectric functions including dissipation, but this led to a very small additional correction. It is very interesting to note that the present result is almost identical to our previous result where the dramatic effect came from dissipation. We are not completely surprised by the result, though. We found in an earlier work[30,31] on the Casimir force between two quantum wells, a problem where the inclusion of spatial dispersion is a necessity, that the contribution from the TE modes dropped out for finite temperature and large separation. Furthermore, the failure of the TE modes to contribute at finite temperature and large separations is not that surprising. For them to contribute the transverse dielectric function has to diverge as ω^{-2} when the frequency goes toward zero. When dissipation is included and spatial dispersion neglected it goes as ω^{-1}. It is not equally obvious when spatial dispersion is included instead of dissipation. Here the formalism becomes more involved. The TE contribution does no longer asymptotically completely vanish but becomes very small as compared to the TM-contribution. It completely vanishes if dissipation is included in the dielectric function the way prescribed by Kliewer and Fuchs.[23] In a more complete expression for the transverse dielectric function the damping parameter should not be a constant, but be frequency and momentum dependent, but there is no reason to believe that this would have any important implications on the present results.

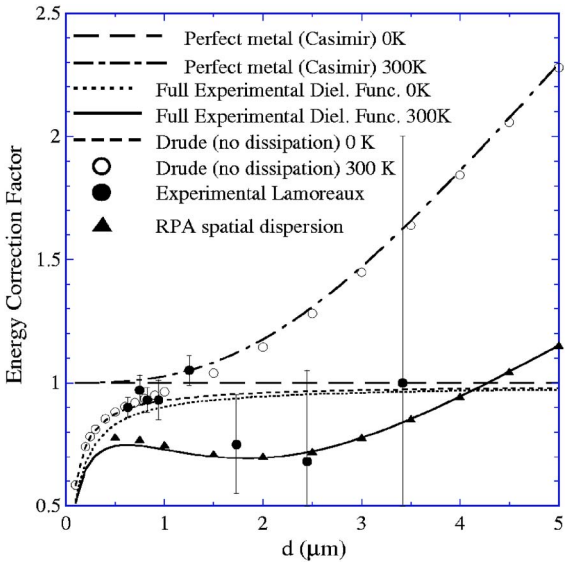

FIG. 5. Energy correction factor as a function of the separation between two gold plates. The long-dashed curve is the Casimir perfect-metal result at zero temperature; it is by definition constant with value unity. The dash-dotted curve is the corresponding result at room temperature. The solid (dotted) curve is the room-temperature (zero-temperature) result from using the experimental dielectric function; these results include the effect of dissipation but neglect the spatial dispersion. The short-dashed curve (open circle) is the simple Drude result at zero temperature (room temperature); these results neglect both dissipation and spatial dispersion. The triangles is the present result including spatial dispersion; dissipation is neglected but the inclusion of dissipation leads to an overlapping result. The solid circles with error bars is the experimental result by Lamoreaux.

In Fig. 6 we show the integrands in the frequency integrals for the TM and TE contributions. We find that the integrand for TE modes becomes very small for small energies or frequencies while the TM integrand saturates at a high value. The zero-frequency value for the TE mode is of the order of 10^{-12}. This drop in value has a negligible effect on the zero-temperature result, since the drop is for very small frequencies; however, for the room-temperature result, it is of the utmost importance. The crosses in the figure show the discrete frequency values entering the frequency summation. There is also one contribution at zero frequency which we cannot indicate in a log-log figure. The zero-frequency value should be multiplied by a factor $1/2$. We see that for 1 μm separation, apart from the zero-frequency contribution, only three points have values big enough to be inside the plot. Thus the zero-frequency contribution is very important for the result. For 5 μm separation no point apart from that at zero frequency gives an important contribution to the result. Here the zero-frequency contributions completely dominates in the net result. Since the integrand for TE modes has a

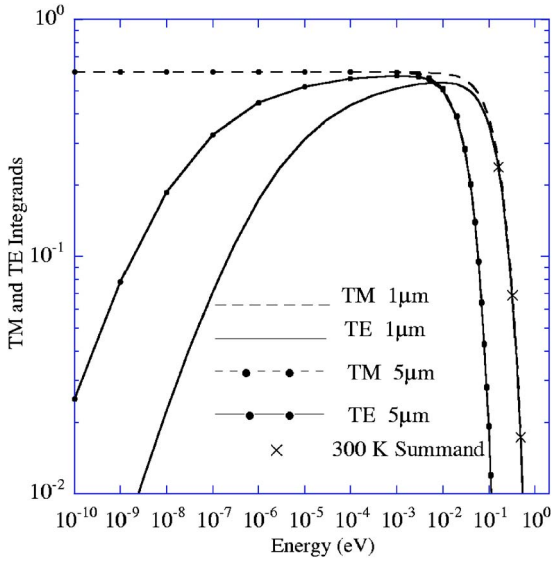

FIG. 6. The frequency integrands after the momentum integration has been performed. The integrands for the TM (TE) contribution at 1 μm separation are represented by a dashed (solid) curve. The curves with data points are the corresponding integrands at 5 μm separation. The crosses indicate the points that contribute to the summation at room temperature.

negligible zero-frequency value only the TM modes contribute to the force at 5 μm separation, at room temperature. At zero temperature both mode types contribute approximately the same to the force at 5 μm separation. The dramatic effects from spatial dispersion are absent at zero temperature. This is in accordance with other recent publications.[11,12] Those authors did not consider the finite temperature effects. There are three important energy scales in the problem: $E_1 \approx 4$ eV is the overall upper limit on the energy contribution on the imaginary frequency axis, determined by the dielectric properties of the materials; $E_2 \approx 3\hbar c/d$ is the energy limit at separation d due to retardation; and $E_3 \approx 2\pi/\beta$ is the first nonzero energy in the discrete frequency summation at finite temperature. The effects we have derived here are in full effect if E_3 is bigger than the smallest of E_1 and E_2. For d smaller than approximately 0.15 μm, E_1 sets the limit. In this case the temperature has to be above 7000 K, so this limit is of minor interest. For larger separations E_2 sets the limit and the full effect is obtained for d of the order of $\frac{1}{2}\hbar\beta c$.

VI. SUMMARY AND CONCLUSIONS

In this work we derived the general conditions for electromagnetic-normal modes at individual planar interfaces and for two half-spaces separated by a vacuum gap; the derivations were performed both with and without spatial dispersion taken into account. We applied the formalism to metallic systems and calculated the van der Waals-Casimir forces between two gold plates. Spatial dispersion was found to have dramatic effects on the finite-temperature Casimir force at large separations, separations larger than $\frac{1}{2}\hbar\beta c$. Half the force was wiped out in this limit. We can conclude that spatial-dispersion and/or dissipation has this very dramatic effect on the force. We conclude by noting that from the present result follows that the Nernst heat theorem[32] is obeyed, even for a system without dissipation. We have earlier shown[15] that our calculations obey this theorem in the presence of dissipation.

ACKNOWLEDGMENTS

Financial support was obtained from the Swedish Research Council.

*Electronic address: bos@ifm.liu.se; www.ifm.liu.se/~boser

[1] B. E. Sernelius, *Surface Modes in Physics* (Wiley-VCH, Berlin, 2001).
[2] D. F. Arago, Mem. Cl. Sci. Math. Phys. Inst. France **1**, 115 (1811).
[3] *Spatial Dispersion in Solids and Plasmas* in the series *Electromagnetic Waves: Recent Developments in Research*, edited by P. Halevi (North-Holland, Amsterdam, 1992), Vol. 1.
[4] *Surface Polaritons: Electromagnetic Waves at Surfaces and Interfaces*, in the series Modern Problems in Condensed Matter Sciences, edited by V. M. Agronovich and D. L. Mills (North-Holland, Amsterdam, 1982), Vol. 1 (series editors V. M. Agranovich and A. A. Maradudin).
[5] See Ref. 4, *Surface Exciton Polariton from an Experimental Viewpoint*, pp. 69–92.
[6] See Ref. 4, *Effects of the Transition Layer and Spatial Dispersion in the Spectra of Surface Polaritons*, pp. 187–238.
[7] W. L. Schaich and W. Chen, Phys. Rev. B **39**, 10714 (1984).
[8] K. L. Kliewer, Surf. Sci. **101**, 57 (1980).
[9] R. Fuchs and K. L. Kliewer, Phys. Rev. B **3**, 2270 (1971).
[10] A. F. Alexandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics* in the series Springer Series in Electrophysics (Springer-Verlag, Berlin, 1984).
[11] R. Esquivel, C. Villarreal, and W. L. Mochan, Phys. Rev. A **68**, 052103 (2003).
[12] R. Esquivel and V. B. Svetovoy, Phys. Rev. A **69**, 062102 (2004).
[13] B. Geyer, G. L. Klimchitskaya, and V. M. Mostepanenko, Phys. Rev. A **67**, 062102 (2003), and references therein.
[14] M. Boström and B. E. Sernelius, Phys. Rev. Lett. **84**, 4757 (2000).
[15] B. E. Sernelius and M. Boström, in *Quantum Field Theory Under the Influence of External Conditions*, edited by K. A. Milton (Rinton Press, Princeton 2004).
[16] J. S. Høye, I. Brevik, J. B. Aarseth, and K. A. Milton, Phys. Rev. E **67**, 056116 (2003); J. S. Høye, I. Brevik, and J. B. Aarseth, *ibid.* **63**, 051101 (2001); I. Brevik, J. B. Aarseth, and J. S. Høye, *ibid.* **66**, 026119 (2002); I. Brevik, J. B. Aarseth, J. S. Høye, and K. A. Milton, in *Quantum Field Theory Under the Influence of External Conditions*, edited by K. A. Milton (Rinton Press, Princeton, 2004).

[17] S. I. Pekar, J. Phys. Chem. Solids **5**, 11 (1958).
[18] R. Ruppin and R. Englman, Phys. Rev. Lett. **53**, 1688 (1984).
[19] M. G. Cottam and D. R. Tilley, *Introduction to Surface and Superlattice Excitations* (Cambridge University Press, Cambridge, 1989).
[20] L. Hedin and S. Lundqvist, Solid State Phys. **23**, 1 (1969).
[21] J. Lindhard, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **28**, 1 (1954).
[22] N. D. Mermin, Phys. Rev. B **1**, 2362 (1970).
[23] K. L. Kliewer and R. Fuchs, Phys. Rev. **181**, 552 (1969).
[24] See the discussion on p. 93 in *Electromagnetic Surface Modes*, edited by A. D. Boardman (Wiley, New York, 1982).
[25] K.-D. Tsuei, E. W. Plummer, and P. J. Feibelman, Phys. Rev. Lett. **63**, 2256 (1989).
[26] K.-D. Tsuei, E. W. Plummer, A. Liebsch, K. Kempa, and P. Bakshi, Phys. Rev. Lett. **64**, 44 (1990).
[27] H. Ishida and A. Liebsch, Phys. Rev. B **54**, 14127 (1996).
[28] V. M. Silkin, E. V. Chulkov, and P. M. Echenique, Phys. Rev. Lett. **93**, 176801 (2004).
[29] S. K. Lamoreaux, Phys. Rev. Lett. **78**, 5 (1997).
[30] B. E. Sernelius and P. Björk, Phys. Rev. B **57**, 6592 (1998).
[31] M. Boström and B. E. Sernelius, Microelectron. Eng. **51–52**, 287 (2000).
[32] W. Nernst, Nachr. Ges. Wiss. Goettingen, Math.-Phys. Kl. **1**, 1 (1906).