A theory of electrodynamic response for bounded metals: surface capacitive effects

Hai-Yao Deng*
School of Physics and Astronomy, Cardiff University, 5 The Parade, Cardiff CF24 3AA, Wales, United Kingdom

We report a general macroscopic theory for the electrodynamic response of semi-infinite metals (SIMs). The theory includes the hitherto overlooked capacitive effects due to the finite spatial extension of a surface. The basic structure of this theory is independent of the particulars of electron dynamics. Analytical expressions have been obtained of the charge density-density response function, which is naturally parsed into two parts. One of them represents a bulk property while the other a pure surface property. We apply the theory to study the responses according to several electronic dynamics models and provide a unified view of their validity and limitations. The models studied include the local dielectric model (DM), the dispersive hydrodynamic model (HDM) and specular reflection model (SRM), as well as the less common semi-classical model (SCM) based on Boltzmann's transport equation. We show that, in terms of their basic equations, the SRM is an extension of the HDM, just as the HDM is an extension of the DM. The SCM improves over the SRM critically through the inclusion of translation symmetry breaking and surface roughness effects. We then employ the response function to evaluate the so-called dynamical structure factor, which plays an important role in particle scattering. As expected, this factor reveals a peak due to the excitation of surface plasma waves (SPWs). Surprisingly, however, the peak is shown to be considerably sharper in the SCM than in other models, indicating an incipient instability of the system according to this model. We also study the distribution of charges induced by a charged particle grazing over a SIM surface at constant speed. This distribution is shown to contain model-specific features that are of immediate experimental interest. This work is expected to find broad applications in optics, plasmonics and other areas such as electron energy loss spectroscopy and accelerator design.

I. INTRODUCTION

Electrodynamic responses, that is, the behaviors of charges in materials and the accompanied electromagnetic fields when subjected to external probes, underlie many physical processes involving the interaction of electron and photon with condensed matter. In principle, the response function can be computed with time-dependent density functional theory (TDDFT)[1] or other formalisms such as Greenwood-Kubo theory. In reality, however, the presence of boundaries (i.e. interfaces and surfaces), which exist in any real materials, makes such computation often unrealistic and impractical. Primarily this is because genuine physical boundaries are complicated and their atomistic profiles are unknown *a priori* whereas in microscopic computation they are usually treated in a highly simplified fashion. On numerous occasions, e.g. in studying optical properties, a microscopic boundary is only of marginal importance and a macroscopic description could be more useful. The obvious path to attaining a macroscopic description is to start with a microscopic model and then proceed to the macroscopic limit, resulting in theories that are nevertheless model specific[2].

To establish a macroscopic theory that is as generic as the microscopic one, a major conceptual obstacle needs to be circumvented, which lies with the macroscopic limit of physical boundaries. Let us take for illustration a surface. On the atomistic scale, the surface layer of a material differs from its bulk interior only quantitatively and all microscopic characteristics – such as the geometric arrangements of atoms and the chemical compositions – smoothly evolve throughout the system without abrupt changes. On a macroscopic scale, however, the surface layer becomes infinitesimally thin regardless of its microscopic details and it is not at all self-evident how and what general physical effects inherited from the microscopic surface should be dealt with.

Traditionally, a macroscopic boundary of vanishing thickness has been treated as a geometric separation and the physical quantities on the opposite interior sides of this separation are then related by boundary conditions. Amongst these are the Maxwell's boundary conditions (MBCs), which directly follow from the fundamental equations of electromagnetism and are the basis of the usual rules governing the reflection and transmission of optical rays. For non-dispersive materials, i.e. those whose electric polarization or current density depends locally on the electric field present in them, MBCs suffice for all purposes. However, for dispersive materials MBCs are well known to be insufficient to determine the solutions. To remedy this deficiency, since 1950s additional boundary conditions (ABCs) have been invoked to supplement the MBCs[3]. These conditions artificially fix the boundary values of certain physical quantities such as the polarization or current density. Despite their widespread use, ABCs lack universality and experimental support. Efforts of justifying them usually start from some microscopic model and the results are specific to the model in use[2]. Some authors showed that conditions equivalent to ABCs could be obtained by use of the extinction theorem[4–7]. They based their results on the so-called 'dielectric approximation', which simply assumed the dispersive constitutive relation of an infinite system extended up to the boundary. This assumption sounds natural but does not take into account genuine boundary effects[8]. It is worth noting that none boundary conditions are needed in microscopic approaches.

In addition to the problem of ABCs, few existing work have discussed the fact that a boundary is not just geometrical but also physical. For example, from a microscopic point of view, the potential in the surface layer differs from the rest and electron waves should be scattered. A macroscopically flat surface can thus appear rough to electron waves that can resolve a distance of the order of a Fermi wavelength. Such scatter-

ing effects break translation symmetry and cannot be incorporated in the dielectric approximation. Moreover, however thin it may appear on a macroscopic scale, a surface layer always represents a region in which physical quantities may experience rapid variations. In particular, charges can flow into and out of the layer leading to capacitive effects. A generic way of handling these unique surface effects in macroscopic theories remains to be recognized.

Our intent here is to put forth a macroscopic theory of electrodynamic response that is applicable to any models of electron dynamics for metals, be they dispersive or non-dispersive. Our theory is based on a straightforward yet general macroscopic description of physical surfaces, which should be valid irrespective of their microscopic profiles, thereby doing away with both MBCs and ABCs just as in the microscopic approaches. This method has recently been used to analyze a simple system (a linear anisotropic dielectric) in a pedagogical way[9]. With this theory, we analyze the electrodynamic responses according to several widely used dispersive or non-dispersive electron dynamics models. Their limitations and relations as well as some longstanding misconceptions about them are critically reviewed and clarified. We then discuss two experimentally interesting quantities: the dynamical structure factor relevant for particle scattering and the distribution of charges induced by a grazing particle that is relevant for surface absorption profile. Applying the theory to metal screening and the propagation of electromagnetic waves will be presented elsewhere.

In this paper we are interested in high-frequency responses, where the ionic motions can be treated as quasi-static and only electronic motions need to be considered.

In the rest of this section, we give an overview of existing work (Sec. I A) and outline our main results (Sec. I B).

A. Overview of the literature

Electrodynamic response may be quantified by the charge density-density response function, which measures the amount of charges induced in a material due to certain probing potential. For infinite systems possessing full translation symmetry, this function has been known in details since the work of Bohm and Pines in the 1950s[10–13]. Their work established the concept of collective electronic oscillations – known as plasma waves or more precisely volume plasma waves (VPWs, sometimes called *bulk plasma waves*) – in the bulk of metals. In reality every system is bounded with surfaces – the hotbed of novel physics and applications[14]. For example, shortly after the discovery of VPWs, it was predicted and later experimentally confirmed that similar oscillations could also be sustained on metal surfaces[15,16]. The study of such surface plasma waves (SPWs) has nowadays grown into a vast field called plasmonics[17–19], which has been pitched as the most viable way toward sub-wavelength control of light-matter interaction. VPWs and SPWs typically dominate the response at frequencies much higher than that of lattice vibrations and other low-energy elementary excitations.

Bounded systems do not possess full translation symmetry and the response function is usually difficult to calculate[1,20–23]. Analytical solutions do not generally exist and an adequate generic understanding remains to be achieved properly taking into account the effects of translation symmetry breaking and surface roughness. Existing work are either macroscopic or microscopic[1], the former based on simple models while the latter relying on computational time-dependent density functional theory. While it accounts for the surface effects in a self-consistent manner and might even provide a microscopic knowledge of the surface itself, the computational approach does not always make transparent the underlying physics and often presumes an ideal surface, such as those modeled by a hard-or-soft-wall-type infinite barrier potential[24]. In addition, it can be computationally expensive for studying realistic aspects of surfaces, e.g. roughness[25]. In recent years, there has seen lots of effort to synergize simple models with density functional theory (in the so-called quantum hydrodynamic model[26–29]) so as to take advantage of both approaches.

Despite the increasing use of computational approaches in electrodynamic response studies[1,23,30–33], the macroscopic approach with simple models continues to be a useful approach and provide additional insights, in particular in the field of applications. The most widely-used amongst existing models include the local dielectric model (DM)[15,34–40], the hydrodynamic model (HDM)[41–47] and the specular reflection model (SRM)[48–51]. These models have existed for a long time and they have been frequently employed to understand surface-related phenomena, examples including the surface energy absorption profile[52], the energy loss spectra of particles scattered off metal surfaces[53–55], the image potential and stopping power[37,56,57], the free energy of metals[58], van der Waals forces and Casimir forces[47], quantum friction and Coulomb drag between relatively moving objects[59–62], ion neutralization spectra[63] and energy dissipation and transport in quantum dots in the proximity of metal surfaces[64–66] as well as photon drag effect (see Ref.[67] and references therein). The DM presumes a local dependence of the electrical current density on the electric field (Drude's law) and is valid only for non-dispersive medium. Where non-local effects are intended, i.e. in dispersive medium, the HDM and the SRM are usually invoked[68,69].

An immediate issue in dealing with dispersive models is that, the models by themselves are insufficient for determining the responses from a macroscopic point of view. Due to spatial dispersion, knowledge must be supplied of the nature of the surfaces to get a unique solution. Historically, this conceptual deficiency has been remedied by a set of what is now known as auxiliary boundary conditions (ABCs)[70], which are imposed to fix the surface values of electric currents or polarization. Very commonly, and overwhelmingly in papers working with the HDM[43,44,46], it has been imposed that no normal electrical current flows in the immediate neighborhood of a surface[1,70]. In the language of electromagnetism, this translates into the vanishing of electrical polarization on surfaces, a condition that was first introduced by Pekar[3,71] in the 1950s and has since been adopted in many variations[72–75]. Nevertheless, it was pointed out long ago that ABCs are physically superficial having no general physical basis[76,77]. Papers trying to justify the ABCs are largely tailored for specific circum-

stances and lack universality[78–81] or based on extinction theorem type development within the dielectric approximation. In a microscopic description that self-consistently takes care of the surfaces, ABCs are obviously superfluous and in fact experimentally refutable[82]. In addition, ABCs are incompatible with the DM and any local models, which are self-sufficient and requires no ABCs. Recently, Henneberger called ABCs *a historical mistake*[83] and proposed a scheme to remove them. Instead of ABCs, he introduced the concept of a surface acting as a radiation source analogous to the charge sheet in the SRM, which is itself controversial[84] and has been refuted by experiments[82].

Beside the electronic models mentioned above, there is a far less common but more accurate model, namely the semi-classical model (SCM)[70,73,74]. This model describes electron dynamics by a semi-classical equation of motion and Boltzmann's transport equation. It is perhaps one of the closest to a rigorous quantum mechanical description[57,85]. The application of this model in bounded systems dated back to the late 1930s, when Fuchs applied it to study the boundary effects on electric conductivity of thin films[86]. In his work, Fuchs introduced a useful parameter p, taking values between zero and unity, to denote the fraction of electrons specularly reflected back off a surface. Circa 1940s, the SCM was used to study anomalous skin effect[87,88] and has since been developed into the standard theory for dealing with this effect[89–91]. In late 1970s, Flores and Garcia were amongst the first to employ it to study electromagnetic responses of dispersive medium on the basis of ABCs[73,74]. As an advantage, the SCM allows one to take care of both translation symmetry breaking as well as surface roughness effects, the latter via the *Fuchs* parameter.

B. Outline of main results

The main purpose of the present work is to derive a macroscopic electrodynamic response theory for semi-infinite metals (SIMs) that is free from the usual boundary conditions, and then employ it to calculate the density-density response function (Sec. II). This is possible thanks to a simple yet general macroscopic description of interfaces possessing whatever microscopic profile. The theory is formulated in a generic form assuming no particulars of electronic dynamics, be they quantum mechanical or classical, local or dispersive. It is valid as long as the length scale of the responses is much bigger than the thickness of the microscopic surface layer so that this layer may be treated as of vanishing thickness, i.e. the macroscopic limit.

With this theory, it is shown that the response function naturally contains two components, one being essentially the same as for an infinite system whereas the other solely due to the presence of surfaces (Secs. II A and II C). It is to the latter that the SPWs contribute. We find that under ABCs the surface contribution would be totally lost and hence no SPWs would exist, in agreement with our recent work[92] showing that the apparent SPW solution admitted in ABC-based HDM is incompatible with that of the DM.

The generality of the theory allows to scrutinize the various model-based macroscopic descriptions under one umbrella and disclose their conceptual relations (Sec. III). Upon inputing a local electron dynamics, the theory expectedly revisits well known results based on the DM. When applied to the classical HDM, our theory yields qualitatively different response due to the surface contribution than the usual treatment of this model based on ABCs imposing no normal current on the surface.

The SRM is subtle in its original design. Nominally, it assumes a specularly reflecting surface and thus no normal current should flow near the surface in apparent conformity with the usual ABCs. No surface contribution and hence no SPWs should occur in this model. Nonetheless, it additionally assumes the existence of a *fictitious* charge sheet located exactly on the surface. As shown in Ref.[93], this charge sheet partially restores the surface contribution and gives rise to SPWs. As far as SPWs are concerned, the SRM is revisited as a direct extension of the HDM in our theory. Despite this, its two basic assumptions are incompatible and the model does not correspond to the specular reflection limit ($p = 1$) of the SCM.

With the SCM, we thoroughly treat the semi-classical response by the theory in Sec. IV. The SCM unveils two interesting yet natural features unseen in other models. Firstly, translation symmetry breaking effects drastically modify the surface part of the response function. Secondly, the function shows dependence on surface roughness by virtue of the *Fuchs* parameter p. In the specular reflection limit, i.e. $p = 1$, the surface contribution disappears and the SRM is not restored, as aforementioned.

Various defining quantities of the theory and models are summarized in Table I, where their relations are made clear.

In connection with the experimental consequences of the symmetry breaking effects, we discuss briefly the energy loss spectra of charged particles reflected off a metal surface in Sec. V. We calculate the dynamical structure factor within the widely used dipole approximation[23,53,94]. It is found that the SPW peak is asymmetric and exceptionally sharper in the SCM than in other models. Actually, its width can possibly be made to vanish by reducing the thermal electronic collision rate, implying that the system contains an instability. It leads to lossless SPWs at the critical point[93], a highly desirable attribute in plasmonics. This finding defies conventional wisdom[95] but is consistent with empirical facts and agrees with our previous work[92,96–98], where it was shown that the decay rate of SPWs is not simply a sum of the thermal collision rate, Landau damping rate and other loss rates such as inter-band absorption rates, but should be from these deducted by a positive-definite term, which is guaranteed by the principle of physical causality.

Another quantity of experimental interest is the spatial distribution of induced charges, which are ultimately responsible for the surface absorption profile and stopping power[100] and may be chartered out directly by near-field-optical microscopy. As an illustration, we have evaluated the distribution of these charges induced by an exterior charged particle grazing over the surface at constant speed [Fig. 1 (a)]. The distribution is shown sensitive to which model is in use, see Fig. 1 (b) for instance and Sec. V for thorough discus-

sions. For example, according to the DM the induced charges should always be symmetrically deployed about the particle along the direction of its motion, while according to the SCM more charges are concentrated in front of the particle. Perpendicular to the direction of motion, the distribution is periodic in all models but with a much shorter period in the SCM.

An online supplemental text[93] has been provided to discuss various issues that could not be accommodated in the main text. These include a phenomenological model[30,101] for partially accounting for the contributions of valence electrons, some numerical details, some properties of SPWs in the SCM and the logical inconsistencies of the conventional SRM.

II. THEORY OF DYNAMICAL RESPONSES

In this section, we derive the macroscopic electrodynamic response theory and calculate the charges induced by external stimuli, and from this the density-density response function is extracted including contributions from both the SPWs and VPWs. The theory is founded on general physical concepts and independent of the particulars of electron dynamics. For the sake of illustration, we shall consider a semi-infinite metal (SIM) with a single macroscopically flat surface. Extension to films and other geometries is straightforward and will be considered elsewhere.

The SIM is assumed to occupy the half-space $z \geq 0$ and interfaces with the vacuum at $z = 0$, as shown in Fig. 1. Throughout the paper, we reserve $\mathbf{r} = (x, y)$ for planar coordinates and $\mathbf{x} = (\mathbf{r}, z)$. A point on the surface is denoted by $\mathbf{x}_0 = (\mathbf{r}, 0)$ and we use t to denote time. The surface may appear rough on the scale of Fermi wavelength and hence cause diffuse scattering of electron waves, but is assumed sufficiently flat on a macroscopic length scale so that the translational symmetry along the surface is preserved.

In studying dynamical responses for bounded medium, it is customary to work directly with the electrostatic potential – or more generally the electromagnetic field in the case of non-negligible retardation effects – and write down its expressions on the vacuum side and the material side separately. ABCs are then invoked together with the usual MBCs – the continuity of both the electrostatic potential and the normal component of the electric displacement field in the electrostatic limit – to join them at the boundary. In what follows we show how a general response theory can be derived without the use of any explicit boundary conditions and other type of *ad hoc* prescriptions such as those of Ref.[83]. To this end, we first need to establish the macroscopic limit of an arbitrary physical interface in a general way. Considering that a real microscopic surface can hardly be specified even for the simplest material, one might deem it hopeless. However, the following elementary analysis suggests otherwise.

Let us imagine bringing two materials (A and B) in contact, and an interfacial layer of thickness d_s – in the order of a few lattice constants – shall form in between (see Fig. 2). We may characterize this layer by a surface potential ϕ_s, which should quickly decay to zero in the bulk regions outside the interfacial layer. The exact microscopic profile of the layer varies from one case to another and can hardly be known *a priori*. Despite this, we may still write down a generic form for the electric current density $\mathbf{j}(\mathbf{x}, t)$ in the whole system including the interfacial layer. To this end, we observe that in the bulk regions where ϕ_s vanishes, the form of $\mathbf{j}(\mathbf{x}, t)$ can be completely determined with the respective dynamic equations for the infinite materials, apart from some parameters (such as the *Fuchs* parameter, see Sec. IV) that encode the effects of surface scattering on the electron waves. Let us denote by $\mathbf{J}_{A/B}(\mathbf{x}, t)$ the values of $\mathbf{j}(\mathbf{x}, t)$ in the bulk region of A/B. Microscopically, $\mathbf{j}$ evolves from $\mathbf{J}_A$ in the bulk region of A, through a rapid but smooth variation in the interfacial layer, to $\mathbf{J}_B$ in the bulk region of B. Formally, we can write for the μ-th component of the current density as

$$j_\mu(\mathbf{x}, t) = J_{A,\mu}(\mathbf{x}, t) w_\mu(z) + J_{B,\mu}(\mathbf{x}, t)(1 - w_\alpha(z)),$$

where the profile functions $w_\alpha(z)$ approach unity in the bulk region of A and zero in that of B. The exact form of $w_\mu(z)$ depends on the microscopic details of the interfacial layer. On the macroscopic length scale of Λ, however, the interfacial layer appears infinitely thin and $w_\alpha(z)$ reduce to the Heaviside step function $\Theta(z)$, where $\Theta(z \geq 0) = 1$ and $\Theta(z < 0) = 0$. In the macroscopic limit, one thus ends up with

$$\mathbf{j}(\mathbf{x}, t) = \mathbf{J}_A(\mathbf{x}, t)\Theta(z) + \mathbf{J}_B(\mathbf{x}, t)(1 - \Theta(z)), \tag{1}$$

which holds valid for any $w_\mu(z)$ and is thus a general and complete macroscopic description of a physical interface, as long as the perturbation on one side does not cause significant responses on the other[102].

To recapitulate, Eq. (1) elegantly captures two important physical consequences of an interface: the rapid variation of the current density through the step function $\Theta(z)$ and the surface scattering effects on electron dynamics through the parameters contained in the bulk values $\mathbf{J}_{A/B}$. These scattering effects – including the symmetry breaking effects – have been ignored in most models except for the SCM. In general $\mathbf{J}_A$ and $\mathbf{J}_B$ are not equal on the interface, as is certainly the case for local dynamics models, and charges can then accumulate in the interfacial layer. Such capacitive effects would be mistakenly erased under usual ABCs, which often dictate continuity of current density across an interface, e.g. the vanishing of the normal component at the metal-vacuum interface.

A. Generic formulation

With the macroscopic limit of physical interfaces, Eq. (1), we now formulate a general theory of electrodynamic response for the SIM.

We started with the fact that, in response to a probing electric field $\mathbf{E}_{\text{probe}}(\mathbf{x}, t)$ an electrical current flows in the metal and charges may appear, whose density we denote by $\rho(\mathbf{x}, t)$. In the jellium model adopted here, $\rho(\mathbf{x}, t) = en(\mathbf{x}, t)$ is simply the *deviation* $n(\mathbf{x}, t)$ of the density of electrons from its mean value n_0. Here e is the charge of an electron. These charges then generate an additional electric field denoted by $\mathbf{E}(\mathbf{x}, t)$. In the bulk region of the metal, the total electric field felt by

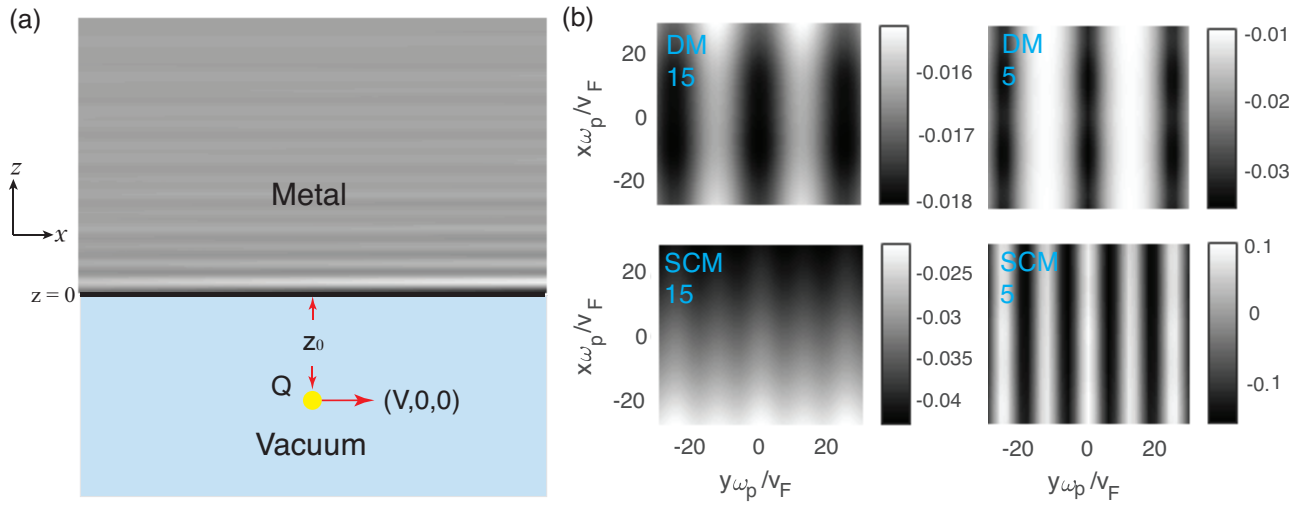

FIG. 1. (a) Sketch of the system: a semi-infinite metal (SIM) occupies the half space $z \geq 0$ and the vacuum occupies the other half. $\mathbf{x} = (\mathbf{r}, z)$ and $\mathbf{r} = (x, y)$. A point on the surface is denoted by $\mathbf{x}_0 = (\mathbf{r}, 0)$. The present work is devoted to deriving a general dynamical response theory for the SIM without suffering from the routinely used boundary conditions. The theory allows us to calculate the charge density $\rho(\mathbf{x}, t)$ induced in the SIM due to the presence of any stimuli. In the example shown in panel (a), a particle of unit charge – indicated by a yellow dot – grazes over the surface at distance z_0 and constant velocity $\mathbf{V} = (V, 0, 0)$, where $V = 10v_F$ for the plot. The gray scale indicates the value of $\rho(\mathbf{x}, t)$ in this example. The planar charge distribution, i.e. $\rho_{\parallel}(\mathbf{r}, t) = \int dz\, \rho(\mathbf{x}, t)$ is displayed in (b) for two models, the DM and the SCM, see Sec. V for discussions and other models. The particle is located at $(0, 0, -z_0)$ for the moment under consideration. The number in each panel indicates the value of $z_0\omega_p/v_F$.

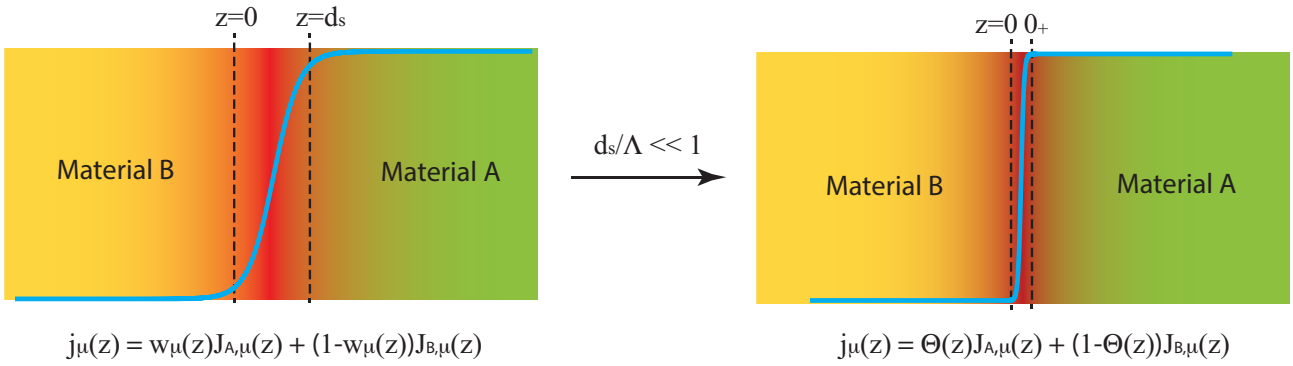

FIG. 2. The macroscopic limit of a physical interface joining materials A and B. On the atomistic scale, the interface has finite thickness d_s (left). The current density j_μ can be related to its values $J_{A/B,\mu}$ in the bulk regions (outside the interfacial layer) via the profile functions $w_\mu(z)$, which approaches unity on side A and zero on side B. On a macroscopic length scale $\Lambda \gg d_s$, the interfacial layer appears infinitely thin and $w_\mu(z)$ reduces to Heaviside step function $\Theta(z)$ (right).

the electrons is $\mathbf{E}_{\rm tot}(\mathbf{x}, t) = \mathbf{E}_{\rm probe}(\mathbf{x}, t) + \mathbf{E}(\mathbf{x}, t)$. In the regime of linear responses, the density of the current flowing in that region then contains two parts, $\mathbf{J}_{\rm tot}(\mathbf{x}, t) = \mathbf{J}_{\rm probe}(\mathbf{x}, t) + \mathbf{J}(\mathbf{x}, t)$, where $\mathbf{J}_{\rm probe}$ and $\mathbf{J}$ are due to $\mathbf{E}_{\rm probe}$ and $\mathbf{E}$, respectively. It should be clear that here the responses (and later the conductivity) are defined with respect to the system's Hamiltonian excluding the long-range Coulomb interaction, which is treated by a self-consistent mean field. This can be justified in the random phase approximation or by the TDDFT.

According to Eq. (1), the current density throughout the entire space including the vacuum can then be written as

$$\mathbf{j}(\mathbf{x}, t) = \Theta(z)\mathbf{J}_{\rm tot}(\mathbf{x}, t), \tag{2}$$

This relation is implicit in any local dielectric models[103]. As to be seen, boundary conditions, i.e. both MBCs and ABCs, are no longer needed. By Eq. (2) charges can accumulate on the surface layer producing capacitive effects, which would be mistakenly excluded under ABCs or other equivalent prescriptions such as in the hard-wall picture often adopted in computational approaches. The need to go beyond the hard-wall picture has recently drawn considerable attention in the computational hydrodynamic approach[26–29] in studying the local

plasmon resonances on metal nano-particles.

Now the equation of continuity can be used to relate ρ and $\mathbf{j}$ as follows

$$\mathcal{D}_t\rho(\mathbf{x},t)+\partial_{\mathbf{x}}\cdot\mathbf{j}(\mathbf{x},t)=0,\quad \mathcal{D}_t=\tau^{-1}+\partial_t. \tag{3}$$

Here a global relaxation term $-\rho(\mathbf{x},t)/\tau$ has been included to account for the relaxation of local non-equilibrium charges [due to finite density deviation $n(\mathbf{x},t)$] due to microscopic electronic collisions driving the system toward thermodynamic equilibrium[57,98] [in which the deviation $n(\mathbf{x},t)$ must vanish]. In terms of $\mathbf{J}_{\rm tot}$, the equation reads

$$\mathcal{D}_t\rho(\mathbf{x},t)+\partial_{\mathbf{x}}\cdot\mathbf{J}_{\rm tot}(\mathbf{x},t)=-\Theta'(z)J_{{\rm tot},z}(\mathbf{x}_0,t), \tag{4}$$

where $\Theta'(z)=d\Theta(z)/dz$. In this equation, we have dropped $\Theta(z)$ on the left hand side to simplify the notation, as is clear that $\mathbf{x}$ represents a point on the metal side. To avoid ambiguity, $\Theta'(z)$ should not be simply identified with the Dirac function $\delta(z)$, because $\int_0^\infty dz\Theta'(z)=1$ by definition but $\int_0^\infty dz\delta(z)=\frac{1}{2}$. The right-hand term of Eq. (4) corresponds to the aforementioned capacitive effects. It plays a critical role in the energy conversion process, which has been overlooked until our recent work[98]. This term was noticed by A. L. Fetter in his study of edge plasmon in confined two-dimension electron gases[104] and also used in Refs.[105] in a different context.

For studying responses, it is convenient to isolate the terms due to the probing field. Thus, we rewrite Eq. (4) as

$$\mathcal{D}_t^2\rho(\mathbf{x},t)+\mathcal{D}_t\partial_{\mathbf{x}}\cdot\mathbf{J}(\mathbf{x},t)=S(\mathbf{x}_0,t)+S_{\rm probe}(\mathbf{x},t), \tag{5}$$

where $S(\mathbf{x}_0,t)=-\Theta'(z)\mathcal{D}_tJ_z(\mathbf{x}_0,t)$ and

$$S_{\rm probe}(\mathbf{x},t)=-\mathcal{D}_t\left[\partial_{\mathbf{x}}\cdot\mathbf{J}_{\rm probe}(\mathbf{x},t)+\Theta'(z)J_{{\rm probe},z}(\mathbf{x}_0,t)\right] \tag{6}$$

denotes the probing source. Introducing the following Fourier transform

$$\rho(\mathbf{x},t)=\sum_{\mathbf{k}}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}}{\sqrt{A}}\rho(z;\mathbf{k},\omega), \tag{7}$$

where A is the surface area used to quantize the in-plane wave vector $\mathbf{k}$, for the charge density, and analogously for all other fields, we can rewrite Eq. (5) as

$$-i\bar{\omega}\nabla\cdot\mathbf{J}(z;\mathbf{k},\omega)-\bar{\omega}^2\rho(z;\mathbf{k},\omega)=S(\mathbf{k},\omega)\Theta'(z)+S_{\rm probe}(z;\mathbf{k},\omega). \tag{8}$$

Here $\bar{\omega}=\omega+i/\tau$, $\nabla=(i\mathbf{k},\partial_z)$ and $S(\mathbf{k},\omega)=i\bar{\omega}J_z(0;\mathbf{k},\omega)$ does not depend on z. In the regime of linear responses considered throughout this paper, we can write $\mathbf{J}(z;\mathbf{k},\omega)$ as a linear functional of $\mathbf{E}(z;\mathbf{k},\omega)$, i.e.

$$J_\mu(z;\mathbf{k},\omega)=\sum_{\nu=x,y,z}\int dz'\sigma_{\mu\nu}(z,z';\mathbf{k},\omega)E_\nu(z';\mathbf{k},\omega), \tag{9}$$

where $\sigma_{\mu\nu}(z,z';\mathbf{k},\omega)$ is the conductivity tensor by definition. The same relation holds valid between $\mathbf{J}_{\rm probe}(z;\mathbf{k},\omega)$ and $\mathbf{E}_{\rm probe}(z;\mathbf{k},\omega)$. Now that $\mathbf{E}(z;\mathbf{k},\omega)$ is also a linear functional of $\rho(z;\mathbf{k},\omega)$ by the laws of electrostatics, we can always define a linear operator $\mathcal{H}$ so that

$$\mathcal{H}\rho(z;\mathbf{k},\omega)=-i\bar{\omega}\nabla\cdot\mathbf{J}(z;\mathbf{k},\omega). \tag{10}$$

With this Eq. (8) becomes

$$\left(\mathcal{H}-\bar{\omega}^2\right)\rho(z;\mathbf{k},\omega)=S(\mathbf{k},\omega)\Theta'(z)+S_{\rm probe}(z;\mathbf{k},\omega). \tag{11}$$

We can do some further transformations by noting that for any quantity existing in the half space a cosine Fourier transform can be defined, i.e.

$$\rho(z;\mathbf{k},\omega)=\frac{2}{\pi}\int_0^\infty dq\ \cos(qz)\,\rho(\mathbf{K},\omega), \tag{12}$$

where $\mathbf{K}=(\mathbf{k},q)$. In terms of $\rho(\mathbf{K},\omega)$, Eq. (11) is rewritten as

$$\int_0^\infty dq'\left\{\mathcal{H}(q,q';\mathbf{k},\omega)-\bar{\omega}^2\delta(q-q')\right\}\rho(\mathbf{K}',\omega) = S(\mathbf{k},\omega)+S_{\rm probe}(\mathbf{K},\omega), \tag{13}$$

where $\mathbf{K}'=(\mathbf{k},q')$, $\mathcal{H}(q,q';\mathbf{k},\omega)$ is the matrix element between the cosine waves $\cos(qz)$ and $\cos(q'z)$, and

$$S_{\rm probe}(\mathbf{K},\omega)=\int_0^\infty dz\ \cos(qz)\,S_{\rm probe}(z;\mathbf{k},\omega). \tag{14}$$

To close Eq. (13), we utilize the fact that $J_z(0;\mathbf{k},\omega)$ and hence $S(\mathbf{k},\omega)$ are also linear functionals of the charge density, i.e.

$$S(\mathbf{k},\omega)=\int_0^\infty dq\,\frac{G(\mathbf{K},\omega)}{K^2}\,\rho(\mathbf{K},\omega) \tag{15}$$

with $K^2=k^2+q^2$ and $k=|\mathbf{k}|$. Here $G(\mathbf{K},\omega)/K^2$ denotes the kernel, which is material and model specific; see what follows.

Equations (13) and (15) comprise a complete dynamical response theory for SIMs, allowing us to determine the induced charges provided $S_{\rm probe}(\mathbf{K},\omega)$ is known. No boundary conditions have been explicitly invoked in this theory. Extension to other geometries such as films and spherical particles will be performed in a separate publication.

B. Induced charge densities

Here we obtain the induced charge densities from the theory derived above.

It is not necessary but useful to simplify the equations in the first place. We can make use of some general properties of $\mathcal{H}(q,q';\mathbf{k},\omega)$ to this end. It is instructive to look at the equations for self-sustained waves in the absence of probing fields, i.e. we leave out $S_{\rm probe}$ from Eq. (13). As shown in Refs.[92,96,97], the resulting equation admits of two types of solutions representing VPWs and SPWs, respectively. Those of VPWs satisfy $S(\mathbf{k},\omega)\equiv 0$, and then the VPW frequencies are obtained as solutions to the secular equation $\left|\mathcal{H}-\bar{\omega}^2\right|=0$. As such, we see that $\mathcal{H}$ contains complete information of VPWs in a SIM. It is reasonable to assume that VPWs are not sensitive to the presence of boundaries, and $\mathcal{H}$ is essentially that of an infinite system. To make this statement accurate, let us analyze the conductivity tensor $\sigma_{\mu\nu}(z,z';\mathbf{k},\omega)$, which contains all information of the electron dynamics of the underlying material. For an infinite system without the surface, the translational symmetry is also preserved along z-axis and

thus $\sigma_{\mu\nu}(z,z';\mathbf{k},\omega)$ depends only on the difference between z and z'. However, for a SIM, the symmetry is broken and it must depend on the coordinates individually. It is then useful to decompose $\sigma_{\mu\nu}$ into two parts, $\sigma_{b,\mu\nu}(z-z';\mathbf{k},\omega)$ and $\sigma_{s,\mu\nu}(z,z';\mathbf{k},\omega)$, where $\sigma_{b,\mu\nu}(z-z';\mathbf{k},\omega)$ is that of the infinite system while $\sigma_{s,\mu\nu}(z,z';\mathbf{k},\omega)$ signifies symmetry breaking effects. By Eq. (10), $\mathcal{H}$ accordingly splits into two parts, $\mathcal{H}_b$ and $\mathcal{H}_s$. Since it is responsible for the properties of VPWs in an infinite (isotropic) system, $\mathcal{H}_b$ must be diagonal in the q-space, i.e. $\mathcal{H}_b(q,q';\mathbf{k},\omega)=\Omega^2(\mathbf{K},\omega)\delta(q-q')$, where $\Omega(\mathbf{K},\omega)$ is a frequency. By virtue of the rotational symmetry of an infinite system, Ω depends on the length but not the direction of $\mathbf{K}$. In the meanwhile, $\mathcal{H}_s$ gives rise to scattering of VPWs, which generally makes a small perturbation of the order kv_F/ω_p, where v_F is the Fermi velocity of the metal and ω_p is the characteristic plasma frequency (see the next subsection), and can be treated perturbatively[96–98]. To the zero-th order in this perturbation, we have

$$\mathcal{H}(q,q';\mathbf{k},\omega)=\Omega^2(K,\omega)\delta(q-q'). \tag{16}$$

Equation (13) then becomes

$$\left[\Omega^2(K,\omega)-\bar{\omega}^2\right]\rho(\mathbf{K},\omega)=S(\mathbf{k},\omega)+S_{\text{probe}}(\mathbf{K},\omega). \tag{17}$$

It is easy to show that the dielectric function of an infinite system is given by

$$\epsilon(K,\omega)=1-\frac{\Omega^2(K,\omega)}{\bar{\omega}^2}. \tag{18}$$

As usual, the zeros of $\epsilon(K,\omega)$ yield the VPW frequencies. The positive-definite quantity $-\text{Im}\left[\epsilon^{-1}(K,\omega)\right]$ is the so-called loss function for an infinite system. Here Im/Re$[f]$ takes the imaginary/real part of an arbitrary quantity f.

Analogously, we may split G, the kernel in Eq. (15), into two parts, G_b and G_s, which originate from $\sigma_{b,\mu\nu}$ and $\sigma_{s,\mu\nu}$, respectively. In all the models to be discussed in this paper, we find that $G_b=-4i\bar{\omega}k\sigma(\omega)$ independent of q, where $\sigma(\omega)$ is the local part of $\sigma_{b,\mu\nu}$, namely $\delta_{\mu\nu}\delta(z-z')\sigma(\omega)$ with $\delta_{\mu\nu}$ being the Kronecker symbol. If inter-band transitions are neglected, one further finds $\sigma(\omega)=(i/\bar{\omega})(\omega_p^2/4\pi)$, which is the Drude conductivity. Thus, we arrive at

$$G(\mathbf{K},\omega)=(k/\pi)\omega_p^2+G_s(\mathbf{K},\omega). \tag{19}$$

As to be seen later, in all the models discussed in this paper, except for the SCM, G_s vanishes.

Combining Eqs. (15) and (17), we easily obtain the density of the induced charges in two components,

$$\rho(\mathbf{K},\omega)=\rho_1(\mathbf{K},\omega)+\rho_2(\mathbf{K},\omega),$$

where ρ_1 stems directly from S_{probe} by Eq. (17), i.e.

$$\rho_1(\mathbf{K},\omega)=\frac{S_{\text{probe}}(\mathbf{K},\omega)}{\Omega^2(K,\omega)-\bar{\omega}^2}=-\frac{S_{\text{probe}}(\mathbf{K},\omega)}{\epsilon(K,\omega)}\frac{1}{\bar{\omega}^2}, \tag{20}$$

and ρ_2 originates from S, which would have been erroneously left out had we imposed that $J_{\text{tot,z}}(0;\mathbf{k},\omega)\equiv 0$ or other ABCs. This part is given by

$$\rho_2(\mathbf{K},\omega)=-\frac{S(\mathbf{k},\omega)}{\epsilon(K,\omega)}\frac{1}{\bar{\omega}^2}=-\frac{\bar{S}_{\text{probe}}(\mathbf{k},\omega)}{\epsilon_s(\mathbf{k},\omega)\epsilon(K,\omega)}\frac{1}{\bar{\omega}^2}. \tag{21}$$

Here ϵ_s and $\bar{S}_{\text{probe}}$ are defined as

$$\epsilon_s(\mathbf{k},\omega)=1+\int_0^\infty dq'\frac{G(\mathbf{K}',\omega)}{\bar{\omega}^2K'^2}\frac{1}{\epsilon(K',\omega)}, \tag{22}$$

which may be called the surface dielectric function, and

$$\bar{S}_{\text{probe}}(\mathbf{k},\omega)=-\int_0^\infty dq'\frac{G(\mathbf{K}',\omega)}{\bar{\omega}^2K'^2}\frac{S_{\text{probe}}(\mathbf{K}',\omega)}{\epsilon(K',\omega)}. \tag{23}$$

By virtue of the rotational symmetry about z-axis, we expect that $\epsilon_s(\mathbf{k},\omega)$ depends on the length of $\mathbf{k}$ but not its direction.

Obviously, $\rho_1(\mathbf{K},\omega)$ features a resonance near the zeros of $\epsilon(K,\omega)$, indicating the excitation of VPWs. On the other hand, $\rho_2(\mathbf{K},\omega)$ contains an additional resonance near the zeros of $\epsilon_s(k,\omega)$. This resonance corresponds to the excitation of SPWs. As shown in Refs.[92,96,97], the SPW dispersion relation is determined by the equation that $\epsilon_s(k,\omega)=0$. Further discussions of this equation and the properties of SPWs are presented in Ref.[93]. Unlike $\epsilon^{-1}(K,\omega)$, the imaginary part of $\epsilon_s^{-1}(k,\omega)$ does not keep a single sign in the entire spectrum of $\omega\geq 0$. As to be seen later, in the vicinity of VPW resonances there is nearly complete cancellation between the responses encoded in ρ_1 and ρ_2 under certain circumstances, leaving only the resonance of the SPWs discernible.

For the sake of completeness, let us also give the electric field generated by the induced charges. The electrostatic potential $\phi(z;\mathbf{k},\omega)$ is given by

$$\phi(z;\mathbf{k},\omega)=\frac{2\pi}{k}\int_{-\infty}^{\infty}dz'e^{-k|z-z'|}\rho(z';\mathbf{k},\omega). \tag{24}$$

The electric field is obtained as $\mathbf{E}(z;\mathbf{k},\omega)=-\nabla\phi(z;\mathbf{k},\omega)$. Explicitly, one finds *in the metal* the projection onto the surface

$$\mathbf{E}_{\parallel}(z;\mathbf{k},\omega)=-i\int_0^\infty dq\ \frac{4\mathbf{k}\rho(\mathbf{K},\omega)}{K^2}\left(2\cos(qz)-e^{-kz}\right) \tag{25}$$

and the normal component

$$E_z(z;\mathbf{k},\omega)=\int_0^\infty dq\ \frac{4k\rho(\mathbf{K},\omega)}{K^2}\left(2\frac{q}{k}\sin(qz)-e^{-kz}\right). \tag{26}$$

These expressions are easily established from the laws of electrostatics.

C. The density-density response function

In this subsection, we discuss two cases of special importance in many applications such as particle and light scattering. The density-density response function will be obtained.

Case (i). We place some charges exterior to the metal and look at the responses of the metal to these charges. Let the density of these charges be $\rho_{\text{ext}}(z;\mathbf{k},\omega)$, which exists only on the vacuum side $z<0$. The probing field is obtained from the corresponding electrostatic potential $\phi_{\text{probe}}(z;\mathbf{k},\omega)$ in the metal. Adapting Eq. (24) to this case, we find

$$\phi_{\text{probe}}(z\geq 0;\mathbf{k},\omega)=\left(e^{-kz}/k\right)\xi(\mathbf{k},\omega), \tag{27}$$

where

$$\xi(\mathbf{k},\omega) = 2\pi \int_{-\infty}^{0} dz\, e^{kz} \rho_{\rm ext}(z;\mathbf{k},\omega).$$

It follows that in the metal

$$\mathbf{E}_{\rm probe}(z;\mathbf{k},\omega) = -\nabla\phi_{\rm probe}(z;\mathbf{k},\omega) = \xi(\mathbf{k},\omega)e^{-kz}(-i\hat{\mathbf{k}},1). \tag{28}$$

Here $\hat{\mathbf{k}} = \mathbf{k}/k$. Note that this field cannot be used to unveil the complete q-resolved profile of the density response of SIMs, as it has a fixed z-dependence of the form e^{-kz}, regardless of the configuration of the exterior charges.

The resulting $S_{\rm probe}$ is proportional to ξ. We can write it as

$$S_{\rm probe}(\mathbf{K},\omega) = B(\mathbf{K},\omega)\xi(\mathbf{k},\omega), \tag{29}$$

where $B(\mathbf{K},\omega)$ is the coefficient. From Eqs. (20) and (21) one finds

$$\rho(\mathbf{K},\omega) = P(\mathbf{K},\omega)\xi(\mathbf{k},\omega), \tag{30}$$

where $P = P_1 + P_2$, with

$$P_1(\mathbf{K},\omega) = -\frac{B(\mathbf{K},\omega)}{\epsilon(K,\omega)}\frac{1}{\bar{\omega}^2} \tag{31}$$

and

$$P_2(\mathbf{K},\omega) = -\frac{\bar{B}(\mathbf{k},\omega)}{\epsilon_s(\mathbf{k},\omega)\epsilon(K,\omega)}\frac{1}{\bar{\omega}^2}. \tag{32}$$

Here

$$\bar{B}(\mathbf{k},\omega) = -\int_0^\infty dq' \frac{G(\mathbf{K}',\omega)}{\bar{\omega}^2 K'^2}\frac{B(\mathbf{K}',\omega)}{\epsilon(K',\omega)}. \tag{33}$$

Note that $B(\mathbf{K},\omega)$ depends on the model of electron dynamics.

Case (ii). We place the metal in an electrostatic potential of the form

$$\phi_{\rm probe}(z;\mathbf{k},\omega) = \varphi(\mathbf{K}',\omega)\cos(q'z)$$

with q' fixed. The corresponding probing field is given by

$$\mathbf{E}_{\rm probe}(z;\mathbf{k},\omega) = \varphi(\mathbf{K}',\omega)\left(-i\mathbf{k}\cos(q'z), q'\sin(q'z)\right). \tag{34}$$

This field implies a probing charge of density

$$\rho_{\rm probe}(z;\mathbf{k},\omega) = (K'^2/4\pi)\varphi(\mathbf{K}',\omega)\cos(q'z),$$

or equivalently

$$\rho_{\rm probe}(\mathbf{K},\omega) = (K'^2/8)\varphi(\mathbf{K}',\omega)\delta(q-q'),$$

which allows us to unveil the q-resolved density responses of a SIM.

Now $S_{\rm probe}$ is proportional to $\varphi(\mathbf{K}',\omega)$, i.e.

$$S_{\rm probe}(\mathbf{K},\omega) = C(\mathbf{K},\mathbf{K}',\omega)\varphi(\mathbf{K}',\omega), \tag{35}$$

where $C(\mathbf{K},\mathbf{K}',\omega)$ is a model-specific coefficient depending on both q and q'. The density of the induced charges can now be written as

$$\rho(\mathbf{K},\omega) = \chi(\mathbf{K},\mathbf{K}',\omega)\varphi(\mathbf{K}',\omega). \tag{36}$$

Of course, $\chi(\mathbf{K},\mathbf{K}',\omega)$ is nothing but the charge density-density response function for a SIM, which is usually studied with the Greenwood-Kubo formalism. It can be parsed as $\chi = \chi_1 + \chi_2$, with

$$\chi_1(\mathbf{K},\mathbf{K}',\omega) = -\frac{C(\mathbf{K},\mathbf{K}',\omega)}{\epsilon(K,\omega)}\frac{1}{\bar{\omega}^2} \tag{37}$$

and

$$\chi_2(\mathbf{K},\mathbf{K}',\omega) = -\frac{\bar{C}(\mathbf{K}',\omega)}{\epsilon_s(\mathbf{k},\omega)\epsilon(K,\omega)}\frac{1}{\bar{\omega}^2}. \tag{38}$$

Here

$$\bar{C}(\mathbf{K}',\omega) = -\int_0^\infty dq \frac{G(\mathbf{K},\omega)}{\bar{\omega}^2 K^2}\frac{C(\mathbf{K},\mathbf{K}',\omega)}{\epsilon(K,\omega)}. \tag{39}$$

The response function in real space, given by

$$\chi(z,z';\mathbf{k},\omega) = \left(\frac{2}{\pi}\right)^2 \times \int_0^\infty dq' \int_0^\infty dq \cos(q'z')\chi(\mathbf{K},\mathbf{K}',\omega)\cos(qz), \tag{40}$$

is more commonly encountered in the literature. One should also see that it is related to the so-called inverse dielectric function $\kappa(z,z';\mathbf{k},\omega)$ by a simple relation: $\nabla^2\kappa(z,z';\mathbf{k},\omega) + 4\pi\chi(z,z';\mathbf{k},\omega) = 0$. In general κ takes on a much more complicated form than χ.

The response function is central to many physical processes. It has been studied mostly by means of first principles computation, in which phenomenological approximations are usually invoked[100] and genuine surface effects are hard to be disclosed systematically. The present theory provides a physically transparent way to address these issues.

An identity. The functions B and C, and hence P and χ are not independent. There is a close relation between them. We notice that the probing potential in case (i), Eq. (27), can be rewritten as

$$\int_0^\infty dq'\, \varphi(\mathbf{K}',\omega)\cos(q'z)$$

with

$$\varphi(\mathbf{K}',\omega) = (2/\pi)\left(\xi(\mathbf{k},\omega)/K'^2\right).$$

The induced charge density for case (i) can then be obtained as an integral over Eq. (36), i.e.

$$\int_0^\infty dq'\, \chi(\mathbf{K},\mathbf{K}',\omega)\varphi(\mathbf{K}',\omega).$$

Equating this with Eq. (30), we arrive at the wanted relation,

$$B(\mathbf{K},\omega) = \frac{2}{\pi}\int_0^\infty \frac{dq'}{K'^2}\, C(\mathbf{K},\mathbf{K}',\omega), \tag{41}$$

or equivalently,

$$P(\mathbf{K},\omega) = \frac{2}{\pi}\int_0^\infty \frac{dq'}{K'^2}\, \chi(\mathbf{K},\mathbf{K}',\omega). \tag{42}$$

This relation shows that χ is more fundamental than P, namely the latter can be completely determined if the former is known while the converse is not true.

Despite this, it is more often the function P that is experimentally and theoretically analyzed, for example in energy losses of ions moving near a surface, in which cases the stimuli penetrate little or not at all into the metal so that case (i) applies. However, in experiments such as electron transmission through metal foils and where penetration is not negligible as well as optical experiments, the full structure of χ should be taken into account. To our knowledge, an analytical expression for χ has not been explicitly noted down even for the simplest model – the DM. In the next subsection, we discuss P and χ for the common models.

III. RESPONSES WITHIN COMMON MODELS

The theory presented in Sec. II is generic and applicable to any electron dynamics models, dispersive or non-dispersive. Different models lead to different expressions for G and Ω as well as B and C. In the literature, there are a few models that have been proposed and widely used for describing electron dynamics in metals. Here we discuss the most common ones, i.e. the DM, the HDM and the SRM, leaving the SCM to be systematically treated in Sec. IV. We consider the responses due to conduction electrons only. The contribution due to valence electrons is briefly discussed in Ref.[93].

In Table I, we summarize the defining quantities for each of the models to facilitate a quick comparison.

A. The local dielectric model (DM)

We begin the survey with the non-dispersive DM. It is the simplest model for discoursing the optical properties of metals and SPWs and often used to benchmark the validity of new methods. It is also popular for understanding electron energy loss spectroscopy and other surface related phenomena[14] such as the photon drag effect[67]. Here we reproduce the results known by this model but also some results which, up to our knowledge, have not been well discussed before. In the literature, the emphasis has been placed on the electromagnetic fields and the metal is viewed simply as a dielectric. Our theory deals with the charges directly.

The DM adopts a purely local relation between the current density and the electric field, i.e. the conductivity tensor given by $\delta_{\mu\nu}\delta(z-z')\sigma_{\rm DM}(\omega)$, with

$$\sigma_{\rm DM}(\omega) = \frac{i}{\bar\omega}\frac{\omega_p^2}{4\pi}. \tag{43}$$

Here $\omega_p = \sqrt{4\pi n_0 e^2/m}$ is the characteristic plasma frequency of a metal, with n_0 being the mean density of conduction electrons while e and m being the effective charge and mass of an electron, respectively. Symmetry breaking effects due to the surface are obviously excluded from this model. Thus,

$$\Omega = \omega_p, \quad G_s = 0, \quad G = (k/\pi)\omega_p^2. \tag{44}$$

The dielectric function $\epsilon(K,\omega)$ then takes on the form

$$\epsilon_{\rm DM}(\omega) = 1 - \omega_p^2/\bar\omega^2,$$

which underlies the usual dielectric theory of metals. The VPW frequency is ω_p by this model. Substituting the expressions of (44) into (22), we find $\epsilon_s(k,\omega)$ given by

$$\epsilon_{\rm s,DM}(\omega) = \frac{1+\epsilon_{\rm DM}(\omega)}{2\epsilon_{\rm DM}(\omega)}, \tag{45}$$

The zero of $\epsilon_{\rm s,DM}$ occurs where

$$\bar\omega = \omega_p/\sqrt{2},$$

which is the usually quoted SPW frequency. SPWs decay in this model at a rate τ^{-1}.

Let us examine the responses to exterior charges as described in Sec. II C. In the first place, we have

$$\nabla\cdot\mathbf{J}_{\rm probe}(z;\mathbf{k},\omega) = 4\pi\sigma_{\rm DM}(\omega)\rho_{\rm ext}(z;\mathbf{k},\omega),$$

which vanishes in the metal by definition. It follows that

$$S_{\rm probe}(\mathbf{K},\omega) = i\bar\omega J_{\rm probe,z}(0;\mathbf{k},\omega) = i\bar\omega\sigma_{\rm DM}(\omega)\xi(\mathbf{k},\omega), \tag{46}$$

where we have used Eq. (28). This leads to

$$B(\mathbf{K},\omega) = i\bar\omega\sigma_{\rm DM}(\omega) = -\frac{\omega_p^2}{4\pi}, \quad P_1(\mathbf{K},\omega) = \frac{1-\epsilon_{\rm DM}}{4\pi\epsilon_{\rm DM}}. \tag{47}$$

Similarly, we find

$$\bar B(\mathbf{k},\omega) = \frac{\omega_p^2}{4\pi}\frac{\omega_p^2}{2\bar\omega^2}\frac{1}{\epsilon_{\rm DM}(\omega)}, \quad P_2(\mathbf{K},\omega) = P_1(\mathbf{K},\omega)\,\frac{\epsilon_{\rm DM}-1}{\epsilon_{\rm DM}+1}. \tag{48}$$

Combining P_1 and P_2, we arrive at

$$P(\omega) := P(\mathbf{K},\omega) = P_1(\mathbf{K},\omega)/\epsilon_{\rm s,DM}(\omega) = \frac{1}{2\pi}\frac{1-\epsilon_{\rm DM}(\omega)}{1+\epsilon_{\rm DM}(\omega)}, \tag{49}$$

We see that, although P_1 features a resonance near the zero of $\epsilon(\omega)$, P does not display such a resonance. Instead, only the resonance near the zero of ϵ_s exists with P. As aforementioned, this is due to the cancellation between P_1 and P_2 near the VPW frequency, as is displayed in the upper panel of Fig. 3 for an illustration. It is clear that ABCs are incompatible with this model.

Equation (49) is one of the most used results for analyzing surface excitations and other surface phenomena such as the energy absorption of grazing particles and photon drag effect.

As for the induced charge density in this case, we see that $\rho(\mathbf{K},\omega)$ does not depend on q in this model, leading to $\rho(z;\mathbf{k},\omega) = \rho_s\Theta'(z)$ purely localized on the surface, where $\rho_s = P(\omega)\xi(\mathbf{k},\omega)$ is the areal surface charge density.

The responses to an electrostatic potential – case (ii) – can be similarly dealt with. By Eq. (34), we deduce that

$$J_{\rm probe,z}(0;\mathbf{k},\omega) = \sigma_{\rm DM}(\omega)E_{\rm probe,z}(0;\mathbf{k},\omega) = 0.$$

In addition,

$$\nabla\cdot\mathbf{J}_{\rm probe}(z;\mathbf{k},\omega) = \sigma_{\rm DM}(\omega)\varphi(\mathbf{K}',\omega)K'^2\cos(q'z).$$

TABLE I. Summary of the defining quantities of various models within the present response theory for SIMs. DM: the classical dielectric (Drude) model. HDM: the hydrodynamic model. SRM: the specular reflection model. SCM: the semi-classical model. Denote by $\rho(\mathbf{x},t)$ the density of the charges induced in the metal by a probing electric field $\mathbf{E}_{\rm probe}(\mathbf{x},t)$, and $\mathbf{E}(\mathbf{x},t)$ the electric field due to the induced charges. The current density in the metal due to $\mathbf{E}(\mathbf{x},t)$ is denoted by $\mathbf{J}(\mathbf{x},t)$. The Fourier transform of $\rho(\mathbf{x},t)$ along the surface, as defined via Eq. (7), is denoted by $\rho(z;\mathbf{k},\omega)$, where $\mathbf{k}$ is the wave vector along the surface and ω the frequency. Similar transforms are defined for other field quantities. A further cosine transform is introduced for $\rho(z;\mathbf{k},\omega)$ via Eq. (12), the q-th component of which is denoted by $\rho(\mathbf{K},\omega)$ with $\mathbf{K}=(\mathbf{k},q)$. The dielectric function of an infinite metal, ϵ is related to Ω by this relation: $\epsilon(K,\omega)=1-\Omega^2(K,\omega)/\bar{\omega}^2$. The dispersion of volume plasma waves (VPWs) is given by $\epsilon(K,\omega)=0$. Meanwhile, G serves as a kernel that plays a role in $J_z(0;\mathbf{k},\omega)=\int_0^\infty dq\left(G(\mathbf{K},\omega)/K^2\right)\rho(\mathbf{K},\omega)$. For all the models other than the SCM, $G=\omega_p^2k/\pi$, whereas for the SCM $G=\omega_p^2k/\pi+G_s$, where G_s is given by Eq. (97). For surface plasma waves (SPWs), the most important quantity is $\epsilon_s(k,\omega)=1-\int_0^\infty(dq/K^2)G(\mathbf{K},\omega)/(\Omega^2(K,\omega)-\bar{\omega}^2)$. The dispersion of SPWs is determined by $\epsilon_s(k,\omega)=0$. The presence of G_s drastically lengthens their lifetime. If the SIM is exposed to a charge of density $\rho_{\rm probe}(z;\mathbf{k},\omega)$ totally residing in the vacuum, one has $\rho(\mathbf{K},\omega)=P(\mathbf{K},\omega)\xi(\mathbf{k},\omega)$, where $\xi(\mathbf{k},\omega)=2\pi\int_{-\infty}^0 dze^{kz}\rho_{\rm probe}(z;\mathbf{k},\omega)$ and $P(\mathbf{K},\omega)=\left[B(\mathbf{K},\omega)+\epsilon_s^{-1}(k,\omega)\bar{B}(\mathbf{k},\omega)\right]/(\Omega^2(K,\omega)-\bar{\omega}^2)$ with $\bar{B}(\mathbf{k},\omega)=\int_0^\infty(dq/K^2)B(\mathbf{K},\omega)/(\Omega^2(K,\omega)-\bar{\omega}^2)$. If the SIM is exposed to an electrostatic potential $\varphi(\mathbf{K}',\omega)\cos(q'z)$, then $\rho(\mathbf{K},\omega)=\chi(\mathbf{K},\mathbf{K}',\omega)\varphi(\mathbf{K}',\omega)$, where $\chi(\mathbf{K},\mathbf{K}',\omega)=\left[C(\mathbf{K},\mathbf{K}',\omega)+\epsilon_s^{-1}(k,\omega)\bar{C}(\mathbf{K}',\omega)\right]/(\Omega^2(K,\omega)-\bar{\omega}^2)$ is the normal density-density response function with $\bar{C}(\mathbf{K}',\omega)=\int_0^\infty(dq/K^2)C(\mathbf{K},\mathbf{K}',\omega)/(\Omega^2(K,\omega)-\bar{\omega}^2)$. The SRM presumes a specularly reflecting surface in the calculation of B and C but not in G, in contrast to its original contrivance. C_s and B_s are given by the second term of Eqs. (104) and (107), respectively. The response functions P and χ are not independent but related by Eq. 42. They are of prime importance in many contexts but have not been analytically amenable until now.

Quantity	DM	HDM	SRM	SCM
$\Omega^2(K,\omega)$	ω_p^2	$\omega_p^2+K^2v_0^2$	$\omega_p^2+4\pi\bar{\omega}\mathbf{K}\cdot\mathbf{F}(\mathbf{K},\mathbf{v})/K^2$	$\omega_p^2+4\pi\bar{\omega}\mathbf{K}\cdot\mathbf{F}(\mathbf{K},\mathbf{v})/K^2$
$G(\mathbf{K},\omega)$	ω_p^2k/π	ω_p^2k/π	ω_p^2k/π	$\omega_p^2k/\pi+G_s(\mathbf{K},\omega)$
$B(\mathbf{K},\omega)$	$-\omega_p^2/4\pi$	$-\omega_p^2/4\pi$	$-\Omega^2(K,\omega)/4\pi$	$-\Omega^2(K,\omega)/4\pi+B_s(\mathbf{K},\omega)$
$C(\mathbf{K},\mathbf{K}',\omega)$	$-(K^2/8)\omega_p^2\delta(q-q')$	$-(K^2/8)\omega_p^2\delta(q-q')$	$-(K^2/8)\Omega^2(K,\omega)\delta(q-q')$	$-(K^2/8)\Omega^2(K,\omega)\delta(q-q')+C_s(\mathbf{K},\mathbf{K}',\omega)$

The corresponding $S_{\rm probe}$ is obtained as

$$S_{\rm probe}(\mathbf{K},\omega)=-\varphi(\mathbf{K}',\omega)\left(K'^2/8\right)\omega_p^2\delta(q-q'), \tag{50}$$

which leads to

$$C(\mathbf{K},\mathbf{K}',\omega)=-\left(K^2/8\right)\omega_p^2\delta(q-q'). \tag{51}$$

Substituting this in Eq. (33), we arrive at

$$\bar{C}(\mathbf{K}',\omega)=\frac{k\omega_p^2}{8\pi}\frac{\omega_p^2}{\bar{\omega}^2}\frac{1}{\epsilon_{\rm DM}(\omega)}. \tag{52}$$

Finally,

$$\chi_1(\mathbf{K},\mathbf{K}',\omega)=\frac{K^2}{8}\frac{1-\epsilon_{\rm DM}}{\epsilon_{\rm DM}}\delta(q-q') \tag{53}$$

and

$$\chi_2(\mathbf{K},\mathbf{K}',\omega)=\frac{k}{4\pi}\frac{\epsilon_{\rm DM}-1}{\epsilon_{\rm DM}+1}\frac{1-\epsilon_{\rm DM}}{\epsilon_{\rm DM}}. \tag{54}$$

Combined, they produce

$$\chi(\mathbf{K},\mathbf{K}',\omega)=\frac{1-\epsilon_{\rm DM}}{\epsilon_{\rm DM}}\left(\frac{K^2}{8}\delta(q-q')-\frac{k}{2}\,P(\omega)\right), \tag{55}$$

This result is not widely known, though an equivalent but much more involved expression has been written down in Ref.[57] for the non-local dielectric function. Most authors have considered only the responses due to SPWs, i.e. the second term in Eq. (55).

Unlike P,χ contains resonances of both VPWs and SPWs. Obviously, χ and P fulfill the relation (42).

B. The hydrodynamic model (HDM)

The DM assumes a local dependence of the current density on the electric field. In recent years there has seen lots of interest in the HDM, which is a slight extension of the DM by inclusion of some non-local effects. There are several paths, which are not always equivalent, to the HDM[106]. Here we use the fluid mechanics approach, by which the current density is given by

$$\mathbf{J}(z;\mathbf{k},\omega)=\frac{i}{\bar{\omega}}\left(\frac{\omega_p^2}{4\pi}\mathbf{E}(z;\mathbf{k},\omega)-v_0^2\nabla\rho(z;\mathbf{k},\omega)\right), \tag{56}$$

where v_0 is a parameter. The first term here is the same as in the DM, while the second one due to inter-electron interactions gives rise to non-local responses. In addition,

$$\mathbf{J}_{\rm probe}(z;\mathbf{k},\omega)=\frac{i}{\bar{\omega}}\frac{\omega_p^2}{4\pi}\mathbf{E}_{\rm probe}(z;\mathbf{k},\omega), \tag{57}$$

which has the same form as in the DM. With these two relations, one can show that

$$\Omega_{\rm HDM}^2(K)=\omega_p^2+v_0^2K^2,\quad G_s=0,\quad G=(k/\pi)\omega_p^2. \tag{58}$$

The dielectric function is then given by[106]

$$\epsilon_{\rm HDM}(K,\omega)=1-\frac{\Omega_{\rm HDM}^2(K)}{\bar{\omega}^2}. \tag{59}$$

The VPW dispersion is given by $\Omega_{\rm HDM}(K)$. The corresponding $\epsilon_s(k,\omega)$ is found to be

$$\epsilon_{s,\rm HDM}(k,\omega)=1+\frac{\omega_p^2}{2\bar{\omega}^2}\frac{k}{\pi}\int_{-\infty}^{\infty}\frac{dq}{K^2}\frac{1}{\epsilon_{\rm HDM}(K,\omega)}, \tag{60}$$

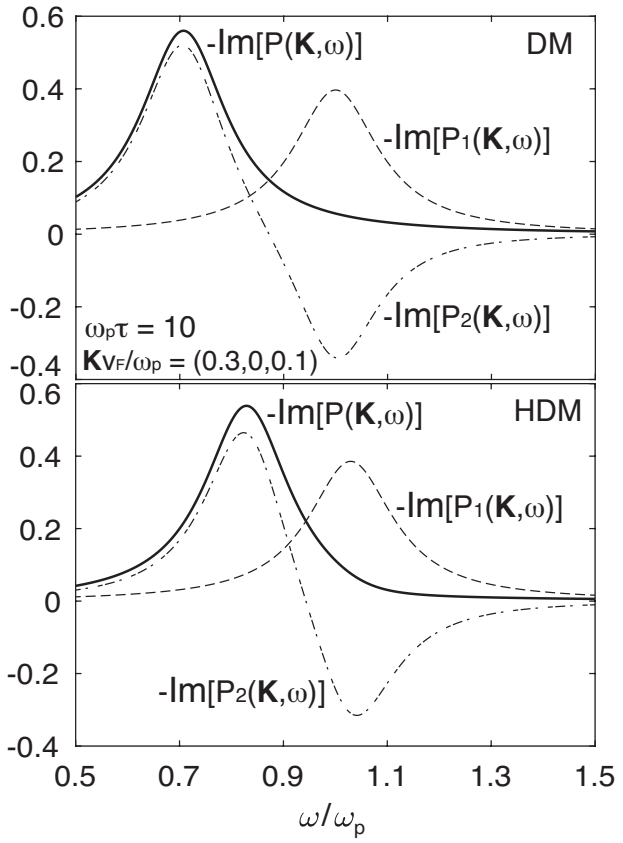

FIG. 3. The function $P(\mathbf{K},\omega) = P_1(\mathbf{K},\omega) + P_2(\mathbf{K},\omega)$ that characterizes the response to exterior charges within the DM (upper panel) and the HDM (lower panel). There is nearly complete cancellation between P_1 and P_2 near the VPW resonances and only the SPW peak appears in P. Parameters are the same in both panels.

whose zeros give the SPW dispersion in the HDM.

Equation (60) recovers $\epsilon_{\rm s,DM}$ in the limit $v_0 = 0$. By solving the equation $\epsilon_{\rm s,HDM} = 0$ we find that the SPW dispersion relation in the HDM, approximately given by

$$\left(\omega_p/\sqrt{2}\right)\left(1 + \alpha k v_0/\omega_p\right)$$

exhibits a linear k dependence. Here α is a constant of the order of unity. As thoroughly discussed in Ref.[92], the widely adopted treatment of SPWs within the HDM is incorrect and the DM cannot be recovered in that treatment.

The responses to exterior charges can easily be obtained using Eq. (57). Obviously $S_{\rm probe}$ and $B(\mathbf{K},\omega)$ are the same as in the DM, see Eqs. (46) and (47), while

$$\bar{B}(k,\omega) = \frac{\omega_p^2}{2\bar{\omega}^2}\frac{\omega_p^2}{4\pi}\frac{k}{\pi}\int_{-\infty}^{\infty}\frac{dq}{K^2}\frac{1}{\epsilon_{\rm HDM}(K,\omega)} = \frac{\omega_p^2}{4\pi}\left(\epsilon_{\rm s,HDM} - 1\right). \tag{61}$$

In obtaining the second equality we have used Eq. (60). We thus find

$$P_1(\mathbf{K},\omega) = \frac{\omega_p^2}{\bar{\omega}^2}\frac{1}{4\pi\epsilon_{\rm HDM}(K,\omega)}, \tag{62}$$

and

$$P_2(\mathbf{K},\omega) = \frac{\omega_p^2}{4\pi\bar{\omega}^2}\frac{1-\epsilon_{\rm s,HDM}(k,\omega)}{\epsilon_{\rm HDM}(K,\omega)\epsilon_{\rm s,HDM}(k,\omega)}. \tag{63}$$

Combined, they yield

$$P(\mathbf{K},\omega) = P_1(\mathbf{K},\omega)/\epsilon_{\rm s,HDM}(k,\omega), \tag{64}$$

which reduces in the limit $v_0 = 0$ to that for the DM. Again there is nearly perfect cancellation between P_1 and P_2 near the VPW resonances, as seen in the lower panel of Fig. 3. The induced charge density $\rho(\mathbf{K},\omega)$ now depends on K via $\epsilon^{-1}_{\rm HDM}(K,\omega)$. For $\omega < \omega_p$, the charges are localized within a layer of thickness around v_0/ω_p.

As for the responses to an electrostatic potential, we see that $C(\mathbf{K},\mathbf{K}',\omega)$ is also the same as in the DM, given by Eq. (51). It follows that

$$\bar{C}(\mathbf{K}',\omega) = \frac{k\omega_p^2}{8\pi}\frac{\omega_p^2}{\bar{\omega}^2}\frac{1}{\epsilon_{\rm HDM}(K',\omega)}. \tag{65}$$

Combining these expressions yields

$$\chi_1(\mathbf{K},\mathbf{K}',\omega) = \frac{K^2}{8}\frac{\omega_p^2}{\bar{\omega}^2}\frac{1}{\epsilon_{\rm HDM}(K,\omega)}\delta(q-q') \tag{66}$$

and

$$\chi_2(\mathbf{K},\mathbf{K}',\omega) = -\frac{\omega_p^2/\bar{\omega}^2}{\epsilon_{\rm HDM}(K,\omega)}\frac{\omega_p^2/\bar{\omega}^2}{\epsilon_{\rm HDM}(K',\omega)}\frac{k/8\pi}{\epsilon_{\rm s,HDM}(k,\omega)}. \tag{67}$$

Combined, they lead to

$$\chi(\mathbf{K},\mathbf{K}',\omega) = \frac{1}{8}\frac{\omega_p^2/\bar{\omega}^2}{\epsilon_{\rm HDM}(K,\omega)} \times\left(K^2\delta(q-q') - \frac{\omega_p^2/\bar{\omega}^2}{\epsilon_{\rm HDM}(K',\omega)}\frac{k/\pi}{\epsilon_{\rm s,HDM}(k,\omega)}\right). \tag{68}$$

Up to our knowledge, these functions have never been discussed in the literature, even though the HDM is a popular model for electron dynamics[44].

C. The specular reflection model (SRM)

In the HDM, Ω is approximated by $\Omega_{\rm HDM}$, which is valid only for small K. The next natural step is to use the exact form of Ω so that the dielectric function $\epsilon(K,\omega)$ becomes exact, while still neglecting the symmetry breaking effects, i.e. one approximates

$$G_s = 0, \quad G = (k/\pi)\omega_p^2. \tag{69}$$

The ensuing $\epsilon_s(k,\omega)$ then takes on the following form

$$\epsilon_{s,\rm SRM}(k,\omega) = 1 + \frac{\omega_p^2}{2\bar{\omega}^2}\frac{k}{\pi}\int_{-\infty}^{\infty}\frac{dq}{K^2}\frac{1}{\epsilon(K,\omega)}. \tag{70}$$

The VPW dispersion relation is obtained by solving the equation that $\bar{\omega} = \Omega(K, \omega)$ while the SPW dispersion relation by the following equation

$$\epsilon_{\rm s,SRM}(k, \omega) = 0. \tag{71}$$

which is nothing but the SRM equation for SPWs first proposed by Ritchie and Marusak[48] in 1966. The present derivation makes it clear that the SRM can be regarded as an extension of the HDM. In contrast to its original contrivance, the SRM does not simply assume a specularly reflecting surface in actuality; otherwise, one would have no surface contribution and Eq. (71) would not have been reached. More discussions on the logical structure of this widely used SRM are given in Ref.[93]. As with the DM and the HDM, the SRM also excludes symmetry breaking effects from G.

The responses within the SRM will be briefly discussed in the next section, in parallel with the SCM. The quantities B and C are quoted here. They are given by

$$C(\mathbf{K}, \mathbf{K}', \omega) = -K^2\Omega^2(K, \omega)\delta(q - q')/8, \tag{72}$$

$$B(\mathbf{K}, \omega) = -\Omega^2(K, \omega)/4\pi, \tag{73}$$

which are direct generalizations of the DM and HDM counterparts. Now

$$\bar{B}(k, \omega) = \frac{1}{4\pi}\frac{\omega_p^2}{2\bar{\omega}^2}\frac{k}{\pi}\int_{-\infty}^{\infty}\frac{dq}{K^2}\frac{\Omega^2(K, \omega)}{\epsilon(K, \omega)}, \tag{74}$$

which may be rewritten as

$$\bar{B}(k, \omega) = \left(\overline{\Omega^2}/4\pi\right)\Big(\epsilon_{\rm s,SRM}(k, \omega) - 1\Big),$$

where $\overline{\Omega^2}$ is defined by

$$\overline{\Omega^2} = \int_{-\infty}^{\infty}\frac{dq}{K^2}\frac{\Omega^2(K, \omega)}{\epsilon(K, \omega)} \Big/ \int_{-\infty}^{\infty}\frac{dq}{K^2}\frac{1}{\epsilon(K, \omega)}, \tag{75}$$

which is plotted in Fig. 4 (b). From these we obtain

$$P_1(\mathbf{K}, \omega) = \frac{1 - \epsilon(K, \omega)}{4\pi\epsilon(K, \omega)} = \frac{\Omega^2(K, \omega)}{\bar{\omega}^2}\frac{1}{4\pi\epsilon(K, \omega)}, \tag{76}$$

$$P_2(\mathbf{K}, \omega) = \frac{\overline{\Omega^2}}{\bar{\omega}^2}\frac{1}{4\pi\epsilon(K, \omega)}\frac{1 - \epsilon_{\rm s,SRM}(k, \omega)}{\epsilon_{\rm s,SRM}(k, \omega)}, \tag{77}$$

which closely resemble those in the HDM. If we approximate $\Omega^2 \approx \overline{\Omega^2}$, this leads to

$$P(\mathbf{K}, \omega) \approx P_1(\mathbf{K}, \omega)/\epsilon_{\rm s,SRM}(k, \omega). \tag{78}$$

This may be a good approximation for small $Kv_F/\bar{\omega}$, where Ω^2 shows little dispersion as discussed in the next section.

IV. RESPONSES BY THE SEMI-CLASSICAL MODEL

In the SCM one calculates the electrical responses due to conduction electrons in terms of a distribution function $f(\mathbf{x}, \mathbf{v}, t)$ defined in the single-particle phase space. Here $\mathbf{v} = (\mathbf{v}_\parallel, v_z)$ denotes the velocity of electrons, where $\mathbf{v}_\parallel = (v_x, v_y)$ is the planar component. As usual, we write the function as a sum of an equilibrium part $f_0(\varepsilon(\mathbf{v}))$ and a non-equilibrium part $g(\mathbf{x}, \mathbf{v}, t)$. $f_0(\varepsilon)$ is taken to be the Fermi-Dirac function at zero temperature. $\varepsilon(\mathbf{v}) = mv^2/2$ is the energy dispersion of the conduction band. Within the relaxation time approximation and the regime of linear responses, the Fourier components of $g(\mathbf{x}, \mathbf{v}, t)$ satisfy the following Boltzmann's equation

$$\Big(\lambda^{-1} + \partial_z\Big)\, g(\mathbf{v}, z; \mathbf{k}, \omega) + e f_0'(\varepsilon)\mathbf{v}\cdot\mathbf{E}(z; \mathbf{k}, \omega)/v_z = 0. \tag{79}$$

Here $\lambda = iv_z/\tilde{\omega}$ with $\tilde{\omega} = \bar{\omega} - \mathbf{k}\cdot\mathbf{v}_\parallel$ and $f_0' = \partial_\varepsilon f_0(\varepsilon)$. The electric field $\mathbf{E}(z; \mathbf{k}, \omega)$ is not specified here: it can be due to the induced charges or the probing field or the total field. As dictated by causality[96], $\gamma_0 = \mathrm{Im}(\bar{\omega})$ must be non-negative and the general solution is then given by

$$g(\mathbf{v}, z; \mathbf{k}, \omega) = e^{-\frac{z}{\lambda}}\left(C_{\mathbf{k}\omega}(\mathbf{v}) - \frac{e f_0'\mathbf{v}}{v_z}\cdot\int_0^z dz'\; e^{\frac{z'}{\lambda}}\;\mathbf{E}(z'; \mathbf{k}, \omega)\right), \tag{80}$$

where $C_{\mathbf{k}\omega}(\mathbf{v}) = g(\mathbf{v}, 0; \mathbf{k}, \omega)$ is the non-equilibrium deviation on the surface to be determined by boundary conditions. We require $g(\mathbf{v}, z; \mathbf{k}, \omega) = 0$ distant from the surface, i.e. $z \to \infty$. For electrons moving away from the surface, $v_z > 0$, this condition is automatically fulfilled. For electrons moving toward the surface, $v_z < 0$, it leads to

$$C_{\mathbf{k}\omega}(\mathbf{v}) = \frac{e f_0'\mathbf{v}}{v_z}\cdot\int_0^\infty dz'\; e^{z'/\lambda}\mathbf{E}(z'; \mathbf{k}, \omega), \quad v_z < 0, \tag{81}$$

yielding

$$g(\mathbf{v}, z; \mathbf{k}, \omega) = \frac{e f_0'\mathbf{v}}{v_z}\cdot\int_z^\infty dz'\; e^{\frac{z'-z}{\lambda}}\;\mathbf{E}(z'; \mathbf{k}, \omega), \quad v_z < 0. \tag{82}$$

To determine $C_{\mathbf{k}\omega}(\mathbf{v})$ for $v_z > 0$, the boundary condition at $z = 0$ has to be used, which, whoever, depends on the surface scattering properties. We adopt a simple picture that was first conceived by Fuchs[86] and afterwards widely used in the study of for instance anomalous skin effect[87,89,90]. According to this picture a fraction p – the *Fuchs* parameter varying between zero and unity – of the electrons impinging on the surface are specularly reflected back, i.e.

$$g(\mathbf{v}, z = 0; \mathbf{k}, \omega) = p\; g(\mathbf{v}_-, z = 0; \mathbf{k}, \omega), \tag{83}$$

where $\mathbf{v}_- = (v_x, v_y, -v_z)$ with $v_z \geq 0$. It follows that

$$C_{\mathbf{k}\omega}(\mathbf{v}) = -p\,\frac{e f_0'\mathbf{v}_-}{v_z}\cdot\int_0^\infty dz'\; e^{-\frac{z'}{\lambda}}\;\mathbf{E}(z'; \mathbf{k}, \omega), \quad v_z \geq 0. \tag{84}$$

Equations (80) - (84) fully specify the distribution function for the electrons due to a field.

The corresponding current density is calculated in the usual way,

$$\mathbf{J}(z; \mathbf{k}, \omega) = \left(\frac{m}{2\pi\hbar}\right)^3\int d^3\mathbf{v}\; e\mathbf{v}\; g(\mathbf{v}, z; \mathbf{k}, \omega). \tag{85}$$

Surface roughness enters the responses through the reflected electrons of fraction p. It is guaranteed that $J_z(0; \mathbf{k}, \omega) = 0$

for specularly reflecting surfaces ($p = 1$). Nevertheless, the charge density is not given by

$$\tilde{\rho}(\mathbf{x}, t) = (m/2\pi\hbar)^3 \, e^{i(kx-\omega t)} \int d^3\mathbf{v} \, e g(\mathbf{v}, z).$$

The reason is because Eq. (79) and hence the as-obtained $g(\mathbf{v}, z)$ is for the bulk region and not valid on the surface[102], since it involves no surface potentials, as explained in Sec. II and in previous work[98]. Actually, $\mathbf{J}(\mathbf{x}, t)$ and $\tilde{\rho}(\mathbf{x}, t)$ obey the equation

$$(\partial_t + 1/\tau)\tilde{\rho}(\mathbf{x}, t) + \partial_{\mathbf{x}} \cdot \mathbf{J}(\mathbf{x}, t) = 0$$

rather than the equation of continuity [c.f. Eq. (4)], thus automatically but incorrectly embodying the condition that $J_z(0) = 0$. This underlies the incorrect conclusion drawn by Harris[85] and calls into question many other works such as Ref.[99]. That Eq. (79) is for the bulk also justifies f_0 being simply the Fermi-Dirac function, since f_0 is the bulk equilibrium distribution without the impact of surface potential. We should also remark that Eq. (79) assumes a global relaxation term. More accurately, it may be replaced with a local relaxation term. However, the difference is a higher-order effect[88], which is negligible in the electrostatic limit concerned in the present work.

A. Expressions for $\Omega(K, \omega)$ and $G(\mathbf{K}, \omega)$

Now we specify to the case where the field in the distribution function is due to the induced charges. We substitute the expressions of $\mathbf{E}(z; \mathbf{k}, \omega)$, i.e. Eqs. (25) and (26) into (80) - (84) and perform the integration over z'. The resulting distribution function $g(\mathbf{v}, z; \mathbf{k}, \omega)$ may be split in two parts, one denoted by $g_b(\mathbf{v}, z)$ and the other by $g_s(\mathbf{v}, z)$. They are given by

$$g_b(\mathbf{v}, z; \mathbf{k}, \omega) = -e f_0' \int_0^\infty dq \, \frac{4\rho_q}{K^2} \times \tag{86}$$
$$\left[F_+(\mathbf{K}, \bar{\omega}, \mathbf{v}) \cos(qz) + iF_-(\mathbf{K}, \bar{\omega}, \mathbf{v}) \sin(qz) - F_0(\mathbf{k}, \bar{\omega}, \mathbf{v}) e^{-kz}\right],$$

where we have introduced the following functions,

$$F_\pm(\mathbf{K}, \bar{\omega}, \mathbf{v}) = \frac{\mathbf{K} \cdot \mathbf{v}}{\bar{\omega} - \mathbf{K} \cdot \mathbf{v}} \pm \frac{\mathbf{K} \cdot \mathbf{v}_-}{\bar{\omega} - \mathbf{K} \cdot \mathbf{v}_-}. \tag{87}$$

$F_\pm$ is an even/odd function of v_z. They signify the bulk responses in the presence of two counter-propagating waves $e^{\pm iqz}$ superposed in/out of phase with equal weights. In addition,

$$F_0(\mathbf{k}, \bar{\omega}, \mathbf{v}) = \frac{\mathbf{k}^* \cdot \mathbf{v}}{\bar{\omega} - \mathbf{k}^* \cdot \mathbf{v}} = \sum_{l=1}^{\infty} \left(\frac{\mathbf{k}^* \cdot \mathbf{v}}{\bar{\omega}}\right)^l, \quad \mathbf{k}^* = (\mathbf{k}, ik), \tag{88}$$

which stems from the exponential term of the electric field. The other part is given by

$$g_s(\mathbf{v}, z; \mathbf{k}, \omega) = \Theta(v_z)(-e f_0') e^{i\frac{\bar{\omega}z}{v_z}} \int_0^\infty dq \frac{4\rho_q}{K^2} \times \tag{89}$$
$$\left[F_0(\mathbf{k}, \bar{\omega}, \mathbf{v}) - pF_0(\mathbf{k}, \bar{\omega}, \mathbf{v}_-) + (p-1)F_+(\mathbf{K}, \bar{\omega}, \mathbf{v})\right].$$

One may also obtain g_b by the arguments of Ritchie and Marusak leading to the SRM[48] or directly by solving Boltzmann's equation for an infinite system. This part gives exactly the responses for an infinite system. It is independent of surface properties, i.e. showing no dependence on the *Fuchs* parameter p, and the electrons incident on the surface (i.e. with $v_z < 0$) and those departing it (i.e. with $v_z > 0$) appear on equal footing in its expression. If we keep only g_b, the SRM equation (71) will be revisited, making it evident that the SRM does not correspond to the limit of $p = 1$ (specularly reflecting surface). Instead, it corresponds to the neglect of g_s. In this sense, 'SRM' is a misnomer for the model.

On the contrary, g_s signifies pure symmetry breaking effects: it exists only for departing electrons, as indicated by the Heaviside function $\Theta(v_z)$ in its expression, and it depends on p and thus reflects on surface scattering properties. Another important feature of g_s lies in its simple dependence on z, i.e. $g_s \propto e^{i\bar{\omega}z/v_z}$. As we reasoned in Refs.[92,96–98], this factor in accord with causality implies $\gamma_0 \geq 0$ and an intrinsic instability of the metal against SPWs only to be stabilized by thermal electronic collisions.

Now we can easily find the current density and the expressions of Ω and G. Let us split the current density in two parts, $\mathbf{J}(z; \mathbf{k}, \omega) = \mathbf{J}_b(z; \mathbf{k}, \omega) + \mathbf{J}_s(z; \mathbf{k}, \omega)$, where $\mathbf{J}_{b/s}(z; \mathbf{k}, \omega)$ are defined via Eq. (85) with $g(\mathbf{v}, z; \mathbf{k}, \omega)$ replaced by $g_{b/s}(\mathbf{v}, z; \mathbf{k}, \omega)$. For small $kv_F/\bar{\omega}$, we may retain only the first term in the series of $F_0(\mathbf{k}, \bar{\omega}, \mathbf{v})$; Actually the next order contribution comes from the third term rather than the second and therefore negligible. We find that

$$\mathbf{J}_b(z; \mathbf{k}, \omega) = \sigma_{\rm DM}(\omega)\mathbf{E}(z; \mathbf{k}, \omega) + \mathbf{J}_{\rm SRM}(z; \mathbf{k}, \omega). \tag{90}$$

Here $\mathbf{J}_{\rm SRM}(z; \mathbf{k}, \omega)$ is responsible for the extension made in the SRM beyond the DM. It is given by

$$J_{{\rm SRM},x/y}(z; \mathbf{k}, \omega) = \int \mathcal{D}q\mathcal{D}^3\mathbf{v} \, v_{x/y} F_+'(\mathbf{K}, \bar{\omega}, \mathbf{v}) \cos(qz), \tag{91}$$

$$J_{{\rm SRM},z}(z; \mathbf{k}, \omega) = i \int \mathcal{D}q\mathcal{D}^3\mathbf{v} \, v_z F_-'(\mathbf{K}, \bar{\omega}, \mathbf{v}) \sin(qz), \tag{92}$$

where we have defined a short-hand

$$\int \mathcal{D}q\mathcal{D}^3\mathbf{v} ... = \left(\frac{m}{2\pi\hbar}\right)^3 \int_0^\infty dq \, \frac{4\rho_q}{K^2} \int d^3\mathbf{v} \left(-e^2 f_0'\right) ...$$

together with these functions

$$F_\pm'(\mathbf{K}, \bar{\omega}, \mathbf{v}) = \frac{1}{2}\left[\frac{(\mathbf{K} \cdot \mathbf{v})^2}{1 - \mathbf{K} \cdot \mathbf{v}/\bar{\omega}} \pm \frac{(\mathbf{K} \cdot \mathbf{v}_-)^2}{1 - \mathbf{K} \cdot \mathbf{v}_-/\bar{\omega}}\right].$$

See that $J_{{\rm SRM},z}(0; \mathbf{k}, \omega) \equiv 0$, which means that $\mathbf{J}_{\rm SRM}$ makes no contribution to G. One thus concludes that

$$G_b = (k/\pi)\omega_p^2$$

as with the DM and other models.

By their definitions, Eqs. (10), (15) and (16), we directly find that

$$\Omega^2(K, \omega) = \omega_p^2 + \frac{4\pi\bar{\omega}\mathbf{K} \cdot \mathbf{F}(\mathbf{K}, \bar{\omega})}{K^2}, \tag{93}$$

where $\mathbf{F}(\mathbf{K},\bar{\omega})$ is an odd function of $\bar{\omega}$ and given by

$$\mathbf{F}(\mathbf{K},\bar{\omega}) = \left(\frac{m}{2\pi\hbar}\right)^3 \int d^3\mathbf{v}\left(-e^2 f_0'\right)\left(\frac{\mathbf{K}\cdot\mathbf{v}}{\bar{\omega}}\right)^2 \frac{\mathbf{v}}{1-\mathbf{K}\cdot\mathbf{v}/\bar{\omega}}. \tag{94}$$

See that $\mathbf{K}\cdot\mathbf{F}$ does not depend on the direction of $\mathbf{K}$. Additionally, we have

$$G_s(\mathbf{K},\omega) = 4i\bar{\omega}\left(\frac{m}{2\pi\hbar}\right)^3 \int_> d^3\mathbf{v}\; v_z\left(-e^2 f_0'\right)\times \left[F_0(\mathbf{k},\bar{\omega},\mathbf{v}) - pF_0(\mathbf{k},\bar{\omega},\mathbf{v}_-) + (p-1)F_+(\mathbf{K},\bar{\omega},\mathbf{v})\right], \tag{95}$$

which strongly depends on p. Here the integral is restricted to $v_z \geq 0$, as indicated by the symbol '>'.

The second term in Eq. (93) is generally complex even in the collision*less* limit where τ^{-1} is vanishingly small, due to a pole at $\bar{\omega} = \mathbf{K}\cdot\mathbf{v}$ in the integrand in $\mathbf{F}$. The imaginary part of Ω^2 gives rise to Landau damping, i.e. the damping due to the excitation of particle-hole pairs. Its real part approximates

$$\omega_p^2 + \frac{3}{5}K^2 v_F^2$$

for small K, which revisits $\Omega_{\rm HDM}$ with $v_0 = \sqrt{\frac{3}{5}}\, v_F$. The integral in the expression of $\mathbf{F}$ can be partially performed. Doing this leads to

$$\Omega^2(K,\omega) = \omega_p^2\left(1 + \frac{3}{2}\frac{Kv_F}{\bar{\omega}}\int_{-1}^{1} dr\,\frac{r^3}{1-rKv_F/\bar{\omega}}\right). \tag{96}$$

It shows that Ω depends on K and ω not individually, but only through the ratio $Kv_F/\bar{\omega}$. In Fig. 4 (a), Ω is plotted, where it is seen that the real (imaginary) part of Ω^2 is even (odd) in ω, a property that can be rigorously proved by use of the relation that $\mathbf{F}(\mathbf{K},\bar{\omega}) + \mathbf{F}(\mathbf{K},-\bar{\omega}) = 0$. The imaginary part displays a minimum on the physical (positive) frequency side, due to particle-hole excitations produced at $\omega = Kv_F$ that is responsible for Landau damping.

A crucial improvement of the SCM over the SRM comes through the quantity $G_s(\mathbf{K},\omega)$. In the SRM and its descendents, $G_s = 0$ and no symmetry breaking effects are present. As shown in Refs.[92,96–98], thanks to G_s, an instability of the metal might be induced at some critical point, where SPWs become lossless with infinitely long lifetime – a highly desirable attribute in plasmonics and other practical areas of SPWs. For small $kv_F/\bar{\omega}$, we may keep only the first term in the series of $F_0(\mathbf{k},\bar{\omega},\mathbf{v})$, and G_s can be rewritten as

$$G_s(\mathbf{K},\omega) = -\frac{1+p}{2}\frac{k}{\pi}\omega_p^2 + 4i\bar{\omega}(p-1)\left(\frac{m}{2\pi\hbar}\right)^3 \int_> d^3\mathbf{v}\left(-e^2 f_0'\right) v_z F_+(\mathbf{K},\bar{\omega},\mathbf{v}). \tag{97}$$

A comparison between this expression and Eq. (95) is displayed in Fig. 5; they agree with each other very well, especially for not so big kv_F/ω. The first term of expression (97) can be absorbed in G_b. It renormalizes the SPW frequencies and renders the latter surface specific, i.e. dependent on the *Fuchs* parameter p. The second term is mostly imaginary and responsible for the aforementioned instability. It is easy to see that $G = 0$ for $p = 1$, as expected of specularly reflecting surfaces. Thus, the SRM is not the same as the limit $p = 1$, in contrast with its intended meanings.

With Ω and G, one can obtain $\epsilon_s(k,\omega)$ using the definition, Eq. (22). The ensuing expression cannot be further simplified and it is thus not repeated here.

B. The functions $\chi(\mathbf{K},\mathbf{K}',\omega)$ and $P(\mathbf{K},\omega)$

To obtain the response functions, let us specify the expressions, (80) – (84) for the electronic distribution to the case where the field represents the probing field. The resulting distribution function is to be called $g_{\rm probe}(\mathbf{v},z;\mathbf{k},\omega)$. Substituting this for g in Eq. (85), one easily obtains $\mathbf{J}_{\rm probe}(z;\mathbf{k},\omega)$ and $S_{\rm probe}(\mathbf{K},\omega)$.

We first establish $\chi(\mathbf{K},\mathbf{K}',\omega)$ by considering the responses in case (ii) described in Sec. II C, to an electrostatic potential. The distribution function follows from Eqs. (80) – (84). It can be written as

$$g_{\rm probe} = g_{\rm prob} + \Theta(v_z)(p-1)g_{\rm pros},$$

where

$$g_{\rm prob}(\mathbf{v},z;\mathbf{k},\omega) = -ef_0'\varphi(\mathbf{K}',\omega)\times \frac{1}{2}\left[F_+(\mathbf{K}',\bar{\omega},\mathbf{v})\cos(q'z) + iF_-(\mathbf{K}',\bar{\omega},\mathbf{v})\sin(q'z)\right] \tag{98}$$

and

$$g_{\rm pros}(\mathbf{v},z;\mathbf{k},\omega) = -\frac{1}{2}ef_0'\varphi(\mathbf{K}',\omega)F_+(\mathbf{K}',\bar{\omega},\mathbf{v})e^{\frac{\bar{\omega}z}{v_z}}. \tag{99}$$

Now $\mathbf{J}_{\rm probe} = \mathbf{J}_{\rm prob} + \mathbf{J}_{\rm pros}$ accordingly splits, where

$$\mathbf{J}_{\rm prob}(z;\mathbf{k},\omega) = \left(\frac{m}{2\pi\hbar}\right)^3 \int d^3\mathbf{v}e\mathbf{v}g_{\rm prob}(\mathbf{v},z;\mathbf{k},\omega) \tag{100}$$

and

$$\mathbf{J}_{\rm pros}(z;\mathbf{k},\omega) = (p-1)\left(\frac{m}{2\pi\hbar}\right)^3 \int_> d^3\mathbf{v}e\mathbf{v}g_{\rm pros}(\mathbf{v},z;\mathbf{k},\omega). \tag{101}$$

By the fact that F_+ is an even function of v_z, one concludes

$$J_{{\rm prob},z}(0;\mathbf{k},\omega) \equiv 0.$$

Straightforward manipulations show that

$$\int_0^\infty dz\cos(qz)\nabla\cdot\mathbf{J}_{\rm prob}(z) = -\varphi(\mathbf{K}',\omega)\frac{K^2\Omega^2(K,\omega)}{8i\bar{\omega}}\,\delta(q-q'). \tag{102}$$

Similarly, we have

$$J_{{\rm pros},z}(0;\mathbf{k},\omega) + \int_0^\infty dz\cos(qz)\nabla\cdot\mathbf{J}_{\rm pros}(z;\mathbf{k},\omega) = \frac{1-p}{4}\varphi(\mathbf{K}',\omega) \times\left(\frac{m}{2\pi\hbar}\right)^3 \int_> d^3\mathbf{v}(-e^2 f_0')v_z F_+(\mathbf{K},\bar{\omega},\mathbf{v})F_+(\mathbf{K}',\bar{\omega},\mathbf{v}). \tag{103}$$

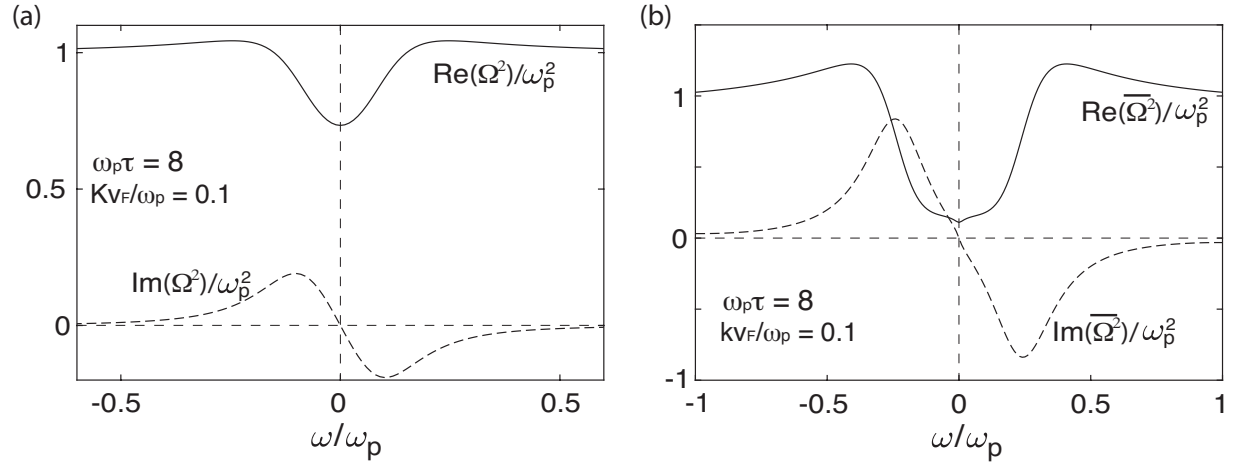

FIG. 4. Frequency dependence of (a) $\Omega^2(K,\omega)$ [Eq. (96)] and (b) $\overline{\Omega^2}(k,\omega)$ [Eq. (75)]. Re$[\Omega^2(K,\omega)]$ is even in ω whereas Im$[\Omega^2(K,\omega)]$ is odd in ω. At $\omega = 0$, Ω is real. Ω^2 depends on K and $\bar{\omega}$ via the combination $Kv_F/\bar{\omega}$, rather than individually. Similar properties hold for $\overline{\Omega^2}(k,\omega)$. For large ω, these two quantities become comparable.

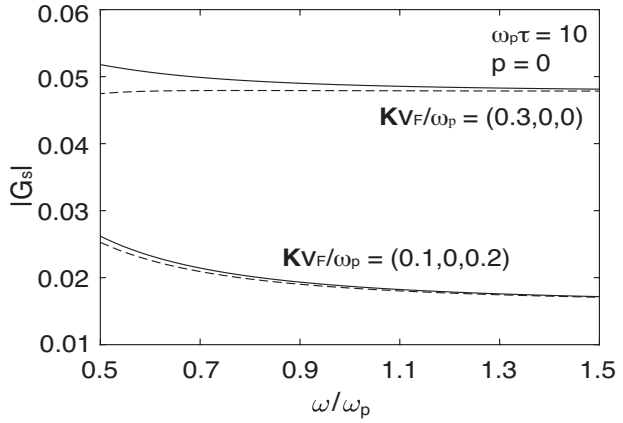

FIG. 5. Illustration of $G_s(\mathbf{K},\omega)$, which contains symmetry breaking effects and disappears from all the models except the SCM. Solid line: Eq. (95). Dashed line: Eq. (97).

With these expressions we can obtain $S_{\text{probe}}(\mathbf{K},\omega)$ by use of its definition and thence

$$C(\mathbf{K},\mathbf{K}',\omega) = -\frac{K^2\Omega^2(K,\omega)}{8}\,\delta(q-q') \tag{104}$$
$$+\frac{1-p}{4}i\bar{\omega}\left(\frac{m}{2\pi\hbar}\right)^3\int_> d^3\mathbf{v}(-e^2f_0')v_zF_+(\mathbf{K},\bar{\omega},\mathbf{v})F_+(\mathbf{K}',\bar{\omega},\mathbf{v}).$$

Inserting this into Eqs. (37) – (39), one obtains the semi-classical response function $\chi(\mathbf{K},\mathbf{K}',\omega)$, which can be written in the following form

$$\chi(\mathbf{K},\mathbf{K}',\omega) = \frac{C(\mathbf{K},\mathbf{K}',\omega)+\epsilon_s^{-1}(k,\omega)\bar{C}(\mathbf{K}',\omega)}{\Omega^2(K,\omega)-\bar{\omega}^2}. \tag{105}$$

with C given by Eq. (104), which further gives $\bar{C}$ via (39).

The responses to exterior charges are encoded in the function $P(\mathbf{K},\omega)$, which is defined in Sec. II C. One can establish P in a similar fashion as we did with χ, i.e. one could first find the corresponding g_{probe} and then uses it to calculate $\mathbf{J}_{\text{probe}}$ and other quantities including P. On the other hand, we can also directly obtain $P(\mathbf{K},\omega)$ from $\chi(\mathbf{K},\mathbf{K}',\omega)$ by means of the relation (42). For this purpose, it suffices to obtain $B(\mathbf{K},\omega)$ from $C(\mathbf{K},\mathbf{K}',\omega)$ via the relation (41). By the method of contour integral, one can easily show that

$$\frac{1}{\pi}\int_0^\infty\frac{dq}{K^2}F_+(\mathbf{K},\bar{\omega},\mathbf{v}) = F_0(\mathbf{k},\bar{\omega},\mathbf{v})/k, \tag{106}$$

with which we immediately arrive at

$$B(\mathbf{K},\omega) = -\frac{\Omega^2(K,\omega)}{4\pi} \tag{107}$$
$$+\frac{1-p}{2}i\bar{\omega}\left(\frac{m}{2\pi\hbar}\right)^3\int_> d^3\mathbf{v}(-e^2f_0')v_zF_+(\mathbf{K},\bar{\omega},\mathbf{v})F_0(\mathbf{k},\bar{\omega},\mathbf{v})/k.$$

Here the first term originates from $\nabla\cdot\mathbf{J}_{\text{pros}}$. Now $P(\mathbf{K},\omega)$ can be directly obtained from these expressions by definition. It can be written as

$$P(\mathbf{K},\omega) = \frac{B(\mathbf{K},\omega)+\epsilon_s^{-1}(k,\omega)\bar{B}(\mathbf{k},\omega)}{\Omega^2(K,\omega)-\bar{\omega}^2}. \tag{108}$$

with B given by Eq. (107), which further gives $\bar{B}$ via (33). An example of P is plotted in Fig. 6 (a). At large K, the SPWs and VPWs are well separated in frequencies and -Im$[P]$ displays two peaks.

Setting $p = 1$ in the expressions of B and C while neglecting G_s, one arrives at the response functions quoted for the SRM, Eqs. (72) and (73). The as-defined SRM, however, is not identical with the usually adopted SRM, see Ref.[93]. In Fig. 6 a comparison is plotted between the SCM [panel (a)] and the SRM [panel (b)].

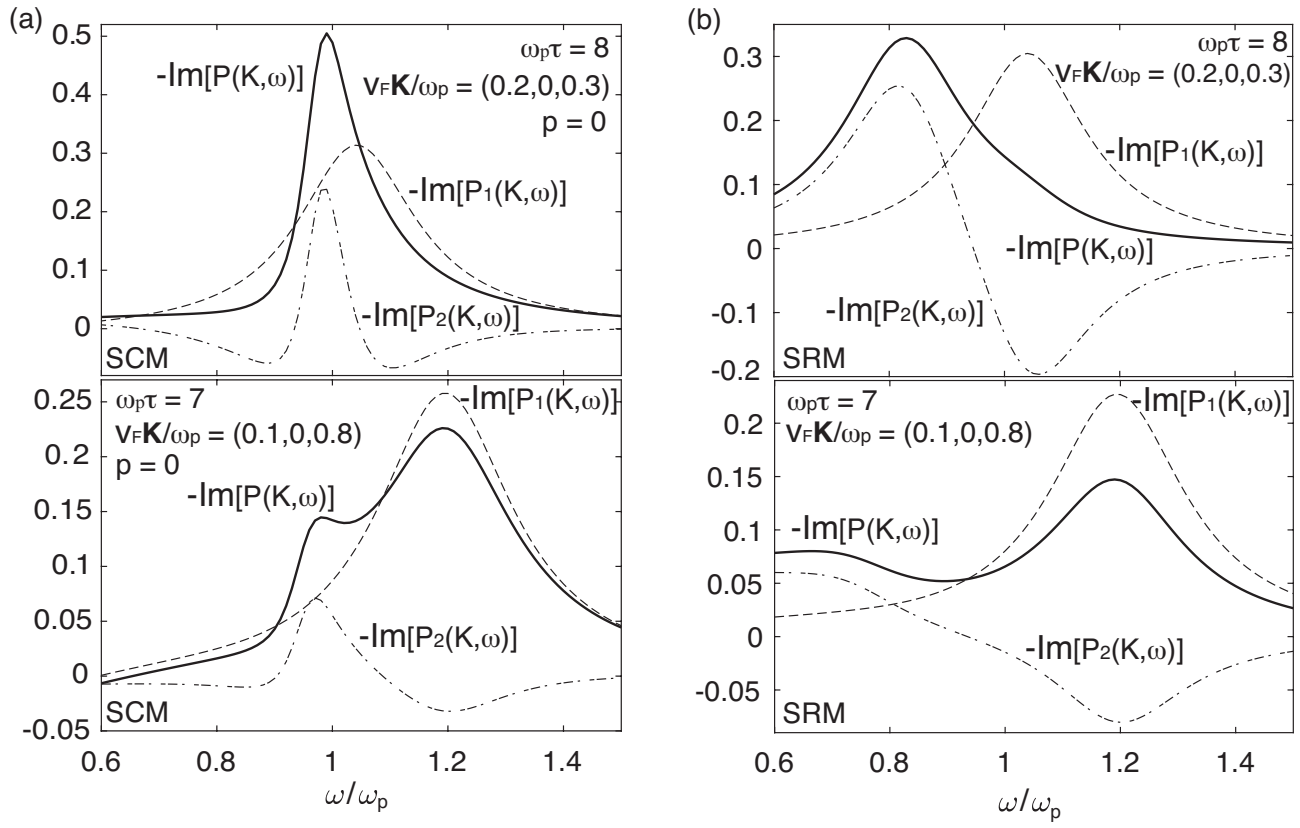

FIG. 6. Frequency dependence of $P(\mathbf{K},\omega) = P_1(\mathbf{K},\omega) + P_2(\mathbf{K},\omega)$ in (a) the SCM and (b) the SRM at various values of $\mathbf{K}$.

V. DISCUSSIONS

We have developed a general dynamical response theory for SIMs. This theory is straightforwardly extendable to other bounded systems such as films and spheres. We have applied it to discuss the responses within several dispersive and non-dispersive common electron dynamics models in addition to the less common SCM. Analytical expressions have been obtained of the density-density response function $\chi(\mathbf{K},\mathbf{K}',\omega)$, which is probed in virtually every physical process involving surfaces, examples including particle scattering[23,53,101] to be discussed in what follows, the scattering of electromagnetic waves[107], photon drag effect[67], secondary electron emission process (e.g. Auger process) and ion neutralization process[63] as well as energy dissipation of objects (e.g. quantum dots and molecules) in the proximity of surfaces[64,65] in addition to quantum forces such as quantum friction and Casimir forces[62]. These processes are interesting in themselves and they underpin many spectroscopies vital for studying the electronic and optical properties of solids. Applying the theory to these physical processes should be a fascinating subject of future study.

Our theory requires neither MBCs nor ABCs, which have been avoided by means of the general macroscopic limit of physical boundaries. The entire issue of ABCs has thence been sidestepped. Introduced over six decades ago and having been adopted in innumerable work, ABCs are widely regarded as superficial without a generic physical basis and should not play any role in a complete theory[21,76]. Our theory reveals that the density response function is comprised of two parts, one of which is directly associated with the excitation of VPWs while the other occurs purely because of the surface capacitive effects and signifies the excitation of SPWs. The ABCs would make the surface part disappear and they are incompatible with non-dispersive models. Our theory calls for a reappraisal of massive experimental data that have been interpreted on the basis of ABCs.

We are aware of some other work aiming to solve the problem of ABCs. As mentioned in the beginning section of this paper, the earliest effort perhaps dated back to 1970s based on Ewald-Oseen extinction theorem within the dielectric approximation, which has recently been further developed by Schmidt *et al.*[6,7]. Another line was taken in the 1990s by Chen *et al.* using their wave-vector-space method[2,103]. In the simplest case of local dielectric models, their approach is actually identical to the present one[103]. In the development of dispersive models appropriate for the media of excitons, their method is microscopic rather than macroscopic[2], allowing them to derive a set of ABCs for the excitons. In addition, K. Henneberger[83] introduced a controversial source term to mimick the surface effects, which in our opinion resembles the fictitious charge sheet in the SRM and may be regarded

as an implicit type of ABCs. Finally, a few years ago[105] M. Apostol and G. Vaman also proposed a method that bypasses the ABCs. These authors based their scheme on the concept of a displacement field that is exclusive to the HDM, which is the only model under their consideration. A generalization of their model may be possible if the displacement field is replaced by a more general concept such as the polarization field. As far as the HDM is concerned, their scheme is similar to the present theory.

In the rest of this section, we employ the theory to evaluate the dynamical structure factor, which plays an important role in particle scattering with metal surfaces and in EELS, and the spatial distribution of charges induced by a charged particle grazing over a metal surface. The main purpose here is to differentiate the various dynamics models. We expect the results to be experimentally interesting.

A. Dynamical structure factor and SPW peak narrowing

$\chi(\mathbf{K},\mathbf{K}',\omega)$ is one of the most fundamental quantities for characterizing the responses of a bounded system and pivotal in the interpretation of a variety of experiments. A systematic analysis of its properties being reserved for a separate publication, here we briefly discuss its use in the study of charged particles (e.g. electrons) reflected off a metal surface. The quantity of interest here is the dynamical structure factor $\mathcal{S}$, which appears in the differential scattering cross section per unit surface area (DCS) in the following manner[101],

$$\text{DCS} \propto \frac{K_f}{K_i}\frac{Q^2}{k^2}\mathcal{S}(\Delta\mathbf{K},\omega), \tag{109}$$

where $\hbar\mathbf{K}_i$ and $\hbar\mathbf{K}_f$ are the incoming and outgoing momenta of the incident particle of charge Q, and $\hbar\Delta\mathbf{K} = \hbar(\mathbf{K}_i - \mathbf{K}_f) = \hbar(\mathbf{k},\Delta k)$ is the momentum exchange during the scattering and $\hbar\omega$ denotes the energy exchange.

In the so-called dipole approximation[53], the particles are assumed to penetrate negligibly into the metal and one has

$$\mathcal{S}(k,\omega) = -\frac{2}{\pi^2}\text{Im}\left[\int_0^\infty \frac{dq}{K^2}\,P(\mathbf{K},\omega)\right]. \tag{110}$$

Here we have suppressed the dependence of $\mathcal{S}$ on Δk. In this approximation, it is P that is directly probed rather than the full spectrum of χ.

In Fig. 7 is exhibited an example of $\mathcal{S}(k,\omega)$, where the left panel is according to the SCM while the right panel to the SRM. The result for the HDM differs only slightly from that for the SRM. In the plots, we have made the decomposition that $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$, where $\mathcal{S}_{1,2}$ are defined via Eq. (110) with P replaced with $P_{1,2}$; see Sec. II. Only the SPW peak is seeable in $\mathcal{S}(k,\omega)$. This peak is asymmetric in the SCM whereas symmetric in other models – a result of symmetry breaking effects in G_s, which strongly modify the shape of $\mathcal{S}_2(k,\omega)$. As seen in the figure, $\mathcal{S}_1$ has almost the same shape in the SCM as in the SRM, while in the SCM $\mathcal{S}_2$ has a much sharper peak that is far closer to the peak in $\mathcal{S}_1$. At small k this asymmetry becomes less pronounced and eventually disappears.

As another consequence of the symmetry breaking effects, the width of the SPW peak appears much smaller in the SCM than in other models. It is even much smaller than $1/\tau$, a scenario inexplicable by the conventional wisdom[108,109]. According to the latter, it can by no means become short of $1/\tau$. This peak narrowing has practical implications for plasmonics and nano photonics, as discussed in recent papers[92,96–98] and briefly recapitulated in Ref.[93].

The SPW peak width can in principle be made as small as desirable due to a criticality in the system. The criticality can be disclosed in $\mathcal{S}(k,\omega)$. For stable systems, $\mathcal{S}$ must stay positive-definite conforming to the fluctuation-dissipation theorem. For a system containing an instability, however, $\mathcal{S}$ crosses zero at the corresponding critical point to assume unphysical negative values[110]. Back to the present case, we note that $\mathcal{S}$ contains two parts $\mathcal{S}_1$ and $\mathcal{S}_2$ canceling each other, as seen in Fig 7 (a). As shown in Ref.[93], upon decreasing $1/\tau$, $\epsilon_s(k,\omega)$ can be made to vanish and hence $\mathcal{S}_2$ can be made singular around the SPW pole whereas $\mathcal{S}_1$ is dominated by Landau damping via the VPW pole and much less affected. As a result, there exist a critical value of τ, across which $\mathcal{S}$ changes sign from positive to negative near the SPW pole, thereby signifying an instability of the system. At the critical point, SPWs are lossless, as should be for any critical phenomena. In Ref.[92], we have put forth a proposal on how to realize this instability. The nature of this criticality is currently under investigation within a quantum mechanical theory.

To gain some insights into the narrowing of the SPW peak, let us examine the limit of small k. It should be cautioned that at very small k retardation effects may play a role and our theory needs to be modified; see Ref.[93,111] for discussions on this matter. For very small k, we note that $k/K^2 \approx \pi\delta(q)$. Using this, we find

$$\mathcal{S}(k,\omega) \approx -\frac{1}{\pi k}\text{Im}\left[P(\mathbf{K}_0,\omega)\right],$$

where $\mathbf{K}_0 = (\mathbf{k},0)$. With the same strategy, we find

$$\bar{B}(k,\omega) = \frac{\pi}{2k}\frac{G(\mathbf{K}_0,\omega)B(\mathbf{K}_0,\omega)}{\omega_p^2 - \bar{\omega}^2}, \tag{111}$$

$$\epsilon_s(k,\omega) = 1 - \frac{\pi}{2k}\frac{G(\mathbf{K}_0,\omega)}{\omega_p^2 - \bar{\omega}^2}. \tag{112}$$

Here we have used $\Omega \approx \omega_p$ for small k. Expressions of G and B can similarly be found for small k. They are given by

$$G(\mathbf{K}_0,\omega) \approx \frac{k\omega_p^2}{\pi}\frac{1-p}{2}\left(1 - i\frac{3kv_F}{\bar{\omega}}\right), \tag{113}$$

$$B(\mathbf{K}_0,\omega) \approx -\frac{\omega_p^2}{4\pi}\left(1 + i\frac{3(1-p)}{8}\frac{kv_F}{\bar{\omega}}\right). \tag{114}$$

Combining the above expressions, we obtain

$$P(\mathbf{K}_0,\omega) \approx \frac{B(\mathbf{K}_0,\omega)}{\omega_p^2 - \bar{\omega}^2}\frac{1}{\epsilon_s(k,\omega)} = \frac{B(\mathbf{K}_0,\omega)}{\omega_s^2\left(1 + i\frac{3(1-p)}{3+p}\frac{kv_F}{\bar{\omega}}\right) - \bar{\omega}^2}. \tag{115}$$

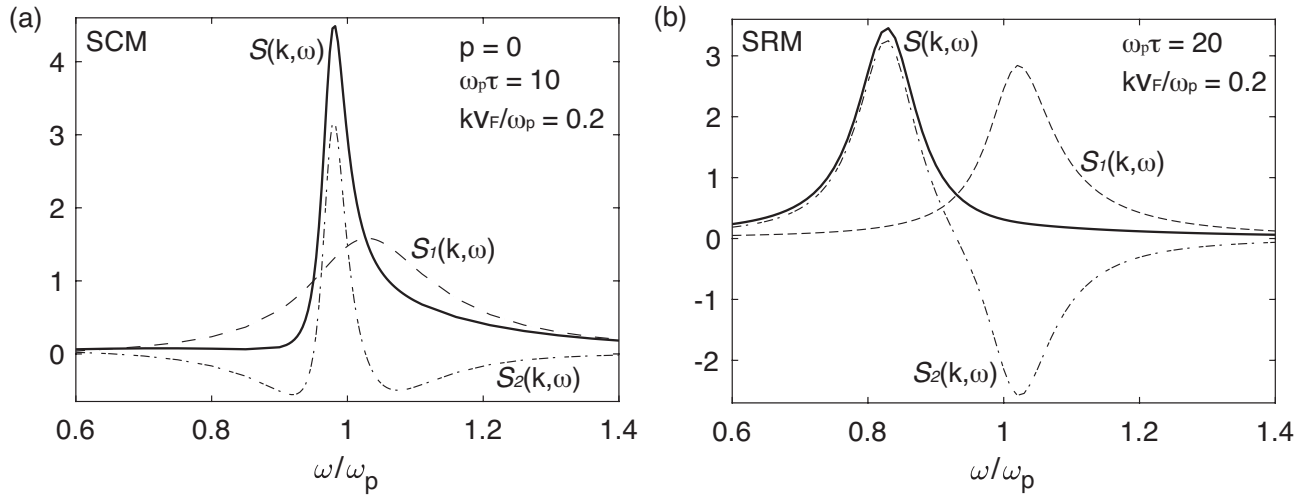

FIG. 7. The dynamical structure factor $\mathcal{S}(k,\omega) = \mathcal{S}_1(k,\omega) + \mathcal{S}_2(k,\omega)$ in the dipole approximation according to (a) the SCM and (b) the SRM. Only the SPW peak is seen in $\mathcal{S}(k,\omega)$. The peak in the SCM is significantly sharper than in the SRM, even though in the latter a bigger value of τ has been used. The curve by the HDM – not shown – is only slightly different from the SRM curve for the same parameters.

Here $\omega_s = \sqrt{\frac{3+p}{4}}\omega_p$. A little more manipulation shows that

$$P(\mathbf{K}_0,\omega) \approx \frac{B(\mathbf{K}_0,\omega)/\left(1+i\frac{3(1-p)}{3+p}\frac{kv_F}{\bar{\omega}}\right)}{\omega_s^2-(\omega+i\gamma)^2},\ \gamma = \frac{1}{\tau}-\gamma_0, \quad (116)$$

where $\gamma_0 = \frac{3(1-p)}{2(3+p)}kv_F$. This expression shows that the effective collision rate γ is reduced relative to its bare value τ^{-1} by an amount of γ_0. This reduction occurs solely because of the imaginary part of G_s, which is absent in other models than the SCM. $P(\mathbf{K}_0,\omega)$ displays a peak at ω_s with width γ, which represents the excitation of SPWs.

As expected, both ω_s and γ depend on surface roughness via the *Fuchs* parameter p. Such dependence is absent from other models than the SCM. A detection (an absence) of this dependence would constitute a strong evidence in support of (against) the SCM. Experimentally, it has been demonstrated that p can be widely tuned in some materials such as copper[112,113].

The long-wavelength SPW frequency in the SCM is $\omega_s \approx 0.87\omega_p$ for diffusely scattering surfaces, which is considerably higher than $0.71\omega_p$ obtained with other models. On the basis of a specific microscopic model within random-phase approximation, Feibelman argued that the SPW frequency should take on the latter value regardless of the microscopic electron density profile near the surface[114]. The solution he found with frequency $0.71\omega_p$ has a constant electrostatic potential and is hence empty of charges, which falls in the category of false solutions mistakenly assigned as standing for SPWs[92]. To discriminate between these two values, a main difficulty lurks in the determination of ω_p. Let us take Al for the sake of illustration. Nominal charge counting gives 15eV for $\hbar\omega_p$ in this metal, whereas first principles computation[115] yields 12.6eV. Now that the measured SPW frequency[16] is 10.7eV in Al, the former would come in favor of $0.71\omega_p$ while the latter of $0.87\omega_p$. This example calls for more effort to be invested in clarifying this issue in the future.

The dipole approximation, despite its widespread use, is incapable of satisfactorily reproducing the experimental observations. In this approximation, $\mathcal{S}(k,\omega)$ displays only the SPW peak, though an additional broad peak due to VPWs has been seen in numerous scattering experiments[23,101]. Several proposals have been evoked to address the discrepancy[101]. We shall address this issue comprehensively elsewhere. In the rest of this section, we discuss the issue in terms of the induced charges.

B. Charges induced by a grazing particle

For simplicity, let us consider a particle of unit charge grazing over a metal surface at distance z_0 and constant velocity $\mathbf{V} = (V,0,0)$, as shown in Fig. 1 (a). The associated charge density is given by $\rho_{\text{probe}}(\mathbf{x},t) = \delta^3(\mathbf{x}-\mathbf{V}t)$, or equivalently

$$\rho_{\text{probe}}(z;\mathbf{k},\omega) = (2\pi/\sqrt{A})\delta(z+z_0)\delta(\omega-k_xV).$$

It follows that

$$\xi(\mathbf{k},\omega) = (4\pi^2/\sqrt{A})e^{-kz_0}\delta(\omega-k_xV).$$

The induced charge density is given by

$$\rho(\mathbf{x},t) = \sum_{\mathbf{k}}\frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{A}}\frac{2}{\pi}\int_0^\infty dq\cos(qz)\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\rho(\mathbf{K},\omega)e^{-i\omega t},$$

which upon using the results in Sec. II C becomes

$$\rho(\mathbf{x},t) = \int d^2\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}(t)}e^{-kz_0}\frac{2}{\pi}\int_0^\infty dq\cos(qz)P(\mathbf{K},k_xV). \quad (117)$$

where the sum over $\mathbf{k}$ has been converted into an integral and $\mathbf{r}(t) = \mathbf{r} - \mathbf{V}_\parallel t$ with $\mathbf{V}_\parallel = (V, 0)$. Without loss of generality, $t = 0$ is taken in all numerical plots. With the expressions of $P(\mathbf{K}, \omega)$ obtained in previous sections, $\rho(\mathbf{x}, t)$ can be evaluated. It is noted that the factor e^{-kz_0} effectively suppresses the contributions from $k \gg 1/z_0$ to the integral over $\mathbf{k}$ in the expression. For large z_0 only components with small k contribute, whereas for small z_0 large-k components also contribute.

In the DM, $P(\mathbf{K}, \omega)$ does not depend on q and hence the induced charge density, which we call $\rho_{\rm DM}(\mathbf{x}, t)$, is completely localized on the surface, i.e. $\rho_{\rm DM}(\mathbf{x}, t) = 2\rho_s(\mathbf{r}, t)\delta(z)$, with the areal density given by

$$\rho_s(\mathbf{r}, t) = \frac{1}{2\pi} \int d^2\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}(t) - kz_0} \frac{\omega_p^2/2}{(k_x V + i/\tau)^2 - \omega_p^2/2}. \quad (118)$$

In the limit $z_0 \gg V\tau$, one may disregard $k_x V$ and

$$\rho_s(\mathbf{r}, t) \approx -\frac{1}{2\pi} \frac{\omega_p^2}{2\tau^{-2} + \omega_p^2} \int d^2\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}(t) - kz_0},$$

which has a circular shape with radius $\sim z_0$. For not so large z_0, the distribution is anisotropic around the grazing particle. An example is shown in Fig. 1 (b), where the DM is contrasted with the SCM (of the diffuse limit $p = 0$) in terms of the planar charge distribution $\rho_\parallel(\mathbf{r}, t) = \int dz\, \rho(\mathbf{x}, t)$, which equals $\rho_s(\mathbf{r}, t)$ in the DM. For small z_0 (the panels with $z_0\omega_p/v_F = 5$), in both models $\rho_\parallel(\mathbf{r}, t)$ is periodic along the y-direction but with a smaller wavelength in the former. For moderate z_0 (the panels with $z_0\omega_p/v_F = 15$), however, $\rho_\parallel(\mathbf{r}, t)$ strongly depends on the model: in the DM it is symmetric about the grazing particle along its motion but in the SCM the charges are more concentrated in front of the particle.

The aforementioned symmetry is preserved in the HDM but not in the SRM, as seen in Fig. 8. In this figure, the panels are organized in eight pairs, each pair consisting of two panels in the same model and with the same z_0. The left panel in a pair shows $\rho(\mathbf{x}_0, t)$ while the right one shows $\rho_\parallel(\mathbf{r}, t)$. For comparison, we have also displayed results for the SCM of the reflection limit $p = 1$. For small z_0, $\rho_\parallel(\mathbf{r}, t)$ exhibits in the SRM, the HDM and the SCM of $p = 1$ the same periodic and symmetric pattern as in the DM, though its magnitude strongly depends on the models. For moderate z_0, $\rho_\parallel(\mathbf{r}, t)$ remains symmetric in the HDM and the SCM of $p = 1$ but not so in the SCM of $p = 0$ and the SRM. In general, $\rho(\mathbf{x}_0, t)$ varies much more mildly than $\rho_\parallel(\mathbf{r}, t)$ along the surface.

The depth dependence of the induced charge density is illustrated in Fig. 9. Here the panels are also grouped in eight pairs, each consisting of two panels in the same model and with the same value of z_0. The left panel in a pair displays the distribution of the induced charges in the plane $y = 0$ while the right one displays $\rho[(\mathbf{r}_0, z), t]$ versus z, where $\mathbf{r}_0 = (0, 0)$. For big z_0, in all models $\rho[(\mathbf{r}_0, z), t]$ decays quickly away from the surface, indicating that the charges are strongly concentrated about the surface. For small z_0, however, $\rho[(\mathbf{r}_0, z), t]$ oscillates in the SCM of $p = 0$. This oscillation stems from symmetry breaking effects encoded in G_s and B_s that are absent from other models, and it is associated with the excitation of VPWs. In the SCM of $p = 1$, $P_2(\mathbf{K}, \omega)$ vanishes and $P(\mathbf{K}, \omega) = (1/4\pi)\Omega^2(\mathbf{K}, \omega)/(\bar{\omega}^2 - \Omega^2(\mathbf{K}, \omega))$. As Ω varies only slightly with q when $\omega_p\tau$ is not very large, the resulting $\rho(\mathbf{x}, t)$ is also largely localized on the surface as seen in this figure, closely resembling that of self-sustained SPWs, whose density is $\rho_{\rm SPW}(\mathbf{K}) = {\rm const}/\epsilon(K, \omega)$, though only VPWs are excited in the limit of $p = 1$.

The induced charge density profiles are of experimental interest for two reasons. Firstly, they may be directly measured[116] to discriminate existing models against one another. In particular, the validity of the SCM can be examined. Secondly, the induced charge density is ultimately responsible for the energy losses experienced by the probing particles. Such losses can be measured to benchmark the models. A systematic study of this issue is beyond the scope of the present paper and will be published elsewhere.

VI. SUMMARY

In summary, we have presented a general macroscopic theory of electrodynamic response for bounded systems without the use of ABCs and MBCs. The theory yields analytical expressions of the density-density response function and sheds fresh light into its mathematical structure and the physical origin behind it. It provides a physically transparent way of evaluating the function either analytically or numerically. Such transparency is not affordable in existing calculations. A unified view has been rendered of various dispersive and non-dispersive models, including the DM, the HDM, the SRM and the SCM. Some long-standing misconceptions regarding these models have been clarified.

According to the SCM, an intrinsic instability of the metal is predicted to occur, as may be revealed as a zero of the dynamical structure factor. This instability can be utilized to drastically reduce the energy losses suffered by SPWs that have so far impeded the progress in the field of plasmonics, as suggested in our previous work.

In contrast with conventional wisdom, we find that a grazing exterior charge can excite volume density waves in a SIM provided the charge is in the vicinity of its surface. We also find that the distribution of induced charges is sensitive to the dynamics model in use. The SCM distinguishes itself from other common models by the inclusion of effects due to translation symmetry breaking and surface roughness. A measurement of the charge distribution may be carried out to examine the validity and limitations of these models.

While it is explicitly developed for metals, in which electrical currents are carried primarily by conduction electrons, the general theory as developed in Sec. II can be adapted to situations where the currents may be of a different nature, e.g. due to excitons.

Addressing a fundamental problem in condensed matter physics and surface science, the theory is expected to be useful in a number of areas including chemistry and nuclear instruments design. Applications in particle scattering and light scattering as well as other phenomena such as quantum forces will be explored in the future.

Acknowledgement. The author is grateful to Dr. T. Philbin

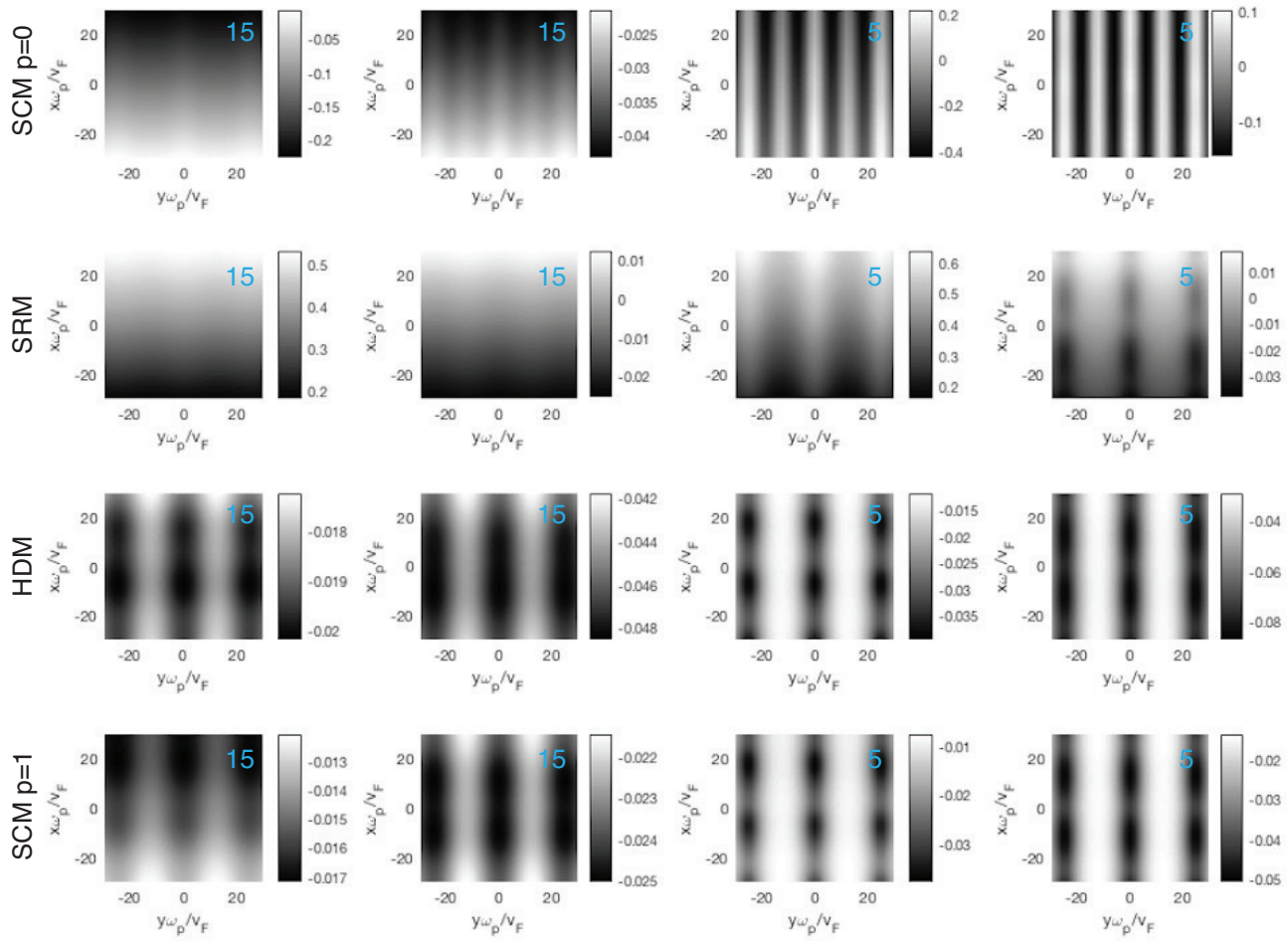

FIG. 8. Planar distribution of the charges induced by a particle of unit charge grazing over the surface at constant speed $V = 10v_F$ and distance z_0, see Fig. 1 (a). The number at the upper right corner of each panel indicates the value of $z_0\omega_p/v_F$. Within each pair of panels of the same z_0 and the same model, the left panel displays $\rho(\mathbf{x}_0, t)$ and the right one displays $\rho_{\|}(\mathbf{r}, t) = \int_0^\infty dz\, \rho(\mathbf{x}, t)$. The particle is located at $(0, 0, -z_0)$ for the moment under consideration. Gray scale indicates their values.

for bringing to his notice the recent work in Refs.[6,7]. He also thanks J. Pendry and M. Apostol for some useful suggestions.

* haiyao.deng@gmail.com

[1] J. M. Pitarke, V. M. Silkin, E. V. Chulkov, and P. M. Echenique,"Theory of surface plasmons and surface-plasmon polaritons." Rep. Prog. Phys. **70**, 1 (2007), and references therein.

[2] B. Chen and D. F. Nelson, "Wave propagation of exciton polaritons by a wave-vector-space method." Phys. Rev. B **48**, 15372 (2007).

[3] S. I. Pekar, "The theory of electromagnetic waves in a crystal in which excitons are produced." Sov. Phys. JETP **6**. 785 (1958).

[4] J. L. Birman and J. J. Sein, "Optics of Polaritons in Bounded Media." Phys. Rev. B **6**, 2482 (1972).

[5] D. N. Pattanayak and E. Wolf, "General form and a new interpretation of the Ewald-Oseen extinction theorem." Opt. Communications **6**, 217 (1972).

[6] R. Schmidt and S. Scheel, "Local density of states near spatially dispersive nanospheres." Phys. Rev. A **93**, 033804 (2016).

[7] R. Schmidt and S. Scheel, "Radiative heat transfer between spatially nonlocally responding dielectric objects." J. Phys. B 51, 044003 (2018).

[8] M. F. Bishop and A. A. Maradudin, "Energy flow in a semi-infinite spatially dispersive absorbing dielectric." Phys. Rev. B **14**, 3384 (1976).

[9] H.-Y. Deng,"Electrostatic responses of anisotropic dielectric films." Eur. J. Phys. **41**, 035203 (2020).

[10] D. Bohm and D. Pines, "A collective description of electron Interactions: III. Coulomb interactions in a degenerate electron gas." Phys. Rev. **92**, 609 (1953).

[11] D. Pines, "A collective description of electron interactions: IV. Electron interaction in metals." Phys. Rev. **92**, 625 (1953).

[12] J. Hubbard, "The dielectric theory of electronic interactions in solids." Proc. Phy. Soc. A **68**, 976 (1955).

[13] A. G. Eguiluz, A. Fleszar and J. A. Gaspar, "On the ab initio eval-

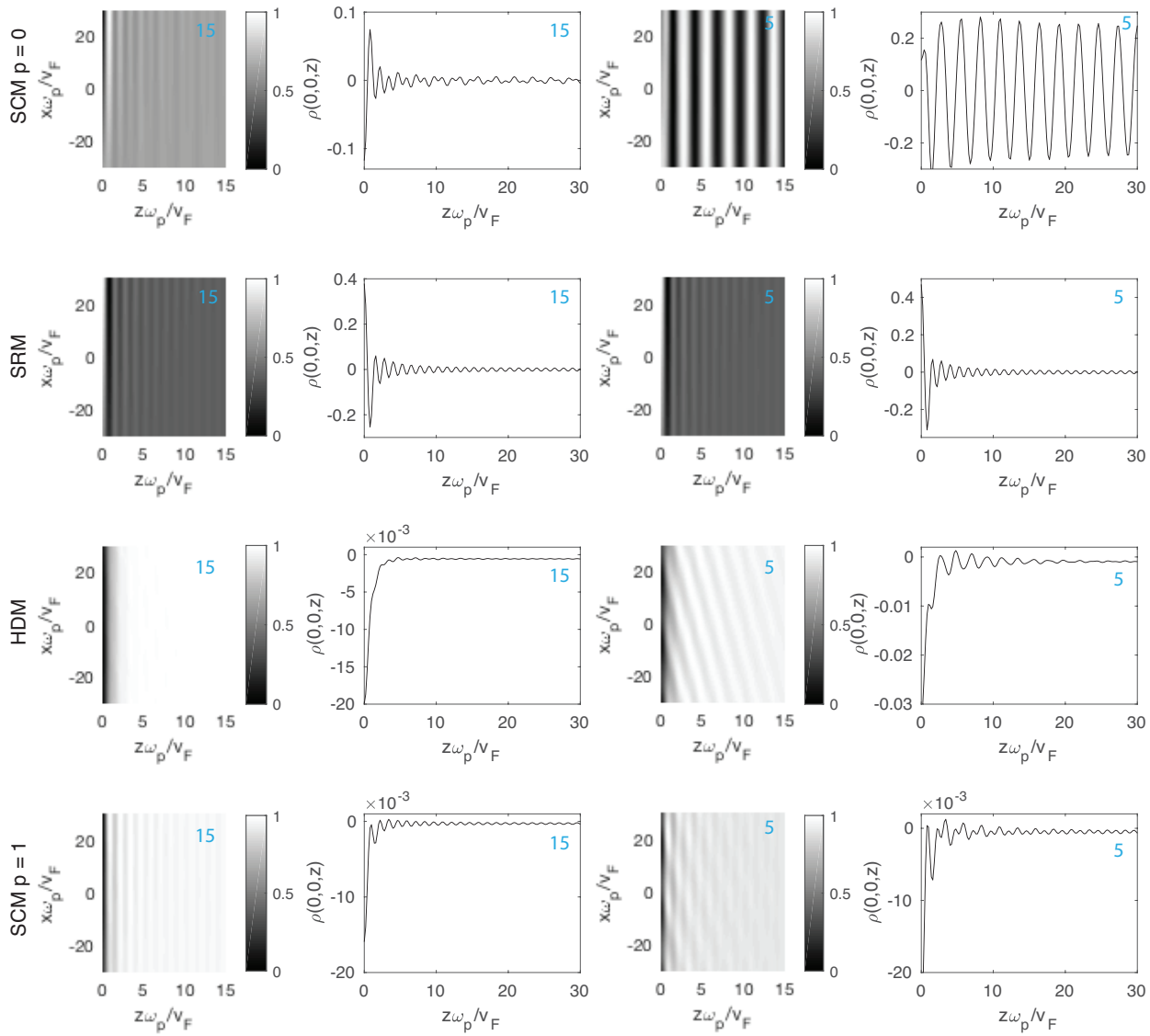

FIG. 9. Depth distribution of the charges induced by the grazing particle. The number at the upper right corner of each panel indicates the value of $z_0\omega_p/v_F$. Within each pair of panels of the same z_0 and the same model, the left panel displays $\rho(\mathbf{x},t)$ in the plane $y = 0$ and the right one displays $\rho[(\mathbf{r}_0,z),t]$, where $\mathbf{r}_0 = (0,0)$. Gray scale indicates their values. The particle is located at $(0,0,-z_0)$ for the moment under consideration.

uation of dynamical electronic response in metals and its comparison with experiment." NIMB **96**, 550 (1995).

[14] J. Kawaji, *Surface science: the first thirty years*, edited by C. B. Duke (North-Holland, New York, 1994)

[15] R. H. Ritchie, "Plasma losses by fast electrons in thin films." Phys. Rev. **106**, 874 (1957).

[16] C. J. Powell and J. B. Swan, "Effect of oxidation on the characteristic loss spectra of aluminum and magnesium." Phys. Rev. **116**, 81 (1959); "Origin of the characteristic electron energy losses in magnesium." ibid 115, 869 (1959).

[17] S. A. Maier, *Plasmonics: fundamentals and applications* (Springer Science & Businnes Media, New York, 2007).

[18] A. V. Zayats, I. S. Igor and A. A. Maradudin, "Nano-optics of surface plasmon polaritons." Phys. Rep. **408**, 131 (2005).

[19] W. L. Barnes, A. Dereux and T. W. Ebbesen, "Surface plasmon subwavelength optics." Nature (London) **424**, 828 (2003).

[20] P. J. Feibelman, "Surface electromagnetic fields." Prog. Surf. Sci. **12**, 287 (1982).

[21] P. Apell, A. Ljungbert and S. Lundgvist, "Non-local optical effects at metal surfaces." Phys. Scripta **30**, 367 (1984).

[22] N. J. M. Horing, E. Kamen and H.-L. Cui, "Inverse dielectric function of a bounded solid-state plasma." Phys. Rev. B **32**, 2184 (1985).

[23] K. -D. Tsuei, E. W. Plummer, A. Liebsch, E. Pehlke, K. Kempa and P. Bakshi, "The normal modes at the surface of simple metals." Surf. Science **247**, 302 (1991).

[24] K. Kempa and W. L. Schaich, "Surface-plasmon dispersion in the infinite-barrier model." Phys. Rev. B **32**, 8375 (1985).

[25] C. A. Ullrich and G. Vignale, "Theory of the linewidth of intersubband plasmons in quantum wells." Phys. Rev. Lett. **87**,

037402 (2001).
[26] G. Toscano, J. Straubel, A. Kwiatkowski, C. Rockstuhl, F. Evers, H. Xu, N. A. Mortensen and M. Wubs, "Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics." Nat. Communications **6**, 7132 (2015)
[27] W. Yan, "Hydrodynamic theory for quantum plasmonics: linear-response dynamics of the inhomogeneous electron gas." Phys. Rev. B **91**, 115416 (2015).
[28] C. Cirací, "Quantum hydrodynamic theory for plasmonics: Impact of the electron density tail." Phys. Rev. B **93**, 205405 (2016).
[29] T. Christensen, W. Yan, A.-P. Jauho, M. Solijacic and N. A. Mortensen, "Quantum corrections in nanoplasmonics: shape, scale and material." Phys. Rev. Lett. **118**, 157402 (2017).
[30] A. Liebsch, "Surface plasmon dispersion of Ag." Phys. Rev. Lett. **71**, 145 (1993).
[31] A. Garcia-Lekue and J. M. Pitarke, "Energy loss of charged particles interacting with simple metal surfaces." Phys. Rev. B **64**, 035423 (2001).
[32] M. Alducin, V. M. Silkin, J. I. Juaristi and E. V. Chulkov, "Energy loss of ions at metal surfaces: band-structure effects." Phys. Rev. A **67**, 032903 (2003).
[33] M. G. Vergniory, V. M. Silkin, I. G. Gurtubay and J. M. Pitarke, "Energy loss of charged particles moving parallel to a magnesium surface: Ab initio calculations." Phys. Rev. B **78**, 155428 (2008).
[34] E. A. Stern and R. A. Ferrell, "Surface plasma oscillations of a degenerate electron gas." Phys. Rev. **120**, 130 (1960).
[35] H. Kanazawa, "On the plasma oscillations in metal foils." Prog. Theor. Phys. **26**, 851 (1961).
[36] R. H. Ritchie and H. B. Eldridge, "Optical emission from irradiated foils. I." Phys. Rev. **126**, 1935 (1962).
[37] W. Ekart, "Simple quantum field theory of the retarded image potential of a slowly moving charge." Solid State Communications **40**, 273 (1981).
[38] M. Babiker, "The Jellium theory of metallic surfaces." Solid State Communications **44**, 1049 (1982).
[39] A. A. Almulhem and M. D. Girardeau, "Theory of ion neutralization at metal surfaces by surface plasmon excitation." Surface Science **210**, 138 (1989).
[40] P. M. Echenique and J. B. Pendry, "Absorption profile at surfaces." J. Phys. C **8**, 2936 (1975).
[41] R. H. Ritchie, "On surface plasma oscillations in metal foils." Prog. Theoret. Phys. (Kyoto) **29**, 607 (1963).
[42] A. J. Bennett, "Influence of the electron charge distribution on surface-plasmon dispersion." Phys. Rev. B **1**, 203 (1970).
[43] J. Harris, "Surface plasmon dispersion: A comparison study of microscopic and hydrodynamics theories." Phys. Rev. B **4**, 1022 (1971).
[44] G. Barton, "Some surface effects in the hydrodynamic model of metals." Rep. Prog. Phys. **42**, 65 (1979).
[45] Y. O. Nakamura, "Quantization of Non-Radiative Surface Plasma Oscillations." Prog. Theoret. Phys. **70**, 908 (1983).
[46] C. Cirací, J. B. Pendry and D. R. Smith, "Hydrodynamic model for plasmonics: a macroscopic approach to a microscopic problem." ChemPhysChem 14, 1109 (2013).
[47] Y. Luo, R. Zhao, and J. B. Pendry, "van der Waals interactions at the nanoscale: The effects of nonlocality." Proc. Natl. Acad. Sci. U.S.A. **111**, 18422 (2014); Y. Luo, A. I. Fernandez-Dominguez, A. Wiener, S. A. Maier, and J. B. Pendry, "Surface plasmons and nonlocality: a simple model." Phys. Rev. Lett. **111**, 093901 (2013).
[48] R. H. Ritchie and A. L. Marusak, "The surface plasmon dispersion relation for an electron gas." Surface Science **4**, 234 (1966).
[49] F. Flores and F. García-Moliner, "Self-energy of a fast-moving charge near a surface." J. Phys. C **12**, 907 (1979).
[50] J. L. Gervasoni and N. R. Arista, "Energy loss and plasmon excitation during electron emission in the proximity of a solid surface." Surf. Sci. **260**, 329 (1992).
[51] F. Yubero, J. M. Sanz, B. Ramskov and S. Tougaard, "Model for quantitative analysis of reflection-electron-energy-loss spectra: Angular dependence." Phys. Rev. B **53**, 9719(1996).
[52] R. Nunez, P. M. Echenique and R. H. Ritchie, "The energy loss of energetic ions moving near a solid surface." J. Phys. C **13**, 4229 (1980).
[53] D. L. Mills, "The scattering fo low energy electrons by electric field fluctuations near crystal surfaces." Surface Science **48**, 59 (1975); H. Ibach and D. L. Mills, *Electron energy loss spectroscopy and surface vibrations* (Academic, New York, 1982).
[54] M. Rocca, "Low-energy EELS investigation of surface electronic excitations on metals." Surf. Sci. Rep. **22**, 1 (1995).
[55] J. I. Juaristi, "Energy loss of ions interacting with metal surfaces." NIMB **230**, 148 (2005).
[56] P. M. Echenique, R. H. Ritchie, N. Barberan and J. Inkson, "Semi-classical image potential at a solid surface." Phys. Rev. B **23**, 6486 (1981).
[57] K. L. Aminov and J. B. Pedersen, "Quantum theory of high-energy electron transport in the surface region." Phys. Rev. B **63**, 125412 (2001).
[58] D. Chan and P. Richmond, "A general theory of free energies of inhomogeneous spatially dispersive media: I." J. Phys. C **8**, 2509 (1975).
[59] J. B. Pendry, "Shearing the vacuum - quantum friction." J. Phys.: Condens. Matter **9**, 10301 (1997).
[60] J. B. Pendry, "Quantum friction - fact or fiction?" New. J. Phys. **12**, 033028 (2010).
[61] A. I. Volokitin and B. N. J. Persson, "Quantum friction." Phys. Rev. Lett. **106**, 094502 (2011).
[62] V. Despoja, P. M. Echenique and M. Sunjic, "Nonlocal microscopic theory of quantum friction between parallel metallic slabs." Phys. Rev. B **83**, 205424 (2011).
[63] R. C. Monreal, "Auger neutralization and ionization processes for charge exchange between slow noble gas atoms and solid surfaces." Prog. Surf. Sci. **89**, 80 (2014), and references therein.
[64] C. Cherqui, *The effect of dynamical image forces on the transport properties of charge carriers and excitons in metal-semiconductor nanostructures* (PhD thesis, University of New Mexico, `http://digitalrepository.unm.edu/phys_etds`, 2014)
[65] A. Vagov, I. A. Larkin, M. D. Croitoru and V. M. Axt, "Role of nonlocality and Landau damping in the dynamics of a quantum dot coupled to surface plasmons." Phys. Rev. B **93**, 195414 (2016).
[66] C. A. Downing and G. Weick, "Topological collective plasmons in bipartite chains of metallic nanoparticle." Phys. Rev. B **95**, 125426 (2017).
[67] J. H. Strait, G. Holland, W. Zhu, C. Zhang, B. R. Ilic, A. Agrawal, D. Pacifili and H. J. Lezec, "Revisiting the photon-drag effect in metal films." Phys. Rev. Lett. **123**, 053903 (2019).
[68] A. Eugiluz, "Density response function and the dynamic structure factor of thin metal films: nonlocal effects." Phys. Rev. B **19**, 1689 (1979).
[69] R. R. Gerhardts, "Surface electromagnetic fields in the semiclassical infinite model: collective and single-particle excitations." Phys. Scripta **28**, 235 (1983).
[70] F. Garcia-Moliner and F. Flores, *Introduction to the theory of solid surfaces* (Cambridge University Press, New York, 1979).
[71] A. A. Rukhadze and V. P. Silin, "Electrodynamics of media with spatial dispersion." Sov. Phys. Usp. **4**, 459 (1961).
[72] A. A. Maradudin and D. L. Mills, "Effect of spatial dispersion on

the properties of a semi-infinite dielectric." Phys. Rev. B **7**, 2787 (1973).
[73] F. Garcia-Moliner and F. Flores, "Classical Electrodynamics of Non-Specular Dielectric Surfaces." Le J. de Physique **38**, 851 (1977).
[74] F. Flores and F. Garcia-Moliner, "Classical electrodynamics of non-specular conducting surfaces." Le J. de Physique 38, 863 (1977).
[75] R. J. Churchill and T. G. Philbin, "Reflection and transmission in nonlocal susceptibility models with multiple resonances." Phys. Rev. B **95**, 205406 (2017).
[76] V. M. Agranovich and V. L. Ginzburg, *Spatial dispersion in crystal optics and the theory of excitons* (Interscience, London, 1966).
[77] F. Richter, D. Semkat and K. Henneberger, "The photon Green's function for bounded media: splitting property and nonequilibrium radiation laws." J. Phys. Conf. Ser. **220**, 012006 (2010).
[78] F. Forstmann and H. Stenschke, "Dispersion of plasmons at the surface of a metal and at the interface between two metals." Phys. Rev. B **17**, 1489 (1978).
[79] P. Halevi, *Spatial dispersion in solids and plasmas* (North-Holland, Amsterdam, 1992).
[80] B. Horovitz and C. Henkel, "Surface plasmons at composite surfaces with diffusive charges." EPL **97**, 57010 (2012).
[81] M. G. Silveirinha, "Boundary conditions for quadrupolar metamaterials." New. J. Phys. **16**, 083042 (2014).
[82] J. Tignon, T. Hasche, D. S. Chemla, H. C. Schneider, F. Jahnke and S. W. Koch, "Unified picture of polariton propagation in bulk GaAs semiconductors." Phys. Rev. Lett. **84**, 3382 (2000).
[83] K. Hennenberger, "Additional boundary conditions: an historical mistake." Phys. Rev. Lett. **80**, 2889 (1998).
[84] D. F. Nelson and B. Chen, "Comment on 'Additional boundary conditions: a historical mistake'." Phys. Rev. Lett. **83**, 1263 (1999); R. Zeyher, ibid **83**, 1264 (1999). See also Hennenberger's reply.
[85] J. Harris, "The effect of short range correlations on surface plasmon dispersion." J. Phys. C **5**, 1757 (1972).
[86] K. Fuchs, "The conductivity of thin metallic films according to the electron theory of metals." Proc. Camb. Phil. Soc. **34**, 100 (1938).
[87] G. E. H. Reuter and E. H. Sondheimer, "The theory of the anomalous skin effect in metals." Proc. R. Soc. London A **195**, 338 (1948).
[88] K. L. Kliewer and R. Fuchs, "Anomalous skin effect for specular electron scattering and optical experiments at non-normal angles of incidence." Phys. Rev. **172**, 607 (1968).
[89] J. M. Ziman, *Electrons and phonons: The theory of transport phenomena in solids* (Oxford University Press, Oxford, 2001).
[90] A. A. Abrikosov, *Fundamentals of the theory of metals* (Elsevier Science Publishers B. V., North-Holland, Amsterdam, 1988).
[91] M. I. Kaganov, G. Y. Lyubarskiy and A. G. Mitina, "The theory and history of the anomalous skin effect in normal metals." Phys. Rep. **288**, 291 (1997).
[92] H.-Y. Deng, "A unified macroscopic theory of surface plasma waves and their losses." New J. Phys. **21**, 043055 (2019).
[93] See the online supplemental information to this paper.
[94] B. N. J. Persson, "Theory of inelastic scattering of slow electrons by molecules adsorbed on metal surfaces." Solid State Communications **24**, 573 (1977).
[95] J. Khurgin, W.-Y. Tsai, D. P. Tsai and G. Sun, "Landau damping and limit to field confinement and enhancement." ACS Phonics **4**, 2871 (2017).
[96] H.-Y. Deng, K. Wakabayashi and C.-H. Lam, "Universal self-amplification channel for surface plasma waves." Phys. Rev. B **95**, 045428 (2017).
[97] H.-Y. Deng, "Theory of nonretarded ballistic surface plasma waves in metal films." Phys. Rev. B **95**, 125442 (2017).
[98] H.-Y. Deng, "Possible instability of the Fermi sea against surface plasma oscillations." J. Phys. Condens. Matter **29**, 455002 (2017).
[99] A. Principi, E. van Loon, M. Polini and M. I. Katsnelson, "Confining graphene plasmons to the ultimate limit." Phys. Rev. B **98**, 035427 (2018).
[100] V. U. Nazarov, J. M. Pitarke, Y. Takada, G. Vignale, and Y.-C. Chang, "Including nonlocality in the exchange-correlation kernel from time-dependent current density functional theory: Application to the stopping power of electron liquids." Phys. Rev. B **76**, 205103 (2007).
[101] V. U. Nazarov, V. M. Silkin and E. E. Krasovskii, "Role of the kinematics of probing electrons in electron energy-loss spectroscopy of solid surfaces." Phys. Rev. B **93**, 035403 (2016).
[102] H.-Y. Deng, "On the electrical conductivity of metals with a rough surface." arXiv: 2001.08639 (2020).
[103] If written in terms of the electric polarization, it reproduces in the local model the basic relation Eq. (3) of this paper: B. Chen and D. F. Nelson, "Wave-vector-space method for wave propagation in bounded media." Phys. Rev. B **48**, 15365 (1993).
[104] A. L. Fetter, "Edge magnetoplasmons in a two-dimensional electron fluid confined to a half-space." Phys. Rev. B **33**, 3717 (1986).
[105] M. Apostol and G. Vaman, "Reflected and refracted electromagnetic fields in a semi-infinite body." Solid State Communications **149**, 1936 (2009); "Electromagnetic field interacting with a semi-infinite plasma." J. Opt. Soc. Am. A **26**, 1747 (2009).
[106] It should be noted that the dielectric function given in Eq. (59) differs from what is usually quoted in the literature (see e.g. Ref.[44]). In the literature it reads $1 - \omega_p^2/(\bar{\omega}^2 - v_0^2K^2)$, which can be derived for an infinite system using Eqs. (56) and (57). These two expressions agree only for small K. The discrepancy arises due to different rationals leading to the HDM. In the fluid-mechanics approach, i.e. Eqs. (56) and (57), non-local effects come through an electronic pressure, i.e. the term with $\nabla\rho$ in Eq. (56), which stems solely from electron-electron interaction. An external electric field contributes to the current via the Drude conductivity not the pressure term. In this sense, the fluid-mechanics HDM is actually a local model. Another approach leading to the HDM, which prescribes a value for v_0 (see Sec. IV), neglects electron-electron interactions. Let us exemplify this approach with an infinite system, for which one can write $\mathbf{J}(\mathbf{K},\omega) = \sigma(\mathbf{K},\omega)\mathbf{E}(\mathbf{K},\omega)$. For small Kv_F/ω, one may expand $\sigma(\mathbf{K},\omega) = \sigma_0(\omega) + (K^2v_F^2/\omega^2)\sigma_1(\omega)$ + higher order terms. The DM is revisited if only the leading term is taken into account. With the quadratic term also considered, one arrives at the HDM. In this approach, an external electric field contributes non-locally to the current in the same way as the induced field. This way, one can show that the dielectric function is given by Eq. (59).
[107] L. D. Landau and E. M. Lifshitz, *Electrodynamics of continuous media* (Pergamon Press, London, 1960).
[108] J. B. Khurgin, "How to deal with the loss in plasmonics and metamaterials" Nat. Photonics **10**, 2 (2015).
[109] M. Rocca, F. Biggio and U. Valbusa, "Surface-plasmon spectrum of Ag(001) measured by high-resolution angle-resolved electron-energy-loss spectroscopy." Phys. Rev. B **42**, 2835 (1990).
[110] M. Saarela, "Density-wave instability in the linear response functions of liquid He and the charged boson gas." Phys. Rev. B **30**, R2925 (1984).
[111] H.-Y. Deng and K. Wakabayashi, "Retardation effects on plasma waves in graphene, topological insulators and quantum wires." Phys. Rev. B **92**, 045434 (2015).

[112] J. M. Purswani and D. Gall, "Electron scattering at single crystal Cu surfaces." Thin Solid Films **516**, 465 (2007).

[113] J. M. Purswani and D. Gall, "Specular electron scattering at single-crystal Cu(001) surfaces." Appl. Phys. Lett. **94**, 252101 (2009).

[114] P. J. Feibelman, "Insensitivity of the infinite-wavelength surface plasmon frequency to the electron density profile." Phys. Rev. B **3**, 220 (1971).

[115] K.-H. Lee and K. J. Chang, "First-principles study of the optical properties and the dielectric response of Al." Phys. Rev. B **49**, 2362 (1994).

[116] F. Huang, V. A. Tamma, Z. Mardy, J. Burdett and H. K. Wickramasinghe, "Imaging Nanoscale Electromagnetic Near-Field Distributions Using Optical Forces." Sci. Rep. **5**, 10610 (2015).

Supplemental Information for "A theory of electrodynamic response for bounded metals: surface capacitive effects"

Hai-Yao Deng*
School of Physics and Astronomy, Cardiff University, 5 The Parade, Cardiff CF24 3AA, Wales, United Kingdom

In this supplemental text, we discuss several issues that have been ignored in the main text: contribution from valence electrons, conductivity tensor in the semi-classical model (SCM), proof that $\chi(\mathbf{K},\mathbf{K}',\omega)$ is the density-density response function, some numerical aspects, properties of surface plasma waves (SPWs) in the SCM and the logical structure of the specular reflection model (SRM).

RESPONSES DUE TO VALENCE ELECTRONS

In the main text we have treated the electrical responses due to conduction electrons with several models. As for the valence electrons, which are important in semi-conductors, we here briefly outline a phenomenological model that has been used widely[1–3]. The observation is that, valence electrons are usually tightly held to their host atoms and the energy bands are largely dispersionless, and thus their dynamical responses should not be much susceptible to the presence of boundaries. One may capture this response by a bulk conductivity $\sigma_p(z,t;\mathbf{k})$, by which the electrical current density due to the valence electrons may be obtained as

$$\begin{aligned}\mathbf{J}_p(z,t;\mathbf{k}) &= \int dt' \int dz' \sigma_p(z-z',t-t';\mathbf{k})\mathbf{E}(z',t';\mathbf{k}),\\ \mathbf{J}_p(z;\mathbf{k},\omega) &= \int dz' \sigma_p(z-z';\mathbf{k},\omega)\mathbf{E}(z';\mathbf{k},\omega)\end{aligned} \tag{1}$$

in response to an electric field $\mathbf{E}(z,t;\mathbf{k})$. In many cases one may neglect all the non-local effects and take $\sigma_p(z-z';\mathbf{k},\omega) \approx \delta(z-z')\sigma_p(\omega)$, where $\sigma_p(\omega)$ can be computed in the atomic limit and modeled with an oscillator model[4]. It accounts for the virtual transitions of valence electrons to the conduction band and is related to the inter-band dielectric function $\epsilon_p(\omega) = 4\pi i\sigma_p(\omega)/\bar{\omega}$, which can be measured for example by means of ellipsometry. $\epsilon_p(\omega)$ contains a real part and an imaginary part. While the real part acts to shield the conduction electrons, the imaginary part – which is positive for stable systems – leads to inter-band absorption. These effects are known to play a significant role in many sorts of materials at optical frequencies.

The inclusion of $\mathbf{J}_p$ in the calculations directly modifies the expressions of only two quantities: $\Omega(K,\omega)$ and $G(\mathbf{K},\omega)$. To $\Omega(K,\omega)$ the term $-\bar{\omega}^2\epsilon_p$ has to be added. Similarly, one must add $-(k/4\pi)\bar{\omega}^2\epsilon_p$ to $G(\mathbf{K},\omega)$. These modified quantities are then used wherever they appear in the theory.

CONDUCTIVITY TENSOR IN THE SCM

Although they are not used in the actual calculations performed in this paper, we show for the sake of completeness how the conductivity tensor $\sigma_{\mu\nu}(z,z';\mathbf{k},\omega)$ in the SCM can be parsed into two components, $\sigma_{b,\mu\nu}(z-z';\mathbf{k},\omega)$ and $\sigma_{s,\mu\nu}(z,z';\mathbf{k},\omega)$, where the former is the same as for an infinite system while the latter due to symmetry breaking effects. To this end, we first write the distribution function as follows

$$g(\mathbf{v},z) = \Theta(v_z)g_>(\mathbf{v},z) + \Theta(-v_z)g_<(\mathbf{v},z), \tag{2}$$

where we have suppressed the explicit dependence on $\mathbf{k}$ and ω for the ease of notation, and

$$g_<(\mathbf{v},z) = ef_0' \int_z^\infty dz' e^{\frac{z'-z}{\lambda}} \frac{\mathbf{v}\cdot\mathbf{E}(z')}{v_z} \tag{3}$$

as well as

$$\begin{aligned}g_>(\mathbf{v},z) &= -ef_0'\\ &\times\left\{\int_0^z dz' e^{\frac{z'-z}{\lambda}} \frac{\mathbf{v}\cdot\mathbf{E}(z')}{v_z} + p\int_0^\infty dz' e^{-\frac{z'+z}{\lambda}} \frac{\mathbf{v}_-\cdot\mathbf{E}(z')}{v_z}\right\}.\end{aligned} \tag{4}$$

Utilizing the identity $\int_0^z = \int_{-\infty}^z - \int_{-\infty}^0$, we can rewrite $g(\mathbf{v},z) = g_b(\mathbf{v},z) + g_s(\mathbf{v},z)$, where

$$\begin{aligned}g_b(\mathbf{v},z) = -ef_0' \int_{-\infty}^\infty dz'\ e^{i\bar{\omega}\left|\frac{z-z'}{v_z}\right|} \frac{\mathbf{v}\cdot\mathbf{E}(z')}{|v_z|} \times\\ \Big[\Theta(z-z')\Theta(v_z) + \Theta(z'-z)\Theta(-v_z)\Big],\end{aligned} \tag{5}$$

which shows exactly the same functional form as for an infinite system, and

$$\begin{aligned}g_s(\mathbf{v},z) &= -ef_0'\Theta(v_z)\\ &\times\int_0^\infty dz' e^{i\frac{\bar{\omega}(z+z')}{v_z}} \frac{\mathbf{v}\cdot\mathbf{E}(-z') + p\mathbf{v}_-\cdot\mathbf{E}(z')}{v_z}.\end{aligned} \tag{6}$$

The above expressions are based on an identity and valid for $\mathbf{E}(z'<0)$ of any mathematical form. To ensure that they agree with those given in the main text, we need to continue the form of $\mathbf{E}(z'\geq 0)$ across the surface well into the vacuum so that no spurious singularity occurs at the surface. With the as-given $g_{b/s}$ it is straightforward to write down the conductivity tensors.

DENSITY-DENSITY RESPONSE FUNCTION

For the sake of completeness, in this appendix we show that the response function $\chi(\mathbf{K},\mathbf{K}',\omega)$ is the density-density response function that is usually calculated by the Greenwood-Kubo formula or *ab initio* methods such as the Time Dependent Density Functional Theory. To this end, we shall work with a quantum mechanical formalism and make use of the correspondence principle.

Let us expose the SIM to a probing electrostatic potential $\varphi(\mathbf{x}, t)$. The Hamiltonian of the system is written as $H_\rho + H'(t)$, where H_ρ describes the SIM in isolation and

$$H'(t) = \int d^3\mathbf{x}\, \varphi(\mathbf{x}, t)\, \hat{\rho}(\mathbf{x}) = \frac{2}{\pi} \sum_{\mathbf{k}} \int_0^\infty dq\, \varphi(\mathbf{K}, t)\, \hat{\rho}^\dagger(\mathbf{K}). \tag{7}$$

contains the effects of the probing field. Here

$$\varphi(\mathbf{K}, t) = \int_0^\infty dz\ \cos(qz)\, \varphi(z, t; \mathbf{k})$$

together with

$$\varphi(z, t; \mathbf{k}) = \int d^2\mathbf{r}\, \varphi(\mathbf{x}, t) e^{-i\mathbf{k}\cdot\mathbf{r}} / \sqrt{A},$$

and similarly for $\hat{\rho}(\mathbf{K})$ – the charge density operator.

Now we specify that $\varphi(z, t; \mathbf{k}) = \varphi(\mathbf{K}', t)\cos(q'z)$, or equivalently

$$\varphi(\mathbf{K}, t) = \frac{\pi}{2}\varphi(\mathbf{K}', t)\delta(q - q'), \tag{8}$$

for which the interaction becomes

$$H'(t) = \sum_{\mathbf{k}} \varphi(\mathbf{K}', t)\hat{\rho}^\dagger(\mathbf{K}'). \tag{9}$$

We suppose that in the remote past the SIM is in the ground state $|0\rangle$ of H_ρ, which we take to be the Fermi sea. At instant t, it is in a state which is denoted by $|t\rangle$, and the charge density then reads $\rho(\mathbf{x}, t) = \langle t|\hat{\rho}(\mathbf{x})|t\rangle$. According to the first-order perturbation theory, we have

$$\rho(\mathbf{x}, t) = \frac{i}{\hbar} \int_{-\infty}^{t} dt' \langle 0|[H'(t'), U^\dagger(t - t')\hat{\rho}(\mathbf{x})U(t - t')]|0\rangle. \tag{10}$$

Here $U(t) = e^{-iH_\rho t/\hbar}$ is the evolution operator and the square bracket takes the commutator between the operators separated by the comma. Taking the Fourier transform and by virtue of the translation symmetry along the surface, we obtain

$$\rho(\mathbf{K}, t) = \int_{-\infty}^{\infty} dt' \chi(\mathbf{K}, \mathbf{K}', t - t')\varphi(\mathbf{K}', t'),$$
$$\rho(\mathbf{K}, \omega) = \chi(\mathbf{K}, \mathbf{K}', \omega)\varphi(\mathbf{K}', \omega), \tag{11}$$

where $\chi(\mathbf{K}, \mathbf{K}', \omega)$ is the Fourier transform of $\chi(\mathbf{K}, \mathbf{K}', t)$ with respect to t and

$$\chi(\mathbf{K}, \mathbf{K}', t) = \Theta(t)\frac{i}{\hbar}\langle 0|[\hat{\rho}^\dagger(\mathbf{K}'), U^\dagger(t)\hat{\rho}(\mathbf{K})U(t)]|0\rangle. \tag{12}$$

Comparing Eq. (11) with Eq. (36) in the main text, we conclude that $\chi(\mathbf{K}, \mathbf{K}', \omega)$ is indeed the density-density response function as claimed.

One can analogously study the responses to an exterior charge and revisit the expression of $P(\mathbf{K}, \omega)$ and the relation Eq. (42) in the main text, which holds valid in the quantum mechanical formalism as well. The calculations are straightforward and not repeated.

NUMERICAL METHODS

Here some details are given about the numerical evaluation of various quantities used in the plots. In evaluating $\Omega(K, \omega)$ by means of Eq. (96) in the main text, we have computed the integral over r by a sum with step $dr = 0.01$ achieving convergence. For integrals over $\mathbf{v}$ with $v_z \geq 0$, we have used spherical coordinates to write $\mathbf{v} = v(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ with $\theta \in [0, \pi/2]$. At zero temperature, the integration over the magnitude v can be directly performed, so that the numerical evaluation boils down to one over the angles, which is approximated by a sum with sampling steps $d\theta = d\varphi = \pi/200$.

For any integral over q running to infinity, we place a cutoff q_c. Namely, $\int_0^\infty dq...$ is replaced by $\int_0^{q_c} dq...$, which is further converted into a sum with step dq. Universal convergence is achieved as soon as q_c gets as large as a few multiples of ω_p/v_F; see Fig. 1 for instance. We have chosen $q_c = 4.5\omega_p/v_F$ in most numerical results presented in this paper. As for dq, we have put $dq = 0.01\omega_p/v_F$ as long as $dq/k \ll 1$. For k comparable to dq, we need to choose a much smaller dq. A more stable evaluation of these integrals is to turn them into this form: $\int dq \frac{k}{k^2+q^2}... = \int \frac{d\tilde{q}}{1+\tilde{q}^2}...$, where $\tilde{q} = q/k$. The factor $1/(1 + \tilde{q}^2)$ effectively puts a cut-off on q, i.e. the contribution from $\tilde{q} > \tilde{q}_c \gg 1$ is automatically suppressed, which ensures the universal convergence with increasing q_c; see Fig. 1 for an illustration. Actually, for very small k, the factor $k/(k^2 + q^2)$ reduces to the Dirac function $\pi\delta(q)$.

Finally, in evaluating the charges induced by a grazing particle, i.e. Eq. (118) in the main text, the integral over $\mathbf{k}$ is restricted to a square $k_{x,y} \in [-k_c, k_c]$, where $k_c = 4\omega_p/v_F$ has been used in numerical plots. This amounts to saying that the grazing particle is not point-like, but instead a spot of size $\sim k_c^{-1}$ in the $x - y$ plane. Namely, $\rho_{\text{probe}}(\mathbf{x}, t) = \delta(z + z_0) \int \frac{d^2\mathbf{k}}{4\pi^2} e^{i\mathbf{k}\cdot\mathbf{r}(t)}$.

SPWS IN THE SCM

Considering the importance of SPWs in the control of light-matter interaction on the scale of nano meters and many other areas of practical significance, we briefly discuss these waves in light of the present theory by the SCM. A systematic investigation has recently been conducted in Refs.[5–8]. Our main purpose here is to extract the frequency and decay rate of SPWs from the profile of $\epsilon_s(k, \omega)$ and comment on their significance. The dependences of the as-extracted quantities on *Fuchs* parameter p and the wavenumber k are displayed in Fig. 2. Within numerical uncertainties they agree well with those obtained in our previous work.

The SPWs are determined by the equation $\epsilon_s(k, \omega) = 0$. In our previous work, we solved this equation directly in the complex ω-plane. Here we consider a faster but less accurate method. Let the solution be written as $w_s(k) = \omega_s(k) - i\gamma(k)$, where ω_s is the SPW frequency and γ the decay rate. If $\left|\epsilon_s(k, \omega)\right|^{-2}$ is plotted versus ω at fixed k, a peak will show up centered about $\omega_s(k)$. The full width at half maximum (FWHM) of this peak is 2γ. This can be shown as follows.

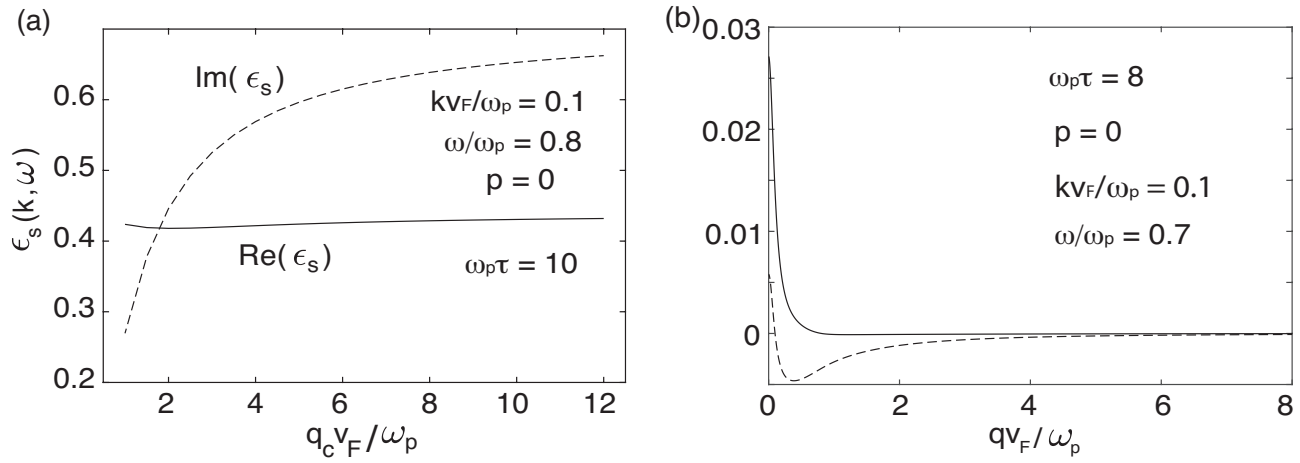

FIG. 1. Convergence upon increasing the cut-off q_c for the integrals over q. (a) Convergence is achieved as long as q_c exceeds a few multiples of ω_p/v_F. (b) The real (solid line) and imaginary (broken line) parts of the integrand in the expression (14), $\frac{1}{1+\tilde{q}^2}\frac{G}{\Omega^2-\bar{\omega}^2}$, at $\omega = 0.7\omega_p$ and $\omega_p\tau = 8$. Inset: the same as the main panel but for a different wavenumber. The integrand quickly decays to zero as long as $\tilde{q} \gg 1$, which provides an effective cut-off and ensures the convergence. In general, universal convergence is achieved when $q_c/k \gg 1$.

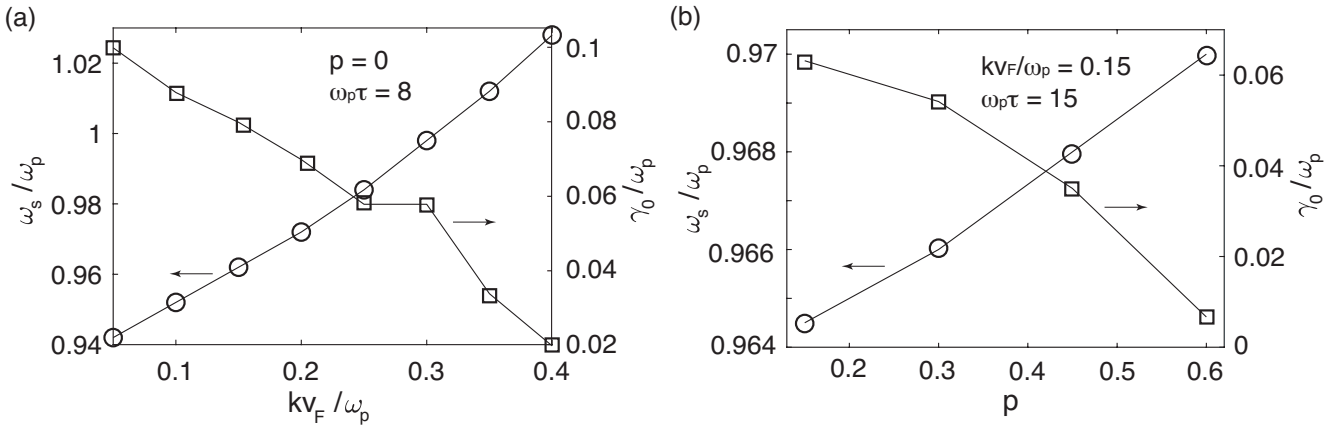

FIG. 2. SPW dispersion $\omega_s(k)$ and decay rate γ in terms of $\gamma_0(k) = \tau^{-1} - \gamma(k)$. The results are calculated according to the SCM with $q_c = 4.5\omega_p/v_F$. Numerical resolution of frequency is $0.002\omega_p$.

Let $\omega = \delta\omega + i\gamma + w_s(k)$. For small $\delta\omega$ and γ, we may then write $\epsilon_s(k,\omega) \approx (\delta\omega + i\gamma)\partial\epsilon_s$, where the derivative is taken at $\omega = w(k)$ in the complex plane. We then obtain

$$\frac{1}{\left|\epsilon_s(k,\omega)\right|^2} \approx \frac{1}{\delta\omega^2+\gamma^2}\frac{1}{\left|\partial\epsilon_s\right|^2}. \tag{13}$$

The maximum occurs at $\delta\omega = 0$, while the half maximum occurs at $\delta\omega = \pm\gamma$. Thus, the FWHM is 2γ. An example is shown in Fig. 3. The method is accurate for sharp peaks.

It is interesting to note that, γ is much smaller than τ^{-1}, or in other words, the SPW lifetime is much longer than τ. This stands in sharp contrast with the conventional wisdom. According to the latter the decay rate should be τ^{-1} plus extra contributions such as from Landau damping – which is in the order of kv_F and significant at short wavelengths – and hence by no means less than τ^{-1}. To directly test our prediction, one must measure separately τ^{-1} and γ. It should be noted that τ^{-1} is not what is usually measured in D.C. transport experiments. It is the relaxation rate at the SPW frequency, of which very few data exist in the literature. Nevertheless, measurements of γ are plenty and provide indirect evidence. For example, the decay rate of SPWs on the surface of a single crystal silver has been measured up to $kv_F/\omega_p \sim 0.1$, where Landau damping rate alone makes up a tenth of ω_p while the observed decay rate stands around a hundredth[9]. This considerable discrepancy can hardly be reconciled with the conventional picture, but is anticipated in our theory.

Further insights into the prolonged lifetime can be gained by looking at the long wavelength limit, for which analytical expressions may be established. We rewrite

$$\epsilon_s(k,\omega) = 1 - \frac{1}{k}\int_0^\infty \frac{d\tilde{q}}{1+\tilde{q}^2}\frac{G(\mathbf{K},\omega)}{\Omega^2(K,\omega)-\bar{\omega}^2}. \tag{14}$$

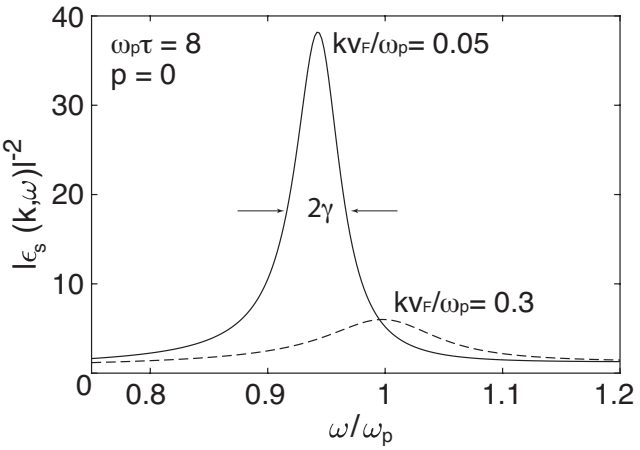

FIG. 3. Extracting the SPW properties from $\left|\epsilon_s(k,\omega)\right|^{-2}$, which features a peak at the SPW frequency ω_s with FWHM equal to 2γ. It is notable that γ is much smaller than τ^{-1}.

For very small k, we make the approximation that $\int_0^\infty d\tilde{q}/(1+\tilde{q}^2)... \to \frac{\pi}{2}\int_{-\infty}^{\infty} dq\ \delta(q)...$ and find

$$\epsilon_s(k,\omega) \approx 1 - \frac{G(\mathbf{K}_0,\omega)}{2k}\frac{\pi}{\omega_p^2-\bar{\omega}^2},$$

where $\mathbf{K}_0 = (\mathbf{k}, 0)$. From this it follows that

$$\gamma_0 \approx -\pi \mathrm{Im}\left[G(\mathbf{K}_0,\omega_s)\right]/4k\omega_s, \tag{15}$$

with ω_s determined by

$$\omega_p^2 - \omega_s^2 - \frac{\pi}{2k}\mathrm{Re}\left[G(\mathbf{K}_0,\omega_s)\right] = 0. \tag{16}$$

Using Eq. (97) in the main text, we find

$$\mathrm{Im}\left[G(\mathbf{K}_0,\omega_s)\right] \approx -\frac{3(1-p)}{2}\frac{k\omega_p^2}{\pi}\frac{kv_F}{\omega_s}, \tag{17}$$

whereby to obtain

$$\gamma_0/\omega_s \approx \frac{3(1-p)}{8}\frac{\omega_p^2}{\omega_s^2}\frac{kv_F}{\omega_s}, \tag{18}$$

which is the same as Eq. (117) in the main text after inserting the expression of ω_s. Existing works have totally missed G_s and failed to reveal the existence of γ_0. Note that in this approximation Landau damping has been automatically suppressed. By means of an energy conservation analysis[8], one can show that Landau damping is always overcompensated by the effects of $\mathcal{H}_s$, which has been ignored in the present study.

It should be cautioned that Eq. (18) is unlikely to be valid for too small k, i.e. $k < \omega_s/c$, where c is the speed of light in vacuum. In such limit, retardation effects shall dominate and the present theory needs to be modified. Generally, $c \sim 100v_F$ in metals, and thus k should not be smaller than $\sim 0.01\omega_s/v_F$ for (18) to make sense.

In closing, we note a size effect that was discussed in Ref.[7]. It was shown that, for metal films of thickness d, γ_0 was strongly reduced when d is in short of the wavelength, i.e. for $kd < 1$, in which case $\gamma_0 \sim kd$. This reduction is ascribed to the factor k/K^2, which strongly suppresses large q contributions, in analogy to the small k limit discussed on the above. For films, the two smallest values of q are zero and π/d, respectively. In the long wavelength limit, virtually only $q = 0$ contributes.

REMARKS ON THE SRM

The usually used SRM assumes that the electrons in the metal are specularly reflected off a surface. One would then expect that the SRM corresponds to the SCM of $p = 1$. However, we have seen in the main text that this is not the case: what the SRM actually does is an extension of the HDM. The reason is simple: the SRM assumes not only (i) a specularly reflecting surface but also (ii) the existence of a fictitious charge sheet located exactly at the surface[10]. It is exactly this charge sheet that produces an artificial extra contribution which closely mimics $P_2(\mathbf{K},\omega)$ in our theory. The purpose here is to discuss this point a little further.

We follow the usual SRM formalism as explained in many papers[10–14]. Let us consider the response to an external distribution of charge $\rho_{\rm ext}(z;\mathbf{k},\omega)$ placed outside the metal. We take $\rho_{\rm ext}(z;\mathbf{k},\omega) = \rho_0(\mathbf{k},\omega)\delta(z-z_0)$ for simplicity, where $z_0 < 0$. In the original SRM, the total potential, $\phi_{\rm tot}(z;\mathbf{k},\omega) = \phi(z;\mathbf{k},\omega) + \phi_{\rm ext}(z;\mathbf{k},\omega)$, where ϕ is the potential produced by the induced charges $\rho(z;\mathbf{k},\omega)$, is written

$$\phi_{\rm tot}(z;\mathbf{k},\omega) = \Theta(z)\phi_m(z;\mathbf{k},\omega) + \Theta(-z)\phi_v(z;\mathbf{k},\omega), \tag{19}$$

where ϕ_m and ϕ_v are the potentials in the so-called pseudo-metal and pseudo-vacuum, respectively. These are further defined by

$$\phi_{m/v}(\mathbf{Q},\omega) = \frac{4\pi}{Q^2\epsilon_{m/v}(\mathbf{Q},\omega)}\left[\rho_{\rm ext}^{m/v}(\mathbf{Q},\omega) + \sigma_s^{m/v}(\mathbf{k},\omega)\right], \tag{20}$$

where $\phi_{m/v}(\mathbf{Q},\omega) = \int_{-\infty}^{\infty} dz\ e^{-iqz}\phi_{m/v}(z;\mathbf{k},\omega)$ is the ordinary Fourier transform, $\mathbf{Q} = (\mathbf{k}, q)$, $\epsilon_m(\mathbf{Q},\omega) = \epsilon(Q,\omega)$, $\epsilon_v(\mathbf{Q},\omega) = 1$ and $\sigma_s^{m/v}(\mathbf{k},\omega)$ is the fictitious surface charge density. In addition, $\rho_{\rm ext}^{m/v}$ is related to $\rho_{\rm ext}$ as follows

$$\rho_{\rm ext}^{m/v}(z;\mathbf{k},\omega) = \Theta(z)\rho_{\rm ext}(\mp z;\mathbf{k},\omega) + \Theta(-z)\rho_{\rm ext}(\pm z;\mathbf{k},\omega). \tag{21}$$

It follows that $\rho_{\rm ext}^{v}(\mathbf{Q},\omega) = 0$ and

$$\rho_{\rm ext}^{m}(\mathbf{Q},\omega) = 2\rho_0(\mathbf{k},\omega)\cos(qz_0). \tag{22}$$

Equations (19) - (21) define the SRM. Requiring the continuity of the dielectric displacement at $z = 0$ leads to

$$\sigma_s^m(\mathbf{k},\omega) = -\sigma_s^v(\mathbf{k},\omega) = \sigma_s(\mathbf{k},\omega). \tag{23}$$

This can be further fixed by requiring the continuity of $\phi_{\rm tot}$ at $z = 0$. One finds

$$\sigma_s(\mathbf{k},\omega) = 2\rho_0(\mathbf{k},\omega)\cosh(kz_0)\epsilon_{\rm s,SRM}^{-1}(\mathbf{k},\omega), \tag{24}$$

where $\epsilon_{\rm s,SRM}$ is given in Sec. 3 in the main text, i.e.

$$\epsilon_{\rm s,SRM}(\mathbf{k},\omega) = 1 + \frac{k}{\pi}\int_{-\infty}^{\infty}\frac{dq}{Q^2\epsilon(Q,\omega)}.$$

The zeros of this quantity determine the SPWs in the SRM. This derivation reveals that the fictitious charge sheet is responsible for the SPWs in the SRM.

One can show that $\phi_m(0;\mathbf{k},\omega) = \phi_v(0;\mathbf{k},\omega)$ and $\partial_z\phi_m(0;\mathbf{k},\omega) = \partial_z\phi_v(0;\mathbf{k},\omega)$, with which we get by Eq. (19) the induced charge density as

$$\rho(z;\mathbf{k},\omega) = \frac{1}{4\pi}\left(k^2 - \partial_z^2\right)\phi_m(z;\mathbf{k},\omega). \tag{25}$$

This can be transformed as

$$\rho(z;\mathbf{k},\omega) = \frac{2}{\pi}\int_0^{\infty} dq\frac{\cos(qz)}{2\epsilon(Q,\omega)}\left[\rho^m_{\rm ext}(\mathbf{Q},\omega) + \sigma_s(\mathbf{k},\omega)\right] \tag{26}$$

from which it follows that

$$\rho(\mathbf{K},\omega) = \frac{\rho_0(\mathbf{k},\omega)}{\epsilon(K,\omega)}\left[\cos(qz_0) + \frac{\cosh(kz_0)}{\epsilon_{\rm s,SRM}(k,\omega)}\right], \tag{27}$$

where we have replaced $\mathbf{Q}$ with $\mathbf{K}$. This expression confirms that the fictitious charge sheet plays a similar role in the SRM as P_2 in our theory. Though this formalism of SRM has been widely used, it should be clear that its two assumptions (i) and (ii) are incompatible. It does not reduce to the DM in the long wavelength limit.

* haiyao.deng@gmail.com

[1] A. Liebsch, "Surface plasmon dispersion of Ag." Phys. Rev. Lett. **71**, 145 (1993).

[2] K. L. Kliewer and R. Fuchs, "Anomalous skin effect for specular electron scattering and optical experiments at non-normal angles of incidence." Phys. Rev. **172**, 607 (1968).

[3] V. U. Nazarov, V. M. Silkin and E. E. Krasovskii, "Role of the kinematics of probing electrons in electron energy-loss spectroscopy of solid surfaces." Phys. Rev. B **93**, 035403 (2016).

[4] C. Cirací, J. B. Pendry and D. R. Smith, "Hydrodynamic model for plasmonics: a macroscopic approach to a microscopic problem." ChemPhysChem 14, 1109 (2013).

[5] H.-Y. Deng, "A unified macroscopic theory of surface plasma waves and their losses." New J. Phys. **21**, 043055 (2019).

[6] H.-Y. Deng, K. Wakabayashi and C.-H. Lam, "Universal self-amplification channel for surface plasma waves." Phys. Rev. B **95**, 045428 (2017).

[7] H.-Y. Deng, "Theory of nonretarded ballistic surface plasma waves in metal films." Phys. Rev. B **95**, 125442 (2017).

[8] H.-Y. Deng, "Possible instability of the Fermi sea against surface plasma oscillations." J. Phys. Condens. Matter **29**, 455002 (2017).

[9] M. Rocca, F. Biggio and U. Valbusa, "Surface-plasmon spectrum of Ag(001) measured by high-resolution angle-resolved electron-energy-loss spectroscopy." Phys. Rev. B **42**, 2835 (1990).

[10] R. H. Ritchie and A. L. Marusak, "The surface plasmon dispersion relation for an electron gas." Surface Science **4**, 234 (1966).

[11] F. Flores and F. García-Moliner, "Self-energy of a fast-moving charge near a surface." J. Phys. C **12**, 907 (1979).

[12] K. L. Aminov and J. B. Pedersen, "Quantum theory of high-energy electron transport in the surface region." Phys. Rev. B **63**, 125412 (2001).

[13] J. L. Gervasoni and N. R. Arista, "Energy loss and plasmon excitation during electron emission in the proximity of a solid surface." Surf. Sci. **260**, 329 (1992).

[14] F. Yubero, J. M. Sanz, B. Ramskov and S. Tougaard, "Model for quantitative analysis of reflection-electron-energy-loss spectra: Angular dependence." Phys. Rev. B **53**, 9719(1996).