This is the accepted manuscript made available via CHORUS. The article has been published as:

Hyperbolic modes of a conductor-dielectric interface

Evgenii E. Narimanov

Phys. Rev. A **99**, 023827 — Published 15 February 2019

DOI: 10.1103/PhysRevA.99.023827

Hyperbolic modes of a conductor-dielectric interface.

Evgenii E. Narimanov
School of Electrical and Computer Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907, USA

Plasmon resonance, with strong coupling of light to electrons at a metal-dielectric interface, allows light confinement and control at subwavelength scale. It's fundamentally limited by the inherent mobility of the free electrons, leading to the corresponding non-locality of the electromagnetic response.[1, 2] We report that this non-locality also results in the formation of a hyperbolic layer near the metal-dielectric interface, with a strong anisotropy of its electromagnetic response. While the resulting "hyperbolic blockade" leads to the suppression of the conventional plasmon resonance, the hyperbolic layer also supports a new class of surface waves, which offer longer propagation distance and stronger field confinement, simultaneously. Furthermore, these "hyper-plasmons" are not limited to the proximity of the plasmon resonance, which extends the operational bandwidth of plasmonic devices.

I. INTRODUCTION

With the ultimate goal of controlling light on a subwavelength scale, the field of nanophotonics generally relies on two main ideas – the plasmon resonance and the use of hyperbolic media. In the former approach, the subwavelength confinement of the electromagnetic field is achieved via the resonant coupling to free charge carriers in a conducting medium,[3] while in the latter it's the result of the extreme anisotropy of the material response that qualitatively changes the nature of the propagating fields.[4] These are generally considered as fundamentally distinct concepts, with their inherent advantages and drawbacks: e.g. plasmonic systems that rely on the properties of a single metal-dielectric interface are simpler to fabricate, but generally limited to the proximity of the corresponding resonance frequency,[3] while the approach based on hyperbolic media offers a broad bandwidth at the expense of highly nontrivial fabrication when the required anisotropy is due to the nanostructuring of the material.[5] However, it is now well understood that the fundamental limits on the light confinement in both cases are defined by the inherent non-locality of the electromagnetic response in the constituent materials, due to e.g. the mobility of the free carriers in conducting materials.[1, 2] In this work, we demonstrate that electromagnetic non-locality leads to an even deeper connection between these two seemingly different concepts of plasmon resonance and hyperbolic media: the inherent mobility of the free charge carriers in a plasmonic material leads to a strong dielectric anisotropy near the metal-dielectric interface, where the corresponding electromagnetic response becomes effectively hyperbolic.

The resulting hyperbolic layer near the metal-dielectric interface supports a new type of surface waves that, compared to the conventional surface plasmons, offer both longer propagation distance and stronger field localization, *at the same time*. This behavior is not limited to the proximity of the plasmon resonance, but – in agreement with the generally broad bandwidth response in hyperbolic media [4] – persists well above the corresponding resonance frequency. Not only does this leads to a dramatic change in the resulting photonic density of states (by several order of magnitude) and consequently in all associated phenomena – from quantum electrodynamics to nonlinear optics to near-field thermal transport, but – by virtue of freeing plasmonics from the proximity of the corresponding plasmon resonance frequency – opens the field to a large class of materials that have not yet been considered in the context of plasmonics.

The existence of the "hyper-plasmon" surface waves with simultaneously long propagation distance and high wavenumbers finally puts to rest the fears of an "ultimate limit field confinement by surface plasmon-polaritons." [6] This limit, while certainly accurate within the framework of an isotropic electromagnetic response theory, appears to be a general artifact of all models that do not properly account for the free electron mobility and the resulting anisotropy of the electromagnetic response of a conducting material.

The predicted phenomena also have broad implications outside the fields of nanophotonics and light-matter interactions in condensed matter physics. In particular, new surface waves predicted in the present work, can be excited at the dielectric elements of research balloons and rocket probes when these enter ionosphere, which would strongly affect electromagnetic measurements and RF communications.

II. THE HYPERBOLIC LAYER

In the local approximation, the electromagnetic response of free carriers to a time-dependent electric field depends on the corresponding frequency and the carrier scattering time, and can be defined in terms of the momentum transfer between the field and the free carriers. However, in close proximity to a high quality metal-dielectric interface that can be considered locally flat, the electron surface reflection will reverse normal to the surface component of the momentum, while leaving its tangential projection intact. As a result, while the specular reflection at the interface will not strongly affect the

electromagnetic response in the tangential direction, its component that is normal to the metal surface, will be substantially altered – leading to a strong anisotropy in this interfacial layer.

In the presence of surface roughness the free carrier reflection is no longer specular,[7] however the effect of the surface scattering on the momentum transfer from the free carriers to the interface (and thus to the entire sample as a whole) is still very different in the normal and tangential directions. As a result, the free carrier electromagnetic response near the conductor - dielectric interface retains its strong anisotropy.

For the field parallel to the surface, the response is similar to that in the bulk medium, and the resulting contribution to the effective dipole moment and the corresponding polarization of the medium, is opposite to the field, just as in the bulk of the material. However, when the field is driving the electron towards the surface, the resulting reflection from the interface reverses the sign of the normal to the interface component of its velocity – and the momentum initially given to the electron by the field, at the reflection is transferred to the crystal as the whole. As a result, compared to the bulk of the material, the electron response in the normal-to-the-interface direction is strongly suppressed. Without the negative contribution of the free electrons, the real part of the permittivity in the normal to the interface direction is now effectively positive – and the thin layer near the surface behaves as if it had negative permittivity parallel to the interface and positive permittivity normal to the interface. A high-quality metal-dielectric surface therefore supports a *hyperbolic layer*.

The formation of the hyperbolic layer relies on high quality of the interface that supports it. While the hyperbolic layer will adiabatically follow a smooth variation of the surface geometry, short-range surface roughness amplitude h that exceeds the characteristic scale of v_F/ω, where v_F is the Fermi velocity of the electrons in the metal, will suppress it. At optical frequencies, this length scale can be on the order of a few nanometers or below, and the formation of the hyperbolic layers that's predicted in the present work, is only expected in high-quality samples with sub-nanometer surface roughness. At lower frequencies however this surface quality requirement is proportionally relaxed – e.g. for mid-IR "designer metals" [8, 9] one needs $h \lesssim 10$ nm.

Note that the conventional hydrodynamical models recently used to account for the free carrier non-locality, generally rely on the material parameters (such as e.g. the phenomenological parameter β in Refs. [1, 2, 10–13]) that are taken from the *bulk* electromagnetic response of the conduction electrons. However, this approximation does not allow to describe the inherent anisotropy of the electromagnetic response in the hyperbolic layer near the metal-dielectric interface. As a result, even though the hydrodynamic models can accurately describe the bulk longitudinal waves with the dispersion band that shows positive slope above the plasma frequency, [12] they do

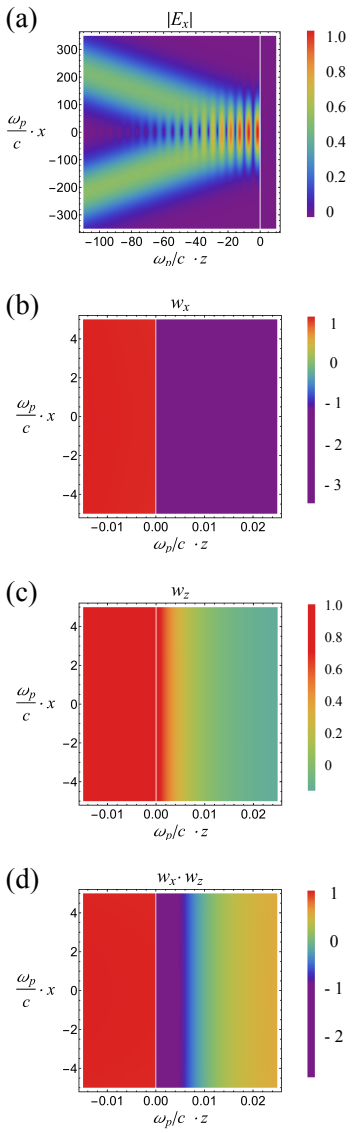

FIG. 1. (Color online) The field and the energy density of a Gaussian beam incident on a metal-dielectric interface, calculated with a full account of the free electron mobility and the resulting non-locality of the electromagnetic response. Panel (a) shows the magnitude of the tangential component of the electric field. Panels (b)-(d) show the corresponding (time-averaged) densities of the tangential electric field w_x (panel (b)) and of the normal to the interface electric field w_z (panel (c)), and the product $w_x w_z$ (panel (d)). The vertical white line indicates the interface $z = 0$. Note clearly visible dielectric region $z < 0$ ($w_x > 0$, $w_z > 0$), metallic region $z \gtrsim 0.01c/\omega_p$ ($w_x < 0$, $w_z < 0$), and the hyperbolic layer $0 < z \lesssim 0.01c/\omega_p$ ($w_x < 0$, $w_z > 0$). The frequency of the incident beam $\omega = 0.5\omega_p$, the electron scattering time $\tau = 18.84/\omega_p$, the crystal lattice permittivity of the conductor $\epsilon_\infty = 12.15$, the permittivity of the dielectric $\epsilon_d = 10.23$, and the Fermi velocity $v_F = 0.00935\ c\ \simeq 2.8 \cdot 10^6$ m/sec; for the plasma wavelength $\lambda_p \equiv 2\pi c/\omega_p = 10\ \mu$m these parameters correspond to the AlInAs/InGaAs material system of Ref. [8]. Note that in this case the electron de Broglie wavelength $\lambda\!\!\!^{-} \simeq 1$ nm, well below the thickness of the hyperbolic layer (~ 20 nm).

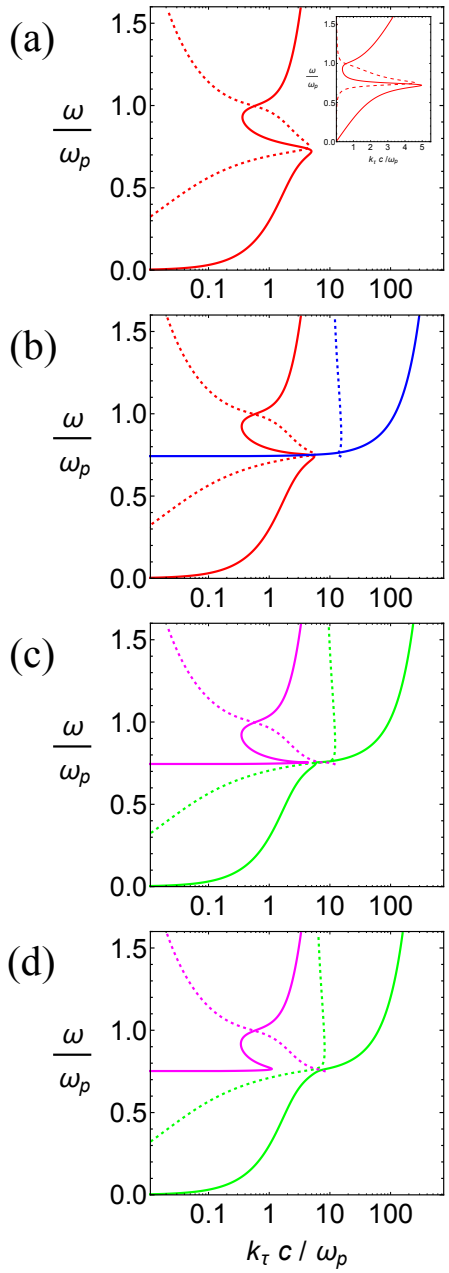

FIG. 2. (Color online) The dispersion of the surface waves at the metal - dielectric interface, with the in-plane momentum k_τ (in units of ω_p/c) and the frequency ω (in units of ω_p). Panel (a) corresponds to the standard result for the Drude metal, with the permittivity $\epsilon_m = \epsilon_\infty \left(1 - \omega_p^2/\left(\omega\left(\omega + i/\tau\right)\right)\right)$, in logarithmic scale (main panel) and linear coordinates (the inset). Panels (b) - (c) show the results for the exact solution, with the ratio of the Fermi velocity to the speed of light in vacuum, $v_F/c = 0.005$ (b), 0.0063 (c), and 0.00935 (d). The material parameters (ϵ_∞, τ, ϵ_d) are the same as in Fig. 1. The red (dark gray) line corresponds to the conventional plasmon, blue (black) – to the hyperbolic mode, green (light gray) – to the hybrid hyper-plasmon, and magenta (gray) curve – to the suppressed resonant plasmon. The solid and dotted lines respectively represent the real and imaginary parts of the in-plane momentum. With the plasma wavelength $\lambda_p = 10$ μm, the doped semiconductor system AlInAs/InGaAs corresponds to the panel (d). Note that in all cases, the wavenumber k_τ is below the Landau damping limit ω/v_F.[42]

not account for the hyper-plasmons that originate from the inherent anisotropy of the hyperbolic layer near the conductor-dielectric interface.

Within the general framework of the hydrodynamic approach, the effect of the spatial nonlocality can also be represented by replacing the nonlocal conductor with a composite material, comprising a thin *isotropic* dielectric layer on top of a local conductor. [13] Although this approach also does not capture the hyper-plasmonic surface waves, this issue can be addressed by the replacement of the isotropic dielectric in the layer by an effective hyperbolic medium. The permittivity components and the thickness of this effective hyperbolic layer can be obtained from the theory introduced in the present work.

On the other hand, new surface waves introduced here in the context of the semiclassical approach, should also emerge in *ab-initio* time-dependent density functional theory of the optical response, both in the case of atomistic first-principle calculations as well as in the jellium limit.

Since the electromagnetic response of free charge carriers is essentially nonlocal, the definition of hyperbolic vs. dielectric vs. metallic response cannot rely on the tensor of the local dielectric permittivity. In principle, it can be formulated in terms of the electric and displacement fields. However, in the presence of the material loss such definition will be complicated and difficult to interpret, as it involves a nontrivial relation between two complex quantities (or equivalently the amplitudes and the phases of the corresponding fields), rather than a simple sign of the relevant permittivity. Instead, we rely on the (time-averaged) real-valued scalar densities

$$w^E_{x,y} = \frac{E_x D_x + E_y D_y}{8\pi}, \quad w^E_z = \frac{E_z D_z}{8\pi}, \tag{1}$$

which represent the response in parallel and normal to the surface directions, respectively.

By definition, in a dielectric $w^E_{x,y} > 0$ and $w^E_z > 0$, in a metal $w^E_{x,y} < 0$ and $w^E_z < 0$, while in a hyperbolic medium $w^E_{x,y}$ and w^E_z have opposite signs. For a dispersion-free material $\mathbf{D} = \epsilon\mathbf{E}$, this reduces to the conventional definition of the dielectric, metallic and hyperbolic media in terms of $\epsilon_{x,y}$ and ϵ_z. However, even when the electromagnetic response is essentially nonlocal, the densities $w_{x,y}$ and w_z in Eqns. (1) can still be calculated from the actual electromagnetic field, and thus represent a useful measure of the electromagnetic response of a non-local medium.

Note that, together with the magnetic density,

$$w^B = \frac{B^2}{8\pi}, \tag{2}$$

$w^E_{x,y}$ and w^E_z add to the standard expression for the electromagnetic energy density of a dispersion-free, lossless medium,[14]

$$w = \frac{B^2 + \mathbf{E}\cdot\mathbf{D}}{8\pi} \equiv w^B + w^E_{x,y} + w^E_z, \tag{3}$$

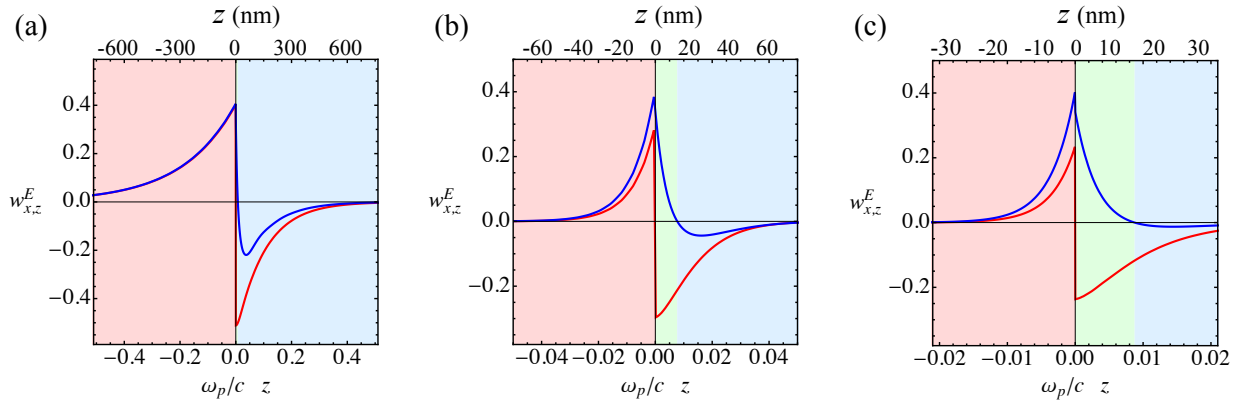

FIG. 3. (Color online) The hyper-plasmonic surface wave profile at the interface of isotropic dielectric with a conducting medium, and its evolution with frequency. The red (gray) and blue (black) lines respectively correspond to the energy densities of the tangential (w_x^E, blue (black) line) and normal to the surface (w_z^E, red (gray) line) components of the electric field, for $\omega = 0.7\omega_p$ (a), $\omega \simeq \omega_p$ (b), $\omega = 1.27\ \omega_p$ (c). The material parameters and the z coordinate in absolute units (at the top) correspond to the AlInAs/InGaAs interface. The light-red (gray) background indicates the dielectric response ($w_x^E > 0$ and $w_z^E > 0$), light-blue (light gray) – to metallic response ($w_x^E < 0$ and $w_z^E < 0$), and light-green (white) – to the hyperbolic layer ($w_x^E < 0$ and $w_x^E > 0$).

However, this connection is only valid in dispersion-free limit, as in the presence of absorption (inherently connected to dispersion via the Kramers-Kronig relations) the electromagnetic energy cannot be defined as a thermodynamic quantity.[14] Even in the limit of infinitesimal absorption, the resulting Brillouin expression for the energy density [14] is not identical to (3), but includes additional terms.[15]

In Fig. 1 we consider a gaussian electromagnetic beam incident onto a half-infinite metal with an atomically flat boundary at $z = 0$, and calculate the actual distribution of the electromagnetic energy density that takes full account of the non-locality of the electron response in the metal. Here, the numerical values for the plasma frequency, electron scattering time etc. correspond to the high-quality interface of doped semiconductor GaInAs with the dielectric AlInAs, the material platform which over the last decade became the system of choice for plasmonics in mid-IR range. [8, 9] While the magnitude of the electric field (see Fig. 1(a)) displays the conventional intensity pattern of the reflected wave, the plots of the local energy density (Fig. 1(b)-(d)) clearly show the presence of the hyperbolic layer at $0 < z \lesssim 0.01c/\omega_p$, where ω_p is the plasma frequency, determined from the metal's bulk response. Note that, the thickness of the hyperbolic layer in this example exceeds the electron de Broglie wavelength by more than an order of magnitude – so that the formation of the hyperbolic layer can be treated within the semiclassical framework.

III. THEORETICAL DESCRIPTION

The actual response to the time-dependent electromagnetic field is defined by the electronic density matrix $\rho_{\mathbf{pp}'}$, governed by the Liouville - von Neumann equation [16] that in the linear response regime reduces to [17, 18]

$$\frac{\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}'} + \hbar\omega}{i\hbar}\,\rho_{\mathbf{pp}'} + \frac{f^{(0)}_{\mathbf{p}} - f^{(0)}_{\mathbf{p}'}}{\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}'}} V^E_{\mathbf{pp}'} = I_{\mathbf{pp}'}\left[\rho\right], \quad (4)$$

where $I_{\mathbf{pp}'}\{\rho\}$ is the collision integral that includes the contributions from both the bulk and the surface scattering of the free carriers, $f^{(0)}_{\mathbf{p}} \equiv f_0(\varepsilon_{\mathbf{p}})$ is the equilibrium (Fermi-Dirac) distribution function, and $V_{\mathbf{pp}'}$ is the matrix element of the spatially dependent amplitude of the electric field $\mathbf{E}(\mathbf{r},t) = \mathbf{E}(\mathbf{r})\exp(-i\omega t)$ that is given by

$$V^E_{\mathbf{pp}'} = \int d\mathbf{r}\ \ \mathbf{j}_{\mathbf{pp}'} \cdot \mathbf{E}(\mathbf{r}), \quad (5)$$

where $\mathbf{j}_{\mathbf{pp}'}$ is the matrix element of the charge carrier current density.[21]

When the relevant "classical" parameters such as the mean-free path $\ell \equiv v_e\tau$ and v_e/ω (where v_e is the typical electron velocity, equal to the Fermi velocity v_F and the thermal velocity v_T for degenerate and non-degenerate statistics respectively) are well above the free carrier de Broglie wavelength $\lambda\!\!\!^{-}$, the Wigner transformation [19, 21] of the density matrix reduces [17, 18, 20] Eqn. (4) to the Boltzmann equation for the charge carrier distribution function $f_{\mathbf{p}}(\mathbf{r})$

$$-i\omega f_{\mathbf{p}}(\mathbf{r}) + \mathbf{v}_{\mathbf{p}} \cdot \nabla f_{\mathbf{p}}(\mathbf{r}) + e\mathbf{E} \cdot \frac{\partial f^{(0)}_{\mathbf{p}}}{\partial \mathbf{p}} = \hat{I}\left[f_{\mathbf{p}}\right], \quad (6)$$

where $\mathbf{v}_{\mathbf{p}} \equiv \partial\varepsilon_{\mathbf{p}}/\partial\mathbf{p}$ is the charge carrier group velocity for the Bloch momentum $\mathbf{p}$, the collision integral $\hat{I}\left[f_{\mathbf{p}}\right]$ includes both the bulk and the surface scattering contributions, and has a highly nontrivial form. However, if the surface roughness h is substantially smaller than v_e/ω and the electron mean free path $\ell = v_e\tau$,[22] the

kinetic equation (6) can be expressed in the conventional form[17, 18, 20]

$$-i\omega f_{\mathbf{p}} + \mathbf{v}_{\mathbf{p}} \cdot \nabla f_{\mathbf{p}} + e\mathbf{E} \cdot \mathbf{v}_{\mathbf{p}} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{p}}} = -\frac{f_{\mathbf{p}} - f_0}{\tau} \quad (7)$$

where the effective relaxation time τ is defined by the bulk scattering, while the effect of the surface is described by the boundary condition on the distribution function at the interface [17, 18] – see Appendix A. For a high-quality interface along one of the symmetry planes of the crystal, the latter reduces to the specular reflection boundary condition at the surface[7, 23–28]

$$f_{\mathbf{p}^-}(\mathbf{r}_s) = f_{\mathbf{p}^+}(\mathbf{r}_s), \quad (8)$$

where $\mathbf{p}^+$ and $\mathbf{p}^-$ are connected by the specular reflection condition, with equal tangential to the surface components $p_\tau^+ = p_\tau^-$, and positive and negative group velocity components in the normal to the interface direction: $\left(\mathbf{v}_{\mathbf{p}^+}\right)_{n_s} > 0$, $\left(\mathbf{v}_{\mathbf{p}^-}\right)_{n_s} < 0$, respectively.[29]

Note that while the standard derivation [10, 11] of the hydrodynamic models [10, 13] for the electromagnetic response of free charge carriers usually follows the application of the Hamilton's principle to the Hohenberg-Kohn ground state Hamiltonian,[30] the hydrodynamic model can also be derived as an approximation for the solution of the kinetic equation (6) based on the method of moments.[31] Such an approximation however neglects the essential anisotropy of the free carrier surface scattering, and the resulting hydrodynamic approach is therefore unable to capture the formation of the hyperbolic surface layer, as well as its implications.

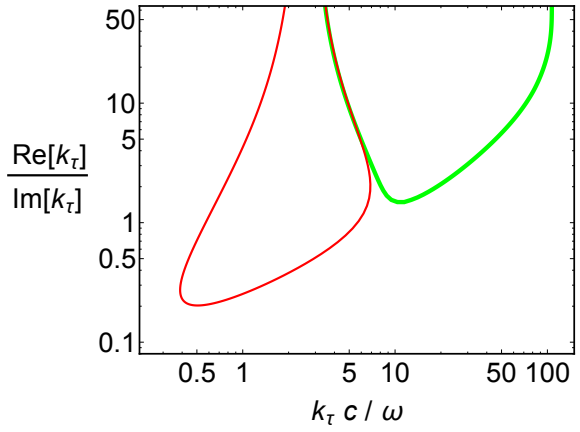

FIG. 4. (Color online) The "figure-of-merit" $\mathrm{Re}\left[k_\tau\right]/\mathrm{Im}\left[k_\tau\right]$ of the surface wave vs. the compression factor $k_\tau/(\omega/c)$, at the AlInAs/InGaAs interface. Red (dark gray) line corresponds to the conventional plasmon and is calculated using the Drude theory (see also Fig. 2(a)), while the green (light gray) curve corresponds to the exact solution for the hyperplasmon (see Fig. 2(e)).

The present semiclassical theory corresponds to a controlled expansion in the small parameter that is equal to the ratio of the electron de Broglie wavelength, $\lambda\!\!\!^{-} \sim \hbar/mv_e$, the characteristic scale of the hyperbolic layer thickness, v_e/ω_p,

$$\frac{\hbar\omega_p}{mv_e^2} \ll 1. \quad (9)$$

When the criterion (9) is satisfied, the quantum phenomena such as finite work function and tunneling leading to the electron "spill-out" – the continuous variation of the electron density from the bulk value to zero), can be safely neglected. However, the density matrix formalism used to derive our semiclassical expressions, also allows to incorporate these effects in the leading-order corrections to the present theory.

Quantitatively, for plasmonic metals we find that the ratio $\hbar\omega_p/mv_e^2$ is barely below unity (~ 0.65 for aluminum, ~ 0.77 for gold, and 0.87 for silver), and the quantum tunneling neglected in the semiclassical theory, leads to the length and energy scales that are similar to those of the hyperbolic layer. As a result, for a quantitative description of the hyper-plasmon waves in metals, quantum corrections to the semiclassical theory may be needed. Note however, that it is also the case of plasmonic metals that leads to the most stringent requirements on the surface roughness – making this regime also the most difficult for experimental study of the new surface waves.

In contrast to this behavior, for transparent conducting oxides such as the indium-tin oxide (ITO) with the plasma frequency in IR range, we find $\hbar\omega_p/mv_F^2 \sim 0.26$, while for doped semiconductors that show the plasmonic behavior, such as InGaAs, $\hbar\omega_p/mv_F^2 \sim 0.15$. Here, we expect the semiclassical theory to yield a fully quantitative description, so that experiments on transparent conducting oxides and semiconductors should show clear manifestations of the predicted behavior.

In the other limit, that of a non-degenerate electron plasma such as the E layer in Earth ionosphere, we find $\hbar\omega_p/mv_T^2 \sim 1.6 \cdot 10^{-5}$. Aside from its practical implications for RF communications and sensing in Earth ionosphere (the problem typically encountered by any space re-entry vehicle), this case should also be considered as a "proof of existence" for the new surface waves, as in this regime no quantum effect should be relevant.

The electromagnetic field at the interface of a dielectric with a conducting medium is defined by the self-consistent solution of the kinetic equation and the surface scattering boundary condition together with the Maxwell equations, where the electron charge and current densities are given by

$$\rho(\mathbf{r}) = 2\int \frac{d\mathbf{p}}{(2\pi\hbar)^3} \cdot \left(f_{\mathbf{p}}(\mathbf{r}) - f_0(\varepsilon_{\mathbf{p}})\right), \quad (10)$$

$$\mathbf{j}(\mathbf{r}) = 2\int \frac{d\mathbf{p}}{(2\pi\hbar)^3} \cdot e\, \mathbf{v}_{\mathbf{p}} f_{\mathbf{p}}(\mathbf{r}), \quad (11)$$

For a high-quality planar surface, [32] the corresponding mathematical problem can be reduced to the system of

two coupled linear integro-differential equations (see Appendix B) that allows an exact analytical solution. For the electric field in the conducting medium we obtain

$$\mathbf{E}_k\left(z>0\right)=\int_{-\infty}^{\infty}\frac{dq}{2\pi}\,\mathbf{e}\left(k,q\right)\,\exp\left(ikx-iqz\right),\quad(12)$$

where

$$\mathbf{e}\left(k,q\right)=\frac{2}{D\left(k,q\right)}\left(\left.\frac{\partial E_x}{\partial z}\right|_{z=+0}-ik\;E_z|_{z=+0}\right)\times\left(\epsilon_{zz}\left(k,q\right)\frac{\omega^2}{c^2}-k^2,0,\nu_{xz}\left(k,q\right)\right),\quad(13)$$

and

$$D\left(k,q\right)=\left(\epsilon_{xx}\left(k,q\right)\frac{\omega^2}{c^2}-q^2\right)\times\left(\epsilon_{zz}\left(k,q\right)\frac{\omega^2}{c^2}-k^2\right)-\nu_{xz}^2\left(k,q\right),\quad(14)$$

with

$$\epsilon_{xx}\left(k,q\right)=\epsilon_\infty-\frac{16\pi ie^2\tau}{\omega}\int_{v_z>0}\frac{d\mathbf{p}}{\left(2\pi\hbar\right)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}}\times v_x^2\,\frac{1-i\omega\tau+ikv_x\tau}{\left(1-i\omega\tau+ikv_x\tau\right)^2+q^2v_z^2\tau^2},\quad(15)$$

and

$$\epsilon_{zz}\left(k,q\right)=\epsilon_\infty-\frac{16\pi ie^2\tau}{\omega}\int_{v_z>0}\frac{d\mathbf{p}}{\left(2\pi\hbar\right)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}}\times v_z^2\,\frac{1-i\omega\tau+ikv_x\tau}{\left(1-i\omega\tau+ikv_x\tau\right)^2+q^2v_z^2\tau^2},\quad(16)$$

and

$$\nu_{xz}\left(k,q\right)=kq-\frac{16\pi e^2\tau^2\omega q}{c^2}\int_{v_z>0}\frac{d\mathbf{p}}{\left(2\pi\hbar\right)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}}\times v_xv_z^2\,\frac{1-i\omega\tau+ikv_x\tau}{\left(1-i\omega\tau+ikv_x\tau\right)^2+q^2v_z^2\tau^2}.\quad(17)$$

Note that the problem of finding the self-consistent solution of Maxwell equations at the planar metal-dielectric interface together with the Boltzmann kinetic equation for free electrons, can be also approached using the surface impedance formalism, [33] pioneered by G. E. H. Reuter and E. H. Sondheimer [28] more than half a century ago. However, while it is in fact possible, [34] an extension of this formalism that would allow to treat the hyper-plasmonic surface waves, was not attempted, and the applications of the surface impedance formalism to surface waves at the metal-dielectric interface remain limited to the regular plasmons. [33]

For a degenerate electron gas [35] we reduce the above expressions to

$$\epsilon_{xx}\left(k,q\right)=\epsilon_\infty-\frac{3\epsilon_\infty}{2}\frac{\omega_p^2}{\omega\left(\omega+i/\tau\right)}\left\{\frac{q^2-2k^2}{\left(k^2+q^2\right)^2}\frac{\left(\omega+i/\tau\right)^2}{v_F^2}+\left(\frac{q^2}{k^2+q^2}+\frac{2k^2-q^2}{\left(k^2+q^2\right)^2}\cdot\frac{\left(\omega+i/\tau\right)^2}{v_F^2}\right)\times\mathcal{F}_0\left(\frac{v_F\sqrt{k^2+q^2}}{\omega+i/\tau}\right)\right\},\quad(18)$$

$$\epsilon_{zz}\left(k,q\right)=\epsilon_\infty-\frac{3\epsilon_\infty}{2}\frac{\omega_p^2}{\omega\left(\omega+i/\tau\right)}\left\{\frac{k^2-2q^2}{\left(k^2+q^2\right)^2}\frac{\left(\omega+i/\tau\right)^2}{v_F^2}+\left(\frac{k^2}{k^2+q^2}+\frac{2q^2-k^2}{\left(k^2+q^2\right)^2}\cdot\frac{\left(\omega+i/\tau\right)^2}{v_F^2}\right)\times\mathcal{F}_0\left(\frac{v_F\sqrt{k^2+q^2}}{\omega+i/\tau}\right)\right\},\quad(19)$$

$$\nu_{xz}\left(k,q\right)=kq\left\{1+\frac{9\epsilon_\infty}{2}\frac{\omega}{\omega+i/\tau}\frac{\omega_p^2}{\left(k^2+q^2\right)c^2}\times\mathcal{F}_1\left(\frac{v_F\sqrt{k^2+q^2}}{\omega+i/\tau}\right)\right\},\quad(20)$$

where ω_p is the standard plasma frequency,

$$\mathcal{F}_0\left(x\right)=\frac{1}{2x}\log\frac{1+x}{1-x},\quad(21)$$

and

$$\mathcal{F}_1\left(x\right)=\frac{1}{x}\left\{\frac{1}{x}+\frac{1}{2}\left(\frac{1}{3}-\frac{1}{x^2}\right)\log\frac{1+x}{1-x}\right\}.\quad(22)$$

The results presented in Fig. 1, were obtained using this solution (see Appendix C).

In the case of a non-degenerate electron gas, we obtain

$$\epsilon_{xx}\left(k,q\right)=\epsilon_\infty-\frac{\epsilon_\infty\,\omega_p^2}{\omega\left(\omega+i/\tau\right)}\left\{\mathcal{F}_2\left(\frac{v_T\sqrt{k^2+q^2}}{\omega+i/\tau}\right)+\frac{k^2}{k^2+q^2}\cdot\mathcal{F}_3\left(\frac{v_T\sqrt{k^2+q^2}}{\omega+i/\tau}\right)\right\},\quad(23)$$

$$\epsilon_{zz}\left(k,q\right)=\epsilon_\infty-\frac{\epsilon_\infty\,\omega_p^2}{\omega\left(\omega+i/\tau\right)}\left\{\mathcal{F}_2\left(\frac{v_T\sqrt{k^2+q^2}}{\omega+i/\tau}\right)+\frac{q^2}{k^2+q^2}\cdot\mathcal{F}_3\left(\frac{v_T\sqrt{k^2+q^2}}{\omega+i/\tau}\right)\right\},\quad(24)$$

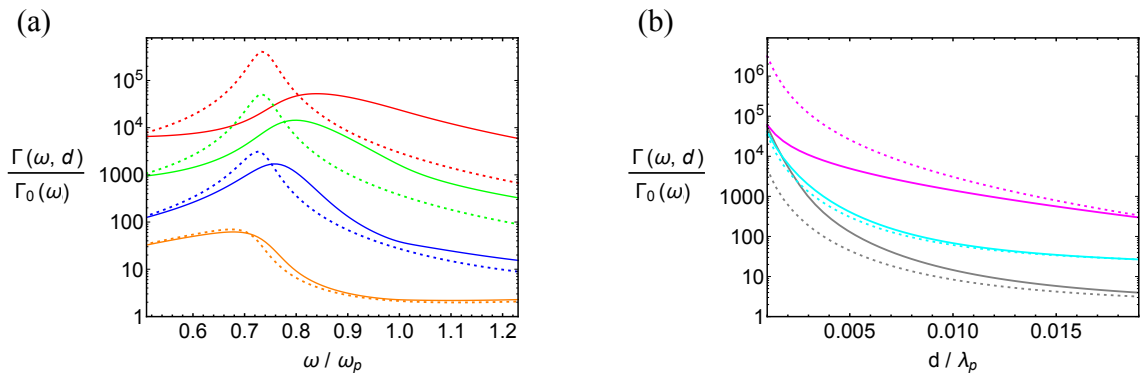

FIG. 5. (Color online) The spontaneous emission rate near the dielectric - doped semiconductor interface, for the AlInAs/InGaAs system, as a function of the frequency (a) and the distance d to the surface (b). The emission rate is normalized to its value in infinite dielectric. Solid lines show the exact solution, while the corresponding dotted lines represent the results of the calculation based on the Drude theory. In panel (a), different colors corresponds to different values of the distance to the interface $\lambda_p/d = 50$ (red (dark gray) curves), 25 (green (light gray)), 10 (blue(black)) and 3 (orange (gray)). In panel (b), different colors correspond to different frequencies, with $\omega = 0.4\omega_p$ (cyan (light gray)), $0.73\omega_p$ (magenta (gray)) and $1.25\omega_p$ (grey). Note that suppression of the plasmon resonance due to the hyperbolic blockade, together with and order of magnitude the enhancement of the spontaneous emission rate above the plasmon resonance frequency seen in panel (a).

and

$$D(k,q) = \epsilon_\infty \left(\frac{\omega}{c}\right)^2 \left\{1 - \frac{\omega_p^2}{\omega(\omega + i/\tau)} \times \mathcal{F}_4\left(\frac{v_T\sqrt{k^2+q^2}}{\omega + i/\tau}\right)\right\} \cdot \left[\epsilon_\infty \left(\frac{\omega}{c}\right)^2 \times \left\{1 - \frac{\omega_p^2}{\omega(\omega + i/\tau)} \mathcal{F}_2\left(\frac{v_T\sqrt{k^2+q^2}}{\omega + i/\tau}\right)\right\} - k^2 - q^2\right], \quad (25)$$

where $v_T \equiv \sqrt{2k_BT/m}$ is the thermal electron velocity, and the functions $\mathcal{F}_2$, $\mathcal{F}_3$ and $\mathcal{F}_4$ are defined as

$$\mathcal{F}_2(x) = \frac{\sqrt{\pi}}{x}\left[\mathrm{Erfi}\left(\frac{1}{x}\right) - i\right]\exp\left(-\frac{1}{x^2}\right), \quad (26)$$

$$\mathcal{F}_3(x) = 2\left(\frac{1}{x^2} - \frac{1}{2}\right)\mathcal{F}_2(x) - \frac{2}{x^2}, \quad (27)$$

$$\mathcal{F}_4(x) = \mathcal{F}_2(x) + \mathcal{F}_3(x). \quad (28)$$

Here, $\mathrm{Erfi}(x)$ is the imaginary error function.

IV. SURFACE WAVES

For a surface state at the metal-dielectric interface, matching the tangential electric field and the normal component of the electric displacement at the interface yields (see Appendix D)

$$\frac{1}{\pi}\int_{-\infty}^{\infty} dq\, \frac{\epsilon_{zz}(k,q)\,\omega^2/c^2 - k^2}{D(k,q)} = -\frac{c^2}{\omega^2}\frac{\kappa_d}{\epsilon_d}, \quad (29)$$

where ϵ_d is the permittivity of the dielectric medium, and

$$\kappa_d = \sqrt{k^2 - \epsilon_d\omega^2/c^2} \quad (30)$$

is the corresponding field decay rate.

For degenerate electron statistics, Eqn. (29) generally has two distinct solutions. For a sufficiently small value of the ratio of the Fermi velocity to the speed of light in vacuum, these correspond to (i) the conventional surface plasmon, and (ii) the hyperbolic wave that is primarily supported by the hyperbolic layer [36] – see Fig. 2(b). Note that the hyperbolic surface wave is only present above the cut-off frequency that is close to that of the standard surface plasmon resonance at the plane interface ω_{sp}, when the bulk (Drude) metal permittivity

$$\epsilon_m(\omega) = \epsilon_\infty\left(1 - \frac{\omega_p^2}{\omega(\omega + i/\tau)}\right) \quad (31)$$

satisfies the resonance condition [3]

$$\epsilon_m(\omega_{sp}) = -\epsilon_d. \quad (32)$$

With the increase of the ratio v_F/c (by e.g. increasing the doping density in a semiconductor) beyond its critical value $(v_F/c)_*$, these two branches of the dispersion diagram undergo an avoided crossing (see Appendix E), so that the "conventional" surface plasmon continuously evolves into the hyperbolic mode (green curve in Fig. 2(c),(d)), while the plasmon resonance, with its peak in the frequency dependence of the in-plane wavenumber (and the corresponding photonic density of states), is strongly suppressed (magenta curve in Fig. 2(c),(d)). The physical origin of this suppression originates from the fact that plasmonic resonance relies on the resonant coupling between the electromagnetic field to the free charges in the immediate vicinity of the interface. The formation of the hyperbolic layer with strongly anisotropic electromagnetic response, no longer allows the resonance condition near the interface, and the conventional plasmon resonance is rapidly suppressed.

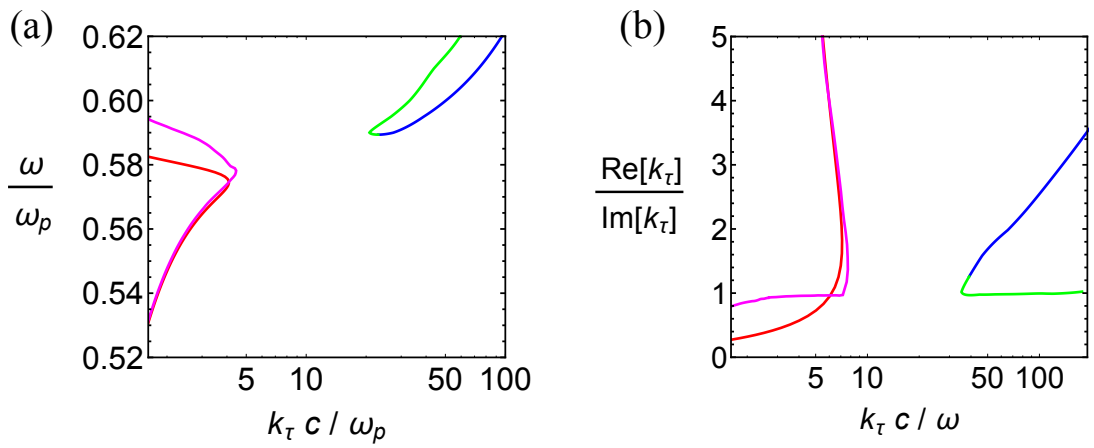

FIG. 6. (Color online) The dispersion of the surface waves at the interface of a dielectric ($\epsilon_d = 2$) with ionized plasma in the atmospheric E-layer (a), and the corresponding "figure of merit" $\mathrm{Re}\,[k_\tau]\,/\mathrm{Im}\,[k_\tau]$ vs. the "compression factor" $k_\tau c/\omega$ diagram (b). The plasma frequency in the E-layer is 5 MHz and the electron scattering time is 3.2 μsec.[40, 41] The red (dark gray) line corresponds to the standard Drude model for the surface plasmon, while the magenta (gray) line corresponds to its exact solution. The green (light gray) and blue (black) curves show the "hyper-plasmons".

One of the main challenges in nanoplasmonics is the inherent trade-off between the contradictory requirements of the surface plasmon propagation and field confinement.[3] In a conventional surface plasmon, an improvement of the "compression factor" [37] k_τ/k_0 (that defines the field confinement) can be generally achieved only at the expense of the smaller propagation distance. This is illustrated by the red curve in Fig. 4, which plots the "figure of merit" $\mathrm{Re}\,[k_\tau]\,/\,\mathrm{Im}\,[k_\tau]$ that represents the propagation distance in units of the plasmon's own wavelength, vs. the compression factor. However, the new "hyper-plasmon" surface wave that is supported by the hyperbolic layer (green curve in Fig. 4) greatly exceeds these values, for both the propagation distance and the compression factor, *simultaneously*.

Due to the singularity in the density of states of a hyperbolic medium,[38] the formation of the hyperbolic layer dramatically changes the photonic density of states near a high-quality metal-dielectric interface, with the resulting effect on all related phenomena – from radiative heat transfer to quantum-electrodynamic effects to Förster energy transfer to nonlinear optics. As an example of this behavior, in Fig. 5 we plot the spontaneous emission rate near the metal-dielectric interface, as a function of frequency (Fig. 5(a)) and the distance to the interface (Fig. 5(b)). Note the dramatic suppression of the conventional plasmon resonance, and the enhancement of the emission rate above the plasmon resonance frequency.

While the Drude theory predicts positive permittivity tensor above the plasma frequency, the inherent non-locality of the electronic response near the metal-dielectric interface dramatically modifies this simple picture. Above the plasma frequency the hyper-plasmon surface wave propagates with the in-plane wavenumber $k_\tau \gg \omega/c$, corresponding to the phase velocity $v_{\mathrm{ph}} \ll c$. For the electrons in the metal, the characteristic velocity $v \sim v_F$ can therefore be on the order of v_{ph}, which results in the Doppler phase shift that is comparable to the actual frequency ω. As a result, even with $\omega > \omega_p$, for an electron that is propagating in the direction close to that of the surface wave, the resulting Doppler-shifted frequency

$$\omega' = \omega - \mathbf{k}\cdot\mathbf{v} \tag{33}$$

can be well below ω_p, thus increasing its negative contribution to the total permittivity

$$\epsilon \simeq \epsilon_\infty \left(1 - \frac{\omega_p^2}{(\omega')^2}\right). \tag{34}$$

Therefore, even when in the stationary frame of reference the frequency ω is well above ω_p, the apparent dielectric permittivity parallel to the surface that corresponds to electromagnetic waves with large wavenumbers, is still negative. As a result, for $k \gg \omega/c$ the hyperbolic layer is still present above ω_p – which explains the continued existence of the hyper-plasmon surface wave at higher frequencies.

In the case of a non-degenerate electron gas at the interface with a dielectric, there are several hyper-plasmon surface waves, in addition to the "conventional" plasmon. This behavior is illustrated in Fig. 6 where we consider the electromagnetic waves at the surface of a dielectric surrounded by free electron plasma of the Heaviside E layer of the Earth ionosphere (at the height of ~ 100 km from the Earth surface).[39] In this case,[40, 41] the plasma frequency is 5 MHz, the electron temperature $T \simeq 750$K corresponds to the thermal velocity $v_T \simeq 0.0005c$, and the electron scattering time $\tau \simeq 3.2$ μsec (so that $\omega_p\tau \simeq 100$). Note the emergence of two separate hyper-plasmon surface waves in a tangent

bifurcation at the frequency $\omega \simeq 0.59\ \omega_p$, and relatively high losses $\mathrm{Re}\left[k_\tau\right]/\mathrm{Im}\left[k_\tau\right] < 10$ due to Laudau damping. [21, 42] With high compression factors ($k_\tau \sim 100\ \omega/c$) corresponding to the effective length on the order of one meter, such modes formed at the dielectric elements of research balloons and rocket probes, can affect the RF communications and measurements in ionosphere.

The *hyperbolic blockade* – the suppression of the plasmon resonance due to the formation of the hyperbolic layer at the metal surface, caused by the inherent non-locality of the free electron electromagnetic response – is the general feature of a high-quality metal-dielectric interface. However, a finite surface roughness leads to an effective averaging of the polarization anisotropy in the hyperbolic layer, and reduces the effect of the hyperbolic blockade. Quantitatively, this corresponds to the *short-range* roughness [43] amplitude h that exceeds the thickness of the hyperbolic layer, $\sim v_F/\omega$. Near the plasma frequency, the hyperbolic layer thickness v_F/ω_p is within a single order of magnitude from the Thomas-Fermi screening length, $v_F/\omega_p = \left(\sqrt{3\epsilon_\infty}/\sqrt[3]{\pi}\right)\ R_{\mathrm{TF}}$. In good plasmonic metals such as silver or gold, the hyperbolic layer thickness can therefore be on the order of a fraction of a nanometer, and the effect of the hyperbolic blockade in all but the highest-quality samples will be negligible.

The situation however is dramatically different in other conductors, such as transparent conducting oxides[44, 45] or doped semiconductors.[8, 9] E.g. in the latter, the thickness of the hyperbolic layer is in the range between 10 nm and 100 nm, and exceeds both the typical roughness in high-quality MBE- or MOCVD-grown samples (generally on the order of a fraction of a nanometer) and the corresponding electron de Broglie wavelength $\lambda\!\!\!^{-}$ by more than an order of magnitude – see the caption of Fig. 1.

Furthermore, the formation of the accumulation / depletion layers at the semiconductor interface can be avoided, or at least substantially reduced by choosing the materials with large bandgap discontinuity (such as e.g. in the AlGaInAs platform). Even if present however, the band curvature can still be accounted for within the same boundary condition formalism [46] and thus will not prevent the hyperbolic blockade and the emergence of the hyper-plasmon surface waves. Charges trapped at the surface states of the semiconductor-dielectric interface, may increase the diffuse component in the electron surface scattering. This however does not remove the inherent anisotropy of the electromagnetic response in the hyperbolic layer. As a result, the effect of the defect states at the interface will be limited to an increase the loss factor of the hyper-plasmon waves. Experiments on doped semiconductor materials should therefore show clear manifestations of the hyperbolic blockade and the emergence of the hyper-plasmon surface waves, predicted in the present work.

The formation of the hyperbolic layers near the metal-dielectric interface both below *and* above the plasma frequency, also offers an entirely new approach for the search of new plasmonic materials. With the requirement for the operation in the proximity to the surface plasmon resonance frequency, the material options for nanoplasmonics remain fairly limited.[44, 45] Although plasmonic bandwidth can be improved by using the metamaterial approach,[47] where one can design and fabricate a metal-dielectric composite that extends the plasmonic behavior to a broader frequency range in a variety of form-factors, from planar metamaterials [48] to core-shell plasmonic particles,[49] this comes at the cost of an increased fabrication complexity,[47] with resulting "hit" in performance due to inevitable disorder at each interface.[50] In contrast to this behavior, the hybrid "hyper-plasmons" introduced in the present work, offer high field compression factors that are not limited to the proximity to the resonance frequency ω_{sp} – and exist well above its value (see Fig. 2). To put it in the context of an actual material platform, the high-quality doped semiconductors originally introduced as plasmonic materials for mid- and far-infrared frequencies,[8] support hyper-plasmons well into the near-IR range.

V. CONCLUSIONS

In this work, we introduced the concept of the hyper-plasmonic surface wave, supported by hyperbolic layers near any high-quality metal-dielectric interface. We presented the theory of this effect that takes full account of the mobility of free charge carriers in plasmonic materials and the corresponding non-locality of the electromagnetic response. For a high-quality planar interface, we obtained the exact solution of the resulting system of coupled integro-differential equations. We demonstrated that hyper-plasmonic surface waves with simultaneously high compression factors and long propagation distance can be supported by an interface of a dielectric with conducting material, well above the corresponding plasma frequency – thus opening the field of plasmonics to many new materials, or extending the applications of existing materials in nanophotonics to shorter wavelength.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation (grant 1629276-DMR), Army Research Office (grant W911NF-14-1-0639) and Gordon and Betty Moore Foundation.

Appendix A: Boundary condition for the charge carriers distribution function.

The effect of the surface can be described by the boundary condition on the distribution function at the

interface, [17, 18] which in the general case can be expressed as

$$f_{\mathbf{p}^-}(\mathbf{r}_s) = \int d\mathbf{p}^+ \; W\left(\mathbf{p}^-, \mathbf{p}^+\right) \; f_{\mathbf{p}^+}(\mathbf{r}_s), \quad \text{(A1)}$$

where the coordinate $\mathbf{r}_s$ corresponds to the surface, $\mathbf{p}^+$ and $\mathbf{p}^-$ are the electron momenta with respectively positive and negative group velocity components in the normal to the interface direction: $\left(\mathbf{v}_{\mathbf{p}^+}\right)_{n_s} > 0$, $\left(\mathbf{v}_{\mathbf{p}^-}\right)_{n_s} < 0$, and the surface scattering indicatrix $W\left(\mathbf{p}^-, \mathbf{p}^+\right)$ can be calculated from first principles.[17, 18]

When the characteristic surface roughness is smaller then both the mean free path $\ell \equiv v_e\tau$ and v_e/ω, Eqn. (A1) can be represented in terms of the specular reflection probability[25, 26] $\mathcal{P}$ as

$$f_{\mathbf{p}^-}(\mathbf{r}_s) = \mathcal{P} \; f_{\mathbf{p}^+}(\mathbf{r}_s) + (1+\mathcal{P})\,\Phi_\varepsilon\left(\varepsilon_{\mathbf{p}^-}\right), \quad \text{(A2)}$$

with $\mathcal{P} = 1$ corresponding to the ideal interface (8) with specular reflection and $\mathcal{P} = 0$ for the opposite limit of diffuse (Lambertian) scattering of the free charge carriers. Here, the function Φ_ε is obtained from the conservation of the electron flux to and from the boundary. The specular reflection probability $\mathcal{P}$ may be treated as a phenomenological parameter, or alternatively calculated quantum-mechanically from the statistical properties of the surface roughness, [7, 17, 18] e.g. when the surface roughness correlation length is smaller than electron be Broglie wavelength λ we find [7]

$$\mathcal{P} = \exp\left(-\frac{16\pi^2 h^2}{\lambda^2}\right). \quad \text{(A3)}$$

For a high-quality interface along one of the symmetry planes of the crystal, Eqn. (A2) reduces to the specular reflection boundary condition at the surface [7, 17, 18]

$$f_{\mathbf{p}^-}(\mathbf{r}_s) = f_{\mathbf{p}^+}(\mathbf{r}_s), \quad \text{(A4)}$$

where $\mathbf{p}^+$ and $\mathbf{p}^-$ are now connected by the specular reflection condition, with equal tangential to the surface components $p_\tau^+ = p_\tau^-$.

Appendix B: Electromagnetic field and charge carrier distribution at the metal-dielectric interface.

The electromagnetic field, and charge and carrier densities near the metal dielectric interface are defined by the self-consistent solution of the system of Maxwell's equations,

$$\text{div}\ \mathbf{D} = 4\pi\rho(\mathbf{r},t) \quad \text{(B1)}$$

$$\text{div}\ \mathbf{B} = 0 \quad \text{(B2)}$$

$$\text{curl}\ \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \quad \text{(B3)}$$

$$\text{curl}\ \mathbf{B} = -\frac{4\pi}{c}\mathbf{j}(\mathbf{r},t) + \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}, \quad \text{(B4)}$$

where the displacement field

$$\mathbf{D} = \epsilon\mathbf{E} = \begin{cases} \epsilon_d\ \mathbf{E}, & z<0 \\ \epsilon_\infty\ \mathbf{E}, & z>0 \end{cases}, \quad \text{(B5)}$$

ϵ_d is permittivity of the dielectric and ϵ_∞ is the "background" permittivity of the crystal lattice in the conductor, while the free charge density $\rho(\mathbf{r},t)$ and the free current density $\mathbf{j}(\mathbf{r},t)$ are defined by the charge carrier distribution function $f_{\mathbf{p}}(\mathbf{r},t)$ via

$$\rho(\mathbf{r},t) = 2\int \frac{d\mathbf{p}}{(2\pi\hbar)^3}\cdot\left(f_{\mathbf{p}}(\mathbf{r}) - f_0\left(\varepsilon_{\mathbf{p}}\right)\right), \quad \text{(B6)}$$

$$\mathbf{j}(\mathbf{r},t) = 2\int \frac{d\mathbf{p}}{(2\pi\hbar)^3}\cdot e\mathbf{v}_{\mathbf{p}} f_{\mathbf{p}}(\mathbf{r},t). \quad \text{(B7)}$$

In the liner response regime, the charge carrier distribution function $f_{\mathbf{p}}(\mathbf{r},t)$ satisfies the Boltzmann kinetic equation

$$\frac{\partial f_{\mathbf{p}}}{\partial t} + \mathbf{v}_{\mathbf{p}}\cdot\nabla f_{\mathbf{p}} + e\mathbf{E}\cdot\mathbf{v}_{\mathbf{p}}\frac{\partial f_0}{\partial \varepsilon_{\mathbf{p}}} = -\frac{f_{\mathbf{p}} - f_0}{\tau}, \quad \text{(B8)}$$

with the boundary condition at the metal-dielectric interface (see also Eqn. (A1)

$$f_{\mathbf{p}}|_{z=0,v_z<0} = \int_{v_z'>0} d\mathbf{p}'\; W(\mathbf{p},\mathbf{p}')\;\; f_{\mathbf{p}'}|_{z=0}. \quad \text{(B9)}$$

When the surface roughness is much smaller than the charge carrier de Broglie wavelength, $h \ll \lambda$, or if $h \simeq \lambda$ and surface roughness correlation length $L \gg \lambda$, Eqn. (B9) reduces to the specular reflection boundary condition (see also Eqns. (8) and (A4))

$$f(v_x, v_y, v_z)|_{z=0} = f(v_x, v_y, -v_z)|_{z=0}. \quad \text{(B10)}$$

For a harmonic wave with the in-plane momentum k in the x-direction,

$$\mathbf{E}(\mathbf{r},t) = (E_x(z), 0, E_z(z))\ \exp(ikx - i\omega t), \quad \text{(B11)}$$

$$\mathbf{B}(\mathbf{r},t) = (0, B(z), 0)\ \exp(ikx - i\omega t), \quad \text{(B12)}$$

$$f_{\mathbf{p}}(\mathbf{r},t) = f_0(\varepsilon) + f(\mathbf{v},z)\ \exp(ikx - i\omega t), \quad \text{(B13)}$$

Note that in the harmonic representation (B11),(B12),(B13), Eqns. (B1),(B2) directly follow from (B3),(B4), and therefore do not represent independent constrains onto the electromagnetic field and the charge carrier distribution function.[52]

Applying curl to (B3), and using (B4), (B7), (B11), (B13), for $z > 0$ we obtain

$$-\frac{\partial^2 E_x}{\partial z^2} + ik\frac{\partial E_z}{\partial z} = \frac{4\pi i\omega}{c^2} j_x + \epsilon_\infty\left(\frac{\omega}{c}\right)^2 E_x, \quad \text{(B14)}$$

$$ik\frac{\partial E_x}{\partial z} + k^2 E_z = \frac{4\pi i\omega}{c^2} j_z + \epsilon_\infty\left(\frac{\omega}{c}\right)^2 E_z, \quad \text{(B15)}$$

where

$$j_{x,z} = 2e\int \frac{d\mathbf{p}}{(2\pi\hbar)^3}\; v_{x,z}\; f(\mathbf{v},z). \quad \text{(B16)}$$

Substituting (B13) into the kinetic equation (B8) and the boundary condition (B10), we obtain

$$\begin{aligned} f(\mathbf{v},z) = & -e\,\frac{\theta(v_z)}{v_z}\frac{\partial f_0}{\partial\varepsilon}\int_0^\infty d\zeta\ \left(v_x E_x(\zeta)+v_z E_x(\zeta)\right) \\ & \times\exp\left(-\frac{\zeta+z}{v_z}\left(\frac{1}{\tau}-i\omega+ikv_z\right)\right) \\ & -e\,\frac{\theta(v_z)}{v_z}\frac{\partial f_0}{\partial\varepsilon}\int_0^z d\zeta\ \left(v_x E_x(\zeta)+v_z E_x(\zeta)\right) \\ & \times\exp\left[\frac{\zeta-z}{v_z}\left(\frac{1}{\tau}-i\omega+ikv_z\right)\right] \\ & +e\,\frac{\theta(-v_z)}{v_z}\frac{\partial f_0}{\partial\varepsilon}\int_z^\infty d\zeta\ \left(v_x E_x(\zeta)+v_z E_x(\zeta)\right) \\ & \times\exp\left[\frac{\zeta-z}{v_z}\left(\frac{1}{\tau}-i\omega+ikv_z\right)\right]. \end{aligned} \tag{B17}$$

Following the approach of Ref. [28], originally developed in the context of the calculation of surface impedance of metals at microwave frequencies, we introduce the auxiliary fields

$$\mathcal{E}_x(z) = E_x(|z|), \tag{B18}$$

and

$$\mathcal{E}_z(z) = E_z(|z|)\,\mathrm{sign}(z), \tag{B19}$$

that represent respectively even- and odd "extension" of the electric field in the conductor ($z>0$) to the entire range $-\infty<z<\infty$.

Substituting (B18) and (B19) together with (B16) and (B17) into (B14) and (B15), we obtain

$$\begin{aligned} & \frac{\partial^2\mathcal{E}_x}{\partial z^2}+\epsilon_\infty\left(\frac{\omega}{c}\right)^2\mathcal{E}_x - ik\frac{\partial\mathcal{E}_z}{\partial z} \\ & \quad = -\frac{4\pi i\omega}{c^2}\int_{-\infty}^\infty d\zeta\ K_{xx}(z-\zeta)\ \mathcal{E}_x(\zeta) \\ & \qquad -\frac{4\pi i\omega}{c^2}\int_{-\infty}^\infty d\zeta\ K_{xz}(z-\zeta)\ \mathcal{E}_z(\zeta), \end{aligned} \tag{B20}$$

and

$$\begin{aligned} & -ik\frac{\partial\mathcal{E}_x}{\partial z}+\left(\epsilon_\infty\left(\frac{\omega}{c}\right)^2-k^2\right)\ \mathcal{E}_z \\ & \quad = -\frac{4\pi i\omega}{c^2}\int_{-\infty}^\infty d\zeta\ K_{zx}(z-\zeta)\ \mathcal{E}_x(\zeta) \\ & \qquad -\frac{4\pi i\omega}{c^2}\int_{-\infty}^\infty d\zeta\ K_{zz}(z-\zeta)\ \mathcal{E}_z(\zeta), \end{aligned} \tag{B21}$$

where

$$\begin{aligned} K_{xx}(u) = & \ 2\int_{v_z>0}\frac{d\mathbf{p}}{(2\pi\hbar)^3}\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\frac{v_x^2}{v_z} \\ & \times\exp\left(-(1-i\omega\tau+ikv_x\tau)\,\frac{|u|}{v_z\tau}\right), \end{aligned} \tag{B22}$$

$$\begin{aligned} K_{xz}(u) = K_{zx}(u) = & \ 2\,\mathrm{sign}(u)\int_{v_z>0}\frac{d\mathbf{p}}{(2\pi\hbar)^3}\left(-\frac{\partial f_0}{\partial\varepsilon}\right)v_x \\ & \times\exp\left(-(1-i\omega\tau+ikv_x\tau)\,\frac{|u|}{v_z\tau}\right), \end{aligned} \tag{B23}$$

$$\begin{aligned} K_{zz}(u) = & \ 2\int_{v_z>0}\frac{d\mathbf{p}}{(2\pi\hbar)^3}\left(-\frac{\partial f_0}{\partial\varepsilon}\right)v_z \\ & \times\exp\left(-(1-i\omega\tau+ikv_x\tau)\,\frac{|u|}{v_z\tau}\right). \end{aligned} \tag{B24}$$

Despite its relative complexity, the system of coupled linear integro-differential equations (B20),(B21) only has difference kernels, and by means of the Fourier transform

$$e_x(k,q) = \int_{-\infty}^\infty dz\ \mathcal{E}_x\exp(iqz), \tag{B25}$$

$$e_z(k,q) = \int_{-\infty}^\infty dz\ \mathcal{E}_z\exp(iqz), \tag{B26}$$

can be reduced to a system of linear algebraic equations.[51] We therefore obtain

$$e_x(k,q) = \frac{2\ A(k)}{D(k,q)}\left(\epsilon_{zz}(k,q)\,\frac{\omega^2}{c^2}-k^2\right), \tag{B27}$$

$$e_z(k,q) = \frac{2\ A(k)}{D(k,q)}\ \nu_{xz}(k,q), \tag{B28}$$

where

$$A(k) = \left.\frac{\partial E_x}{\partial z}\right|_{z=+0} - ik\ E_z|_{z=+0}, \tag{B29}$$

$$\begin{aligned} D(k,q) = & \left(\epsilon_{xx}(q)\,\frac{\omega^2}{c^2}-q^2\right) \\ & \times\left(\epsilon_{zz}(q)\,\frac{\omega^2}{c^2}-k^2\right)-\nu_{xz}^2(k,q), \end{aligned} \tag{B30}$$

and

$$\begin{aligned} \epsilon_{xx}(k,q) = & \ \epsilon_\infty - \frac{16\pi ie^2}{\omega}\int_0^\infty du\cos(qu)\int_{v_z>0}\frac{d\mathbf{p}}{(2\pi\hbar)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}} \\ & \times\frac{v_x^2}{v_z}\ \exp\left(-(1-i\omega\tau+ikv_x\tau)\,\frac{u}{v_z\tau}\right), \end{aligned} \tag{B31}$$

$$\begin{aligned} \epsilon_{zz}(k,q) = & \ \epsilon_\infty - \frac{16\pi ie^2}{\omega}\int_0^\infty du\cos(qu)\int_{v_z>0}\frac{d\mathbf{p}}{(2\pi\hbar)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}} \\ & \times v_z\ \exp\left(-(1-i\omega\tau+ikv_x\tau)\,\frac{u}{v_z\tau}\right), \end{aligned} \tag{B32}$$

$$\begin{aligned} \nu_{xz}(k,q) = & \ kq - \frac{16\pi e^2}{\omega}\int_0^\infty du\ \sin(qu)\int_{v_z>0}\frac{d\mathbf{p}}{(2\pi\hbar)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}} \\ & \times v_x\ \exp\left(-(1-i\omega\tau+ikv_x\tau)\,\frac{u}{v_z\tau}\right). \end{aligned} \tag{B33}$$

For $z > 0$, the auxiliary field $\mathcal{E}$ is identical to the actual electric field $\mathbf{E}$, and Eqns. (B25) - (B33) therefore offer the exact analytical solution for the electric field in the metal:

$$\mathbf{E}\left(z>0\right)=\int_{-\infty}^{\infty}\frac{dq}{2\pi}\ \mathbf{e}\left(k,q\right)\ \exp\left(-iqz\right). \quad \text{(B34)}$$

The amplitude $A(k)$ in Eqn. (B29) is defined by the values of the normal component of the electrical field $E_z|_{z=+0}$ and the normal derivative of the tangential electric field $\partial E_x/\partial z|_{z=+0}$ at the boundary. These magnitudes depend of the electric field in the dielectric ($z < 0$), and are obtained from the continuity of the tangential components of the electrical field and the normal components of the displacement vector

$$E_x|_{z=-0}=E_x|_{z=+0}\,, \quad \text{(B35)}$$
$$\epsilon_d\ E_z|_{z=-0}=\epsilon_\infty\ E_z|_{z=+0}\,, \quad \text{(B36)}$$

where ϵ_d is the permittivity of the dielectric.

Finally, the u-integration in Eqns. (B31),(B32),(B33) can be performed analytically, which yields

$$\epsilon_{xx}\left(k,q\right)=\epsilon_\infty-\frac{16\pi ie^2\tau}{\omega}\int_{v_z>0}\frac{d\mathbf{p}}{\left(2\pi\hbar\right)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}}\times v_x^2\ \frac{1-i\omega\tau+ikv_x\tau}{\left(1-i\omega\tau+ikv_x\right)^2+q^2v_z^2\tau^2}, \quad \text{(B37)}$$

$$\epsilon_{zz}\left(k,q\right)=\epsilon_\infty-\frac{16\pi ie^2\tau}{\omega}\int_{v_z>0}\frac{d\mathbf{p}}{\left(2\pi\hbar\right)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}}\times v_z^2\ \frac{1-i\omega\tau+ikv_x\tau}{\left(1-i\omega\tau+ikv_x\tau\right)^2+q^2v_z^2\tau^2}, \quad \text{(B38)}$$

$$\nu_{xz}\left(k,q\right)=kq-\frac{16\pi e^2\tau^2\omega q}{c^2}\int_{v_z>0}\frac{d\mathbf{p}}{\left(2\pi\hbar\right)^3}\frac{\partial f_0}{\partial\varepsilon_{\mathbf{p}}}\times v_xv_z^2\ \frac{1-i\omega\tau+ikv_x\tau}{\left(1-i\omega\tau+ikv_x\tau\right)^2+q^2v_z^2\tau^2}. \quad \text{(B39)}$$

Appendix C: The reflection amplitude at the planar metal-dielectric boundary.

For a given in-plane momentum k, the electric electric field in the dielectric ($z < 0$) with the permittivity ϵ_d can be expressed as

$$\mathbf{E}\left(\mathbf{r},t\right)=E_+\left(1,0,-\frac{k}{\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}}\right)\times\exp\left(ikx+i\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}\ z-i\omega t\right)+E_-\left(1,0,\frac{k}{\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}}\right)\times\exp\left(ikx-i\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}\ z-i\omega t\right), \quad \text{(C1)}$$

leading to the corresponding magnetic field

$$\mathbf{B}\left(\mathbf{r},t\right)=\frac{c}{i\omega}\ \mathrm{curl}\ \mathbf{E}=\hat{\mathbf{y}}\left[E_+\exp\left(i\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}z\right)-E_+\exp\left(i\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}z\right)\right]\times\frac{\epsilon_d\ \omega/c}{\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}}\exp\left(ikx-i\omega t\right). \quad \text{(C2)}$$

Therefore the electromagnetic wave impedance [52, 53] in the $z = -0$ plane

$$Z|_{z=-0}\equiv\left.\frac{E_x}{B_y}\right|_{z=+0}=\frac{r+1}{r-1}\ \frac{\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}}{\epsilon_d\ \omega/c}, \quad \text{(C3)}$$

where the reflection coefficient

$$r\equiv\frac{E_+}{E_-}. \quad \text{(C4)}$$

On the other hand, from Eqns. (B27)-(B34) the tangential electric field at the metal side of the interface

$$E_x|_{z=+0}=\left(\left.\frac{\partial E_x}{\partial z}\right|_{z=+0}-ik\ E_z|_{z=+0}\right)\times\frac{1}{\pi}\int_{-\infty}^{\infty}dq\ \frac{\epsilon_{zz}\left(k,q\right)\left(\omega/c\right)^2-k^2}{D\left(k,q\right)}, \quad \text{(C5)}$$

while the magnetic field

$$\mathbf{B}|_{z=+0}=\left.\frac{c}{i\omega}\ \mathrm{curl}\ \mathbf{E}\ \right|_{z=+0}=\frac{c}{i\omega}\ \hat{\mathbf{y}}\left(\left.\frac{\partial E_x}{\partial z}\right|_{z=+0}-ik\ E_z|_{z=+0}\right), \quad \text{(C6)}$$

so that the corresponding wave impedance in the $z = +0$ plane

$$Z|_{z=+0}\equiv\left.\frac{E_x}{B_y}\right|_{z=+0}=\frac{i\omega}{\pi c}\int_{-\infty}^{\infty}dq\ \frac{\epsilon_{zz}\left(k,q\right)\left(\omega/c\right)^2-k^2}{D\left(k,q\right)}. \quad \text{(C7)}$$

From Eqns. (C3) and (C7) for the reflection coefficient r we therefore obtain

$$r=-1+2\left\{1+i\ \frac{\epsilon_d\left(\omega/c\right)^2}{\sqrt{\epsilon_d\left(\omega/c\right)^2-k^2}}\times\frac{1}{\pi}\int_{-\infty}^{\infty}dq\ \frac{\epsilon_{zz}\left(k,q\right)\left(\omega/c\right)^2-k^2}{D\left(k,q\right)}\right\}^{-1}. \quad \text{(C8)}$$

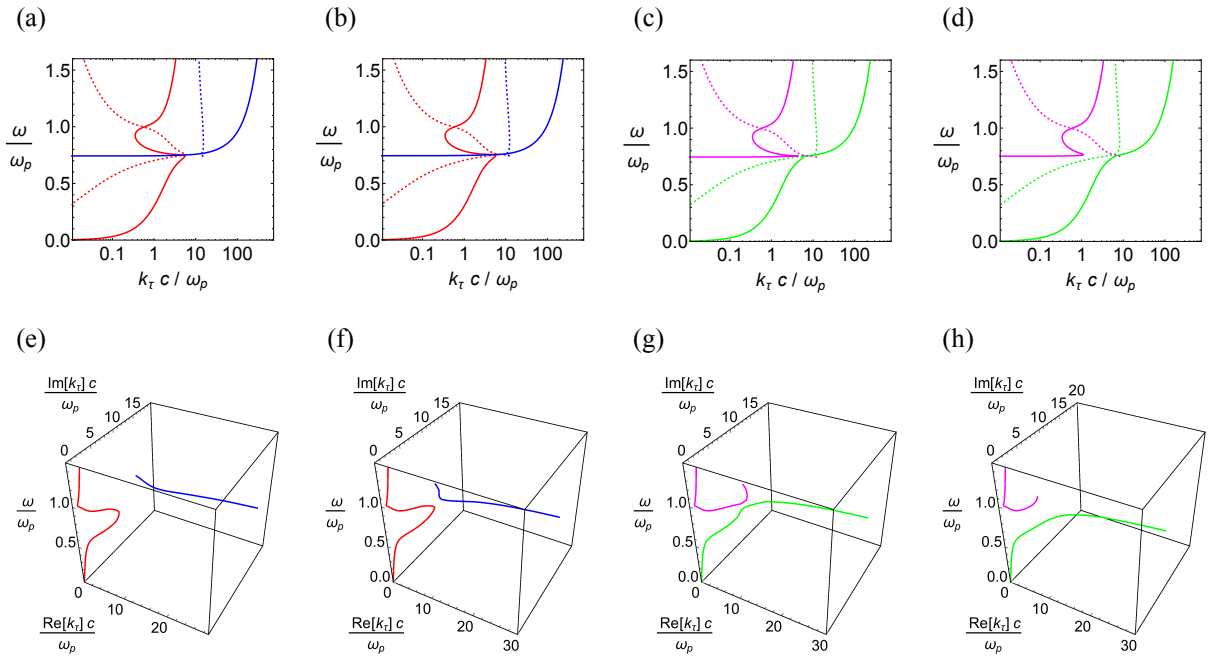

FIG. 7. (Color online) The evolution of the "crossing" of the plasmonic (red (dark gray) line) and the hyperbolic mode (blue (black)), into the hybrid hyper-plasmonic (green (light gray) line) and "residual" (magenta (gray)) modes, with the increase of v_F/c. Panels (a,e): $v_F/c = 0.005$, panels (b,f): $v_F/c = 0.0062$, panel (c,g): $v_F/c = 0.0063$, panels (d,h): $v_F/c = 0.00935$. Other material parameters (ϵ_∞, τ, ϵ_d) correspond to the semiconductor system AlInAs/InGaAs, and are the same as in Fig. 1. With the plasma wavelength $\lambda_p = 10$ μm, the doped semiconductor system AlInAs/InGaAs corresponds to the panels (d,h). The dotted lines in panels (a) - (d) represent the imaginary parts of the in-plane momentum k_τ.

Appendix D: Surface waves at the metal-dielectric interface.

For a surface wave at the metal-dielectric interface with the in-plane momentum $k > \sqrt{\epsilon_d}\ \omega/c$, the electric field in the dielectric half-space $z < 0$ is given by

$$\mathbf{E}(\mathbf{r},t) = E_0 \left(1, 0, -\frac{ik}{\sqrt{k^2 - \epsilon_d (\omega/c)^2}}\right) \times \exp\left(ikx + \sqrt{k^2 - \epsilon_d (\omega/c)^2}\ z - i\omega t\right), \quad \text{(D1)}$$

while the corresponding magnetic field

$$\mathbf{B}(\mathbf{r},t) = \hat{\mathbf{y}}\ E_0 \frac{i\epsilon_d\ \omega/c}{\sqrt{k^2 - \epsilon_d (\omega/c)^2}} \times \exp\left(ikx + \sqrt{k^2 - \epsilon_d (\omega/c)^2}\ z - i\omega t\right). \quad \text{(D2)}$$

The wave impedance at $z = -0$ is therefore given by

$$Z|_{z=-0} \equiv \left.\frac{E_x}{B_y}\right|_{z=-0} = \frac{\sqrt{k^2 - \epsilon_d (\omega/c)^2}}{i\ \epsilon_d\ \omega/c}. \quad \text{(D3)}$$

From Eqns. (C7) and (D3)

$$\frac{1}{\pi}\int_{-\infty}^{\infty} dq\ \frac{\epsilon_{zz}(k,q)(\omega/c)^2 - k^2}{D(k,q)} = -\frac{\epsilon_d (\omega/c)^2}{\sqrt{k^2 - \epsilon_d (\omega/c)^2}}, \quad \text{(D4)}$$

which defines the dispersion law of the surface wave $\omega(k)$.

Appendix E: "Crossing" to "Avoided Crossing" crossover

The dispersion equation for the surface modes at the conductor-dielectric interface, Eqn. (29) for a degenerate electron gas generally has two distinct solutions. For a sufficiently small value of the ratio of the Fermi velocity to the speed of light in vacuum, these correspond to the conventional surface plasmon (red (dark gray) curve in Fig. 7 (a),(b) and (e),(f)), and the hyperbolic wave that is primarily supported by the hyperbolic layer (blue (black) curve in see Fig. 7 (a),(b) and (e),(f)). In this regime, there is a large difference in the lifetimes of the "plasmonic" and the "hyperbolic" surface waves, so the seemingly un-avoided crossing in the plot of the real parts of the wavenumber and the frequency in Fig. 7 (a),(b) is a direct consequence of this behavior – in the full phase

space (see Fig. 7 (e),(f)) these two modes actually stay far apart from each other.

With the increase of the ratio v_F/c (by e.g. increasing the doping density in a semiconductor) however, the corresponding lifetimes approach each other, and at the critical value of v_F/c the "plasmonic" and the "hyperbolic" modes finally become degenerate. At higher v_F/c the dispersion diagram shows an avoided crossing – as seen in Fig. 7 (c),(g) calculated for the value of v_F/c just above this critical point. From now on, with an increase of the frequency, the "conventional" surface plasmon continuously evolves into the hyperbolic mode – see the evolution of the magenta curve in Fig. 7 (c),(d) and (g),(h). At the same time, the standard plasmonic resonance, which generally manifests itself by the peak in the frequency dependence of the in-plane wavenumber (and the corresponding photonic density of states), is strongly suppressed – see the behavior of the magenta curve in Fig. 7(c),(d) and note the use of the logarithmic scale.

[1] C. Ciraci, R. T. Hill, J. J. Mock, Y. Urzhumov, A. I. Fernández-Dominguez, S. A. Maier, J. B. Pendry, A. Chilkoti, and D. R. Smith, "Probing the Ultimate Limits of Plasmonic Enhancement," Science **337** (6098), 1072 - 1074 (2012).

[2] G. Toscano, J. Straubel, A. Kwiatkowski, C. Rockstuhl, F. Evers, H. Xu, N. A. Mortensen, M. Wubs, "Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics," Nature Communications **6**, 7132 (2015).

[3] S. A. Maier, "Plasmonics: Fundamentals and Applications," (Springer; 1st edition, 2007).

[4] V. A. Podolskuiy and E. E. Narimanov, "Strongly anisotropic waveguide as a nonmagnetic left-handed system," Phys. Rev. B **71**, 201101(R) (2005).

[5] S. S. Kruk, Z. J. Wong, E. Pshenay-Severin, K. O'Brien, D. N. Neshev, Yu. S. Kivshar and X. Zhang, "Magnetic hyperbolic optical metamaterials," Nature Communications **7**, 11329 (2016).

[6] J. B. Khurgin, "Ultimate limit of field confinement by surface plasmon polaritons," Faraday Discuss. **178**,109 - 122 (2015).

[7] J. Ziman, "Electrons and Phonons: The Theory of Transport Phenomena in Solids ," (Oxford University Press; reprint edition, 2001)

[8] A. J. Hoffman, L. V. Alekseyev, S. S. Howard, K. J. Franz, D. Wasserman, V. A. Podolskiy, E. E. Narimanov, D. L. Sivco and C. Gmachl, "Negative refraction in semiconductor metamaterials," Nature Materials **6**, 948 (2007).

[9] Y. Zhong, S. Malagari, T. Hamilton, and D. Wasserman, "Mid-Infrared Plasmonic Materials," J. Nanophotonics **9**(1), 093791 (2015).

[10] A. D. Boardman, "Hydrodynamic theory of plasmon-polaritons on plane surfaces," in *Electromagnetic Surface Modes*, (Wiley, 1982).

[11] A. Eguiluz, J. J. Quinn, "Hydrodynamics model for surface plasmons in metals and degenerate semiconductors," Phys. Rev. B **14**, 1347 (1976).

[12] C. Ciraci, J. B. Pendry, and D. R. Smith, "Hydrodynamic Model for Plasmonics: A Macroscopic Approach to a Microscopic Problem," ChemPhysChem **14**, 1109 - 1116 (2013).

[13] 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).

[14] L. D. Landau, L. P. Pitaevskii, and E.M. Lifshitz, "Electrodynamics of Continuous Media," (Butterworth-Heinemann; 2nd edition, 1984).

[15] L. Brillouin, Comptes Rendus hebdomadaires des Sances de lAcadmie des Sciences, Paris **173**, 1167 (1921).

[16] H. Breuer, F. Petruccione, "The theory of open quantum systems," (Oxford University Press; 1st edition, 2002).

[17] V.I. Okulov, V. V. Ustinov, "Boundary condition for the distribution function of conduction electrons scattered by a metal surface," Sov. Phys. JETP **40** (3), 584 - 590 (1975).

[18] V. I. Okulov, and V. V. Ustinov, "Surface Scattering of Conductivity Electrons and Kinetic Phenomena in Metals," Fizika Nizkikh Temperatur **5** (3), 213 - 251 (1979).

[19] E. Wigner, "On the Quantum Correction for Thermodynamic Equilibrium," Physical Review **40** (5), 749 (1932).

[20] W. Kohn and J. M. Luttinger, "Quantum Theory of Electrical Transport Phenomena," Phys. Rev. **108**, 590 - 611 (1957).

[21] L. D. Landau and L. M. Lifshitz, "Quantum Mechanics," (Butterworth-Heinemann; 3rd edition, 1981).

[22] Here there is no restriction on the relative value of the surface roughness h as compared to the de Broglie wavelength $\lambda\!\!\!^{-}$.

[23] A. F. Andreev, "Interaction of Conducting Electrons with the Metal Surface," Uspekhi Fiz. Nauk **105** (1), 113 - 124 (1971).

[24] S.B. Soffer, "Statistical Model for the Size Effect in Electrical Conduction," J. Appl. Phys. **38** (4), 1710 - 1715, (1967).

[25] K. Fuchs, "The Conductivity of Thin Metallic Films According To the Electron Theory of Metals," Proc Cambridge Phil. Soc. **34** (1), 100 - 108 (1938).

[26] E.H. Sondheimer, "The Mean Free Path of Electrons in Metals," Advances in Physics bf 50 (6), 499 - 537 (2001).

[27] E. A. Kaner, A. A. Krokhin, N. M. Makarov, and V. A. Yampolskii, "Surface absorption of electromagnetic waves in metals by random boundary inhomogeneities," Sov. Phys. JETP **52** (5), 938 - 944 (1980).

[28] G. E. H. Reuter and E. H. Sondheimer, "The Theory of the Anomalous Skin Effect in Metals," Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **195** (1042) , 336 - 364 (1948).

[29] Note that for an anisotropic Fermi surface, $\mathbf{p}_n^-$ is not necessarily equal to $-\mathbf{p}_n^+$. [7]

[30] P. Hohenberg, and W. Kohn, "Inhomogeneous electron gas," Phys. Rev. **136**, 864 (1964).

[31] T. Christensen, "From Classical to Quantum Plasmonics in Three and Two Dimensions," (Springer, 2017).

[32] An exact solution can be also obtained for the opposite case of diffuse scattering. While qualitatively similar, the

resulting expressions show a higher degree of algebraic complexity, and will be given elsewhere.

[33] G. W.Ford and W.H.Weber, Phys. Rep. **113**, 195 (1984).

[34] E. Narimanov, unpublished.

[35] The inequality $\varepsilon_F \gg k_B T$ is generally well satisfied for plasmonic materials even at room temperature: while $k_B T \simeq 20$ meV, the Fermi energy is usually in the excess of an electron-volt.

[36] Note that, while a conventional hyperbolic waveguide supports multiple guided modes, [4] these correspond to progressively higher wavenumbers k_τ. In a conducting medium however, these higher-order modes with $k \gg v_F/\omega$ are strongly suppressed due to Landau damping. [42]

[37] N. Rivera, I. Kaminer, B. Zhen, J. D. Joannopoulos, M. Soljačić, "Shrinking light to allow forbidden transitions on the atomic scale," Science **353** (6296), 263 - 269 (2016).

[38] I. I. Smolyaninov, and E. E. Narimanov, "Metric Signature Transitions in Optical Metamaterials," Phys. Rev. Lett. 105, 067402 (2010).

[39] K.-I. Oyama, T. Abe, H. Mori, and J. Y. Liu, "Electron temperature in nighttime sporadic E layer at midlatitude," Ann. Geophys. **26**, 533 - 541 (2008).

[40] K. E. Oughstun, "Electromagnetic and Optical Pulse Propagation 2," (Springer, 2009).

[41] M. A. Messier, "A standard ionosphere for the study of electromagnetic pulse propagation," Tech. Rep. Note 117, Air Force Weapons Laboratory, Albuquerque, NM, 1971.

[42] L. P. Pitaevskii, and E.M. Lifshitz, "Physical Kinetics," (Butterworth-Heinemann; 1st edition, 1981).

[43] A slowly-varying (on the scales of the electron de Broglie wavelength and the Thomas-Fermi screening length $R_{\rm TF}$) surface roughness will simply lead to the hyperbolic layer adiabatically following its landscape, and will not suppress the hyperbolic blockade.

[44] J. Kim, G. V. Naik, N. K. Emani, U. Guler, A. Boltasseva, "Plasmonic Resonances in Nanostructured Transparent Conducting Oxide Films," IEEE Journal of Selected Topics in Quantum Electronics **19**, 4601907, (2013).

[45] G. V. Naik, V. M. Shalaev, A. Boltasseva, "Alternative Plasmonic Materials: Beyond Gold and Silver," Advanced Materials **25** (24), 32643294 (2013).

[46] V. A. Kozlov, E. E. Narimanov, K. A. Sakharov, "Possibility of determining the surface band curvature in semimetals," JETP Letters **53** (5), 267 - 270 (1991)

[47] W. Cai and V. Shalaev, "Optical Metamaterials: Fundamentals and Applications," (Springer, 1st edition, 2009).

[48] Z. Liu, M. D. Thoreson, A. V. Kildishev, and V. M. Shalaev, "Translation of nanoantenna hot spots by a metal-dielectric composite superlens," Appl. Phys. Lett. **95**, 033114 (2009).

[49] S. J. Oldenburg, R.D. Averitt, S.L. Westcott and N.J. Halas, "Nanoengineering of optical resonances," Chem. Phys. Lett. **288**, 243 - 247 (1998).

[50] H. Reddy, U. Guler, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Temperature-dependent optical properties of gold thin films," Optical Materials Express **6** (9), 2776 - 2802 (2016).

[51] P. M. Morse and H. Feshbach, "Methods of Theoretical Physics," part II (Feshbach Publishing, 2004).

[52] S. Ramo, J. R. Whinnery, T. Van Duzer, "Fields and Waves in Communication Electronics," (John Wiley & Sons.; 3rd edition, 1994).

[53] S. A. Schelkunoff, "The Impedance Concept and Its Application to Problems of Reflection, Refraction, Shielding and Power Absorption," Bell Syst. Tech. J. **17** (1), 17 - 48 (1938).