

# Nonlocal Effects in Plasmonic Nanostructures' Optical Response and Electron Scattering

Author: Jiantao Kong

Persistent link: <http://hdl.handle.net/2345/bc-ir:108243>

This work is posted on [eScholarship@BC](#),  
Boston College University Libraries.

---

Boston College Electronic Thesis or Dissertation, 2018

Copyright is held by the author. This work is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0>).

# Nonlocal Effects in Plasmonic Nanostructures' Optical Response and Electron Scattering

Jiantao Kong

A dissertation  
submitted to the Faculty of  
the Department of Physics  
in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy in Physics

Boston College  
Morrissey College of Arts and Sciences  
Graduate School

October 2018



# NONLOCAL EFFECTS IN PLASMONIC NANOSTRUCTURES' OPTICAL RESPONSE AND ELECTRON SCATTERING

Jiantao Kong

Advisor: Prof. Krzysztof Kempa

## Abstract:

Nonlocal effects, the wavenumber dependence in a medium's response to external disturbance, is treated in this thesis. Numerical computation methods to include nonlocal effects in plasmonic nanostructures' electromagnetic response are discussed, and applications of plasmonics to a few other fields are elaborated. First, a computation scheme is proposed to extend conventional finite-difference time-domain (FDTD) methods to nonlocal domain. An effective film whose response is derived from Feibelman's d-function formalism is to replace the highly non-uniform metal surfaces in simulations. It successfully produces numerical results of plasmonic resonance shift and field enhancement which agrees with the experimental data to first order. This scheme is still classical, thus very fast compared to the other first principle quantum methods such as density functional theory. Then electron's scattering rate in an effective medium with plasmonic nanostructures embedded-in, in random phase approximation, is developed, with the wavenumber dependence in the medium's response accounted. Utilizing this calculation scheme of electron's scattering rate, further specific applications are following. We show by simulation of the plasmonic nanostructures and calculation of the electron scattering rates that hot-electron plasmon-protection (HELPP) effects can protect the extra energy of hot electrons from being dissipated as heat. This can be a prototype of the 3<sup>rd</sup> generation solar cells. In another application, we investigate the electron polar-optical-phonon (POP) scattering in heavily-doped semiconductors when plasmonic nanostructures are embedded-in. We show that electron-POP scattering can be significantly suppressed compared to that of bulk semiconductors. In the third application, we propose the plasmonic multiple exciton generation (PMEG) scheme, with simulations and calculations, showing that the efficiency of multiple exciton generation in conventional semiconductors could be enhanced significantly with proper designed plasmonic nanostructures embedded-in or attached-adjacent.

## Acknowledgements

I would like to thank my teachers, fellow graduate students, friends and my family for all their help and support during my years of graduate study.

First of all, I would like to express my sincere gratitude to my advisor Prof. Krzysztof Kempa for the continuous support of my PhD research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not imagine having a better advisor and mentor for my PhD study.

I would also like to thank the other members of my thesis committee: Prof. Kevin Bedell, Prof. David Broido, and Prof. Michael Naughton, for their insightful comments and encouragement, and for the fruitful collaborations in research projects in the past years.

I would like to thank my fellow PhD students and friends at BC Physics for the stimulating discussions, and for all the fun we have had in the years. I have learned a lot from many of them. Another big thanks goes to all the staff members at BC physics, for their help in the past years.

I would like to thank my parents and my sister. Without their persistent support and care, I could not have achieved anything.

## Publications

1. **J. Kong**, A. Shvonski, K. Kempa, "Nonlocal response with local optics", *Physical Review B* 97, 165423 (2018)
2. **J. Kong**, A. H. Rose, C. Yang, X. Wu, J. M. Merlo, M. J. Burns, M. J. Naughton, and K. Kempa, "Hot electron plasmon-protected solar cell", *Optics Express* 23, A1087 (2015)
3. **J. Kong**, T. Sun (co-first), F. Cao, Y. Li, J. Bao, J. Gao, G. Zhou, C. Guo, Z. Ren, and K. Kempa, "A metamaterial-plasmonic scheme based on a random metallic network for controlling thermal emission", *Physica Status Solidi A* 215, 1800206 (2018)
4. A. Shvonski, **J. Kong**, K. Kempa, "Nonlocal extensions of the electromagnetic response of plasmonic and metamaterial structures", *Physical Review B* 95, 045149 (2017)
5. C. Yang, J. M. Merlo, **J. Kong**, Z. Xian, B. Han, G. Zhou, J. Gao, M. J. Burns, K. Kempa, and M. J. Naughton, "All-solution processed, template-based, scalable Ag network transparent conductor", *Physica Status Solidi A*, 1700504 (2017)
6. R. Li, Q. Peng, B. Han, Y. Ke, X. Wang, X. Lu, X. Wu, **J. Kong**, Z. Ren, E. M. Akinoglu, M. Giersig, G. Zhou, J. Liu, K. Kempa, and J. Gao, "Plasmonic refraction-induced ultrahigh transparency of highly conducting metallic networks", *Laser and Photonics Reviews* 10, 465 (2016)
7. J. He, P. Shafer, T. Mion, V. Tra, Q. He, **J. Kong**, Y. Chuang, W. Yang, M. Graf, J. Lin, Y. Chu, E. Arenholz, and R. He, "Observation of a three-dimensional quasi-long-range electronic supermodulation in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  /  $\text{La}_{0.7}\text{Ca}_{0.3}\text{MnO}_3$  heterostructures", *Nature Communications* 7, 10852 (2016)

## Submitted

8. X. Wu, **J. Kong** (co-first), D. Broido, K. Kempa, "Tailoring electron scattering with plasmonic metamaterial structures", *Materials Today Physics*
9. **J. Kong**, X. Wu, M. J. Naughton, and K. Kempa, "Plasmonic multiple exciton generation", arXiv 1806.10259
10. A. Shvonski, **J. Kong**, K. Kempa, "Plasmon-polaron of the topological metallic surface states".

## Table of Contents

Chapter 1 Introduction and overview.....	1
1.1 The nonlocal problem .....	1
1.2 Review of classical models .....	4
References .....	7
Chapter 2 Effective film approach for nonlocal electromagnetic response of plasmonic nanostructures .....	8
2.1 Nonlocal problem in FDTD simulations, and the recent effective film approaches .....	9
2.2 Nonlocal to local mappings .....	13
2.3 Test of causality and sum rules .....	16
2.4 Application in producing accurate resonance shifts .....	21
2.5 Application in producing qualitatively correct field enhancement .....	24
2.6 Conclusion .....	26
Appendices .....	27
References .....	36
Chapter 3 Electron self-energy and scattering rate in a many-particle system .....	39
3.1 Quasi particle Green's function and its lifetime .....	40
3.2 Formulation in random phase approximation .....	42
3.3 A quick application in calculating electron mean free path in metals .....	46
3.4 A compact formula for $T=0$ limit in a Lorentz oscillator medium .....	51
Appendices .....	54
References .....	59
Chapter 4 Hot electron plasmon-protected solar cells and nonlocal treatment in the calculation .....	60
4.1 Introduction to hot electrons and the HELPP idea .....	61
4.2 Theoretical models for the scattering rate in the effective medium .....	64
4.3 Simulation, and calculation of the scattering rates .....	68

4.4 Application to photovoltaics .....	73
4.5 Conclusion .....	79
References .....	80
Chapter 5 Electron polar-optical-phonon scattering rates in heavily-doped semiconductors with plasmonic nanostructures embedded-in .....	83
5.1 Scattering rate with bare polar optical phonons .....	84
5.2 Electron-POP scattering with Thomas-Fermi screening .....	87
5.3 Electron scattering in metamaterials case one .....	91
5.4 Electron scattering in metamaterials case two .....	95
5.5 Comprehensive discussions taking electronic and phononic perturbations on equal footing and methods to distinguish them .....	98
5.6 Effects of adding metallic nanostructure in the “unscreened approximation”	105
Appendices .....	108
References .....	117
Chapter 6 Plasmon-enhanced multiple exciton generation .....	118
6.1 Introduction .....	119
6.2 Scattering rate calculation scheme revisit .....	121
6.3 Simulation and calculation results .....	123
6.4 Dependence on the gap size .....	127
6.5 Comments on fabrications .....	130
6.6 Conclusion .....	131
References .....	132

## CHAPTER 1 Introduction and Overview

### 1.1 The nonlocal problem

Description of propagation of disturbance in a medium, such as electromagnetic wave, must include the spatial and temporal variables,  $\mathbf{r}$  and  $t$ . Often, Fourier transform is performed, and the variables become  $\mathbf{q}$  the wavenumber, and  $\omega$  the frequency. Response of the medium to an external impulse can be parametrized by the so-called response function which depends on the two variables  $\mathbf{q}$  and  $\omega$ . Q-dependence herein is termed nonlocality. It basically says that, response of the medium at a particular point not only depends on the material properties at that very point, but also depends on properties of other points within the medium.

In fact, in describing a medium's electromagnetic response, the macroscopic Maxwell equations are solved, accompanied by the constitutive relations involving the electric permittivity (dielectric function)  $\epsilon(\mathbf{q},\omega)$  (and the magnetic permeability  $\mu(\mathbf{q},\omega)$ ). In many scenarios the  $\omega$ -dependence is sufficient for the description and the  $q$ -dependence is unimportant and neglected, which is called local approximation. For example, metal's optical response near its plasma frequency  $\omega_p$  can be very well described by the Drude model  $\epsilon = 1 - \omega_p^2/\omega^2$ , which is  $\omega$ -dependent only, i.e. local. In many other scenarios, the  $q$ -dependence is important or even dominant, for example the dielectric function of an electron gas at the low frequency limit, the Thomas-Fermi model  $\epsilon = 1 + q_{TF}^2/q^2$ , where  $q_{TF}$  is the Thomas-Fermi screening wavenumber. The only- $q$  dependence shows clearly the dominance of nonlocal effects, in which case the  $\omega$ -dependence doesn't matter at all. A phenomenological hybrid of the above two is the hydrodynamic approximation  $\epsilon = 1 -$

$\omega_p^2/(\omega^2 - \beta q^2)$ . It reduces correctly to the Drude model in the long wavelength limit ( $q \rightarrow 0$ ), and to the Thomas-Fermi model at the static limit ( $\omega \rightarrow 0$ ).

In general, the constitutive relation in a linear medium's electromagnetic response is expressed as

$$\mathbf{D}(\mathbf{r}, \omega) = \epsilon_0 \int d\mathbf{r}' \vec{\epsilon}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{E}(\mathbf{r}', \omega) \quad (1.1)$$

where  $\vec{\epsilon}$  is the dielectric tensor. There is also a similar relation between magnetic quantities. If the medium is *homogeneous* and *isotropic*, the relation is simplified to

$$\mathbf{D}(\mathbf{r}, \omega) = \epsilon_0 \int d\mathbf{r}' \epsilon(|\mathbf{r} - \mathbf{r}'|, \omega) \mathbf{E}(\mathbf{r}', \omega) \quad (1.2)$$

After Fourier transforming the space variables Eq. (1.2) reads

$$\mathbf{D}(\mathbf{q}, \omega) = \epsilon_0 \epsilon(\mathbf{q}, \omega) \mathbf{E}(\mathbf{q}, \omega) \quad (1.3)$$

Q-dependence in the dielectric function above originates from the  $\mathbf{r}'$  dependence in Eq. (1.1), and thus have the name nonlocality.

Eq. (1.3) basically is the photon's perspective, when it is incident on a medium. There's an equivalent electron's perspective, too, if we are to discuss one electron's behavior inside a medium. The zeroth order interaction when an external electron is injected into a many-particle system is the bare Coulomb interaction  $v_q$ . Taking into account screening by other particles, the effective interaction would in general become

$$v_{eff} = \frac{v_q}{\epsilon(\mathbf{q}, \omega)} \quad (1.4)$$

where  $\epsilon(\mathbf{q}, \omega)$  includes information about all characteristic modes that many-particle system can sustain.  $\omega$ -dependence here relates to the energy conservation in a scattering event, while the  $\mathbf{q}$ -dependence parametrizes the conservation of momenta.

This thesis aims at adequate description of the problem of  $q$ -dependence (nonlocality), in Eqs. (1.3) and (1.4), for effective media made of plasmonic nanostructures. Chapter 2 proposes a numerical scheme to extend the conventional finite-difference time-domain (FDTD) method, to effectively account for the nonlocality. Chapter 3 gives general derivations of the electron's scattering rate in a many-particle system, stressing the importance of the  $q$ -dependence in a few specific situations. Chapter 4 extends and applies the results of Chapter 3, proposing the so-called hot electron plasmonic protection (HELPP) scheme as a 3<sup>rd</sup> generation solar cell concept, involving simulations and calculations, as well as special treatment of the nonlocal effects in the response. Chapter 5 extends and applies the results of Chapter 3 to calculation of the electron - polar optical phonons (POP) scattering rates in heavily doped semiconductors, modified by embedding into them properly designed metallic nanostructures. The role of the response's nonlocality is again emphasized. Chapter 6 applies the Chapter 3 results, investigating the so-called plasmonic multiple exciton generation (PMEG) scheme, as a novel 3<sup>rd</sup> generation solar cell concept.

## 1.2 Review of classical models

This section reviews models of dielectric function (electric permittivity) which will be frequently visited and used throughout this thesis. There are numerous standard models based on classical electrodynamic theory or on many-particle quantum theory. The very first is the Drude model for simple metals [1, 2]

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)} \quad (1.5)$$

where

$$\omega_p = \sqrt{n_0 e^2 / \varepsilon_0 m} \quad (1.6)$$

is the plasma frequency determined by the free electron density  $n_0$ ,  $m$  is the free electron mass, and  $\gamma$  is the damping parameter. Taking into account polarization of the inner ion core [3, 4], equation (1.5) modifies to

$$\varepsilon(\omega) = \varepsilon_{core}(\omega) - \frac{\omega_p^2}{\omega(\omega + i\gamma)} \quad (1.7)$$

Drude model is local, i.e., it is the  $q = 0$  (long wavelength) limit. The other extreme, the static ( $\omega = 0$ ) limit is Thomas-Fermi approximation [2],

$$\varepsilon(q, \omega = 0) = 1 + \frac{q_{TF}^2}{q^2} \quad (1.8)$$

where  $q_{TF}$  is the Thomas-Fermi screening wavenumber. Next model is the hydrodynamic approximation (HDA), which treats the electron gas as a classical fluid that obeys the Bernoulli equation [5, 6, 7]

$$n \frac{\partial \mathbf{v}}{\partial t} + n(\mathbf{v} \cdot \nabla) \mathbf{v} + \gamma n \mathbf{v} = -\frac{ne}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) - \frac{\nabla p}{m} \quad (1.9)$$

leading to

$$\varepsilon(q, \omega) = \varepsilon_{core}(\omega) - \frac{\omega_p^2}{\omega(\omega + i\gamma) - \beta q^2} \quad (1.10)$$

where  $\beta$  comes from the electron pressure term in Eq. (1.9), and can be parametrized, or given quantum mechanically as

$$\beta = \frac{3}{5} v_F^2 \quad (1.11)$$

Many-particle quantum theory, within the so-called random-phase-approximation (RPA), leads to the Lindhard function [8, 9, 10]

$$\varepsilon(q, \omega) = 1 - \frac{4\pi e^2}{\varepsilon_b q^2} \Pi_{3D} \quad (1.12)$$

where

$$\Pi_{3D} = -\frac{mk_F}{2\pi^2 \hbar^2} \left[ 1 + \sum_{s=-1}^1 \frac{s}{2Q} (1 - k_s^2) \ln \left( \frac{k_s + 1}{k_s - 1} \right) \right] \quad (1.13)$$

and

$$k_s = \frac{sQ}{2} - \frac{\Omega}{2Q}, \quad Q = \frac{q}{k_F}, \quad \Omega = \frac{\hbar(\omega + i\gamma)}{E_F} \quad (1.14)$$

Lindhard function can also be extended as temperature dependent [9]

$$\Pi_{3D}^T = -\frac{mk_F}{2\pi^2 \hbar^2} \sum_{s=-1}^1 \frac{s}{2Q} \left[ \int_0^\infty \Delta_T(k) \ln \left( \frac{k_s + k}{k_s - k} \right) dk \right] \quad (1.15)$$

where

$$\Delta_T(k) = \frac{2\pi k \beta e^{-k^2/T}}{1 + \beta e^{-k^2/T}}, \quad \beta = \exp\left(\frac{\mu(T)}{T}\right) \quad (1.16)$$

It can be shown [11] that Lindhard function reduces to the Drude formula Eq. (1.5) in the long wavelength limit (small  $q$  but finite  $\omega$ ), and to the Thomas-Fermi result Eq. (1.8) at the static limit (vanishing  $\omega$  but finite  $q$ ). It provides also a  $q$ -dependent extension, which

for large  $\omega$  and small  $q$  is consistent with the HDA formula Eq. (1.10), and in particular it yields Eq. (1.11). Improvements to the RPA Lindhard formula include those by Hubbard, Mermin, Sjolander, etc. [see chapter 5 of Ref. 11], but are less convenient to use than Eq. (1.12).

## References:

- [1] P. Drude, Ann. Phys. 306, p566 (1900)
- [2] N. W. Ashcroft and N. D. Mermin, Solid State Physics, Brooks/Cole Cengage Learning (1976)
- [3] H. Ehrenreich and H. R. Philipp, Phys. Rev. 128, 1622 (1962)
- [4] D. Pines, Elementary Excitations in Solids, Benjamin, New York (1964)
- [5] A. D. Boardman, Electromagnetic Surface Modes, Wiley, New York (1982)
- [6] F. Forstmann and R. Gerhardt, Metal Optics Near the Plasma Frequency, Springer Tracts in Modern Physics, Vol. 109, Springer, Berlin, Heidelberg (1986)
- [7] C. Ciraci, J. B. Pendry and D. R. Smith, “Hydrodynamic Model for Plasmonics: A Macroscopic Approach to a Microscopic Problem”, ChemPhysChem 14, 1109 (2013)
- [8] J. Lindhard, Kgl. Danske Vidensk. Selesk. Mat. Fys. Medd. 28 (1954)
- [9] J. Cen, “Current driven plasma instabilities in layered solid state systems”, Ph.D. dissertation, Boston College (1991)
- [10] Y. Wang, E. W. Plummer and K. Kempa, “Foundations of Plasmonics”, Advances in Physics 60:5, 799-898 (2011)
- [11] G. D. Mahan, Many-Particle Physics, 2<sup>nd</sup> edition, Plenum Press, New York (1990)

## CHAPTER 2

### Effective Film Approach for Nonlocal Electromagnetic Response of Plasmonic Nanostructures

**Outline:** For plasmonic systems too small for classical, local simulations to be valid, but too large for *ab initio* calculations to be computationally feasible, we developed in this chapter a practical approach — a nonlocal-to-local mapping that enables the use of a modified local system to obtain the response due to nonlocal effects to lowest order, at the cost of higher structural complexity. In this approach, the nonlocal surface region of a metallic structure is mapped onto a local dielectric film, mathematically preserving the nonlocality of the entire system. The most significant feature of this approach is its full compatibility with conventional, highly efficient FDTD simulation codes. Our optimized choice of mapping is based on the Feibelman’s *d*-function formalism, and it produces an effective dielectric function of the local film that obeys all required sum rules, as well as the Kramers-Kronig causality relations. We demonstrate the power of this approach combined with an FDTD scheme, in a series of comparisons with experiments and *ab initio* DFT calculations from literature, for structures with dimensions from the sub-nano to micro range. This project was concluded as a publication: J. Kong, A. Shvonski and K. Kempa, “Nonlocal response with local optics”, Phys. Rev. B 97, 165423 (2018).

## 2.1. Nonlocal problem in FDTD simulations, and the recent effective film approaches

Conventional, numerical methods of studying plasmonic and nanophotonic systems, such as the finite-difference time-domain (FDTD) method [1], are based on solving macroscopic Maxwell's equations on a grid of points in the real space,

$$\begin{aligned}\nabla \cdot \mathbf{D} &= \rho_f & \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \cdot \mathbf{B} &= 0 & \nabla \times \mathbf{H} &= \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t}\end{aligned}\tag{2.1}$$

At each grid point two material parameters are defined: the electric permittivity  $\varepsilon$  and the magnetic permeability  $\mu$ . In the most efficient and popular FDTD codes (MEEP, CST and COMSOL) [2], these are functions of the propagating wave frequency ( $\omega$ ) only, and the important dependency on the wave vector  $q$  of the propagating wave (see Chapter 1 discussions) is typically neglected. This is equivalent to the assumption of locality: the response of the medium at a point  $r$  in space is assumed to depend only on the material response at the same point  $r$ . However, nonlocal effects might become important, in particular in the highly non-uniform regions of samples. Ways to treat them have been proposed, beginning with the pioneering work of Pekar [3], as well as in the follow up works in the 60-ties [4] and the 80-ties [5], and more recently [6-11].

There is no easy way to include nonlocality in the FDTD numerical codes. However, for simple interfaces this can be accomplished with the HDA [12]. Similarly, one can employ the quantum mechanical  $d$ -function formalism of Feibelman [13], which also provides a simple, analytical expression for the reflection coefficient, valid for small  $q$ ,

$$r_p = \frac{\varepsilon - p_m/p_v + i(\varepsilon - 1) \left[ (Q^2/p_v) d_{\perp} - p_m d_{\parallel} \right]}{\varepsilon + p_m/p_v - i(\varepsilon - 1) \left[ (Q^2/p_v) d_{\perp} + p_m d_{\parallel} \right]} \quad (2.2)$$

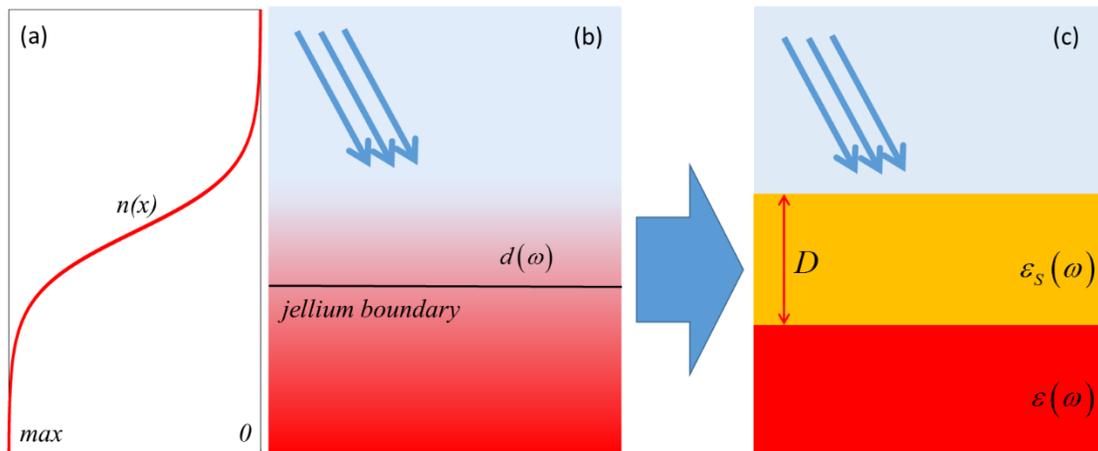
In addition to  $\varepsilon$  and  $\mu$ , there are two surface dielectric response functions ( $d$ -functions), which also depend only on  $\omega$  [14]. Moreover, by a proper choice of the reference frame (with zero at the jellium edge),  $d_{\parallel} = 0$ , and

$$d_{\perp}(\omega) = \frac{\int dz \cdot z \delta\rho(z, 0, \omega)}{\int dz \cdot \delta\rho(z, 0, \omega)} \quad (2.3)$$

Similar to  $\varepsilon$  and  $\mu$ , these functions can be calculated in an *ab initio* manner for a given metal [15-17], and then used in simulating the response of metal-containing structures via a Fresnel-like strategy. However, while the  $d$ -function formalism was very useful in understanding the basic physics of surface plasmons on simple metallic surfaces and simple large-scale metallic structures, treatment of more complex systems or ultra-small nanostructures cannot benefit directly from this approach. As pointed out above, implementing this approach into highly efficient FDTD codes is difficult. In principle, nonlocality can be fully included in more fundamental treatments of the material response, such as Density Functional Theory (DFT) [18]. However, due to limitations on computer power, this *ab initio* approach can handle only ultra-small structures, of hundreds of atoms at most (i.e. a few nm in diameter). Therefore, there is a need for a theory capable of handling nonlocal effects in nanostructures with intermediate sizes, at least for small  $q$ .

In an attempt to respond to this need, a recent work [19] proposed an idea of mapping the nonlocal surface region of a metal into a fictitious local dielectric effective film. The nonlocality of the surface region was modeled using the hydrodynamic approximation. This mapping was shown to be valid only for very small  $q$  and  $\omega$  [20]. In

an earlier work, we employed a similar mapping, but using instead the  $d$ -function formalism [21]. Both works produced mapping relations that were not unique and, therefore, additional constraints are needed. In Refs. [19-21] no effort was made to make the local dielectric functions of the fictitious films physical; these were just mapping functions. However, physicality (at least causality) of the dielectric functions is implied by the FDTD schemes, which supposed to benefit from the nonlocal to local mapping. In this chapter, we show how to develop a mapping process, which is fully causal (i.e. physical), as well as satisfies various sum rules. We show that our mapping works quantitatively in a much larger region of phase-space ( $q - \omega$ ), than the earlier mappings [19, 21]. We demonstrate, that with our mapping, conventional FDTD simulations acquire nonlocal capability, sufficient for a quantitative description of various nanoscopic systems.



**Fig. 2.1** (a) Diffusive character of electron density profile  $n(x)$  at a metal surface. (b) Gradient of the red color represents the diffusive profile of electrons confined by the jellium (the uniform positive charge, extending below the jellium boundary). (c) Local model of the metal surface, with a hypothetical film (orange color) of thickness  $D$  replacing the non-uniform electron density profile in the transition region. This figure is from: J. Kong, A. Shvonski and K. Kempa, “Nonlocal response with local optics”, Phys. Rev. B 97, 165423 (2018).

## 2.2 Nonlocal to local mappings

Consider an interface between metal and vacuum where huge electron density gradients exist. While, in the classical description, metal electrons are sharply confined to the bulk, quantum mechanics predicts the electron density near the metal surface to have a smooth profile,  $n(x)$ , as shown schematically in Figs. 2.1(a) and (b). The transition region thickness is of order  $D$ . Such a smooth electron density profile can be calculated self-consistently in the jellium model [22]. The electromagnetic response for a metallic surface with such a profile has been calculated [15-17], and allowed determination of the  $d$ -functions for many metals. The  $d$ -function formalism provides closed formulas for long-wavelength, nonlocal extensions of local Fresnel formulas; thus, it is a perfect platform to develop a nonlocal-to-local mapping.

In our earlier work [21] we based the mapping on the Feibelman formula [13] for the reflection coefficient of a p-polarized light [Eq. (2.2)]. By comparing this expression to the corresponding small- $D$  expansion of the local, Fresnel reflection coefficient for the model shown in Fig. 2.1(c), we obtained the effective dielectric function of the fictitious film [21]

$$\varepsilon_s(\omega) = \frac{D\varepsilon(\omega)}{[\varepsilon(\omega) - 1]d(\omega)} \quad (2.4)$$

where  $\varepsilon(\omega)$  is the Drude dielectric function of the metal and  $d(\omega)$  is the  $d$ -function. Even though  $\varepsilon_s$  obtained from Eq. (2.4) follows from a correct, formal mapping procedure, it is unphysical. This is due to an *ad-hoc* choice,  $\varepsilon_s(\omega) \propto D$ , employed in the derivation of Eq. (2.4). Clearly,  $\varepsilon_s$  diverges for large  $\omega$ , since  $\varepsilon(\omega) \rightarrow 1$ , and  $d(\omega) \propto \int_{-\infty}^{\infty} z\delta\rho(z)dz \rightarrow 0$

(the induced charge density  $\delta\rho$  vanishes for sufficiently large frequency). Moreover  $\varepsilon_s$  violates causality. To see this, we examine the corresponding susceptibility,

$$\chi_s(\omega) = \frac{\varepsilon_s(\omega) - 1}{4\pi} = \frac{D[1 + 4\pi\chi(\omega)]}{4\pi\chi(\omega)d(\omega)} \quad (2.5)$$

Causality requires that proper response functions (e.g. susceptibility), when analytically continued into the upper complex half-plane of variable  $z = \text{Re}(\omega) + i \text{Im}(\omega)$ , are analytic and vanish for  $|z| \rightarrow \infty$  and  $\text{Im}(\omega) > 0$  [23]. Since the Drude susceptibility,  $\chi(\omega)$ , and  $d(\omega)$  are causal, they satisfy these conditions, and so from Eq. (2.5) we have, for  $|z| \rightarrow \infty$  and  $\text{Im}(\omega) > 0$ ,

$$\chi_s(z) = \frac{D[1 + 4\pi\chi(z)]}{4\pi\chi(z)d(z)} \rightarrow \infty \quad (2.6)$$

This is a clear violation of the causality condition.

To eliminate this problem, we have investigated several alternative effective response functions within Feibelman's formalism, requiring that, in addition to the required nonlocal-to-local mapping, they lead to causal  $\varepsilon_s$ . We identified only one, physical  $\varepsilon_s$ , based on Liebsch's formula for the generalized reflection coefficient [24],

$$g(q, \omega) \approx \frac{\varepsilon(\omega) - 1}{\varepsilon(\omega) + 1 - 2q\varepsilon(\omega)d(\omega)} \quad (2.7)$$

The local Fresnel reflection coefficient for the model structure shown in Fig. 2.1(c) is [25, 26]

$$g(q, \omega) = \frac{\varepsilon'(q, \omega) - 1}{\varepsilon'(q, \omega) + 1} \quad (2.8)$$

where

$$\varepsilon'(q, \omega) = \varepsilon_s(\omega) \frac{\varepsilon(\omega) + \varepsilon_s(\omega) + [\varepsilon(\omega) - \varepsilon_s(\omega)] e^{-2qD}}{\varepsilon(\omega) + \varepsilon_s(\omega) - [\varepsilon(\omega) - \varepsilon_s(\omega)] e^{-2qD}} \quad (2.9)$$

It can be shown (see Appendix at the chapter end) that for  $Dq \ll 1$ , and with

$$\varepsilon_s(\omega) = \varepsilon(\omega) \left[ 1 - \frac{d(\omega)}{D} \right] \quad (2.10)$$

Eq. (2.8) reduces to Eq. (2.7). Another derivation of this result is also provided in the Appendix of this chapter.

### 2.3 Test of causality and sum rules

The effective dielectric function of the fictitious film, given by Eq. (2.10), is causal. The corresponding expression for the susceptibility is

$$\chi_s(\omega) = \frac{\varepsilon_s(\omega) - 1}{4\pi} = \frac{1}{4\pi} \left\{ [1 + 4\pi\chi(\omega)] \left[ 1 - \frac{d(\omega)}{D} \right] - 1 \right\} \quad (2.11)$$

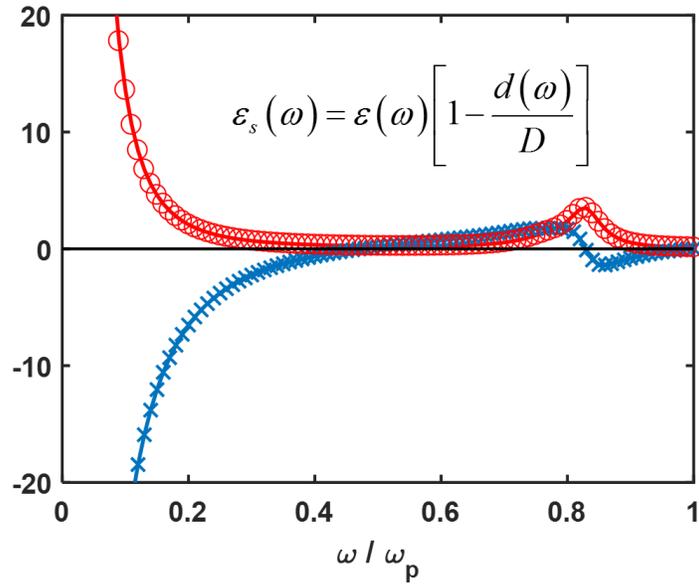
Clearly,  $\chi_s(\omega)$  vanishes for  $|z| \rightarrow \infty$  and  $\text{Im}(\omega) > 0$ . In addition, since  $\chi(\omega)$  and  $d(\omega)$  are causal,  $\chi_s(\omega)$  also is analytic in the upper complex half-plane of variable  $z$ . Thus, all the conditions for the causality of  $\chi_s(\omega)$  are satisfied and Eq. (2.10) represents a causal (i.e. physical) dielectric function.

To further illustrate that causality is satisfied, we checked that the Kramers-Kronig (KK) relations are satisfied numerically [23]. These follow the causality condition and have the following form,

$$\text{Re}[\varepsilon_s(\omega)] = 1 + \frac{2}{\pi} P \int_0^{\infty} \text{Im}[\varepsilon_s(\omega')] \frac{\omega'}{\omega'^2 - \omega^2} d\omega' \quad (2.12)$$

$$\text{Im}[\varepsilon_s(\omega)] = -\frac{2}{\pi} P \int_0^{\infty} \text{Re}[\varepsilon_s(\omega')] \frac{\omega}{\omega'^2 - \omega^2} d\omega' \quad (2.13)$$

where the integrals are evaluated along the real frequency axis, and  $P$  indicates the principle value integral. To check Eq. (2.12), we used Eq. (2.10) first to calculate  $\text{Im}[\varepsilon_s(\omega')]$ . Then, we numerically evaluated the integral on the right-hand side for a given  $\omega$ , and thus obtained the numerical values of  $\text{Re}[\varepsilon_s(\omega)]$ . Here, we used the following explicit expression for the Drude dielectric function:  $\varepsilon(\omega) = 1 - \omega_p^2 / \omega(\omega + i\gamma)$ , with  $\omega_p$  the plasma



**Fig. 2.2** Effective dielectric function  $\varepsilon_s$  of the fictitious film shown in Fig. 2.1(c), plotted versus  $\omega/\omega_p$ . Inset shows the analytical expression for  $\varepsilon_s$ , also listed as Eq. (2.10) in text. Solid lines represent the analytical evaluation of this equation:  $\text{Re}[\varepsilon_s]$  (blue) and  $\text{Im}[\varepsilon_s]$  (red). Symbols represent the corresponding numerical results based on Kramers-Kronig relations:  $\text{Re}[\varepsilon_s]$  (blue crosses) and  $\text{Im}[\varepsilon_s]$  (red circles). See text for details. This figure is from: J. Kong, A. Shvonski and K. Kempa, “Nonlocal response with local optics”, Phys. Rev. B 97, 165423 (2018).

frequency for potassium, the phonon-impurity scattering rate  $\gamma = 0.05\omega_p$ , and  $d(\omega)$  taken from Ref. [15] for potassium, and then fitted to a Lorentzian form:

$$d(\omega) = \sum_{s=1}^2 \frac{A_s}{\omega_s^2 - \omega(\omega + i\gamma_s)} \quad (2.14)$$

where  $A_1 = 0.367$ ,  $A_2 = 0.078$ ,  $\omega_1 = 0.829$ ,  $\omega_2 = 0.774$ ,  $\gamma_1 = 0.063$ , and  $\gamma_2 = 0.097$ . The numerically obtained values of  $\text{Re}[\varepsilon_s(\omega)]$  can be compared with the exact values of  $\text{Re}[\varepsilon_s(\omega)]$ , obtained directly from Eq. (2.10). Fig. 2.2 shows that, as expected, the two values are essentially identical. The same procedure was used to check the second KK relation (Eq. 2.13), confirming validity of this condition as well. This confirms that the dielectric function given by Eq. (2.10) is causal. Eq. (2.10) is the *main result* of this project.

In addition to satisfying KK relations, the dielectric function given by Eq. (2.10) has the characteristic Drude-like form with a reduced, effective plasma frequency  $\bar{\omega}_p$ , as expected for a metal film modeling the diffusive region, where the electron density is reduced. Also, we have checked numerically that the dielectric function satisfies the following sum rules [27]:

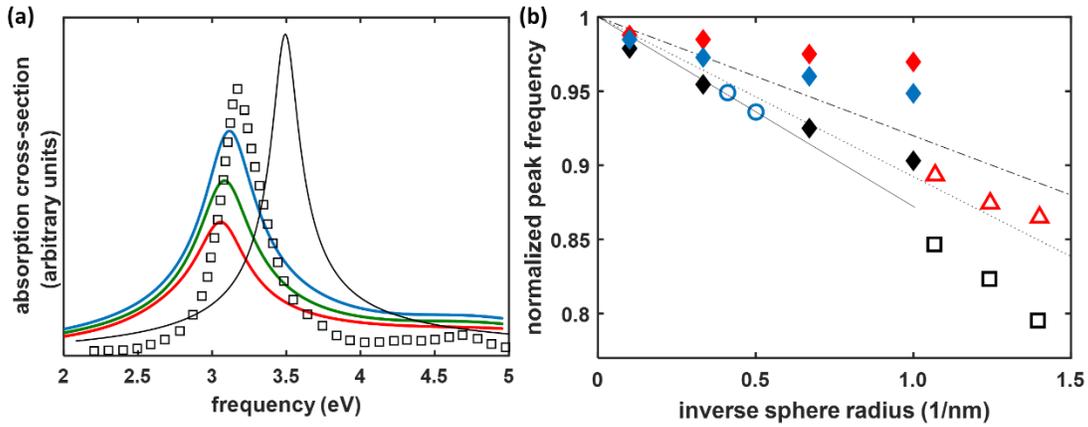
$$\int_0^{\infty} \omega \text{Im}[\varepsilon_s(\omega)] d\omega = \frac{\pi}{2} \bar{\omega}_p^2 \quad (2.15)$$

$$\int_0^{\infty} \omega \text{Im}[\varepsilon_s(\omega)]^{-1} d\omega = -\frac{\pi}{2} \bar{\omega}_p^2 \quad (2.16)$$

$$\int_0^{\infty} \text{Im}[\varepsilon_s(\omega)]^{-1} \frac{d\omega}{\omega} = -\frac{\pi}{2} \quad (2.17)$$

While the last sum rule (Eq. 2.17) is exactly and unconditionally satisfied, the first two (Eqs. 2.15 and 2.16) yield identical, reduced effective plasma frequency equal to

$\bar{\omega}_p = 0.7448\omega_p$ . This frequency is larger than the zero-crossing of  $\text{Re}[\varepsilon_s(\omega)]$  in Fig. 2.2, which occurs at  $\omega \approx 0.5\omega_p$ . In interpreting this discrepancy one should keep in mind that, while the sum rules are strictly valid only for a homogeneous electron gas, this is not the case in the surface region described by the fictitious film.



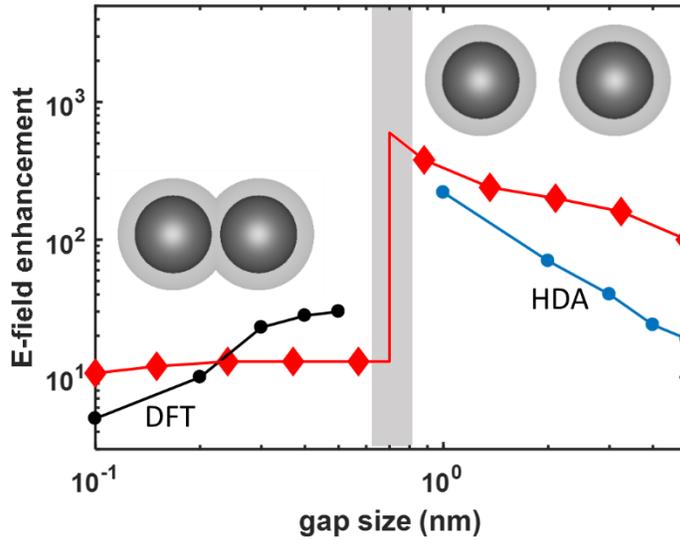
**Fig. 2.3** (a) Plasmonic response spectrum (absorption cross-section in arbitrary units) of a sodium sphere of radius 1.4 nm. Classical Mie resonance is shown as the black solid line. Open squares are *ab initio* DFT results adapted from Ref. [28]. Red/Green/Blue solid lines are from our effective film scheme with film thickness equal to 0.3 nm, 0.4 nm, 0.5 nm, respectively. (b) Normalized frequencies of plasmon resonances (peak positions) for sodium spheres versus the inverse sphere radius. Effective film approach results: red/black/blue diamonds with the corresponding film thickness 0.1 nm, 0.3 nm, 0.9 nm, respectively. Hollow circles, triangles and squares are experimental data adapted from Ref. [29]. Solid, dotted and dash-dotted lines are theoretical models discussed in Ref. [29]. This figure is from: J. Kong, A. Shvonski and K. Kempa, “Nonlocal response with local optics”, *Phys. Rev. B* 97, 165423 (2018).

## 2.4 Application 1: producing accurate resonance shifts in ultra-small nanoparticles

To demonstrate the strength and robustness of the mapping based on Eq. (2.10), in this section and the next we apply it to a few prototypical plasmonic systems, and compare with data available in the literature. First, we study the plasmonic resonance of sodium spheres/particles, which is known to be redshifted from its classical prediction, the Mie resonance given by  $\omega_{Mie} = \omega_p / \sqrt{3}$  [24]. In addition, this redshift increases with decreasing particle size. The Mie resonance is shown in Fig. 2.3(a) as the black solid line. The *ab initio* DFT calculation for the same system [28], yields black squares in Fig. 2.3(a) — a redshifted resonance. Implementing our effective film approach, i.e. coating the sphere with the film of a metal-like material having dielectric function given by Eq. (2.10), and performing the conventional FDTD simulation, yields color lines in Fig. 2.3(a), in a good agreement with the DFT result for a range of the film thickness. This insensitivity of our simulation outcome to the film thickness choice is an important feature, since the film thickness is an adjustable parameter, constrained loosely only by the size of the diffusive region at the metal surface.

In Fig. 2.3(b), we plot the normalized plasmon resonance frequencies for sodium spheres versus the inverse sphere radius. We compare here our effective film scheme (for 3 different film thicknesses), with experimental results and other theoretical studies reported in Ref. [29]. The results for the chosen film thickness of 0.3 nm agree best with those from experiments and advanced theories. This is consistent with the average thickness of the diffusive region  $D$ , as obtained from microscopic models [22], and shows that it can be used as an appropriate choice of the fictitious film thickness. The negative

slope of the curve in Fig. 2.3(b) is identical to the characteristic negative slope of surface plasmons on a planar sodium surface. This is a universal property for all alkali metals [30]; a one-to-one correlation between the slopes is facilitated by the simple surface plasmon “whispering gallery” relation  $2\pi R = \lambda$  [31], which reduces to  $q \approx 1/R$ .



**Fig. 2.4** Electric field enhancement half-way between two sodium spheres vs the inter-sphere gap. Each sphere has diameter  $D = 2.8$  nm, and thickness of the surface region (shown gray in the insets)  $d = 0.35$  nm. The frequency of the driving field is  $\hbar\omega = 2.75$  eV. Red diamonds represent our simulations, based on the effective (fictitious) film scheme. Black circles represent the *ab initio* DFT calculations [28], available only in the quantum regime. Blue circles show results of the hydrodynamic approximation [8], valid only in the classical regime. The insets show sketches of the spheres at an overlap (left) and well-spaced (right) cases, with diffusive regions marked gray. This figure is from: J. Kong, A. Shvonski and K. Kempa, “Nonlocal response with local optics”, Phys. Rev. B 97, 165423 (2018).

## 2.5 Application 2: producing qualitatively correct field enhancement in ultra-small nanoparticles

As another test of our scheme, we applied it to a previously studied case of the electric field enhancement in-between two, closely spaced sodium spheres, tuned out of the plasmonic (Mie) resonance. Each sphere has diameter  $D = 2.8$  nm, and thickness of the surface region (shown gray in the insets)  $d = 0.35$  nm. The frequency of the driving field is  $\hbar\omega = 2.75eV$ . Fig. 2.4 shows the electric field enhancement half-way between the two sodium spheres vs the inter-sphere gap size. Our scheme results are shown as red diamonds. The blue circles represent calculations based on the hydrodynamic approximation (HDA) taken from Ref. [8]. These are valid only for gap sizes much greater than the width of surface regions. The black circles represent *ab initio* DFT calculations taken from Ref. [28], which due to numerical difficulties are possible only for ultra-small spheres and gaps. Thus, *only our results* span the small to large gap sizes.

Our results show a sharp decrease in the field enhancement for gap sizes less than  $2d$ . This step-like decrease is physical, and the result of the presence of a surface region (of thickness  $\sim d$ ) of reduced electron density, present at metallic surfaces. As long as the two spheres are farther than  $2d$  apart, the mid-point (where field is calculated) lies in vacuum, and thus the electric field is only remotely (weakly) screened by the polarized spheres. However, once the surface regions begin to overlap for gap sizes  $< 2d$ , the mid-point probes this overlap region of (on average) high electron density, and thus a very strong shielding takes place. In our simulations, as discussed above, the surface region is represented by a constant, reduced electron density, step-like region (marked gray in the insets).

Consequently, the step-like discontinuity of the electric field enhancement at gap size of  $2d$ , is a result of this step-like model of the electron density. In more realistic calculations, in which the gradual surface electron density can be accommodated, a more gradual (but still quite sharp) transition of the field enhancement is expected. The most important conclusion here is, that our simulations, valid everywhere, agree (to within an order of magnitude) with the available calculation results from literature (DFT and HDA), in their respective regions of validity.

## 2.6 Conclusion

In conclusion, we proposed an effective film approach for plasmonic and nanophotonic studies, which accounts for nonlocal effects caused by the non-uniform region at metal surfaces. The approach requires adding a fictitious, but physical, thin film with a specified dielectric response. In our scheme this dielectric function is related to the  $d$ -function. Adapted into classical FDTD calculation schemes, this approach produces a moderate  $q$  extension of all nonlocal response characteristics, while avoiding the burdensome computations that are usually required by DFT for intermediate-size nano-structures. Specifically, we demonstrated that this method produces a quantitatively correct resonance shift and qualitatively correct field enhancement for small sodium spheres. While these examples have primarily focused on plasmonic systems with highly symmetric (e.g. spherical) geometries, we stress that this effective film approach may be applied to structures of any shape—the method simply requires coating all metal surfaces with a properly calculated thin film, and then performing the usual, local FDTD simulation. Thus, this effective film approach extends local FDTD calculations to a moderate  $q$  nonlocal regime.

## Appendices

(1) Derivation detail of equation (2.10):

In  $d$ -parameter formalism, surface response function accurate at small  $q$  (but non-retarded regime) of a semi-infinite Drude metal can be written as [24]

$$g(q, \omega) \approx \frac{\varepsilon(\omega) - 1}{\varepsilon(\omega) + 1 - 2q[\varepsilon(\omega)d_{\perp}(\omega) + d_{\parallel}(\omega)]} \quad (\text{A2.1})$$

On the other hand, in classical local optics, the surface response function of a semi-infinite Drude metal with an additional surface layer of thickness  $D$  and dielectric function  $\varepsilon_S(\omega)$ , sketched as in Fig. 2.1(c), is [24, 25, 26]

$$g(q, \omega) = \frac{\varepsilon'(q, \omega) - 1}{\varepsilon'(q, \omega) + 1} \quad (\text{A2.2})$$

$$\varepsilon'(q, \omega) = \varepsilon_S \frac{\varepsilon + \varepsilon_S + (\varepsilon - \varepsilon_S)e^{-2qD}}{\varepsilon + \varepsilon_S - (\varepsilon - \varepsilon_S)e^{-2qD}}$$

At optical wavelengths and a thickness of  $D$  about 1nm or less (typical length scale of the inhomogeneous, diffusive surface region), the following is true

$$qT \ll 1$$

Thus we can safely expand,

$$\begin{aligned} & \varepsilon'(q, \omega) \\ &= \varepsilon_S \left[ 1 + \frac{2(\varepsilon - \varepsilon_S)e^{-2qD}}{\varepsilon + \varepsilon_S - (\varepsilon - \varepsilon_S)e^{-2qD}} \right] \\ &\approx \varepsilon_S \left[ 1 + \frac{2(\varepsilon - \varepsilon_S)(1 - 2qD)}{\varepsilon + \varepsilon_S - (\varepsilon - \varepsilon_S)} \right] \\ &= \varepsilon_S + (\varepsilon - \varepsilon_S)(1 - 2qD) \end{aligned} \quad (\text{A2.3})$$

So

$$\begin{aligned}
g(q, \omega) &= \frac{\varepsilon_s + (\varepsilon - \varepsilon_s)(1 - 2qD) - 1}{\varepsilon_s + (\varepsilon - \varepsilon_s)(1 - 2qD) + 1} \\
&\approx \frac{\varepsilon_s + (\varepsilon - \varepsilon_s) - 1}{\varepsilon_s + \varepsilon - \varepsilon_s - 2qD\varepsilon + 2qD\varepsilon_s + 1} = \frac{\varepsilon - 1}{\varepsilon + 1 - 2qD(\varepsilon - \varepsilon_s)}
\end{aligned} \tag{A2.4}$$

Matching this local optics form (A2.4) with the nonlocal optics form (A2.1), we see, if

$$\begin{aligned}
2qD(\varepsilon - \varepsilon_s) &= 2q[\varepsilon(\omega)d_{\perp}(\omega) + d_{\parallel}(\omega)] \\
D(\varepsilon - \varepsilon_s) &= \varepsilon d_{\perp} + d_{\parallel}
\end{aligned}$$

the two results would be identical, for any (small)  $q$  value. That is to say, substituting

$$\varepsilon_{surface} = \varepsilon_{bulk} - \frac{1}{D}[\varepsilon_{bulk}d_{\perp} + d_{\parallel}] \tag{A2.5}$$

into local optics scheme would produce the same result as nonlocal optics.

(2) A different derivation of equation (2.10):

From another different way, the much more general results for a layered system [32], can actually reduce almost exactly to equation (2.10) if their more delicate considerations are dropped.

$$d_{\perp} - d_{\parallel} = \frac{\varepsilon_a}{\varepsilon_{im} - \varepsilon_a} \left[ (\varepsilon_{im} - \varepsilon_{if}) \left( \frac{1}{\varepsilon_{if}} - \frac{1}{\varepsilon_a} \right) d + \frac{i}{p_{lm}} \frac{1}{\varepsilon_{bm}} (\varepsilon_{im} - \varepsilon_{bm}) \right] \tag{A2.6}$$

Expressed in consistent symbols as in this chapter and neglect the hydrodynamic consideration,

$$d_{\perp} = \frac{1}{\varepsilon - 1} \left[ (\varepsilon - \varepsilon_s) \left( \frac{1}{\varepsilon_s} - 1 \right) D + \frac{i}{\infty} \frac{1}{\varepsilon_{bm}} (\varepsilon_{im} - \varepsilon_{bm}) \right]$$

Simplified to a quadratic equation,

$$\varepsilon_s^2 - \left[ (\varepsilon + 1) + (\varepsilon - 1) \frac{d_{\perp}}{D} \right] \varepsilon_s + \varepsilon = 0 \tag{A2.7}$$

It has two roots (with only one approximation in the solving procedure  $\varepsilon+1 \approx \varepsilon-1$ , ie  $|\varepsilon| \gg 1$ , which is true for Drude metal)

$$\varepsilon_s \approx \varepsilon + (\varepsilon - 1) \frac{d_{\perp}}{D} \quad \text{and} \quad \varepsilon_s \approx 1 \quad (\text{A2.8})$$

The first root in (A2.8) is actually equation (A2.5) above (noticing that the sign convention for  $d$ -function is opposite here), while the second is trivially vacuum. Again, response of the hypothesized film is in general metallic, with  $d$ -function corrections to some extent.

**(3)** Second order  $q$  extension for surface plasmon dispersion relation (quadratic extension to Liebsch self-energy ansatz) (this theory produces both the monopole and multipole branches as the same time):

This part discusses a phenomenological nonlocal theory for the surface plasmon dispersion relation, based on Liebsch's self-energy ansatz, to second order in  $q$ -dependence. The results were included in our publication [21].

Based on Ref. [33], Drude model is extended with Liebsch's self-energy ansatz,

$$\begin{aligned} \varepsilon(\omega, q) &= 1 - \frac{\omega_p^2}{\omega^2 + \Sigma(\omega, q)} \\ \Sigma(\omega, q) &\equiv qd(\omega)(\omega_p^2 - \omega^2) \end{aligned} \quad (\text{A2.9})$$

Plugging into the surface response function

$$g(\omega, q) = \frac{\varepsilon(\omega, q) - 1}{\varepsilon(\omega, q) + 1} = \frac{\omega_s^2}{\omega_s^2 - \omega^2 - \Sigma(\omega, q)} \quad (\text{A2.10})$$

it gives the dispersion condition

$$\omega_s^2 - \omega^2 - \Sigma(\omega, q) = 0 \quad (\text{A2.11})$$

which produces the initial negative slope at  $q \sim 0$ .

Besides the first order term introduced by Liebsch [33], we add a second order term to the self-energy (B, the pressure coefficient, is a constant to be found later)

$$\Sigma(\omega, q) \equiv qd(\omega)(\omega_p^2 - \omega^2) - B\omega_p^2 \cdot q^2 \quad (\text{A2.12})$$

Now the dispersion condition becomes

$$\omega_s^2 - \omega^2 - qd(\omega)(\omega_p^2 - \omega^2) + B\omega_p^2 \cdot q^2 = 0 \quad (\text{A2.13})$$

Write both  $\omega$  and d-function as full complex (damping  $\gamma \ll \omega$  is a tiny positive number)

$$\omega \rightarrow \omega + i\gamma \quad d(\omega) = d_R(\omega) + id_I(\omega)$$

Then the dispersion condition would be

$$\begin{aligned} 0 &= \omega_s^2 - \omega^2 - qd(\omega)(\omega_p^2 - \omega^2) + B\omega_p^2 \cdot q^2 \\ 0 &= \omega_s^2 - (\omega + i\gamma)^2 - q(d_R + id_I)\left[\omega_p^2 - (\omega + i\gamma)^2\right] + B\omega_p^2 \cdot q^2 \\ 0 &= \omega_s^2 - (\omega^2 - \gamma^2 + 2i\gamma\omega) - q(d_R + id_I)\left[\omega_p^2 - (\omega^2 - \gamma^2 + 2i\gamma\omega)\right] + B\omega_p^2 \cdot q^2 \\ 0 &= \omega_s^2 - \omega^2 + \gamma^2 - 2i\gamma\omega - q(d_R + id_I)\left[\omega_p^2 - \omega^2 + \gamma^2 - 2i\gamma\omega\right] + B\omega_p^2 \cdot q^2 \end{aligned} \quad (\text{A2.14})$$

Solving from real part = 0, with  $\gamma^2$  terms dropped

$$\begin{aligned} 0 &\approx \omega_s^2 - \omega^2 - qd_R(\omega_p^2 - \omega^2) - 2qd_I\gamma\omega + B\omega_p^2 \cdot q^2 \\ 0 &\approx B\omega_p^2 \cdot q^2 - q\left[d_R(\omega_p^2 - \omega^2) + 2d_I\gamma\omega\right] + (\omega_s^2 - \omega^2) \end{aligned}$$

Re-organize, to a more compact form

$$\begin{aligned} 0 &\approx B \cdot q^2 - q\left[d_R(1 - \Omega^2) + 2d_I\Gamma\Omega\right] + \left(\frac{1}{2} - \Omega^2\right) \\ \Omega &\equiv \frac{\omega}{\omega_p}, \quad \Gamma \equiv \frac{\gamma}{\omega_p} \quad \text{setting as } 0.05 \end{aligned} \quad (\text{A2.15})$$

For a specified  $\omega$ , we can look up  $d_R(\omega)$  and  $d_I(\omega)$  in the data, Fig. 2.5(a) extracted from [15], which were calculated from LDA, and then solve for q (generally there will be two q's for each  $\omega$ ). Raw results are shown as in Fig. 2.5(b) (negative q region should be

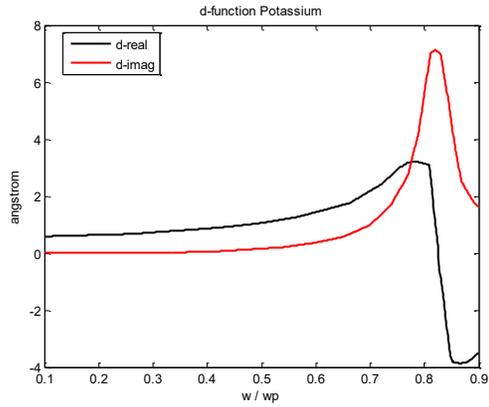
removed; imaginary  $q$  for smaller  $\omega$  lower than  $\sim 0.65$  already removed). Zooming-in to the small window where experimental data located are shown as in Fig. 2.5(c).

Fitting the lower curve (surface plasmon monopole branch) to experimental data and LDA calculation [33, Fig. 1] or [34] narrows the parameter  $B \sim 4.2$ .

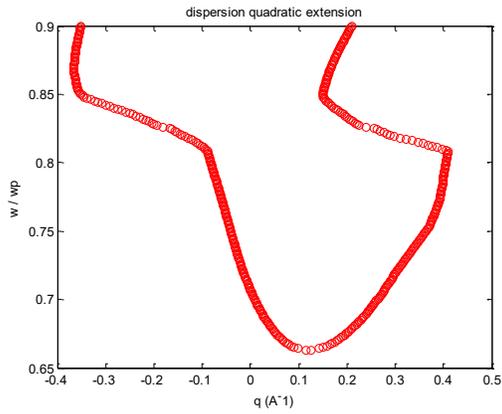
$$B\omega_p^2 \cdot q^2 \approx 4.2\omega_p^2 \cdot q^2 = 4 \times \frac{3}{5} V_F^2 \cdot q^2$$

While the lower curve (monopole branch) is fitted, the upper curve (multipole branch) is produced automatically, which mathematically is due to the pole structure of the Feibelman  $d$ -function. A final comparison plot is the graph shown in Fig. 2.6.

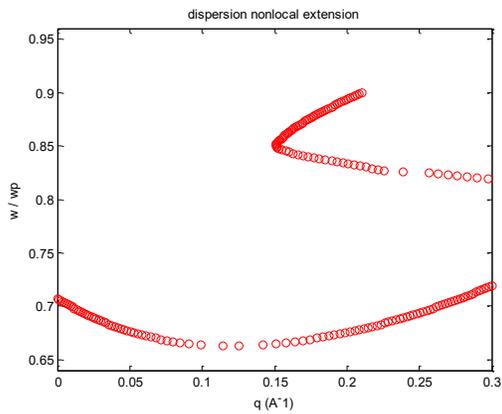
(a)



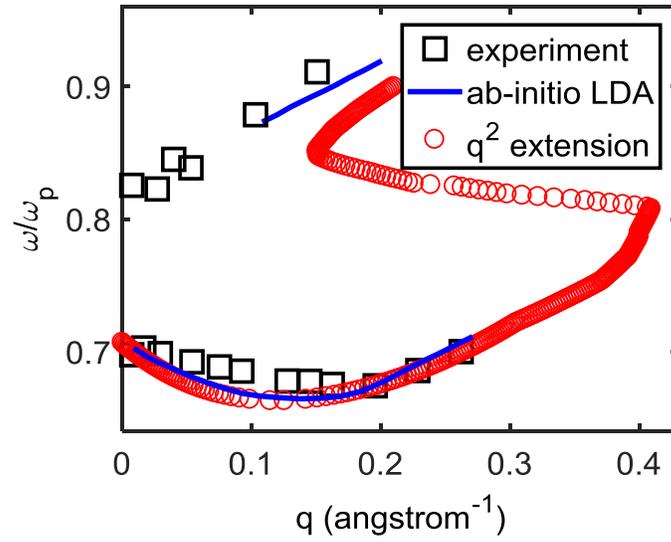
(b)



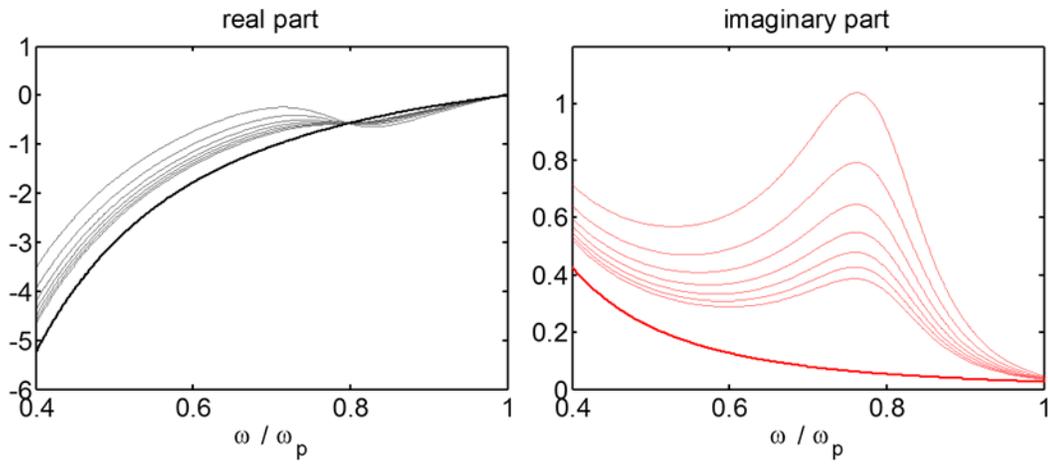
(c)



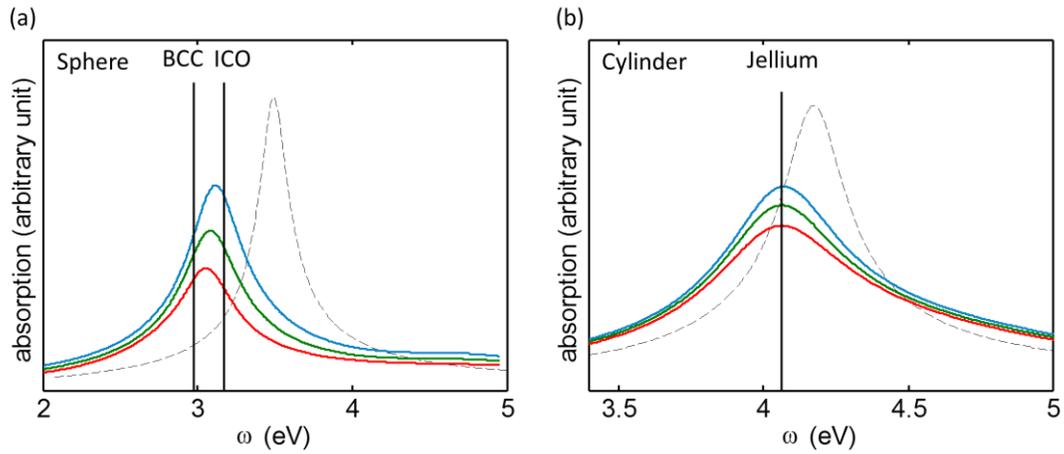
**Fig. 2.5** (a) Real and imaginary parts of d-function of potassium extracted from Ref. [15]; (b) Raw results after solving the surface plasmon dispersion condition (A2.15); (c) Zoom-in of (b) to the parameter window where experiment and computation data are located.



**Fig. 2.6** Fitting the quadratic extension of surface plasmon dispersion Eq. (A2.15) to ab-initio calculation [33] and experiment data [34] of the lower monopole branch. The upper multipole branch is produced automatically.



**Fig. 2.7** A family of curves of  $\epsilon_{\text{film}}$  as thickness varies, for potassium. Thick lines are Drude model. Thin lines are  $\epsilon_{\text{film}}$  for a series of film thicknesses ranging from 3 Angstroms to 9 Angstroms. It can be seen that Drude is the infinite-thickness limit of  $\epsilon_{\text{film}}$ , which is true according to Eq. (2.10).



**Fig. 2.8** Additional simulation results of resonance shift of simple metal structures. **(a)** Resonance shift of sodium spheres. ‘BCC’ and ‘ICO’ data are from [7]. Red, green and blue curves are our simulations of film thickness 3, 4 and 5 Angstroms. Dashed line is classical result calculated with Drude model. **(b)** Resonance shift of sodium cylinders. ‘Jellium’ data is from [28]. Red, green and blue curves are our simulations of film thickness 7, 8 and 9 Angstroms. Dashed line is classical result calculated with Drude model.

## References:

- [1] A. Taflove and S. C. Hagness, "Computational Electrodynamics: The Finite-Difference Time-Domain Method", Artech, Norwood MA (1995)
- [2] MIT Electromagnetic Equation Propagation, Massachusetts Institute of Technology, <http://ab-initio.mit.edu/wiki/index.php/Meep>; CST Microwave Studio, the Computer Simulation Technology AG, <https://www.cst.com/products/cstmws>; COMSOL Multiphysics Modeling Software, <https://www.comsol.com/>
- [3] S. I. Pekar, "Theory of electromagnetic waves in a crystal with excitons," J. Phys. Chem. Solids **5**, 11–22 (1958)
- [4] J. J. Hopfield and D. G. Thomas, "Theoretical and experimental effects of spatial dispersion on the optical properties of crystals," Phys. Rev. **132**, 563–572 (1963)
- [5] V. M. Agranovich and V. L. Ginzburg, Crystal Optics with Spatial Dispersion, and Excitons, Springer-Verlag, (1984)
- [6] N. A. Mortensen, S. Raza, M. Wubs, T. Sondergaard and S. I. Bozhevolnyi, "A generalized non-local optical response theory for plasmonic nanostructures", Nat. Comm. (2014), DOI: 10.1038/ncomms4809
- [7] T. V. Teperik, P. Nordlander, J. Aizpurua and A. G. Borisov, "Robust Subnanometric Plasmon Ruler by Rescaling of the Nonlocal Optical Response", Phys. Rev. Lett. **110**, 263901 (2013)
- [8] C. Ciraci, R. T. Hill, J. J. Mock, Y. Urzhumov, A. I. Fernandez-Dominguez, S. A. Maier, J. B. Pendry, A. Chilkoti and D. R. Smith, "Probing the Ultimate Limits of Plasmonic Enhancement", Science **337**, 1072 (2012)
- [9] F. Javier Garcia de Abajo, "Nonlocal Effects in the Plasmons of Strongly Interacting Nanoparticles, Dimers, and Waveguides", J. Phys. Chem. C **112**, 17983 (2008)
- [10] Y. Luo, R. Zhao and J. B. Pendry, "van der Waals interactions at the nanoscale: The effects of nonlocality", Proc. National Academy of Sciences **111**(52), 18422 (2014)
- [11] J. M. McMahon, S. K. Gray and G. C. Schatz, "Nonlocal Optical Response of Metal Nanostructures with Arbitrary Shape", Phys. Rev. Lett. **103**, 097403 (2009)
- [12] F. Forstmann and R. Gerhardts, "Metal Optics Near the Plasma Frequency", Springer Tracts in Modern Physics, Vol. **109**, Springer, Berlin, Heidelberg (1986)
- [13] P. J. Feibelman, "Surface Electromagnetic Fields", Prog. in Surf. Sci **12**, 287 (1982)
- [14] *footnote*: for jellium metals one of the  $d$ -functions can be made to vanish, by choosing the origin at the jellium edge.

- [15] A. Liebsch, "Dynamical screening at simple-metal surfaces", *Phys. Rev. B* **36**(14), 7378 (1987)
- [16] K. Kempa and W. L. Schaich, "Calculation of corrections to Fresnel optics from density response", *Phys. Rev. B* **34**(2), 547 (1986)
- [17] P. Apell, A. Ljungbert and S. Lundqvist, "Non-Local Optical Effects at Metal Surfaces", *Phys. Scripta* **30**, 367 (1984)
- [18] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", *Phys. Rev* **140**, A1133 (1965)
- [19] 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)
- [20] W. L. Schaich, "Comment on Surface Plasmons and Nonlocality: A Simple Model", *Phys. Rev. Lett.* **115**, 239401 (2015)
- [21] A. Shvonski, J. Kong and K. Kempa, "Nonlocal extensions of the electromagnetic response of plasmonic and metamaterial structures", *Phys. Rev. B* **95**, 045149 (2017)
- [22] N. D. Lang and W. Kohn, "Theory of Metal Surfaces: Charge Density and Surface Energy", *Phys. Rev. B* **1**(12), 4555 (1970)
- [23] F. W. Byron and R. W. Fuller, "Mathematics of Classical and Quantum Physics", Dover Publication, New York (1992)
- [24] A. Liebsch, "Electronic Excitations at Metal Surfaces", Springer Science+Business Media, New York (1997)
- [25] E. A. Stern and R. A. Ferrel, "Surface Plasma Oscillations of a Degenerate Electron Gas", *Phys. Rev* **120**, 130 (1960)
- [26] J. W. Gadzuk, "Coupled Surface-Plasmon Modes in Metal-Thin-Film-Vacuum Sandwiches", *Phys. Rev. B* **1**, 1267 (1970)
- [27] G. D. Mahan, "Many-Particle Physics", Plenum Press, New York (1990)
- [28] P. Zhang, J. Feist, A. Rubio, P. Garcia-Gonzalez, F. J. Garcia-Vidal, "Ab initio nanophotonics: The impact of atomic structure", *Phys. Rev. B* **90**, 161407(R) (2014)
- [29] T. Reiners, C. Ellert, M Schmidt and H. Haberland, "Size Dependence of the Optical Response of Spherical Sodium Clusters", *Phys. Rev. Lett.* **74**(9), 1558 (1995)
- [30] K-D Tsuei, E. W. Plummer and P. J. Feibelman, "Surface-Plasmon Dispersion in Simple Metals", *Phys. Rev. Lett.* **63**(20), 2256 (1989)

- [31] Y. Wang, E. W. Plummer and K. Kempa, "Foundation of Plasmonics", *Advances in Physics* **60** (2011) page799-898
- [32] K. Kempa and R. R. Gerhardts, "Nonlocal effects in ellipsometry of metallic films on metals", *Surface Science* 150, 157-172 (1985)
- [33] A. Liebsch, "Surface-plasmon dispersion and size dependence of Mie resonance", *Phys. Rev. B* 48, 11317 (1993)
- [34] K-D Tsuei, E. W. Plummer, A. Liebsch, K. Kempa and P. Bakshi, "Multipole plasmon modes at a metal surface", *Phys. Rev. Lett.* 64, 44 (1990)

## CHAPTER 3

### Electron Self-Energy and Scattering Rate in a Many-Particle System

**Outline:** This chapter contains theoretical derivations on electron lifetime in a many-particle system with the random phase approximation (RPA), and calculations of electron's mean-free-path in bulk aluminum as an example, which was included in the publication: Phys. Rev. B 95, 045149 (2017), Shvonski, Kong and Kempa, "Nonlocal extensions of the electromagnetic response of plasmonic and metamaterial structures". Next three chapters extend these general results further to applications in different specific many-particle systems, including conventional bulk materials and plasmonic nanostructured materials where proper nonlocal response  $\epsilon(\omega, \mathbf{q})$  are formulated.

### 3.1. Quasi particle Green's function and its lifetime

In the study of many-particle quantum systems [1-6], the concept of quasi particle was formulated,

$$\text{real particle} + \text{'coat' or 'cloud' of other particles} = \text{quasi particle}$$

Quasi particle has its new energy instead of the original free (bare) particle energy, and the difference is termed as the 'self-energy',

$$E_{quasi} - E_{bare} = E_{self}$$

The Green's function (also called propagator) is

$$G_{quasi\ particle}^+(\mathbf{k}, t_2 - t_1) = -iZ_{\mathbf{k}} \exp[-iE_{\mathbf{k}}'(t_2 - t_1)] \exp\left[-\frac{(t_2 - t_1)}{\tau_{\mathbf{k}}}\right] \quad (3.1)$$

where  $\mathbf{k}$  is the quantum number specifying the state,  $E_{\mathbf{k}}'$  is the quasi particle energy,  $\tau_{\mathbf{k}}$  is its lifetime in the exponential decay, and  $Z_{\mathbf{k}}$  is an occupation factor  $\leq 1$  for fermions. For a free particle in a non-interacting system, the Green's function would just be equation (3.1), but without the exponential decay factor. After Fourier transform from time domain to frequency domain equation (3.1) reads

$$G_{quasi\ particle}^+(\mathbf{k}, \omega) = \frac{Z_{\mathbf{k}}}{\omega - E_{\mathbf{k}}' + i\tau_{\mathbf{k}}^{-1}} \quad (3.2)$$

On the other hand, Dyson's equation expresses the propagator with the self-energy part  $\Sigma(\mathbf{k}, \omega)$  explicitly, [see Ref. 7 chapter 10 for a clear discussion]

$$G(\mathbf{k}, \omega) = \frac{1}{\omega - E_{\mathbf{k}} - \Sigma(\mathbf{k}, \omega) + i0^+} \quad (3.3)$$

The above two should have the same pole position,

$$\omega_{pole} = E_{\mathbf{k}}' - i\tau_{\mathbf{k}}^{-1} = E_{\mathbf{k}} + \text{Re}[\Sigma] + i \text{Im}[\Sigma] \quad (3.4)$$

and thus the scattering rate is

$$\gamma_k = \tau_k^{-1} = -\text{Im}[\Sigma] \quad (3.5)$$

This and the above equations are expressed with the convention  $\hbar=1$ . All following are in standard SI units.

### 3.2 Formulation in the random phase approximation

In general, the electron scatters with other electrons and ions in a many-particle system via the dressed combined interaction [4, 7, 8], which diagrammatically is

The diagrammatic equation (3.6) shows a thick horizontal line with a textured pattern on the left, followed by an equals sign. To the right of the equals sign are three terms: a wavy horizontal line, a plus sign, and a circular loop diagram. The loop diagram consists of two horizontal lines (one solid, one textured) connected at their ends by two vertical lines, forming a circle.

$$\text{Diagram} = \text{Wavy Line} + \text{Loop Diagram} \quad (3.6)$$

After translation, it reads

$$V_{eff}(q, \omega) = V_{el}(q, \omega) + V_{ph}(q, \omega) \quad (3.7)$$

where  $V_{el}(q, \omega)$  and  $V_{ph}(q, \omega)$  are the dressed Coulomb and phonon interactions respectively.

In the random phase approximation, and the point-ion, long wavelength approximation for ions, the screened Coulomb interaction is

$$V_{el}(q, \omega) = \frac{V_q}{\varepsilon(q, \omega)} \quad (3.8)$$

and the phonon dressed Frohlich interaction is

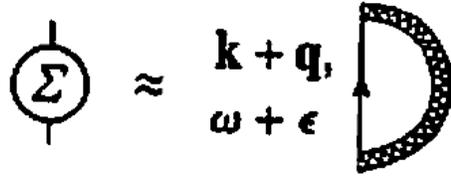
$$V_{ph}(q, \omega) = \frac{\Omega_q |g_q / \varepsilon(q, \omega)|^2}{\omega^2 - \Omega_q^2 / \varepsilon(q, \omega) + i\delta} \quad (3.9)$$

where the bare Coulomb interaction  $V_q$  and  $g_q$  are functions of  $q$  (length of  $\mathbf{q}$ ) alone,  $\Omega_q$  is the longitudinal optical phonon frequency (plasma frequency of the ionic “plasma”), and  $\delta$  is a positive infinitesimal. This well-known formula contains also the BCS interaction, with the Frohlich term becoming negative (attraction) for  $\{ \omega < \Omega_q / \sqrt{\varepsilon(q, \omega)} \text{ and real positive } \varepsilon(q, \omega) \}$ , leading to superconductivity for sufficiently large  $\Omega_q g_q^2$  to overcome the Coulomb interaction (repulsion).

The out-scattering (i.e., scattering away) rate of an electron from the state  $|\mathbf{k}\rangle$  to all possible other states  $|\mathbf{k}-\mathbf{q}\rangle$  in the electron gas, due to phonon or electron (or plasmon) excitations with wave vectors  $\mathbf{q}$ , is given by the imaginary part of the electron self-energy (in full SI units)

$$\gamma_{\mathbf{k}}^{\text{eff}} \approx -\frac{2}{\hbar} \text{Im}[\Sigma(k, E_k)] \quad (3.10)$$

which in turn is given by



$$\Sigma \approx \begin{matrix} \mathbf{k} + \mathbf{q}, \\ \omega + \epsilon \end{matrix} \quad (3.11)$$

In the Matsubara notation this diagram translates into

$$\Sigma(i\omega_n, \mathbf{k}) = \int \frac{d^3q}{(2\pi)^3} \frac{1}{\beta} \sum_m V_{\text{eff}}(i\omega_m, \mathbf{q}) G_0(i\omega_n - i\omega_m, \mathbf{k} - \mathbf{q}) \quad (3.12)$$

The free-electron propagator is

$$G_0(i\omega_n - i\omega_m, \mathbf{k} - \mathbf{q}) = \frac{1}{i\omega_n - i\omega_m - E_{\mathbf{k}-\mathbf{q}} + \mu} \quad (3.13)$$

and the Matsubara frequencies are  $\omega_m = 2m\pi / \beta$ ,  $\omega_n = (2n+1)\pi / \beta$ , with  $m$  and  $n$  integers ( $\omega_n$ , fermion, measured from chemical potential;  $\omega_m$ , boson, the energy transferred [9]). Using the Lehmann representation for  $V_{\text{eff}}$  [see for example Ref. 10]

$$V_{\text{eff}}(i\omega_m, \mathbf{q}) = -\frac{1}{\pi} \int \frac{\text{Im}[V_{\text{eff}}(\omega', \mathbf{q})]}{i\omega_m - \omega'} d\omega' \quad (3.14)$$

we get [see appendix (1) at the chapter end for detail steps]

$$\begin{aligned}
\Sigma(i\omega_n, k) &= -\frac{1}{\pi} \int \frac{d^3q}{(2\pi)^3} \int d\omega' \text{Im}[V_{\text{eff}}(\omega', q)] \frac{1}{\beta} \sum_m \frac{1}{i\omega_m - \omega'} \frac{1}{i\omega_n - i\omega_m - E_{k-q} + \mu} \\
&= +\frac{1}{\pi} \int \frac{d^3q}{(2\pi)^3} \int d\omega' \text{Im}[V_{\text{eff}}(\omega', q)] \frac{[n_B(\omega') + n_F(E_{k-q} - \mu)]}{i\omega_n + \omega' - E_{k-q} + \mu}
\end{aligned} \tag{3.15}$$

where the Bose-Einstein distribution function is

$$n_B(\omega) = \frac{1}{\exp(\beta\omega) - 1} \tag{3.16}$$

and the Fermi-Dirac distribution function is

$$n_F(\omega) = \frac{1}{\exp(\beta\omega) + 1} \tag{3.17}$$

Analytically continuing back, we get

$$\Sigma(\bar{\omega}, k) = +\frac{1}{\pi} \int d\omega' \int \frac{d^3q}{(2\pi)^3} \text{Im}[V_{\text{eff}}(\omega', q)] \frac{[n_B(\omega') + n_F(E_{k-q} - \mu)]}{\bar{\omega} + \omega' - E_{k-q} + \mu} \tag{3.18}$$

Integral using the Lehmann representation again, but backwards,

$$\begin{aligned}
\Sigma(\omega, k) &= +\int \frac{d^3q}{(2\pi)^3} V_{\text{eff}}(E_{k-q} - \mu - \omega, q) [n_B(E_{k-q} - \mu - \omega) + n_F(E_{k-q} - \mu)] \\
\text{Im}\Sigma(\omega, k) &= +\int \frac{d^3q}{(2\pi)^3} \text{Im}[V_{\text{eff}}(E_{k-q} - \mu - \omega, q)] [n_B(E_{k-q} - \mu - \omega) + n_F(E_{k-q} - \mu)]
\end{aligned} \tag{3.19}$$

$$\text{Im}\Sigma(E_k - \mu, k) = +\int \frac{d^3q}{(2\pi)^3} \text{Im}[V_{\text{eff}}(E_{k-q} - E_k, q)] [n_B(E_{k-q} - E_k) + n_F(E_{k-q} - \mu)] \tag{3.20}$$

For the in-scattering processes of state  $E_k$  from all other possible states  $E_{k+q}$ , self-energy could be expressed symmetrically as (could of course be derived from first principles)

$$\text{Im}\Sigma_{\text{in}}(E_k - \mu, k) = +\int \frac{d^3q}{(2\pi)^3} \text{Im}[V_{\text{eff}}(E_k - E_{k+q}, q)] [n_B(E_k - E_{k+q}) + n_F(E_k - \mu)] \tag{3.21}$$

Thus, the total scattering rate (net count), in principle, is

$$\gamma_k^{eff} \approx -\frac{2}{\hbar} \left\{ \text{Im}[\Sigma_{out}(k, E_k)] - \text{Im}[\Sigma_{in}(k, E_k)] \right\} \quad (3.22)$$

This is the final result of derivations of this part, and with the effective interaction given by equations (3.7)-(3.9), gives a complete account of the electron scattering in RPA, at arbitrary temperature ( $\beta = 1/kT$ ). [Sign convention: Imaginary part of self-energy, always negative; An excitation imposed on the medium means negative  $\omega$ , i.e.  $E_{k-q} - E_k \equiv \hbar\omega$ ; An excitation emitted by the medium means positive  $\omega$ ; same convention as Mahan, see page 480 in Ref. 4]

From this chapter and on, equation (3.20) is applied multiple times, together with nonlocal models for the response of bulk or plasmonic nanostructured materials  $\epsilon(\omega, q)$ , to calculate electrons' lifetime and thus scattering rates: (1) mean-free-path in bulk aluminum. (2) hot electron plasmon-protected solar cell. (3) tailoring electron-phonon scattering with plasmonic nanostructures. (4) plasmonic multiple exciton generation.

### 3.3 A quick application to calculate electron mean free path in metals

In this part, as a quick application of the general results obtained above, we calculate the electron mean-free-path in metals (e.g., aluminum), and compare with data in the literature.

In metals, since the free electron density is very high ( $\sim 10^{23} \text{cm}^{-3}$ ), Fermi energy  $E_F$  is much greater than the thermal fluctuation  $k_B T$  at room temperature, and thus there is no significant difference between zero-temperature RPA  $\epsilon$  (Lindhard function) or temperature dependent RPA  $\epsilon$  (see Chapter 1 for the formulation). In order to calculate the temperature dependent RPA  $\epsilon$ , accurate numerical values for chemical potential  $\mu$  of an electron gas of varied density are included (see Appendix at the end of this chapter). Calculation scheme is following Ref. [11, page76 and Appendix A therein].

For aluminum (data from [12]), free electron density  $n=18.1 \times 10^{22} \text{cm}^{-3}$  at 300K,  $E_F=11.7 \text{eV}$ ,  $T_F=13.6 \times 10^4 \text{K}$ , converted  $\omega_p=15.8 \text{eV}$ , ratio over  $E_F$  equals 1.35. Fig. 3.1(a) are color maps from Lindhard function (see Chapter 1), left shows  $\text{Im}[1/\epsilon]$ , right shows  $\text{Im}[\epsilon]$ , where damping parameter is set as 1% of Fermi energy. Both the single particle continuum (SPC) and the bulk plasma mode can be clearly seen, and slope of the SPC boundary is  $v_F$ . The second is the hydrodynamic model,

$$\epsilon(\omega, q) = 1 - \frac{\omega_p^2}{\omega^2 - \beta q^2}$$

where  $\beta=3v_F^2/5$  is used. Color maps are shown in Fig. 3.1(b); left shows  $\text{Im}[1/\epsilon]$ , right shows  $\text{Im}[\epsilon]$ , damping parameter 1% of Fermi energy. Only the bulk plasma mode is visible  $\omega = \sqrt{\omega_p^2 + \beta q^2}$ , and the asymptotic slope is as expected  $\sqrt{3/5}v_F$ , but it extends in the phase space without limit since there's no SPC to dissipate it (Landau damping). Map of

$\text{Im}[\epsilon]$  just shows the pole position, which is  $\omega = \sqrt{3/5}v_F \cdot q$ . The third is the Drude model, damping parameter is set as 1% of Fermi energy as well. Fig. 3.1(c) left shows  $\text{Im}[1/\epsilon]$  where a nondispersive plasma mode is visible  $\omega=\omega_p$ , right shows  $\text{Im}[\epsilon]$  which is most intensive as  $\omega$  approaches zero.

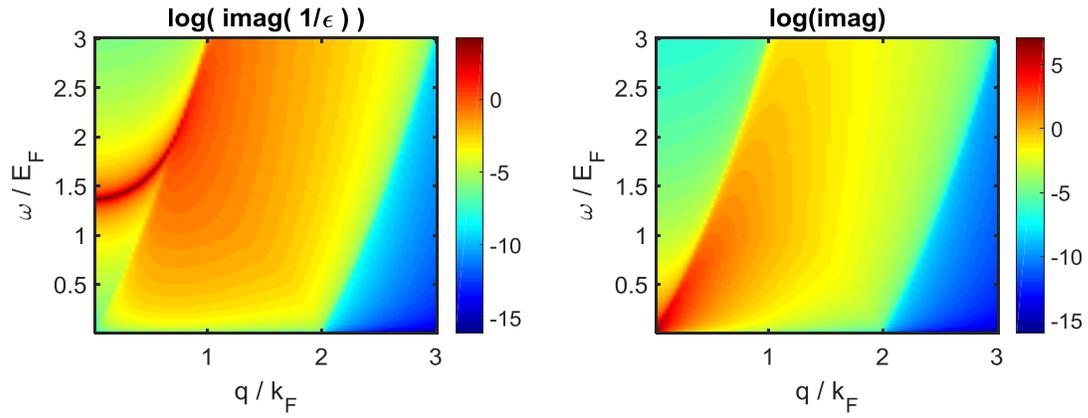
Fig. 3.2 is the result of this section, by plugging in Drude model, hydrodynamic model and Lindhard function as  $\epsilon(\omega, q)$  in Eq. (3.20). Horizontal axis is electron's energy measured from  $E_F$ , vertical axis is its mean-free-path in angstrom, converted from scattering rate by

$$mfp = \frac{v}{\gamma_{eff}} \quad (3.23)$$

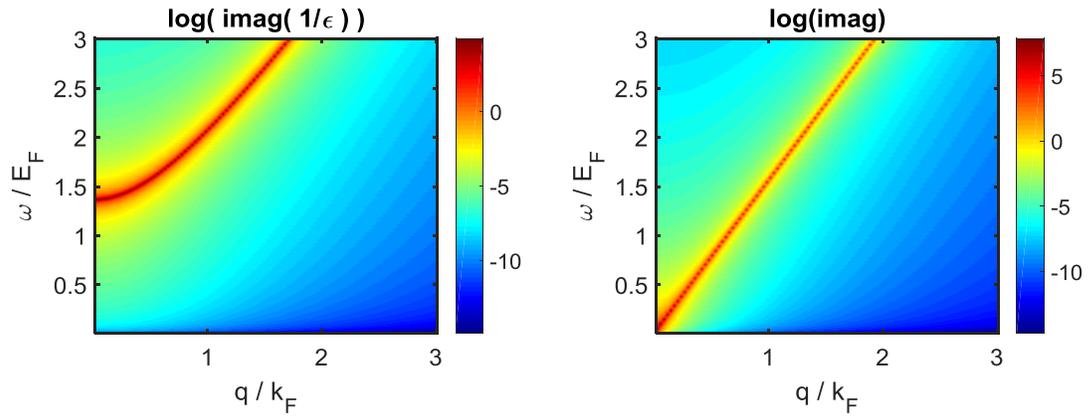
Yellow triangles are data from Ref. [13]. We see that nonlocality matters a lot, especially at the low energy region. Due to the lack of  $q$ -dependence in  $\epsilon$ , Drude model deviates far from the others. The reason why hydrodynamic model and Lindhard function give similar results is, their color maps of  $\text{Im}[1/\epsilon]$  are quantitatively similar at those points where the most probable transitions occur.

Temperature dependence in the RPA  $\epsilon$  is checked. In Fig. 3.3 left,  $T = 300$  K is applied, it is literally indistinguishable from the zero-temperature color map shown in Fig. 3.1; right  $T = 20000$  K, SPC boundary blurred, but no significant difference overall.

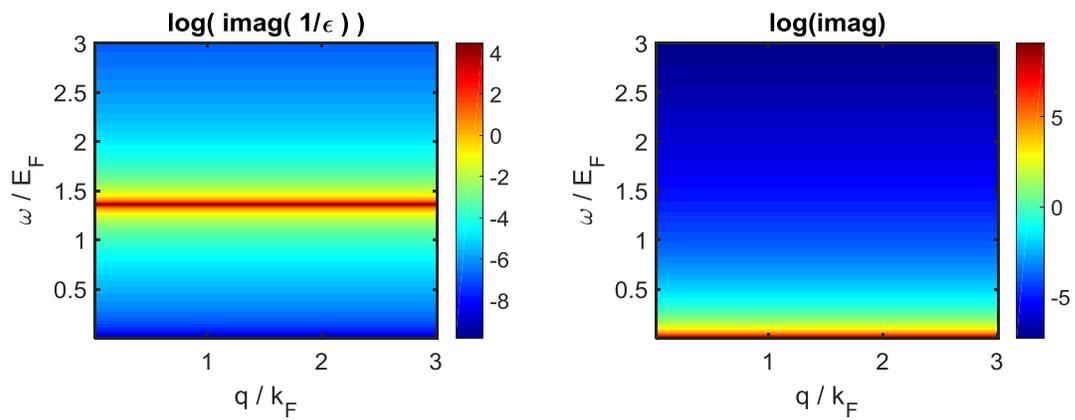
(a) Lindhard



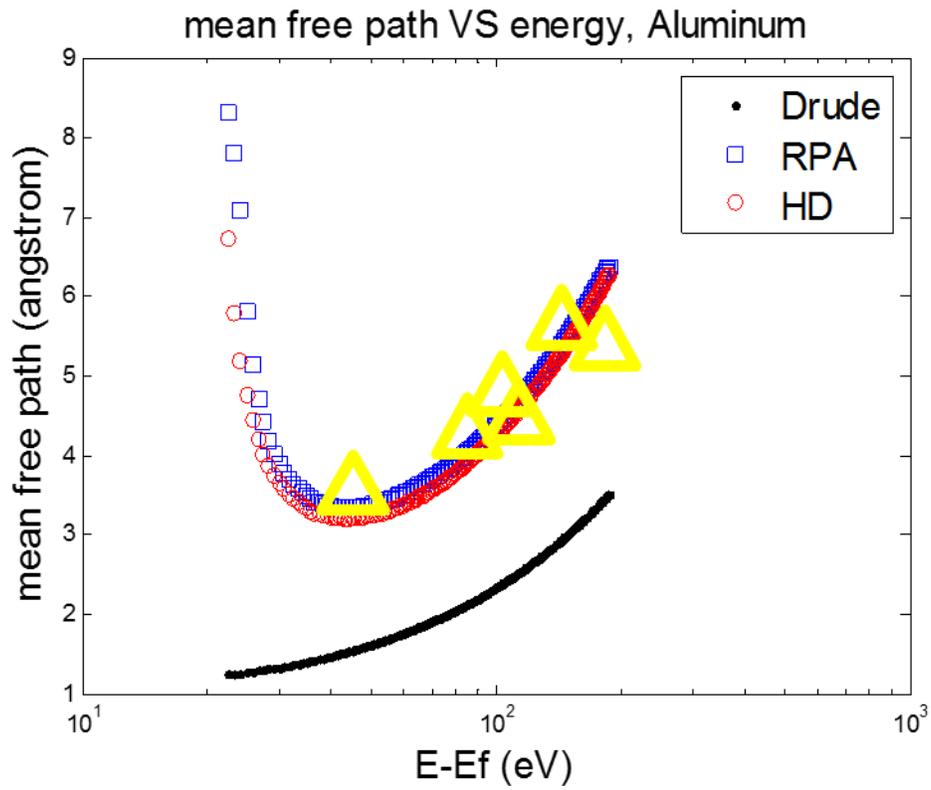
(b) Hydrodynamic



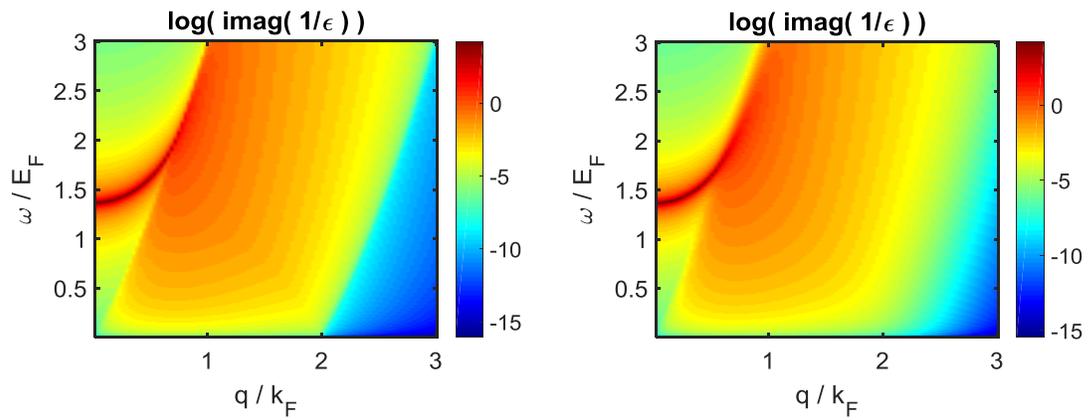
(c) Drude



**Fig. 3.1** Log-scale color maps of  $\text{Im}[1/\epsilon]$  and  $\text{Im}[\epsilon]$  of aluminum, of (a) Lindhard function; (b) Hydrodynamic model; (c) Drude model.



**Fig. 3.2** Numerical results for electron's mean free path in aluminum, calculated with different models of the dielectric function. Yellow triangles are experimental data from Ref. [13].



**Fig. 3.3** Log-scale color maps of  $\text{Im}[1/\epsilon]$  of temperature dependent Random Phase Approximation  $\epsilon$  of aluminum. Left,  $T = 300$  K; Right,  $T = 20000$  K.

### 3.4 A compact formula for T=0 limit in a Lorentz oscillator medium

In the situation of absolute zero temperature, or in metals that  $E_F$  is much higher than the thermal fluctuation  $k_B T$  at room temperature, a compact analytical result can be obtained, if the models for  $\varepsilon$  are as simplified, for example, as a Lorentz oscillator (see Ref. [14] for discussions on Lorentzian).

In the T approaches zero limit, we can use the following identity (there are alternative forms too, see appendix at the end of this chapter)

$$n_B(x) = -n_F(x) = -\Theta(-x) \quad (3.24)$$

valid for  $x \neq 0$ , and where  $\Theta(x)$  is the Heaviside step function. Then one obtains from equation (3.20) [same sign convention as in Ref. [4] Mahan (5.8.8)]

$$\gamma_k^{eff} = -\frac{2}{\hbar} \int \frac{d^3 q}{(2\pi)^3} V_q \operatorname{Im}[1/\varepsilon(E_{k-q} - E_k, q)] [\Theta(-E_{k-q} + E_k) - \Theta(-E_{k-q} + \mu)] \quad (3.25)$$

For an electron gas (or an equivalent electronic/plasmonic system), a typical local dielectric function (generalized Drude) could be written as a Lorentzian [see appendix at the end of this chapter for derivations of (3.27)]

$$\varepsilon(\omega) = \varepsilon_b + \frac{\omega_p^2}{\omega_r^2 - \omega(\omega + i\gamma)} \quad (3.26)$$

$$\operatorname{Im}\left[\frac{1}{\varepsilon}\right] \rightarrow \frac{\omega_1^2 - \omega_r^2}{\varepsilon_b \cdot 2\omega_1} (-\pi)\delta(\omega - \omega_1) \quad (3.27)$$

where  $\omega_1 = \sqrt{\omega_r^2 + \omega_p^2/\varepsilon_b}$ , corresponding to the longitudinal plasma mode in such medium.

Substituting it into equation (3.25), yields [15]

$$\gamma_k^{\text{eff}} = \frac{\sqrt{2E_k/m}}{2a^*} f\left(\frac{E_k}{E_0}\right) \Theta\left(\frac{E_k}{E_0} - 1\right) \quad (3.28)$$

where  $f(x) = \frac{2}{x} \ln(\sqrt{x} + \sqrt{x-1})$ ,  $E_0 = \hbar\sqrt{\omega_r^2 + \omega_p^2/\varepsilon_b}$ ,  $a^* = \frac{\varepsilon_b^2 E_0^2}{me^2 \omega_p^2} \cdot 4\pi\varepsilon_0$ . This compact

analytic result is used in the following chapters, whenever there is a well-defined plasma mode (e.g., localized surface plasmon mode) present.

If setting  $\omega_r$  to zero and  $\varepsilon_b$  to 1,  $\varepsilon(\omega)$  in (3.26) reduces to simple Drude model, and the result (3.28) reduces to Ref. [4, Mahan] (5.8.10) for the free electron gas. If nonlocality is further into account, for example the hydrodynamic extension,

$$\varepsilon(q, \omega) = \varepsilon_b + \frac{\omega_p^2}{\omega_r^2 - \omega(\omega + i\gamma) - \beta q^2}$$

or Lindhard function of  $\varepsilon$ , the integral has to be done numerically as we did in the previous section.

### **3.5 Other scattering mechanisms**

For ionized defect, the effective potential would be that of a point charge  $Q$ , screened by the other electrons. Analytical expression would be similar as in the electron-electron scattering part. To summarize, all kinds of scattering center could be interpreted as an external perturbation to the electron gas, i.e. an effective interaction, so could be calculated by this formalism. See [11, chapter 6] and [4] for a comprehensive review of all possible scattering mechanisms.

## Appendices:

(1) Matsubara frequencies summation, from Refs. [4] and [9]

$$\sum (i\omega_n) = -\frac{1}{\beta} \sum_{i\omega_m} \frac{1}{i\omega_n + i\omega_m - \varepsilon} \frac{1}{i\omega_m - \Omega} = \frac{n_F(\varepsilon) + n_B(\Omega)}{i\omega_n - \varepsilon + \Omega}$$

Using the above

$$\begin{aligned} & \frac{1}{\beta} \sum_m \frac{1}{i\omega_m - \omega'} \frac{1}{i\omega_n - i\omega_m - E_{k-q} + \mu} \\ &= \frac{-1}{\beta} \sum_m \frac{1}{-i\omega_m + \omega'} \frac{1}{i\omega_n - i\omega_m - E_{k-q} + \mu} \\ &= \frac{-1}{\beta} \sum_{m^*} \frac{1}{i\omega_{m^*} - (-\omega')} \frac{1}{i\omega_n + i\omega_{m^*} - (E_{k-q} - \mu)} \\ &= \frac{n_F(E_{k-q} - \mu) + n_B(-\omega')}{i\omega_n - (E_{k-q} - \mu) + (-\omega')} \end{aligned}$$

(2)  $\text{Im}[1/\varepsilon]$  for simple Drude model, referring to Ref. [4] p480 and p466

$$\lim_{q \rightarrow 0} \varepsilon(q, \omega) = 1 - \frac{\omega_p^2}{\omega^2}$$

$$\lim_{q \rightarrow 0} \frac{1}{\varepsilon(q, \omega)} = \frac{\omega^2}{\omega^2 - \omega_p^2} = 1 + \frac{\omega_p^2}{\omega^2 - \omega_p^2}$$

$$= 1 + \frac{\omega_p}{2} \left( \frac{1}{\omega - \omega_p} - \frac{1}{\omega + \omega_p} \right)$$

$$\stackrel{\text{retarded}}{\Rightarrow}_{\omega \rightarrow \omega + i\delta} 1 + \frac{\omega_p}{2} \left( \frac{1}{\omega - \omega_p + i\delta} - \frac{1}{\omega + \omega_p + i\delta} \right)$$

[Sokhotsky's formula, Dirac-delta-function]

$$\lim_{\varepsilon \rightarrow 0^+} \frac{1}{x \pm i\varepsilon} = P.V. \frac{1}{x} \mp i\pi\delta(x)$$

$$\lim_{\delta \rightarrow 0^+} \text{Im} \left[ \frac{1}{\varepsilon(q, \omega)} \right] = \frac{\omega_p}{2} (-\pi) \left[ \delta(\omega - \omega_p) - \delta(\omega + \omega_p) \right]$$

(3) generalized Drude model

$$\begin{aligned} \varepsilon(\omega) &= \varepsilon_b + \frac{\omega_p^2}{\omega_r^2 - \omega^2} \\ \frac{1}{\varepsilon(\omega)} &= \frac{\omega_r^2 - \omega^2}{\varepsilon_b(\omega_r^2 - \omega^2) + \omega_p^2} = \frac{\omega_r^2 - \omega^2 + \omega_p^2/\varepsilon_b - \omega_p^2/\varepsilon_b}{\varepsilon_b(\omega_r^2 - \omega^2) + \omega_p^2} \\ &= \frac{1}{\varepsilon_b} + \frac{\omega_p^2/\varepsilon_b^2}{\omega^2 - \omega_r^2 - \omega_p^2/\varepsilon_b} \\ &= \frac{1}{\varepsilon_b} + \frac{\omega_p^2}{\varepsilon_b^2} \frac{1}{2\omega_0} \left[ \frac{1}{\omega - \omega_0} - \frac{1}{\omega + \omega_0} \right], \quad \omega_0 \equiv \sqrt{\omega_r^2 + \omega_p^2/\varepsilon_b} \end{aligned}$$

Retarded extension, using result in (2)

$$\lim_{\delta \rightarrow 0^+} \text{Im} \left[ \frac{1}{\varepsilon(q, \omega)} \right] = \frac{\omega_p^2}{\varepsilon_b^2} \frac{1}{2\omega_0} (-\pi) \left[ \delta(\omega - \omega_0) - \delta(\omega + \omega_0) \right]$$

(4) composition Dirac delta function, referring to [16]

$$\delta(\alpha x) = |\alpha|^{-n} \delta(x), \text{ n is the dimension}$$

$$\delta[g(x)] = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|}$$

(5) distribution factor identities

$$N_0 \frac{(1 - f_0(\varepsilon + \hbar\omega_0))}{(1 - f_0(\varepsilon))} = N_0 + f_0(\varepsilon + \hbar\omega_0)$$

$$(N_0 + 1) \frac{(1 - f_0(\varepsilon - \hbar\omega_0))}{(1 - f_0(\varepsilon))} = N_0 + 1 - f_0(\varepsilon - \hbar\omega_0)$$

$$n_\eta(-x) = -\eta - n_\eta(x)$$

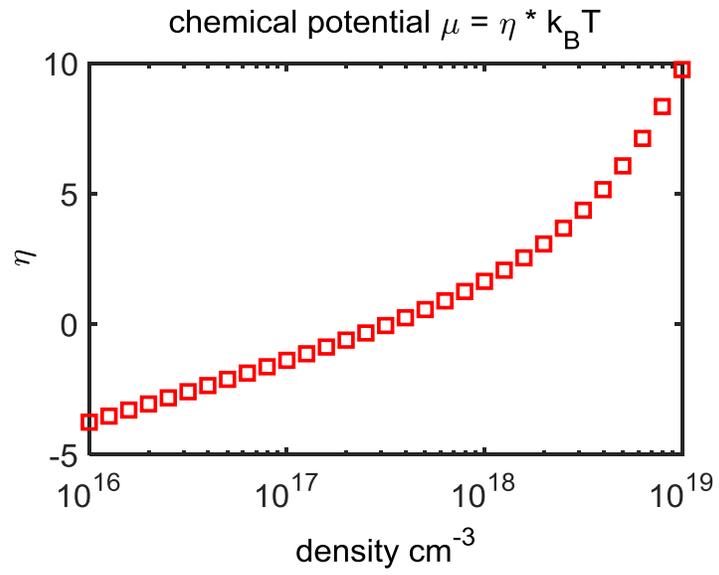
$$n_\eta(x) \equiv \frac{1}{e^{\beta x} - \eta}, \quad \eta = \pm 1 \text{ for Boson / Fermion}$$

$$\begin{aligned}
& N_0 \frac{(1 - f_0(\varepsilon))}{(1 - f_0(\varepsilon - \hbar\omega_0))} \\
&= N_0 \frac{1 - \frac{1}{\exp[\beta(\varepsilon)] + 1}}{1 - \frac{1}{\exp[\beta(\varepsilon - \hbar\omega_0)] + 1}} \\
&= N_0 \frac{(\exp[\beta(\varepsilon - \hbar\omega_0)] + 1)(\exp[\beta\varepsilon] + 1) - \exp[\beta(\varepsilon - \hbar\omega_0)] - 1}{(\exp[\beta(\varepsilon - \hbar\omega_0)] + 1)(\exp[\beta\varepsilon] + 1) - \exp[\beta(\varepsilon)] - 1} \\
&= N_0 \frac{\exp[\beta(2\varepsilon - \hbar\omega_0)] + \exp[\beta(\varepsilon)]}{\exp[\beta(2\varepsilon - \hbar\omega_0)] + \exp[\beta(\varepsilon - \hbar\omega_0)]} \\
&= \frac{1}{\exp[\beta(\hbar\omega_0)] - 1} \frac{\exp[\beta(\varepsilon - \hbar\omega_0)] + 1}{\exp[\beta(\varepsilon - \hbar\omega_0)] + \exp[\beta(-\hbar\omega_0)]} \\
&= \frac{1}{\exp[\beta(\hbar\omega_0)] - 1} \frac{\exp[\beta(\varepsilon)] + \exp[\beta(\hbar\omega_0)]}{\exp[\beta(\varepsilon)] + 1} \\
&= \frac{1}{\exp[\beta(\hbar\omega_0)] - 1} \frac{\exp[\beta(\varepsilon)] + 1 + \exp[\beta(\hbar\omega_0)] - 1}{\exp[\beta(\varepsilon)] + 1} \\
&= \frac{1}{\exp[\beta(\hbar\omega_0)] - 1} \left[ 1 + (\exp[\beta(\hbar\omega_0)] - 1) \cdot f_0(\varepsilon) \right] \\
&= N_0 + f_0(\varepsilon)
\end{aligned}$$

$$\begin{aligned}
& (N_0 + 1) \frac{(1 - f_0(\varepsilon))}{(1 - f_0(\varepsilon + \hbar\omega_0))} \\
&= \left( \frac{1}{\exp[\beta(\hbar\omega_0)] - 1} + 1 \right) \left[ 1 + (\exp[\beta(-\hbar\omega_0)] - 1) \cdot f_0(\varepsilon) \right] \\
&= N_0 + 1 + \frac{\exp[\beta(\hbar\omega_0)]}{\exp[\beta(\hbar\omega_0)] - 1} (\exp[\beta(-\hbar\omega_0)] - 1) f_0(\varepsilon) \\
&= N_0 + 1 - f_0(\varepsilon)
\end{aligned}$$

$$N_0 \frac{f_0(x)}{f_0(x+y)} = \frac{1}{e^{\beta y} - 1} \frac{e^{\beta(x+y)} + 1}{e^{\beta(x)} + 1} \stackrel{\text{skip}}{=} 1 + N_0 - f_0(x)$$

$$\begin{aligned} (N_0 + 1) \frac{f_0(x)}{f_0(x-y)} &= \frac{e^{\beta y}}{e^{\beta y} - 1} \frac{e^{\beta(x-y)} + 1}{e^{\beta(x)} + 1} = \frac{1}{e^{\beta y} - 1} \frac{e^{\beta(x)} + 1 - 1 + e^{\beta y}}{e^{\beta(x)} + 1} \\ &= \frac{1}{e^{\beta y} - 1} \left[ 1 + \frac{e^{\beta y} - 1}{e^{\beta(x)} + 1} \right] = N_0 + f_0(x) \end{aligned}$$



**Fig. 3.4** Chemical potential of free electron gas of varied densities. Calculation scheme is following Ref. [11].

## References:

- [1] A. A. Abrikosov, L. P. Gorkov and I. Y. Dzyaloshinskii, Quantum Field Theoretical Methods in Statistical Physics, Pergamon Press (1965)
- [2] D. Pines, Elementary Excitations in Solids, Perseus Books Publishing (1966)
- [3] A. L. Fetter and J. D. Walecka, Quantum Theory of Many-Particle Systems, McGraw-Hill Book Company (1971)
- [4] G. D. Mahan, Many-Particle Physics, Plenum Press, New York (1990)
- [5] X-G. Wen, Quantum Field Theory of Many-Body Systems, Oxford University Press (2004)
- [6] A. Altland and B. Simons, Condensed Matter Field Theory, 2<sup>nd</sup> edition, Cambridge University Press (2010)
- [7] R. Mattuck, A Guide to Feynman Diagrams in the Many-Body Problem, 2<sup>nd</sup> edition, Dover (1992)
- [8] T. D. Schultz, Quantum Field Theory and the Many-Body Problem, Gordon and Breach, New York (1964)
- [9] N. Nagaosa, Quantum Field Theory in Condensed Matter Physics, Springer (1999)
- [10] K. Kempa, P. Bakshi, J. Engelbrecht and Y. Zhou, Phys. Rev. B 61, 11083 (2000), Equation (3)
- [11] B. R. Nag, Electron Transport in Compound Semiconductors, Springer (1980)
- [12] N. W. Ashcroft and N. D. Mermin, Solid State Physics, Brooks/Cole Cengage Learning (1976)
- [13] Tracy, J. Vac. Sci. Technol. 11, 280 (1974)
- [14] H. Ibach and H. Luth, Solid-State Physics, 4<sup>th</sup> edition, Springer (2009)
- [15] K. Kempa, “Plasmonic protection of the hot-electron energy,” Phys. Status Solidi RRL 7(7), 465–468 (2013); (erratum) *ibid*, 7(12), 1112 (2013)
- [16] F. W. Byron and R. W. Fuller, Mathematics of Classical and Quantum Physics, Dover, New York (1992)

## CHAPTER 4

### Hot Electron Plasmon-Protected Solar Cells and Nonlocal Treatment in the Calculation

**Outline:** This chapter extends the general results of electron scattering rate in a medium obtained in chapter 3, combines with FDTD simulations and nonlocal modeling of the response, and applies them in a project proposing a concept for third generation solar cells, the hot electron plasmon-protected (HELPP) solar cell. In this cell, a thin-film, plasmonic metamaterial structure acts as both an efficient photon absorber in the visible frequency range and a plasmonic resonator in the IR range, the latter of which absorbs and protects against phonon emission the free energy of the hot electrons in an adjacent semiconductor junction. It is shown that in this structure, electron–plasmon scattering is much more efficient than electron–phonon scattering in cooling-off hot electrons, and the plasmon-stored energy is recoverable as an additional cell voltage. This project was concluded as a publication: J. Kong, A. H. Rose, C. Yang, X. Wu, J. M. Merlo, M. J. Burns, M. J. Naughton, and K. Kempa\*, “A hot electron plasmon-protected solar cell”, *Optics Express* 23, A1087 (2015).

## 4.1 Introduction to “hot” electron, the Shockley-Queisser limit and the HELPP idea

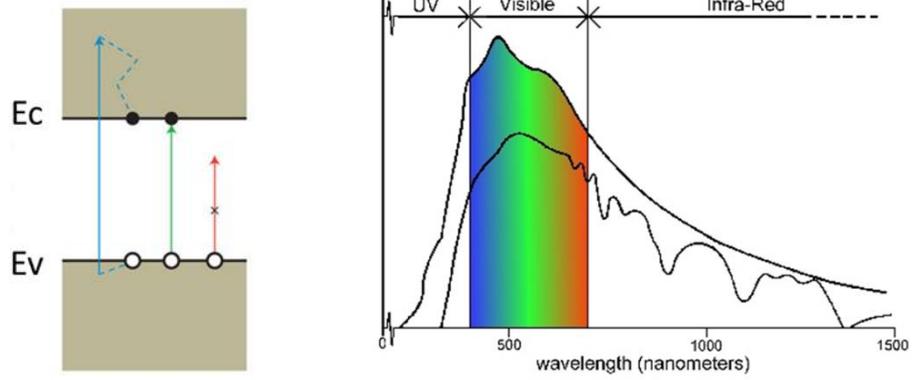
Charge carriers in semiconductors having excess energy (*i.e.* above the conduction band minimum for electrons and below the valence band maximum for holes) are referred to as “hot”. For example, electrons photo-excited from a valence band deep into a conduction band are considered hot electrons, as shown in Fig. 4.1.

Effects of hot electrons have been studied and utilized for more than half a century in a variety of electronic devices, from Gunn diodes to integrated circuits [1-11]. In conventional solar cells, hot electrons rapidly and irreversibly lose their excess energy (defined as the difference between the energies of their occupied states and that of the bottom of the conduction band, with a similar relationship for holes and the valence band) to phonons (heat) [2], which leads to the Shockley-Queisser limit for single junction cell efficiency [12]. The amount of the energy lost to heat in a conventional cell often exceeds that harvested in the form of electricity. For example, commercially available, high efficiency crystalline silicon solar cells convert 20-25% of absorbed sunlight into electricity, but more than 30% into heat via hot electrons. Many concepts have been proposed to harvest or convert this hot electron energy into usable form, but none have been experimentally verified or demonstrated to date [13].

Recently, a scheme was proposed to reduce the number of irreversibly dissipative electron-phonon scattering events in a solar cell, by providing an energy-dissipation channel into plasmons in an adjacent or embedded plasmonic structure [14]. In this scheme, the hot electron free energy remains reversibly “protected” in a collective electronic degree

of freedom. This hot electron plasmon protection (HELPP) mechanism, which relies on electron-plasmon scattering occurring on a time scale sufficiently smaller than phonon emission by either plasmons or hot electrons, was theoretically supported by a simple model calculation [14] [see chapter 3 section 4 for the analytic formula]. The effect is somewhat similar to the Förster resonance energy transfer (FRET) known in molecular physics, describing non-radiative (dipole-dipole) energy transfer between an excited electronic state of a donor molecule and single electron excitation of a nearby acceptor molecule [15]. It is distinct from schemes in which subgap photons are absorbed by a plasmonic metal followed by energy transfer to a semiconductor [16-18], as well as from up-conversion mechanisms [19]. HELPP can also be viewed as an inverse of plasmon resonance energy transfer (PRET) [20], where localized plasmon oscillations in metallic nanoparticles non-radiatively excite electrons in adjacent (a few nanometers away) molecules.

Here in this chapter, it is confirmed by detailed simulations that the HELPP mechanism can be achieved in an ultrathin metamaterial structure, which acts simultaneously as a plasmonic resonator in the infrared (IR) and a broad-band visible absorber. Electrically disconnecting the two could lead to a tandem hot electron photovoltaic device in which the absorber acts as a conventional solar cell, and the plasmonic resonator (which stores the hot electron energy) could be made a part of a Schottky cell, yielding an additional voltage originating from the hot electron energy.



**Fig. 4.1** Left panel: in conventional single-gap solar cells, “hot” electrons are generated when photons of energy greater than the gap are absorbed; Right panel shows the solar radiation spectrum.

## 4.2 Theoretical models for the scattering rate in the effective medium

According to the general derivations in Chapter 3, the scattering rate of an electron in a medium from a state  $E_k$  to all possible other states due to single particle and collective (plasmon) excitations, was given by Eq. (3.20) [21-23].  $V_{eff}(q, \omega)$  in Eq. (3.20) is the dressed combined interaction. It can be written as a simple (but exact) sum of the Coulomb and phonon (Fröhlich) terms (3.7-3.9) [23]. It is in the random phase approximation for electrons (1<sup>st</sup> term), and the point-ion, long wavelength approximation for ions (2<sup>nd</sup> term).

The simple, additive form of Eq. (3.7) allows one to write the total scattering rate as  $\gamma_k = \gamma_{el-pl} + \gamma_{el-ph}$ , where  $\gamma_{el-pl}$  is given by Eq. (3.20) with  $V_{eff}(q, \omega)$  given by only the first (Coulomb) term in Eq. (3.7), and  $\gamma_{el-ph}$  given by only the second (Fröhlich) term. Here, the subscripts *el*, *ph*, and *pl* refer to electron, phonon and plasmon, respectively. Since  $\gamma_{el-ph}$  for systems of interest here has been studied in detail elsewhere [24, 25], we focus on calculating only  $\gamma_{el-pl}$ .

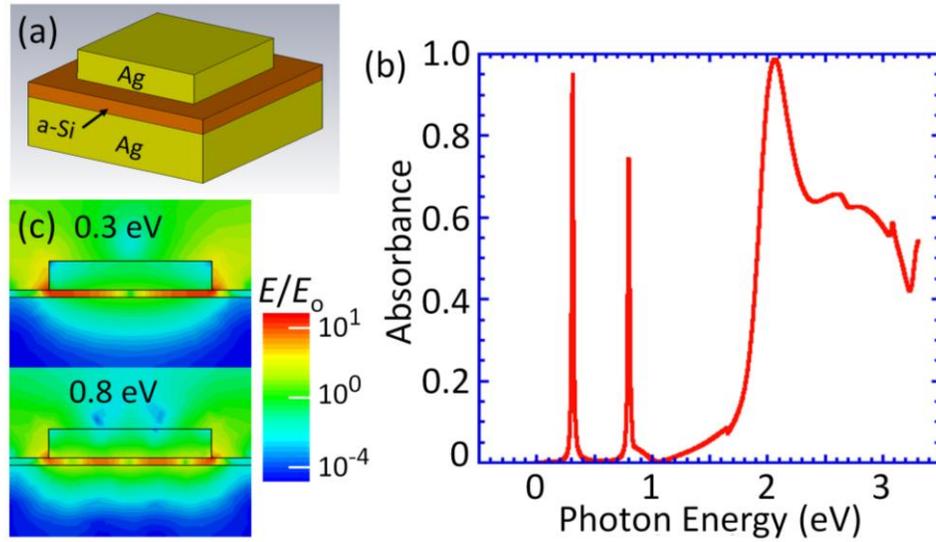
Clearly, this calculation requires knowledge of the effective dielectric function of a given structure. In Ref. [14], a simple model structure was considered in which a semiconductor film was assumed to be coupled so strongly to a metallic, plasmonic resonator that the system could be described as an “effective medium” with an effective relative dielectric function of the general form [see chapters 3 and 5 for discussions on response function of typical  $\varepsilon$ ]:

$$\varepsilon(\omega) = \varepsilon_b + \sum_{m=1}^M \frac{\omega_{pm}^2}{\omega_{rm}^2 - \omega(\omega + i\bar{\gamma}_{el-ph}) - \Sigma(q)} \quad (4.1)$$

where it was assumed for simplicity that  $M = 1$  (only one, dominant localized plasmon resonance), and the self-energy  $\Sigma(q)$  and the electron-phonon scattering rate in the metal vanish ( $\Sigma(q)=0$  and  $\bar{\gamma}_{el-ph} \rightarrow 0^+$ ). Here, each plasmon resonance, indexed by mode  $m$ , is characterized by a modal plasma frequency  $\omega_{pm}$  and a localized plasmon modal resonance frequency  $\omega_{rm}$  within the Drude-Lorentz model and  $\epsilon_b$  is the core or bound electron dielectric constant. When Eqs. (3.7-3.9) are inserted into Eq. (3.20), calculations [14] showed that plasmon emission by a hot electron in a semiconductor could occur at the rate of  $\gamma_{el-pl} \approx 10^{14} \text{ s}^{-1}$ , which is faster than the typical phonon emission rate in the semiconductor,  $\gamma_{el-ph} \approx 10^{13} \text{ s}^{-1}$ . In such a case, the hot electron “cools down” to the bottom of the conduction band by emitting plasmons in the collector/plasmon resonator, rather than by emitting phonons in the semiconductor medium. However, this is a necessary but not sufficient condition for the HELPP effect to occur. In addition, one must arrange for the plasmons in the metal to decay into phonons sufficiently slowly. This indeed is expected to be the case, since typical plasmon-phonon scattering rates in metals are  $\gamma_{pl-ph} \approx \bar{\gamma}_{el-ph} \approx 10^{13} \text{ s}^{-1}$ .

It is clear that HELPP depends on strong interactions between the semiconductor, where the electron transitions occur, and the plasmonic resonator. This depends on, among other things, their relative proximity. PRET studies [20] have shown that one needs distances of the order of 10 nm to achieve strong coupling. Therefore, the semiconductor film, e.g. the active region of a solar cell, must be very thin, on a similar 10 nm scale. Solar cells of such thickness are indeed feasible; highly-efficient light trapping in active media of the order of 10 nm thickness has been demonstrated with plasmonic metamaterial

schemes [26-28]. Here, we propose a solar cell based on such a scheme, modified to assure also a contribution from HELPP.



**Fig. 4.2** (a) Schematic of a unit cell ( $300 \text{ nm} \times 300 \text{ nm}$ ) of a square array of plasmonic resonators in a HELPP solar cell structure. The thickness of each layer is: macroscopically-thick Ag bottom electrode,  $10 \text{ nm}$   $a\text{-Si}$  absorber, and  $40 \text{ nm}$  Ag top plasmonic resonator ( $200 \text{ nm} \times 200 \text{ nm}$ ). (b) Simulated absorbance spectrum of the structure, showing two plasmon resonances in the infrared. For the purposes of this paper, this scale for the resonance peaks near  $0.3$  and  $0.8 \text{ eV}$  refers to energies above (below) a semiconductor absorber's conduction (valence) band edge. (c) Electric field distribution (scaled to the incident field magnitude) in a cross-section of the unit cell (through the unit center, parallel to the unit side), at the frequencies of the main plasmon resonance peaks at  $0.3$  and  $0.8 \text{ eV}$ . Field enhancements up to  $E/E_0 \sim 30$  (*i.e.* intensity gain  $\sim 30^2 \sim 1,000$ ) are realized within the photovoltaic absorber. This figure is from: J. Kong, A. H. Rose, C. Yang, X. Wu, J. M. Merlo, M. J. Burns, M. J. Naughton, and K. Kempa\*, "A hot electron plasmon-protected solar cell", *Optics Express* 23, A1087 (2015).

### 4.3. Simulation, and calculation of the scattering rates

We consider first a version of the solar cell which enables highly efficient light absorption and HELPP action, but not yet having optimized electronic recovery of the hot electron energy. A schematic of this basic HELPP structure is shown in Fig. 4.2(a), which is a unit cell of a periodic square array with lattice constant 300 nm. It consists of a metallic (Ag) back electrode, an active semiconductor layer (ultrathin *a*-Si *p-i-n* junction) of total thickness 10 nm, and a top, square metallic (Ag) plasmonic resonator with thickness 40 nm and side width 200 nm.

The absorption spectrum of the structure, calculated by using the FDTD method [29-32] and intrinsic *a*-Si as the absorber, is shown in Fig. 4.2(b). It has a broad band in the visible frequency range that assures high photon absorption. This is the signature of the structure's metamaterial action, similar to that reported in Refs. [26-28]. The spectrum has also two sharp absorbance peaks in the infrared, near 0.3 eV and 0.8 eV; these are the plasmon resonances that enable the HELPP effect. The corresponding electric field distribution, calculated at the frequencies of those main plasmonic peaks, is shown in Fig. 4.2(c). Strong concentration of the field (intensity enhancement over incident value  $(E/E_0)^2 \sim 30^2 \sim 10^3$ ) within the semiconductor film is clearly visible. These plasmon resonances are designed to be excited by hot electrons, and the excitation rate can be calculated from Eq. (3.20). However, photons from the IR tail of the solar spectrum could also directly excite these resonances, further improving the efficiency of a tandem configuration discussed below. We have also simulated the effect of positioning a point dipole directly in the absorber layer, and found that, as expected, this also excites the plasmon resonances in the IR, with comparable efficiency.

To calculate the various scattering rates in this structure, we need to obtain its aforementioned effective dielectric function. To achieve this, we first model the system as a three-layer problem: a semi-infinite Ag substrate, an effective coating of total thickness  $d$ , and vacuum. Such a model has been shown [33] to be a good approximation for the optical response of thin discontinuous films, provided that the effective dielectric function of the film is extended into the nonlocal domain, as discussed and applied below. The effective coating, which consists of the absorber and Ag nanostructures of Fig. 4.2(a), is described in the IR via a single effective dielectric function  $\varepsilon(\omega)$  of the form of Eq. (4.1) with two Lorentzian terms ( $M = 2$ ), associated with the two HELPP-enabling plasmon resonances of Fig. 4.2(b). The parameters of this Lorentzian expansion are chosen so that the reflectance  $R$  of the model three-layer system [34] closely fits that obtained from the simulation in Fig. 4.2(b).

$$R = |r|^2 = \left| \frac{(1 - \sqrt{\varepsilon}) - (1 + \sqrt{\varepsilon}) \exp(i4\pi nd / \lambda)}{(1 + \sqrt{\varepsilon}) + (1 - \sqrt{\varepsilon}) \exp(i4\pi nd / \lambda)} \right|^2 \quad (4.2)$$

Note that the simulated reflectance  $R = 1 - A$  can be immediately obtained from the absorbance  $A$  in Fig. 4.2(b). The resulting fit is shown in Fig. 4.3(a), comparing well with calculations from the model.

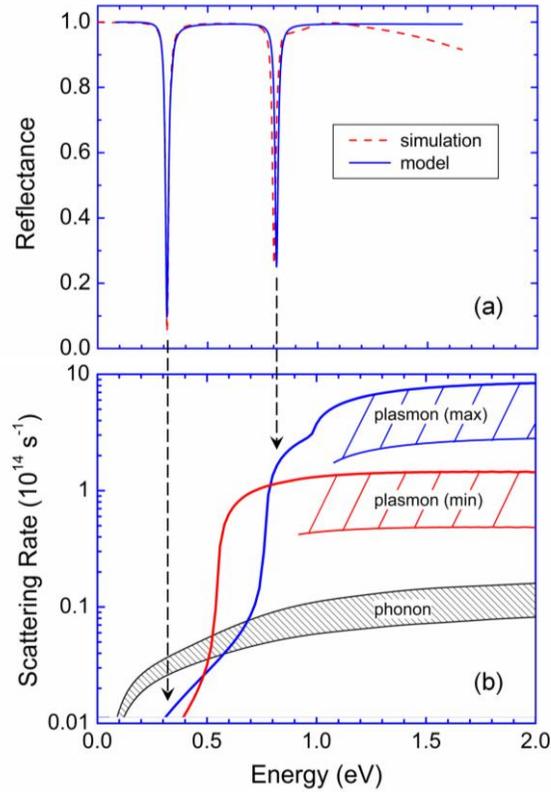
Before using it to calculate the scattering rate, the local effective dielectric function  $\varepsilon(\omega)$  obtained from the fit must be extended into the nonlocal domain, *i.e.* it must be made dependent on the wavevector  $q$  (see Chapter 1 for general discussions). This is because while the optical response is limited to only very small momenta, much less than the Fermi momentum ( $q \ll k_F$ ), the plasmon scattering of hot electrons involves both small and large momentum transfers ( $q$  of order  $k_F$ ). The most efficient way to accomplish this analytical

extension is to use the  $d$ -function formalism of Feibelman [35, 36] (see also Chapter 2 discussions). This combines the simplicity of the Fresnel analysis with a rigorous quantum mechanical treatment by introducing surface response functions  $d(\omega)$ , which have been systematically calculated for various metals [37, 38]. Liebsch [39] has shown that the self-energy in Eq. (4.1) can be expressed in terms of the Feibelman  $d$ -function as:

$$\Sigma(q) = |q|D(\omega)\omega_s^2 \approx |q|D(\omega_s)\omega_s^2 \quad (4.3)$$

$$\omega_s = \omega_p / \sqrt{1 + \text{Re}(\epsilon_b)} \quad (4.4)$$

where  $\omega_s$  is the surface plasmon frequency, and Eq. (4.3) is inserted into Eq. (4.1) to yield the non-locally extended dielectric function, which subsequently is inserted into Eqs. (3.7) and (3.20) to yield  $\gamma_{el-pl}$ , shown in Fig. 4.3(b). Due to the simplicity of the Lorentzian formula, part of Eq. (3.20) can be done analytically (see chapter 3 section 4 and chapter 5 section 1). One result shown (in red, marked “min”) could be considered a worst-case scenario, using the largest possible value of the only adjustable parameter,  $\epsilon_b = 10$ , which assumes that  $\epsilon_b$  equals the maximum bulk value in silver [40]. We also show a more realistic scenario range (in blue, marked “max”) which uses a bulk vacuum-weighted background dielectric constant of  $\epsilon_b = 3$ .



**Fig. 4.3 (a)** Reflectance of the structure shown in Fig. 4.2: simulations (dashed line) and model three-layer system calculations (solid line). **(b)** Calculated minimum (red) and maximum (blue) ranges of scattering rates for hot electrons with plasmons (corresponding to the upper and lower limits, respectively, for the background permittivity), and the reported range of scattering rates [24,25] for hot electrons with phonons (black). Arrows indicate connections between plasmonic absorbance resonances and enhancements in hot electron-plasmon scattering. This figure is from: J. Kong, A. H. Rose, C. Yang, X. Wu, J. M. Merlo, M. J. Burns, M. J. Naughton, and K. Kempa\*, “A hot electron plasmon-protected solar cell”, *Optics Express* 23, A1087 (2015).

Finally, these ranges can be compared to a range for the electron-phonon scattering rate, obtained from Refs. [24-25]. Figure 4.3(b) demonstrates that in the typical range of hot electron energies (say, 0.7 – 2 eV, corresponding to the energy range of solar radiation minus a typical PV absorber band gap) one has at least  $\gamma_{el-pl}/\gamma_{el-ph} > 1$  (though it could reach  $\sim 10^2$ , Fig. 4.3(b)), broadly confirming the earlier, simple model result. Throughout, we refer to the hot electron energy  $E_{hot}$  as that in excess of the conduction band minimum. The most important feature for the HELPP effect is the *hot-electron relaxation (i.e. complete cooling-off) time*,  $\tau^c$ ; this is the time required to cool a hot electron from its initial, high photo-excited energy state down to the bottom of the conduction band. This time depends on the quantum energy exchanged at each scattering event. While the typical phonon energy is in the 50 meV range, the plasmon energy is a few hundred meV, and so about factor of ten larger. Therefore, it takes about 10 times fewer plasmon scattering events than phonon events to completely cool off a hot electron. Taking this effect into account, we estimate that for an average hot-electron energy of about 1 eV, which coincides with  $\gamma_{el-pl}/\gamma_{el-ph} \gg 1$ , the corresponding cooling time ratio is  $\tau_{el-pl}^c/\tau_{el-ph}^c \ll 1$ . Thus, *electron-plasmon scattering dominates carrier relaxation*.

#### 4.4 Application to photovoltaics

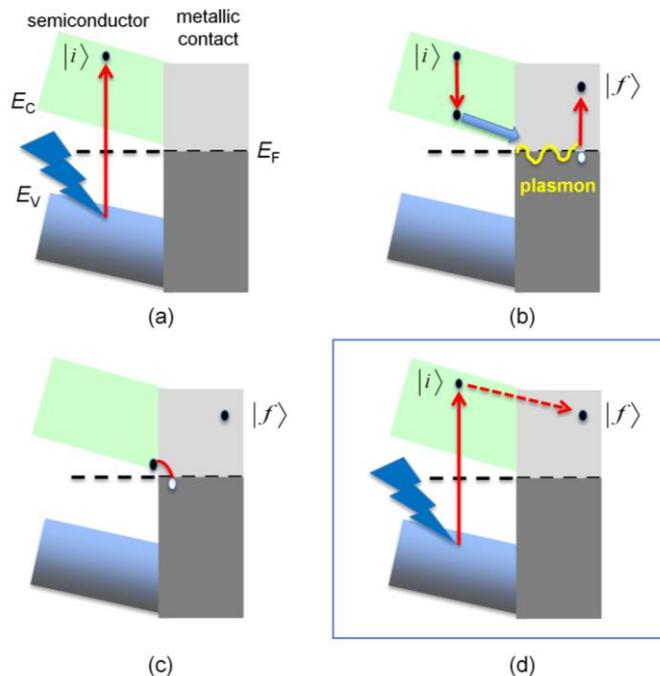
With a PV junction representing the semiconductor film, the structure of Fig. 4.2(a), after coating with a transparent conductor, is a functioning solar cell, with the periodic, plasmonic structure also acting as one charge collector of the cell. The HELPP action will lead to an increased voltage of the cell relative to that without HELPP. To see this, consider a hot electron generated in the conduction band of the semiconductor, as in Fig. 4.4(a). In the absence of HELPP, this electron would first lose its entire free energy to phonon scattering on a femtosecond time scale, and then, as it drifts towards the collector, would continue to thermalize with the lattice, as well as with other electrons. With HELPP, the process is similar, except the initial “cooling” of the electron occurs via plasmon emission, with the plasmons resonating in the plasmonic collector, Fig. 4.4(b). Any electron arriving at the metallic plasmonic collector rejoins the free-energy stored initially in the form of a plasmon. These plasmons rapidly undergo Landau damping (LD), *i.e.*, they turn into single particle excitations (electron-hole pairs), such that the plasmon energy turns into the electron-hole pair energy, Fig. 4.4(b). This damping is rapid, comparable to that of plasmon-electron scattering [33, 41], such that the hot electron energy remains in the electronic degree of freedom, *i.e.* is protected from phonon emission. Initially photo-excited electrons arriving at the collector recombine with the hole of this electron-hole pair, Fig. 4.4(c), and an excited hot-electron in the collector is the final product of the process, Fig. 4.4(d). The end result is thus identical to that of *a hot-electron never losing its free energy on its way to the collector.*

HELPP, therefore, is a mechanism that effectively extends the lifetime (or the mean free path) of a hot electron. In recent work, it was demonstrated that in ultrathin *p-i-n a-Si*

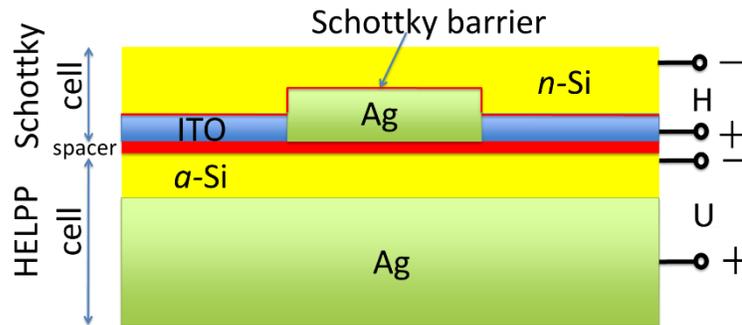
junctions, hot electrons contributed to small voltage increase to the overall open circuit voltage [42]. We anticipate a similar but significantly stronger effect to be observable in the proposed HELPP structure.

The efficiency of this effect can potentially be improved by modifying the structure so that the plasmonic resonators are electrically separated from the semiconductor. This can be achieved by placing a very thin dielectric spacer between the semiconductor and the array of metallic resonators (red layer in the schematic in Fig. 4.5). This should eliminate any deleterious, direct electron-electron scattering between the electrons in the semiconductor and the plasmonic reservoir, making the HELPP effect much stronger. The hot-electron energy stored in the plasmonic reservoirs can potentially be detected via small radiation resulting from the plasmon resonators, or can be directly extracted from the generated LD electron-hole pairs. These can be spatially separated in a metal-insulator-metal (MIM) structure [43, 44], or at Schottky barriers formed at the interface of the plasmonic nanostructures and the semiconductor [41]. In either case, an electron of the LD electron-hole pair is excited above the asymmetric potential barrier, leaving behind a hole in the nanostructure, and producing a contact voltage due to separation of the carriers. As such, one could prepare a type of *tandem, high efficiency hot electron cell*, as depicted in Fig. 4.5. The first cell of this tandem is our basic HELPP cell shown in Fig. 4.2, but with an insulating spacer between the PV junction and the Ag plasmonic resonators (thick red layer), as well as ITO. This cell would produce the conventional photovoltage, as indicated in the band diagram in Fig. 4.6. The second cell consists of the plasmonic reservoir array coated with a thin semiconductor (*e.g.* *n*-type Si or *a*-Si) to form a Schottky barrier with the metal of the reservoir and ITO. This cell produces a voltage that is directly proportional

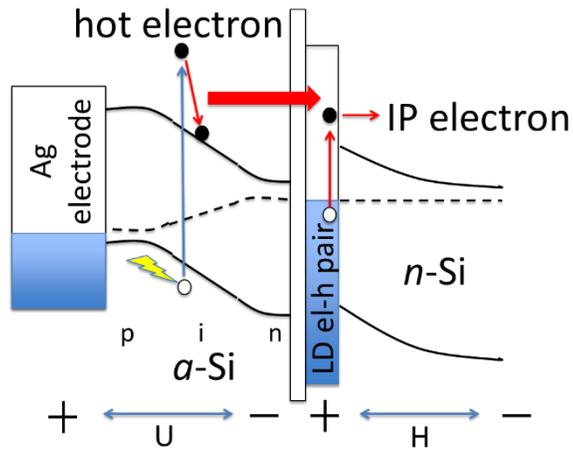
to the hot electron energy. Connected in series, this tandem combination would lead to a hot electron solar cell with efficiency exceeding the Shockley-Queisser limit. Current matching can be achieved simply by controlling the surface area of each cell in the tandem.



**Fig. 4.4** Schematic of HELPP action. **(a)** A semiconductor absorbs a photon of energy greater than the band gap  $E_C - E_V$ , exciting an electron high into the conduction band. **(b)** Prior to losing the above-gap excess energy to phonons/heat, this hot electron resonantly exchanges that excess energy with a plasmon mode in a proximate plasmonic metamaterial structure having effective Fermi energy  $E_F$ . This plasmon Landau damps into an electron-hole pair in the metal (e.g. Ag). **(c)** The originally photo-excited electron recombines with the plasmon-excited hole at  $E_F$ , leaving behind a high energy electron to be harvested as current, and at higher voltage than that conventionally determined by the semiconductor band gap. These 3 steps can equivalently be described by the process in **(d)**. This figure is from: J. Kong, A. H. Rose, C. Yang, X. Wu, J. M. Merlo, M. J. Burns, M. J. Naughton, and K. Kempa\*, "A hot electron plasmon-protected solar cell", *Optics Express* 23, A1087 (2015).



**Fig. 4.5** Schematic of a tandem hot electron cell, with a thin plasmonic resonator-embedded Schottky junction generating voltage  $H$  (for hot) separated from an ultrathin solar cell ( $a$ -Si shown) generating voltage  $U$ . This figure is from: J. Kong, A. H. Rose, C. Yang, X. Wu, J. M. Merlo, M. J. Burns, M. J. Naughton, and K. Kempa\*, “A hot electron plasmon-protected solar cell”, *Optics Express* 23, A1087 (2015).



**Fig. 4.6** Band diagram of tandem cell of Fig. 4.5. IP is inverse photoemission process for hot electron energy release from plasmonic resonator. LD refers to Landau damping. This figure is from: J. Kong, A. H. Rose, C. Yang, X. Wu, J. M. Merlo, M. J. Burns, M. J. Naughton, and K. Kempa\*, “A hot electron plasmon-protected solar cell”, *Optics Express* 23, A1087 (2015).

## 4.5 Conclusion

It is demonstrated via simulation and calculation a solar cell based on hot electron plasmon protection (HELPP). A thin-film metamaterial structure acts as both an efficient photon absorber in the visible range and a plasmonic resonator in the IR, the latter of which absorbs the free energy of the hot electrons in an adjacent semiconductor junction. With a combination of FDTD simulations, nonlocal effective medium modeling, and quantum-mechanical calculations, we show that hot electron–plasmon scattering is much more efficient in cooling-off these electrons than conventional phonon emission, and that the plasmon-stored energy is recoverable as an additional cell voltage. The proposed structure could become a prototype of a high efficiency solar cell.

## References

- [1]. J. B. Gunn, “Microwave oscillations of current in III–V semiconductors,” *Solid State Commun.* 1(4), 88–91 (1963).
- [2]. R. T. Ross and A. J. Nozik, “Efficiency of hot-carrier solar energy converters,” *J. Appl. Phys.* 53(5), 3813–3818 (1982).
- [3]. H. Kroemer, “Theory of the Gunn effect,” *Proc. IEEE* 52, 1736 (1964).
- [4]. D. Frohman-Bentchkowsky, “Memory behavior in a floating-gate avalanche-injection MOS (FAMOS) structure,” *Appl. Phys. Lett.* 18(8), 332–334 (1971).
- [5]. M. Heiblum, M. I. Nathan, D. C. Thomas, and C. M. Knoedler, “Direct observation of ballistic transport in GaAs,” *Phys. Rev. Lett.* 55(20), 2200–2203 (1985).
- [6]. C. Rauch, G. Strasser, K. Unterrainer, W. Boxleitner, K. Kempa, and E. Gornik, “Ballistic electron transport in vertical biased superlattices,” *Physica E* 2(1–4), 282–286 (1998)
- [7]. A. Othonos, “Probing ultrafast carrier and phonon dynamics in semiconductors,” *J. Appl. Phys.* 83(4), 1789–1830 (1998).
- [8]. J. R. Goldman and J. A. Prybyla, “Ultrafast dynamics of laser-excited electron distributions in silicon,” *Phys. Rev. Lett.* 72(9), 1364–1367 (1994).
- [9]. S. M. Sze, *High-Speed Semiconductor Devices* (Wiley, 1990).
- [10]. K. M. Kramer and W. N. G. Hitchon, *Semiconductor Devices: A Simulation Approach* (Prentice Hall, 1997).
- [11]. J. A. Kash and J. C. Tsang, “Watching chips work: Picosecond hot electron light emission from integrated circuits,” *J. Cryst. Growth* 210(1–3), 318–322 (2000).
- [12]. W. Shockley and H. J. Queisser, “Detailed balance limit of efficiency of p-n junction solar cells,” *J. Appl. Phys.* 32(3), 510–519 (1961).
- [13]. M. A. Green, *Third Generation Photovoltaics: Advanced Solar Energy Conversion* (Springer, 2006).
- [14]. K. Kempa, “Plasmonic protection of the hot-electron energy,” *Phys. Status Solidi RRL* 7(7), 465–468 (2013); (erratum) *ibid* 7(12), 1112 (2013).
- [15]. P. C. Cheng, “The contrast formation in optical microscopy,” in J. B. Pawley (ed.) *Handbook of Biological Confocal Microscopy*, 3rd edn. (Springer, 2006), pp. 162–206.
- [16]. C. Clavero, “Plasmon-induced hot-electron generation at nanoparticle/metaloxide interfaces for photovoltaic and photocatalytic devices,” *Nat. Photonics* 8, 95–103 (2014).

- [17]. H. Chalabi and M. L. Brongersma, “Plasmonics: harvest season for hot electrons,” *Nat. Nanotechnol.* 8, 229–230 (2013).
- [18]. M. L. Brongersma, N. J. Halas, and P. Nordlander, “Plasmon-induced hot carrier science and technology,” *Nat. Nanotechnol.* 10, 25–34 (2015).
- [19]. D. J. Farrell, H. Sodabanlu, Y. Wang, M. Sugiyama, and Y. Okada, “Can a hot-carrier solar cell also be an efficient up-converter?” *IEEE J. Photovolt.* 5(2), 571–576 (2015).
- [20]. G. L. Liu, Y. T. Long, Y. Choi, T. Kang, and L. P. Lee, “Quantized plasmon quenching dips nanospectroscopy via plasmon resonance energy transfer,” *Nat. Methods* 4(12), 1015–1017 (2007).
- [21]. G. D. Mahan, *Many-Particle Physics* (Plenum, 1981).
- [22]. J. J. Quinn and R. A. Ferrell, “Electron self-energy approach to correlation in a degenerate electron gas,” *Phys Rev* 112(3), 812–827 (1958).
- [23]. R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem* (McGraw-Hill, 1976).
- [24]. O. D. Restrepo, K. Varga, and S. T. Pantelides, “First-principles calculations of electron mobilities in silicon: Phonon and Coulomb scattering,” *Appl. Phys. Lett.* 94, 212103 (2009).
- [25]. D. Rideau, W. Zhang, Y. M. Niquet, C. Delerue, C. Tavernier, and H. Jaouen, “Electron-phonon scattering in Si and Ge: From bulk to nanodevices,” 2011 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) Conference Publication (2011) pp. 47–50.
- [26]. Y. Wang, T. Sun, T. Paudel, Y. Zhang, Z. Ren, and K. Kempa, “Metamaterial-plasmonic absorber structure for high efficiency amorphous silicon solar cells,” *Nano Lett.* 12, 440–445 (2012).
- [27]. F. Ye, M. J. Burns, and M. J. Naughton, “Embedded metal nanopatterns as a general scheme for enhanced broadband light absorption,” *Phys. Status Solidi A* 212, 561–565 (2014);
- [28]. F. Ye, M. J. Burns, and M. J. Naughton, “Embedded metal nanopatterns for near-field scattering-enhanced optical absorption,” 209, 1829–1834 (2012).
- [29]. A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech House, 1995) Norwood, MA.
- [30]. X. Wang and K. Kempa, “Negative refraction and subwavelength lensing in a polaritonic crystal,” *Phys. Rev. B* 71(23), 233101 (2005).

- [31]. Computer Simulation Technology, [www.cst.com](http://www.cst.com)
- [32]. A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson, "MEEP: A flexible free-software package for electromagnetic simulations by the FDTD method," *Comp. Phys. Commun.* 181, 687–702 (2010).
- [33]. M. Scharfe, R. Porath, T. Ohms, M. Aeschlimann, J. R. Krenn, H. Ditlbacher, F. R. Aussenegg and A. Liebsch, "Do Mie plasmons have a longer lifetime on resonance than off resonance?" *Appl. Phys. B* 73(4), 305–310 (2001).
- [34]. O. S. Heavens, *Optical Properties of Thin Solid Films* (Dover, 1965).
- [35]. P. J. Feibelman, "Microscopic calculation of electromagnetic fields in refraction at a jellium-vacuum interface," *Phys. Rev. B* 12(4), 1319–1336 (1975).
- [36]. P. J. Feibelman, "Surface electromagnetic fields," *Prog. Surf. Sci.* 12(4), 287–407 (1982).
- [37]. A. Liebsch, "Dynamical screening at simple-metal surfaces", *Phys. Rev. B* 36(14), 7378–7387 (1987).
- [38]. K. Kempa, A. Liebsch and W. L. Schaich, "Comparison of calculations of dynamical screening at jellium surfaces," *Phys. Rev. B* 38(17), 12645–12648 (1988).
- [39]. A. Liebsch, "Surface-plasmon dispersion and size dependence of Mie resonance: Silver versus simple metals", *Phys. Rev. B* 48(15), 11317–11328 (1993).
- [40]. P. B. Johnson and R. W. Christy, "Optical constants of the noble metals," *Phys. Rev. B* 6(12), 4370–4379 (1972).
- [41]. M. W. Knight, H. Sobhani, P. Nordlander, and N. Halas, "Photodetection with active optical antennas," *Science* 332(6030), 702–704 (2011).
- [42]. K. Kempa, M. J. Naughton, Z. Ren, A. Herczynski, T. Kirkpatrick, J. Rybczynski, and Y. Gao, "Hot electron effect in nanoscopically thin photovoltaic junctions," *Appl. Phys. Lett.* 95, 233121 (2009).
- [43]. F. Wang and N. A. Melosh, "Plasmonic energy collection through hot carrier extraction," *Nano Lett.* 11(12), 5426–5430 (2011).
- [44]. F. Wang and N. A. Melosh, "Power-independent wavelength determination by hot carrier collection in metal-insulator-metal devices," *Nat. Commun.* 4, 1711 (2013).

## CHAPTER 5

### **Electron Polar-Optical-Phonon Scattering Rates in Heavily-Doped Semiconductors with Plasmonic Nanostructures Embedded-in**

**Outline:** This chapter extends the general results of electron lifetime in many-particle systems obtained in Chapter 3, and develops calculation scheme for the electron-phonon (specifically polar optical phonon POP) scattering rates in a newly proposed plasmonic metamaterial - traditional bulk semiconductor with metallic nanostructures embedded-in. Materials discussed in this chapter contributed to the paper submitted to Materials Today Physics: Wu#, Kong#, Broido and Kempa\*, “Tailoring the electron-phonon interaction with metallic plasmonic structures”.

## 5.1 Scattering rate with bare polar optical phonons (POP)

The physical picture of the general results obtained in Chapter 3 is, free electron gas subject to external or internal perturbation  $V_{\text{eff}}$ . Adapted into this chapter, carriers in the conduction band of a semiconductor form a free electron gas, and this gas is then subject to the perturbation due to the polar optical phonons (POP), with  $V_{\text{eff}} = V_{\text{q}} / \epsilon_{\text{POP}}$ .

In intrinsic (and very low doping) polar semiconductors, since the electron gas is dilute, the dielectric function is dominated by polarization due to phonons, and is given by [1, 2] (In this chapter,  $\omega_L$  is short for  $\omega_{\text{LO}}$ , the longitudinal optical phonon frequency, and  $\omega_T$  is short for  $\omega_{\text{TO}}$ , the transverse optical phonon frequency)

$$\epsilon(\omega) = \epsilon_{\infty} \frac{\omega_L^2 - \omega^2}{\omega_T^2 - \omega^2} \quad (5.1)$$

In order to calculate electron's lifetime in such medium, we can just plug this in the general result (3.20) of Chapter 3. Equation (5.1) could be written as a Lorentz oscillator [3] similar as in the electron case

$$\epsilon(\omega) = \epsilon_{\infty} \frac{\omega_T^2 - \omega^2 + \omega_L^2 - \omega_T^2}{\omega_T^2 - \omega^2} = \epsilon_{\infty} + \frac{\epsilon_{\infty}(\omega_L^2 - \omega_T^2)}{\omega_T^2 - \omega^2} \quad (5.2)$$

With the damping rate extension,

$$\omega^2 \rightarrow \omega(\omega + i\gamma), \quad \gamma \rightarrow 0^+$$

$\text{Im} \left[ \frac{1}{\epsilon} \right]$  could then be simplified as the delta function form as well (see chapter appendix)

$$\text{Im} \left[ \frac{1}{\epsilon} \right] = \frac{[\epsilon_{\infty}(\omega_L^2 - \omega_T^2)]}{(\epsilon_{\infty})^2} \frac{1}{2\omega_L} (-\pi) [\delta(\omega - \omega_L) - \delta(\omega + \omega_L)]$$

$$= \frac{(\omega_L^2 - \omega_T^2)}{\varepsilon_\infty} \frac{1}{2\omega_L} (-\pi) [\delta(\omega - \omega_L) - \delta(\omega + \omega_L)] \quad (5.3)$$

Inserting (5.3) into equation (3.20) yields a golden-rule-like expression (here, terms with  $\mathbf{k}-\mathbf{q}$  means a transition from  $\mathbf{k}$  to  $\mathbf{k}-\mathbf{q}$  by emitting a phonon mode  $\mathbf{q}$ ; terms with  $\mathbf{k}+\mathbf{q}$  means a transition from  $\mathbf{k}$  to  $\mathbf{k}+\mathbf{q}$  by absorbing a phonon mode  $\mathbf{q}$ )

$$\begin{aligned} \gamma_k^{eff} &= \frac{-2}{\hbar} \int \frac{d^3q}{(2\pi)^3} V_q [n_B(E_{k\mp q} - E_k) + n_F(E_{k\mp q} - \mu)] \text{Im} \left[ \frac{1}{\varepsilon} \right] \\ &= \frac{-2}{\hbar} \int \frac{d^3q}{(2\pi)^3} V_q \left[ \begin{array}{l} n_B(E_{k\mp q} - E_k) \\ + n_F(E_{k\mp q} - \mu) \end{array} \right] (-\pi) \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} [\delta(\omega - \omega_L) - \delta(\omega + \omega_L)] \\ &= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{d^3q}{(2\pi)^3} V_q \left[ \begin{array}{l} n_B(E_{k\mp q} - E_k) \\ + n_F(E_{k\mp q} - \mu) \end{array} \right] \cdot \left[ \begin{array}{l} \delta(E_{k\mp q} - E_k - \hbar\omega_L) \\ - \delta(E_{k\mp q} - E_k + \hbar\omega_L) \end{array} \right] \\ &= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{d^3q}{(2\pi)^3} V_q \left\{ \begin{array}{l} - [n_B(-\hbar\omega_L) + n_F(E_{k-q} - \mu)] \delta(E_{k-q} - E_k + \hbar\omega_L) \\ + [n_B(\hbar\omega_L) + n_F(E_{k+q} - \mu)] \delta(E_{k+q} - E_k - \hbar\omega_L) \end{array} \right\} \\ &= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{d^3q}{(2\pi)^3} V_q \left\{ \begin{array}{l} [1 + n_B(\hbar\omega_L) - n_F(E_{k-q} - \mu)] \delta(E_{k-q} - E_k + \hbar\omega_L) \\ + [n_B(\hbar\omega_L) + n_F(E_{k+q} - \mu)] \delta(E_{k+q} - E_k - \hbar\omega_L) \end{array} \right\} \end{aligned} \quad (5.4)$$

Integrate on final states  $\mathbf{f}$ , instead of on  $\mathbf{q}$ ,

$$V_q = \frac{1}{4\pi\varepsilon_0} \frac{4\pi e^2}{q^2} = \frac{e^2}{\varepsilon_0} \frac{1}{k^2 + f^2 - 2kf \cos \theta} \quad (5.5)$$

and define  $F_\pm$  such that (“F” stands for final. Sign convention is, final state – initial state,

$$\frac{\hbar^2 f^2}{2m} - \frac{\hbar^2 k^2}{2m} \equiv \hbar\omega)$$

$$\frac{\hbar^2 F_\pm^2}{2m} - \frac{\hbar^2 k^2}{2m} = \pm \hbar\omega_L$$

Then the integral could be carried out as (see end chapter appendix for detail steps)

$$\begin{aligned}
\gamma_k^{\text{eff}} &= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{d^3q}{(2\pi)^3} V_q \left[ n_B(E_{k\mp q} - E_k) + n_F(E_{k\mp q} - \mu) \right] \left[ \delta(\omega - \omega_L) - \delta(\omega + \omega_L) \right] \\
&= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{f \sin\theta d\phi f d\theta df}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{k^2 + f^2 - 2kf \cos\theta} \left[ \frac{n_B(E_{k\mp q} - E_k)}{+n_F(E_{k\mp q} - \mu)} \right] \left[ \frac{\delta(f - F_+)}{\hbar F_+/m} - \frac{\delta(f - F_-)}{\hbar F_-/m} \right]
\end{aligned} \tag{5.6}$$

Result is (out-scattering, unscreened)

$$= \frac{\hbar\omega_L}{\hbar^2} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right) \frac{e^2}{4\pi\varepsilon_0} \sqrt{\frac{m}{2E}} \left\{ \begin{aligned} &\ln \left( \frac{\sqrt{E} + \sqrt{E - \hbar\omega_L}}{\sqrt{E} - \sqrt{E - \hbar\omega_L}} \right) \left[ 1 + n_B(\hbar\omega_L) - n_F(E - \hbar\omega_L - \mu) \right] \\ &+ \ln \left( \frac{\sqrt{E + \hbar\omega_L} + \sqrt{E}}{\sqrt{E + \hbar\omega_L} - \sqrt{E}} \right) \left[ n_B(\hbar\omega_L) + n_F(E + \hbar\omega_L - \mu) \right] \end{aligned} \right\} \tag{5.7}$$

## 5.2 Electron-POP scattering rate with Thomas-Fermi screening

The above equation (5.7) is the basic raw result for conduction band electron gas subject to perturbation by polar optical phonons. Next-step correction is to take into account the screening of the other carriers, which will become important when the polar semiconductor is gradually doped making the carrier density higher and higher. A common practice in the literature is the Thomas-Fermi approximation [4], screening the vertex of Frohlich interaction (the matrix element). In fact, Frohlich interaction - the phonon mediated electron-electron interaction, is second order, thus there are two vertices (see introduction part in Chapter 3), making the screening follows  $1/\epsilon^2$ , instead of  $1/\epsilon$ . (Mahan [5] chapter 6 has a clear discussion too) In this case, the bare Coulomb interaction would become [see appendix at this section end for a brief “justification” for this approximation]

$$V_q = \frac{1}{4\pi\epsilon_0} \frac{4\pi e^2}{q^2} \left( \frac{1}{1+q_{TF}^2/q^2} \right)^2 = \frac{e^2}{\epsilon_0} \frac{k^2 + f^2 - 2kf \cos \theta}{(k^2 + f^2 - 2kf \cos \theta + q_{TF}^2)^2} \quad (5.8)$$

Define (same form as in Refs. [6, 7])

$$I_{POP}(k, F_{\pm}, q_{TF}) \equiv \int_{-1}^1 \frac{1 + F_{\pm}^2/k^2 + 2xF_{\pm}/k}{(1 + F_{\pm}^2/k^2 + 2xF_{\pm}/k + q_{TF}^2/k^2)^2} dx \quad (5.9)$$

And the result further is

$$\gamma_{k-out}^{eff} = \frac{\hbar\omega_L}{\hbar^2} \left( \frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon(0)} \right) \frac{e^2}{4\pi\epsilon_0} \frac{m}{\hbar} \times \left\{ \begin{array}{l} \frac{F_-}{k^2} I_{POP}(k, F_-, q_{TF}) [1 + n_B(\hbar\omega_L) - n_F(E - \hbar\omega_L - \mu)] \\ + \frac{F_+}{k^2} I_{POP}(k, F_+, q_{TF}) [n_B(\hbar\omega_L) + n_F(E + \hbar\omega_L - \mu)] \end{array} \right\}$$

$$\begin{aligned}
&= \frac{\hbar\omega_L}{\hbar^3} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right) \frac{me^2}{4\pi\varepsilon_0} \frac{1}{k^2} \\
&\quad \times \left\{ \begin{aligned} &F_- I_{POP}(k, F_-, q_{TF}) [1 + n_B(\hbar\omega_L) - n_F(E - \hbar\omega_L - \mu)] \\ &+ F_+ I_{POP}(k, F_+, q_{TF}) [n_B(\hbar\omega_L) + n_F(E + \hbar\omega_L - \mu)] \end{aligned} \right\} \quad (5.10)
\end{aligned}$$

This is identical as the out-scattering result in Refs. [6, 7] from different approaches.

In-scattering rate (3.21) could be carried out similarly. Define  $\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 I_\pm^2}{2m} = \mp \hbar\omega_L$

("I" stands for initial. Sign convention is, final state – initial state,  $\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 I^2}{2m} \equiv \hbar\omega$ )

$$\gamma_{k-in}^{eff} = \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{d^3 q}{(2\pi)^3} V_q [n_B(E_k - E_{k\pm q}) + n_F(E_k - \mu)] [\delta(\omega - \omega_L) - \delta(\omega + \omega_L)] \quad (5.11)$$

Result (in-scattering, unscreened) is similar to that of out-scattering,

$$\frac{\hbar\omega_L}{\hbar^2} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right) \frac{e^2}{4\pi\varepsilon_0} \sqrt{\frac{m}{2E}} \left\{ \begin{aligned} &\ln \left( \frac{\sqrt{E + \hbar\omega_L} + \sqrt{E}}{\sqrt{E + \hbar\omega_L} - \sqrt{E}} \right) [1 + n_B(\hbar\omega_L) - n_F(E - \mu)] \\ &+ \ln \left( \frac{\sqrt{E} + \sqrt{E - \hbar\omega_L}}{\sqrt{E} - \sqrt{E - \hbar\omega_L}} \right) [n_B(\hbar\omega_L) + n_F(E - \mu)] \end{aligned} \right\} \quad (5.12)$$

With Thomas-Fermi screening (5.8), result is

$$\begin{aligned}
\gamma_{k-in}^{eff} &= \frac{\hbar\omega_L}{\hbar^3} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right) \frac{me^2}{4\pi\varepsilon_0} \frac{1}{k^2} \\
&\quad \times \left\{ \begin{aligned} &k_- I_{POP}(k, k_-, q_{TF}) [n_B(\hbar\omega_L) + n_F(E - \mu)] \\ &+ k_+ I_{POP}(k, k_+, q_{TF}) [1 + n_B(\hbar\omega_L) - n_F(E - \mu)] \end{aligned} \right\} \quad (5.13)
\end{aligned}$$

$I_{POP}$  is the same form as out-scattering part defined in Eq. (5.9).

From the point of view of classical transport theory, in the calculation scheme presented in Chapter 3, there is no electric field or temperature gradient applied, so it is all about EQUILIBRIUM state. When field is turned on and the system gets stable eventually,

it is then called STEADY state, or STATIONARY state (the whole idea of Boltzmann equation is to solve for this steady state). In equilibrium state, the out-scattering and in-scattering cancel out, it is the so called "principle of detailed balance" (see Ref. [8] Eq. (7.3.8) on page 265, and Ref. [7] page 136). In steady state when there's field applied, left hand side of Ref. [8, Ziman] (7.3.5) is non-zero, so right hand side has to be non-zero too. In other words, the field is accelerating the electrons (changing the distribution), and scattering(collison) is changing the distribution at the same time. These two processes cancel each other as Boltzmann equation requires, producing the steady state distribution  $f_k(r)$ , rather than the equilibrium state distribution  $f_k^0(r)$ .

In the calculation scheme presented in Chapter 3, equilibrium of the electron gas is one of the initial assumptions. The core formula for in-scattering, equation (3.21), was set up by symmetric consideration. In fact, the implied physics is, "principle of detailed balance". It might be a good topic to extend the Chapter 3 electron self-energy (lifetime) calculation scheme, to non-equilibrium situations [9].

### **Section Appendix: a justification of the Thomas-Fermi approximation used**

Refer to [5, page565], the screened phonon interaction in heavily-doped polar semiconductor is

$$V_{sc-ph} = \frac{V_{ph}}{\epsilon^2 (1 - V_{ph} P/\epsilon)} \quad (5.14)$$

where  $\varepsilon$  is the contribution due to electron gas, P is the electronic polarization part, and the bare phonon interaction is [5, page562]

$$V_{ph} = -v_q \frac{L^2}{L^2 - \omega^2} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right) = v_q \left( \frac{1}{\varepsilon_{ph}} - \frac{1}{\varepsilon_\infty} \right) \quad (5.15)$$

where  $\varepsilon_{ph} \equiv \varepsilon_\infty \frac{T^2 - \omega^2}{L^2 - \omega^2}$

Ignoring electronic polarization part P in this phononic calculation, i.e. setting P = 0,

$$\begin{aligned} V_{sc-ph} &= \frac{V_{ph}}{\varepsilon^2 (1 - V_{ph} P / \varepsilon)} \approx \frac{V_{ph}}{\varepsilon^2} \\ &= \frac{1}{\varepsilon^2} v_q \left( \frac{1}{\varepsilon_{ph}} - \frac{1}{\varepsilon_\infty} \right) \\ &\approx \frac{1}{\varepsilon^2} v_q \left( \frac{1}{\varepsilon_{ph}} \right) \end{aligned} \quad (5.16)$$

With the above two approximations, we justify the screened phonon interaction used in equation (5.8), in getting the same electron-POP out-scattering rate as Ref. [6] and Ref. [7, (6.65) (7.30)]. We can also make use of the full version of equation (5.14) where electronic polarization part P is fully considered, which will be discussed in the next few sections, numerically.

### 5.3 Electron scattering in metamaterials (case 1)

#### ---plasmonic resonance near LO phonons

Embedding nano/micro-sized metallic structures in dielectric materials (e.g. polar semiconductors) makes an effective medium. Both the added structure resonance and the original POP resonance as independent contributions in the total dielectric response could be written as (see Refs. [10, 11] for discussions on the superposition or additivity) Lorentz oscillators, each stands for an independent energy dissipation channel,

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{S_1^2}{R_1^2 - \omega^2} + \frac{S_2^2}{R_2^2 - \omega^2} \quad (5.17)$$

Schematically it looks like Fig. 5.1 (numbers are arbitrary; see appendix for detail).

Response function of such an  $\varepsilon$  can be written as

$$\begin{aligned} \text{Im} \left[ \frac{1}{\varepsilon(\omega)} \right] &= \frac{(P_1^2 - R_1^2)(P_1^2 - R_2^2)}{\varepsilon_\infty (P_1^2 - P_2^2)} \frac{(-\pi)}{2P_1} [\delta(\omega - P_1) - \delta(\omega + P_1)] \\ &\quad + \frac{(P_2^2 - R_1^2)(P_2^2 - R_2^2)}{\varepsilon_\infty (P_2^2 - P_1^2)} \frac{(-\pi)}{2P_2} [\delta(\omega - P_2) - \delta(\omega + P_2)] \\ &= \text{Im} \left[ \frac{1}{\varepsilon(\omega)} \right]_1 + \text{Im} \left[ \frac{1}{\varepsilon(\omega)} \right]_2 \end{aligned} \quad (5.18)$$

where  $P_1^2$  and  $P_2^2$  are roots of

$$\varepsilon_\infty (R_1^2 - \omega^2)(R_2^2 - \omega^2) + S_1^2 (R_2^2 - \omega^2) + S_2^2 (R_1^2 - \omega^2) = 0$$

Simple linear form of (5.18) makes the further calculation straight forward. Plugging first term in general results (3.20) yields out-scattering by channel-one (unscreened),

$$\begin{aligned}
[\gamma_{k-out}^{eff}]_1 &= \frac{-2}{\hbar} \int \frac{d^3q}{(2\pi)^3} V_q [n_B(E_{k\mp q} - E_k) + n_F(E_{k\mp q} - \mu)] \text{Im} \left[ \frac{1}{\varepsilon} \right]_1 \\
&= \frac{-2}{\hbar} \int \frac{d^3q}{(2\pi)^3} V_q \left[ \frac{n_B(E_{k\mp q} - E_k)}{+n_F(E_{k\mp q} - \mu)} \right] \frac{(P_1^2 - R_1^2)(P_1^2 - R_2^2)}{\varepsilon_\infty (P_1^2 - P_2^2)} \frac{(-\pi)}{2P_1} [\delta(\omega - P_1) - \delta(\omega + P_1)] \\
&= \dots \\
&= \frac{\hbar P_1}{\hbar^2} \left( \frac{P_1^2 - R_2^2}{P_1^2 - P_2^2} \right) \left( \frac{P_1^2 - R_1^2}{\varepsilon_\infty \cdot P_1^2} \right) \frac{e^2}{4\pi\varepsilon_0} \sqrt{\frac{m}{2E}} \left\{ \begin{aligned} &\ln \left( \frac{\sqrt{E} + \sqrt{E - \hbar P_1}}{\sqrt{E} - \sqrt{E - \hbar P_1}} \right) [1 + n_B(\hbar P_1) - n_F(E_k - \hbar P_1 - \mu)] \\ &+ \ln \left( \frac{\sqrt{E + \hbar P_1} + \sqrt{E}}{\sqrt{E + \hbar P_1} - \sqrt{E}} \right) [n_B(\hbar P_1) + n_F(E_k + \hbar P_1 - \mu)] \end{aligned} \right\}
\end{aligned} \tag{5.19}$$

Similarly, out-scattering by channel-two (unscreened),

$$\begin{aligned}
[\gamma_{k-out}^{eff}]_2 &= \frac{-2}{\hbar} \int \frac{d^3q}{(2\pi)^3} V_q [n_B(E_{k\mp q} - E_k) + n_F(E_{k\mp q} - \mu)] \text{Im} \left[ \frac{1}{\varepsilon} \right]_2 \\
&= \frac{\hbar P_2}{\hbar^2} \left( \frac{P_2^2 - R_1^2}{P_2^2 - P_1^2} \right) \left( \frac{P_2^2 - R_2^2}{\varepsilon_\infty \cdot P_2^2} \right) \frac{e^2}{4\pi\varepsilon_0} \sqrt{\frac{m}{2E}} \left\{ \begin{aligned} &\ln \left( \frac{\sqrt{E} + \sqrt{E - \hbar P_2}}{\sqrt{E} - \sqrt{E - \hbar P_2}} \right) [1 + n_B(\hbar P_2) - n_F(E_k - \hbar P_2 - \mu)] \\ &+ \ln \left( \frac{\sqrt{E + \hbar P_2} + \sqrt{E}}{\sqrt{E + \hbar P_2} - \sqrt{E}} \right) [n_B(\hbar P_2) + n_F(E_k + \hbar P_2 - \mu)] \end{aligned} \right\}
\end{aligned} \tag{5.20}$$

With Thomas-Fermi screening, equation (5.8), similar integral-factor (as in the bare POP case) shows in the results. For channel-one,

$$\begin{aligned}
[\gamma_{k-out}^{eff}]_1^{TF-scr} &= \frac{\hbar P_1}{\hbar^3} \left( \frac{P_1^2 - R_2^2}{P_1^2 - P_2^2} \right) \left( \frac{P_1^2 - R_1^2}{\varepsilon_\infty \cdot P_1^2} \right) \frac{me^2}{4\pi\varepsilon_0} \frac{1}{k^2} \\
&\times \left\{ \begin{aligned} &F_1^- I_{POP}(k, F_1^-, q_{TF}) [1 + n_B(\hbar P_1) - n_F(E_k - \hbar P_1 - \mu)] \\ &+ F_1^+ I_{POP}(k, F_1^+, q_{TF}) [n_B(\hbar P_1) + n_F(E_k + \hbar P_1 - \mu)] \end{aligned} \right\}
\end{aligned} \tag{5.21}$$

$I_{POP}$  is the same definition as in equation (5.9).  $F_1^\pm$  (final states) are defined as

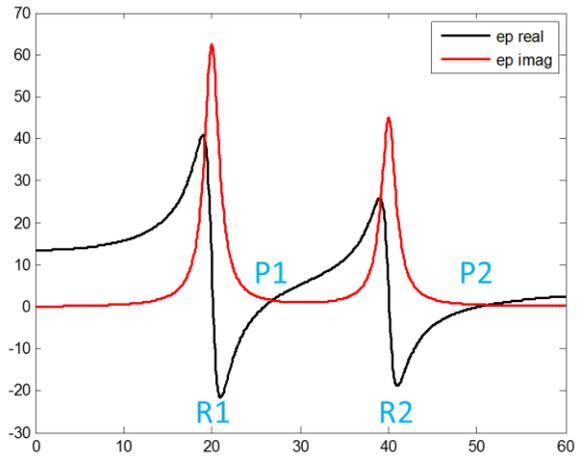
$$\frac{\hbar^2 (F_1^\pm)^2}{2m} - \frac{\hbar^2 k^2}{2m} = \pm \hbar P_1$$

Similarly, for channel-two,

$$\begin{aligned}
& \left[ \gamma_{k-out}^{eff} \right]_2^{TF-scr} \\
&= \frac{\hbar P_2}{\hbar^3} \left( \frac{P_2^2 - R_1^2}{P_2^2 - P_1^2} \right) \left( \frac{P_2^2 - R_2^2}{\epsilon_\infty \cdot P_2^2} \right) \frac{m e^2}{4\pi \epsilon_0} \frac{1}{k^2} \\
& \quad \times \left\{ \begin{aligned} & F_2^- I_{POP} \left( k, F_2^-, q_{TF} \right) \left[ 1 + n_B(\hbar P_2) - n_F(E_k - \hbar P_2 - \mu) \right] \\ & + F_2^+ I_{POP} \left( k, F_2^+, q_{TF} \right) \left[ n_B(\hbar P_2) + n_F(E_k + \hbar P_2 - \mu) \right] \end{aligned} \right\}
\end{aligned} \tag{5.22}$$

Also, applying to equation (3.21) will produce the corresponding in-scattering rates. Expressions would be similar as (5.21) and (5.22).

Back to the purpose of the project of this chapter, i.e., suppressing electron-phonon scattering in thermoelectric materials, observing above, we designate  $(R_1, P_1)$  as the POP resonance of the semiconductor (e.g.,  $Mg_2Si$ ), and  $(R_2, P_2)$  as the plasmonic resonance we are adding to the bulk  $Mg_2Si$ . It can then be seen, if setting  $R_2$  at  $P_1$ , (5.21) will vanish, i.e., electron-POP scattering will be totally suppressed. In other words, scattering by channel-one will be totally replaced by that of the added channel-two. This mixture/interference mechanism was also briefly mentioned in Ref. [5, Mahan] on page 563. It could be a general engineering trick, with possible applications in many other fields.



**Fig. 5.1** Schematic of real and imaginary parts of a 2-Lorentzian dielectric function. R stands for resonance. P stands for plasmon. (numbers are arbitrary; see chapter Appendix for detail of the response function of such  $\epsilon$ ).

## 5.4. Electron scattering in metamaterials (case 2)

### ----plasmonic resonance at much higher energy than LO phonons

Continuing discussion in the above section, if the added plasmonic resonance is at much higher energy than the LO phonons, the effect of the plasmonic resonance on electron-phonon interaction in low energy range would just be an approximately constant *screening/shielding*.

Simulations and extractions show (by Xueyuan Wu of Kempa Group), when ignoring the phonons (i.e., embedding the metallic structures in constant background), the plasmonic resonance could be written as (R stands for “resonance”, S stands for “strength”.)

$$\varepsilon(\omega) = \varepsilon_{\infty} + \frac{S^2}{R^2 - \omega^2} \quad (5.23)$$

Bringing back the polar optical phonons,

$$\varepsilon(\omega) = \varepsilon_{\infty} + \frac{\varepsilon_{\infty}(\omega_{LO}^2 - \omega_{TO}^2)}{\omega_{TO}^2 - \omega^2} + \frac{S^2}{R^2 - \omega^2} \quad (5.24)$$

At lower energy where the phonons located ( $\omega \ll R$ ),

$$\begin{aligned} \varepsilon(\text{low } \omega) &\approx \varepsilon_{\infty} + \frac{\varepsilon_{\infty}(\omega_{LO}^2 - \omega_{TO}^2)}{\omega_{TO}^2 - \omega^2} + \frac{S^2}{R^2 - 0^2} \\ &= \varepsilon_{\infty}' + \frac{\varepsilon_{\infty}(\omega_{LO}^2 - \omega_{TO}^2)}{\omega_{TO}^2 - \omega^2} \\ \varepsilon_{\infty}' &\equiv \varepsilon_{\infty} + a \equiv \varepsilon_{\infty} + \frac{S^2}{R^2} \end{aligned} \quad (5.25)$$

producing an effectively greater epsilon-infinity. Response function of such a medium is (see chapter appendix for detail)

$$\lim_{\delta \rightarrow 0^+} \text{Im} \left[ \frac{1}{\varepsilon(q, \omega)} \right] = \frac{\varepsilon_\infty (\omega_{LO}^2 - \omega_{TO}^2)}{(\varepsilon_\infty')^2} \frac{1}{2\omega_0} (-\pi) [\delta(\omega - \omega_0) - \delta(\omega + \omega_0)] \quad (5.26)$$

where

$$\omega_0 = \sqrt{\omega_{TO}^2 + \frac{\varepsilon_\infty (\omega_{LO}^2 - \omega_{TO}^2)}{\varepsilon_\infty'}} \quad (5.27)$$

is the plasmon-screened phonon frequency (similar results like Ref. [12]), shifted from  $\omega_{LO}$ . Physically, after bringing in the metallic structures, we have two kinds of *elementary excitations* in the many-particle system [11]. If taking into account the screening(dressing) of one to the other, we can treat these two *elementary excitations* independently. A brief math manipulation,

$$\begin{aligned} \varepsilon(\text{low } \omega) &\approx \varepsilon_\infty' + \frac{\varepsilon_\infty (\omega_{LO}^2 - \omega_{TO}^2)}{\omega_{TO}^2 - \omega^2} \\ &= \frac{\varepsilon_\infty' (\omega_{TO}^2 - \omega^2) + \varepsilon_\infty (\omega_{LO}^2 - \omega_{TO}^2)}{\omega_{TO}^2 - \omega^2} \\ &= \varepsilon_\infty' \frac{(\omega_{TO}^2 - \omega^2) + \frac{\varepsilon_\infty}{\varepsilon_\infty'} (\omega_{LO}^2 - \omega_{TO}^2)}{\omega_{TO}^2 - \omega^2} \\ &= \varepsilon_\infty' \frac{\left[ \sqrt{\omega_{TO}^2 + \varepsilon_\infty / \varepsilon_\infty'} (\omega_{LO}^2 - \omega_{TO}^2) \right]^2 - \omega^2}{\omega_{TO}^2 - \omega^2} \\ &= \varepsilon_\infty' \frac{\omega_0^2 - \omega^2}{\omega_{TO}^2 - \omega^2} \end{aligned} \quad (5.28)$$

Comparing (5.28) with the original POP equation (5.1), we find it only needs a simple parameter replacement in the previous bare electron-POP scattering equations (5.7-11) (one-term Lorentzian model)

$$\varepsilon_\infty \rightarrow \varepsilon_\infty' \quad \omega_{LO} \rightarrow \omega_0 \quad \omega_{TO} \rightarrow \omega_{TO} \quad (5.29)$$

to get all the results. Note that transverse-optical phonon frequency does not change, which is discussed in Refs. [7, 10]. Also, when the plasmonic resonance frequency  $R$  is really approaching infinity,

$$R, S \gg \omega_{LO}, \omega_{TO} \quad \text{while} \quad S^2/R^2 = a$$

the above result (5.29) is actually consistent with the two-term Lorentzian model results.

Equation (5.21) in the two-term Lorentzian scheme is

$$\begin{aligned} \left[ \mathcal{Y}_{k-out}^{eff} \right]_1^{TF-scr} &\sim \\ &\left( \frac{P_1^2 - R_2^2}{P_1^2 - P_2^2} \right) \left( \frac{P_1^2 - R_1^2}{\epsilon_\infty \cdot P_1^2} \right) \approx \left( \frac{0 - R_2^2}{0 - P_2^2} \right) \left( \frac{\omega_0^2 - R_1^2}{\epsilon_\infty \cdot \omega_0^2} \right) \approx \left( \frac{R_2^2}{R_2^2 + S_2^2/\epsilon_\infty} \right) \left( \frac{\left[ T^2 + \epsilon_\infty/\epsilon_\infty'(L^2 - T^2) \right] - T^2}{\epsilon_\infty \cdot \omega_0^2} \right) \\ &= \left( \frac{1}{1 + a/\epsilon_\infty} \right) \left( \frac{\epsilon_\infty/\epsilon_\infty'(L^2 - T^2)}{\epsilon_\infty \cdot \omega_0^2} \right) = \left( \frac{\epsilon_\infty}{\epsilon_\infty + a} \right) \left( \frac{(L^2 - T^2)}{\epsilon_\infty' \cdot \omega_0^2} \right) = \frac{\epsilon_\infty}{\epsilon_\infty'} \frac{(L^2 - T^2)}{\epsilon_\infty' \cdot \omega_0^2} \end{aligned} \quad (5.30)$$

corresponding part of equation (5.10) in the one-term scheme when applying (5.29) is

$$\begin{aligned} \left( \frac{1}{\epsilon_\infty} - \frac{1}{\epsilon(0)} \right) &= \frac{1}{\epsilon_\infty'} - \frac{1}{\epsilon_\infty' \omega_0^2/T^2} = \frac{\epsilon_\infty' \omega_0^2/T^2 - \epsilon_\infty'}{\epsilon_\infty' \epsilon_\infty' \omega_0^2/T^2} = \frac{\omega_0^2 - T^2}{\epsilon_\infty' \omega_0^2} \\ &= \frac{T^2 + \epsilon_\infty/\epsilon_\infty'(L^2 - T^2) - T^2}{\epsilon_\infty' \omega_0^2} = \frac{\epsilon_\infty (L^2 - T^2)}{\epsilon_\infty' \epsilon_\infty' \omega_0^2} \end{aligned} \quad (5.31)$$

the equivalence of (5.30) and (5.31) means this quantitative analysis is (at least) self-consistent.

## 5.5 Comprehensive discussions taking electronic and phononic perturbation on equal footing and methods to distinguish them

From electrons' point of view, the effective interaction between them in a medium can be written as [5, 12-19]

$$V_{eff}(q, \omega) = \frac{4\pi e^2}{q^2 \epsilon_{total}(q, \omega)} = \frac{v_q}{\epsilon_{total}} \quad (5.32)$$

where  $\epsilon_{total}$  packs up all the mechanisms: direct electron-electron Coulomb interaction, phonon mediated electron-electron interaction (Frohlich), etc. The physical picture is, free electron gas plus perturbations of different level of sophistication. In a heavily doped polar semiconductor mainly comprised of electrons and polar optical phonons, total dielectric function can be written as [2, 5, 10, 12] (superposition approximation of independent polarization components)

$$\epsilon_{total}(q, \omega) = \epsilon_{\infty} - v_q P(q, \omega) + \epsilon_{\infty} \frac{\omega_{LO}^2 - \omega_{TO}^2}{\omega_{TO}^2 - \omega^2} \quad (5.33)$$

where  $\epsilon_{\infty}$  is due to interband transitions at much higher energy than  $\sim$ meV; second term is due to electron gas in conduction band; third term due to phonons. Plugging this form into general result (3.20), TOTAL scattering rate would be obtained, but effects from different mechanisms are not distinguishable. In order to separate them, manipulate the diagram of the total effective interaction [5, 11, 20]

$$V_{eff}(q, \omega) = \frac{v_q}{\epsilon_{total}} = \frac{v_q^{\infty}}{\epsilon_{el}(q, \omega)} + V_{sc-ph}(q, \omega) \quad (5.34)$$

where  $\epsilon_{el}$  is the Coulomb part due to the electron gas only, defined as

$$\varepsilon_{el}(q, \omega) = 1 - v_q^\infty P(q, \omega) \quad (5.35)$$

and  $V_{sc-ph}$  is the "screened phonon part". It is formally named so, but not a mathematically well-defined quantity, which may give rise to "negative damping rate" [13]. In RPA, it contains a phonon self-energy that arises from the polarization of the electron gas,

$$V_{sc-ph} = \frac{V_{ph}}{\varepsilon_{el}^2 (1 - V_{ph} P / \varepsilon_{el})} = \frac{M_q^2}{\varepsilon_{el}^2} D = \frac{M_q^2}{\varepsilon_{el}^2} \frac{D^{(0)}}{1 - M_q^2 D^{(0)} P / \varepsilon_{el}} \quad (5.36)$$

in which the unscreened phonon interaction is

$$V_{ph} = M_q^2 D^{(0)} \quad (5.37)$$

matrix element is ("bare Frohlich")

$$M_q^2 = v_q \frac{\omega_{LO}}{2} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right) \quad (5.38)$$

and the unscreened phonon Green function is

$$D^{(0)} = \frac{2\omega_{LO}}{\omega^2 - \omega_{LO}^2} \quad (5.39)$$

In the previous sections confirmed analytic results on electron-POP scattering, we actually used a straightforward "unscreened approximation" [5, Mahan, section 6.3], by dropping the electron-polarization correction to the phonon Green function D in equation (5.36) (see also Section 5.1 appendix), i.e.

$$V_{sc-ph} \approx \frac{M_q^2}{\varepsilon_{el}^2} D^{(0)} = \dots = \frac{v_q}{\varepsilon_{el}^2} \frac{1}{\varepsilon_\infty \frac{\omega^2 - \omega_{LO}^2}{\omega_{LO}^2 - \omega_{TO}^2}} \quad (5.40)$$

$\varepsilon_{el}$  was replaced by Thomas-Fermi form (static limit of RPA, it is numerically critical that there's no  $\omega$  dependence in  $\varepsilon_{el}$ ). And at poles  $\pm\omega_{LO}$ , it is equivalent to

$$V_{sc-ph} = \frac{v_q}{\epsilon_{el}^2 \left[ \epsilon_\infty \frac{\omega_{LO}^2 - \omega^2}{\omega_{TO}^2 - \omega^2} \right]} = \frac{v_q}{\epsilon_{el}^2 [\epsilon_{LST}]} = \frac{v_q}{\left[ 1 + q_{TF}^2 / q^2 \right]^2 [\epsilon_{LST}]} \quad (5.41)$$

Physically this approximation means, electron cloud only screens matrix elements (the diagram vertex, i.e., the interaction itself), but doesn't alter the LO phonon frequency. The result was equation (5.10).

As a recent new approach to distinguish the effects from different mechanisms, Ref. [21] proposed the "phonon dissipation weight factor"

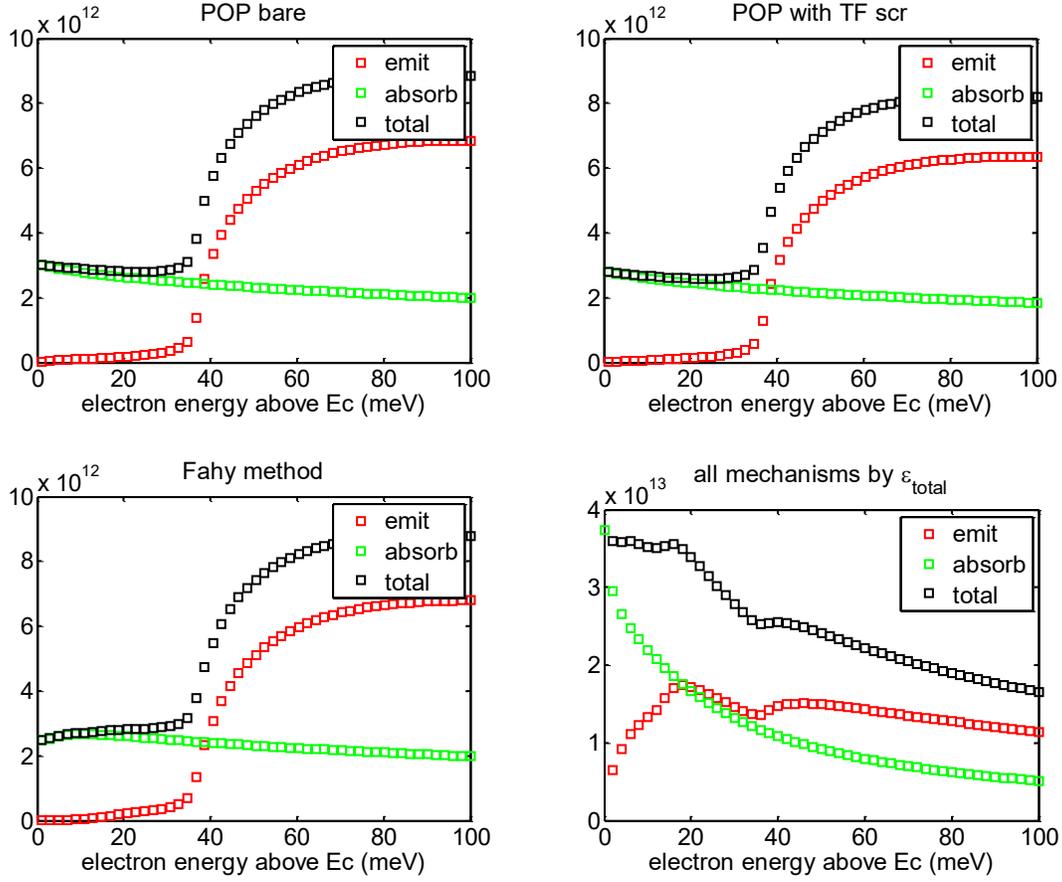
$$factor = \frac{\text{Im}(\epsilon_{phonon})}{\text{Im}(\epsilon_{total})} \quad (5.42)$$

to "pick out" the phonon contribution from the total scattering rate

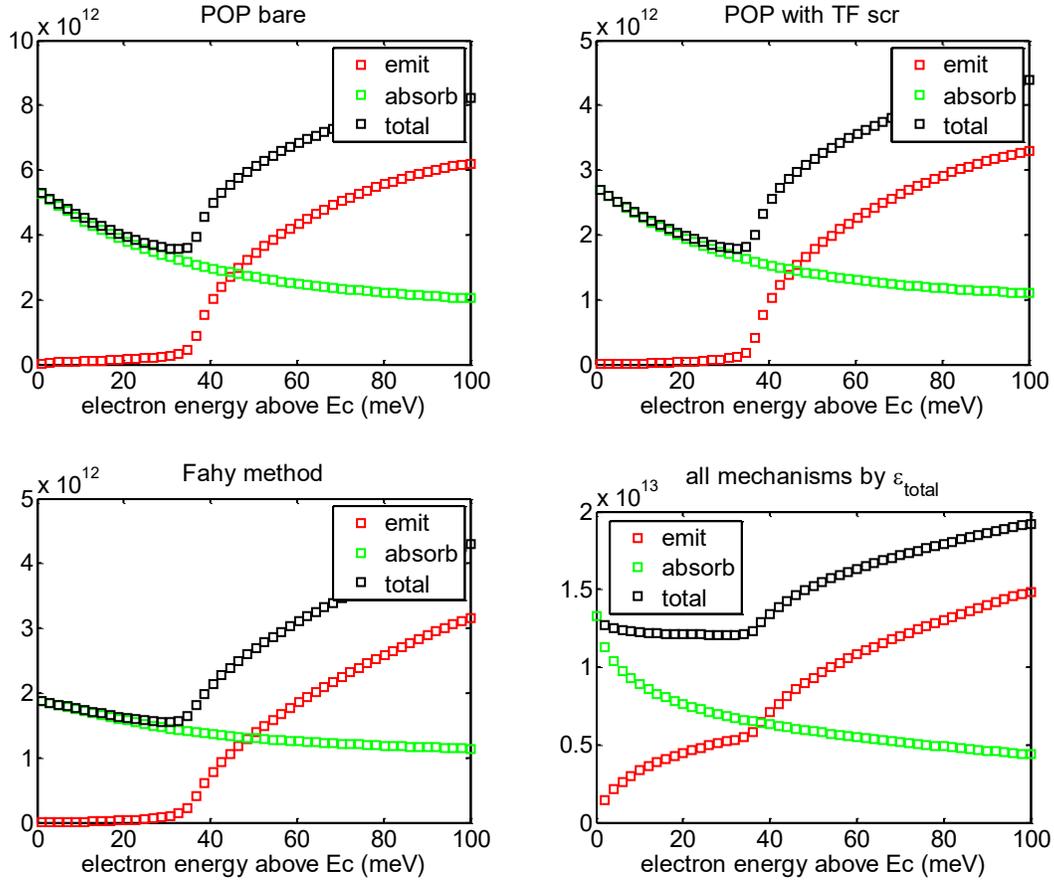
$$phonon\ contribution \sim \frac{\text{Im}(\epsilon_{phonon})}{\text{Im}(\epsilon_{total})} \times \text{Im} \left[ \frac{1}{\epsilon_{total}} \right] \quad (5.43)$$

where  $\epsilon_{total}$  is as in Eq. (5.33). It actually gives quantitatively similar results as Thomas-Fermi screening approach, as shown numerically later. Remarks on Fahy method [21] (equation numbers are as in their paper): (a) It seems the factor (Eq. 31 in [21]) is not only defined for the initial step for the iteration, but also applied to all later steps. In other words, the effective scattering rate (Eq. 33 in [21]) replaces the total scattering rate  $W^0$  in (Eq. 24 in [21]) when the coupled Eqs. (22-24 in [21]) are solved. Otherwise, their method would produce the same huge scattering by electronic modes as seen in my plot of the total scattering rate, especially in the low density cases. (b) If the factor (Eq. 31 in [21]) is not to be applied to all steps in the iteration but only at the initial guess, why bother defining it and elaborating the physical meaning? The iteration method, if robust, should be insensitive to the initial guess any way. (c) My calculation corresponds to Ref. [21]'s

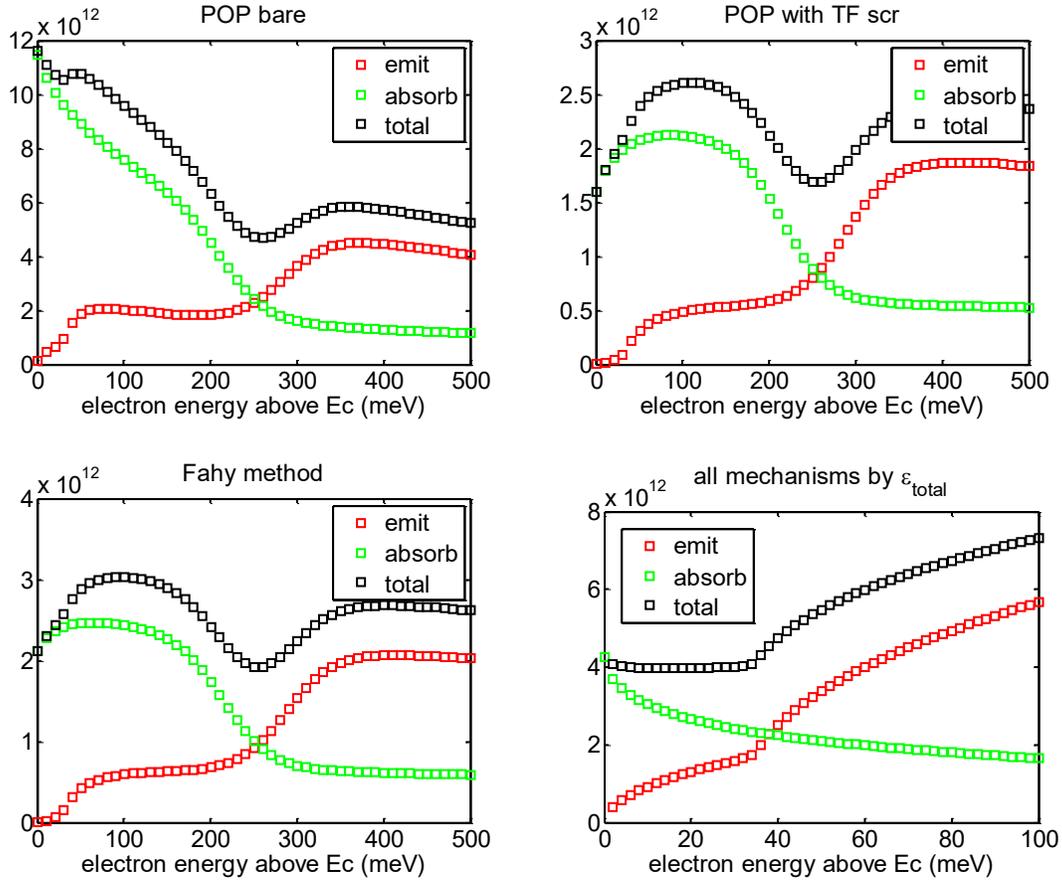
relaxation time approximation (RTA). Their calculations beyond RTA are supposed to be more accurate.



**Fig. 5.2** Numerical results for GaAs system, low density case: electron density  $n = 2.3 \times 10^{16} \text{ cm}^{-3}$ , effective mass = 0.0665, temperature  $T = 300\text{K}$ , chemical potential  $\mu = -2.92 k_B T$ ,  $\epsilon_{inf} = 10.92$ ,  $\omega_{TO} = 33.3 \text{ meV}$ ,  $\omega_{LO} = 36.4 \text{ meV}$ ,  $E_F = 4.4 \text{ meV}$ ,  $\omega_{P0} = 22 \text{ meV}$ ,  $\omega_P = 6.6 \text{ meV}$ . In this case, plasma frequency is much lower than phonons', electron gas is very dilute, so screening is weak. There's not much difference among the "bare", "Thomas-Fermi screening" and "Fahy method".



**Fig. 5.3** Numerical results for GaAs system, medium density case: electron density  $n = 5 \times 10^{17} \text{ cm}^{-3}$ ,  $E_F = 34.5 \text{ meV}$ ,  $\omega_{P0} = 102 \text{ meV}$ ,  $\omega_P = 31 \text{ meV}$ . In this case, plasma frequency is in the vicinity of phonons', Fahy method successfully overcomes the numerical difficulty caused by RPA electronic screening, and gives similar results as Thomas-Fermi screening.



**Fig. 5.4** Numerical results for GaAs system, high density case: electron density  $n = 10^{19}$   $\text{cm}^{-3}$ ,  $E_F = 254.5$  meV,  $\omega_{p0} = 455$  meV,  $\omega_p = 138$  meV. In this case, plasma frequency is much higher than phonons', Fahy method still gives similar results as Thomas-Fermi screening. Since  $\omega_p = 138$  meV is much high, there's no collective electronic mode that electron can be scattered by, and the effect of the electron gas would be just an approximately constant screening. When evaluating RPA around phonon frequency, it is basically the static limit, like Thomas-Fermi (see Chapter 1 discussions).

## 5.6 Effects of adding metallic nanostructures in the "unscreened approximation"

### ---- a simple scheme for modified Thomas-Fermi screening

Based on discussions in Section 5.3, this section talks about the effect of plasmonic nanostructures by adding it to the total epsilon from the beginning, and treat it on equal footing with the conventional Thomas-Fermi screening. We bring in an additional plasmonic resonance  $R$  to the system by adding metallic nanostructures, such that

$$\omega_{LO} \ll R \ll \text{interband transition}$$

From optical simulations, (embedding the metallic nanostructures in a medium where there's only screening from interband  $\epsilon_\infty$  but nothing else)

$$\epsilon(\omega) = \epsilon_\infty + \frac{S^2}{R^2 - \omega^2} \quad (5.44)$$

then, by the same superposition logic [10, 11], the total dielectric function would be

$$\begin{aligned} \epsilon_{total}(q, \omega) &= \epsilon_\infty - v_q P(q, \omega) + \frac{S^2}{R^2 - \omega^2} + \epsilon_\infty \frac{\omega_{LO}^2 - \omega_{TO}^2}{\omega_{TO}^2 - \omega^2} \\ &= \epsilon_\infty + \left[ -v_q P(q, \omega) + \epsilon_{mn}(\omega) \right] + \epsilon_\infty \frac{\omega_{LO}^2 - \omega_{TO}^2}{\omega_{TO}^2 - \omega^2} \end{aligned} \quad (5.45)$$

Square bracket shows the new *electronic contribution*. In static limit, the first term in bracket reduces to Thomas-Fermi, and the second term reduces to a constant  $\epsilon_{mn}(\omega \rightarrow 0) = S^2/R^2$  [Section 5.3]. Use the same "unscreened approximation" (5.40) for the phonon Green's function, equation (5.8) with metallic nanostructures taken into account would become

$$V_{sc-ph} = \frac{v_q}{\left[ 1 + q_{TF}^2/q^2 + \epsilon_{mn}(\omega \rightarrow 0)/\epsilon_\infty \right]^2 \left[ \epsilon_\infty \frac{\omega_{LO}^2 - \omega^2}{\omega_{TO}^2 - \omega^2} \right]} \quad (5.46)$$

Corresponding equation (5.10) would be the same, except for a new form of

$$I_{POP}^{new}(k, F_{\pm}, q_{TF}) \equiv \int_{-1}^1 \frac{1 + F_{\pm}^2/k^2 + 2xF_{\pm}/k}{\left[ \left(1 + F_{\pm}^2/k^2 + 2xF_{\pm}/k\right) \left[1 + \varepsilon_{mn}(\omega \rightarrow 0)/\varepsilon_{\infty}\right] + q_{TF}^2/k^2 \right]^2} dx \quad (5.47)$$

Thus, for the Thomas-Fermi screening method, it only needs a substitution to have the effects of plasmonic nanostructures taken into account

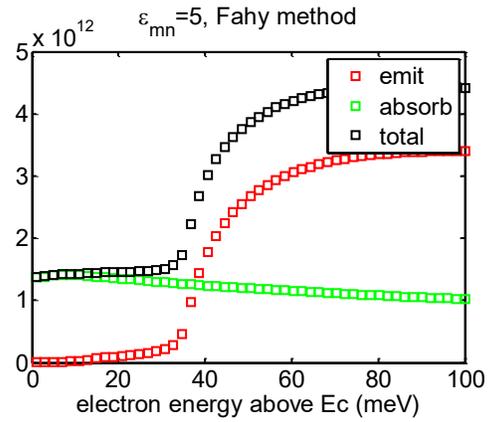
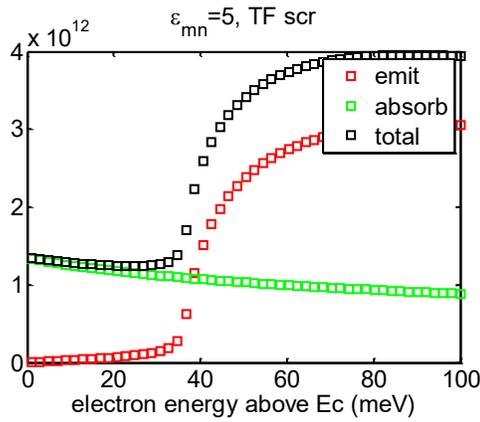
$$\varepsilon_{el} \leftarrow 1 + \frac{q_{TF}^2}{q^2} + \frac{\varepsilon_{mn}}{\varepsilon_{\infty}} \quad (5.48)$$

For the Fahy method [21], it needs a substitution

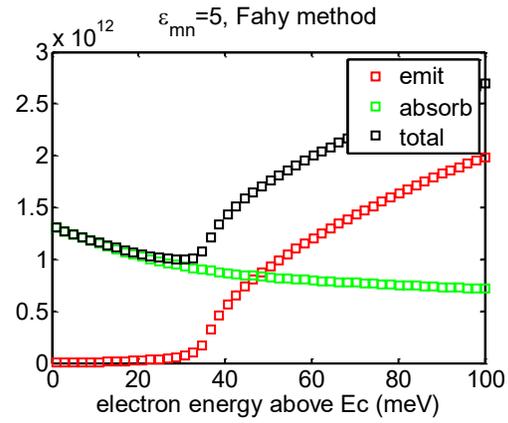
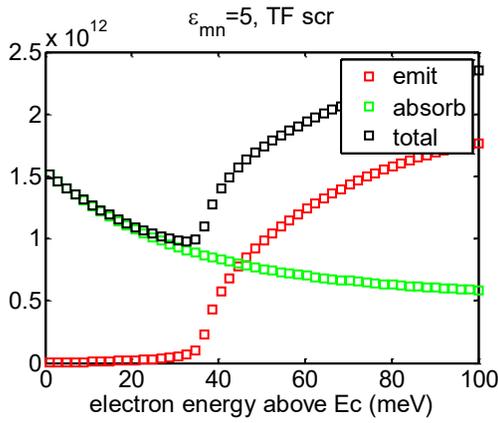
$$\varepsilon_{total} \leftarrow \varepsilon_{total} + \varepsilon_{mn} \quad (5.49)$$

Fig. 5.5 shows the numerical results for GaAs of different carrier densities.

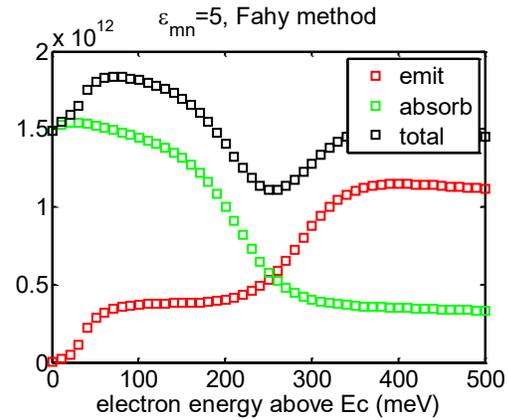
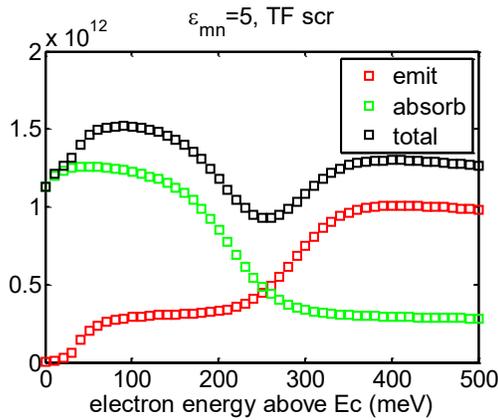
(1) Low density,  $n = 2.3 \times 10^{16} \text{ cm}^{-3}$



(2) Medium density,  $n = 5 \times 10^{17} \text{ cm}^{-3}$



(3) High density,  $n = 1 \times 10^{19} \text{ cm}^{-3}$



**Fig. 5.5** Numerical results for GaAs of different carrier densities after plasmonic nanostructures are embedded, with an example of  $\epsilon_{mn} \approx 5$ .

## Appendices

(1) Derivation detail of equation (5.6):

$$\begin{aligned}
\gamma_k^{\text{eff}} &= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{d^3q}{(2\pi)^3} V_q \left[ n_B(E_{k\mp q} - E_k) + n_F(E_{k\mp q} - \mu) \right] \left[ \delta(\omega - \omega_L) - \delta(\omega + \omega_L) \right] \\
&= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \frac{f \sin \theta d\varphi f d\theta df}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{k^2 + f^2 - 2kf \cos \theta} \left[ \frac{n_B(E_{k\mp q} - E_k)}{+n_F(E_{k\mp q} - \mu)} \right] \left[ \frac{\delta(f - F_+)}{\hbar F_+/m} - \frac{\delta(f - F_-)}{\hbar F_-/m} \right] \\
&= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \left\{ \begin{aligned} & - \frac{F_- \sin \theta d\varphi F_- d\theta}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{k^2 + F_-^2 - 2kF_- \cos \theta} \frac{n_B(-\hbar\omega_L) + n_F(E_k - \hbar\omega_L - \mu)}{\hbar F_-/m} \\ & + \frac{F_+ \sin \theta d\varphi F_+ d\theta}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{k^2 + F_+^2 - 2kF_+ \cos \theta} \frac{n_B(\hbar\omega_L) + n_F(E_k + \hbar\omega_L - \mu)}{\hbar F_+/m} \end{aligned} \right\} \\
&= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \int \left\{ \begin{aligned} & \frac{F_- \sin \theta 2\pi F_- d\theta}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{k^2 + F_-^2 - 2kF_- \cos \theta} \frac{1 + n_B(\hbar\omega_L) - n_F(E_k - \hbar\omega_L - \mu)}{\hbar F_-/m} \\ & + \frac{F_+ \sin \theta 2\pi F_+ d\theta}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{k^2 + F_+^2 - 2kF_+ \cos \theta} \frac{n_B(\hbar\omega_L) + n_F(E_k + \hbar\omega_L - \mu)}{\hbar F_+/m} \end{aligned} \right\} \\
&= \frac{2\pi}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \left\{ \begin{aligned} & \frac{F_- 2\pi F_-}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{2kF_-} \ln \left( \frac{k + F_-}{k - F_-} \right)^2 \frac{1 + n_B(\hbar\omega_L) - n_F(E_k - \hbar\omega_L - \mu)}{\hbar F_-/m} \\ & + \frac{F_+ 2\pi F_+}{(2\pi)^3} \frac{e^2}{\varepsilon_0} \frac{1}{2kF_+} \ln \left( \frac{k + F_+}{k - F_+} \right)^2 \frac{n_B(\hbar\omega_L) + n_F(E_k + \hbar\omega_L - \mu)}{\hbar F_+/m} \end{aligned} \right\} \\
&= \frac{1}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot 2\omega_L} \left\{ \begin{aligned} & \frac{1}{(2\pi)} \frac{e^2}{\varepsilon_0} \frac{1}{2k} \ln \left( \frac{k + F_-}{k - F_-} \right)^2 \frac{1 + n_B(\hbar\omega_L) - n_F(E_k - \hbar\omega_L - \mu)}{\hbar/m} \\ & + \frac{1}{(2\pi)} \frac{e^2}{\varepsilon_0} \frac{1}{2k} \ln \left( \frac{k + F_+}{k - F_+} \right)^2 \frac{n_B(\hbar\omega_L) + n_F(E_k + \hbar\omega_L - \mu)}{\hbar/m} \end{aligned} \right\} \\
&= \frac{1}{\hbar} \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot \omega_L} \left\{ \begin{aligned} & \frac{1}{(4\pi)} \frac{e^2}{\varepsilon_0} \ln \left( \frac{\sqrt{E} + \sqrt{E - \hbar\omega_L}}{\sqrt{E} - \sqrt{E - \hbar\omega_L}} \right) \frac{1 + n_B(\hbar\omega_L) - n_F(E_k - \hbar\omega_L - \mu)}{\sqrt{2mE}/m} \\ & + \frac{1}{(4\pi)} \frac{e^2}{\varepsilon_0} \ln \left( \frac{\sqrt{E + \hbar\omega_L} + \sqrt{E}}{\sqrt{E + \hbar\omega_L} - \sqrt{E}} \right) \frac{n_B(\hbar\omega_L) + n_F(E_k + \hbar\omega_L - \mu)}{\sqrt{2mE}/m} \end{aligned} \right\} \\
&= \frac{\hbar\omega_L}{\hbar^2} \left( \frac{\omega_L^2 - \omega_T^2}{\varepsilon_\infty \cdot \omega_L^2} \right) \frac{e^2}{4\pi\varepsilon_0} \sqrt{\frac{m}{2E}} \left\{ \begin{aligned} & \ln \left( \frac{\sqrt{E} + \sqrt{E - \hbar\omega_L}}{\sqrt{E} - \sqrt{E - \hbar\omega_L}} \right) \left[ 1 + n_B(\hbar\omega_L) - n_F(E_k - \hbar\omega_L - \mu) \right] \\ & + \ln \left( \frac{\sqrt{E + \hbar\omega_L} + \sqrt{E}}{\sqrt{E + \hbar\omega_L} - \sqrt{E}} \right) \left[ n_B(\hbar\omega_L) + n_F(E_k + \hbar\omega_L - \mu) \right] \end{aligned} \right\}
\end{aligned}$$

(2) Two-term Lorentzian details:

$$\begin{aligned}
\varepsilon(\omega) &= \varepsilon_\infty + \frac{S_1^2}{R_1^2 - \omega^2} + \frac{S_2^2}{R_2^2 - \omega^2} \\
&= \frac{\varepsilon_\infty (R_1^2 - \omega^2)(R_2^2 - \omega^2) + S_1^2 (R_2^2 - \omega^2) + S_2^2 (R_1^2 - \omega^2)}{(R_1^2 - \omega^2)(R_2^2 - \omega^2)} \\
&= \frac{\varepsilon_\infty (\omega^2 - P_1^2)(\omega^2 - P_2^2)}{(\omega^2 - R_1^2)(\omega^2 - R_2^2)}
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\varepsilon(\omega)} &= \frac{(\omega^2 - R_1^2)(\omega^2 - R_2^2)}{\varepsilon_\infty (\omega^2 - P_1^2)(\omega^2 - P_2^2)} = \frac{(\omega^2 - R_1^2)(\omega^2 - R_2^2)}{\varepsilon_\infty (P_1^2 - P_2^2)} \left[ \frac{1}{(\omega^2 - P_1^2)} - \frac{1}{(\omega^2 - P_2^2)} \right] \\
&= \frac{(\omega^2 - R_1^2)(\omega^2 - R_2^2)}{\varepsilon_\infty (P_1^2 - P_2^2)} \left[ \frac{1}{2P_1} \left( \frac{1}{\omega - P_1} - \frac{1}{\omega + P_1} \right) - \frac{1}{2P_2} \left( \frac{1}{\omega - P_2} - \frac{1}{\omega + P_2} \right) \right]
\end{aligned}$$

Retarded extension, using results above

$$\text{Im} \left[ \frac{1}{\varepsilon(\omega)} \right] = \frac{(\omega^2 - R_1^2)(\omega^2 - R_2^2)}{\varepsilon_\infty (P_1^2 - P_2^2)} \left[ \frac{(-\pi)}{2P_1} [\delta(\omega - P_1) - \delta(\omega + P_1)] - \frac{(-\pi)}{2P_2} [\delta(\omega - P_2) - \delta(\omega + P_2)] \right]$$

Reducing further

$$\begin{aligned}
\text{Im} \left[ \frac{1}{\varepsilon(\omega)} \right] &= \frac{(P_1^2 - R_1^2)(P_1^2 - R_2^2)}{\varepsilon_\infty (P_1^2 - P_2^2)} \frac{(-\pi)}{2P_1} [\delta(\omega - P_1) - \delta(\omega + P_1)] \\
&\quad + \frac{(P_2^2 - R_1^2)(P_2^2 - R_2^2)}{\varepsilon_\infty (P_2^2 - P_1^2)} \frac{(-\pi)}{2P_2} [\delta(\omega - P_2) - \delta(\omega + P_2)]
\end{aligned}$$

$P_1^2$  and  $P_2^2$  are roots of

$$\varepsilon_\infty (R_1^2 - \omega^2)(R_2^2 - \omega^2) + S_1^2 (R_2^2 - \omega^2) + S_2^2 (R_1^2 - \omega^2) = 0$$

re-organize

$$\begin{aligned}
&\varepsilon_\infty (R_1^2 - \omega^2)(R_2^2 - \omega^2) + S_1^2 (R_2^2 - \omega^2) + S_2^2 (R_1^2 - \omega^2) \\
&= \varepsilon_\infty \omega^4 - \left[ \varepsilon_\infty (R_1^2 + R_2^2) + S_1^2 + S_2^2 \right] \omega^2 + \left[ \varepsilon_\infty R_1^2 R_2^2 + S_1^2 R_2^2 + S_2^2 R_1^2 \right] \\
&= \varepsilon_\infty \omega^4 - B \cdot \omega^2 + C
\end{aligned}$$

thus,

$$P_{1,2}^2 = \frac{B \mp \sqrt{B^2 - 4\varepsilon_\infty C}}{2\varepsilon_\infty}$$

where

$$B = \varepsilon_\infty (R_1^2 + R_2^2) + S_1^2 + S_2^2$$

$$C = \varepsilon_\infty R_1^2 R_2^2 + S_1^2 R_2^2 + S_2^2 R_1^2$$

Further approximations can be made.

(3) Attempt to resolve the “N<sub>0</sub> / N<sub>0</sub>+1 inconsistency” in an earlier calculation, following Prof. Broido’s derivations [6], equation numbers are that of his notes2.

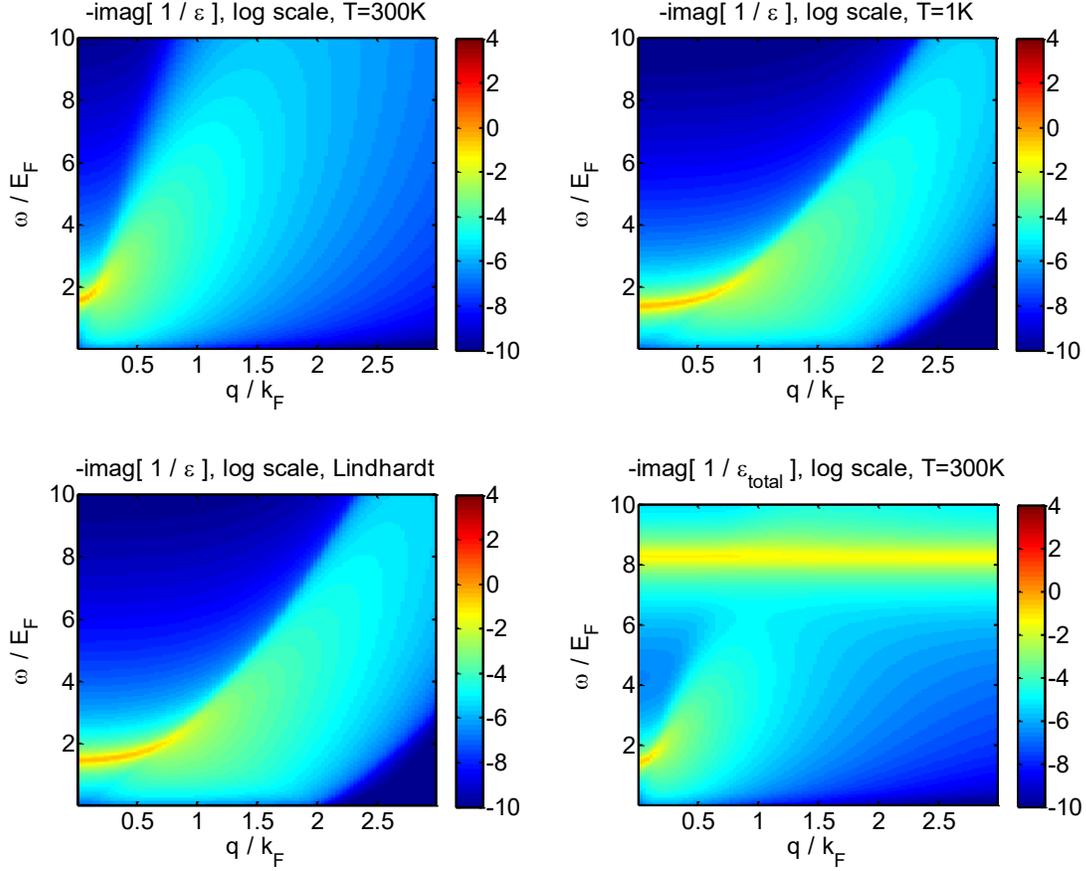
$$\mathbf{v} \cdot \left( -e\mathbf{E} \frac{\partial f}{\partial \varepsilon} + \frac{\partial f}{\partial T} \nabla T \right) = \frac{\partial f}{\partial t} \Big|_{coll} \quad (1)$$

$$\frac{\partial f}{\partial t} \Big|_{coll} = \sum_{\mathbf{k}'} [f_{\mathbf{k}'}(1 - f_{\mathbf{k}})S(\mathbf{k}', \mathbf{k}) - f_{\mathbf{k}}(1 - f_{\mathbf{k}'})S(\mathbf{k}, \mathbf{k}')] \quad (3)$$

$$f_0(\varepsilon)(1 - f_0(\varepsilon'))S(\mathbf{k}, \mathbf{k}') = f_0(\varepsilon')(1 - f_0(\varepsilon))S(\mathbf{k}', \mathbf{k}) \quad (4a)$$

$$f_0(\varepsilon)(1 - f(\varepsilon'))S(\mathbf{k}, \mathbf{k}') = f_0(\varepsilon')(1 - f(\varepsilon))S(\mathbf{k}', \mathbf{k}) \quad (4b)$$

Equation (4a) is what implied by equilibrium distribution. Substituting it in Boltzmann equation will produce of course same value for out-scattering and in-scattering, as my calculation scheme does. Delicacy of the Boltzmann equation was to solve self-consistently for the steady distribution  $f$  (not equilibrium distribution  $f_0$ ), by using (4b). Remark: terms involving Bose-Einstein distribution factors N<sub>0</sub>/N<sub>0</sub>+1, root back from the phonon creation and annihilation operations. [7, Nag page103]



**Fig. 5.6** Temperature dependent RPA epsilon of homogeneous electron gas: Refer to [22], these are plots of the GaAs low density case, electron density  $n = 2.3 \times 10^{16} \text{ cm}^{-3}$ , effective mass = 0.0665, temperature  $T = 300 \text{ K}$ , chemical potential  $\mu = -2.92 k_B T$ ,  $\epsilon_{\text{inf}} = 10.92$ ,  $\omega_{\text{TO}} = 33.3 \text{ meV}$ ,  $\omega_{\text{LO}} = 36.4 \text{ meV}$ ,  $E_F = 4.4 \text{ meV}$ ,  $\omega_{\text{P0}} = 22 \text{ meV}$ ,  $\omega_{\text{P}} = 6.6 \text{ meV}$ . We can see that the strong mode in Lindhard colormap locates at  $1.5E_F$ .

(4) Discussions on high temperature limit:

In high temperature approximation  $k_B T \gg \Delta E$

$$\left[ n_B(E_{k-q} - E_k) + n_F(E_{k-q} - \mu) \right] \approx \frac{k_B T}{E_{k-q} - E_k}$$

Substituting in Eq. (3.20) yields

$$\gamma_k \approx \frac{\sqrt{2}e^2 k_B T \sqrt{m}}{\hbar^2 \sqrt{E_0}} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon(0)} \right)$$

where  $E_0 = \hbar\omega_L$ , and  $\varepsilon(0) = \varepsilon_\infty \frac{\omega_L^2}{\omega_T^2}$  is the low frequency limit value. It shows a scaling

with temperature, as other results in the literature.

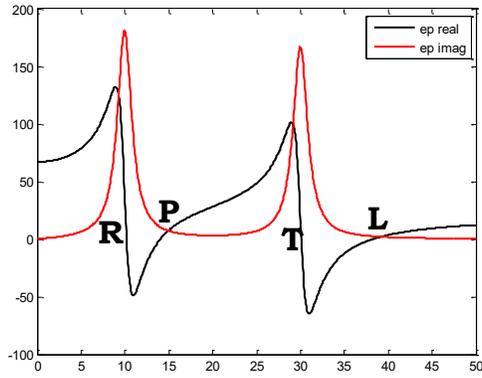
Further if one adds on the electronic component, e.g., an plasmonic excitation, making the (local) response of the medium as

$$\begin{aligned} \varepsilon(q, \omega) &= \varepsilon_{el}(q, \omega) + \varepsilon_{ph}(q, \omega) = \varepsilon_b + \frac{\omega_p^2}{\omega_r^2 - \omega^2} + \varepsilon_\infty \frac{\omega_L^2 - \omega^2}{\omega_T^2 - \omega^2} \\ &= \varepsilon'_\infty + \frac{\omega_p^2}{\omega_r^2 - \omega^2} + \frac{\varepsilon_\infty (\omega_L^2 - \omega_T^2)}{\omega_T^2 - \omega^2} \end{aligned}$$

Substituting in Eq. (3.20), yields the total electron scattering by both electron(plasmon) and phonon excitations, in high temperature approximation

$$\gamma_k^{eff} = \frac{\sqrt{2}e^2 k_B T \sqrt{m}}{\hbar^2 \varepsilon'_\infty} \left[ \frac{(P^2 - R^2)(P^2 - T^2)}{(P^2 - L^2)P^{5/2}} + \frac{(L^2 - R^2)(L^2 - T^2)}{(L^2 - P^2)L^{5/2}} \right]$$

where R P T L are sketched as below. (Numbers are not to scale; Values of P and L could be expressed.)



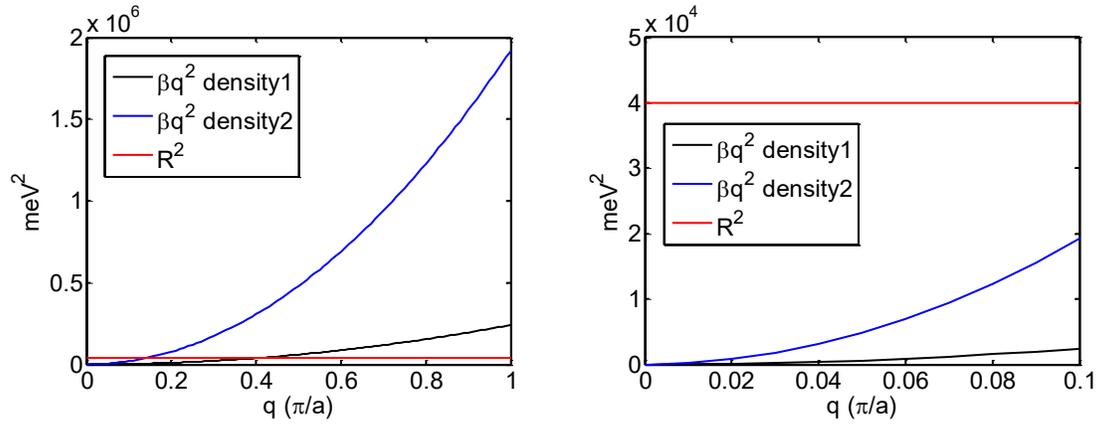
**Fig. 5.7** Schematic of real and imaginary parts of the total dielectric function, when adding a plasmonic resonance (R, P) whose energy is below that of the phonon resonance (T, L) to the polar semiconductors (numbers are arbitrary).

(5) A brief check of the nonlocal effects from HDA point of view:

This notes addresses the nonlocal effects in the response of the added plasmonic resonance (metallic nanostructure) from the hydrodynamic approximation. A usual way of extension [Chapter 1 and Chapter 4] is

$$\varepsilon_{mn}(\omega) = \frac{S^2}{R^2 - (\omega^2 - \beta q^2)} , \quad \beta = \frac{3}{5} v_F^2$$

where  $v_F$  is Fermi velocity which in our system corresponds to the carrier density. Plots in Fig. 5.8 show comparison of the  $\beta q^2$  term with a plasmonic resonance  $R$  at 200meV. Fig. 5.8 left shows  $q$  range up to the Brillouin zone boundary (lattice constant assumed 5 Angstrom). Fig. 5.8 right zooms in to the  $q$  range of the electron gas in the system. Note that,  $k_F$  for  $n_1 = 2.3 \times 10^{16} \text{ cm}^{-3}$  is about  $9 \times 10^7 / \text{m} = 0.014 \pi / a$ ,  $k_F$  for  $n_2 = 5 \times 10^{17} \text{ cm}^{-3}$  is about  $2.5 \times 10^8 / \text{m} = 0.04 \pi / a$ . Most of the carriers are distributed not too far away from  $k_F$ , making typical momentum transfer in a scattering event within  $2k_F$ . We see from the right plot that the  $\beta q^2$  term is smaller than 10% of the  $R^2$  term, thus safe to neglect.



**Fig. 5.8** Comparison of the  $\beta q^2$  term with a plasmonic resonance  $R$  at 200meV. Left shows  $q$  range up to the Brillouin zone boundary (lattice constant assumed 5 Angstrom). Right zooms in to the  $q$  range of the electron gas in the system.

(6) Other forms of approximation more accurate than the "unscreened approximation", comments on Ref. [9]: In the initial general derivation, full version RPA electronic part was used, but in later numerical carry-out, static-RPA, i.e. Thomas-Fermi-like form was used. That's why there's no huge scattering by the plasmon-like mode as can be seen in my results by full RPA electronic part, especially in the cases of low carrier densities. A numerical fact is, as long as the electronic part has no  $\omega$ -dependence, we are fine, even using the full version Frohlich term including the phonon self-energy correction. As also discussed/confirmed by Das Sarma [13, 15], poles of  $1/\epsilon^2$  are much more complicated than poles of  $1/\epsilon$ .

## References:

- [1] N. W. Ashcroft and N. D. Mermin, Solid State Physics, Brooks/Cole Cengage Learning (1976), chapter 27
- [2] B. B. Varga, Phys. Rev. 137, A1896 (1965)
- [3] H. Ibach and H. Luth, Solid-State Physics, 4<sup>th</sup> edition, Springer (2009)
- [4] J. O. Sofo and G. D. Mahan, Phys. Rev. B 49, 4565 (1994)
- [5] G. D. Mahan, Many-Particle Physics, Plenum Press, New York (1990), chapter 6 has a clear discussion
- [6] Notes from Prof. David Broido, Boston College
- [7] B. R. Nag, Electron Transport in Compound Semiconductors, Springer (1980)
- [8] J. M. Ziman, Electrons and Phonons, Oxford University Press (1962)
- [9] B. A. Sanborn, Phys. Rev. B 51, 14247 (1995), and references therein
- [10] H. Ehrenreich and H. R. Philipp, Phys. Rev. 128, 1622 (1962)
- [11] D. Pines, Elementary Excitations in Solids, Benjamin, New York (1964)
- [12] M. E. Kim, A. Das and S. D. Senturia, Phys. Rev. B 18, 6890 (1978)
- [13] R. Jalabert and S. Das Sarma, Phys. Rev. B 41, 3651 (1990)
- [14] A. F. J. Levi, J. R. Hayes, P. M. Platzman and W. Wiegmann, Phys. Rev. Lett. 55, 2071 (1985)
- [15] B. Y. K. Hu and S. Das Sarma, Phys. Rev. B 44, 8319 (1991)
- [16] J. F. Young and P. J. Kelly, Phys. Rev. B 47, 6316 (1993)
- [17] P. Nozieres and D. Pines, Theory of Quantum Liquids, Advanced Books Classic Series, Westview, Boulder (1999)
- [18] H. Nyquist, Phys. Rev. 32, 110 (1928)
- [19] H. B. Callen and T. A. Welton, Phys. Rev. 83, 34 (1951)
- [20] R. Mattuck, A Guide to Feynman Diagrams in the Many-Body Problem, 2<sup>nd</sup> edition, Dover (1992)
- [21] A. Hauber and S. Fahy, Phys. Rev. B 95, 045210 (2017)
- [22] J. Cen, “Current driven plasma instabilities in layered solid state systems”, Ph.D. dissertation, Boston College (1991)

## CHAPTER 6

### Plasmon-Enhanced Multiple Exciton Generation

**Outline:** This chapter extends the general results of electron lifetime in a many-particle system obtained in Chapter 3, combines with the idea of HELPP discussed in Chapter 4, and applies them in a project proposing a new approach for multiple exciton generation (MEG) for solar cells. It is shown that bi-exciton formation can be highly efficient in a solar cell with the semiconductor absorber filled with an array of metallic nanoparticles having plasmonic resonance tuned to the semiconductor gap energy. This process can be viewed as plasmon-enhanced multiple exciton generation (PMEG), with the resulting cell efficiency exceeding the Shockley–Queisser limit. It is demonstrated that efficiency of the PMEG process increases with decreasing of the semiconductor gap size, by considering in detail three systems with gradually decreasing gap size: GaAs, Si and Ge. This project was concluded as a preprint on arXiv, and is being submitted to journals: Jiantao Kong, Xueyuan Wu, Xin Wang, Michael J. Naughton, and Krzysztof Kempa, “Plasmonic multiple exciton generation”, arXiv (2018), [arxiv.org/abs/1806.10259](https://arxiv.org/abs/1806.10259)

## 6.1 Introduction

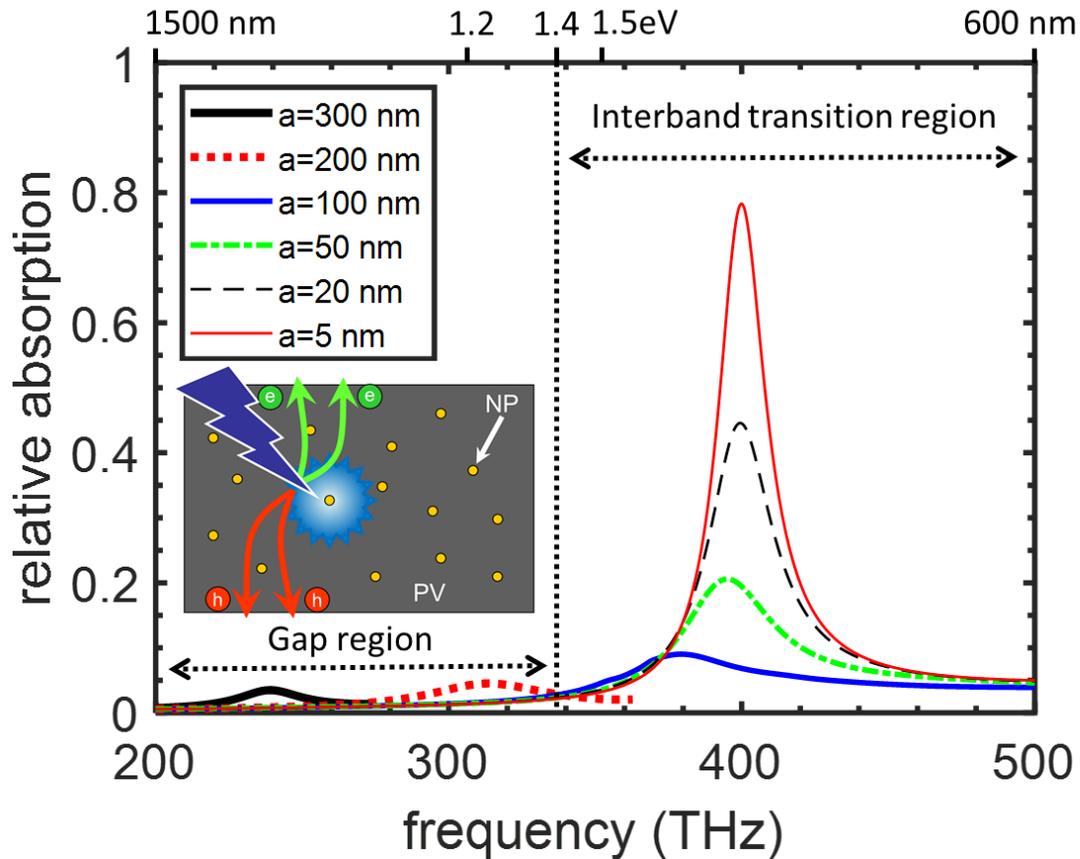
Electrons or holes in semiconductors, excited into their respective conduction and valence bands away from the thermal equilibrium distributions, are referred to as “hot” (see also discussions in Chapter 4). Effects of hot electrons have been studied and utilized for more than half a century in a variety of electronic devices, from Gunn diodes to integrated circuits [1-10]. In conventional solar cells, hot electrons rapidly and irreversibly lose their excess or “hot” energy to phonons (heat), which leads to the Shockley-Queisser limit for single junction cell efficiency [11]. The amount of the energy lost to heat in a conventional solar cell actually exceeds that harvested in the form of usable electricity. For example, commercially available, high efficiency crystalline silicon solar cells convert 20-25% of absorbed sunlight into electricity, but more than 30% into heat via hot electrons. Many concepts have been proposed to harvest or convert this hot electron energy into usable form, but none have been experimentally verified or demonstrated to date [11]. One of the seminal concepts proposed for so called 3<sup>rd</sup>-generation solar photovoltaics (PV) involves harvesting the excess energy of these hot electrons before it is dissipated as heat [12], with theoretical efficiency limits of over 60%. This is posited to be achievable by first somehow eliminating the phonon scattering in the active region, and then extracting the hot electrons through narrow band energy filters at absorber-electrode contacts, assuring isentropic cooling. However, this is far from a trivial proposition, and no successful solar cell based on this idea has been developed. While early investigations found some evidence for hot electron injection into an electrolyte [13], and the hot electron contribution to the photovoltage was recently demonstrated [14], there remains limited experimental evidence of improved photovoltaic performance via hot electrons, despite many decades of research.

In another important scheme to recover the hot electron energy, it was envisioned that a single photon in a solar cell could generate two or more electron-hole pairs (physically-separated excitons), instead of the canonical single pair. This is the multi-exciton generation (MEG) concept [15, 16, 17], known to be vanishingly small in bulk materials in the frequency range of interest to photovoltaics. It has been demonstrated in laser spectroscopic [16, 17] and photocurrent [18] studies that, in semiconductor nanoparticles (NPs), it can become significant.

Recently, a plasmonic, 3<sup>rd</sup> generation PV scheme is proposed, by providing an efficient energy-dissipation channel into plasmons in an adjacent or embedded plasmonic structure [19] (see also Chapter 4). In this scheme, the hot electron free energy remains reversibly “protected” in a collective electronic degree of freedom. This hot electron plasmon protection (HELPP) mechanism, which relies on electron-plasmon scattering occurring on a time scale sufficiently smaller than phonon emission by either plasmons or hot electrons, was theoretically supported by a simple model calculation [19] (see also Chapter 3 Section 4). Here in this chapter, we describe a way to combine the HELPP concept with MEG, a process which can be viewed as plasmon-enhanced multiple exciton generation (PMEG).

## 6.2 Scattering rate calculation scheme revisited

The MEG theory often breaks the process into two steps: first, an incoming photon excites a single exciton, with hot carriers participating; second, this exciton, before emitting phonons, decays into multiple excitons via Coulomb scattering [20]. Instead of employing Fermi's golden rule to estimate the decay rate of excitons (hot electrons and holes) to bi-excitons, we calculate the hot electron scattering rate exactly, including secondary excitons as a part of the single particle excitation continuum. According to discussions in Chapter 3, the scattering rate of an electron in a semiconductor matrix from a state  $E_{\mathbf{k}}$  to all other possible states due to single particle and collective (plasmon) excitations, in RPA [21] is given by Eq. (3.20). This calculation requires knowledge of the effective dielectric function of a given structure. In a simple, single Lorentzian approximation, the dielectric function can be written as Eq. (3.26) [22] (see Chapter 3 Appendix for response function of such dielectric form). In Section 3.4 we show that the single Lorentzian form leads to a simple analytic formula for the scattering rate (3.28) [23]. Eq. (3.28) can be used as guidance for more rigorous calculations/simulations, and it shows, as expected, that the scattering vanishes for electron energies lower than the localized-plasmon energy  $E_{\mathbf{k}} < \hbar\omega_p$ , and also that it increases rapidly with increasing plasmonic oscillator strength  $\omega_p$ .



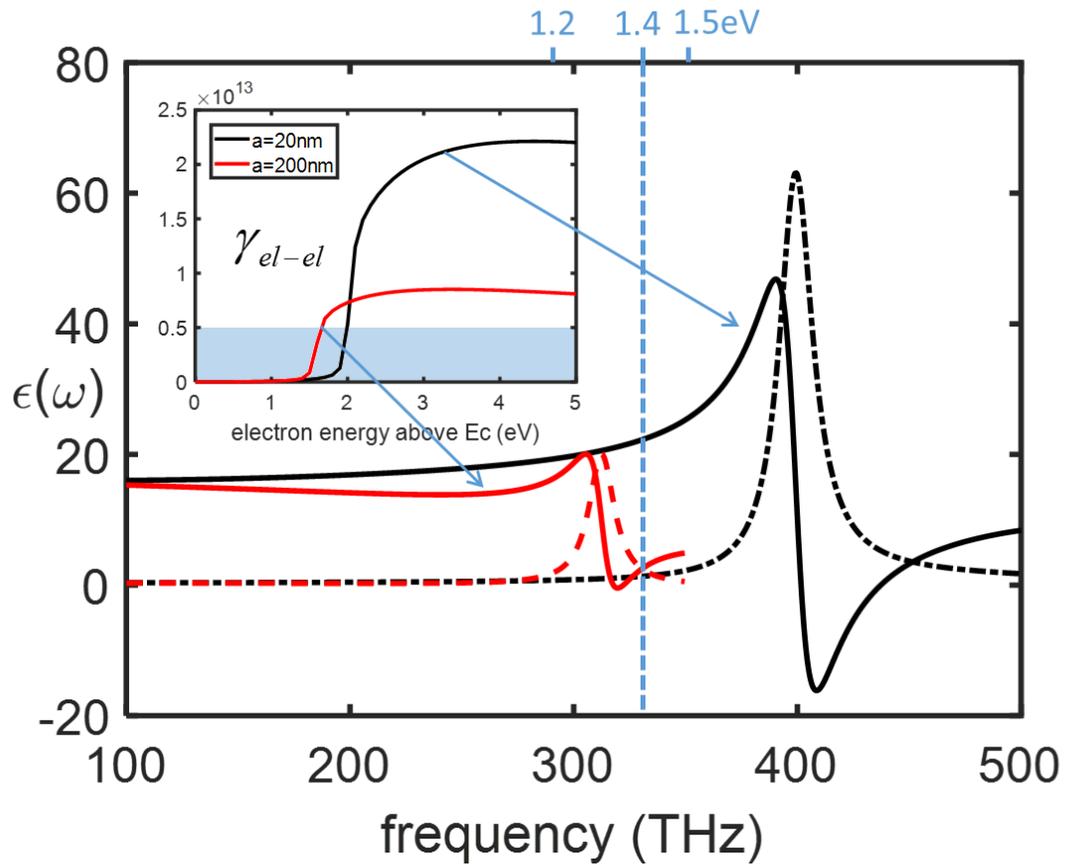
**Fig. 6.1** Relative (per unit volume of the absorber) absorption spectra of GaAs absorber filled with cubic array of Ag NPs (period  $a$ , diameter  $D = a/3$ ). Strong plasmonic absorbance is seen around 400 THz. Inset shows basic PMEG scheme: incident high energy photon (dark blue) interacts with NP to establishes a surface plasmon, whose strong  $E$ -field (gradient blue) excites a bi-exciton, which separates into two electron-hole pairs, which drift/diffuse via an inferred  $p$ - $n$  junction. Top axis shows relevant energy and wavelength scales. This figure is from: J. Kong, X. Wu, X. Wang, M. J. Naughton, and K. Kempa, “Plasmonic multiple exciton generation”, arXiv (2018), arxiv.org/abs/1806.10259.

### 6.3 Simulation and calculation results

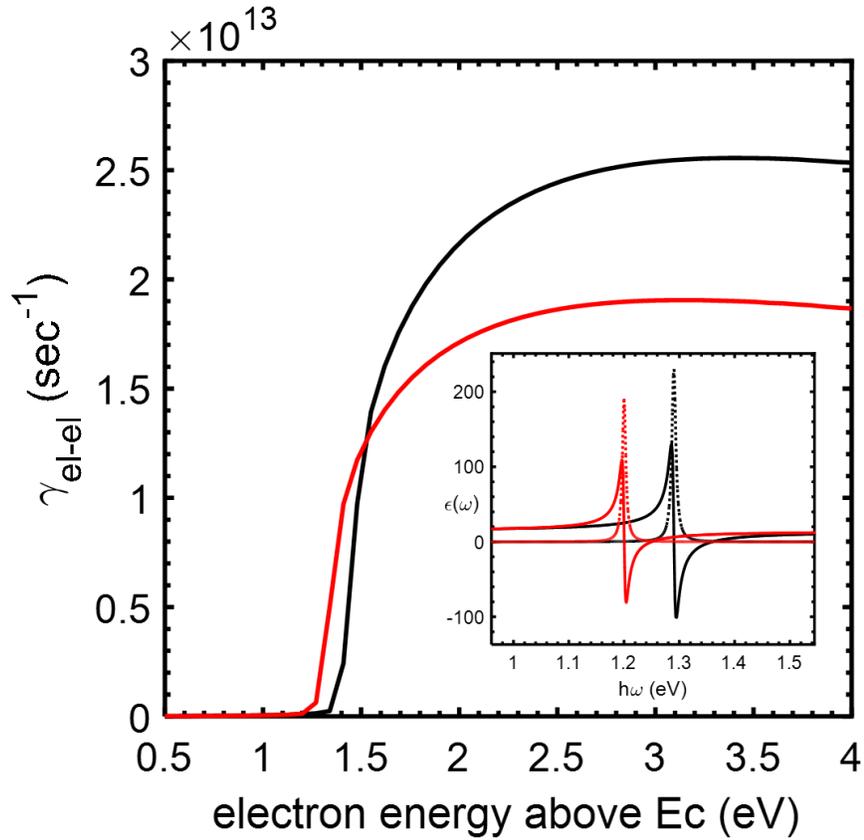
Consider now a PV absorber filled with an array of simple spherical metal NPs (*i.e.* nanospheres), as depicted in the inset to Fig. 6.1. We chose the NPs to be in a cubic lattice of period  $a$  and the nanosphere diameter  $D = a/3$ , so that the projected area fraction remains unchanged as we vary  $a$ . The relative absorption (per unit volume of the absorber), as simulated in FDTD [24, 25] for crystalline GaAs semiconductor and Ag NPs, is shown in Fig. 6.1, for six values of  $a$ .

Fig. 6.1 shows that the frequency of the plasmonic absorption increases with decreasing  $a$ , and saturates  $\sim 400$  THz. This behavior reflects the well-known dispersion relation of a surface plasmon induced on the surface of the metallic sphere; changing the sphere diameter changes an effective surface plasmon quasi-momentum according to the “whispering gallery” mode condition [26, 22]  $q \approx 2/D$  (see discussions on this  $q$ -dependence in Chapter 2). The plasmonic absorption peak strengths rapidly increases once the peak frequency enters the intersubband transition region above the gap energy of 1.4 eV ( $\sim 340$  THz). In this region, massive generation of interband transitions (*i.e.* excitons) by decaying hot electrons is also expected, as will be demonstrated below. The absorption spectrum for each value of  $a$  is dominated by a single plasmonic resonance, and so one could use Eq. (3.26) as a simple model of the dielectric function, and then use Eq. (3.28) to estimate of the scattering rate. For an accurate analysis, we extract the effective dielectric function of the medium by the method described in detail in [26], and then use the exact result from Eq. (3.20) to obtain the scattering rate. The extracted single Lorentzian

dielectric functions for  $D = 67$  nm and 6.7 nm are shown in Fig. 6.2. The inset shows the corresponding scattering rates vs. hot electron energy. For the smaller spheres, intersubband transitions are possible (producing secondary excitons), and the scattering rates of hot electrons with energies 2.5 eV and more above the conduction band edge exceed  $2 \times 10^{13} \text{ s}^{-1}$  (*i.e.* faster than 50 ps). This rate is larger than the phonon cooling rate in GaAs of  $\sim 0.5 \times 10^{13} \text{ s}^{-1}$  (*i.e.* 200 ps) [27]. This is the rate at which the hot electrons cool to the bottom of the conduction band, which requires many electron-phonon scattering events; the energy of a single phonon is only  $\sim 36$  meV, such that more than 50 scattering events are needed to completely cool a hot electron with energy 2 eV. The shaded area in the inset in Fig. 6.2 shows an estimated cooling rate. For larger spheres ( $D = 67$  nm), with resonances below the energy gap, no secondary excitons are generated, only plasmons at a smaller rate.



**Fig. 6.2** Extracted effective dielectric function of the GaAs absorber filled with a cubic array of Ag nanospheres (each with diameter  $D = a/3$ ) for two nanosphere sizes  $D = 6.7$  nm (black), and  $D = 67$  nm (red). The inset shows the corresponding electron- electron scattering rates. The shaded area represents the rates of electron-phonon scattering processes. This figure is from: J. Kong, X. Wu, X. Wang, M. J. Naughton, and K. Kempa, “Plasmonic multiple exciton generation”, arXiv (2018), arxiv.org/abs/1806.10259.



**Fig. 6.3** Calculated electron-electron scattering rates for a crystalline Si absorber filled with a cubic array of Ag nanospheres (of diameter  $D = a/3$ ), with  $a=200$  nm (black curve) and  $a =230$  nm (red curve). The inset shows the corresponding extracted effective dielectric functions, used to obtain the scattering rates. This figure is from: J. Kong, X. Wu, X. Wang, M. J. Naughton, and K. Kempa, “Plasmonic multiple exciton generation”, arXiv (2018), [arxiv.org/abs/1806.10259](https://arxiv.org/abs/1806.10259).

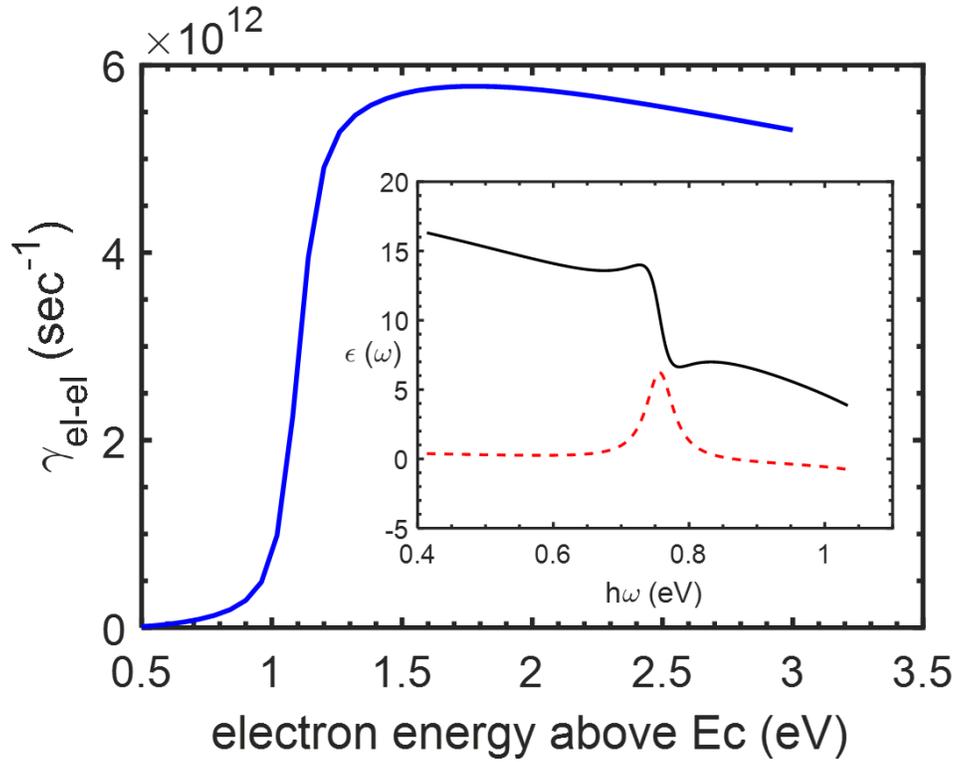
## 6.4 Dependence on the gap size

As the efficiency of PMEG diminishes with increasing gap size, only hot electrons with energy greater than the gap can generate secondary excitons. In fact, GaAs is not an optimal material for PMEG solar cells. The maximum value of the hot electron energy generated by 1AM1.5 solar radiation (as measured from the top of the valence band) is about 3.4 eV [28], and so we estimate that in GaAs, the hot electrons reach only about  $3.4 \text{ eV} - 1.4 \text{ eV} = 2 \text{ eV}$  into the conduction band. However, Fig. 6.2 shows that significant (exceeding the phonon scattering rate) plasmon generation occurs for hot electrons with energy  $> 2 \text{ eV}$ , such that only a small fraction of photo-generated hot electrons can generate secondary excitons. Nevertheless, GaAs is a good material to demonstrate the PMEG effect by using laser illumination.

Next, we investigated crystalline Si. Employing the same procedure as for GaAs, we obtained the result shown in Fig. 6.3. The scattering rates are shown in the main part of the figure, for two NP diameters,  $D = 67$  and  $76 \text{ nm}$ . In this case, we have the solar radiation-induced hot electron bandwidth equal to  $3.4 \text{ eV} - 1.1 \text{ eV} = 2.3 \text{ eV}$ . For the larger diameter sphere, we obtain a significant scattering rate ( $\sim 1.5 \times 10^{13} \text{ sec}^{-1}$ ) already for  $1.3 \text{ eV}$ , which exceeds that of the electron-phonon cooling rate ( $< 10^{13} \text{ sec}^{-1}$ ). Thus, in this case, a reasonably large portion of the hot electron distribution,  $\sim 43\%$ , is available for PMEG recovery. Thus, crystalline Si is a viable material for both PMEG demonstration and a PMEG solar cell.

Semiconductors with even smaller gaps, such as Ge (0.68 eV) or InAs (0.32 eV), should further improve the efficiency of PMEG. As an example, we consider Ge in Fig.

6.4, for NPs with  $D = 33.3$  nm. The scattering rate has a maximum near 1.5 eV, representing the PMEG. Since in this case the range of hot electrons induced by a 1-sun illumination is  $3.4 \text{ eV} - 0.7 \text{ eV} = 2.7 \text{ eV}$  (as measured from the bottom of the conduction band), a large fraction of hot electrons (more than 50%), with energies ranging from 1.3 eV to 2.7 eV, can produce the secondary electrons. The electron-phonon scattering rate in Ge is  $\sim 10^{14} \text{ sec}^{-1}$  [29], and the corresponding cooling rate (in view of the single phonon emission energy of  $\sim 20 \text{ meV}$  [30]) is  $\sim 10^{12} \text{ sec}^{-1}$ , much lower than the electron-electron scattering rate. Thus, we can conclude that Ge could be used as a practical platform for PMEG cells.



**Fig. 6.4** Calculated electron- electron scattering rates for a Ge absorber filled with a cubic array of Ag nanospheres (each with diameter  $D = a/3 = 33.3$  nm). The inset shows the corresponding extracted effective dielectric function, used to obtain the scattering rates. This figure is from: J. Kong, X. Wu, X. Wang, M. J. Naughton, and K. Kempa, “Plasmonic multiple exciton generation”, arXiv (2018), [arxiv.org/abs/1806.10259](https://arxiv.org/abs/1806.10259).

## 6.5 Fabrications

Finally, we comment on possible methods of developing arrays of NPs inside and active absorber volume. Wet chemistry-processed semiconductors are perhaps the easiest, as embedding can be achieved by simply mixing the NPs with the semiconductor. Embedding NPs into amorphous semiconductors processed by PECVD (a-Si and a-Ge) can be also obtained relatively easy by the layer-by-layer processing [31], or co-sputtering of a metal and semiconductor, followed by thermal processing [32]. Embedding plasmonic NPs into crystalline semiconductors is much more challenging. Most promising are crystalline NPs of silicides, which are plasmonic (metallic) with plasma energies in the 3 eV range [33], and so similar to Ag or Au. Most importantly, silicides are nearly lattice matched to Si, so they can be epitaxially grown on Si [34], and vice versa [35]. Many of the silicide NPs are also compatible with Ge, opening an avenue to PMEG solar cells. Another emerging technology is NP implantation, which allows deposition of NP growth seeds into semiconductors by ion implantation, and subsequent NP growth from those seeds during annealing, which restores crystalline structure [36].

## 6.6 Conclusion

In conclusion, it is shown that photoexcitation of two-pair bi-excitons can be protected against phonon emission, and therefore be a statistically likely event, if the semiconductor is filled with metallic NPs having plasmonic resonance tuned to the semiconductor gap energy. The bi-exciton formation process then results from a rapid sequence of two events: (i) initial exciton generation by the incoming photon, and (ii) the second exciton generation by the plasmon-stimulated hot electron's decay. This process can be viewed as plasmon-enhanced multiple exciton generation, PMEG. The universality of this effect provides a new paradigm in the development of ultrahigh efficiency solar cells, beyond the Shockley–Queisser limit. It is also demonstrated that PMEG solar cells benefit from smaller gap semiconductors, by consideration in detail three systems: large gap GaAs, intermediate gap *c*-Si and low gap Ge. While the first can be used only to demonstrate the PMEG process, the latter two could provide a possible platform for PMEG solar cells.

## References:

- [1] J. B. Gunn, "Microwave oscillation of current in III-V semiconductors", *Solid State Commun.* 1, 88-91 (1963).
- [2] H. Kroemer, "Theory of the Gunn effect", *Proc. IEEE* 52, 1736 (1964).
- [3] D. Frohman-Bentchkowsky, "Memory behavior in a floating-gate avalanche-injection MOS (FAMOS) structure", *Appl. Phys. Lett.* 18, 332-334 (1971).
- [4] M. Heiblum, M. I. Nathan, D. C. Thomas, and C. M. Knoedler, "Direct observation of ballistic transport in GaAs", *Phys. Rev. Lett.* 55, 2200-2203 (1985).
- [5] C. Rauch, G. Strasser, K. Unterrainer, W. Boxleitner, K. Kempa, and E. Gornik, "Ballistic electron spectroscopy of vertical biased superlattices", *Physica E* 2, 282-286 (1998).
- [6] A. Othonos, "Probing ultrafast carrier and phonon dynamics in semiconductors", *J. Appl. Phys.* 83, 1789-1830 (1998).
- [7] J. R. Goldman and J. A. Prybyla, "Ultrafast dynamics of laser-excited electron distributions in silicon", *Phys. Rev. Lett.* 72, 1364-1367 (1994).
- [8] S. M. Sze, "Microwave diodes", Chap. 9 in *High-Speed Semiconductor Devices*, Ed. by S. M. Sze (Wiley, New York 1990).
- [9] K. M. Kramer and W. N. G. Hitchon, *Semiconductor Devices: A Simulation Approach* (Prentice Hall PTR, Upper Saddle River, New Jersey, 1997).
- [10] J. A. Kash and J. C. Tsang, "Watching chips work: picosecond hot electron light emission from integrated circuits", *J. Cryst. Growth* 210, 318-322 (2000).
- [11] M. A. Green, *Third Generation Photovoltaics: Advanced Solar Energy Conversion* (Springer-Verlag, Berlin, 2006).
- [12] R. T. Ross and A. J. Nozik, "Efficiency of hot-carrier solar energy converters", *J. Appl. Phys.* 53, 3813-3818 (1982).
- [13] J. A. Turner, A.J. Nozik, "Evidence for hot-electron injection across p-GaP/electrolyte junctions", *Appl. Phys. Lett.* 41, 101-103 (1982).
- [14] K. Kempa, M. J. Naughton, Z. F. Ren, A. Herczynski, T. Kirkpatrick, J. Rybczynski, and Y. Gao, "Hot electron effect in nanoscopically thin photovoltaic junctions", *Appl. Phys. Lett.* 95, 233121 (2009).
- [15] M. C. Beard, J. M. Luther, O. E. Semonin, A. J. Nozik, "Third generation photovoltaics based on multiple exciton generation in quantum confined semiconductors", *Acc. Chem. Res.* 46, 1252-1260 (2013).

- [16] R. Schaller, V. Klimov, "High efficiency carrier multiplication in PbSe nanocrystals: Implications for solar energy conversion", *Phys. Rev. Lett.* 92, 186601 (2004).
- [17] M. C. Beard, J. M. Luther, A. J. Nozik, "The promise and challenge of nanostructured solar cells", *Nature Nanotechnol.* 9, 951-954 (2014).
- [18] O. E. Semonin, J. M. Luther, S. Choi, H.-Y. Chen, J. Gao, A. J. Nozik, and M. C. Beard, "Peak external photocurrent quantum efficiency exceeding 100% via MEG in a quantum dot solar cell", *Science* 334, 1530-1533 (2011).
- [19] J. Kong, A. H. Rose, C. Yang, J. M. Merlo, M. J. Burns, M. J. Naughton, K. Kempa, "A hot electron plasmon-protected solar cell", *Opt. Express* 23, A1087-A1095 (2015).
- [20] M. Voros, G. Galli and G. T. Zimanyi, "Colloidal nanoparticles for intermediate band solar cells", *ACS Nano* 9, 6882-6890 (2015).
- [21] G. D. Mahan, *Many-Particle Physics* (Plenum Press, New York 1981).
- [22] Y. Wang, E. Plummer and K. Kempa, "Foundations of plasmonics", *Adv. Phys.* 60, 799-898 (2011).
- [23] K. Kempa, "Plasmonic protection of the hot-electron energy", *Phys. Stat. Solidi RRL* 7, 465-468 (2013).
- [24] A. Taflove and S. C. Hagness, *Computational electrodynamics: The finite-difference time-domain method* (Artech, Norwood, MA 1995).
- [25] CST Microwave Studio, Computer Simulation Technology AG, <https://www.cst.com/products/cstmws>
- [26] D. R. Smith, S. Schultz, P. Markoš, and C. M. Soukoulis, "Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients", *Phys. Rev. B* 65, 195104 (2002).
- [27] M. Bernardi, D. Vigil-Fowler, C. S. Ong, J. B. Neaton, and S. G. Louie, "Ab initio study of hot electrons in GaAs", *PNAS* 112, 5291-5296 (2015).
- [28] M. Bernardi, D. Vigil-Fowler, J. Lischner, J. B. Neaton, and S. G. Louie, "Ab initio study of hot carriers in the first picosecond after sunlight absorption in silicon", *Phys. Rev. Lett.* 112, 257402 (2014).
- [29] N. Tandon, J. D. Albrecht and L. R. Ram-Mohan, "Electron-phonon interaction and scattering in Si and Ge: Implications for phonon engineering", *J. Appl. Phys.* 118, 045713 (2015).
- [30] R. T. Payne, "Phonon energies in germanium from phonon-assisted tunneling", *Phys. Rev. B* 139, A570-A582 (1965).

- [31] Y. Zhang, B. Cai and B. Jia, "Ultraviolet plasmonic aluminium nanoparticles for highly efficient light incoupling on silicon solar cells", *Nanomaterials* 6, 95 (2016).
- [32] R. Sachan, C. Gonzalez, O. Dyck, Y. Wu, H. Garcia, S. J. Pennycook, P. D. Rack, G. Duscher, and R. Kalyanaraman, "Enhanced absorption in ultrathin Si by NiSi<sub>2</sub> nanoparticles", *NME1* 2, 11-19 (2012).
- [33] R. Soref, R. E. Peale, and W. Buchwald, "Longwave plasmonics on doped silicon and silicides", *Opt. Express* 16, 6507-6514 (2008).
- [34] J. J. Chen, "Metal silicides: An integral part of microelectronics", *JOM* 57, 24-31 (2005).
- [35] J.-Y. Veuillen, "Growth of silicon thin films on erbium silicide by solid phase epitaxy", *J. Appl. Phys.* 75, 223-226 (1994).
- [36] H. W. Seo, Q.Y. Chen, I.A. Rusakova, Z.H. Zhang, D. Wijesundera, S.W. Yeh, X. M. Wang, L.W. Tu, N. J. Ho, Y. G. Wu, H. X. Zhang, W. K. Chu, "Formation of silver nanoparticles in silicon by metal vapor vacuum arc ion implantation", *Nucl. Instrum. Methods Phys. Res. B* 292, 50-54 (2012).
-