



From Classical to Quantum Plasmonics in Three and Two Dimensions

Christensen, Thomas; Mortensen, N. Asger; Wubs, Martijn; Jauho, Antti-Pekka

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Christensen, T., Mortensen, N. A., Wubs, M., & Jauho, A-P. (2015). From Classical to Quantum Plasmonics in Three and Two Dimensions. Technical University of Denmark (DTU).

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

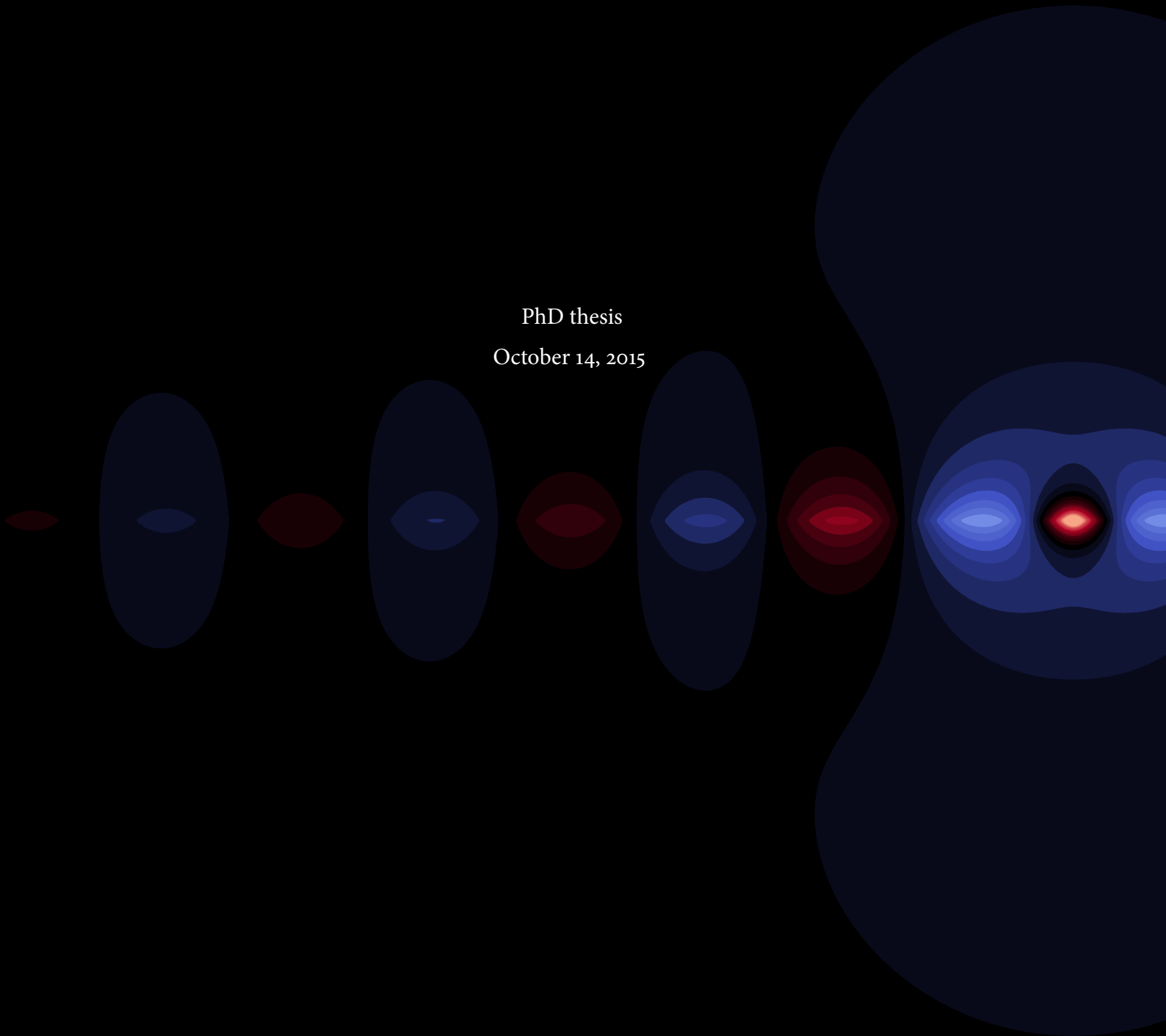
- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FROM CLASSICAL TO QUANTUM PLASMONICS
IN THREE AND TWO DIMENSIONS

THOMAS CHRISTENSEN

PhD thesis
October 14, 2015



COVER IMAGE

Cross-sectional view (xz -plane) of the induced potential across a graphene sheet (lying in the xy -plane) due to a traversing electron of velocity $v\hat{z}$. A plasmon of frequency ω and characteristic momentum ω/v is excited, propagating and decaying away from the impact-point of the electron (right-hand side).

Color-scale ranges from negative (blue) through zero (black) to positive (red). The technical details of the calculation are detailed in [Appendix A](#).

TYPOGRAPHY, FIGURES, AND CALCULATIONS

This thesis was typeset in \LaTeX using a modified version of the ArsClassica style with typefaces MinionPro, Iwona, and MyriadPro. All calculations were implemented and performed in Matlab. Figures were produced in Matlab (data representation), PovRay (ray-tracing), and TikZ (sketches and illustrations).

Thomas Christensen: *From classical to quantum plasmonics in three and two dimensions*, PhD thesis.

© October 14, 2015– version 1.0

SUPERVISORS

N. Asger Mortensen,
Martijn Wubs,
Antti-Pekka Jauho.

UNIVERSITY

Technical University of Denmark.
Department of Photonics, Structured Electromagnetic Materials, and Center for Nanostructured Graphene.

God made the bulk, but the devil invented the surfaces

— Wolfgang Pauli (1900 – 1958)

ABSTRACT

This thesis presents theoretical results for the description and understanding of plasmons in three- and two-dimensional platforms, with a special emphasis on the evolution from classical to nonclassical behavior as the optical and structural length scales are reduced towards the intrinsic scales of the electronic plasma. The content is divided into two parts and split by dimensionality.

First, following a general introduction to the topic of plasmonics in three dimensions, we review the fundamental shortcomings of the conventional classical approach, finding its limitations to fall in four categories. We center our efforts on the deficiencies originating in the disregard of nonlocality, and explore its inclusion by means of a hydrodynamic model, which accounts to lowest order for the momentum-dispersion of the dielectric response. Concretely, we apply the hydrodynamic framework to the half-space, thin film, and spherical geometries. In the latter case, we extend the understanding of hydrodynamics beyond the dipolar regime, by contrasting optical probes of far- and near-field character. For short probe-to-surface separations, we establish that near- and far-field measurements provide significantly dissimilar weighting of spectral features, with particular importance to the excitation of multipole plasmons. Moreover, for these multipole plasmons, we find a hydrodynamic shift which increases with multipole order. This shift removes the singular classical pile-up of multipoles near the planar surface plasmon frequency. Complementing these considerations, we present results arising from an experimental collaboration, in which, using electron-energy loss spectroscopy, the impact of higher-order multipoles are identified in embedded silver nanoparticles with radii down to 4 nm. Finally, in recognizing the limitations of the hydrodynamic model, we propose the outline of an extension of the Feibelman d -parameter approach to arbitrary geometries. Formally, this extension achieves a simultaneous first-order account of spill-out, nonlocality, and Landau damping, by instating a natural division between electronic and optical aspects.

Our treatment of two-dimensional plasmonics centers on the platform of graphene. After a short review of graphene's intrinsic electronic and optical properties, we introduce and explicate the main characteristic features of graphene plasmonics. We tabulate and discuss the resonance conditions and properties for the extended sheet, half-sheet, ribbon, disk, and regular polygons. In addition, we consider the existence of graphene plasmons in the non-planar geometry of a coated nanosphere. Proceeding to a consideration of effects beyond the conventional approach, we present first an adaptation of the hydrodynamic model to graphene. Next, we review the recently introduced tight-binding approach to the quantum plasmonic response of graphene nanostructures. We demonstrate how this frequency-domain method finds an equivalent implementation in the time-domain. Using this method and a related Dirac equation approach, we investigate the role of edge states in graphene, finding significant differences between the plasmonic properties of armchair- and non-armchair-terminated nanostructures. Lastly, we discuss the influence of a nonlinear Kerr interaction on the plasmons supported by a graphene nanoribbon, finding redshifting behavior relative to the linear case. We offer a straightforward and general perturbative understanding of this prediction, which depends only on the plasmon's spatial inhomogeneity and geometry-averaged field intensity.

RESUMÉ

Denne afhandling præsenterer teoretiske resultater for beskrivelsen og forståelsen af plasmoner i tre og todimensionelle platforme, med en særlig vægt på overgangen fra klassisk til ikke-klassisk opførsel ved reduktionen af optiske og strukturelle længdeskalaer mod de intrinsiske skalaer i det elektroniske plasma. Indholdet er opdelt efter dimensionalitet i to bestanddele.

Efter en generel introduktion til emnet plasmonik i tre dimensioner, redegør vi for de fundamentelle begrænsninger af den konventionelle klassiske tilgang, og placerer disse mangler i fire kategorier. Vi centrerer vores indsats mod afhjælpningen af de begrænsninger, der skyldes tilsidesættelsen af ikke-lokalitet, og undersøger dens inklusion ved en hydrodynamisk model, som til laveste orden redegør for impulsafhængigheden af det dielektriske respons. Konkret anvender vi det hydrodynamiske rammeværk i et halv-rum, en tynd-film og en sfære. I sidstnævnte tilfælde udvider vi forståelsen af hydrodynamikkens indflydelse forbi det dipolare regime, ved at sammenligne det optiske respons i fjern- og nærfelt. For korte sonde-til-overflade afstande etablerer vi, at fjern- og nærfelts målinger udviser en væsentligt forskellige vægtning af de spektrale egenskaber, med særlig indflydelse på excitationen af multipol plasmoner. For disse multipol plasmoner finder vi ydermere et hydrodynamisk skift, der stiger med multipol-orden. Dette skift fjerner den singulære sammensmeltning af multipoler i den klassiske beskrivelse ved den plane overflade plasmons egenfrekvens. Disse overvejelser komplementeres af eksperimentelle resultater fra et elektron energitabs spektroskopi studium, hvori vi påviser indflydelsen af højere-ordens multipoler i indlejrede sølvpartikler af radier ned til 4 nm. Endeligt, med begrænsningerne af den hydrodynamiske model i mente, foreslår vi en udvidelse af Feibelman d -parameter tilgangen til generelle geometrier. Denne udvidelse opnår formelt en simultan førsteordens beskrivelse af udspildning, ikke-lokalitet og Landau dæmpning. Dette opnås i praksis ved indførslen af en naturlig opdeling af elektroniske og optiske aspekter.

Vores behandling af todimensionelle plasmoner fokuserer på platformen udgjort af grafen. Efter en kort gennemgang af grafens intrinsiske elektroniske og optiske egenskaber, introducerer og udlægger vi de primære karakteristiske særpræg i grafen plasmonik. Vi tabulerer og diskuterer resonansbetingelserne samt egenskaberne af det uendelige ark, det halve ark, båndet, skiven og de regulære polygoner. Derudover undersøger vi eksistensen af grafen plasmoner i den ikke-planare konfiguration udgjort af en belagt sfære. Derefter forsætter vi overvejelsen af effekter hinsides den konventionelle tilgang, og præsenterer i denne relation først en tilpasning af den hydrodynamiske model til grafen. Efterfølgende redegør vi for den nyligt indførte tætte-binding tilgang til det kvantemekaniske plasmoniske respons af grafen nanostrukturer. Vi demonstrerer, hvordan denne frekvens-domæne beskrivelse kan implementeres på ækvivalent vis i tids-domænet. Ved anvendelse af denne metode og en relateret Dirac-lignings tilgang, udforsker vi rollen af kanttilstande i grafen og finder herved væsentlige forskelle mellem de plasmoniske egenskaber i nanostrukturer termineret med armstol og ikke-armstol kanter. Slutteligt diskuterer vi indflydelsen af en ikke-lineær Kerr interaktion for plasmoner understøttet af et nanobånd, og forudsiger røds-kiftende opførsel relativt til den lineære beskrivelse. Vi tilbyder en ligefrem og generel perturbativ forståelse af denne forudsigelse, som kun afhænger af plasmonens rumlige inhomogenitet samt domæne-integrerede felt intensitet.

PREFACE

This thesis is submitted in partial fulfillment of the requirements to receive a Doctorate of Philosophy (PhD) at the Technical University of Denmark. The work detailed herein was carried out over the span of three years (October 1, 2012 – October 14, 2015) at the Department of Photonics Engineering and at the Center for Nanostructured Graphene in the group of Structured Electromagnetic Materials. The project was supervised by Prof. N. Asger Mortensen, Assoc. Prof. Martijn Wubs, and Prof. Antti-Pekka Jauho.

The present text embodies my best attempt at the presentation of a selection of topics which have occupied me in the last three years. I have tried hard to provide a clear and interesting account of the subject matter, though not a complete one. Whether the presentation as a whole is ultimately satisfactory I leave entirely to reader.



Thomas Christensen

Technical University of Denmark
Kgs. Lyngby
October 14, 2015

ACKNOWLEDGEMENTS

Over the course of this PhD study I have had the good fortune to amass a debt of gratitude. First and foremost I am grateful to my supervisors – in particular, I am grateful to Asger for providing me with this opportunity and for trusting me with free reigns; to Martijn for his always careful consideration and open mind; and to Antti for his valiant efforts towards maturing my scientific taste and prose – and to them all for their continued encouragement, positive outlook, and sterling scientific input.

More broadly, I would like to thank the members, current and past, of the Structured Electromagnetic Materials group from whom I have learned much. I single out particularly my early office-mates Søren Raza and Jeppe S. Clausen, and thank them both for guiding me into the PhD-arena and for many good discussions – Søren I additionally thank for several fruitful collaborations. I am also thankful to Weihua Wang for his willingness to include me in his work and ideas, which helped me gain traction at the beginning of my studies. To Nicolas Stenger I express my appreciation especially for an experimental perspective. As a last SEM-entrant, I would express my sincere gratitude to Wei Yan for our daily interactions, countless insights, infectious drive, and more generally for an explicit and substantial partaking in my education.

For visits abroad I wish to thank Meir Orenstein and Ofir Sorias for hosting me at Technion for two months; especially Ofir for including me in everything outside of Technion as well. Similarly, I thank Javier Garcia de Abajo and Joel Cox for a one-week stay at ICFO, and for allowing me a unique opportunity to benchmark and verify numerical calculations. That I did not manage to materialize publishable content from these excellent interactions constitutes one of my few real regrets in this project.

To my thesis committee, specifically to Prof. Nader Engheta, Prof. Frank H.L. Koppen, and Prof. Ulrik L. Andersen, I wish to express my thanks for taking time out of their busy schedule to consider this thesis, for reading it, and for attending my defense.

With my academic debts of gratitude thus reviewed in shorter terms than it deserves, I move finally to my everyday, personal debt. First, I extend my thanks to family and friends for their interest in my life and well-being and for distractions from physics. Penultimately, I express my profound thankfulness to my parents, for their steadfast support, hands-on council, and for teaching me that the worthwhile paths in life seldom coincide with the easy ones. Specific thanks are also due to my father for proofreading this thesis; naturally, I stand responsible for the remaining errors. Lastly, my immeasurable gratitude and love to Sofie, for taking care of quite literally *everything* outside of this PhD and above all for her love and affection, without which I would be much less.

LIST OF PUBLICATIONS

With the exception of [Publication H](#) all papers have been published during the author's PhD studies. [Publications A–G](#) are discussed in this thesis while [Publications A, D, E](#), and [F](#) constitute efforts primarily spearheaded by the author (and are appended to this thesis in [Part IV](#)).

- A. [T. Christensen](#), W. Yan, A.-P. Jauho, M. Wubs, and N.A. Mortensen, *Kerr nonlinearity and plasmonic bistability in graphene nanoribbons*, [Physical Review B](#) **92**, 121407(R) (2015)
- B. S. Raza, S. Kадkhodazadeh, [T. Christensen](#), M. Di Vece, M. Wubs, N.A. Mortensen, and N. Stenger, *Multipole plasmons and their disappearance in few-nanometre silver nanoparticles*, [Nature Communications](#) **6**, 8788 (2015, in print). [[arXiv:1505.00594](#)]
- C. W. Wang, [T. Christensen](#), A.-P. Jauho, K.S. Thygesen, M. Wubs, and N.A. Mortensen, *Plasmonic eigenmodes in individual and bow-tie graphene nanotriangles*, [Scientific Reports](#) **5**, 9535 (2015).
- D. [T. Christensen](#), A.-P. Jauho, M. Wubs, and N.A. Mortensen, *Localized plasmons in graphene-coated nanospheres*, [Physical Review B](#) **91**, 125414 (2015).
- E. [T. Christensen](#), W. Wang, A.-P. Jauho, M. Wubs, and N.A. Mortensen, *Classical and quantum plasmonics in graphene nanodisks: Role of edge states*, [Physical Review B](#) **90**, 241414(R) (2014).
- F. [T. Christensen](#), W. Yan, S. Raza, A.-P. Jauho, N.A. Mortensen, and M. Wubs, *Nonlocal response of metallic nanospheres probed by light, electrons, and atoms*, [ACS Nano](#) **8**, 1745 (2014).
- G. S. Raza, [T. Christensen](#), M. Wubs, S.I. Bozhevolnyi, and N.A. Mortensen, *Nonlocal response in thin-film waveguides: Loss versus nonlocality and breaking of complementarity*, [Physical Review B](#) **88**, 115401 (2013).
- H. J.R. de Lasson, [T. Christensen](#), J. Mørk, and N. Gregersen, *Modeling of cavities using the analytic modal method and an open geometry formalism*, [JOSA A](#) **29**, 1237 (2012).

CONTENTS

1	INTRODUCTION	1
1.1	Outline	2
I	PLASMONICS IN THREE DIMENSIONS	5
2	FUNDAMENTALS OF PLASMONICS	7
2.1	Optical response	7
2.1.1	Foundations of macroscopic electrodynamics	7
2.1.2	Optical excitations	9
2.2	Plasmonics	9
2.2.1	The local response approximation	9
2.2.2	The Drude response model	10
2.2.3	Classical plasmonics and the role of geometry	12
2.3	Probing plasmonic effects	16
2.3.1	Far-field measurements	16
2.3.2	Near-field measurements	17
2.4	Summary	19
3	NONCLASSICAL PLASMONICS	21
3.1	Shortcomings of the classical theory	21
3.1.1	The Feibelman d -parameter	24
3.2	Beyond the Drude model	25
3.2.1	The hydrodynamic model	25
3.2.2	The Lindhard model	29
3.2.3	GNOR and Kreibig broadening	29
3.2.4	Applicability of the hydrodynamic model	30
3.3	Nonlocality at planar interfaces	31
3.3.1	SPP dispersion and the single interface	31
3.3.2	Thin-film waveguides	32
3.4	Nonlocality in nanospheres	34
3.4.1	Mie–Lorenz theory	35
3.4.2	Multipole plasmons	36
3.4.3	Role of the measurement probe	39
3.5	Multipole plasmons in embedded silver nanoparticles	42
3.6	Outlook	46
3.7	Summary	48
II	PLASMONICS IN TWO DIMENSIONS	49
4	ELECTRONIC PROPERTIES OF GRAPHENE	51
4.1	Electronic properties of graphene	51
4.1.1	Structural composition	51
4.1.2	Tight-binding approach	52
4.1.3	The Dirac limit	54
4.1.4	Miscellaneous auxiliary properties	54
4.2	Optical response of graphene	55
4.2.1	Nonlocal density-density response	55
4.2.2	Local-response conductivity	57
4.3	Summary	59
5	CLASSICAL GRAPHENE PLASMONICS	61
5.1	Influence of dimensionality	61
5.2	Surface plasmons in extended graphene	63

5.2.1	Reflection coefficients and surface currents	63
5.2.2	TM graphene plasmons	64
5.2.3	TE graphene plasmon polaritons	68
5.3	Plasmons in nanostructured graphene	69
5.3.1	Nonretarded formulation	69
5.3.2	Role of geometry	71
5.4	Nonplanar geometries: graphene-coated sphere	78
5.5	Summary	82
6	NONCLASSICAL GRAPHENE PLASMONICS	83
6.1	Hydrodynamic description	83
6.2	Quantum graphene plasmonics	86
6.2.1	Edge states in graphene nanostructures	86
6.2.2	Tight-binding and the RPA	89
6.2.3	Continuum Dirac approach	94
6.3	Nonlinear response	97
6.4	Summary	99
7	OUTLOOK AND CONCLUSIONS	101
III APPENDICES 105		
A	INTERACTION OF FAST ELECTRONS WITH GRAPHENE	107
A.1	Potential and screened interaction	107
A.2	Nonretarded electron energy loss	108
B	EXISTENCE CRITERIA: TE PLASMON POLARITON	111
C	EIGENDECOMPOSITION IN 2D NANOSTRUCTURES	113
D	NONINTERACTING DENSITY-DENSITY RESPONSE	115
E	DISCRETIZED UNITARY TIME-OPERATOR	117
E.1	Bare evolution thresholding	118
IV SELECTED PUBLICATIONS 119		
	PHYS. REV. B 92 , 121407(R) (2015)	121
	PHYS. REV. B 91 , 125414 (2015)	135
	PHYS. REV. B 90 , 241414(R) (2014)	145
	ACS NANO 8 , 1745 (2014)	167
BIBLIOGRAPHY 197		

ACRONYMS

Acronym	Expanded form
ID	one-dimensional
2D	two-dimensional
3D	three-dimensional
ABC	additional boundary condition
AC	armchair
BC	boundary condition
BEM	boundary element method
CBP	confined bulk plasmon
DFT	density functional theory
DOS	density of states
EEL	electron energy loss
EELS	electron energy loss spectroscopy
EHP	electron-hole pair
FBZ	first Brillouin zone
FDTD	finite-difference time domain
FEM	finite-element method
FWHM	full width at half maximum
GIEM	Green function integral equation method
GNOR	generalized nonlocal optical response
GP	graphene plasmon
GPP	graphene plasmon polariton
HDM	hydrodynamic model
HO	higher-order
LDOS	local density of states
LRA	local response approximation
LSP	localized surface plasmon
RPA	random phase approximation
SERS	surface-enhanced Raman spectroscopy
SNOM	scanning near-field optical microscopy
SP	surface plasmon
SPP	surface plasmon polariton
STEM	scanning transmission electron microscopy
TB	tight-binding
TDDFT	time-dependent density functional theory
TE	transverse electric

... continued on the following page

...continued from the previous page

Acronym	Expanded form
TEM	transmission electron microscopy
TM	transverse magnetic
xc	exchange-correlation
ZZ	zigzag

1

INTRODUCTION

Plasmons are collective excitations in the free-electron plasma of conductive materials, such as e.g. noble metals. Classically, they are characterized by an oscillation of the free-electron plasma, sustained by inter-electronic interaction in a positive ionic background. The interaction is mediated in the nonretarded case by the Coulomb force, and in a full electrodynamic view, by the Maxwell equations. It is a testament to their remarkable potential, that their unique attributes have been exploited in numerous cases long before their theoretical description. For instance, cranberry glass (occasionally known as gold ruby glass), which have existed for centuries, even millennia, owes its bright red color to the presence of nanoscopic gold colloids that support so-called localized surface plasmons [1]. The same mechanism is responsible for the vivid colors of certain stained glass mosaics and ancient glazed lustre ceramics [2].

Despite their historical applications, the understanding of plasmons – and of the above-noted examples – is more recent, unfolding primarily from the turn of the 20th century. Notable early contributions include Mie's 1908 account of scattering in colloidal metal solutions [3] and Zsigmondy's 1902 identification and characterization of individual gold colloids [4, 5]. Their theoretical description was given a full microscopic account in 1952 by Pines and Bohm [7], in their introduction of the so-called random-phase approximation. Subsequent theoretical consideration of plasmons in thin films by Ritchie [8] paved the way for the concept of surface plasmons.

Today, the field of *plasmonics* is, as the name suggests, concerned with the study of plasmons for technological and fundamental purposes [9–11]. The study has unveiled a range of truly extraordinary features, particularly in the subbranch of nanoplasmonics, which centers on the properties of plasmons in nanostructures [12, 13]. Among these feats are extreme light enhancements [14], subwavelength light-concentration far beyond the diffraction limit [15], and exceptional absorbing and scattering properties [16]. These features have allowed plasmonics to extend into and influence several related fields, e.g. in sensing [17], metamaterials [18], and antenna-design [19].

The long-standing interest in plasmons has accelerated in the last two decades, partly due to advances in nanotechnology, especially in fabrication and characterization aspects, and partly due to a number of galvanizing developments. In more recent years, we highlight two developments – which, as it happens, constitute the focus of this thesis – namely, the emerging interest in nonclassical, or quantum, corrections to the classical theory of plasmons, and the study of plasmons in the two-dimensional material graphene. In the following, we discuss briefly the motivation underlying each of these outgrowths.

The first-noted development, nonclassical plasmonics, has been spurred on, essentially, by ever increasing capabilities in nanofabrication [21] and optical characterization [22]. For instance, it is today possible to fabricate structures with characteristic features in the ~ 1 nm range [23, 24], and to investigate their plasmonic response with simultaneous spatial and spectral distinction [25, 26]. The conventional theoretical approach to plasmonics relies on classical macroscopic electrodynamics combined with local spectrally dispersive material response functions. Eventually, as the relevant optical and structural length scales approach the intrinsic quantum mechanical scales of the electronic gas itself, this classical account deteriorates. Efforts to extend the classical framework, i.e. include quantum effects, come in essentially two flavors, distinguished by their quantization of either (a) the electromagnetic field [29, 30],

For a while, following the 1929 classical description by Tonks and Langmuir [6], plasmons were known as Langmuir waves.

The 1998 observations of enhanced transmission through subwavelength hole-arrays [20], constitutes an oft-noted example of such a galvanizing and stimulating event.

The intersection of these extensions, i.e. simultaneous accounting of optical and electronic quantization, poses a formidable challenge, which, however, has seen progress recently [27, 28].

constructing a bridge between plasmonics and quantum optics [31] or (b) the states of the free-electron plasma, connecting then with the language of condensed matter physics [32]. The nonclassical extensions relevant to the small-scale limit are primarily of the second type, and are the kind discussed in this thesis. More specifically, we will focus predominantly on the exploration of certain semi-classical corrections, which, although invariably less complete than full quantum accounts, have the crucial benefits of interpretational clarity and computational scalability.

The second-noted development, plasmonics in graphene, has revitalized the interest of the 1960's, 70's and 80's in two-dimensional plasmons [33]. The catalyst, of course, has been the synthesis of graphene [34–36], the two-dimensional allotrope of carbon which make up the planar building blocks of graphite. At the outset of this project, graphene plasmonics represented a comparatively young, largely virginal field [37–39]; here, at the end of the project, the field has matured tremendously and expanded in several multifaceted directions [40–46]. Our considerations of graphene plasmonics in this thesis will pertain predominantly to their intrinsic properties, i.e. not to any particular application. It is the author's hope that the treatment here might prove useful e.g. to newcomers to the field, or, perhaps, provide a complementary perspective to those already in it. In connecting our efforts on nonclassical and two-dimensional plasmonics, we will, as the last component of the thesis, also find opportunity to examine examples of nonclassical effects which are unique to graphene plasmonics.

In accordance with the prevailing preferences of the plasmonic community we employ the SI unit-system throughout this thesis, except for energies which we state in units of eV.

1.1 OUTLINE

The thesis is divided into two parts, [Parts I](#) and [II](#), which center, respectively, on plasmonics in three and two dimensions. In either case, we introduce first the classical understanding, and proceed subsequently to the exploration of nonclassical corrections. For concreteness, we provide a short overview of each chapter below:

CHAPTER 2 | FUNDAMENTALS OF PLASMONICS Establishes the core elements of conventional three-dimensional plasmonics, introduces the characteristic attributes and classifications of plasmons, and finally indicates a selection of approaches to their practical investigation and detection.

CHAPTER 3 | NONCLASSICAL PLASMONICS Discusses the shortcomings of the conventional, classical approach to plasmonics in the small-scale and large-momentum limits; develops and applies the hydrodynamic model, a low-momentum description of nonlocality; investigates theoretically and experimentally the role of multipole plasmons in near-field measurements; and closes with an outlook to a broader and more general perspective of nonclassical effects.

CHAPTER 4 | ELECTRONIC PROPERTIES OF GRAPHENE Introduces graphene as a material platform, presenting first its fundamental electronic properties, and next its low-energy optical response functions.

CHAPTER 5 | CLASSICAL GRAPHENE PLASMONICS Describes the general role of dimensionality in plasmonics, introduces the general features and characteristics of classical graphene plasmonics in nanostructured and extended graphene, and closes with a consideration of graphene plasmonics in a non-planar configuration.

CHAPTER 6 | NONCLASSICAL GRAPHENE PLASMONICS Presents an adaptation of the hydrodynamic model to graphene; discusses a quantum description of graphene plasmonics, based on a tight-binding random-phase approximation; explores the role of edge states and develops a semi-classical account for their role; and lastly examines the influence of a nonlinear Kerr effect in nanoribbons.

CHAPTER 7 | OUTLOOK AND CONCLUSIONS Reexamines the overall content of the thesis, and suggests a number of perspectives on the results obtained and their relations to the efforts of the nanoplasmonic community as a whole.

Part I

PLASMONICS IN THREE DIMENSIONS

2

FUNDAMENTALS OF PLASMONICS

One particularly attractive aspect of plasmonics is that its characteristic features can be deduced from just classical electrodynamics and a simple picture of the electron gas, known as the Drude model. This fact is only made all the more pleasing by the breadth of possible viewpoints and interpretations, ranging from the very classical to the very quantum. For instance, a classical macroscopic view of a plasmon depicts a spatially oscillating charge density, self-sustained by its coupling with the electromagnetic field. Conversely, the quantum state of a plasmon portray a collective superposition of electron-hole pairs oscillating about the Fermi surface. To appreciate the quantum picture fully, however, it is pertinent to acquaint ourselves first with the classical picture, which, moreover has the great advantage of practical applicability. This is purpose of the present chapter – aspects of the quantum picture are explored in [Chapters 3 and 6](#).

We first review the necessary components of electrodynamics and establish a practical notion of optical excitations. Next, we specialize to plasmonic excitations, and discuss their classical features. Finally, we cover theoretical aspects of techniques that probe the properties of plasmons in practice.

2.1 OPTICAL RESPONSE

A classical description of plasmons require two components: geometry and material. The former, geometry, involves shape and scale and is included principally by an electrodynamic consideration, whereas the latter, material, accounts for the response properties of the constituent electronic ingredients, constituting a branch of condensed matter theory. In this section, we review the aspects pertaining to geometry, namely electrostatics. Later, in [Section 2.2.2](#) we return to the material aspects.

2.1.1 Foundations of macroscopic electrodynamics

The macroscopic Maxwell equations in space-time coordinates (\mathbf{r}, t) read as [\[47\]](#)

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\partial_t \mathbf{B}(\mathbf{r}, t), \quad (1a)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \mathbf{J}_f(\mathbf{r}, t) + \partial_t \mathbf{D}(\mathbf{r}, t), \quad (1b)$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \rho_f(\mathbf{r}, t), \quad (1c)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0, \quad (1d)$$

for the electric \mathbf{E} -, magnetic \mathbf{B} -, displacement \mathbf{D} -, and magnetizing \mathbf{H} -fields, driven by the *free* charge densities ρ_f and currents \mathbf{J}_f , which is decomposed into external and induced constituents according to $\rho_f \equiv \rho_{\text{ext}} + \rho_{\text{ind}}$ and $\mathbf{J}_f \equiv \mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{ind}}$. The densities and currents are linked by the continuity equation

$$\nabla \cdot \mathbf{J}_{[\dots]}(\mathbf{r}, t) + \partial_t \rho_{[\dots]}(\mathbf{r}, t) = 0, \quad (2)$$

applicable to each constituent $[\dots] = \{f, \text{ext}, \text{ind}\}$ individually. While external sources, ρ_{ext} and \mathbf{J}_{ext} , are assumed known *a priori*, the induced counterparts, ρ_{ind} and \mathbf{J}_{ind} , arise from the response of free charge-carriers, such as electrons in the conduction band of metals.

Eqs. (1) collect the historical Faraday law (a), the Maxwell–Ampère law (b), and the electric and magnetic Gauss laws (c) and (d).

The Maxwell equations can in fact be written in just three equations, namely Eqs. (2), (1a) and (1b) from which Eqs. (1c) and (1d) follow [48].

The electric and magnetic fields \mathbf{E} and \mathbf{B} constitute the fundamental quantities – though macroscopically averaged – through which the displacement and magnetizing field \mathbf{D} and \mathbf{H} , as well as \mathbf{J}_{ind} , are assumed to depend. In particular, \mathbf{D} accounts for the bound polarization \mathbf{P} via $\mathbf{D} \equiv \epsilon_0 \mathbf{E} + \mathbf{P}$, while \mathbf{H} accounts for magnetization \mathbf{M} through $\mathbf{H} \equiv \mu_0^{-1} \mathbf{B} - \mathbf{M}$. We will not consider materials with finite magnetization and hence take $\mathbf{H} = \mu_0^{-1} \mathbf{B}$ throughout. Consequently, what remains is only to establish a link between \mathbf{D} and \mathbf{J}_{ind} and the fundamental quantities. For linear media, this link is provided by the spatially and temporally dispersive constitutive relations

$$\mathbf{D}(\mathbf{r}, t) = \epsilon_0 \iint \epsilon_{\text{B}}(\mathbf{r}, \mathbf{r}'; t - t') \mathbf{E}(\mathbf{r}', t') \, d\mathbf{r}' dt', \quad (3a)$$

$$\mathbf{J}_{\text{ind}}(\mathbf{r}, t) = \iint \sigma(\mathbf{r}, \mathbf{r}'; t - t') \mathbf{E}(\mathbf{r}', t') \, d\mathbf{r}' dt', \quad (3b)$$

expressed in terms of the bound-response dielectric function ϵ_{B} and the free-carrier conductivity σ , and integrated over $\mathbf{r}' \in \mathbb{R}^3$ and $t' \in \mathbb{R}$ (assuming appropriate causal cutoffs at $t - t' < 0$). Evidently, the response depends on the electric field at all previous times and throughout space, reflecting aspects of memory and nonlocality, respectively. The former, i.e. temporal dispersion, assumes the form of a convolution in all time-independent systems where the response can only depend on the time-difference $t - t'$. Accordingly, the time-integral is a product in the frequency domain (in which we shall reside almost exclusively) cf. the convolution theorem

$$\mathbf{D}(\mathbf{r}, \omega) = \epsilon_0 \int \epsilon_{\text{B}}(\mathbf{r}, \mathbf{r}'; \omega) \mathbf{E}(\mathbf{r}', \omega) \, d\mathbf{r}', \quad (4a)$$

$$\mathbf{J}_{\text{ind}}(\mathbf{r}, \omega) = \int \sigma(\mathbf{r}, \mathbf{r}'; \omega) \mathbf{E}(\mathbf{r}', \omega) \, d\mathbf{r}'. \quad (4b)$$

Conversely, the spatial integration is not necessarily a convolution, and a momentum (\mathbf{k} -)representation is correspondingly not necessarily useful. In particular, the microscopic lattice itself breaks full translational symmetry, allowing only partial decomposition in Brillouin-restricted momenta \mathbf{k} and reciprocal lattice vectors \mathbf{G} and \mathbf{G}' . In addition, and perhaps of more practical significance, the symmetry is generally broken in finite, structured systems by the existence of material boundaries. Even so, owing to the quality of the jellium ansatz, a homogeneous assumption – paired with appropriate boundary conditions (BCs) at material boundaries – will generally suffice.

The partition into bound and free currents is in principle arbitrary, and is furthermore not a particularly useful distinction in plasmonics. Accordingly, it is desirable to combine *all* induced polarizations and currents in a *total* dielectric function $\epsilon(\mathbf{r}, \mathbf{r}'; \omega)$. To do so, we observe that combination of the frequency-domain form of Eq. (1b) and (4) yields (letting $\partial_t \rightarrow -i\omega$ under Fourier transformation)

$$\nabla \times \mathbf{H}(\mathbf{r}, \omega) = \mathbf{J}_{\text{ext}}(\mathbf{r}, \omega) - i\omega \tilde{\mathbf{D}}(\mathbf{r}, \omega), \quad (5)$$

with an amalgamate displacement field $\tilde{\mathbf{D}}(\mathbf{r}) \equiv \mathbf{D}(\mathbf{r}) + \frac{i}{\omega} \mathbf{J}_{\text{ind}}(\mathbf{r})$ which in turn defines the total dielectric function $\tilde{\mathbf{D}}(\mathbf{r}) \equiv \epsilon_0 \int \epsilon(\mathbf{r}, \mathbf{r}') \mathbf{E}(\mathbf{r}') \, d\mathbf{r}'$

$$\epsilon(\mathbf{r}, \mathbf{r}'; \omega) = \epsilon_{\text{B}}(\mathbf{r}, \mathbf{r}'; \omega) + \frac{i\sigma(\mathbf{r}, \mathbf{r}'; \omega)}{\epsilon_0 \omega}. \quad (6)$$

We note that the new displacement field $\tilde{\mathbf{D}}$ fulfills Maxwell's equations Eqs. (1) in the frequency domain under the substitutions $\mathbf{D} \rightarrow \tilde{\mathbf{D}}$, $\mathbf{J}_f \rightarrow \mathbf{J}_{\text{ext}}$, and $\rho_f \rightarrow \rho_{\text{ext}}$ as a direct consequence of Eq. (2). Accordingly, it is customary to redefine \mathbf{D} to correspond to $\tilde{\mathbf{D}}$; we will follow this convention henceforth. This choice is beneficial because all induced quantities are then contained in a single field, rather than two. Moreover, the total dielectric function still expresses the partitioning into bound polarization, ϵ_{B} , and free currents, σ , analogizing aspects of insulator-like and conductor-like response, respectively.

Anisotropy can be included in Eqs. (3) by considering tensorial forms of the response functions ϵ_{B} and σ .

We define Fourier transforms: $f(t) \equiv \int f(\omega) e^{-i\omega t} \, d\omega$ and $g(\mathbf{r}) \equiv \int g(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \, d\mathbf{k}$ with appropriate inverse relations.

Equivalently, in the time-domain: $\partial_t \tilde{\mathbf{D}} \equiv \partial_t \mathbf{D} + \mathbf{J}_{\text{ind}}$.

2.1.2 Optical excitations

To introduce the concept of an optical excitation, we first contemplate a general linear relation between a total or induced signal s and an external stimuli s_{ext} conveyed by a frequency-dependent response operator $R(\omega)$ (with explicit declaration of the frequency-dependence of s and s_{ext} omitted but implicitly assumed)

$$s = R(\omega)s_{\text{ext}}. \quad (7a)$$

We propose the following definition of an optical excitation, characterized by its frequency ω_n

$$\text{An optical excitation } \omega_n \text{ corresponds to poles of } R(\omega), \text{ or, more generally, to zeros of the (functional) determinant of } R^{-1}(\omega). \quad (7b)$$

Accordingly, associated optical excitation modes s_n correspond to the nullspace of $R^{-1}(\omega_n)$, or equivalently, to nontrivial solutions of $R^{-1}(\omega)s = 0$. The definition posed by Eq. (7b) is slightly unconventional and perhaps less familiar than the usual definition associated with the eigensolutions of the wave equation; nevertheless, it is equivalent. This is clear if the response operator $R(\omega)$ is identified as the dyadic Green function $\vec{\mathbf{G}}(\omega)$. To demonstrate this, we note that the E-field wave equation is obtained by combining Eqs. (1a), (1b), and (4)

$$[\nabla \times \nabla \times - k_0^2 \hat{\varepsilon}] \mathbf{E}(\mathbf{r}, \omega) = i\omega\mu_0 \mathbf{J}_{\text{ext}}(\mathbf{r}, \omega), \quad (8)$$

with free-space wavenumber $k_0 \equiv \omega/c$ expressed in terms of the speed of light c , and with the short-hand notation $\hat{\varepsilon} \mathbf{E}(\mathbf{r}, \omega) \equiv \int \varepsilon(\mathbf{r}, \mathbf{r}'; \omega) \mathbf{E}(\mathbf{r}', \omega) d\mathbf{r}'$ to account for spatial nonlocality. In the absence of external currents, $\mathbf{J}_{\text{ext}} = 0$, Eq. (8) constitutes a generalized (though still Hermitian, in the absence of loss) eigenvalue problem with real eigenvalues k_0^2 and eigenfunctions $\mathbf{E}(\mathbf{r})$, orthogonal over \mathbf{r} with respect to a weighting $\hat{\varepsilon}$ [50]. Evidently, Eq. (8) can be brought to the form of Eq. (7a) by introducing a response operator $R(\omega) = [\nabla \times \nabla \times - k_0^2 \hat{\varepsilon}]^{-1} \equiv \vec{\mathbf{G}}(\omega)$, which is just the dyadic Green function defined in operator notation, in which case $\mathbf{E} \propto \vec{\mathbf{G}}(\omega) \mathbf{J}_{\text{ext}}$. Since the Green function can always be expressed in a spectral representation [51–53], we are ensured that the poles coincide with the eigenvalues of the wave equation.

The physical interpretation of Eqs. (7a) and (7b) is rather intuitive. Specifically, it is clear that a pole of the response operator allows a finite signal s even in the presence of infinitesimal external stimuli $s_{\text{ext}} \rightarrow 0$; in other words, the optical excitations can be considered self-sustaining and intrinsic to the system defined by $\varepsilon(\mathbf{r}, \mathbf{r}'; \omega)$. In fact, though the appropriate formal identification of $R(\omega)$ lies with the dyadic Green function – ultimately allowing the concept of quasi-normal modes [53] – any optical response function will generally suffice. In the remainder of this thesis, we will see that this is the case in a range of common response functions, such as the reflection coefficient r as relevant to plane interfaces, or in scattering amplitudes t_l (or, equivalently, in multipolar polarizabilities α_l) as relevant to e.g. spherical structures.

2.2 PLASMONICS

2.2.1 The local response approximation

Though the preceding discussion is entirely general, it also conceals an essential difficulty; the full nonlocal and inhomogeneous dielectric function $\varepsilon(\mathbf{r}, \mathbf{r}'; \omega)$ is not usually known, nor is it easy to measure directly, and its dimensional attributes, indexed by seven coordinates $(\mathbf{r}, \mathbf{r}', \omega)$, are formidable. Very significant simplifications can be

The optical excitations deduced from Eq. (7b) are not generally guaranteed to be bound excitations; such a distinction can be made by using an optical potential $\Delta\varepsilon(\mathbf{r})$ relative to the background [49].

The equivalent real-space definition reading $[\nabla \times \nabla \times - k_0^2 \hat{\varepsilon}] \vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}'; \omega) = \delta(\mathbf{r} - \mathbf{r}')$.

In the presence of material loss, or indeed even of radiation loss, the optical excitation frequencies are complex; the imaginary part provides the excitation's inverse life-time.

made, however, under the assumption of short-range nonlocality and inhomogeneity relative to the wavelengths and geometric feature sizes under considerations. What is more, these simplifications lead to surprisingly accurate conclusions in an enormous range of circumstances. In the following we indicate how this approach, referred to as the local response approximation (LRA), relates with the microscopic picture of Eqs. (4).

Firstly, if $\mathbf{E}(\mathbf{r}')$ varies slowly around points \mathbf{r} , one may reasonably replace it by its Taylor approximation

$$E_i(\mathbf{r}') = E_i(\mathbf{r}) + \sum_j \frac{\partial E_i(\mathbf{r})}{\partial x_j} (x_j - x'_j) + \frac{1}{2} \sum_{j,l} (x_j - x'_j) \frac{\partial^2 E_i(\mathbf{r})}{\partial x_j \partial x_l} (x_l - x'_l) + \dots, \quad (9)$$

expressed in Cartesian coordinates $\mathbf{r}^{(l)} = [x_1^{(l)}, x_2^{(l)}, x_3^{(l)}]$ and with summation indices appropriately restricted. If $\varepsilon(\mathbf{r}, \mathbf{r}')$ is a short-range and local function in the coordinate $|\mathbf{r} - \mathbf{r}'|$ – crucially with a range much shorter than any characteristic length $E(\mathbf{r}')$ – the displacement field can then be written as a series in the field inhomogeneity [54]

Here assuming, additionally, full isotropy and correspondingly scalar $\varepsilon(\mathbf{r}, \mathbf{r}'; \omega)$.

$$D_i(\mathbf{r}, \omega) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j_1, \dots, j_n} \varepsilon_{j_1, \dots, j_n}^{[n]}(\mathbf{r}, \omega) \frac{\partial^n}{\partial x_{j_1} \dots \partial x_{j_n}} E_i(\mathbf{r}, \omega), \quad (10a)$$

expressed in terms of the n th order moment of the dielectric function, constituting a rank n tensor in the coordinate indices j_1, \dots, j_n

$$\varepsilon_{j_1, \dots, j_n}^{[n]}(\mathbf{r}, \omega) \equiv \int \varepsilon(\mathbf{r}, \mathbf{r}'; \omega) [(\mathbf{r} - \mathbf{r}')_{j_1} \dots (\mathbf{r} - \mathbf{r}')_{j_n}] d\mathbf{r}'. \quad (10b)$$

Evidently, the $n = 0$ truncation of Eq. (10a), which we shall refer to as the LRA, leads to a local dielectric function

$$\varepsilon(\mathbf{r}, \mathbf{r}'; \omega) \simeq \delta(\mathbf{r} - \mathbf{r}') \varepsilon_{\text{LRA}}(\mathbf{r}, \omega), \quad \varepsilon_{\text{LRA}}(\mathbf{r}, \omega) \equiv \int \varepsilon(\mathbf{r}, \mathbf{r}'; \omega) d\mathbf{r}', \quad (11)$$

entailing $\mathbf{D}(\mathbf{r}, \omega) \simeq \varepsilon_{\text{LRA}}(\mathbf{r}, \omega) \mathbf{E}(\mathbf{r}, \omega)$. In scenarios with multiple different dielectric constituents, the LRA is conventionally combined with a so-called hard wall approximation. This approximation constitutes a spatially piecewise description of $\varepsilon_{\text{LRA}}(\mathbf{r}, \omega)$, whereby each distinct material region is assigned the spatially constant value of $\varepsilon_{\text{LRA}}(\omega)$ as obtained from an extended bulk system. Accordingly, the hard wall $\varepsilon_{\text{LRA}}(\mathbf{r}, \omega)$ exhibits discontinuous jumps at material boundaries.

The remarkable accuracy of the hard wall LRA, Eq. (11), as established by a century of successful local response electromagnetism, vis-à-vis its crudeness relative to its full expression is a testament to the rapid convergence of $\varepsilon_{\text{LRA}}(\mathbf{r}, \omega)$ towards its bulk value, and to the truly short nonlocal range of $\varepsilon(\mathbf{r}, \mathbf{r}'; \omega)$. We return to this characteristic range in Section 3.2.1. First, however, we explore the frequency-behavior of $\varepsilon_{\text{LRA}}(\omega)$ in free-electron metals and examine its crucial role in plasmonics.

2.2.2 The Drude response model

The free conduction electrons of a metal are responsible for the majority of their optical characteristics, such as their large reflectivity. Accordingly, metals are most simply (and most often) described in a one-band model, accounting for just the conduction band. Within the jellium ansatz, i.e. the assumption of a homogeneous ionic background, the conduction electrons of bulk metals then exhibit parabolic energy-dispersion $\epsilon_{\mathbf{k}} = \hbar^2 k^2 / 2m_{\text{eff}}$ with effective mass m_{eff} . Besides m_{eff} only one additional parameter, the Fermi level ϵ_{F} – or equivalently, the free-electron equilibrium density $n_0 = (3\pi^2)^{-1} (2m_{\text{eff}}\epsilon_{\text{F}}/\hbar^2)^{3/2}$ – then specifies the problem.

The heralded Drude model solves that problem in the LRA limit. There exists a multitude of approaches to derive the Drude model, ranging in complexity from simple

Comment on applicability of jellium ansatz. The specification is occasionally made in terms of the Wigner-Seitz radius r_s , defined as the radius of a spherical volume containing on average one electron $\frac{4\pi}{3} r_s^3 = 1/n_0$ [32].

damped harmonic oscillator models [10, 55], to integrating the semi-classical Boltzmann equation in the relaxation-time approximation [56], and further still from the low momentum limits of the Lindhard formula [57]. Regardless of method, the result provides an LRA expression for the conductivity

$$\sigma_{\text{LRA}}(\omega) = \frac{i\varepsilon_0\omega_p^2}{\omega + i\gamma}, \quad (12)$$

defined in terms of the important plasma frequency $\omega_p^2 \equiv n_0 e^2 / \varepsilon_0 m_{\text{eff}}$ and a phenomenological loss-rate γ . Accordingly, the associated LRA dielectric function, see Eq. (6), reads

$$\varepsilon_{\text{LRA}}(\omega) = \varepsilon_{\text{B}}(\omega) - \frac{\omega_p^2}{\omega(\omega + i\gamma)}. \quad (13)$$

Of course, in this form, no real restrictions have been made on the actual frequency-dependence – an arbitrary dependence can still be achieved via $\varepsilon_{\text{B}}(\omega)$. Nevertheless, it is an implicit condition of the Drude-like response that $\varepsilon_{\text{B}}(\omega)$ in practice has the character of a high-frequency correction (or at least local in frequency). Specifically, the bound dielectric function is assumed to account for e.g. interband transitions and valence-band screening. Typically, interband transitions are the dominant corrections, and are modelled in practice by summed Lorentzian oscillator terms $\varepsilon_{\text{B}}(\omega) = \sum_n f_n \omega_n^2 / [\omega_n^2 - \omega(\omega + i\gamma_n)]$ with oscillator-strengths f_n , transition frequencies ω_n , and decay-rates γ_n .

When combined with a Drude model, such a model is typically referred to as a Lorentz-Drude model.

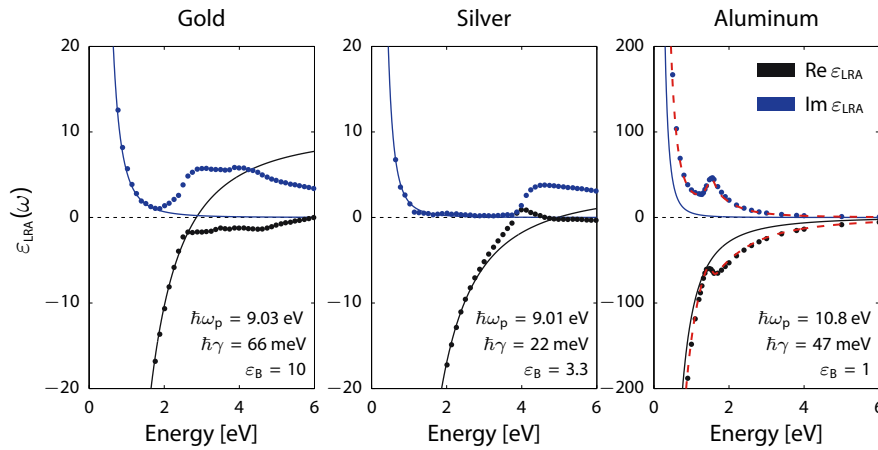


Figure 1: The Drude model in comparison with measured metals for gold, silver, and aluminum. Experimental data – obtained from the measurements of Johnson and Christy [58] for gold and silver and from Rakić [59] for aluminum – indicated in markers with corresponding Drude response in unbroken lines (parameters indicated in text-inset). For gold and silver, the plasma frequencies applied in the Drude fit reflect their actual values; conversely, for aluminum a fitted value is used, differing from its actual value of 14.98 eV. For aluminum, in dashed red is shown also a Drude-Lorentz model with four fitted oscillator terms [60].

The applicability of the Drude model is examined in Figure 1 and compared with experimental data for three metals (with the LRA dielectric function in turn inferred from reflection and transmission measurements of thin films). Evidently, the Drude model provides an excellent qualitative fit of the essential features for energies below the onset of interband excitations.

Not just metals but also strongly doped semiconductors are well described by a Drude model [61, 62] – even graphene, a semi-metal with nonparabolic dispersion, exhibits a (modified) Drude response, as we return to in Part II.

2.2.3 Classical plasmonics and the role of geometry

Having established the formal basis for a classical treatment of light-matter interaction in metals, we can now, finally, embark on the exploration and classification of various types of plasmons, distinguished essentially by the geometry of the underlying system.

BULK PLASMONS As the name suggests, bulk plasmons subsist in *bulk* metals – though also in restricted geometries, with minor modifications [63]. Arguably, they constitute the most fundamental plasmonic excitation as they exist independently of any particular geometric engineering. They are longitudinal excitations that correspond to zeros of the dielectric function. In particular, observe that Gauss law, Eq. (1c), for a translationally invariant, isotropic, and uniform jellium [in which $\epsilon(\mathbf{r}, \mathbf{r}') = \epsilon(\mathbf{r} - \mathbf{r}')$] obeys $\mathbf{k} \cdot \epsilon(\mathbf{k}, \omega) \mathbf{E}(\mathbf{k}, \omega) = \rho_{\text{ext}}(\mathbf{k}, \omega)$ in momentum-frequency space. Accordingly, for a longitudinal field, i.e. for \mathbf{E} parallel with \mathbf{k} , the total field amplitude $E(\mathbf{k}, \omega)$ is given by

$$E(\mathbf{k}, \omega) = -ik^{-1} \epsilon^{-1}(\mathbf{k}, \omega) \rho_{\text{ext}}(\mathbf{k}, \omega). \quad (14)$$

According to our definition of optical excitations Eqs. (7), the poles of $\epsilon^{-1}(\mathbf{k}, \omega)$ consequently indicate longitudinal excitations. Indeed, within the LRA Drude model, Eq. (13), these excitations evidently occur at $\omega \simeq \omega_p + \frac{1}{2}\gamma$ (assuming $\epsilon_B = 1$ and negligible loss). By merit of their co-occurrence with the plasma frequency ω_p these excitations are therefore known as bulk (or occasionally volume) plasmons.

More generally, bulk plasmons emerge as quasiparticle solutions in condensed matter theory, exhibiting also momentum dispersion [32, 57]. For a lossless free-electron jellium, within the random phase approximation (RPA), the bulk plasmons disperse as

$$\omega(\mathbf{k}) = \sqrt{\omega_p^2 + \frac{3}{5}v_F^2 k^2 + \mathcal{O}(k^4)}, \quad (15)$$

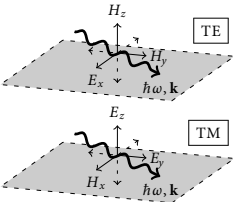
valid for small momentum $k/k_F \ll 1$, where k_F denotes the Fermi momentum, which in turn relates to the Fermi energy $\epsilon_F = \hbar^2 k_F^2 / 2m_{\text{eff}}$ and Fermi velocity $v_F = \hbar k_F / m_{\text{eff}}$. In passing, we note that the factor $\frac{3}{5}v_F^2$ constitutes the essential extension of the Drude model known as the hydrodynamic model; we return to this in Chapter 3.

From the perspective of the random phase approximation, the bulk plasmon represents collective, bosonic electron-hole pair fluctuations correlated and sustained by the Coulomb interaction. The existence – and their quantization as befits bosons – was first experimentally demonstrated by Ruthemann [64] in 1948 in various metals, but is perhaps most beautifully exhibited in the 1962-measurements in aluminum by Marton *et al.* [65]. The low- k dispersion was investigated in magnesium by Chen [66], affirming the factor 3/5 to within 10% relative deviation; in each of the three instances utilizing electron energy loss spectroscopy of thin-films.

SURFACE PLASMON POLARITONS The bulk plasmons, while fundamental, are of only meager optical interest since their polarization is strictly longitudinal, while the free electromagnetic field is transverse – thus precluding coupling between the two. Near surfaces, however, the breaking of translational invariance allows a mixed polarization state of simultaneously longitudinal and transverse character (relative to the surface). This polarization is of course that associated with the transverse magnetic (TM) state. Conversely, the transverse electric (TE) state exhibits no plasmonic features.

In the spirit of identifying excitations from their emergence as poles of a response function, it is natural to examine the Fresnel reflection coefficients. Specifically, consider an incident TM-polarized plane wave $\mathbf{E}_{\text{TM}}^{\text{inc}} \exp[i(\mathbf{k}^{\text{inc}} \cdot \mathbf{r} - \omega t)]$ impinging on an

The full dispersion is obtained from the zeros of the Lindhard dielectric function $\epsilon_L(k, \omega) = 1 - V(k)\chi_0(k, \omega)$, with Coulomb interaction $V(k) = e^2/\epsilon_0 k^2$ and noninteracting density-density response $\chi_0(k, \omega)$ given by products of fractions and logarithms of k [32, 57].



TE and TM polarization are commonly referenced as s- and p-polarization, respectively, indicating parallel and senkrecht (German; perpendicular) polarization relative to the plane of incidence.

interface separating two semi-infinite media characterized by local dielectric functions ε^+ and ε^- (above and below, respectively). The reflection coefficients links the reflected field $\mathbf{E}_{\text{TM}}^{\text{ref}} \exp[i(\mathbf{k}^{\text{ref}} \cdot \mathbf{r} - \omega t)]$ with the incident via their amplitudes $\mathbf{E}_{\text{TM}}^{\text{ref}} = r_{\text{TM}} \mathbf{E}_{\text{TM}}^{\text{inc}}$. The incident light is specularly reflected, and the associated propagation constants are accordingly $\mathbf{k}^{\text{inc}} = [0, k_{\parallel}, k_{\perp}^+]^T$ and $\mathbf{k}^{\text{ref}} = [0, k_{\parallel}, -k_{\perp}^+]^T$. Similarly, a field $\mathbf{E}_{\text{TM}}^{\text{tra}} \exp[i(\mathbf{k}^{\text{tra}} \cdot \mathbf{r} - \omega t)]$ is transmitted across the interface with $\mathbf{k}^{\text{tra}} = [0, k_{\parallel}, k_{\perp}^-]^T$. Note that the parallel component of propagation, k_{\parallel} , is conserved under both reflection and transmission by merit of the continuity of the parallel component of \mathbf{E} . The total propagation constant, however, *does* change across the interface, equaling $k^+ \equiv |\mathbf{k}^{\text{ref,inc}}| = k_0 \sqrt{\varepsilon^+}$ and $k^- \equiv |\mathbf{k}^{\text{tra}}| = k_0 \sqrt{\varepsilon^-}$. The perpendicular component $(k_{\perp}^{\pm})^2 = (k^{\pm})^2 - k_{\parallel}^2$ is similarly not conserved (and coincidentally implies Snell's law of refraction). Finally, the reflection coefficient r_{TM} is uncovered after applying the remaining BCs

$$r_{\text{TM}} = \frac{\varepsilon^- k_{\perp}^+ - \varepsilon^+ k_{\perp}^-}{\varepsilon^- k_{\perp}^+ + \varepsilon^+ k_{\perp}^-}. \quad (16)$$

By Eqs. (7) the associated excitations correspond to the poles of r_{TM} , i.e. to the condition $\varepsilon^+/\varepsilon^- = -k_{\perp}^+/k_{\perp}^-$. Evidently, since k_{\perp}^{\pm} are positive (real or imaginary) quantities by the conventional choice of square-root branch-cut, a solution is possible only if ε^- and ε^+ are of opposite sign. This sign-condition can be fulfilled at the interface between a metal and a dielectric, by assigning a Drude model to ε^- and a positive value to ε^+ . Indeed, under this assumption, the previous condition can be inverted by squaring, revealing a simple dispersion relation between k_{\parallel} and ω

$$k_{\parallel} = k_0 \sqrt{\frac{\varepsilon^-(\omega) \varepsilon^+}{\varepsilon^-(\omega) + \varepsilon^+}}, \quad (17)$$

with the assumption of frequency-(in)dependent response for the Drude component $\varepsilon^-(\omega)$ (dielectric component ε^+) made explicit. An additional condition, $\varepsilon^-(\omega) + \varepsilon^+ < 0$, is immediately apparent: otherwise the excitation will decay along the interface even in the absence of material loss. Under these conditions, the perpendicular components of the propagation become (positively) imaginary since $k_{\parallel} > k_0$. As consequence, the excitation exhibits a z -dependence of $e^{\mp \text{Im}(k_{\perp}^{\pm})z}$ (for $\pm z > 0$) and is thus exponentially localized to the interface. Accordingly, it is designated as the surface plasmon polariton (SPP).

In Figure 2 we plot the dispersion relation of the solutions to Eq. (17) for a simple Drude model $\varepsilon^-(\omega) = \varepsilon_{\text{B}}^- - \omega_{\text{p}}^2/\omega^2$ with constant ε_{B}^- and ε^+ . A number of observations are in order:

1. In the lossless case, the SPP branch exists from zero frequency up to the nonretarded surface plasmon frequency $\omega_{\text{sp}} = \omega_{\text{p}}/\sqrt{\varepsilon_{\text{B}}^- + \varepsilon^+}$. Near ω_{sp} the excitation is entirely longitudinal, and k_{\parallel} , and hence the degree of confinement, diverges. At low frequencies, the SPP is loosely bound and predominantly transversely polarized; it follows, and essentially resembles, the polariton branch of the dielectric medium. At the midpoint between these two extremes both polaritonic and plasmonic features co-exist in nearly equal measure.
2. Above the plasma frequency ω_{p} a radiative solution, known as the Brewster mode, exists. Though a mathematical solution of Eq. (17) it does not correspond to a pole, but rather a zero, of Eq. (16) – its emergence as a solution of Eq. (17) occurs because of the squaring procedure precluding it.
3. Inclusion of loss causes back-bending of the SPP branch, which furthermore becomes complex. The imaginary part of k_{\parallel} indicates the inverse decay length of the SPP. Evidently, the decay length is minimal near ω_{sp} . This constitutes the famously adversarial relationship in plasmonics between confinement ($\propto \text{Re } k_{\parallel}$) and loss ($\propto \text{Im } k_{\parallel}$).

The corresponding TE reflection coefficient is $r_{\text{TE}} = \frac{k_{\perp}^+ - k_{\perp}^-}{k_{\perp}^+ + k_{\perp}^-}$.

The polariton-addition indicates the mixing with the free polariton field in the dielectric medium, i.e. the inclusion of retardation. In the nonretarded limit it is omitted; the $k_{\parallel}/k_0 \rightarrow \infty$ solution of Eq. (17) is thus simply denoted the SP [9].

Back-bending is absent if ω rather than k_{\parallel} is chosen complex [67, 68], in this case, the relevant decay involves times rather than lengths.

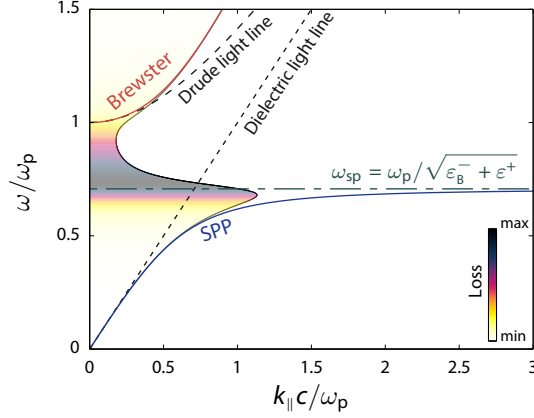


Figure 2: Solutions of the SPP dispersion equation in Eq. (17) for $\epsilon^+ = \epsilon_B^- = 1$. Blue and red full lines indicate solutions in a lossless Drude model, with physical adherence to the SPP and Brewster mode, respectively. Indicated in dashed are the light lines of the Drude and dielectric materials, k^- and k^+ , respectively. Finally $\text{Re } k_||$ is indicated for a Drude model with finite loss $\gamma/\omega_p = 0.1$ by the delimitation of the color-shaded area. The color-range reflects the relative magnitude of the imaginary part of $\text{Im } k_||$.

4. The SPP branch is generally below the dielectric light line, and so cannot be excited by propagating plane waves. Accordingly, special coupling techniques are necessary: traditionally, this is achieved by prism or grating coupling [10] – or alternatively by utilizing exciting fields with evanescent components, such as those associated with dipoles or traveling electrons, see Section 2.3.

LOCALIZED SURFACE PLASMONS Though the SPP of planar surfaces can couple to light via the coupling techniques just described, this complication is undeniably a practical setback. This setback is circumvented in truly finite structures where the geometric truncation itself provides the missing momentum necessary to couple with light. This (effective) momentum is on the order an inverse characteristic length of the structure.

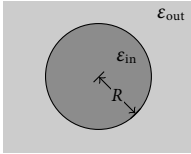
As a concrete and prototypical example, consider a spherical object of radius R , characterized by dielectric functions ϵ_{in} and ϵ_{out} inside and outside the sphere, respectively. The dipole polarizability α_1 links the induced dipole moment p to the incident field amplitude E_0 via $p = \epsilon_0 \epsilon_{\text{out}} \alpha_1 E_0$. Its generalization to higher multipole orders $l \geq 1$, known as the multipolar polarizability α_l , similarly connects multipolar moments and the incident field. For the geometry outlined here, it takes the form [69–71]

$$\alpha_l = 4\pi R^{2l+1} \frac{l(\epsilon_{\text{in}} - \epsilon_{\text{out}})}{l\epsilon_{\text{in}} + (l+1)\epsilon_{\text{out}}}, \quad (18)$$

given in the nonretarded limit (valid for spheres small relative to the excitation wavelength); its retarded generalization is treated by Mie–Lorenz theory, see Section 3.4.1. Once again, application of the excitation-condition Eqs. (7) readily reveals that the l th multipole plasmon of the sphere, characterized by frequency ω_l , fulfills

$$l\epsilon_{\text{in}}(\omega_l) + (l+1)\epsilon_{\text{out}} = 0, \quad (19)$$

where, as before, an assumed frequency-dependence of the interior medium has been explicitly emphasized. The dipole resonance, i.e. the $l = 1$ resonance, is occasionally referred to as the Fröhlich condition $\epsilon_{\text{in}}(\omega_1) = -2\epsilon_{\text{out}}$ [10, 16]. In clear analogy with the SPP case, either medium must exhibit a negative dielectric function. For a lossless Drude metal, see Eq. (13), embedded in a constant dielectric medium, the associated



Schematic of geometry.

Solving the Poisson eq. $\nabla \cdot [\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = 0$ for the potential $\phi(\mathbf{r})$ via expansion in and matching of multipole solutions of the Laplace eq. leads to Eq. (18).

dipole resonance frequency is $\omega_1 = \omega_p / \sqrt{\epsilon_B + 2\epsilon_{\text{out}}}$. Conversely, the $l \rightarrow \infty$ limit is $\omega_\infty = \omega_p / \sqrt{\epsilon_B + \epsilon_{\text{out}}}$ which recovers the SP resonance frequency of a flat interface since the local curvature of the sphere is approximately flat relative to the large- l plasmon wavelength, which is $\sim 2\pi R/l$. Regardless of multipole order, the excitation is tightly bound to the surface of the sphere, decaying approximately as $r^{-(l+2)}$ in the radial coordinate r outside the sphere. Accordingly, by the apparent resemblance with the characteristics of the plasmons discussed so far, these excitations are known as localized surface plasmons (LSPs). The terminology applies quite generally to any restricted geometry, even for cylinders with only partial restriction.

Of course, a simple complementary class of plasmons arise when the frequency-dependence of the interior and exterior media is swapped, i.e. if we assume Drude-like $\epsilon_{\text{out}}(\omega)$ and dielectric-like ϵ_{in} . Such plasmons, though arguably still contained in the LSP class, are often referred to as void plasmons [72]. In the important nonretarded limit, the complementarity between the inverse cases of void (–) and “nonvoided” (+) geometries, is dictated by a simple sum rule derived by Apell *et al.* [73]. For vacuum adjacent to a simple Drude metal without bound polarization ($\epsilon_B = 1$), it reads

$$\omega_-^2 + \omega_+^2 = \omega_p^2, \quad (20)$$

valid for single-interface restrictions and applicable to each valid quantum number separately, such as momentum \mathbf{k} or multipole order l . It is rewarding to verify the validity of this sum rule, e.g. for the SP of a flat vacuum-Drude interface where $\omega_\pm^2 = \omega_p^2/2$ or for a vacuum-Drude sphere where $\omega_{l+}^2 = \omega_p^2 \frac{l}{2l+1}$ and $\omega_{l-}^2 = \omega_p^2 \frac{l+1}{2l+1}$.

Localized surface plasmons have a privileged status in plasmonics, because their spectrum contains modes with nonvanishing dipole moments, and so allows coupling to plane wave excitations. Moreover, the coupling is surprisingly strong, to the point that a metallic nanoparticle may absorb (or scatter) more light than incident on it [16]. Indeed, the apparent cross-section of each resonance is principally limited only by the resonance-wavelength λ , through the so-called single-channel limit $3\lambda^2/2\pi$ (for dipolar excitations) [74]. In practice, however, the maximum absorption is limited by the degree of loss in the metal, i.e. by $\text{Im } \epsilon$. We return to this subject in Section 2.3.1.

COUPLED AND GENERAL VARIANTS In addition to the canonical examples of plasmons just examined, a veritable zoo of variants exist. Though generally of a more complicated nature, be it geometrically or in number of elements, most of their qualitative features can be appreciated from the examples just considered. A common feature of the more esoteric subspecies of plasmons is the coupling of SPPs and LSPs. This is essentially the case e.g. for gap and channel plasmons [75], lattice plasmons of any kind [76] such as spoof plasmon polaritons [77], or coupled nanoparticle dimers [78–80]. The coupling mechanism is provided by the free electromagnetic field; for large separations of elements dipole coupling will typically suffice. For very small separations, all multipole orders contribute – in fact, within the classical LRA treatment, divergences appear in touching configurations because a physical cutoff of very-large-order modes is absent [80]. The resolution of this issue has partially motivated many of the beyond-LRA efforts in recent years [69, 81–87].

Lastly, we touch briefly upon systems with less symmetry than those just examined. Broadly speaking, analytical expressions are unobtainable in systems that do not, in one way or another, possess a simple representation in either of the orthogonal coordinate systems, such as polar, spherical, or ellipsoidal coordinates. Though the class of solvable problems has been slightly enlarged by the techniques of transformation optics [88, 89], the vast majority of imaginable configurations require numerical treatments. A range of computational methods proliferate in plasmonics [90]. In passing, we mention a selection of these methods: the finite-difference time domain (FDTD) method [91], the finite-element method (FEM) [92], the Green function in-

The general sum rule, not restricted to specific choices of the dielectric or bound response, is $\Delta(\omega_-) + \Delta(\omega_+) = 1$ with $\Delta(\omega) \equiv \epsilon^d / [\epsilon^d - \epsilon^m(\omega)]$ for dielectric functions ϵ^d and $\epsilon^m(\omega)$ corresponding to the dielectric and metal regions, respectively [73]. This can readily be derived from the LRA-limit of Eq. (66a).

The single-channel limit is ubiquitous also outside plasmonics, e.g. in atomic physics in the interaction between classical radiation and two-level systems [47].

tegral equation method (GIEM) [93], its sibling; the discrete-dipole approximation (DDA) [94], and the boundary element method (BEM) [95]. FEM and FDTD share a common attribute in the use of a finite-size simulation domain, and are arguably the most widespread approaches, owing to their availability in commercial software such as Comsol, Lumerical, and CST. While this thesis will focus almost exclusively on systems with sufficient symmetry that analytical solutions exist, we will have opportunity to apply the BEM, as implemented in the MNPBEM toolbox [96, 97], in Section 3.5.

2.3 PROBING PLASMONIC EFFECTS

Having established the basic properties and classifications of plasmons, we are now equipped to discuss a selection of practical approaches to their experimental investigation. Already it should be clear that plasmons are associated with prominent spectral features and considerable induced physical quantities, cf. the pole-interpretation Eqs. (7). In this section, we quantify this notion further by considering the spectral fingerprints of plasmons – particularly LSPs – in far- and near-field measurements.

2.3.1 Far-field measurements

The interaction of nanoparticles, and their associated LSPs, with plane-wave excitations account for the majority of their suggested applications. Indeed, this is not too surprising, since the transverse spatial extent of conventional light-sources is limited by diffraction, i.e. by the Airy disk radius on the order of the wavelength λ [55], with most beam-widths far exceeding that. Conversely, LSP-supporting nanoparticles are deeply subwavelength. This syllogism renders plane-wave approximations – applied to plasmonic considerations – excellent in lens-based systems.

The interaction of plane waves and LSPs can be quantified by scattered, absorbed, or extinct intensity, angle-integrated in full or partial windows. Considering all angles, it is advantageous to introduce optical cross-sections of scattering, absorption, and extinction kind, defined by $\sigma_{\text{sca,abs,ext}} \equiv W_{\text{sca,abs,ext}}/I_0$, i.e. as the ratio of scattered, absorbed, or extinct power $W_{\text{sca,abs,ext}}$ to incident intensity I_0 . The latter, extinction, defines the sum of the former contributions, i.e. $\sigma_{\text{ext}} = \sigma_{\text{sca}} + \sigma_{\text{abs}}$. Thus, W_{ext} describes the “missing” power “behind” the scattering object, and σ_{ext} is recognized as the finite-scatterer analogue of $1 - T = R + A$ for transmission T , reflection R , and absorption A at a planar interface. The so-called optical theorem [47] concretizes these considerations, linking rigorously the extinct power to the forward scattering amplitude.

For the purpose of illustration, let us consider the cross-sections in the nonretarded limit. Assuming a scatterer embedded in a lossless medium of dielectric function ϵ^d , with associated bulk momentum $k^d = \sqrt{\epsilon^d}k_0$, and that only a single dipole resonance contributes (characterized by its polarizability α_1), the scattering and absorption cross-sections are [98]

$$\sigma_{\text{sca}} = (6\pi)^{-1}(k^d)^4|\alpha_1|^2, \quad \sigma_{\text{abs}} = k^d \text{Im } \alpha_1. \quad (21)$$

An important conclusion follows from these formulas when applied e.g. to a small spherical particle of dielectric function ϵ with $\alpha_1 = 4\pi R^3(\epsilon - \epsilon^d)/(\epsilon + 2\epsilon^d)$. Specifically, absorption is dominant relative to scattering in the small particle and large wavelength limits since $\sigma_{\text{abs}} \sim R^3/\lambda$ while $\sigma_{\text{sca}} \sim R^6/\lambda^4$ (ignoring material-dispersion). Accordingly, in very small particles extinction and absorption are practically equivalent (coincidentally, this equivalence is exact in a strictly nonretarded approach, constituting thereby a violation of the optical theorem; the inconsistency can, however, be remedied by including a retardation reaction term [55, 100, 101]).

As first noted by Rayleigh [99], the scaling $\sigma_{\text{sca}} \sim \lambda^{-4}$ qualitatively accounts for the color of the sky: light scattering off airborne molecules is strongest at short wavelengths, thus entailing a blue appearance.

Lastly, we note that Eq. (21) predicts a maximum absorption peak limited only by the ratio $\varepsilon^d/\text{Im } \varepsilon$, or, in a pure Drude model by ω_p/γ . Similarly, the resonance's full width at half maximum is just γ for nearly real (i.e. lossless) bound polarization. These considerations clearly demonstrate a key challenge faced by nanoplasmonics: to maximize optical response, material losses must be minimized.

2.3.2 Near-field measurements

A notable downside to far-field measurements is that they are often performed on ensembles of nanoparticles. As a result, the measured properties account also for inhomogeneities in the ensemble-distribution, which incur e.g. additional broadening – and, at sufficiently short interparticle distances, also introduce non-negligible electromagnetic coupling. Moreover, far-field measurements do not reveal information about local optical quantities related to the field pattern. A number of near-field techniques exist, however, that overcome these deficiencies, allowing both spectroscopic and spatial quantification of plasmonic properties of individual nanoparticles. Of these we highlight in particular electron energy loss spectroscopy (EELS) and fluorescence microscopy techniques, whose associated calculable quantities – the electron energy loss (EEL) probability $\Gamma(\mathbf{r}_{\parallel}, \omega)$ and electric local density of states (LDOS) $\rho_{\hat{\mathbf{n}}}^E(\mathbf{r}, \omega)$, respectively – we elaborate on in the following.

Even if we do not here discuss them further, at least two other near-field techniques warrant mention. Firstly, scanning near-field optical microscopy (SNOM) [102] has seen widespread use in the optical regime for metal plasmonics [103–105], particularly in studies of SPP propagation, and more recently its extension to the infrared facilitated the first experimental near-field investigations of graphene plasmons [106–109], a subject which we return to in Part II. Secondly, surface-enhanced Raman spectroscopy (SERS) [17] utilizes the very large local fields of plasmons through an approximate $\propto |\mathbf{E}(\mathbf{r}, \omega)/E_0|^4$ enhancement scaling of the Raman signal, allowing spatial mapping of the hot spots of the plasmonic intensity [110].

Local density of states: spontaneous emission

The partial electric LDOS $\rho_{\hat{\mathbf{n}}}^E(\mathbf{r}, \omega)$ measures, as the name suggests, the available density of optical states of frequency ω at point \mathbf{r} in the polarization direction $\hat{\mathbf{n}}$. Formally, it is associated with the imaginary part of the $\hat{\mathbf{n}}$ -projected dyadic Green function $\vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}; \omega)$, such that [55]

$$\rho_{\hat{\mathbf{n}}}^E(\mathbf{r}, \omega) = \frac{6\omega}{\pi c^2} \text{Im} \left[\hat{\mathbf{n}} \cdot \vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}; \omega) \cdot \hat{\mathbf{n}} \right]. \quad (22)$$

The orientation-averaged LDOS $\rho^E(\mathbf{r}, \omega)$ is obtained by simply replacing the projection $\hat{\mathbf{n}} \cdot \vec{\mathbf{G}} \cdot \hat{\mathbf{n}}$ by the trace $\frac{1}{3} \text{Tr} \vec{\mathbf{G}}$.

An important feature of the LDOS in plasmonic and photonic applications is its link to decay dynamics of emitters in inhomogeneous optical environments. Several possible emitter types proliferate in nanophotonics, e.g. atoms, molecules, quantum dots, and nitrogen-vacancy centers. Importantly, they can all be treated as two-level systems under suitable conditions. In turn, optically small two-level systems can be treated as dipole emitters, characterized by their transition frequency ω_{12} and their dipole transition matrix element $\mathbf{p}_{12} = |\mathbf{p}_{12}| \hat{\mathbf{n}} = \langle 1 | \hat{\mathbf{p}} | 2 \rangle$ between excited $|2\rangle$ and ground $|1\rangle$ states (with quantum mechanical dipole moment operator $\hat{\mathbf{p}} = -e\hat{\mathbf{r}}$). The spontaneous emission rate γ_{12} between excited and ground state, which depends on ω_{12} and \mathbf{p}_{12} , is not an immutable property of the emitter, but is very much susceptible to the environment. This dependency is described by the electric LDOS in what is known as the Weisskopf–Wigner theory [112], equivalent to the weak-coupling limit of field-

A full discussion of the optical LDOS requires consideration of both electric and magnetic contributions [111]. The former, however, is arguably more important than the latter, at least in a nanophotonic context, since it relates directly with the properties of electric dipoles and thus real emitters.

The operator form of the dyadic Green function was noted in Section 2.1.2. Its real-space form is obtained by taking appropriate matrix elements $\langle \mathbf{r} | \vec{\mathbf{G}}(\omega) | \mathbf{r}' \rangle = \vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}'; \omega)$, such that $\nabla \times \nabla \times \vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}'; \omega) - \varepsilon k_0^2 \vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}'; \omega) = \delta(\mathbf{r} - \mathbf{r}')$.

matter interaction. Specifically, for an emitter positioned at \mathbf{r} the spontaneous decay rate becomes [55]

$$\gamma_{12} = \frac{\pi\omega_{12}}{3\hbar\epsilon_0} |\mathbf{p}_{12}|^2 \rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}, \omega_{12}). \quad (23)$$

Decay enhancement (or quenching) is a cornerstone of cavity quantum electrodynamics, wherein it is known as the Purcell factor.

The decay enhancement relative to its rate in vacuum can consequently be obtained as the ratio between the LDOS of the actual optical environment and vacuum

$$\frac{\gamma_{12}}{\gamma_{12}^0} = \frac{\rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}, \omega_{12})}{\rho_0^{\text{E}}(\omega_{12})}, \quad (24)$$

In evaluating Eq. (24) for practical systems it is often advantageous to make use of the following classical relation: $\frac{\rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}, \omega_{12})}{\rho_0^{\text{E}}(\omega_{12})} = \frac{6\pi\epsilon_0 \text{Im}[\mathbf{p}_c^ \cdot \mathbf{E}(\mathbf{r})]}{|\mathbf{p}_c|^2 k_0^3}$ with \mathbf{E} denoting the total field due to a classical dipole \mathbf{p}_c at point \mathbf{r} , with evaluation at frequency ω_{12} understood [55].*

where $\rho_0^{\text{E}}(\omega_{12}) = \omega_{12}^2/\pi^2 c^3$ denotes the (\mathbf{r} -independent) vacuum LDOS. As we will see in Section 3.4.3 the electric LDOS can be extremely large in the vicinity of metallic nanoparticles particularly near plasmonic resonances, and accordingly associates with concomitantly extreme decay dynamics. One of the earliest experimental verifications of this concept is due to Drexhage [113], who studied the decay of fluorescent dyes at variable distances from different substrates, such as silver and gold. Experimental developments since then have significantly increased the control of surface-to-emitter distances, with demonstrations of nanometer [114] and even sub-nanometer control [115].

Electron energy loss spectroscopy

Transmission electron microscopy (TEM) has been a tremendous technological driver in a host of nanotechnology-related disciplines. Though the prevailing application-area of electron microscopy is as a nanoscopic visualization tool, its offspring, EELS, also provides a powerful basis for spatial and spectroscopic investigations of optical excitations [22, 116, 117]. The degree of spatial control is arguably unparalleled in this context, comfortably achieving Å-scale resolutions in e.g. scanning TEM (STEM), and the spectral resolution ($\lesssim 0.1$ eV) is steadily improving with developments in electron monochromators. Here we describe the basic theoretical principle of EELS from an electromagnetic perspective.

An operating STEM emits tightly collimated electrons with energies in the 40 – 300 keV range [117]. The basic principle of EELS involves measuring the tiny loss, on the order of a few eV, experienced by an electron in the beam as it passes a polarizable object, such as a nanoparticle. The origin of this loss is a back-action onto the traveling electron by the field it induces in the optical environment. To make these considerations explicit, we consider a single electron in the beam, traveling in vacuum with velocity $\mathbf{v} = v\hat{\mathbf{z}}$ and associated time-dependent path $\mathbf{r}_e(t) = \mathbf{v}t$, impacting the xy -plane in origo. In the frequency-domain, the electron is associated with a bare field [22, 47]

$$\mathbf{E}_{\text{ext}}(\mathbf{r}, \omega) = \frac{ek_v}{2\pi\epsilon_0 v\gamma} e^{ik_v z} \left[\frac{i}{\gamma} K_0 \left(\frac{k_v r_{\parallel}}{\gamma} \right) \hat{\mathbf{z}} - K_1 \left(\frac{k_v r_{\parallel}}{\gamma} \right) \hat{\mathbf{r}}_{\parallel} \right], \quad (25)$$

with coordinates $\mathbf{r} = r_{\parallel} \hat{\mathbf{r}}_{\parallel} + z\hat{\mathbf{z}}$ (\mathbf{r}_{\parallel} denotes in-plane coordinates along $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$), momentum transfer $k_v \equiv \omega/v$, Lorentz contraction factor $\gamma = 1/\sqrt{1-v^2/c^2}$, and modified Bessel functions $K_{0,1}$. From the interaction of this bare field and the optical environment, an induced field \mathbf{E}_{ind} is generated. In turn, \mathbf{E}_{ind} works back on the electron cf. the Lorentz force law, such that the electron experiences an energy loss ΔE given by

$$\Delta E = e \int_{-\infty}^{\infty} \mathbf{E}_{\text{ind}}[\mathbf{r}_e(t), t] \cdot \mathbf{v} dt. \quad (26)$$

Implicit in Eq. (26) is the assumption that the electron's path $\mathbf{r}_e(t)$ is not appreciably modified by the interaction with the induced field, i.e. an assumption of linear response, justified by the enormous differences in total (hundreds of keV) and lost (few eV) energy.

The EEL probability $\Gamma(\omega)$ is introduced as the frequency-components of this loss, i.e. defined by $\Delta E \equiv \int_{-\infty}^{\infty} \hbar\omega\Gamma(\omega) d\omega$. Accordingly, $\Gamma(\omega)$ reflects the probability of losing

energy $\hbar\omega$ during passage. By substituting $\mathbf{E}_{\text{ind}}(\mathbf{r}, t)$ by its spectral decomposition and making use of its realness, forcing $\mathbf{E}_{\text{ind}}(\mathbf{r}, \omega) = \mathbf{E}(\mathbf{r}, -\omega)^*$, one finds

$$\Gamma(\omega) = \frac{e}{\pi\hbar\omega} \int_{-\infty}^{\infty} \text{Re} \left[\hat{\mathbf{z}} \cdot \mathbf{E}_{\text{ind}}(\mathbf{r}_{\parallel} + z\hat{\mathbf{z}}, \omega) e^{-ik_v z} \right] dz, \quad (27)$$

where we have allowed for arbitrary in-plane impacts \mathbf{r}_{\parallel} . As such, the EEL probability can be determined by calculating the induced field due to the excitation Eq. (25) along the electron's trajectory. In this sense, the problem is reduced to its bare electromagnetic content: obtain the induced field due to Eq. (25) in a specified geometry.

As a final comment on the general theory of EELS, we note a rather pleasing connection with the LDOS due to García de Abajo and Kociak [118]. Concretely, the induced field can be expressed as $\mathbf{E}_{\text{ind}}(\mathbf{r}, \omega) = i\omega\mu_0 \int \vec{\mathbf{G}}_{\text{ind}}(\mathbf{r}, \mathbf{r}'; \omega) \mathbf{J}_e(\mathbf{r}') d^3\mathbf{r}'$ in terms of the induced dyadic Green function $\vec{\mathbf{G}}_{\text{ind}}$ of the optical environment and the current density $\mathbf{J}_e(\mathbf{r}') = -e\delta(\mathbf{r}'_{\parallel} - \mathbf{r}_{\parallel}) e^{ik_v z'} \hat{\mathbf{z}}$ associated with the electron. Combining this with Eq. (27) then allows

$$\begin{aligned} \Gamma(\omega) &= \frac{e^2\mu_0}{\pi\hbar} \text{Im} \left[\hat{\mathbf{z}} \cdot \langle k_v | \vec{\mathbf{G}}_{\text{ind}}(\mathbf{r}_{\parallel}, \mathbf{r}_{\parallel}; \omega) | k_v \rangle \cdot \hat{\mathbf{z}} \right] \\ &= \frac{e^2}{6\varepsilon_0\hbar\omega} \langle k_v | \rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}_{\parallel}, \mathbf{r}_{\parallel}; \omega) | k_v \rangle, \end{aligned} \quad (28a)$$

where the two-variable Fourier transform of an operator f with real-space representation $f(z, z') \equiv \langle z | f | z' \rangle$ is defined by $\langle k | f | k' \rangle = \iint_{-\infty}^{\infty} \langle k | z \rangle \langle z | f | z' \rangle \langle z' | k' \rangle dz dz' = \iint_{-\infty}^{\infty} e^{-ikz} f(z, z') e^{ik'z'} dz dz'$, and where the definition of the LDOS, Eq. (22), has been applied, revealing the interrelation between the EEL probability and the Fourier transform of a generalized LDOS along the electron trajectory

$$\langle k_v | \rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}_{\parallel}, \mathbf{r}_{\parallel}; \omega) | k_v \rangle = \iint_{-\infty}^{\infty} e^{-ik_v z} \rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}_{\parallel}, \mathbf{r}_{\parallel}; z, z'; \omega) e^{ik_v z'} dz dz'. \quad (28b)$$

It is important to note that the interrelation is not literal [119], in the sense that the quantity $\rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}_{\parallel}, \mathbf{r}_{\parallel}; z, z'; \omega)$ is not actually the LDOS $\rho_{\mathbf{n}}^{\text{E}}(\mathbf{r}; \omega)$ being rather a *nonlocal* density of states defined by analogy with Eq. (22), for non-identical coordinate evaluation of the dyadic Green function. A strict connection can be made, however, in the case of systems with translational invariance along z [76].

We return to the differences and similarities between measurements in Section 3.4.3, where we compare theoretically the extinction cross-section, EEL probability, and LDOS for a metallic nanosphere. Furthermore, in Section 3.5 we consider the experimental use of EELS in embedded silver nanoparticles.

2.4 SUMMARY

The present chapter has served to introduce the core elements of plasmonics, and some of the techniques used to study it. The necessary foundations of optical response and optical excitations, of which plasmons constitute a subset, were introduced in Section 2.1. Subsequently, we reviewed the basics of plasmonics in Section 2.2, introducing the main workhorses of the field, namely the LRA and the Drude model, and subsequently investigated their implications in different geometries, allowing us to classify and describe distinct plasmonic excitations. Lastly, in Section 2.3 we discussed theoretically the means to probe plasmons, particularly the localized kind, in far- and near-field setups, exemplified by the quantities of optical cross-section, LDOS, and EEL probability.

The identification with a generalized LDOS requires replacing $\vec{\mathbf{G}}_{\text{ind}}$ by $\vec{\mathbf{G}} = \vec{\mathbf{G}}_0 + \vec{\mathbf{G}}_{\text{ind}}$, with vacuum contribution $\vec{\mathbf{G}}_0$. The replacement makes no change to $\Gamma(\omega)$ because the $\vec{\mathbf{G}}_0$ contribution vanishes when integrated, reflecting that the electron does not incur loss on itself in vacuum.

3

NONCLASSICAL PLASMONICS

The treatment of plasmonics introduced in the preceding chapter is classical in character: it applies the macroscopic Maxwell equations to homogeneous media whose boundaries are well-defined and induced quantities respond to perturbations in a point-by-point fashion, i.e. locally. It ignores all intrinsic length scales of the electron gas, and is generally justified in doing so, given their smallness relative to both optical and structural equivalents. This relative smallness, however, is challenged by experimental developments which today allow the exploration of plasmonic properties in truly nanometric structures [21–24]. These capabilities have made pertinent the kind of scale-related limiting-type questions, with which we will concern ourselves in this chapter: roughly, how and why does the classical LRA approach deteriorate?

The chapter is organized as follows: first, we examine the general causes for the deterioration of the classical theory. Next, we specialize to just one of these causes, nonlocality, and consider its inclusion by means of a hydrodynamic model. Concretely, we apply this model to both planar and spherical configurations, focusing in the latter case especially on the impact of nonlocality in near-field vis-à-vis far-field measurements. Emerging from this consideration is an appreciation of the so-called multipole plasmons. Departing slightly from nonclassical theme of the chapter, we next analyze and observe their impact in an experimental context. Finally, in closing the chapter, we offer the outline of a general approach, which remedies to first-order several of the deficiencies of the LRA.

In fairness, aspects of quantum mechanics can be embedded in the classical approach through the frequency dependence of $\epsilon(\omega)$, e.g. by dispersion due to interband transitions etc. – the nonclassical elements discussed here extend beyond those implicit aspects.

3.1 SHORTCOMINGS OF THE CLASSICAL THEORY

The vast, varied, and indeed flourishing literature on LRA-based plasmonics is an undeniable and impressive testament to the accuracy of the LRA, and more generally to the enormous insight, both qualitative and quantitative, that can be achieved on the basis of simple conceptual approximations. Even so, the LRA and its conventional combination with the hard wall assumption of spatially piecewise dielectric functions exhibit a number of unphysical consequences and similarly omit some important physical phenomena. Exploring these consequences is profitable for mainly two reasons. Firstly, by identification of the limitations of the predominant framework, new ideas and approaches to the development of still better frameworks can be provoked. Secondly, by identifying e.g. the first-order corrections to the existing approximations, we obtain an even more general appreciation of their surprising success in the first place.

Because there are rather many shortcomings of the conventional approach, and because we do not aim to remedy them all, let us first summarize the main issues, such that we are not later led to assign an undeserving explanatory power to individual extensions – i.e. let us categorize which issues an extension *can* and *cannot* cure.

DISCONTINUITY The standard hard wall approach assumes a discontinuous transition $\epsilon_{\text{LRA}}(\mathbf{r}) = \epsilon_1 \mathbb{1}_{\Omega_1}(\mathbf{r}) + \epsilon_2 \mathbb{1}_{\Omega_2}(\mathbf{r})$ between two media restricted to adjacent regions Ω_1 and Ω_2 with local dielectric function ϵ_1 and ϵ_2 , respectively. At the interface, indicated by boundaries $\partial\Omega_1$ and $\partial\Omega_2$ separated by an infinitesimal distance but contained in their parent region, the normal component of the displacement field must be continuous, i.e. $\epsilon_1 \mathbf{E}_\perp(\partial\Omega_1) = \epsilon_2 \mathbf{E}_\perp(\partial\Omega_2)$; consequently, the normal component of \mathbf{E} is discontinuous there. For a metal $\epsilon_1(\omega) =$

The indicator function $\mathbb{1}_\Omega(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \in \Omega \\ 0 & \text{if } \mathbf{r} \notin \Omega \end{cases}$ is the generalization of the Heaviside step-function onto arbitrary regions Ω .

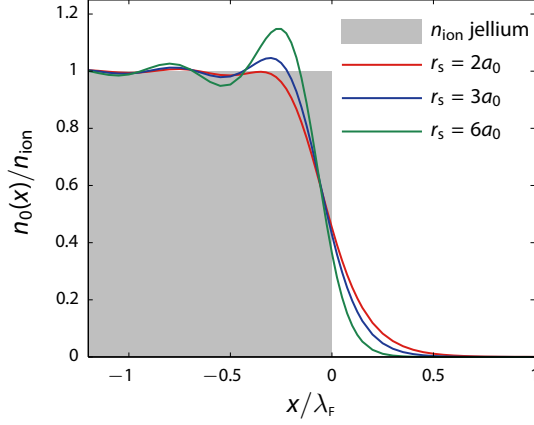


Figure 3: The electron density $n_0(x)$ relative to the corresponding (ion) jellium density n_{ion} in the vicinity of a flat interface, where the latter steps discontinuously. Coordinates are normalized to the Fermi wavelength $\lambda_F = 2\pi/k_F$. The bulk electron density is specified via the Wigner–Seitz radius r_s in units of the Bohr radius a_0 . Besides electronic spill-out for $x > 0$, Friedel oscillations are evident for $x < 0$ (being, essentially, a physical manifestation of Gibb’s phenomenon from Fourier analysis). Graphs reproduced from the tabulated data of Lang and Kohn [120], calculated in the jellium approximation of density functional theory (DFT) with Wigner’s xc-potential.

$\varepsilon_{\text{B1}} + i\sigma/\varepsilon_0\omega$ adjacent to an insulator with $\sigma = 0$, this discontinuity incurs a divergent contribution to the induced charge density [54]

$$\rho_{\text{ind}}(\mathbf{r}) = \varepsilon_0[\varepsilon_2\mathbf{E}_\perp(\partial\Omega_2) - \varepsilon_{\text{B1}}\mathbf{E}_\perp(\partial\Omega_1)]\delta(\mathbf{r} - \partial\Omega_1), \quad (29)$$

The infinite localization of ρ_{ind} also leads to an overestimation of field-enhancement in the immediate vicinity of the interface [82], to the point of divergence in the touching configuration, as appreciable from plate-capacitor reasoning [80].

as can be derived from the continuity equation $\nabla \cdot \mathbf{J}_{\text{ind}} = i\omega\rho_{\text{ind}}$. Quite obviously, such a divergence is unphysical – rather, we would expect a smooth, if peaked, behavior with a width on the order of a characteristic length of the electron gas, e.g. the Fermi length. Of course, if \mathbf{E} is tangential to the surface everywhere, i.e. transverse electric (TE) or just normally incident, this issue is circumvented – unfortunately, the induced charge is then similarly lacking, and plasmonic effects absent.

Finally, there are a number of conceptual worries. Firstly, the spill-out of the electronic gas beyond the range of the ionic lattice is completely disregarded – accordingly, fields obtained in the spill-out vicinity of the interface via a step-ansatz cannot be trusted. Secondly, it is unclear *where* the step should be positioned relative to the ionic background; this issue has bearing particularly on questions pertaining to optically induced transitions of molecules on metallic surfaces. Thirdly, and finally, the discontinuity of \mathbf{E}_\perp and the associated divergent localization of ρ_{ind} is, in the words of Feibelman [54], an extremely potent provider of momentum, cf. Fourier analysis. Physical processes that depend on momentum-transfer via the interface field may then very probably be overestimated with the step-ansatz.

It is worth noting an issue at the cross-section of locality and discontinuity – namely, that the latter cannot be physically remedied without the inclusion of the former; indeed, such attempts introduce nonphysical “bulk-like” excitations (at all frequencies!) in the spill-out region corresponding to values of $\varepsilon(\mathbf{r}, \omega) = 0$ [121].

LOCALITY Having already developed the meaning of locality in Section 2.2.1, let us recollect some consequences of its omission. Specifically, recall the underlying assumption of the LRA: the contribution of terms $\propto \partial^n \mathbf{E}$ to \mathbf{D} is negligible relative to terms $\propto \mathbf{E}$, see Eqs. (10), or, put differently, the variation of $\mathbf{E}(\mathbf{r}')$ is slow relative to the variation of $\varepsilon(\mathbf{r}, \mathbf{r}')$ in \mathbf{r}' . To overstep this assumption requires excitations of either large momentum or small spatial extent. Notably, both of these attributes are defining qualities of plasmonics. Indeed, as discussed above,

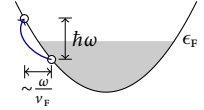
the \mathbf{E} -field is discontinuous at the interface and so cannot be considered slowly varying there. Moreover, by excluding features of nonlocality all reference to intrinsic length scales of the bulk electron gas is lost – in turn, physical cutoffs associated with such length scales are missing in LRA. This has particular impact on predictions of the LRA in scenarios of high momentum or small spatial features.

INCOMPLETENESS In principle, *all* optical excitations should feature prominently in $\varepsilon(\mathbf{r}, \mathbf{r}'; \omega)$ if built from a full set of many-body solutions to the Schrödinger equation. In practice this is impossible in all but the simplest systems, and we account only for a partial, incomplete range of phenomena. Certainly, a simple Drude model omits much, ranging from the basic, such as resonant interband transitions, to the elaborate, such as excitons. Midway between the two – and of particular importance to plasmonics – are effects explainable in a single-particle framework. One such effect is photoexcitation of electrons, i.e. the creation of electron-hole pairs (EHPs) by photon-absorption [122]. Though usually hindered by the large momentum-mismatch between electrons and photons, a plasmon can readily supply the necessary momentum $\sim \omega/v_F$ for transitions of energy $\hbar\omega$. The process, also known as Landau damping, siphons energy from the electric field and contributes to increased damping of high-momentum plasmons. The LRA, being momentum-independent, cannot account for photoexcitation and consequently underestimates plasmon damping. An effective solution is occasionally adopted through the inclusion of an additional momentum-dependent (or, equivalently, size-dependent) loss-channel [123, 124]; we return to this approach in Section 3.2.3.

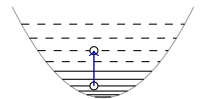
In a similar vein, the LRA overlooks the existence of longitudinal \mathbf{E} -fields. Indeed, in the LRA, the \mathbf{E} -field is constrained by transversality $\varepsilon \nabla \cdot \mathbf{E} = 0$, such that $\nabla \cdot \mathbf{E} = 0$ unless $\varepsilon = 0$. More generally, however, longitudinal waves are allowed at all frequencies if we account for nonlocal response. Furthermore, transverse and longitudinal fields may couple at material boundaries. The longitudinal components are [unsurprisingly cf. the discussion pertaining to Eq. (14)] of particular importance to properties of the bulk plasmon: they e.g. allow their spatial quantization in finite structures [63]. We return to this point in Section 3.4.

QUANTUM SIZE EFFECTS As geometric scales are reduced ever further, the assumption of a continuum of electron states eventually succumbs, necessitating consideration of the quantized nature of the electron eigenspectrum. Modifications arising from this quantization is usually referred to as quantum size effects [125–127]. The length scale beneath which such effects are *certainly* important, L_{qse} , can be roughly estimated from a ‘particle in a box’-consideration, yielding $L_{\text{qse}} \sim \pi\sqrt{3\hbar/2m_{\text{eff}}\omega}$ for perturbations (or Fermi level) of energy $\hbar\omega$. For electrons under optical excitation the scale is on the order of 0.5 to 1 nm. One consequence of quantization is the fragmentation of the main spectral peak of the classical LSP into several subpeaks [128, 129]. A related effect is one due to atomistic features, i.e. sensitivity to the exact atomic structure and termination, which significantly perturbs the jellium picture predictions in very small metallic clusters [130]. Such effects are especially pronounced in graphene due in part to large Fermi wavelengths and the existence of edge states; we return to this in Section 6.2.

In closing, we comment on relevant approaches to account for quantum size effects. Though a number of heuristic approaches exist [25, 62, 131], fully satisfactory treatments generally require both inclusion of nonlocality and at least partial treatment of the electronic many-body problem. One approach is facilitated by the RPA, which is constructed by means of a single-particle eigenspectrum, either in wave function or density representation. Particularly widespread are



Excitation of an EHP in a single semi-filled metallic band.



Quantized spectrum.

DFT and TDDFT rely formally on the Hohenberg-Kohn [132] and Runge-Gross [133] theorems, respectively. Practically, both are available in a number of implementations, see e.g. Octopus, GPAW, and VASP.

ab initio calculations relying on the framework of DFT [134] and its dynamic correspondent time-dependent density functional theory (TDDFT) [135].

3.1.1 The Feibelman d -parameter

Before proceeding to discuss concrete models and their predictions, we pause to introduce a powerful conceptual idea pioneered by Feibelman [54]. To motivate this idea, we consider a planar interface oriented normal to x , illuminated by a plane wave. As just discussed, the interface and its response in general necessitates a rather complicated and multifaceted treatment, with distinct roles played by nonlocality, spill-out, single-particle excitations, etc. From a pragmatic perspective, however, the *optical* response depends only on the induced density $\rho_{\text{ind}}(x)$. Furthermore, since the induced density is generally strongly peaked in the immediate vicinity of the material interface (i.e. at the termination of the ionic background, assumed restricted to $x \leq x_{\text{ion}}$), see Eq. (29), a multipole moment expansion of ρ_{ind} is attractive. The zeroth moment is just $\int_{-\infty}^{\infty} \rho_{\text{ind}}(x) dx$. The first (normalized) moment is now known as the Feibelman d -parameter

$$d_{\perp} \equiv \frac{\int_{-\infty}^{\infty} (x - x_{\text{ion}}) \rho_{\text{ind}}(x) dx}{\int_{-\infty}^{\infty} \rho_{\text{ind}}(x) dx}, \quad (30)$$

which, crucially, is an implicit function of the perturbation that induced ρ_{ind} , entailing e.g. frequency- and momentum-dependence. The d -parameter evidently measures the centroid of induced charge relative to the ionic interface. For instance, within the LRA treatment one readily finds $d_{\perp}^{\text{LRA}} = 0$ cf. Eq. (29). Accounting for spill-out and nonlocality allows induced densities that proliferate on either side of the interface, corresponding to plasmon spill-in ($\text{Re } d_{\perp} < 0$) or spill-out ($\text{Re } d_{\perp} > 0$).

The zeroth and first moments yield monopole and dipole contributions to the reflected field. Considering reflection of a wave impinging from vacuum onto a medium with LRA dielectric function $\varepsilon(\omega)$, defined from Eq. (11), the monopole yields the classical nonretarded reflection coefficient $r_0^{\text{TM}}(\omega) = [\varepsilon(\omega) - 1]/[\varepsilon(\omega) + 1]$, obtained from the $k_0/k_{\parallel} \rightarrow 0$ limit of Eq. (16). The dipole contribution modifies this result in a rather appealing manner, namely [136, 137]

$$r_0^{\text{TM}}(k_{\parallel}, \omega) = \frac{\varepsilon(\omega) - 1 + [\varepsilon(\omega) - 1]k_{\parallel}d_{\perp}}{\varepsilon(\omega) + 1 - [\varepsilon(\omega) - 1]k_{\parallel}d_{\perp}}. \quad (31)$$

It is instructive to consider the modifications to the usual SP resonance, discussed in Section 2.2.3, which, for $\varepsilon(\omega) = 1 - \omega_p^2/\omega^2$, resonates at $\omega_{\text{sp}}^{\text{clas}} = \omega_p/\sqrt{2}$. The corresponding corrected resonance, ω_{sp} , follows from Eqs. (7) and (31)

$$\omega_{\text{sp}} = \frac{\omega_p}{\sqrt{2}} \sqrt{1 - k_{\parallel}d_{\perp}} \simeq \frac{\omega_p}{\sqrt{2}} \left(1 - \frac{1}{2}k_{\parallel}d_{\perp}\right). \quad (32)$$

From this result follows a number of simple and general conclusions [54]:

1. The real part of d_{\perp} shifts the resonance relative to its classical counterpart: plasmon spill-in ($\text{Re } d_{\perp} < 0$) entails a blueshift, while plasmon spill-out ($\text{Re } d_{\perp} > 0$) entails a redshift.
2. The imaginary part of d_{\perp} induces additional broadening. This broadening corresponds to the creation of electron-hole pairs in the interface region.
3. At normal incidence, $k_{\parallel} = 0$, the dipole contribution vanishes and the classical result is recovered.

For clarity, the (k_{\parallel}, ω) -dependence of d_{\perp} has been suppressed in Eq. (31) but is implicitly assumed. As a result, Eq. (32) defines an implicit equation in ω_{sp} rather than an explicit expression.

A curious consequence of d_{\perp} is the existence of an additional surface mode, historically termed the multipole plasmon mode of a flat interface (not to be confused with its spherical-geometry counterpart), coinciding with poles in $d_{\perp}(k_{\parallel}, \omega)$. Its experimental existence was verified by Tsuei et al. [137] by angle-resolved electron scattering.

A natural follow-up consideration is whether such concepts generalize to curved geometries. Indeed they do, as discussed by Apell and Ljungbert [138, 139] for a sphere of ionic radius R . The d -parameter is necessarily modified by the shift from Cartesian x -restriction to radial r -restriction, hence motivating the introduction of $d_r \equiv \int_0^\infty r^2 (r - R) \rho_{\text{ind}}(r) dr / \int_0^\infty r^2 \rho_{\text{ind}}(r) dr$. Even so, its actual value is not significantly different from d_\perp for reasonable radii of curvature: indeed, an expansion in the curvature $1/R \simeq 0$ yields $d_r \simeq d_\perp + \mathcal{O}(\lambda_e/R)$ [139], with λ_e denoting a relevant length scale of the electronic gas, such as the screening length. The classical multipolar polarizability, see Eq. (18), is then modified by the accounting of the dipole correction due to d_r , yielding for a $\varepsilon(\omega)$ -sphere in vacuum [138] (frequency-dependence suppressed)

$$\alpha_l = 4\pi R^{2l+1} \frac{l(\varepsilon - 1) + l(\varepsilon - 1)k_l d_r}{l\varepsilon + l + 1 - (l + 1)(\varepsilon - 1)k_l d_r}, \quad (33)$$

with the effective momentum $k_l \equiv l/R$, and with the d_r -corrections approaching the same form as in Eq. (31) for $l \rightarrow \infty$, as expected. Again, a modified resonance condition can be deduced, yielding for the pure Drude case

$$\omega_l^{\text{LSP}} = \omega_p \sqrt{\frac{l}{2l + 1}} \sqrt{1 - k_{l+1} d_r}, \quad (34)$$

whose interpretation follows Eq. (32)'s fully.

3.2 BEYOND THE DRUDE MODEL

As should be clear from the preceding considerations, a veritable plethora of effects conspire to muddle the conclusions of the hard wall LRA approach, particularly under high-momentum, small-feature size, or near-field scenarios. Though sophisticated numerical methods, in particular TDDFT in frequency- and time-domains, can shed light on all aspects simultaneously, their computational nature complicates direct physical interpretation. Furthermore, the correlation between computational effort and system size in such *ab initio* approaches is fiercely unfavorable, limiting their practical application to systems of significant symmetry or of truly nanoscopic extent.

At the opposite end of the spectrum are semi-classical models that account partially for a selection of shortcomings of the classical theory. Though one cannot hope for *generally* faithful conclusions from semi-classical models across all materials and circumstances, their comparative simplicity provides a window to physical insight. In this section, and indeed in the remainder of this chapter, we will restrict our efforts to remedying a single of the issues raised in Section 3.1: nonlocality, and even then only to first order. Accordingly, issues related to discontinuity and quantum size-effects persist; though incompleteness may be at least partially resolved. Even so, the corrections arising from the first-order accounting of nonlocality, known as the hydrodynamic model (HDM), are striking in their own right. Moreover, under certain experimental circumstances they are the dominant corrections to the hard wall LRA. Below we first provide a semi-classical introduction to the HDM in Section 3.2.1, followed by a comparison with its rigorous generalization, the Lindhard model, in Section 3.2.2; in Section 3.2.3 we detail how a recent extension of the HDM incorporates Kreibig damping, and finally discuss the practical applicability of the HDM to real metals in Section 3.2.4.

3.2.1 The hydrodynamic model

The HDM can be derived in a variety of manners. One popular approach relies on Hamilton's principle applied to the Hohenberg–Kohn ground state Hamiltonian [132],

The key approximation of DFT, and TDDFT by extension, namely the xc potential, has only modest impact on predicted plasmonic properties, that chiefly depend on self-consistent Hartree coupling. The omission of the xc potential in dynamic aspects constitutes the RPA [85, 140].

partitioned into contributions of classical (classical kinetic energy, inter-electronic Hartree interaction, and electron-ion interaction) and quantum origin (xc terms and internal kinetic energy) [141–143], with the latter accounted for through a density-dependent functional $G[n]$. The hydrodynamic model then results from neglecting xc effects, accounting only for quantum terms due to internal kineticism. An advantage of this approach is that it is generalizable by inclusion of ever more corrections to $G[n]$, eventually leading to orbital-free DFT [144]. Here, however, we will follow a fully semi-classical approach, based on the Boltzmann equation. We do this partly to underscore the semi-classical aspects of the HDM; partly to showcase an alternative approach, which is perhaps less well-known to the nanoplasmonic community.

This choice corresponds to including only the Thomas–Fermi term in $G[n] \propto \frac{\hbar^2}{m_{\text{eff}}} \int n^{5/3}(\mathbf{r}) d^3\mathbf{r}$.

MOMENTS OF THE BOLTZMANN EQUATION The Boltzmann equation for the one-particle distribution function $f_{\mathbf{p}}(\mathbf{r}, t)$ in canonical position-momentum coordinates \mathbf{q} and \mathbf{p} , respectively, at time t reads [145]

$$[\partial_t + \mathbf{v}_{\mathbf{p}} \cdot \nabla_{\mathbf{r}} + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}}] f_{\mathbf{p}}(\mathbf{r}, t) = \mathcal{K}[f_{\mathbf{p}}(\mathbf{r}, t)], \quad (35)$$

with velocity $\mathbf{v}_{\mathbf{p}} = \mathbf{p}/m_{\text{eff}}$ and collision operator \mathcal{K} . Macroscopic external forces are included via $\dot{\mathbf{p}}$ which, by virtue of the total derivative, may depend on (\mathbf{r}, t) in addition to \mathbf{p} . In the presence of an electric field $\mathbf{E}(\mathbf{r}, t) = -\nabla_{\mathbf{r}}\phi(\mathbf{r}, t)$, the Lorentz force law requires $\dot{\mathbf{p}} = e\nabla_{\mathbf{r}}\phi(\mathbf{r}, t)$. The momentum-moments of the microscopic one-particle distribution function *define* the macroscopic observables, such as the particle density $n(\mathbf{r}, t)$, velocity flow $\mathbf{v}(\mathbf{r}, t)$, pressure-tensor $\mathbf{P}(\mathbf{r}, t)$, energy density, and so forth. For our considerations it suffices to introduce the former two:

$$n(\mathbf{r}, t) \equiv \int f_{\mathbf{p}}(\mathbf{r}, t) d^3\mathbf{p}, \quad n(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t) \equiv \int \mathbf{v}_{\mathbf{p}} f_{\mathbf{p}}(\mathbf{r}, t) d^3\mathbf{p}. \quad (36)$$

The 1- and $\mathbf{v}_{\mathbf{p}}$ -moments of Eq. (35) correspondingly provides equations of motion for $n(\mathbf{r}, t)$ and $\mathbf{v}(\mathbf{r}, t)$. After some labor, and using that the moments of 1 and \mathbf{p} of the collisional integral vanish, due to particle and momentum conservation, one finds (omitting explicit declaration of variable-dependence and restoring $\nabla \equiv \nabla_{\mathbf{r}}$)

$$\partial_t n + \nabla \cdot (n\mathbf{v}) = 0, \quad (37a)$$

$$m_{\text{eff}} n [\partial_t + (\mathbf{v} \cdot \nabla)] \mathbf{v} = -en\mathbf{E} - \nabla \cdot \mathbf{P}, \quad (37b)$$

which are recognized as the continuity and Cauchy momentum equations (from which the Navier–Stokes equations stem), expressing the conservation of integrated density and momentum. The last term of Eq. (37b) conceals the primary deficiency of this approach: \mathbf{P} corresponds to a 2nd order moment of $f_{\mathbf{p}}(\mathbf{r}, t)$. Accordingly, \mathbf{P} is governed by the equation resulting from the 2nd order moment of Eq. (35), which in turn couples to a 3rd order moment of $f_{\mathbf{p}}(\mathbf{r}, t)$; in general the k th moment equation couples to a $k + 1$ th moment. In other words, the equations do not close, and an appropriate phenomenological truncation is required. The hydrodynamic equations – as applied in nanoplasmonics – emerges from the ansatz that \mathbf{P} is diagonal, which is reasonable if $n(\mathbf{r}, t)$ is not too far from equilibrium, and that it accounts only for degeneracy pressure, arising from the Pauli exclusion principle, leading to

$$\mathbf{P} = P\mathbf{I} = \frac{\hbar^2}{m_{\text{eff}}} \frac{(3\pi^2)^{2/3}}{5} n^{5/3} \mathbf{I}, \quad (38)$$

as can be verified, by taking the (negative) volume-derivative of the total internal energy of a noninteracting Fermi gas of fixed density and volume; this amounts to a Thomas–Fermi treatment.

LINEARIZED RESPONSE The hydrodynamic equations derived thus far, Eqs. (37), contain nonlinearities in every term except $\partial_t n$. Though these nonlinearities hold interesting consequences [147–149], the main plasmonic relevance of Eqs. (37) lie in the

We restrict our consideration to the nonretarded regime: the linearized hydrodynamic equations (in the absence of driving \mathbf{B} -fields) do not differ when retardation is included [143].

Specifically, $\mathbf{P}(\mathbf{r}, t) \equiv m_{\text{eff}}^{-1} \int \mathbf{p}\mathbf{p} f_{\mathbf{p}+m_{\text{eff}}\mathbf{u}}(\mathbf{r}, t) d^3\mathbf{p}$ (with $\mathbf{p}\mathbf{p} \equiv \mathbf{p} \otimes \mathbf{p}$) measures the local pressure due to fluctuations in a co-moving frame.

In principle, the truncation can be made rigorous rather than phenomenological by decoupling higher order moment equations through a gradient-expansion around a local equilibrium; the procedure is known as the Chapman–Enskog method [145, 146].

linearized regime. The linearization is achieved by expanding in (static and homogeneous) equilibrium quantities, such as n_0 , and (dynamic) induced quantities, such as n_{ind} . Additionally, we assume the system is initially unperturbed, such that the \mathbf{E} -field is inherently first-order, and at rest such that $\mathbf{v} = \mathbf{v}_{\text{ind}}$, in which case Eqs. (37) and (38) combine to yield

$$\partial_t \rho_{\text{ind}} + \nabla \cdot \mathbf{J}_{\text{ind}} = 0, \quad (39a)$$

$$\partial_t \mathbf{J}_{\text{ind}} = \varepsilon_0 \omega_p^2 \mathbf{E} - \beta^2 \nabla \rho_{\text{ind}}, \quad (39b)$$

where the charge density $\rho_{\text{ind}} \equiv -en_{\text{ind}}$, charge current $\mathbf{J}_{\text{ind}} \equiv -en_0 \mathbf{v}_{\text{ind}}$, and plasma-frequency $\omega_p^2 = e^2 n_0 / \varepsilon_0 m_{\text{eff}}$ have been reintroduced. Moreover, we have introduced the hydrodynamic velocity $\beta^2 \equiv \frac{1}{3} v_F^2$, which is fixed by the Fermi energy ϵ_F (or density n_0) through the Fermi velocity v_F via $\epsilon_F = \frac{1}{2} m_{\text{eff}} v_F^2 = (\hbar^2 / 2m_{\text{eff}}) (3\pi^2 n_0)^{2/3}$. Finally, Eqs. (39a) and (39b) can be combined to yield closed equations for \mathbf{J}_{ind} or n_{ind}

$$\left[\partial_t^2 - \beta^2 \nabla (\nabla \cdot) \right] \mathbf{J}_{\text{ind}} = \varepsilon_0 \omega_p^2 \partial_t \mathbf{E}, \quad (40a)$$

$$\left[-\partial_t^2 + \beta^2 \nabla^2 \right] \rho_{\text{ind}} = \varepsilon_0 \omega_p^2 \nabla \cdot \mathbf{E}. \quad (40b)$$

Finally, incorporating the phenomenological relaxation rate γ and converting to the frequency domain, we uncover

$$\left[1 + \frac{\beta^2}{\omega(\omega + i\gamma)} \nabla (\nabla \cdot) \right] \mathbf{J}_{\text{ind}} = \sigma_{\text{LRA}} \mathbf{E}, \quad (41)$$

which is recognized as a (hydrodynamic) generalization of Ohm's law, reducing to the conventional LRA-Drude form of Eq. (12) in the $\beta \rightarrow 0$ limit.

CORRECTED β -PREFACTOR Before proceeding we correct an implicit assumption in the pressure-ansatz, Eq. (38), which is unsuitable for plasmonics. Specifically, the ansatz tacitly relies on a quasi-static picture of the electron gas, i.e. it assumes decay rates much larger than oscillation frequencies $\gamma \gg \omega$. In fact, in all cases of plasmonic interest the opposite is true $\omega \gg \gamma$. Under these conditions, the motion of the induced density is essentially constrained to follow the perturbation (along the direction of \mathbf{E}), i.e. the motion is one-dimensional rather than three-dimensional. As pointed out by Jackson [150] and reiterated by Fetter [151], this entails that the low- and high-frequency limits of β differ by

$$\beta^2 = \begin{cases} \frac{1}{d} v_F^2 & \text{for } \gamma \gg \omega, \\ \frac{3}{d+2} v_F^2 & \text{for } \omega \gg \gamma, \end{cases} \quad (42)$$

for a d -dimensional gas. The intermediate frequency problem was later resolved by Halevi [152] by employing the Mermin-corrected relaxation-time approximation [153], yielding a frequency dependent β -factor

$$\beta^2(\omega) = \frac{\frac{3}{5}\omega + \frac{1}{3}i\gamma}{\omega + i\gamma} v_F^2, \quad (43)$$

for the three-dimensional case; evidently, the limiting cases of Eq. (42) are recovered. For simplicity, however, we will employ the simple high-frequency limit in our work with the hydrodynamic model, i.e. $\beta^2 = \frac{3}{5} v_F^2$ for the three-dimensional case.

LONGITUDINAL AND TRANSVERSE DIELECTRIC RESPONSE As the penultimate step in our treatment of the fundamentals of the HDM, we consider the implications of Eq. (41) for the total dielectric function. In particular, including also the bound

Unsurprisingly, the corresponding two-dimensional result is [154]

$$\beta^2(\omega) = \frac{\frac{3}{4}\omega + \frac{1}{2}i\gamma}{\omega + i\gamma} v_F^2.$$

The splitting's validity is guaranteed by the Helmholtz decomposition theorem.

Note that the hydrodynamic bulk plasmon associated with the zeros of Eq. (44a) agrees with the RPA result of the free-electron gas (i.e. where $\epsilon_b = 1$) previously stated in Eq. (15) up to second order in k^2 . We return to this point in Section 3.2.2.

A second ABC is required for the case of two adjacent hydrodynamic metals, see e.g. Ref. 158. Our considerations will, however, be restricted to the simpler metal-dielectric interface.

response via ϵ_B , see Section 2.1.1, the previous considerations can be reformulated to apply instead to the displacement field \mathbf{D} . The reformulation is achieved by splitting the field into longitudinal and transverse components, \mathbf{E}_L and \mathbf{E}_T , which are characterized by their irrotational (curl-free) and solenoidal (divergence-free) qualities, $\nabla \times \mathbf{E}_L = 0$ and $\nabla \cdot \mathbf{E}_T = 0$, respectively. These fields correspond to oscillations parallel or perpendicular to the propagation direction. A similar decomposition can naturally be made for the current \mathbf{J}_{ind} . Accordingly, Eq. (41) decouples into a pair of uncoupled equations for the longitudinal and transverse components. Finally, by a momentum transform, a decomposition of the displacement field can be achieved via $\mathbf{D}_L(k, \omega) = \epsilon_L^{\text{HDM}}(k, \omega)\mathbf{E}_L(k, \omega)$ and $\mathbf{D}_T(k, \omega) = \epsilon_T^{\text{HDM}}(\omega)\mathbf{E}_T(k, \omega)$ with [142, 155]

$$\epsilon_L^{\text{HDM}}(k, \omega) = \epsilon_B(\omega) - \frac{\omega_p^2}{\omega(\omega + i\gamma) - \beta^2 k^2}, \quad (44a)$$

$$\epsilon_T^{\text{HDM}}(\omega) = \epsilon_B(\omega) - \frac{\omega_p^2}{\omega(\omega + i\gamma)}. \quad (44b)$$

Notably, the HDM imparts nonlocality only to the longitudinal response as evidenced by the k -independence of Eq. (44b).

ADDITIONAL BOUNDARY CONDITIONS It is a curious fact, that the choice of appropriate BCs for the HDM (as applied to electrodynamics) has led to confusion at least twice historically. Specifically, the choice was a matter of some contention during the 1970's and early 1980's until a consistent choice was distinguished [142, 156–158]. Surprisingly, following the resurgence of the HDM in recent years, the issue reemerged sporadically [81, 159–163].

The need for an additional boundary condition (ABC), in addition to the standard BCs derivable by pill-box arguments applied to the macroscopic Maxwell equations, is apparent from the fact that a new longitudinal wave is supported in the HDM [or, alternatively, by the differential nature of Eq. (41), which incurs additional undetermined integration constants]. For a hard wall boundary between a hydrodynamic metal and a dielectric, i.e. for step-like equilibrium densities $n_0(\mathbf{r}) = n_0 \mathbb{1}_\Omega(\mathbf{r})$, the correct ABC follows directly from integration of the continuity equation, Eq. (39a), which, when combined with charge conservation and the divergence theorem requires that

$$\mathbf{J}_{\text{ind}} \cdot \hat{\mathbf{n}} = 0, \quad (45)$$

everywhere on the boundary $\mathbf{r} \in \partial\Omega$ (with associated normal $\hat{\mathbf{n}}$). Put into words, the normal component of the free-electron current density is continuous across the boundary (being zero in the dielectric region). Occasionally this statement is rephrased by noting that $\mathbf{D} \cdot \hat{\mathbf{n}}$ is required to be continuous. \mathbf{D} contains contributions from *both* bound and free response, see Section 2.1.1, such that $\mathbf{D}^m = \epsilon_B^m \mathbf{E} + i\omega^{-1} \mathbf{J}_{\text{ind}}$ and $\mathbf{D}^d = \epsilon^d \mathbf{E}$ in the metal and dielectric regions, respectively. Matching $(\mathbf{D}^m - \mathbf{D}^d) \cdot \hat{\mathbf{n}} = 0$ for $\mathbf{r} \in \partial\Omega$ with Eq. (45) then demands that [164, 165]

$$\epsilon_B^m \mathbf{E}(\mathbf{r} \in \partial\Omega^m) \cdot \hat{\mathbf{n}} = \epsilon^d \mathbf{E}(\mathbf{r} \in \partial\Omega^d) \cdot \hat{\mathbf{n}}, \quad (46)$$

with $\partial\Omega^{m,d}$ the metal- or dielectric-oriented equivalent of $\partial\Omega$, respectively. Accordingly, the discontinuity of \mathbf{E}_\perp noted in Section 3.1 persists, but is modified and slightly lessened. *Only* in the case of a pure Drude metal $\epsilon_B^m = 1$ adjacent to vacuum $\epsilon^d = 1$ is the issue fully resolved (more generally, if $\epsilon_B^m = \epsilon^d$). Consequently, an ABC corresponding to continuous \mathbf{E}_\perp across interfaces apply only under this rather strong constraint, underscoring the inappropriateness of its use to actual metals – despite examples of past transgressions [81, 161–163].

3.2.2 The Lindhard model

It is interesting to compare the HDM with a member of a “higher” hierarchy of response formalism. The natural first candidate is the RPA result for the free-electron gas. For the longitudinal response in three dimensions this (analytical) result is known as the Lindhard dielectric function [32, 56, 57]. Its derivation follows from the combination of the noninteracting density-density response χ^0 and the RPA approximation $\epsilon^{\text{RPA}} = 1 - e^2 V \chi^0$ with Coulomb interaction V . Though an analytical form exists, it is for our purposes more appealing to work with its integral form

$$\begin{aligned}\epsilon_L^{\text{RPA}}(k, \omega) &= 1 - \frac{e^2}{\epsilon_0 k^2} \frac{2}{\mathcal{V}} \sum_{\mathbf{q}} \frac{f_{\mathbf{q}+\mathbf{k}} - f_{\mathbf{q}}}{\epsilon_{\mathbf{q}+\mathbf{k}} - \epsilon_{\mathbf{q}} - \hbar(\omega + i\eta)} \\ &= 1 + \frac{2e^2}{\epsilon_0 k^2} \frac{2}{\mathcal{V}} \sum_{\mathbf{q}} f_{\mathbf{q}} \frac{\epsilon_{\mathbf{q}+\mathbf{k}} - \epsilon_{\mathbf{q}}}{(\epsilon_{\mathbf{q}+\mathbf{k}} - \epsilon_{\mathbf{q}})^2 - \hbar^2(\omega + i\eta)^2},\end{aligned}\quad (47)$$

with electron energies $\epsilon_{\mathbf{q}} = \hbar^2 q^2 / 2m_{\text{eff}}$, associated Fermi–Dirac occupation functions $f_{\mathbf{q}} = \{1 + \exp[(\epsilon_{\mathbf{q}} - \epsilon_F) / k_B T]\}^{-1}$ at temperature T , infinitesimal loss $\eta \rightarrow 0^+$, sample volume \mathcal{V} , and where the second equality follows from inversion ($\pm \mathbf{q}$) symmetry. It is a simple matter next to expand this result around small momentum-transfers q , allowing [57]

$$\begin{aligned}\epsilon_L^{\text{RPA}}(k, \omega) &= 1 - \frac{1}{(\omega + i\eta)^2} \frac{e^2}{\epsilon_0 m_{\text{eff}}} \frac{2}{\mathcal{V}} \sum_{\mathbf{q}} f_{\mathbf{q}} - \frac{1}{(\omega + i\eta)^4} \frac{e^2}{\epsilon_0 m_{\text{eff}}} k^2 \frac{2}{\mathcal{V}} \sum_{\mathbf{q}} f_{\mathbf{q}} v_{\mathbf{q}}^2 - \dots \\ &= 1 - \frac{\omega_p^2}{(\omega + i\eta)^2} \left[1 - \frac{3}{5} \frac{v_F^2}{(\omega + i\eta)^2} k^2 - \dots \right],\end{aligned}\quad (48)$$

with the last step facilitated by identifying $\frac{2}{\mathcal{V}} \sum_{\mathbf{q}} f_{\mathbf{q}} = n_0$ and calculating $\frac{2}{\mathcal{V}} \sum_{\mathbf{q}} f_{\mathbf{q}} v_{\mathbf{q}}^2 = \frac{3}{5} n_0 v_F^2$ in the low-temperature limit ($v_{\mathbf{q}} = \hbar q / m_{\text{eff}}$). Ignoring loss, it is evident that this result agrees to second order with its HDM-equivalent in Eq. (44a). More precisely, the HDM result represents the [0/2] Padé approximation of Eq. (47) [168]. In this sense, we concludingly summarize our introduction to the HDM by noting that the HDM (i) represents the lowest order correction to the Drude model, (ii) is consistent with the Lindhard model, (iii) is equivalent with a Thomas–Fermi treatment, and (iv) only introduces nonlocality in the longitudinal response.

3.2.3 Kreibig broadening and generalized nonlocal optical response

The nonretarded description of plasmons in very small (or high momentum) nanoparticles predicts scale-independent resonance properties. As we will see in Sections 3.3 and 3.4 the introduction of nonlocality in the form of the HDM modifies this conclusion. Concretely, the HDM predicts a shifting of the resonance frequency: for instance, in nanospheres of radius R the shift is $\propto \beta/R$. The resonance width (or, equivalently, the effective loss), however, is largely unaffected because β is a real parameter [predominately so at least, cf. Eq. (43)]. Nevertheless, we *do* expect a size- or momentum-dependent loss, because large momenta, either real or effective, allow non-vertical photoexcitation of EHPs in the surface region as discussed in Section 3.1. Indeed, the effect of size-dependent damping is well-established experimentally, wherein it is known as Kreibig damping [123, 169]. A pragmatic remedy for nanospheres is to introduce size-dependence in the bulk Drude loss via

$$\gamma \rightarrow \gamma + A v_F / R. \quad (49)$$

Mathematically, Eq. (49) is simply a first order Taylor expansion in v_F / R with coefficient A . It is an experimental fact, corroborated also by TDDFT calculations [170],

Named after the Danish physicist Jens Lindhard [166].

The real-space equivalent of the present momentum-expansion procedure is also possible, though more laborious in its nonlocal consequences, see LRA-restricted efforts in Ref. [167].

If η is taken as a (phenomenologically introduced) finite, but small “microscopic” decay rate, its correspondence with the prior “macroscopic” decay rate γ is achieved by fixing $\gamma = 2\eta$.

The modification suggests a deceptively simple classical interpretation: the electrons, moving at velocity v_e , “bounce” off the spherical walls at a rate $\sim v_e / R$ thereby contributing additional decoherence. This is oversimplified, however: electrons are not balls and the origin of the increased loss is the creation of EHPs.

that A is on the order of unity for the dipole resonances of spheres [123, 169]; this is at least qualitatively appreciable from the observation that the nanosphere curvature provides an effective momentum $\sim 1/R$, while a small EHP transition of energy $\hbar\omega$ requires momentum $\sim \omega/v_F$, see Section 3.1.

The primary drawback of the Kreibig model, Eq. (49), is that it applies only to the dipole LSP of spheres: its generalization to e.g. structures with sharp tips, multiple scattering scenarios, or even to non-dipolar sphere-LSPs is, a priori, unclear. One approach to circumvent this excessive specificity, is suggested by the $\propto \beta/R$ frequency shifts of spheres in the HDM: a complex β parameter would naturally incur size-dependent broadening, being generalizable to arbitrary structures through Eq. (41). This is, in rough measure, the motivation for the recently introduced generalized non-local optical response (GNOR) approach [86, 87, 165]. Specifically, within GNOR the β -factor of HDM is replaced by a generalized parameter η , incorporating both the degeneracy-pressure of the HDM and Kreibig broadening

$$\eta^2 = \beta^2 + D(\gamma - i\omega). \quad (50)$$

Here D denotes a “diffusion” length, which originates in the semi-classical approach taken to derive GNOR. Though the semi-classical picture involves diffusion, it is clear that the actual quantum mechanical situation cannot be understood satisfactorily except in terms of electronic transitions. As such, we shall make no further attempt to do so, but simply note that D in principle is material-dependent, and that it can be obtained from measurements by appropriate fitting [87]. Alternatively, from a small-loss nonretarded expansion of the GNOR sphere dipole resonance, reached in Eq. (62), combined with Eq. (49), one finds that D , up to orders of unity, equals $D \sim Av_F^2/\omega_p$. We will use this approximate form, as our considerations will be limited to strictly principal consequences of the GNOR model.

3.2.4 Applicability of the hydrodynamic model

Finally, we offer a few short comments on the applicability of the HDM to actual metals. A meaningful discussion requires us to state in advance a cardinal feature of the HDM, which we shall establish more thoroughly in the following section. Specifically, a hydrodynamic approach, which retains the ansatz of “hard” separations between material regions, i.e. ignores spill-out, invariably leads to resonances that shift toward higher energies than the classical LRA-counterpart, i.e. blueshift. Experimentally, this is indeed also observed for e.g. silver [25, 26, 171–174]. In the archetypal free-electron metals, such as the alkalis and aluminum, which are only vaguely screened by lower-lying orbitals, the opposite is true, and the resonances redshift [175, 176]. A different perspective on this difference is provided by the Feibelman d -parameter discussed in Section 3.1.1: in the blueshifting case, the plasmon spills inward, while, in the redshifting case, it spills outward. The origin of these disparities has been discussed by Liebsch [177, 178], and assigned essentially to the presence vis-à-vis absence of significant background screening. In this sense, the noble metals, which exhibit significant background screening from the d -orbitals, constitute examples expected to be in line with hydrodynamic predictions.

Ultimately, the strength of the HDM rests predominantly in its appropriate account of *bulk* properties, as evidenced e.g. by its low-momentum agreement with the Lindhard model cf. Section 3.2.2. In the following, we proceed with this and the above facts in mind, recognizing that the HDM only describes a partial account of the full picture. In some cases, such as noble metals, this partial picture will contain the decisive and dominant elements – while, in others, which we venture to avoid, it will not.

3.3 NONLOCALITY AT PLANAR INTERFACES

As a concrete application of the HDM, we consider here systems with translational symmetry in the xy -plane. Specifically, we first re-examine the SPP dispersion of the metal-dielectric (MD) interface, previously considered within the LRA in [Section 2.2.3](#), and subsequently discuss the metal-dielectric-metal (MDM) and dielectric-metal-dielectric (DMD) double-interface problems, which support even and odd gap plasmons. Our focus in this section rests on the hydrodynamic impact on dispersion properties relative to the LRA, i.e. on the interrelation of plasmon energies $\hbar\omega$ and their momentum k_{\parallel} . Results discussed herein relate directly to [Publication G](#).

3.3.1 SPP dispersion and the single interface

The single interface is most transparently approached through the consideration of reflection coefficients. In particular, in complete analogy with the LRA result of [Eq. \(16\)](#), we consider the reflection coefficients of a plane wave incident from a dielectric onto a metal, with the latter described by the HDM. For concreteness, we label bulk-property variables by their adherence to the dielectric (d) or metallic (m) side, i.e. dielectric functions $\varepsilon^{\text{d,m}}$ and their associated total, parallel, and perpendicular momenta $(k^{\text{d,m}})^2 = \varepsilon^{\text{d,m}} k_0^2 = k_{\parallel}^2 + (k_{\perp}^{\text{d,m}})^2$. The reflection coefficients for TE and transverse magnetic (TM) polarization can be derived following the previously discussed LRA approach, extended only to account for the hydrodynamic Ohm's law of [Eq. \(41\)](#) and the associated ABC of [Eq. \(45\)](#) [or, equivalently, [\(46\)](#)], yielding [[179](#), [180](#)]

$$r_{\text{TE}}^{\text{HDM}} = \frac{k_{\perp}^{\text{d}} - k_{\perp}^{\text{m}}}{k_{\perp}^{\text{d}} + k_{\perp}^{\text{m}}}, \quad r_{\text{TM}}^{\text{HDM}} = \frac{\varepsilon^{\text{m}} k_{\perp}^{\text{d}} - \varepsilon^{\text{d}} k_{\perp}^{\text{m}} (1 + \delta^{\text{HDM}})}{\varepsilon^{\text{m}} k_{\perp}^{\text{d}} + \varepsilon^{\text{d}} k_{\perp}^{\text{m}} (1 + \delta^{\text{HDM}})}, \quad (51\text{a})$$

where the non-LRA aspects of the HDM is contained in the correction

$$\delta^{\text{HDM}} = \frac{k_{\parallel}^2}{k_{\perp}^{\text{HDM}} k_{\perp}^{\text{m}}} \frac{\varepsilon_{\text{B}}^{\text{m}} - \varepsilon^{\text{m}}}{\varepsilon_{\text{B}}^{\text{m}}}, \quad (51\text{b})$$

which in turn references the momentum of the hydrodynamic pressure waves via $(k_{\perp}^{\text{HDM}})^2 = (k^{\text{HDM}})^2 - k_{\parallel}^2$, with $(k^{\text{HDM}})^2 = k_{\beta}^2 \varepsilon^{\text{m}} / \varepsilon_{\text{B}}^{\text{m}}$ and $k_{\beta}^2 = \omega(\omega + i\gamma) / \beta^2$.

Evidently, the HDM does not modify the TE reflection coefficient. This is not surprising since the TE polarization-direction is parallel with the interface, which prohibits coupling with the longitudinal response of the HDM. Since the transverse response of the HDM mirrors that of the LRA no change is introduced, cf. [Eqs. \(44\)](#). Indeed, this is a rather general feature of the HDM: it does not perturb the response of TE waves (a statement relevant primarily in structures of symmetry, such as planes, cylinders, or spheres).

Conversely, the TM response is modified via the nonlocal correction δ^{HDM} . The correction is generally negative for Drude-like materials near the SPP frequency. Accordingly, it shifts the SPP frequency of the interface to higher values. To be explicit, the resonance condition for the TM SPP follows immediately from poles of $r_{\text{TM}}^{\text{HDM}}$ cf. [Eqs. \(7\)](#), allowing the dispersion equation [[155](#), [180](#)]

$$1 = -\frac{\varepsilon^{\text{d}} \kappa^{\text{m}}}{\varepsilon^{\text{m}} \kappa^{\text{d}}} (1 + \delta^{\text{HDM}}), \quad (\text{MD}) \quad (52)$$

with new momenta κ^j defined by $k_{\perp}^j \equiv i\kappa^j$ with $j \in \{\text{d, m, HDM}\}$, highlighting that the excitations are surface bound such that k_{\perp}^j is predominantly an imaginary quantity near resonance.

Allowing for both loss and retardation, the dispersion equation, [Eq. \(52\)](#), represents a radical equation in the complex frequency k_{\parallel} . We shall not repeat the details here

Correspondence with the notation of [Eq. \(16\)](#) is obtained by changing superscripts according to $\{\overset{+}{-}\} \leftrightarrow \{\overset{\text{d}}{\text{m}}\}$.

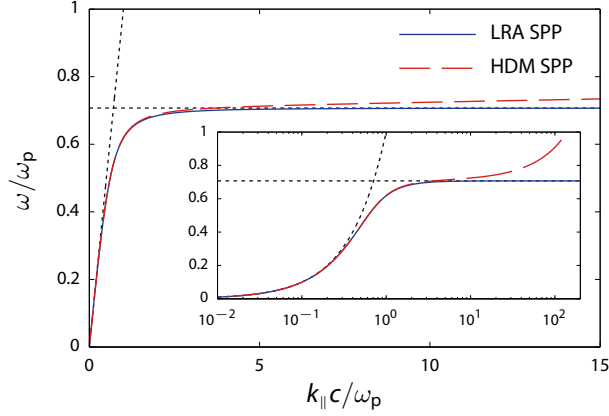


Figure 4: Dispersion of the SPP at the MD-interface between a pure, lossless Drude metal ($\epsilon_B^m = 1$ and $\gamma = 0$) and vacuum ($\epsilon^d = 1$), with frequencies and momenta normalized through ω_p . The pressure velocity is chosen as $\beta = 1.39 \times 10^6$ m/s, representative for silver ($\hbar\omega_p = 9.01$ eV) [60]. The inset depicts the same dispersion, but in a semilogarithmic view, highlighting the large-momentum discrepancy between the LRA and HDM.

An attractive side effect of the k_{\parallel} -dependent blueshift of the HDM is the introduction of a lower bound for group velocity of the lossless SPP (which vanishes at ω_{sp} in the LRA) – or, from a different perspective, the regularization of the otherwise divergent density of states.

but note that it can be converted to a third order polynomial equation in k_{\parallel}^2 by repeated squaring. Accordingly, the solutions, though complicated, have an analytical form. We explore the dispersion in Figure 4 in the absence of loss. In concord with our earlier considerations of the sign of δ^{HDM} , we observe that the HDM dispersion increasingly blueshifts relative to the LRA result as k_{\parallel} increases. That the HDM dispersion blue- rather than redshifts is not unexpected cf. our previous considerations of the Feibelman d -parameter: indeed, in the hard wall HDM d_{\perp} must exhibit a negative real part, since the induced charge cannot spill-out, thereby inducing a blueshift of the resonance cf. Eq. (32). Concretely, the hydrodynamic d -parameter can be derived from Eq. (30), yielding $d_{\perp}^{\text{HDM}} = i/k^{\text{HDM}}$ [54], which, in the simplifying $\gamma = 0$ and $\epsilon_B^m = 1$ case reduces to $d_{\perp}^{\text{HDM}} = -\beta/(\omega_p^2 - \omega^2)^{1/2}$, i.e. a purely negative quantity below the plasma-frequency. Moreover, the value of d_{\perp}^{HDM} indicates that the induced charge in the HDM is smeared over a length scale comparable to an inverse Thomas–Fermi wavenumber k_{TF} [since $k_{\text{TF}}^2 = 3\omega_p^2/v_F^2$], in sharp contrast to the singular distribution found in the LRA, see Eq. (29).

For non-negligible loss the HDM also exhibits the back-bending familiar from the LRA, previously discussed in Figure 2. However, in contrast to the LRA, back-bending is only exhibited beyond a certain loss-threshold, delimited roughly by $\gamma/\omega_p \lesssim \beta/c$ – a fact intuitively appreciable since the group-velocity of the HDM does not reach similarly low extremes as the LRA, and so is less loss-sensitive. The finer details of the transition between nonlocality- and loss-dominated regimes are interesting albeit rather esoteric, involving the merging and splitting of the SPP and Brewster mode, and will not be discussed further herein but is analyzed numerically in Publication C.

Besides the one- and two-interface structures considered here, the HDM has been applied in finite and infinite multilayers by several authors [164, 180–183], with implications particularly for metamaterial homogenization constructs.

3.3.2 Thin-film waveguides

In closing our treatment of planar structures, we briefly consider the MDM and DMD configurations of channel or slab width w . The dispersion equation of the MDM is particularly straightforward to obtain, following directly from requiring in-channel

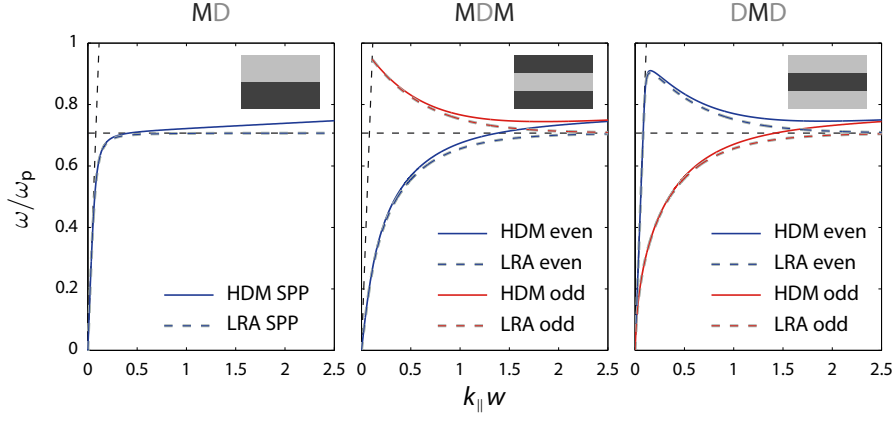


Figure 5: Dispersion of the interface-, channel-, and slab-plasmons corresponding to the MD, MDM, and DMD configurations, respectively (see insets and titles), with the latter two calculated at $w = 5$ nm. Material parameters are as in Figure 4. Light line and single-interface SP frequency indicated in all cases by dotted black lines.

round-trip phase-accumulation $\exp(ik_{\perp}^d w) r_{\text{TM}}^{\text{HDM}} = \pm 1$, with associated even (+) and odd (−) channel (or gap-)plasmons (referenced relative to \mathbf{E}), yielding [155, 180]

$$\tanh\left(\frac{\kappa^d w}{2}\right) = -\frac{\varepsilon^d \kappa^m}{\varepsilon^m \kappa^d} (1 + \delta^{\text{HDM}}), \quad (\text{even, MDM}) \quad (53a)$$

$$\coth\left(\frac{\kappa^d w}{2}\right) = -\frac{\varepsilon^d \kappa^m}{\varepsilon^m \kappa^d} (1 - \delta^{\text{HDM}}). \quad (\text{odd, MDM}) \quad (53b)$$

In comparison, the DMD requires more labor since both longitudinal and transverse must be included in the solution-approach, the result of which is (with symmetry considerations applied to the \mathbf{H} -field) [155]

$$\coth\left(\frac{\kappa^m w}{2}\right) = -\frac{\varepsilon^m \kappa^d}{\varepsilon^d \kappa^m} - \delta^{\text{HDM}} \coth\left(\frac{\kappa^{\text{HDM}} w}{2}\right), \quad (\text{odd, DMD}) \quad (54a)$$

$$\tanh\left(\frac{\kappa^m w}{2}\right) = -\frac{\varepsilon^m \kappa^d}{\varepsilon^d \kappa^m} - \delta^{\text{HDM}} \tanh\left(\frac{\kappa^{\text{HDM}} w}{2}\right). \quad (\text{even, DMD}) \quad (54b)$$

Naturally, in the large-width limit, the dispersion equations Eqs. (53) and (54) reduce to the MD result of Eq. (52) since $\tanh x$ and $\coth x$ tend to unity for $x \rightarrow \infty$. The opposite limit of vanishing w , perhaps unsurprisingly, is less well-behaved. For example, within the LRA the fundamental mode of the MDM, Eq. (53a), is singular, diverging with decreasing w as $k_{\parallel} \propto 1/w$ in the nonretarded limit. This behavior is regularized in the MDM by a cutoff $\Delta_{\text{MDM}} \equiv 2i\varepsilon^d(\varepsilon^m - \varepsilon_B^m)/\varepsilon^m \varepsilon_B^m k^{\text{HDM}}$ such that $k_{\parallel} \propto 1/(w + \Delta_{\text{MDM}})$. Unfortunately, a similar regularization does not arise for the DMD scenario, whose mending likely requires inclusion of spill-out.

The dispersion of MD, MDM, and DMD is compared in Figure 5. Evidently, the behavior of the fundamental (i.e. energetically lowest) mode is similar across all three configurations. In fact, even the higher-energy mode of the MDM and DMD are similar at larger k_{\parallel} . In fact, this similarity is perfect in the nonretarded LRA limit, and is known as Babinet's principle of complementary structures. Evidently, this complementary symmetry is a corollary of Apell's nonretarded void-nonvoid sum-rule previously discussed in Eq. (20). Naturally, the symmetry is imperfect when retardation is accounted for (since $k_{\perp}^d \neq k_{\perp}^m$ when $k_0/k_{\parallel} \rightarrow 0$). However, as evident from the nonretarded limits of Eqs. (53) and (54) (and as shown explicitly in Publication C by numerical means), the symmetry is also broken by the HDM, even in the nonretarded limit.

CBPs were first observed experimentally in thin-films in 1971: in Ag by Lindau and Nilsson [184] and in K by Anderegg et al. [185] – and more recently by Özer et al. [186] in Mg films, down to three atomic monolayers. In all cases, the HDM predictions agree remarkably well with experiment, e.g. quantified by an inter-CBP resonance-spacing $\propto \beta^2/w^2$.

In either case the breaking of complementary symmetry is due to the introduction of an additional length scale: $1/k_0$ and $1/k_\beta$ for retardation and the HDM, respectively.

An effect which has not been discussed here is that of confined bulk plasmons (CBPs), corresponding, essentially, to standing-wave oscillations of longitudinal polarization in a finite metallic region, such as in the DMD structure. CBPs provide some of the so-far strongest experimental support for the core physics of the HDM, largely uncluttered by competing mechanics, and is arguably the HDM's most unique hallmark. We return to this subject in the following section, in our treatment of nanospheres, and provide also an approximate relation for the spectral positions of CBPs in general structures.

3.4 NONLOCALITY IN NANOSPHERES

The prominent position held by LSPs in plasmonics is, as previously noted in [Section 2.2.3](#), largely due to their direct and sizable coupling with plane wave excitations. This coupling can be investigated by ensemble-measurements in homogeneous distributions of nanostructures or nanoparticles, e.g. in dilute suspensions of either gaseous, aqueous, or solid kind [187, 188], or in planar arrays [76]. Correlating such measurements with single particle properties necessitates very dilute and very homogeneous distributions to avoid excessive inhomogeneous broadening due to size-, material-, and environment-related ensemble-fluctuations. Generally this represents a challenge, though some far-field techniques – such as dark-field microspectroscopy [189–191] – can investigate LSPs on a single-particle basis. This deficiency of far-field spectroscopy, though not unsurmountable, is compounded by a weak interaction between plane waves and higher-order (HO) multipole resonances in truly subwavelength structures, which effectively limits investigations to the dipolar sort. On the other hand, near-field techniques – such as SNOM [102], fluorescence microscopy [192, 193], or EELS [22, 116, 117] – inherently probe single-particle LSP properties. Moreover, they may sample beyond-dipolar properties and provide additional control through a variable particle-probe distance. These techniques have flourished over several decades, with metal-plasmonics applications ranging from studies of general nanostructures [194–196] to simple spherical nanoparticles [25, 26, 172, 197, 198]. The latter type has attracted much attention in the small-radius limit [25, 26, 171–174], owing to measurements of non-classical size-dispersion with decreasing radius in both ensemble and single-particle studies.

In this section we investigate the HDM in spherical nanostructures. Our motivation for this is partly the noted measurements of blueshifted resonances and partly the desire to establish in simple terms the impact of nonlocality in finite nanostructures. In this examination of nanospheres, we will simultaneously examine three distinct quantities that pertain directly to measurements in the far- and near-field. Specifically, we contrast the response of one far-field quantity, the extinction cross-section σ_{ext} and two near-field quantities, namely electron energy loss spectroscopy Γ (EELS) and local density of states ρ^{E} (LDOS). The three examples correspond to illumination by plane-, cylinder-, and spherical-like waves, respectively. The concurrent comparison of each of these measurements will allow us to highlight the characteristics of each, and, importantly, to underscore that near-field measurements of nanostructures are much richer than a dipole-only interpretation would suggest. The material discussed in this section derives from [Publication F](#).

3.4.1 Mie–Lorenz theory

The appropriate framework for the study of scattering by spherical structures is that of Mie–Lorenz theory. We will not review the mathematical foundation of the approach – mainly because it is available in great detail in textbooks on electromagnetism [47, 202, 203] and light-scattering in particular [98, 101, 188] – but will instead focus on introducing the key concepts necessary for the appreciation of the culminating result: the Mie–Lorenz coefficients t_l^{TE} and t_l^{TM} .

At the core of Mie–Lorenz theory are vector wave functions. In particular, three classes of vector wave function exist: two solenoidal (transverse) $\mathbf{M}_\nu(\mathbf{r})$ and $\mathbf{N}_\nu(\mathbf{r})$ describing, respectively, TE and TM polarized waves, and one irrotational (longitudinal) $\mathbf{L}_\nu(\mathbf{r})$. Together, they form a complete orthogonal basis for the electric field \mathbf{E} , allowing expansion in the composite mode-index ν

$$\mathbf{E}(\mathbf{r}) = \sum_\nu a_\nu \mathbf{M}_\nu(\mathbf{r}) + b_\nu \mathbf{N}_\nu(\mathbf{r}) + c_\nu \mathbf{L}_\nu(\mathbf{r}), \quad (55)$$

with associated expansion coefficients a_ν , b_ν , and c_ν for the TE, TM, and longitudinal polarization, respectively. For systems described in spherical coordinates $\mathbf{r} = [r, \theta, \varphi]^T$ the composite index ν separates into angular momentum components $l \in [1, \infty[$ and $m \in [-l, l]$. The spherical vector wave functions themselves are defined through a (frequency-dependent) momentum k and a scalar generating function $\psi_{lm}(\mathbf{r})$, which is a solution of the Helmholtz equation $\nabla^2 \psi_\nu(\mathbf{r}) + k^2 \psi_\nu(\mathbf{r}) = 0$, such that

$$\mathbf{M}_{lm}(\mathbf{r}) = \nabla \times \mathbf{r} \psi_{lm}(\mathbf{r}), \quad (56a)$$

$$\mathbf{N}_{lm}(\mathbf{r}) = k^{-1} \nabla \times \nabla \times \mathbf{r} \psi_{lm}(\mathbf{r}), \quad (56b)$$

$$\mathbf{L}_{lm}(\mathbf{r}) = k^{-1} \nabla \psi_{lm}(\mathbf{r}). \quad (56c)$$

The choice of momentum k is dictated by the underlying medium and wave-type. When considering a spherical system of dielectric exterior and a metallic interior (characterized by ε^d and $\{\varepsilon^m, \beta\}$, respectively), the previously defined momenta k^d , k^m , and k^{HDM} constitute the relevant selection. Adherence of momentum to a particular vector wave will be indicated by corresponding superscripts ‘d’, ‘m’ and ‘HDM’. Lastly, the generating function $\psi_{lm}(r, \theta, \varphi) = z_l(kr) P_l^m(\cos \theta) e^{im\varphi}$ (with P_l^m denoting the associated Legendre polynomials) specifies the in- or outgoing character of the vector waves through z_l , being either a spherical Bessel function j_l (ingoing) or a spherical Hankel function of the first kind $h_l^{(1)}$ (outgoing). The choice of either will be indicated by superscripts ‘in’ and ‘out’, respectively.

With these matters established, it is clear that we may expand an ingoing, exciting field (impinging from the dielectric region) \mathbf{E}^{ex} and its outgoing, scattered field \mathbf{E}^{sc} by

$$\mathbf{E}^{\text{ex}}(\mathbf{r}) = \sum_{lm} a_{lm}^{\text{ex}} \mathbf{M}_{lm}^{\text{d},\text{in}}(\mathbf{r}) + b_{lm}^{\text{ex}} \mathbf{N}_{lm}^{\text{d},\text{in}}(\mathbf{r}), \quad (57a)$$

$$\mathbf{E}^{\text{sc}}(\mathbf{r}) = \sum_{lm} a_{lm}^{\text{sc}} \mathbf{M}_{lm}^{\text{d},\text{out}}(\mathbf{r}) + b_{lm}^{\text{sc}} \mathbf{N}_{lm}^{\text{d},\text{out}}(\mathbf{r}), \quad (57b)$$

for $r > R$, with R denoting the particle radius. Note the absence of the vector waves \mathbf{L}_{lm} , which are only relevant in regions that support longitudinal waves, such as the metallic region. Accordingly, for the interior region $r < R$, an ingoing, transmitted field \mathbf{E}^{tr} exists, with an expansion of the form

$$\mathbf{E}^{\text{tr}}(\mathbf{r}) = \sum_{lm} a_{lm}^{\text{tr}} \mathbf{M}_{lm}^{\text{m},\text{in}}(\mathbf{r}) + b_{lm}^{\text{tr}} \mathbf{N}_{lm}^{\text{m},\text{in}}(\mathbf{r}) + c_{lm}^{\text{tr}} \mathbf{L}_{lm}^{\text{HDM},\text{in}}(\mathbf{r}). \quad (57c)$$

With the proper expansions established, the remaining task involves only the determination of the interrelation of the expansion coefficients. Such interrelations can be

It is curious to note that Mie–Lorenz theory commonly is attributed only to Mie [3] – in fact, Lorenz [199] made an equivalent contribution 18 years earlier. Also of note is a solution by Debye [200], published one year after Mie’s. For a historical account, see Ref. [201].

The basis constituted by the wave vector functions in their spherical form is usually known as the multipole basis.

discerned by application of BCs and use of modal orthogonality. Of particular interest is the relation between exciting and scattered coefficients, reading

$$a_{lm}^{\text{sc}} = t_l^{\text{TE}} a_{l'm'}^{\text{ex}} \delta_{ll'} \delta_{mm'}, \quad b_{lm}^{\text{sc}} = t_l^{\text{TM}} b_{l'm'}^{\text{ex}} \delta_{ll'} \delta_{mm'}, \quad (58)$$

where the Kronecker deltas $\delta_{ll'}$ and $\delta_{mm'}$ express the joint spherical symmetry of system and vector waves. The proportionality constants $t_l^{\text{TE,TM}}$ define the Mie–Lorenz coefficients, which in the HDM takes the following form, first derived by Ruppin [204] and more recently resurfaced by David and García de Abajo [205]

$$t_l^{\text{TE}} = \frac{-j_l(x^{\text{m}})[x^{\text{d}} j_l(x^{\text{d}})]' + j_l(x^{\text{d}})[x^{\text{m}} j_l(x^{\text{m}})]'}{j_l(x^{\text{m}})[x^{\text{d}} h_l^{(1)}(x^{\text{d}})]' - h_l^{(1)}(x^{\text{d}})[x^{\text{m}} j_l(x^{\text{m}})]'}, \quad (59a)$$

$$t_l^{\text{TM}} = \frac{-\epsilon^{\text{m}} j_l(x^{\text{m}})[x^{\text{d}} j_l(x^{\text{d}})]' + \epsilon^{\text{d}} j_l(x^{\text{d}})[x^{\text{m}} j_l(x^{\text{m}})]'(1 + \delta_l^{\text{HDM}})}{\epsilon^{\text{m}} j_l(x^{\text{m}})[x^{\text{d}} h_l^{(1)}(x^{\text{d}})]' - \epsilon^{\text{d}} h_l^{(1)}(x^{\text{d}})[x^{\text{m}} j_l(x^{\text{m}})]'(1 + \delta_l^{\text{HDM}})}, \quad (59b)$$

with a hydrodynamic correction

$$\delta_l^{\text{HDM}} \equiv \frac{(l+1)j_l(x^{\text{m}})}{[x^{\text{m}} j_l(x^{\text{m}})]'} \frac{l j_l(x^{\text{HDM}})}{x^{\text{HDM}} j_l'(x^{\text{HDM}})} \frac{\epsilon^{\text{m}} - \epsilon_{\text{B}}^{\text{m}}}{\epsilon_{\text{B}}^{\text{m}}}, \quad (59c)$$

and with normalized momenta $x^{\text{d,m,HDM}} = k^{\text{d,m,HDM}} R$, and where primes indicate derivatives with respect to x . The multipole-correction in Eq. (59c) evidently bears significant resemblance to its planar sibling in Eq. (51b), both in overall structure and in its modification of the LRA result, albeit with modifications due to the curvature of the interface. In further correspondence with the planar case and with our general expectations, the hydrodynamic correction does not feature in the TE coefficients and the TM coefficients tend to their LRA-counterparts as $\beta \rightarrow 0$ where $\delta_l^{\text{HDM}} \rightarrow 0$.

The associated nonretarded limit, as expressed by the multipolar HDM polarizability α_l^{HDM} , can naturally be derived from the retarded framework of Mie–Lorenz theory. The transition from the retarded $t_l^{\text{TE,TM}}$ to the nonretarded α_l^{HDM} involves only the TM coefficients, and is formally obtained by (!! denotes the double factorial)

$$\begin{aligned} \alpha_l^{\text{HDM}} &= -4\pi i \frac{l[(2l+1)!!]^2}{(l+1)(2l+1)} \lim_{k_0 R \rightarrow 0} \left[\frac{t_l^{\text{TM}}}{(k^{\text{d}})^{2l+1}} \right] \\ &= 4\pi R^{2l+1} \frac{l\epsilon^{\text{m}} - l\epsilon^{\text{d}}(1 + \tilde{\delta}_l^{\text{HDM}})}{l\epsilon^{\text{m}} + (l+1)\epsilon^{\text{d}}(1 + \tilde{\delta}_l^{\text{HDM}})}, \end{aligned} \quad (60)$$

with the nonretarded HDM correction $\tilde{\delta}_l^{\text{HDM}} = \frac{l j_l(x^{\text{HDM}})}{x^{\text{HDM}} j_l'(x^{\text{HDM}})} \frac{\epsilon^{\text{m}} - \epsilon_{\text{B}}^{\text{m}}}{\epsilon_{\text{B}}^{\text{m}}}$.

3.4.2 Multipole plasmons

The poles of α_l^{HDM} evidently provide us with the hydrodynamic generalization of the nonretarded multipole plasmon condition of Eq. (19), while the poles of t_l^{TM} provide the retarded equivalent. The main difference between the retarded and nonretarded plasmon conditions is that the latter yields radius-independent resonances in the LRA, while the former redshifts (relative to the constant nonretarded energies) for increasing radius. The nonretarded LRA multipole plasmons are governed by Eq. (19), which establishes an explicit expression for the LRA resonance frequencies ω_l^{LRA} in the low-loss limit

$$\omega_l^{\text{LRA}} = \frac{\omega_{\text{p}}}{\sqrt{\epsilon_{\text{B}}^{\text{m}} + \frac{l+1}{l} \epsilon^{\text{d}}}}. \quad (61)$$

The HDM-equivalent of Eq. (61), following from the poles of α_l^{HDM} , is less palatable in the sense that it is not directly invertible in the resonance frequency ω_l^{HDM} . An approximate relation can be discerned, however, by making a pole-approximation around the

The small-argument limiting forms of the spherical Bessel and Hankel functions are applied here, reading:
 $j_l(x) \approx \frac{x^l}{(2l+1)!!}$ and
 $h_l^{(1)}(x) \approx \frac{x^l}{(2l+1)!!} - \frac{i(2l-1)!!}{x^{l+1}}$
for $x \ll \sqrt{l+1}$.

For dispersive dielectric surroundings or bound polarization (i.e. for ϵ^{d} or $\epsilon_{\text{B}}^{\text{m}}$ frequency-dependent), Eq. (61) constitutes an implicit equation in ω_l^{LRA} .

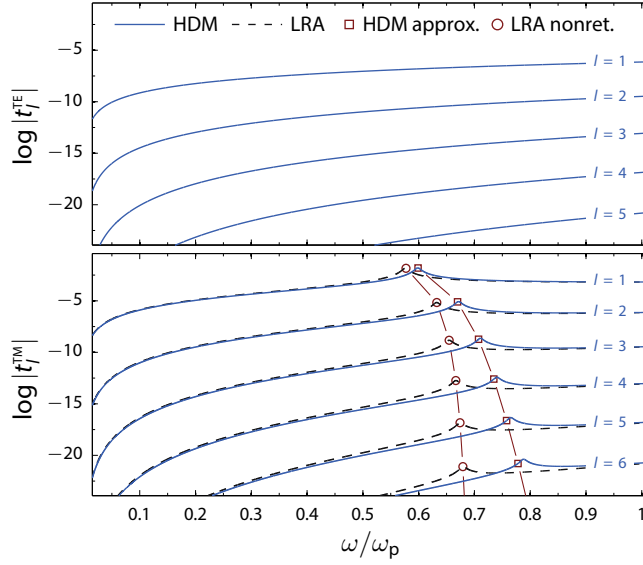


Figure 6: Absolute value of the HDM TE and TM Mie-Lorenz coefficients, t_l^{TE} and t_l^{TM} , as a function of frequency. A free-electron ($\epsilon_B^m = 1$) nanosphere of radius $R = 2.5$ nm, plasma frequency $\hbar\omega_p = 10$ eV, and loss-rate $\hbar\gamma = 0.1$ eV is considered, assumed embedded in vacuum ($\epsilon^d = 1$). The LRA TM Mie-Lorenz coefficient is indicated in dashed gray for comparison. The approximate relations of Eqs. (61) and (62) are indicated by circular and square markers, respectively.

LRA frequencies, i.e. by evaluating δ_l^{HDM} at the LRA frequency ω_l^{LRA} , and expanding around $k_\beta R \gg 1$, allowing [63, 206]

$$\omega_l^{\text{HDM}} \simeq \omega_l^{\text{LRA}} + \frac{\beta}{R} \sqrt{\frac{l(l+1)\epsilon^d}{4\epsilon_B^m}}. \quad (62)$$

In Fig. 6 we examine the frequency dependence of the TE and TM coefficients for a nanoscopic sphere, $R = 2.5$ nm, where the nonretarded considerations just discussed apply. As expected, the TE polarization is void of resonances, while the multipole plasmons of TM kind are well-described by the analytical expressions of Eqs. (61) and (62). A rather drastic difference between the nonretarded large- l limits of the LRA and HDM is also observed, namely that the LRA resonances tend toward the planar interface SP resonance $\omega_{l \rightarrow \infty}^{\text{LRA}} = \omega_{\text{sp}} = \omega_p/\sqrt{2}$ (for a free-electron gas in vacuum), while the HDM resonances extend beyond this limit without bound as evident from Eq. (62). This difference marks a regularizing feature of the HDM since it removes the pile-up of LRA multipoles near $\omega_p/\sqrt{2}$ – which itself is responsible for extreme, even divergent LDOS in the near-surface vicinity. This regularization has significant implications on LRA predictions of extraordinary emitter-dynamics in the near-surface region [207, 208], since it provides a cutoff for the LDOS, thereby limiting coupling efficiency between emitter and nanostructure. It is worth noting that this feature does not hinge upon the relative smallness of the structure: the extension of the multipoles beyond the $\omega_p/\sqrt{2}$ limit is guaranteed, since the strength of the hydrodynamic correction relates not only with $1/R$ but also with multipole order l through an effective momentum $\sim l/R$. Accordingly, we observe that the HDM modifies the LRA conclusions not only in nanoscale structures, but also in measurements with nanoscale separations between probe and surface. We further quantify and expand on this observation in the following section.

The GNOR model naturally also applies to Eq. (62) following a substitution $\beta \rightarrow \eta$. Interestingly, this highlights the l -dependent generalization of Kreibig's dipolar considerations: within GNOR the size-dependent broadening $\propto A v_F/R$ exhibits an approximate $A \propto \sqrt{l(l+1)}$ scaling with multipole order. This additional consideration only exacerbates the reduction of LDOS vis-à-vis HDM and LRA predictions.

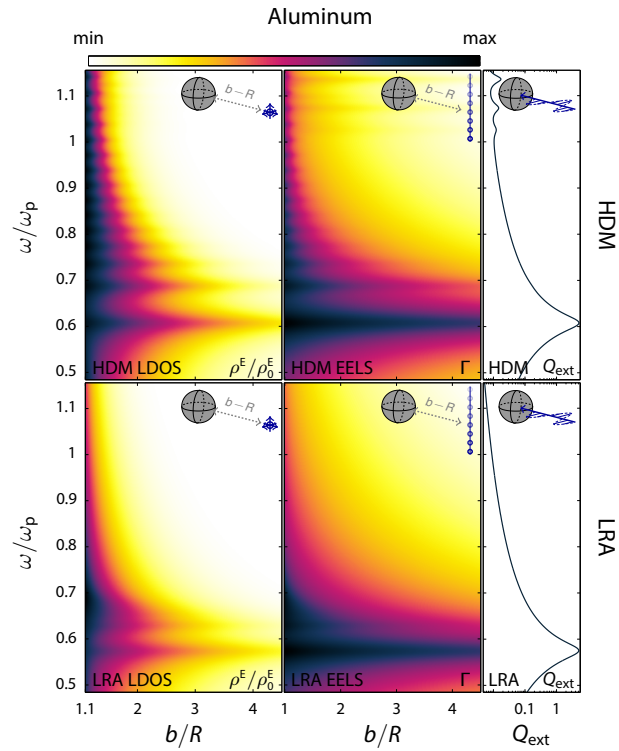
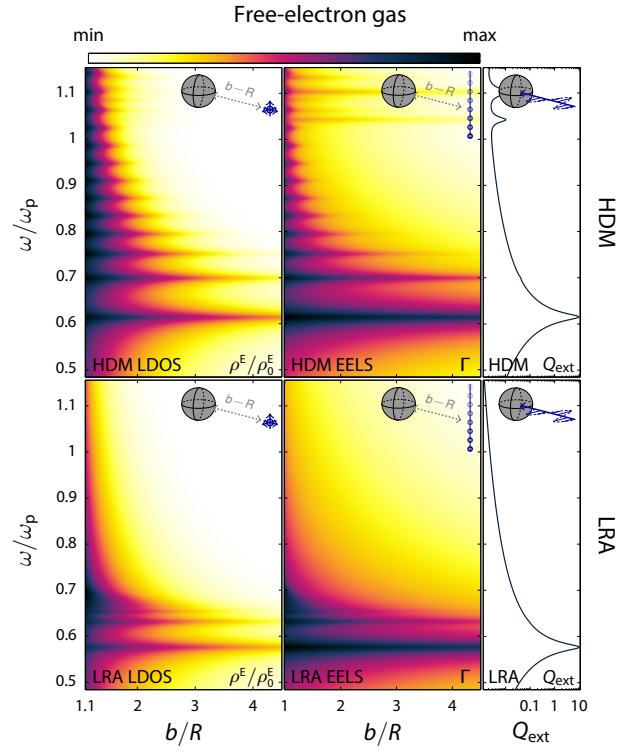


Figure 7: LDOS enhancement ρ^E/ρ_0^E , EEL probability Γ , and extinction efficiency Q_{ext} calculated within the HDM and LRA, as indicated in labels, for a nanosphere of radius $R = 1.5$ nm in vacuum, with distinct metal properties in (a) and (b). Colormaps represent the relative intensity of ρ^E/ρ_0^E and Γ in a logarithmic scale.

3.4.3 Role of the measurement probe

The salient strength of the Mie–Lorenz coefficients is perhaps most clearly displayed in considerations of measurements. Specifically, since $t_l^{\text{TE, TM}}$ specify the induced response outside the sphere, any linear measurement of an induced quantity, say \mathcal{O} , can be expressed as a linear superposition of $t_l^{\text{TE, TM}}$ and coefficients $\mathcal{O}_{lm}^{\text{TE, TM}}$ which relate with the measurement type and excitation field. More precisely, for the quantities considered here, the expansion takes the form

$$\mathcal{O} = \sum_{lm} \mathcal{O}_{lm}^{\text{TE}} \text{Re } t_l^{\text{TE}} + \mathcal{O}_{lm}^{\text{TM}} \text{Re } t_l^{\text{TM}}. \quad (63)$$

We will here examine three distinct quantities (whose significance was previously introduced in [Section 2.3](#)):

1. The extinction cross-section $\sigma_{\text{ext}}(\omega)$ (depicted by its area-normalized equivalent, the extinction efficiency $Q_{\text{ext}} = \sigma_{\text{ext}}/\pi R^2$). [98]
2. The EEL probability $\Gamma(\omega)$ of an electron with velocity $v\hat{\mathbf{z}}$ impacting the xy -plane at a distance $b > R$ from origo (where the sphere is centered). We take as electron velocity $v \approx 0.695c$, corresponding to an electron energy of 200 keV. [22, 209]
3. The electric LDOS $\rho^{\text{E}}(\omega)$ measured at a distance $b > R$ from the sphere center (depicted by the orientation-averaged LDOS-enhancement $\rho^{\text{E}}/\rho_0^{\text{E}}$ relative to the vacuum value ρ_0^{E}). [210–213]

The expansion of these quantities in Mie–Lorenz coefficients were obtained in earlier LRA studies and are also reported in the Supporting Information of [Publication F](#). As such, we will not repeat them here, but have highlighted pertinent references in the margin above. Importantly, the LRA-expansions transfer fully to the HDM when considering measurements made outside the metallic region – indeed, the expansions apply to arbitrary electromagnetic problems of spherical symmetry (such as e.g. multilayer shells), provided the Mie–Lorenz coefficients are suitably modified.

A full comparison of extinction, EELS, and LDOS within the HDM and LRA is presented in [Figure 7](#), with simultaneous consideration of frequency and separation dependency. We discuss its interpretation below.

MULTIPOLE PEAKS Let us examine initially the main features by below the plasma-frequency. In that frequency-range, the extinction cross-section exhibits a single peak associated with the dipole LSP – the primary difference between LRA and HDM is a blueshift. This picture is unchanged for the EEL probability and LDOS for large probe-surface separations $b/R \gg 1$. As the normalized distance b/R is reduced, however, the spectra develop significant differences: the increasingly inhomogeneous nature of the excitation field, as experienced by the nanosphere, facilitates the excitation of higher multipoles. Accordingly, for very small b/R the LRA near-field measurements exhibit large peaks near the planar surface plasmon resonance ω_{sp} . Conversely, for the HDM the pile-up of multipoles is absent, with each exhibiting a minimum separation $\sim \beta/R$ cf. [Eq. \(62\)](#), and their influence extend beyond ω_{sp} . On this note, we highlight that the experimental observation of multipoles beyond the $l \rightarrow \infty$ limit at ω_{sp} has been reported by vom Felde *et al.* [214] in EELS performed on ensembles of potassium clusters (radius 1 – 2 nm) embedded in magnesium oxide. There, the blueshift into the classically quiet region was tentatively attributed to quantum size effects, through qualitative comparison with TDDFT-calculations in few-atom (<100) jellium clusters. The present hydrodynamic analysis demonstrates that the blueshift might just as well be assigned a nonlocal origin.

A natural question is whether the HDM-separation between multipoles exceed practically attainable loss-rates, rendering then their individual signatures experimentally accessible. Considering e.g. aluminum, see [Figure 7b](#), which has rather low loss

The formal correspondence between σ_{ext} and $\rho^{\text{E}}/\rho_0^{\text{E}}$ in the $b/R \gg 1$ limit can be discerned by comparison of the associated limits of $\mathcal{O}_{lm}^{\text{TE, TM}}$. The expansion of Γ similarly achieves formal equivalence with σ_{ext} in the $v/c \rightarrow 1$ limit. Details available in the Supporting Information of [Publication F](#).

and only modest bound polarization, this would appear to be feasible; the peaks remain well-separated and distinguishable. Even so, several factors conspire to modify this conclusion appreciably, effectively rendering observation of distinct HO multipoles either extremely challenging or outright impossible. A first basic modifier pertains to the choice of aluminum. In fact, the dipole LSP plasmon of aluminum is not well-described by HDM, being known to experimentally redshift due to appreciable plasmon spill-out [175]. Even so, we expect that nonlocality will eventually come to dominate the effects of spill-out, such that the charge-centroid of the HO multipoles gradually recedes inwards as l increases – a consideration which is supported by TDDFT-calculations of d_{\perp} at large momentum [178] – rendering the HDM-predictions of increasing multipole separation and eventual protraction above the $l \rightarrow \infty$ limit qualitatively applicable. A more severe modifier arises due to Kreibig broadening and its generalization as described by the GNOR model. Specifically, since not only the resonance shift but also the size-dependent broadening scale with $\sim l\beta/R$ the HO multipoles are eventually damped beyond observability. These considerations are quantitatively demonstrated in Figure 8, depicting GNOR calculations of the free-electron gas. It is worth noting that the GNOR and LRA results differ significantly, even in the presence of additional size-dependent broadening, exemplified e.g. by a large broadband response extending above ω_{sp} for $b/R \sim 1$. The difference can readily be appreciated by extending our previous HDM considerations to include size-dependent broadening, which simply smear out the peaks observed in the HDM.

In summary, these considerations suggest an additional experimental earmark of nonclassical nanoplasmonic features, beyond those due to a shifting dipole resonance: namely scrutiny of response-behavior at short probe-to-surface distances, with particular attention to single or agglomerate peaks extending beyond the LRA $l \rightarrow \infty$ limit at ω_{sp} . Moreover, our comparisons highlight the necessity of accounting for not just

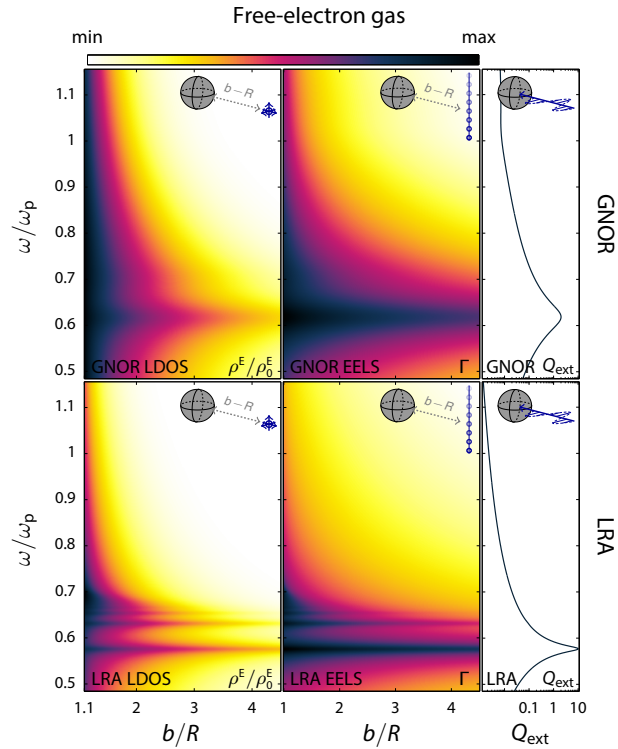


Figure 8: GNOR calculations for the setup considered previously in Figure 7a. The GNOR diffusion parameter D is here chosen by dimensional analysis as v_F^2/ω_p .

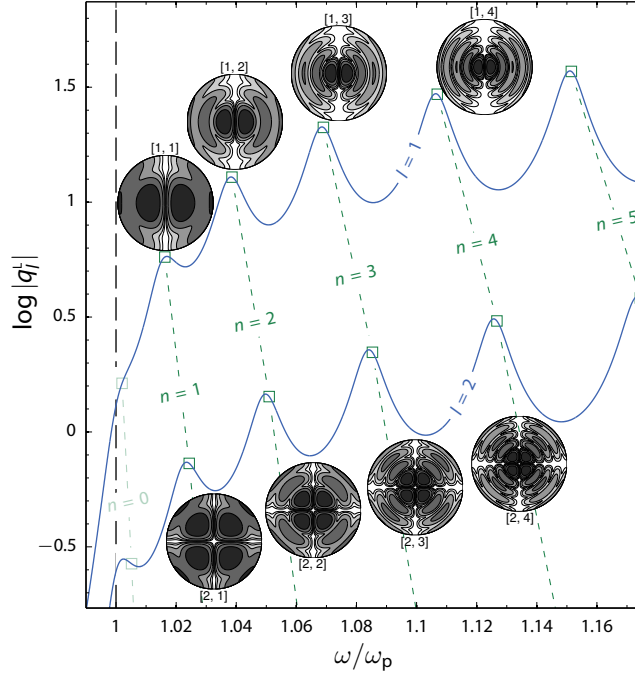


Figure 9: Absolute value of the Mie-Lorenz transmission coefficients q_l^T as a function of frequency. The coefficients give the coupling amplitude between transmitted longitudinal and incident TM multipoles via $c_{lm}^{\text{tr}} = q_l^T b_{l'm'}^{\text{ex}} \delta_{ll'} \delta_{mm'}$, see Eqs. (57). Setup parameters are identical to those in Figure 6. Shown are the dipolar and quadrupolar coefficients q_1^T and q_2^T in blue. The approximate resonance frequencies of Eq. (64) are given in green markers. The absence of the $n = 0$ CBP is apparent. Insets depict logarithmic contour plots (in an xz -plane) of the absolute value of the induced charge density of the CBPs with $[l, n]$ indices labeled (contours separated by factors of 2).

the dipole but also HO multipoles in EELS and LDOS scenarios with $b/R \sim 1$. This latter point is particularly pertinent to recent EELS studies in nanospheres [25, 26]. In those studies the electron beam was positioned exactly at the edge of the nanosphere, i.e. with $b/R \approx 1$. This choice is made for practical reasons, with an aim to cultivate large LSP losses relative to the zero-energy loss peak. Even so, measurements were compared only with the dipolar mode, though HO modes are appreciably excited under those circumstances as revealed by our preceding discussion. A danger then, is the misidentification of several merged HO modes as a *single* dipole mode – such a misidentification would overestimate possible resonance shifts and widths.

CONFINED BULK PLASMON PEAKS Finally, we briefly discuss the peaks observed above ω_p within the HDM in Figure 7. These peaks correspond to the excitation of CBPs. A detailed account of their properties, dipolar as well as multipolar, is given in Publication F and we shall not repeat the full analysis here. Rather, we will focus on a simple, approximate result for their dispersion, derivable from the HDM by treating the CBPs as strictly longitudinal oscillations, thereby neglecting their transverse components which are only minor. Under this assumption, the v th CBP resonance ω_v^{CBP} of a general (hydrodynamically treated) finite geometry Ω is given by

$$\omega_v^{\text{CBP}} (\omega_v^{\text{CBP}} + i\gamma) = \frac{\omega_p^2}{\epsilon_B^{\text{m}}} + u_v^2 \frac{\beta^2}{L^2}, \quad (64)$$

where L denotes a characteristic length of the structure (e.g. the radius R in a sphere), and with eigenvalues u_v of the Helmholtz equation $(\nabla_{\mathbf{x}}^2 + u_v^2)\psi_v(\mathbf{x}) = 0$ in L -scaled

The emergence of the – to electromagnetism otherwise uncommon – Neumann BC, emanates from the hydrodynamic ABC of vanishing normal current, see Eq. (45).

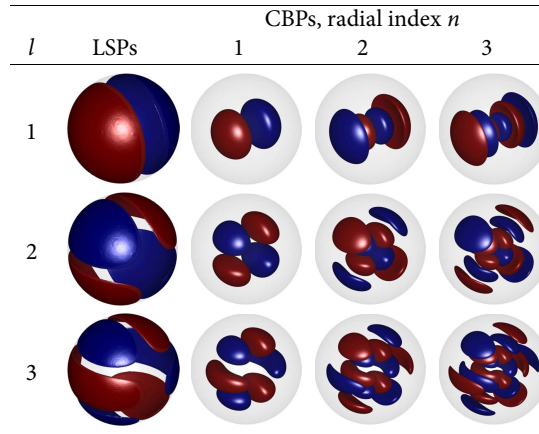


Figure 10: Isosurfaces of the induced density (real part) for LSPs and CBPs calculated in the HDM for a free-electron gas nanosphere (outline indicated in gray). Isovalues are drawn at plus/minus (red/blue) twice the mean of the absolute value of the induced charge density.

coordinates $\mathbf{x} = \mathbf{r}/L$ with Neumann BCs applied at the boundary $\mathbf{x} \in \partial\Omega/L$. For the spherical system, the composite index ν separates into multipole and radial indices, l and n , respectively, with the noted eigenvalues corresponding to n th root of the derivative of the l th-order spherical Bessel function, i.e. $j'_l(u_{ln}) = 0$ (the first roots, $n = 0$, however, are spurious and nonphysical; their emergence arise due to the neglect of transverse field components).

The quality of this approximation can be established by comparison with full calculations of the Mie–Lorenz transmission coefficient q_l^t into the longitudinal L_{lm} vector waves: this is shown in Figure 9 along with their associated induced charge distributions. The confined standing wave nature of the CBPs are clearly expressed. For the sake of further illustration we contrast the nature of the induced charge of the LSPs and CBP modes in Figure 10. Though the induced density of the HDM LSPs is not δ -like – as in the LRA – it is still strongly surface bound, in stark contrast to that of the CBPs.

The planar DMD system of Section 3.3 similarly supports CBPs: Eq. (64) applies there with parameters $L = w$ and $u_n = n\pi$.

3.5 OBSERVATION OF MULTIPOLE PLASMONS IN EMBEDDED SILVER NANOPARTICLES

In this section we depart from the main focus of this chapter, i.e. nonclassical features in plasmonics, to discuss an experimental EELS investigation of HO multipoles in embedded silver nanospheres. The investigation naturally ties in with the concepts discussed in the previous section, and so, we suggest, justifies this departure. The results are presented in full in Publication B [215] – we here discuss a subset of these results, particularly those that pertain to our preceding theoretical deliberations. The study is a result of a collaborative project with experimentalists; the author has not contributed directly in experimental aspects, rather contributions were of computational, theoretical, and interpretational kind.

A central issue for the practical observation of multipoles is interband dispersion. Specifically, in good plasmonic materials such as silver or gold, strong interband dispersion (and concomitant large deviations from ideal Drude behavior) at larger frequencies generally dampens HO multipoles, rendering their observation challenging. To circumvent this issue, particles can be encapsulated in an embedding material with

a sufficiently high dielectric constant, affording a general redshift of all resonances cf. Eq. (61), to push the HO multipoles out of the interband region. Following this reasoning, we embedded silver nanoparticles (in the radius range 1 – 20 nm) in a silicon nitride coating, see Figure 11. A further advantage of encapsulation is the prevention of silver oxide and sulfide formation at the nanoparticle surface. The solitary drawback of encapsulation is a reduced inter-multipole spectral spacing $\Delta\omega_l = \omega_{l+1} - \omega_l$, which nevertheless still allows identification of agglomerate resonances above the dipole. We note that similar embedding approaches were used in earlier ensemble studies [173, 214]. Samples were prepared by magnetron sputtering under vacuum conditions, and their morphology subsequently characterized by STEM, see supplementary information of Publication B, verifying the quality of the conformal coating.

Before proceeding we discuss the simulation approach to model the EELS signal $\Gamma(\omega)$ of single embedded nanoparticles. To account for the full electromagnetic environment simulations are performed using the retarded BEM in the four-potential (ϕ, \mathbf{A}) formulation [95], as provided by the Matlab toolbox MNPBEM [96] with its recent EELS extension [97]. The simulated spectra are of the LRA-kind: a departure from our previous focus necessitated by the capabilities of MNPBEM (though HDM extensions of the BEM and related surface GIEM are possible, see [206]). Meshing-wise, we discretize the surfaces of nanoparticle (experimentally icosahedral, treated in simulations as spheres) and embedding layer in non-uniform triangular and quadrilateral polygonal elements ($\sim 5500 - 8000$ faces in total, depending on particle radii). The in principle infinite horizontal extent of the embedding is circularly truncated to a diameter d_{mem} (ranging from 225 – 300 nm, until convergence); no vertical boundaries are included at the truncation, paralleling their absence in the non-truncated case. The conformal cladding of the nanosphere is modeled by a smoothed spherical cap, with smoothing profile specified by $r_s(z)$ incorporating a characteristic length l_s . Finally, the material dispersion of silver is interpolated from measured data [58], while the cladding's dielectric constant $\epsilon_{\text{SiN}_x} = 3.2$ is estimated by fitting of the simulated and measured dipole energy at large particle radii.

In Fig. 12(a) we compare the measured EEL spectrum of a single nanoparticle of radius $R = 20 \pm 0.2$ nm with simulations for an impact parameter $b = 19.5 \pm 0.25$ nm. The measured signal clearly display two resonances (below the effective plasma frequency around 3.8 eV), a dipole at 2.75 eV and what we interpret as an agglomerate HO multipole peak around 3.25 eV. Simulated measurements, of which we also display a convoluted version which mimics the finite experimental energy-resolution of 0.15 eV, agree qualitatively with these findings: specifically, the convoluted spectra similarly display two primary peaks due to the dipole and agglomerate HO modes, re-

We do not discuss the mathematical underpinnings of the potential-based BEM here (see instead Ref. 95), but return to its nonretarded sibling in Section 3.6.

The indicated radius uncertainty reflects the deviation from absolute circularity, as measured by the difference between major and minor axis of a fitted ellipse, i.e. it does include instrumental uncertainty.

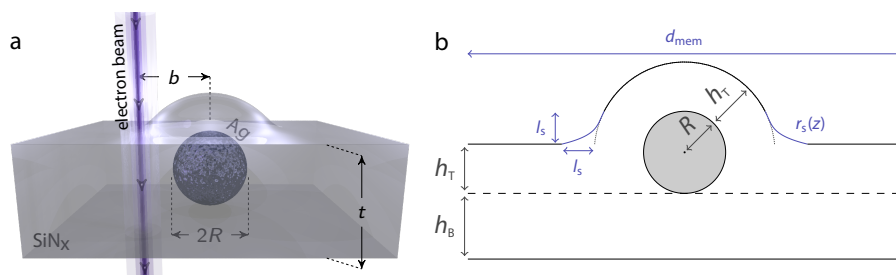


Figure 11: Sketches of a silver (Ag) nanoparticle, embedded in a silicon nitride (SiN_x) conformal coating, excited in (a) by an electron beam with impact parameter b . The coating thickness $t = h_B + h_T$ is estimated from EELS data as 35 nm, with contributions $h_B = 20$ nm and $h_T = 15$ nm highlighted in (b). Additional mesh-related parameters l_s , $r_s(z)$, and d_{mem} are indicated in blue in (b), pertaining to the BEM-calculations.

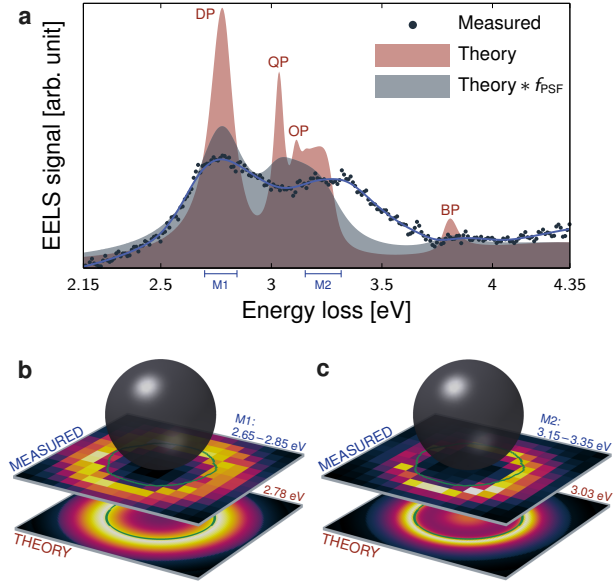


Figure 12: EELS signal of single embedded nanoparticle, $R = 20 \pm 0.2$ nm. (a) EEL spectrum at fixed electron impact position $b = 19.5 \pm 0.25$ nm. Curves are normalized to equal integrated areas. Experimental data (zero-energy loss peak removed) indicated in black markers with associated spline-fit in blue. BEM calculations are shown in red areas with labeled dipole (DP), quadrupole (QP), and octopole (OP) resonances. Gray areas depict the simulations convoluted with a Lorentzian point-spread function f_{PSF} of full width at half maximum 0.15 eV, matching the experimental energy resolution. (b-c) EELS intensity maps across the nanoparticle, comparing experimental [at energies-intervals M1 and M2, as indicated in (a)] and theoretical maps [at the BEM DP and QP energies]. The relative intensity is compared, ranging from dark (minimum) to bright (maximum).

spectively. The primary discrepancy between measured and simulated (LRA) spectra is a blueshift of the measured HO multipoles and its high-energy tail. This shift finds a natural, qualitative explanation in the hydrodynamic l -dependent resonance shifts discussed in Eq. (62) and Section 3.4.3.

The spatial dependence of the probe-position is examined in Figures 12(b-c) which display EELS maps at fixed energy for the dipole and HO multipole peaks. The simulated and measured maps are in good agreement, both underscoring the tighter confinement of the HO multipoles. This tighter confinement can be understood qualitatively from the $|r - R|^{-(l+2)}$ radial decay-behavior of the induced field of the l th order multipole in conjunction with the short radial extent of the exciting electron beam. Quantitatively, it can be understood from the nonretarded multipolar expansion of the EEL probability of a sphere in vacuum (for $b > R$) [22, 209, 216], which can be simplified in the limit $\omega b/v \ll 1$

The asymptotic result is achieved by retaining only the most divergent term, $m = \pm l$, of the Coulomb interaction $K_m^2(\omega b/v)$ which is further simplified by noting the $x \ll 1$ expansion of the modified Bessel function $K_m^2(x) \simeq [2^{l|m|-1}(|m|-1)!x^{-l|m|}]^2$, valid for $m \neq 0$.

$$\begin{aligned} \Gamma_{\text{sph}}^{\text{nr}}(\omega) &= \alpha_v \frac{4R}{\pi v} \sum_{l=1}^{\infty} \sum_{m=-l}^l \left(\frac{\omega R}{v}\right)^{2l} K_m^2\left(\frac{\omega b}{v}\right) \frac{\text{Im}[\tilde{\alpha}_l(\omega)]}{(l+m)!(l-m)!} \\ &\simeq \alpha_v \frac{2R}{\pi v} \sum_{l=1}^{\infty} n_l \left(\frac{R}{b}\right)^{2l} \text{Im}[\tilde{\alpha}_l(\omega)], \end{aligned} \quad (65)$$

with free-electron effective fine-structure constant $\alpha_v \equiv e^2/4\pi\epsilon_0\hbar v$, dimensionless multipolar polarizability $\alpha_l(\omega) \equiv 4\pi R^{2l+1}\tilde{\alpha}_l(\omega)$, see Eq. (18), and coefficients $n_l \equiv 2^{2l}[(l-1)!]^2/(2l)!$ that approach $n_l \simeq \sqrt{\pi}/l^{3/2}$ asymptotically for $l \gg 1$ cf. Stirling's formula. It is clear from Eq. (65) that the contribution of the l th multipole to $\Gamma_{\text{sph}}^{\text{nr}}(\omega)$

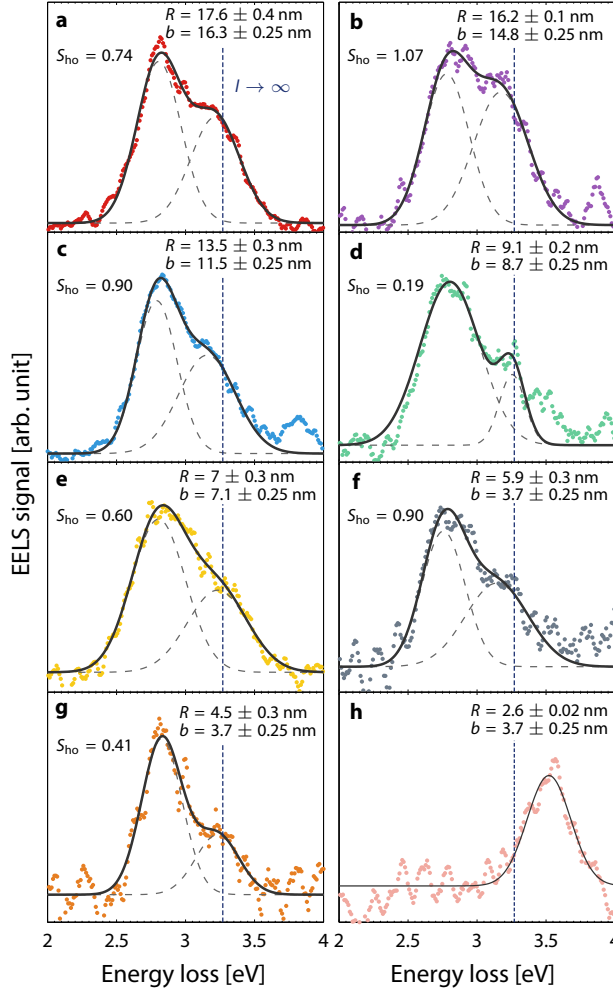


Figure 13: EEL spectra of embedded silver nanoparticle for radii ranging from (a) $R = 17.6 \pm 0.4$ nm to (h) $R = 2.6 \pm 0.2$ nm, acquired for glancing electron impacts, with R and b indicated in insets. Dotted blue lines mark the LRA $l \rightarrow \infty$ SP limit estimated from a bulk embedding treatment. (a-g) A least squares fit of the sum of two Gaussian functions are indicated in gray dashed lines, and their sum in black full lines, whose quality relative to the measured data (markers) underpins the applicability of a two-peak treatment. The spectral weight S_{HO} , defined as the ratio of integrated HO and dipolar Gaussian fits, is given in each panel. (h) Below $R \lesssim 4$ nm we find only a single dominant peak, tentatively attributed as a dipole resonance cf. its excitation for external ($b > R$) impacts. In this size-regime a large blueshift relative to LRA predictions is observed.

exhibits an approximate $(R/b)^{2l}$ decay outside the sphere: accordingly, confinement increases with multipole order l , as confirmed experimentally by [Figures 12\(b-c\)](#).

Observation of quadrupolar and octopolar multipoles in silver particles was previously reported by Yamamoto *et al.* [197] who, using cathode luminescence, found HO multipoles in individual nanospheres in the size-range $2R \gtrsim 140 - 285$ nm, i.e. well into the regime of significant retardation effects. A natural question is whether HO multipoles can be experimentally discerned in truly nanoscopic particles. Evidently, [Figure 12](#) affirms this in the positive sense for a nanoparticle of $R \approx 20$ nm. The progression towards smaller radii is explored further in [Figure 13](#) for particles down to $R \approx 2.6$ nm. We find that multipole plasmons can be identified in a two-peak model all the way down to $R \approx 4.5$ nm. Indeed, the spectral weight of the HO peak is apprecia-

We estimate the $l \rightarrow \infty$ limit through a bulk embedding model, Eq. (61), with associated effective bulk dielectric constant $\epsilon_{\text{SiN}_x}^{\text{bulk}} = 3.3$ obtained by fitting to the measured dipole energy for particle radii 4 – 10 nm. Note the good agreement between fitted values obtained from full BEM calculations, $\epsilon_{\text{SiN}_x} = 3.2$, and $\epsilon_{\text{SiN}_x}^{\text{bulk}} = 3.3$ of the simpler bulk approach.

ble, approaching that of the dipole at several radii. This supports our previous remark that single-peak interpretations of EEL spectra [25, 26, 172] acquired at glancing impact $b \sim R$ is questionable – though we concede that HO multipolar effects will be less pronounced in non-embedded scenarios. Finally, we note that the HO peaks extend into the energy-region beyond the LRA $l \rightarrow \infty$ SP limit, estimated at $\omega_{\text{sp}} \approx 3.27$ eV, in conflict with LRA predictions and in qualitative accordance with the HDM – though to a lesser degree than predicted therein.

In summary, we find in this experimental study that HO multipole LSPs feature prominently in the spectrum of embedded nanoparticles all the way down to $R \approx 4$ nm. These agglomerate peaks, composed of several merged multipoles, furthermore exhibit a blueshifted peak-structure exceeding that predicted by our LRA BEM simulations. Further discussion, particularly of the size-dispersion of the dipolar peaks in the small-radius region and its possible interpretation within HDM and GNOR approaches, is presented in [Publication B](#).

3.6 OUTLOOK

One of the greatest strengths of the HDM – and by extension the GNOR – is its comparative simplicity relative to e.g. fully computational approaches such as TDDFT. Importantly, this simplicity allows analytical considerations, and an associated ease of interpretation, as discussed in the preceding sections. Conversely, this simplicity is arguably also the greatest weakness of the HDM. Comparing the physical content of the HDM with our list of LRA-shortcomings from [Section 3.1](#), it is clear that rather much is omitted. The most severe omission is undoubtedly the neglect of spill-out, but the neglect of xc-effects is similarly troubling. Recent developments have relaxed the hard wall assumption [217–219], and even included additional contributions to the pressure term, [Eq. \(38\)](#), such as the von Weizsäcker correction [220] and the xc-potential. Importantly, these augmentations demonstrate that the HDM can be generalized to predict not only blueshifts but also redshifts, when pertinent. Even so, the simplicity – and certainly the analytical insight – of the HDM is largely forfeited with these extensions that eventually resemble orbital-free TDDFT more than the HDM.

Ideally, a quantum-nanoplasmonic approach should embody two qualities: (i) it should account for the nonlocal and inhomogeneous material-response correctly to a controllable degree, i.e. its description of material-response should derive from first principles, and (ii) it should be computationally efficient, generalizable to arbitrary structures, and physically transparent. It appears to the author that this can be achieved only by a meaningful division of labor. In fact, our discussion of the Feibelman d -parameter in [Section 3.1.1](#) already suggests rather strongly the appropriate division: expand the plasmon’s induced density $\rho_{\text{ind}}(\mathbf{r})$, which is spatially inhomogeneous and surface-peaked, in monopole and dipole contributions at the ionic boundary, accounting for the LRA and leading order quantum effects, respectively. Feibelman [54] indicated how this could be done for the flat interface, with a dipole term $\propto d_{\perp}$, see [Eq. \(30\)](#), and Apell and Ljungbert [138, 139] demonstrated that a similar approach could be applied to the sphere, albeit with a dipole term $\propto d_r$. A key insight in the latter extension was the suggestion that $d_r \simeq d_{\perp}$ for small λ_e/R [139], with λ_e and R denoting the electronic screening length and the local radius of curvature, respectively.

Very recently, these ideas have been generalized to arbitrary structures and to spatially nonlocal d -parameters by Yan *et al.* [136], building, essentially, on the relative invariance of d_{\perp} , d_r , and similarly constructed d -parameters. In closing our considerations of nonclassical plasmonics, we outline a particularly straightforward variant of this approach in the nonretarded and local- d limit, amounting to a simple extension of the nonretarded BEM [221]. Importantly, the approach facilitates a clear separation

of the optical geometry (quasistatic integral equation) and electronic bulk and surface properties (ϵ_{LRA} and d -parameter), thus achieving the above-noted division.

The nonretarded BEM has at its center the surface charge density $\sigma(\mathbf{r}, \omega)$, from which e.g. the potential can be deduced. The approach can be extended to include a dipole-correction by accounting for a dipole density, oriented normal to the surface, equaling $d_\Omega(\omega)\sigma_{\text{ind}}(\mathbf{r}, \omega)$. Here d_Ω denotes the geometry-specific d -parameter in the local limit, while σ_{ind} denotes the induced surface charge due to free-electron contributions. The induced rather than the total surface charge σ is needed because d_Ω does not include bound contributions, cf. Eq. (30), an exclusion justified by their inherently rigid and confined nature. With this augmentation, a dipole-extended nonretarded BEM can be derived, reading (suppressing frequency-dependence)

$$\begin{aligned} \Lambda\sigma(\mathbf{r}) &= \oint_{\partial\Omega} [\hat{\mathbf{n}} \cdot \nabla g(\mathbf{r}, \mathbf{r}')] \sigma(\mathbf{r}') d^2\mathbf{r}' \\ &+ d_\Omega w_\sigma \lim_{\delta \rightarrow 0} \oint_{\partial\Omega} [\hat{\mathbf{n}} \cdot \nabla \nabla' g(\mathbf{r} + \delta\hat{\mathbf{n}}, \mathbf{r}') \cdot \hat{\mathbf{n}}'] \sigma(\mathbf{r}') d^2\mathbf{r}', \end{aligned} \quad (66a)$$

for a geometry Ω with ionic boundary $\mathbf{r} \in \partial\Omega$, associated outward normal-vector $\hat{\mathbf{n}}^{(l)}$ at point $\mathbf{r}^{(l)}$, scalar Green function $g(\mathbf{r}, \mathbf{r}') \equiv 1/|\mathbf{r} - \mathbf{r}'|$, total-to-induced charge density proportionality factor $\sigma_{\text{ind}} \equiv w_\sigma \sigma$ equaling $w_\sigma = (\epsilon^{\text{m}} - \epsilon_{\text{B}}^{\text{m}})/(\epsilon^{\text{m}} - \epsilon^{\text{d}})$, Hessian analogue $\nabla \nabla' \equiv \nabla \otimes \nabla'$, and eigenvalue

$$\Lambda \equiv 2\pi \frac{\epsilon^{\text{d}} + \epsilon^{\text{m}}}{\epsilon^{\text{d}} - \epsilon^{\text{m}}}, \quad (66b)$$

for a metallic interior ϵ^{m} (with bound contribution $\epsilon_{\text{B}}^{\text{m}}$) and dielectric exterior ϵ^{d} . Without the second right-hand term in Eq. (66a), this is just the conventional nonretarded BEM [221], from which e.g. the planar and spherical interface LRA (L)SP conditions follow [$\Lambda_0^{\text{SP}} = 0$ and $\Lambda_0^{\text{LSP}} = -2\pi/(2l+1)$, respectively]. Including the dipole term, we recover – and simultaneously generalize to non-unity values of ϵ^{d} and $\epsilon_{\text{B}}^{\text{m}}$ – the previous isolated results for the planar [54] and spherical system [138, 139], via

$$\Lambda^{\text{SP}} = -2\pi k_{\parallel} d_{\perp} w_\sigma, \quad \Lambda^{\text{LSP}} = -\frac{2\pi}{2l+1} (1 + 2lk_{l+1} d_r w_\sigma), \quad (67)$$

with $k_l \equiv l/R$. These results readily replicate Eqs. (32) and (34) in the special case $\epsilon^{\text{d}} = 1$ and $\epsilon^{\text{m}}(\omega) = 1 - \omega_{\text{p}}^2/\omega^2$. The correspondence demonstrates the validity of Eq. (66a).

Though the dipole-extended BEM is not analytically solvable in general geometries, one can still obtain a result akin to Eqs. (32) and (34) by applying perturbation theory to the n th eigensolution, denoted $\{\Lambda^n, \sigma^n(\mathbf{r})\}$, by expanding around the classical LRA eigensolution, $\{\Lambda_0^n, \sigma_0^n(\mathbf{r})\}$, obtained from Eq. (66a) with $d_\Omega = 0$. The classical charge densities $\sigma_0^n(\mathbf{r})$ are biorthogonal to their corresponding potentials $\phi_0^n(\mathbf{r}) = (4\pi\epsilon_0)^{-1} \int_{\partial\Omega} g(\mathbf{r}, \mathbf{r}') \sigma_0^n(\mathbf{r}') d^2\mathbf{r}'$, such that $\int_{\partial\Omega} \phi_0^{n*}(\mathbf{r}) \sigma_0^m(\mathbf{r}) d^2\mathbf{r} = \langle \phi_0^m | \sigma_0^n \rangle \propto \delta_{nm}$. Using this, we derive the first-order correction $d_\Omega \Lambda_1^n$ to Λ^n as

$$\Lambda^n = \Lambda_0^n + d_\Omega w_\sigma \Lambda_1^n + \mathcal{O}(d_\Omega^2), \quad (68a)$$

$$\Lambda_1^n = \frac{\langle \phi_0^n | D | \sigma_0^n \rangle}{\langle \phi_0^n | \sigma_0^n \rangle} = \frac{\lim_{\delta \rightarrow 0} \oint_{\partial\Omega} \phi_0^{n*}(\mathbf{r}) D_\delta(\mathbf{r}, \mathbf{r}') \sigma_0^n(\mathbf{r}') d^2\mathbf{r}' d^2\mathbf{r}}{\oint_{\partial\Omega} \phi_0^{n*}(\mathbf{r}) \sigma_0^n(\mathbf{r}) d^2\mathbf{r}}, \quad (68b)$$

where the operator D is defined by its action $\langle \mathbf{r} | D | \sigma \rangle \equiv \lim_{\delta \rightarrow 0} \oint_{\partial\Omega} D_\delta(\mathbf{r}, \mathbf{r}') \sigma(\mathbf{r}') d^2\mathbf{r}'$ with elements $D_\delta(\mathbf{r}, \mathbf{r}') \equiv \hat{\mathbf{n}} \cdot \nabla \nabla' g(\mathbf{r} + \delta\hat{\mathbf{n}}, \mathbf{r}') \cdot \hat{\mathbf{n}}'$. This remarkable result demonstrates that one can obtain the leading-order *nonclassical* correction from the *classical* solution and d_Ω alone! The correction Λ_1^n plays the role of an effective momentum. Its length scale can be discerned by introducing dimensionless coordinates $\tilde{\mathbf{r}} = \mathbf{r}/L$ through a geometric scale L , which renders $\Lambda_1^n = \tilde{\Lambda}_1^n/L$ with dimensionless $\tilde{\Lambda}_1^n$ obtained from $\tilde{\mathbf{r}}^{(l)}$ -integration of Eq. (68b). The special case of a free-electron gas adjacent to a vacuum scenario ($\epsilon_{\text{B}}^{\text{m}} = \epsilon^{\text{d}} = 1$), provides a particularly transparent formula

The limiting procedure for the dipole correction reflects the infinitesimal (but nonzero) extent of a dipole – the procedure can be interpreted as a “directional” Cauchy principal value, oriented along the surface normal. In addition, we implicitly assume ordinary in-plane principle values applying as $\mathbf{r}' \rightarrow \mathbf{r}$ in both integral terms.

The derivation of Eq. (67) requires some labor, but can be obtained by starting from ansatzes $\sigma_{\text{sp}}(\mathbf{r}) \propto \exp(ik_{\parallel}x)$ and $\sigma_{\text{LSP}}(\mathbf{r}) \propto P_l(\cos\theta)$ applied to Eq. (66a), whose integration is possible though cumbersome, whereafter only the $\delta \rightarrow 0$ limit remains.

Perturbation theory in biorthogonal systems matches the regular kind, requiring only a change of bra to its biorthogonal partner, see e.g. Ref. 222.

for the corrected frequency ω^n , expressible by the LRA frequency ω_0^n and the noted eigenvalues

$$\omega^n(\omega^n + i\gamma) = \omega_0^n(\omega_0^n + i\gamma) \left(1 + \frac{\tilde{\Lambda}_1^n}{1 + \Lambda_0^n} \frac{d_\Omega}{L} \right). \quad (69)$$

The only remaining conceptual worry is the evaluation of d_Ω which in principle depends on both ω and Ω . The dependence on Ω is weak, however, as has already been discussed above. Specifically, when the radius of curvature of Ω is large relative to the screening length the geometry is, from an electronic perspective, essentially planar and $d_\Omega \simeq d_\perp$. These considerations are supported by recent (sodium-)calculations demonstrating a relatively k_\parallel -invariant $d_\perp(k_\parallel, \omega)$ for $k_\parallel \lesssim 0.1k_F$ [136], with k_F denoting Fermi momentum. In summary, d_Ω can be excellently approximated by its planar, local equivalent $d_\perp(k_\parallel \rightarrow 0, \omega)$, evaluated at the resonance frequency [being in a pole approximation just the LRA correspondent ω_0^n]. Crucially, d_\perp can be extracted accurately and efficiently from TDDFT treatments of the semi-infinite half-space, e.g. within the adiabatic local density approximation. In this way, the first three shortcomings of the LRA (discontinuity, locality, and incompleteness) can be sidestepped, and the desired division of labor is automatically instated. For these reasons, we believe the above approach holds significant promise for providing a rigorous and straightforward quantum-to-classical bridge, connecting the sometimes disconnected efforts of the nanoplasmonic and condensed matter communities.

A solitary deficiency of the dipole approach is that CBPs cannot be captured, since they are not surface-localized.

3.7 SUMMARY

In this chapter, we have discussed a few of the myriad effects that exist in nanoplasmonics beyond the LRA. The shortcomings of the hard wall LRA were discussed in [Section 3.1](#), falling under the categories of discontinuity, locality, incompleteness, and quantum size effects. Motivated by its interpretational power, the section concluded by introducing the Feibelman d -parameter, encompassing elements pertaining to the first three shortcomings. Next, in an effort to explore a single of these deficiencies, nonlocality, we introduced the hydrodynamic extension of the LRA Drude model, the HDM, in [Section 3.2](#), discussing its derivation, key properties, relation with the Lindhard model, extensions to include Kreibig broadening, and lastly its applicability to real materials. With the general framework of the HDM established, we considered its predictions in two analytically soluble cases, namely planar systems in [Section 3.3](#) and nanospheres in [Section 3.4](#). In the latter case, we also examined the role of the measurement probe, particularly its vicinity to the metallic surface, and found that resonances beyond the dipole, i.e. HO multipoles, were responsible for the primary differences between near- and far-field measurements. [Section 3.5](#) extended these considerations by experimental EELS measurements of embedded nanospheres, demonstrating the prominent role of HO multipoles in EEL spectra acquired at glancing electron impacts. Finally, looking forward and beyond the HDM we considered in [Section 3.6](#) a generalization of Feibelman's approach to arbitrary geometries, which allowed us to conveniently separate the roles of optical geometry and electronic surface response.

Part II

PLASMONICS IN TWO DIMENSIONS

4

ELECTRONIC PROPERTIES OF GRAPHENE

Although graphene in no sense constitutes a solitary entry in two-dimensional plasmonics, it does in many ways pose as a uniquely interesting entry, and inarguably is the entry of principal current interest [40–46]. It is the purpose of this chapter to introduce the basic properties of graphene as necessary for a treatment of its plasmonic properties. Invariably, such an objective involves the condensed matter properties of graphene’s conduction electrons, which, coincidentally, are responsible for an appreciable portion of graphene’s fame. The literature on these aspects is comprehensive and today largely well-established [223]. As a consequence, we aim here only to provide a cursory overview, seeking essentially to equip ourselves with a set of tools necessary for further exploration.

The chapter is composed of two components. First, we introduce graphene briefly, and proceed immediately to a discussion of its unique electronic properties. Subsequently, equipped with an understanding of graphene’s single-particle features, we approach the topic of its low-energy optical response.

4.1 ELECTRONIC PROPERTIES OF GRAPHENE

While graphene was experimentally discovered in just 2004(–2005) [34–36], interest in its theoretical properties predate this turning point by several decades: initially, graphene played the role of a construct in considerations of (intercalated) graphitic systems [228–231], and later found similar use in the study of carbon nanotubes [232–234]. The standard approach to graphene, via a tight-binding (TB) description, is today essentially equivalent to its 1947-treatment by Wallace [228]. Below we briefly review the results of the TB approach to the energy dispersion of extended graphene, emphasizing the important low-energy features near the Dirac point, and establish finally some useful auxiliary quantities. First, however, we recall some basic properties of graphene’s structure.

Although the Nobel-awarded 2004 paper [34] today often is juxtaposed with the “discovery” of graphene, its introduction was far more gradual, and indeed older, than that [224, 225], see e.g. the comments of Refs. 226 and 227.

4.1.1 Structural composition

Structurally, graphene is composed of carbon atoms distributed on a hexagonal lattice, as indicated in Figure 14a, with inter-atomic spacing $a_{CC} \approx 1.42 \text{ \AA}$. The lattice itself is decomposable into a triangular Bravais lattice with a two-atom basis, contained e.g. in a rhombic unit cell, spanned by lattice-vectors $\mathbf{a}_1 = \frac{1}{2}a_{LC}[\sqrt{3}, 1]^T$ and $\mathbf{a}_2 = a_{LC}[0, 1]^T$ with lattice-constant $a_{LC} = \sqrt{3}a_{CC} \approx 2.46 \text{ \AA}$. Conceptually, the lattice can be considered composed of two triangular sublattices, denoted A and B, as indicated explicitly in Figure 14a. The corresponding reciprocal lattice is spanned by $\mathbf{b}_1 = \frac{4\pi}{\sqrt{3}}a_{LC}^{-1}[1, 0]^T$ and $\mathbf{b}_2 = \frac{2\pi}{\sqrt{3}}a_{LC}^{-1}[-1, \sqrt{3}]^T$, and is shown in Figure 14b. The associated first Brillouin zone (FBZ) is similarly depicted with four points of high symmetry indicated. The six so-called Dirac points \mathbf{K} delineate the boundaries of the FBZ, and come in two inequivalent variants $\mathbf{K}^\pm = \frac{4\pi}{3}a_{LC}^{-1}[0, \pm 1]^T$ (the remaining four being connected by reciprocal lattice vectors) and are of singular importance to our low-energy considerations.

All vectorial quantities in this chapter are implicitly in-plane, i.e. two-dimensional; in Chapter 5 we introduce additional notation to distinguish vectors of different dimensionality.

The M-point designates a van Hove singularity; its importance, however, is minor for our low-energy considerations.

The orbital electron configuration of graphene is $1s^2 2s^1 2p^3$, departing from the $1s^2 2s^2 2p^2$ configuration of atomic carbon due to orbital mixing with neighboring

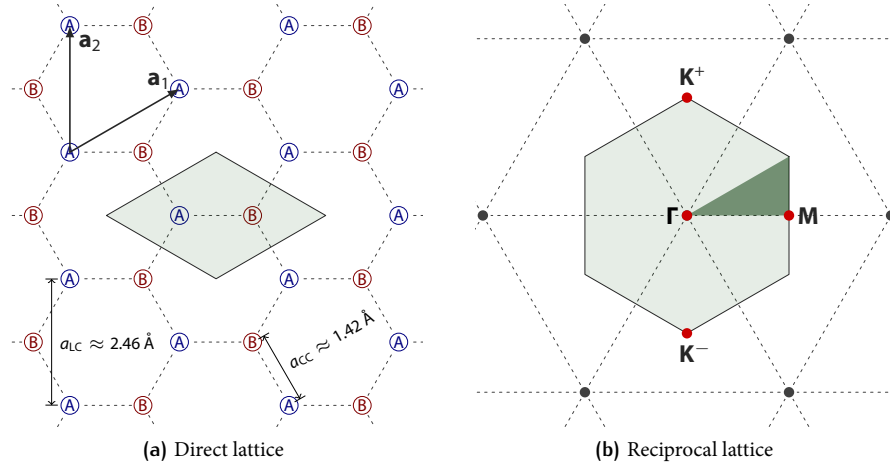


Figure 14: The direct and reciprocal lattices of graphene. (a) The lattice is decomposed in A- and B-sites, with inter-site spacing a_{CC} and lattice-constant a_{LC} . In addition to the lattice vectors $\mathbf{a}_{1,2}$ we show in shaded green one realization of the unit cell. (b) Besides the reciprocal lattice in black markers, we depict in shaded green the FBZ, with high-symmetry points in red markers and the irreducible FBZ in dark green.

atoms that promote an electron from the $2s$ orbital to the $2p$ orbital. The remaining $2s$ orbital hybridizes with the half-filled $2p_x$ and $2p_z$ orbitals, and form strong in-plane σ bonds which are responsible for the hexagonal structure of graphene (and for its mechanical properties). The electrons in the half-filled $2p_z$ orbital, however, do not contribute significantly to bonding because the orbital is oriented perpendicularly to the plane. Instead, these electron are responsible for the conductive properties, forming, as we will see, an occupied valence band and an empty conduction band, occasionally known as the π and π^* bands. In the following, we examine these bands from the perspective of a TB treatment applied to the $2p_z$ orbitals.

4.1.2 Tight-binding approach

For an extended system with discrete translational symmetry, the wave function must necessarily obey Bloch's theorem. Accordingly, we split the $2p_z$ wave function into two distinct Bloch functions, $|A_{n\mathbf{k}}\rangle$ and $|B_{n\mathbf{k}}\rangle$, belonging to the A- and B-sublattice, respectively, and with an anticipated dependence on band-index n and momentum \mathbf{k} . Next, we expand each sublattice Bloch function in sublattice-specific localized Wannier states, $|l^A\rangle$ and $|l^B\rangle$, according to $|A_{n\mathbf{k}}\rangle = \sum_l \psi_{l,n\mathbf{k}}^A |l^A\rangle$ and $|B_{n\mathbf{k}}\rangle = \sum_l \psi_{l,n\mathbf{k}}^B |l^B\rangle$, with l extending over the *triangular* lattice sites (with positions \mathbf{r}_l) and with expansion coefficients $\psi_{l,n\mathbf{k}}^{A,B} = \psi_{n\mathbf{k}}^{A,B} e^{i\mathbf{k}\cdot\mathbf{r}_l}$ with spatial dependence following from Bloch's theorem. In brief terms, the TB approach consists in transforming the above expansions to a coupled system with associated eigenenergy $\epsilon_{n\mathbf{k}}$, reading as

$$\epsilon_{n\mathbf{k}} \psi_{l,n\mathbf{k}}^A = -t_{AB} \sum_{l' \in \text{nn}_{lA}} \psi_{l',n\mathbf{k}}^B, \quad (70a)$$

$$\epsilon_{n\mathbf{k}} \psi_{l,n\mathbf{k}}^B = -t_{AB} \sum_{l' \in \text{nn}_{lB}} \psi_{l',n\mathbf{k}}^A, \quad (70b)$$

Different values of the hopping parameter proliferate: we adopt the value of the prevailing theoretical review [223].

written here in the nearest-neighbor approximation, with $\text{nn}_{lA(B)}$ denoting the three nearest B-site (A-site) neighbors of an A-site (B-site) lattice-point l . The coupling strength is governed by the so-called hopping parameter $t_{AB} = 2.8$ eV, which express the overlap between neighboring Wannier orbitals. Upon inserting the Bloch form of

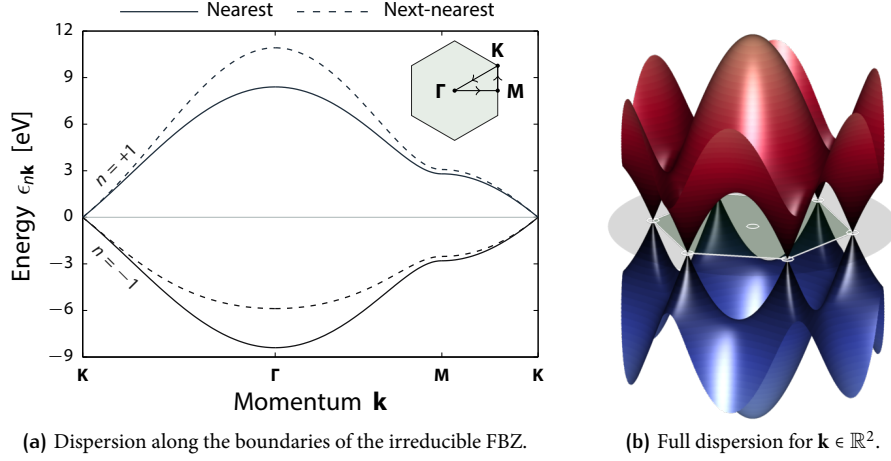


Figure 15: The energy dispersion of graphene's $2p_z$ electrons within a TB treatment with hopping parameter $t_{AB} = 2.8$ eV. (a) Dispersion along the boundary of the irreducible FBZ, see inset, in nearest and next-nearest neighbor approximations, in the latter case with hopping parameter $t_{AA'} = t_{AB}/10$. (b) Dispersion in three-dimensional view; the FBZ is indicated in shaded green and delimited by white lines.

$\psi_{l,n\mathbf{k}}^{A,B}$ the above equations reduce to a simple Hamiltonian form $\mathbf{H}_{\mathbf{k}}\psi_{n\mathbf{k}} = \epsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}$ for the undetermined expansion coefficients $\psi_{n\mathbf{k}} \equiv [\psi_{n\mathbf{k}}^A, \psi_{n\mathbf{k}}^B]^T$ with

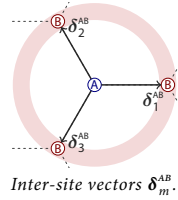
$$\mathbf{H}_{\mathbf{k}} = -t_{AB} \begin{bmatrix} 0 & f_{\mathbf{k}} \\ f_{\mathbf{k}}^* & 0 \end{bmatrix}. \quad (71)$$

The characteristic function $f_{\mathbf{k}} \equiv \sum_{m=1}^3 e^{i\mathbf{k} \cdot \boldsymbol{\delta}_m^{AB}}$ sums the Bloch phase differences between neighboring A- and B-sites. For the lattice orientation in Figure 14a the inter-site vectors $\boldsymbol{\delta}_m^{AB}$ take the values $\boldsymbol{\delta}_1^{AB} \equiv a_{CC}[1, 0]^T$, $\boldsymbol{\delta}_2^{AB} \equiv \frac{1}{2}a_{CC}[-1, \sqrt{3}]^T$, and $\boldsymbol{\delta}_3^{AB} \equiv \frac{1}{2}a_{CC}[-1, -\sqrt{3}]^T$.

The Hamiltonian can be conveniently expressed as $\mathbf{H} = -t_{AB}\boldsymbol{\sigma} \cdot \mathbf{R}_{\mathbf{k}}$ in a basis of Pauli matrices $\boldsymbol{\sigma} \equiv [\sigma_x, \sigma_y, \sigma_z]^T$ with real coefficients $\mathbf{R}_{\mathbf{k}} \equiv [\text{Re } f_{\mathbf{k}}, -\text{Im } f_{\mathbf{k}}, 0]^T$. The desired object of the present analysis, the eigenenergies $\epsilon_{n\mathbf{k}}$, then follow immediately as

$$\epsilon_{n\mathbf{k}} = nt_{AB}|f_{\mathbf{k}}|, \quad (72)$$

with band-index $n = \pm 1$. The resulting energy-dispersion with momentum is illustrated in Figure 15. The most striking feature is the emergence of cone-like bands which touch at the Dirac points \mathbf{K} : i.e. the bands are gap-free and degenerate at \mathbf{K} . Since the $2p_z$ orbital is half-filled at equilibrium, the lower band ($n = -1$) is completely filled (at zero temperature) while the upper band ($n = +1$) is completely empty; accordingly, they correspond to the valence and conduction bands, respectively. In Figure 15a we examine additionally the influence of extending the TB treatment to the next-nearest neighbors [223] by including an additional hopping parameter $t_{AA'}$, chosen heuristically as $t_{AB}/10$. Evidently, although adjustments appear away from the Dirac points, particularly in the form of band-asymmetry, this additional interaction retains the characteristic cone-like dispersion near \mathbf{K} . In fact, these low-energy features are robust, as can be shown rigorously by the $\mathbf{k} \cdot \mathbf{p}$ method using the symmetry groups of the lattice [229]. Accordingly, the nearest-neighbor TB approach is entirely representative and fully adequate in the low-energy region. With this in mind, we next examine more closely these unique low-energy features.



Inter-site vectors $\boldsymbol{\delta}_m^{AB}$.

These properties earn graphene the classification of 'semi-metal'.

We mention one minor modification, namely the opening of a very small bandgap ~ 1 μeV when spin-orbit interaction is included [235].

4.1.3 The Dirac limit

The low-energy limit of the energy-dispersion, Eq. (72), follows by expanding around either of the Dirac points \mathbf{K}^\pm : for the governing phase term $f_{\mathbf{k}}$ this produces to first order $f_{\mathbf{k}} = \frac{3}{2}a_{\text{CC}}(\mathbf{k} - \mathbf{K}^\pm)[i, \mp 1]^T + \mathcal{O}(|\mathbf{k} - \mathbf{K}^\pm|^2/a_{\text{CC}}^2)$. By introducing relative momenta $\mathbf{q} \equiv \mathbf{k} - \mathbf{K}^\pm$ the famously linear dispersion emerges

$$\epsilon_{n\mathbf{q}} = n\hbar v_{\text{F}} q, \quad (73)$$

with Fermi velocity $v_{\text{F}} \equiv 3a_{\text{CC}}t_{\text{AB}}/2\hbar \approx 9.1 \times 10^5$ m/s. The difference from the standard parabolic dispersion $\epsilon_q \propto q^2$ is striking, and, indeed, responsible for the majority of the unique electronic features in graphene. At a basic level, it is immediately clear that graphene requires a distinct perspective, since e.g. the conventional effective mass $m_{\text{eff}} \equiv \hbar^2(\partial_q^2 \epsilon_q)^{-1}$ loses its meaning. The linear low-energy behavior of Eq. (73) is conventionally referred to as Dirac dispersion or as the Dirac approximation. The name is owed to the formal equivalence between the low-energy Hamiltonian of graphene and the Dirac equation of massless particles, differing only in their characteristic velocities [230]. We return to some aspects of this Dirac description later in Section 6.2.3.

The Dirac limit constitutes the main workhorse of graphene plasmonics for at least two reasons. Firstly, it allows significant analytical headway, e.g. in the shape of closed-form optical response functions. Secondly, and more importantly, the achievable doping levels of graphene essentially remain within the validity-range of the Dirac approximation. Roughly speaking, the approximation remains excellent for energies up to $\lesssim 2$ eV, whereafter corrections due to e.g. trigonal warping, an $\mathcal{O}(q^2)$ -effect, must be included. In practice, we will not venture beyond frequencies exceeding ~ 1 eV, i.e. we restrict ourselves to THz domain.

4.1.4 Miscellaneous auxiliary properties

In closing this section, we summarize two useful electronic properties of graphene within the Dirac approximation. Firstly, we discuss the electronic density of states (DOS) and secondly relate the carrier density and Fermi energy.

DENSITY OF STATES The DOS is defined by a straightforward summation of states $\text{DOS}(\epsilon) \equiv \mathcal{A}^{-1} \sum_{s\kappa n\mathbf{q}} \delta(\epsilon - \epsilon_{n\mathbf{q}})$, with sample area \mathcal{A} , and spin- and valley-summations s and κ . These two sums are immediately resolved, cf. the degeneracies of $\epsilon_{n\mathbf{q}}$, yielding each a factor 2. The mirror symmetry of the Dirac dispersion $\epsilon_{n\mathbf{q}} = n\hbar v_{\text{F}} q$ reduces the n summation to a factor $\delta_{\text{sign}(\epsilon), n}$, such that $\text{DOS}(\epsilon) = 4\mathcal{A}^{-1} \sum_{\mathbf{q}} \delta(|\epsilon| - \hbar v_{\text{F}} q)$. The remaining \mathbf{q} -summation is transformed by the usual large-sample rule $\sum_{\mathbf{q}} \rightarrow (2\pi)^{-2} \mathcal{A} \int d^2\mathbf{q}$, yielding finally

$$\text{DOS}(\epsilon) = \frac{2|\epsilon|}{\pi\hbar^2 v_{\text{F}}^2}. \quad (74)$$

We highlight that the DOS vanishes at the Dirac points: as we discuss next, this allows sizable changes of the Fermi energy with modest carrier densities.

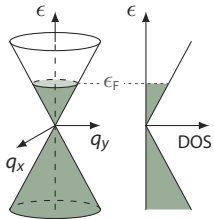
CARRIER DENSITY & FERMİ ENERGY The total area-normalized number of states in the conduction band, i.e. the carrier density n_0 , is $n_0 \equiv \int_0^\infty f(\epsilon) \text{DOS}(\epsilon) d\epsilon$, that is, the integrated DOS weighted by the Fermi-Dirac occupation $f(\epsilon)$. In the important low-temperature limit (and positive Fermi energy ϵ_{F}) the occupation reduces to a step-function $\theta(\epsilon_{\text{F}} - \epsilon)$, yielding the carrier density

$$n_0 = \int_0^{\epsilon_{\text{F}}} \text{DOS}(\epsilon) d\epsilon = \frac{\epsilon_{\text{F}}^2}{\pi\hbar^2 v_{\text{F}}^2} = \frac{k_{\text{F}}^2}{\pi}, \quad (75)$$

Though we here make a distinction between \mathbf{k} and \mathbf{q} , we will generally prefer the Dirac perspective; accordingly, both \mathbf{k} and \mathbf{q} are henceforth implicitly measured relative to a Dirac point.

We warn in advance, that the Dirac equation studied in Section 6.2.3 does not follow directly from our present considerations, corresponding instead to a rotated and reflected lattice, preferred for reasons of algebraic simplicity.

Six Dirac points reside on the boundary of the FBZ, each including a third of their vicinity inside the FBZ: thus, the valley degeneracy sums to $6 \times \frac{1}{3} = 2$.



Energy dispersion and DOS in the Dirac limit.

with Fermi momentum defined by $\hbar v_F k_F \equiv \epsilon_F$. Equivalently, the Fermi energy (momentum) is fixed by the carrier density via $\epsilon_F = \hbar v_F \sqrt{\pi n_0}$ ($k_F = \sqrt{\pi n_0}$).

It is worthwhile to consider briefly the relation between dopant electrons and Fermi energies in a concrete scenario, e.g. in a $10 \times 10 \text{ nm}^2$ sample: there, just a single dopant electron procures $\epsilon_F \approx 0.1 \text{ eV}$ (or, equivalently, $n_0 = 10^{12} \text{ cm}^{-2}$), while 14 electrons afford $\epsilon_F \approx 0.4 \text{ eV}$ (or, $n_0 = 1.4 \times 10^{13} \text{ cm}^{-2}$). These large shifts of Fermi energy with modest doping equip graphene with one of its most attractive technological traits, namely tunability. Relative to e.g. bulk metals, the degree of tunability is impressive: for a $10 \times 10 \times 10 \text{ nm}^3$ gold cube, a Fermi energy shift of, say 0.1 eV , requires $\sim 1.5 \times 10^3$ dopant electrons. Shortly put, graphene's tunability is superior to noble metals' for the compound reasons of reduced dimensionality, unique DOS, and a low (zero) equilibrium Fermi energy.

Experimentally, the doping of graphene can be achieved either by chemical means [236] or by electrostatic gating [237]. While the former is more accessible, the latter undoubtedly holds the greatest potential for in-situ modulation. Although Fermi energies exceeding 1 eV have been realized [238], the values attained in most plasmonics-oriented studies have been lower [106, 239, 240]; recent efforts, however, are gradually closing this gap [241].

Even for thin-films, e.g. a $10 \times 10 \times 0.3 \text{ nm}^3$ gold film, one requires ~ 50 electrons for a 0.1 eV Fermi energy shift.

4.2 OPTICAL RESPONSE OF GRAPHENE

The dominant theoretical framework for the response of graphene is, just as for metals, the RPA. In the RPA the induced response to a perturbation is obtained self-consistently from the *total* perturbation, i.e. the sum of external and induced quantities, coupled with a *noninteracting* response function. In the Dirac limit, the relevant response functions, namely the density-density response and the conductivity, can be derived analytically. In this section we present and summarize the most relevant results. It is not, however, the purpose of this section to present full derivations; they are lengthy and of mostly technical interest.

In the following we consider first the nonlocal density-density response function in the low-temperature limit, and subsequently the local conductivity in both finite- and low-temperature limits. Along the way, we comment also on the appropriate treatment of decay in graphene.

Throughout this section we assume $\epsilon_F > 0$; hole-doped equivalents follow by replacing $\epsilon_F \rightarrow |\epsilon_F|$.

The derivations omitted here were previously provided in full detail in the author's MSc thesis.

4.2.1 Nonlocal density-density response

The noninteracting density-density response function $\chi^0(q, \omega)$ links the induced density n_{ind} with the total potential ϕ via $n_{\text{ind}}(q, \omega) = \chi^0(q, \omega)\phi(q, \omega)$. For a system of discrete translational symmetry $\chi^0(q, \omega)$ follows from the result first derived by Ehrenreich and Cohen [242] (derivable by the self-consistent field method or from the Kubo formalism)

$$\chi^0(\mathbf{q}, \omega) = \frac{g_d}{\mathcal{A}} \sum_{nn'\mathbf{k}} \frac{f_{n\mathbf{k}} - f_{n'\mathbf{k}+\mathbf{q}}}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}+\mathbf{q}} + \hbar(\omega + i\eta)} |\langle n\mathbf{k} | e^{-i\mathbf{q}\cdot\mathbf{r}} | n'\mathbf{k} + \mathbf{q} \rangle|^2, \quad (76)$$

with Fermi-Dirac functions $f_{n\mathbf{k}}$ evaluated at energies $\epsilon_{n\mathbf{k}}$ and with a graphene-specific degeneracy factor $g_d = 4$ from spin- and valley-contributions. The expression is a generalization of Lindhard's result for the free-electron gas [166] to systems described by Bloch states $|n\mathbf{k}\rangle$. The primary revision is the introduction of a general (as opposed to parabolic) energy dispersion and a form factor $\langle n\mathbf{k} | e^{-i\mathbf{q}\cdot\mathbf{r}} | n'\mathbf{k} + \mathbf{q} \rangle$, which for graphene reduces to $\langle n\mathbf{k} | e^{-i\mathbf{q}\cdot\mathbf{r}} | n'\mathbf{k} + \mathbf{q} \rangle = \psi_{n\mathbf{k}}^\dagger \psi_{n'\mathbf{k}+\mathbf{q}}$; in the Dirac limit, the squared form factor simplifies to just $|\langle n\mathbf{k} | e^{-i\mathbf{q}\cdot\mathbf{r}} | n'\mathbf{k} + \mathbf{q} \rangle|^2 = \frac{1}{2} \left[1 + nn' \frac{\mathbf{k}\cdot(\mathbf{k}+\mathbf{q})}{|\mathbf{k}||\mathbf{k}+\mathbf{q}|} \right]$. With these matters

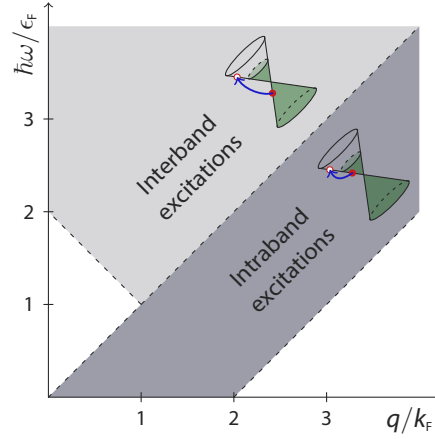


Figure 16: Regions of Landau damping in extended graphene, corresponding to values of $\chi^0(q, \omega) > 0$, are highlighted by shaded areas. Conversely, white areas correspond to regions devoid of Landau damping. The transition type is indicated explicitly in text labels and schematically by insets.

established, the remaining obstacles are of a purely mathematical kind. In the low-temperature limit, the obstacles can be overcome, as shown by Wunsch *et al.* [243] and Hwang and Das Sarma [244], through an appropriate partitioning of terms and subsequent use of residue theory, resulting for $\eta \rightarrow 0^+$ in

$$\chi^0(q, \omega) = \text{DOS}(\epsilon_F) \left(-1 + \frac{1}{8} F(q, \omega) \left\{ \mathcal{W}(\Omega_+) - i\pi - \theta(\Omega_- + 1) \mathcal{W}(\Omega_-) - \theta(-\Omega_- - 1) [\mathcal{W}(-\Omega_-) - i\pi] \right\} \right), \quad (77a)$$

$$\text{with } F(q, \omega) \equiv \frac{\tilde{q}^2}{\sqrt{\tilde{\omega}^2 - \tilde{q}^2}}, \quad \mathcal{W}(x) = x\sqrt{x^2 - 1} - \ln(x + \sqrt{x^2 - 1}), \quad (77b)$$

and $\Omega_{\pm} \equiv (\tilde{\omega} \pm 2)/\tilde{q}$ expressed in normalized quantities $\tilde{q} \equiv q/k_F$ and $\tilde{\omega} \equiv \hbar\omega/\epsilon_F$. We will not discuss the values attained by $\chi^0(q, \omega)$ here – they hold little value on their own, requiring a self-consistent consideration – but we briefly comment on the (q, ω) -pairs where $\text{Re } \chi^0(q, \omega) > 0$, i.e. we discuss the regions of Landau damping. Concretely, two distinct types of Landau transitions arise, of intraband and interband kind, as shown in Figure 16. The latter type is particularly important because it remains in the local limit, i.e. for vanishing momenta: specifically, for $\hbar\omega > 2\epsilon_F$ vertical interband transitions are allowed. We also note the existence of a triangular region without Landau transitions for small momenta and frequencies – fortunately, as we will see in Section 5.2.2, this region coincides with graphene’s plasmon dispersion.

Lastly, we introduce a minor extension of Eq. (77a) related to inclusion of a Drude decay rate γ (whose value we return to shortly). The addition of a Drude decay is motivated, as in metals, by the existence of scattering mechanisms besides Landau damping, e.g. due to impurities. Although the relaxation-time approximation $\omega \rightarrow \omega + i\gamma$ is often used, and will indeed be used also by us in Section 6.2, it is strictly speaking an erroneous approach because it conflicts with the continuity equation and so fails to obey particle conservation. The correct approach was outlined by Mermin [153] in what is today known as the Mermin-corrected relaxation-time approximation,

$$\chi_{\text{RTA}}^0(q, \omega) = \frac{(1 + i\gamma/\omega)\chi^0(q, \omega + i\gamma)}{1 + (i\gamma/\omega)[\chi^0(q, \omega + i\gamma)/\chi(q, 0)]}. \quad (78)$$

We note a detail of technical importance in the application of Eq. (78): specifically, evaluation of $\chi^0(q, z)$ for complex z does *not* follow from a naïve $\omega \rightarrow z$ analytical

Mermin’s critical insight was the observation that scattering incurs relaxation not to a global equilibrium but to a local equilibrium which accounts for the influence of a nonzero perturbing field.

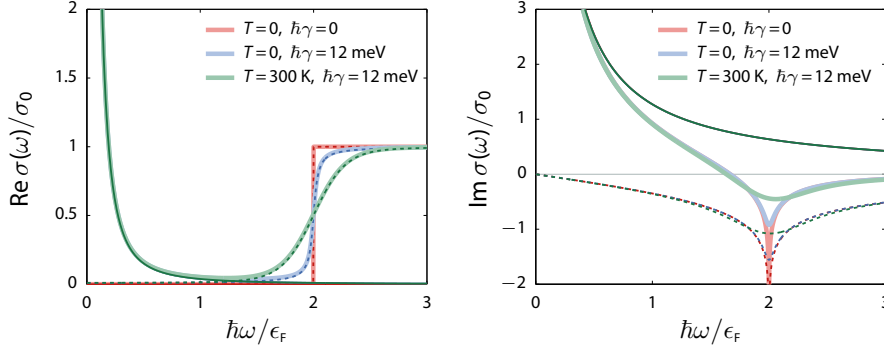


Figure 17: Real and imaginary parts of the local conductivity of extended graphene (normalized to $\sigma_0 \equiv e^2/4\hbar$). Intra- and interband contributions are indicated by thin unbroken and dotted lines, respectively, while their sum is shown in faded thick unbroken lines. Three distinct scenarios are considered, as indicated in legend-entries, and calculated by means of Eqs. (80) and (81) with $\epsilon_F = 0.4$ eV.

continuation of Eq. (77a) (whose derivation relies explicitly on taking the limit $\eta \rightarrow 0^+$). The appropriate analytical continuation, however, is given in Ref. 245. We will have opportunity to apply the Mermin-correction later in Figures 20 and 21.

4.2.2 Local-response conductivity

Our final purpose in this chapter is the consideration of the *surface* conductivity of graphene. We return to its explicit connection with macroscopic fields in Section 5.2.1, and here observe merely that it links the induced surface current \mathbf{K} with the total in-plane field \mathbf{E}_{\parallel} . Its nonlocal form follows immediately from our earlier considerations of $\chi^0(q, \omega)$, to which it is linked through [32]

$$\sigma(q, \omega) = \frac{ie^2\omega}{q^2} \chi^0(q, \omega), \quad (79)$$

as can be derived from the continuity equation. Typically, we will be interested predominately in the local response, $q/k_F \rightarrow 0$, limit of the conductivity. This limit can be deduced from a Taylor expansion of Eq. (76) to second order in $q/k_F \rightarrow 0$ combined with Eq. (79), which lead to a temperature-dependent conductivity $\sigma(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega)$ with distinct contributions due to intra- and interband transitions [39, 246]

$$\sigma_{\text{intra}}(\omega) = \frac{2ie^2k_B T}{\pi\hbar^2(\omega + iy)} \ln \left[2 \cosh \left(\frac{\epsilon_F}{2k_B T} \right) \right], \quad (80a)$$

$$\sigma_{\text{inter}}(\omega) = \frac{e^2}{4\hbar} \left[H\left(\frac{1}{2}\hbar\omega\right) + \frac{4i}{\pi} \hbar(\omega + iy) \int_0^\infty \frac{H(\epsilon) - H\left(\frac{1}{2}\hbar\omega\right)}{\hbar^2(\omega + iy)^2 - 4\epsilon^2} d\epsilon \right], \quad (80b)$$

using a (regular) relaxation-time approximation for the decay rate γ [247], and where $H(\epsilon)$ denotes the occupation difference between energies $\mp\epsilon$ at temperature T

$$H(\epsilon) \equiv f(-\epsilon) - f(\epsilon) = \frac{\sinh(\epsilon/k_B T)}{\cosh(\epsilon_F/k_B T) + \cosh(\epsilon/k_B T)}. \quad (80c)$$

The low-temperature limit of Eqs. (80) is of particular interest since the Fermi temperatures $T_F \equiv \epsilon_F/k_B$ generally exceed room-temperature by orders of magnitude (e.g. $T_F \approx 4.6 \times 10^3$ K for $\epsilon_F = 0.4$ eV), rendering the limit excellently representative. The

The link ignores retardation effects since $\chi^0(q, \omega)$ is a nonretarded construct; as a consequence it does not distinguish between transverse and longitudinal perturbations.

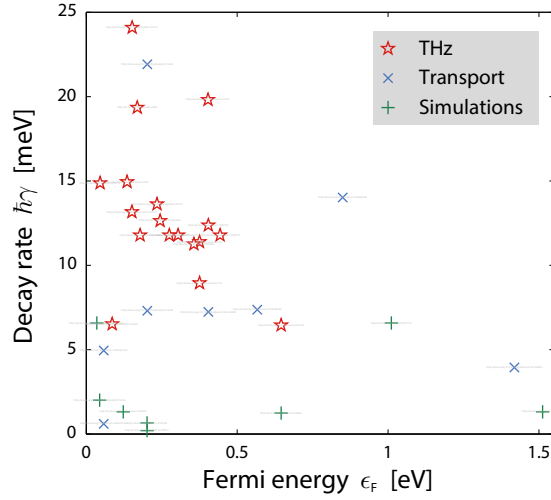


Figure 18: Measurements of the Drude decay rate γ compared with doping level, quantified by ϵ_F , as obtained in direct-current transport and optical THz measurements. The figure is reconstructed from the data collected by Tassin *et al.* [248]. A selection (mirroring Ref. 248) of values employed in simulations is also indicated, highlighting a slight mismatch between current capabilities and theoretical assumptions. In each marker is embedded a clickable link to the original publication.

limit can be obtained either from the $\epsilon_F/k_B T \rightarrow 0$ limit of Eqs. (80), or, as we will do later in Section 6.1, from the low-momentum limit of Eq. (77a), yielding

$$\sigma_{\text{intra}}(\omega) = \frac{ie^2\epsilon_F}{\pi\hbar^2(\omega + iy)}, \quad (81a)$$

$$\sigma_{\text{inter}}(\omega) = \frac{e^2}{4\hbar} \left[\theta(\hbar\omega - 2\epsilon_F) + \frac{1}{\pi} \ln \left| \frac{2\epsilon_F - \hbar\omega}{2\epsilon_F + \hbar\omega} \right| \right], \quad (81b)$$

The intraband term's small temperature dependence can be appreciated from the $x = T_F/T \ll 1$ expansion $\ln(2 \cosh \frac{1}{2}x) \simeq \frac{1}{2}x + e^{-x} + \dots$, revealing [when comparing with Eq. (80a)] that the finite-temperature corrections to Eq. (81a) are exponentially damped.

with Drude loss retained in just the intraband term. The low- and finite-temperature conductivities are compared in Figure 17. Generally, we observe that the intraband term is largely temperature-invariant, whereas the interband term exhibits appreciable dependence near the onset of vertical Landau damping at $\hbar\omega = 2\epsilon_F$, due to a temperature-induced lessening of the otherwise strict Pauli blocking. For low frequencies, i.e. for $\hbar\omega \lesssim \epsilon_F$, the loss is predominately of the Drude kind; this is the most important loss type for our considerations, because the plasmons in graphene reside in the region where $\text{Im } \sigma(\omega) > 0$, as we will see in Section 5.2.

The relation is obtained by comparing the dc-conductivity definition of the mobility, $\sigma = n_0 e \mu_e$, with the low-frequency intraband Drude conductivity $\sigma(\omega \rightarrow 0) = e^2 \epsilon_F / \pi \hbar^2 \gamma$.

DRUDE LOSS We end this chapter with a brief consideration of the *value* of the decay rate γ . Specifically, being of a strictly phenomenological character, its value must be estimated from experimental data. In Figure 18 we depict measured (and inferred) decay rates collected from a selection of experimental studies of graphene [248]. Two types of measurements are included, optical measurements in the THz regime and direct-current transport measurements of the electron mobility μ_e . The latter connects with a low-frequency estimate of the Drude decay rate via $\gamma = e v_F^2 / \epsilon_F \mu_e$. A significant spread in decay rates is evident, but at least one significant trend is discernible. Specifically, the transport measurements achieve a generally lower loss than the corresponding THz measurements. The latter find values lying predominantly in the range $0.1 \text{ eV} \lesssim \epsilon_F \lesssim 0.5 \text{ eV}$ and $11 \text{ meV} \lesssim \hbar\gamma \lesssim 15 \text{ meV}$. In this thesis we will generally choose values $\epsilon_F = 0.4 \text{ eV}$ and $\hbar\gamma = 12 \text{ meV}$ to reflect appropriately the current state-of-the-art. Nevertheless, a subset of the measurements indicate that much lower losses in

principle are achievable under optimal conditions (e.g. in suspended ultraclean samples). Occasionally, we will consider lower losses with this potential in mind.

Finally, we caution that a simple relaxation-time approximation cannot fully capture the manifold decay mechanisms relevant to graphene plasmonics. One important omission is hybridization with substrates, which impacts both the electronic and optical properties. In the optical respect, it facilitates e.g. the hybridization between graphene plasmons and phonon modes of a polar substrate [240, 249, 250]. Although this type of hybridization can be accounted for by an appropriate screening of the Coulomb interaction, we will in this thesis opt to focus strictly on the intrinsic properties of graphene's plasmons. In this endeavor we content ourselves with the limitations of a relaxation-time approximation.

4.3 SUMMARY

This chapter has served to provide a minimal introduction to graphene and to its electronic and optical properties in the low-energy domain. In [Section 4.1](#) we first briefly reviewed the TB approach to the energy dispersion of graphene's carriers, obtaining in the low-energy limit the famous linear Dirac dispersion. Subsequently, in [Section 4.2](#) we provided a summary of the associated low-energy optical response properties of graphene, specifically the density-density response function and the surface conductivity. Equipped with these tools, we proceed in the following chapters to explore the interesting topic of graphene plasmonics.

5

CLASSICAL GRAPHENE PLASMONICS

It is perhaps not surprising, given our considerations of metal plasmonics in local and nonlocal frameworks in [Chapters 2 and 3](#), that the main characteristics of graphene plasmonics find excellent descriptions in a classical framework. Of course, a classical treatment cannot account for all details; in [Chapter 6](#) we return to elements that lie beyond the classical realm, i.e. to nonclassical features. To appreciate these subtler elements, however, necessitates an established understanding of the classical predictions. It is the purpose of this chapter to introduce such an understanding.

The classical approach is in principle straightforward, combining just two components: one to account for the material response of graphene, i.e. linking the induced current and total field by the conductivity via $\mathbf{K} = \sigma \mathbf{E}_{\parallel}$, and one electrodynamic consideration to selfconsistently couple induced currents and total electric fields, i.e. application of Maxwell's (or, in the nonretarded limit, Coulomb's) equations jointly with the the first element. With these rather plain constituents the whole of classical graphene plasmonics can be assembled. Nevertheless, these ingredients, when joined under appropriate circumstances, far exceed the sum of their parts, just as in metal plasmonics. For instance, while extended undoped graphene exhibits an absorptance of $\approx 2.3\%$ [\[251\]](#), doped graphene may enhance this feat by orders of magnitude; an incontestably impressive feat for a one-atom thick material.

Though we in this chapter focus on graphene, many of our conclusions are readily transferable in the broader context of conductive two-dimensional interfaces. With this in mind, we open this chapter with a general consideration of the role of dimensionality in plasmonics. Subsequently, specializing in earnest to graphene, we examine the existence of plasmons in extended sheets of doped graphene. Finally, following an extensive discussion of plasmons in nanostructured graphene, we explore plasmons in curved graphene, concretely in graphene-coated nanospheres.

5.1 INFLUENCE OF DIMENSIONALITY

The arrival of atomic monolayers and their subsequent enthusiastic exploration for plasmonic applications has arguably established a conceptual juxtaposition, bordering sometimes on perceived equivalence, between atomic monolayers, graphene in particular, and two-dimensional (2D) plasmonics. Of course, graphene does not represent the historical roots of 2D plasmonics, nor is it the first platform investigated experimentally. Theoretically, the topic dates back at least to Stern's [1967](#) treatment of the 2D electron gas [\[252\]](#), with e.g. the classical treatment by Fetter [\[151\]](#) also deserving mention. The first experimental observation of 2D plasmons was achieved in surface states atop liquid helium by Grimes and Adams [\[253\]](#) in [1976](#), with several investigations following suit [\[33, 254\]](#), e.g. extending the observations to semiconductor inversion layers [\[255, 256\]](#), and to edge-localized 2D plasmons [\[257, 258\]](#).

Undoubtedly, the present-day focus on atomic monolayers is fully merited by their unique technological potential. Even so, to theoretically appreciate the particular fingerprints of graphene in 2D plasmonics vis-à-vis other material systems, it will aid us to establish which properties pertain to graphene's electronic attributes and which are simply a matter of dimensionality. In doing so, we will also have an opportunity to appreciate the differences between three-, two-, and even one-dimensional (3D, 2D, and

Graphene plasmonics, with very few exceptions, require nonzero doping – the premise of our investigation consequently relies on nonzero doping, i.e. $\epsilon_F \neq 0$.

For instance, in a recent review of two-dimensional nanophotonics [\[43\]](#), just two citations precede the [2004](#) discovery of graphene, both in reference to bulk properties.

d	$g_d(q)$	$\omega_d^2(q)/\omega_d^2$	$\omega_d(q)$ dispersion	size-dispersion
3D	$4\pi/q^2$	1	constant	constant
2D	$2\pi/q$	$q/2$	$\propto \sqrt{q}$	$\propto \sqrt{1/L}$
1D	$2K_0(qa)$	$q^2 K_0(qa)/2\pi$	$\propto q\sqrt{-\ln qa}$	$\propto L^{-1}\sqrt{\ln L/a}$

Table 1: Plasmon dispersion in a d -dimensional extended free-electron gas in the $q \ll k_{\beta_d}$ limit. Given for each dimensionality is the scaled Coulomb interaction $g_d(q)$, the normalized plasmon frequency $\omega_d^2(q)/\omega_d^2$, the corresponding low-momentum dispersion behavior of $\omega_d(q)$, and finally its analogue behavior in finite systems of characteristic extent L , obtained by the phenomenological replacement $q \sim 1/L$. The 1D case further utilizes the asymptotic relation $K_0(x) \simeq -\ln x$ applicable for $x \ll 1$.

See similar treatments
in Refs. 259 and 260.

1D) plasmonics. This is the purpose of the present section: deduce the low-momentum behavior of the plasmon dispersion in each dimension.

PLASMON DISPERSION Within the RPA the (nonretarded) plasmon dispersion of an extended d -dimensional system is deducible from just the Coulomb interaction $V_d(q)$ and the noninteracting density-density response $\chi_d^0(q, \omega)$, in momentum-frequency space (q, ω) . Specifically, the dispersion follows from the zeros of dielectric function, i.e. from $1 - V_d(q)\text{Re}\chi_d^0(q, \omega) = 0$. Though graphene exhibits a q -dependent form factor, cf. Eq. (76), we here specialize to the case of a free-electron gas with unity form factor and a single parabolic band; the end-result, apart from prefactors, remains applicable also to graphene. In this case, χ_d^0 can be expanded in the manner previously expounded in Eq. (47), such that (\mathcal{V}_d denoting a d -dimensional sample volume)

$$\chi_d^0(q, \omega) = \frac{2}{\mathcal{V}_d} \sum_{\mathbf{k} \in \mathbb{R}^d} \frac{f_{\mathbf{k}+\mathbf{q}} - f_{\mathbf{k}}}{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar(\omega + i\eta)} = \frac{q^2}{\omega^2} \frac{n_{0d}}{m_{\text{eff}}} + \frac{q^4}{\omega^4} \frac{n_{0d}\beta_d^2}{m_{\text{eff}}} + \mathcal{O}\left(\frac{q^6}{\omega^6}\right), \quad (82)$$

written in the low-loss $\eta \rightarrow 0^+$ limit, with equilibrium densities n_{0d} (per unit length, area, and volume in $d = 1, 2, 3$, respectively), and pressure velocities β_d^2 given by the high-frequency case of Eq. (42). Importantly, the form of $\chi_d^0(q, \omega)$ is dimensionally invariant, except for trivial dependencies through n_{0d} and β_d^2 . As a consequence, the plasmon dispersion $\omega_d(q)$ simplifies to

$$\omega_d^2 \simeq \frac{\omega_d^2 q^2 g_d(q)}{4\pi} \left(1 + \frac{q^2}{k_{\beta_d}^2}\right), \quad (83)$$

with $g_d(q)$ denoting the d -dimensional momentum-transform of the scaled Coulomb interaction $g(\mathbf{r}, \mathbf{r}') \equiv 1/|\mathbf{r} - \mathbf{r}'|$, hydrodynamic momentum $k_{\beta_d} \equiv \omega/\beta_d$, and $\omega_d^2 \equiv e^2 n_{0d}/\epsilon_0 m_{\text{eff}}$, which is simply the plasma frequency ω_p^2 in 3D – but not in 2D and 1D where its unit is $\text{m}^{3-d}/\text{s}^2$. Equation (83) highlights that dimensionality only affects the dispersion through $g_d(q)$.

The momentum transforms in three, two, and one dimensions can be carried out analytically, and the associated dispersion deduced: the results are summarized in Table 1. The 3D case reiterates the conclusions of Chapters 2 and 3: plasmonic excitations are momentum-, and size-independent in the nonretarded LRA limit. The picture is qualitatively modified in 2D, wherein a \sqrt{q} momentum-dependence, or equivalently, a $\sqrt{1/L}$ size-dependence, is exhibited. The 1D case is even further altered, requiring also the introduction of an additional length scale a . This scale is needed because the 1D momentum-transform of the Coulomb interaction is divergent. Accordingly, a small-scale cutoff is introduced on physical grounds by noting the necessarily nonzero transverse extent of any 1D system, instated in practice via the substitution

The 1D transform can
be performed using
 $\int_{-\infty}^{\infty} \frac{e^{-iqx}}{\sqrt{x^2+a^2}} dx = 2K_0(qa)$.
In polar coordinates the
2D transform requires
 $\int_0^{2\pi} e^{-ix\cos\theta} d\theta = 2\pi J_0(x)$
and $\int_0^{\infty} J_0(qr) dr = q^{-1}$.
Finally, by considering a
Yukawa potential $e^{-q_V r}/r$
in the $q_V \rightarrow 0^+$ limit,
the 3D transform can be
deduced in spherical
coordinates using
 $\int_0^{\pi} e^{-ix\cos\theta} \sin\theta d\theta = 2 \text{sinc } x$
and
 $\int_0^{\infty} e^{-q_V r} \sin qr dr = \frac{q}{q^2+q_V^2}$.

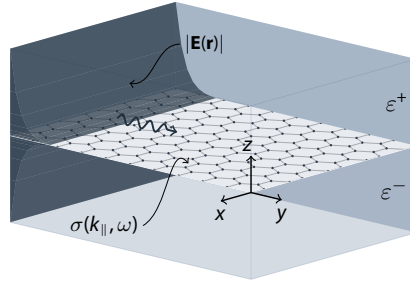
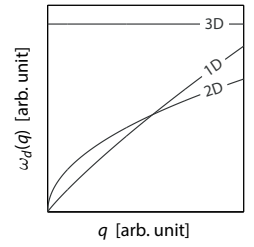


Figure 19: An extended graphene sheet, characterized by its conductivity $\sigma(k_{\parallel}, \omega)$, sandwiched between super- and substrates ϵ^{\pm} . A propagating graphene SPP is indicated by its spatial field-dependence.

$g(x) = |x|^{-1} \rightarrow (x^2 + a^2)^{-1/2}$. We will encounter a 1D plasmon only once in this thesis, namely the low-momentum monopole mode of a nanoribbon; in that case a finds a natural replacement in the ribbon width.

As we will derive rigorously in the following section, graphene also exhibits the square-root dependence found here for the parabolic free-electron gas. An interesting corollary of this conclusion is that 2D systems such as graphene, unlike 3D metallic structures, will exhibit beyond-LRA corrections $\propto q^{3/2}$ and $\propto 1/L^{3/2}$ rather than the $\propto 1/L$ corrections discussed in Section 3.4.2. Conversely, in the low-momentum LRA limit, the primary differences between graphene and parabolic 2D systems are deferred to the dispersion prefactor, i.e. to density-dependence.



Qualitative plasmon dispersion versus dimensionality.

5.2 SURFACE PLASMONS IN EXTENDED GRAPHENE

In this section, we will examine the plasmonic properties of a single extended graphene sheet. Though our subsequent discussion of graphene plasmonics in general structures, see Section 5.3, naturally encompasses several of the results obtained in the present section, we believe it worthwhile to elucidate here the fundamental features by a familiar technique, namely coupling of field-components by BCs. Having obtained the relevant governing formula, we investigate the plasmonic response of graphene in its local and nonlocal descriptions. Finally, we also touch upon the concept of TE plasmon polaritons in graphene [261].

5.2.1 Reflection coefficients in the presence of a surface current

As in our introduction to the classical SPP in Section 2.2.3 – and in the spirit of the pole condition of Eqs. (7) – we here choose to approach graphene plasmons (GPs) from the perspective of reflection coefficients. The setup in mind is indicated in Figure 19: a single graphene sheet lying in the $z = 0$ plane sandwiched between dielectrics ϵ_{\pm} . Given graphene's extreme thinness, on the order of the van der Waals diameter of carbon $h_g \approx 3.4 \text{ \AA}$ [262], relative to any optical quantity we expect to encounter, we model its presence simply as a 2D current-carrying sheet. The current is of the induced kind, induced by the perturbing *total* field $\mathbf{E}(\mathbf{r}, \omega)$. It is quantified by the nonlocal surface conductivity $\sigma(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; \omega)$ and the surface current $\mathbf{K}(\mathbf{r}_{\parallel}, \omega) = \int \sigma(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; \omega) \mathbf{E}_{\parallel}(\mathbf{r}'_{\parallel}, \omega) d^2 \mathbf{r}'_{\parallel}$, with \parallel -subscripted variables indicating in-plane quantities oriented along $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ and implicitly evaluated in $z = 0$. The surface current modifies the conventional two-

The surface-current approach is analogous to the dipole approximation familiar from field-emitter interactions; its validity is guaranteed by the complete fulfilment of the condition $h_g/\lambda \ll 1$ for wavelengths λ up to the ultraviolet.

interface BCs: by applying the standard pillbox arguments while accounting for an interfacial current $\mathbf{K}(\mathbf{r}_\parallel, \omega)\delta(z)$ the BCs generalize to [55]

$$\hat{\mathbf{n}} \times (\mathbf{E}^+ - \mathbf{E}^-) = \mathbf{0}, \quad \hat{\mathbf{n}} \times (\mathbf{H}^+ - \mathbf{H}^-) = \mathbf{K}, \quad (\text{tangential}) \quad (84a)$$

$$\hat{\mathbf{n}} \cdot (\mathbf{D}^+ - \mathbf{D}^-) = 0, \quad \hat{\mathbf{n}} \cdot (\mathbf{H}^+ - \mathbf{H}^-) = \rho_\parallel, \quad (\text{normal}) \quad (84b)$$

with normal-vector $\hat{\mathbf{n}} = \hat{\mathbf{z}}$, short-hand notation $\mathbf{E}^\pm \equiv \mathbf{E}(\mathbf{r}_\parallel + \hat{\mathbf{z}}0^\pm, \omega)$ indicating evaluation at an infinitesimal distance above or below the interface, and surface charge density ρ_\parallel related to \mathbf{K} by the continuity equation $\nabla \cdot \mathbf{K} + \partial_t \rho_\parallel = 0$, both implicitly evaluated at \mathbf{r}_\parallel .

With the modified BCs in hand, we desire to determine the reflection coefficients, r_{TE} and r_{TM} , for TE and TM plane waves, respectively, oscillating at frequency ω impinging from the $z > 0$ region with in-plane momentum component k_\parallel . Their deduction follows by the same analysis as the conventional Fresnel coefficients, Eq. (16), albeit with adjusted BCs, (84); after some algebra the results are [39, 263]

$$r_{\text{TM}} = \frac{\varepsilon^- k_\perp^+ - \varepsilon^+ k_\perp^- + \sigma k_\perp^- k_\perp^+ / \varepsilon_0 \omega}{\varepsilon^- k_\perp^+ + \varepsilon^+ k_\perp^- + \sigma k_\perp^- k_\perp^+ / \varepsilon_0 \omega}, \quad (85a)$$

$$r_{\text{TE}} = \frac{k_\perp^+ - k_\perp^- - \mu_0 \omega \sigma}{k_\perp^+ + k_\perp^- + \mu_0 \omega \sigma}, \quad (85b)$$

with (k_\parallel, ω) -dependence implicitly understood for σ (and in principle also for ε^\pm), and with normal momentum components $(k_\perp^\pm)^2 = \varepsilon^\pm k_0^2 - k_\parallel^2$ as in Eq. (16). With the reflection coefficients established we are now in a position to investigate the possible plasmonic excitations. The TM case presents by far the most compelling plasmonic features and is discussed in the following section. We return to the TE case in Section 5.2.3.

5.2.2 TM graphene plasmons

For $k_\parallel > \varepsilon^\pm k_0$ the normal component k_\perp^\pm turns imaginary, indicating the necessarily bound nature of excitations in this regime. Specifically, for propagation along the x -direction a field profile $\propto \exp(ik_\parallel x \mp \kappa^\pm z)$ is exhibited, with $\kappa^\pm = ik_\perp^\pm$, as indicated schematically in Figure 19. The TM SPP associated with this profile follows directly from the poles of Eq. (85a), yielding the dispersion equation [37, 39]

$$\frac{\varepsilon^-}{\sqrt{k_\parallel^2 - \varepsilon^- k_0^2}} + \frac{\varepsilon^+}{\sqrt{k_\parallel^2 - \varepsilon^+ k_0^2}} = \frac{\sigma(k_\parallel, \omega)}{i\varepsilon_0 \omega}. \quad (86)$$

This dispersion equation fully accounts for retardation; hence the addendum of polariton to its naming.

NONRETARDED LIMIT As it turns out, the polaritonic aspects of Eq. (86) are of very little significance in nearly all relevant scenarios. Accordingly, a nonretarded approach is sufficient, derivable immediately as the $k_0/k_\parallel \rightarrow 0$ limit of Eq. (86), allowing a much simpler dispersion equation

$$\omega = \frac{\sigma(k_\parallel, \omega)}{2i\varepsilon_0 \bar{\varepsilon}} k_\parallel, \quad (87)$$

where the introduction of an effective dielectric cladding $\bar{\varepsilon} \equiv (\varepsilon^+ + \varepsilon^-)/2$ reflects the fact that the sub- and superstrates are on equal footing in the nonretarded limit, since the field profile simplifies to $\propto \exp(ik_\parallel x \mp k_\parallel z)$ there.

More generally the effective dielectric cladding imparts a rescaling of the in-plane Coulomb interaction $V(\mathbf{r}_\parallel, \mathbf{r}'_\parallel) \rightarrow \bar{\varepsilon}^{-1} V(\mathbf{r}_\parallel, \mathbf{r}'_\parallel)$ in the nonretarded limit. This rescaling is rigorous; we revisit it in Section 5.3.

It is worthwhile to examine the origin of the extraordinary quality of the nonretarded approach in graphene by an explicit calculation. At low frequencies and momenta the conductivity can be approximated by its lossless local intraband approximation $\sigma(\omega) \simeq ie^2\epsilon_F/\pi\hbar^2\omega$, in which case the full retarded dispersion, Eq. (86), can be solved exactly in the special case $\epsilon^\pm = \bar{\epsilon}$, allowing

$$k_{\parallel} = k_0 \sqrt{\left(\frac{\bar{\epsilon} \hbar \omega}{2 \alpha \epsilon_F}\right)^2 + \bar{\epsilon}} \simeq k_{\text{GP}}^{\text{intra}} + \alpha \frac{\epsilon_F}{\hbar c}, \quad (88)$$

expressed in terms of the nonretarded intraband GP momentum

$$k_{\text{GP}}^{\text{intra}} = \frac{2\pi\epsilon_0\bar{\epsilon}(\hbar\omega)^2}{e^2\epsilon_F}, \quad (89)$$

and where the right-hand side of Eq. (88) follows upon expansion around $\alpha\epsilon_F/\hbar\omega \ll 1$. Since we will generally be concerned with excitation energies $\hbar\omega$ comparable to the Fermi energy ϵ_F and since the fine-structure constant $\alpha \approx 1/137$ is small, the expansion is excellently approximated by its zeroth order term. In other words, the nonretarded limit is superb so long that $\alpha\epsilon_F/\hbar\omega \ll 1$, with retardation-corrections appearing only at second order in this very small parameter. Accordingly, it is in practice fully sufficient to approach graphene plasmonics entirely from the nonretarded limit.

DISPERSION PROPERTIES Equation (89) provides the nonretarded intraband result. Agreement with the dimensional considerations of Section 5.1 is readily observed by shifting the perspective to a momentum-dependent resonance frequency

$$\omega_{\text{GP}}^{\text{intra}} = \frac{1}{\hbar} \sqrt{\frac{e^2\epsilon_F}{2\pi\epsilon_0\bar{\epsilon}}} k_{\parallel}. \quad (90)$$

It is instructive to compare this result with that of the 2D parabolic gas, in which the resonance $\omega_{2\text{D}} \propto (n_0 k_{\parallel})^{1/2}$ exhibits a square-root scaling not only with momentum but also with the 2D equilibrium density, cf. Table 1. By contrast, GPs exhibit a fourth-root scaling $\omega_{\text{GP}}^{\text{intra}} \propto n_0^{1/4} k_{\parallel}^{1/2}$, since $\epsilon_F \propto n_0^{1/2}$. The distinct cases can be at least mnemonically joined by the introduction of a density-dependent graphene Drude mass $m_{\text{eff}}^{\text{g}} = \hbar k_F/v_F \propto n_0^{1/2}$ (unrelated with the effective band mass) allowing a pragmatic approach to the transfer of parabolic Drude considerations to graphene by simple substitution $m_{\text{eff}} \rightarrow m_{\text{eff}}^{\text{g}}$ [264].

Beyond the intraband model, the plasmon dispersion must be solved numerically. Figure 20 presents the result of such a numerical study, performed in the zero-loss zero-temperature limits of the local intraband, intra- and interband, and full nonlocal response functions. We observe that the intraband approximation is in excellent agreement with the latter two tiers of description for energies below ϵ_F . Above the Fermi energy, interband dispersion introduces considerable redshifting relative to the intraband predictions. The “redness” of the shift is a simple consequence of the interband conductivity exhibiting negative imaginary part for all energies. On this note, we comment that physical solutions to Eq. (87) exist only if $\text{Im } \sigma > 0$: while this condition is true for all energies in the intraband model, the inclusion of interband dispersion restricts the region of positivity to $\hbar\omega/\epsilon_F \lesssim 1.6671$. This explains the apparent k_{\parallel} -divergence of the full local model at this energy. Finally, we observe that the full nonlocal treatment blueshifts slightly relative to the local predictions, a feature qualitatively appreciable from hydrodynamic considerations, and derivable by series expansion of $\sigma(k_{\parallel}, \omega)$ in k_{\parallel}/k_F [265]. Eventually, the nonlocal dispersion departs considerably from its local counterparts, with a qualitative transition emerging after the plasmon branch enters the region of interband Landau damping. In this region, the plasmon decays via generation of EHPs across the Dirac point and is very rapidly

The opposite limit, i.e. $\hbar\omega \ll \alpha\epsilon_F$, though of little technological or plasmonic importance is noteworthy at least for the sake of completeness. There, as was the case also for metal SPPs, the graphene SPP eventually exhibits predominately polaritonic properties, with dispersion $k_{\parallel} \simeq k_0\sqrt{\bar{\epsilon}} + \mathcal{O}[(\hbar\omega/\alpha\epsilon_F)^2]$. In other words, the dispersion is ultimately linear for $\omega \rightarrow 0$.

For nonvanishing but small loss, the plasmon frequency $\omega_{\text{GP}}^{\text{intra}}$ acquires a finite imaginary part $\simeq i\gamma/2$.

In practice, the dispersion is obtained by numerically minimizing $|1 - V(k_{\parallel})\text{Re}\chi^0(k_{\parallel}, \omega)|$ over real, positive $\{k_{\parallel}, \omega\}$.

To include loss it is necessary to solve $1 - V(k_{\parallel})\chi^0(k_{\parallel}, \omega) = 0$ in $\{k_{\parallel}, \omega\}$ with either k_{\parallel} or ω complex, rather than $1 - V(k_{\parallel})\text{Re}\chi^0(k_{\parallel}, \omega) = 0$ in real variables. Beyond a perturbative approach [37], which is inapplicable in Landau regions, this has not been achieved in nonlocal treatments. Arguably, a quantitative resolution of this issue is of modest practical worth: plasmon properties are certainly poor in these regions.

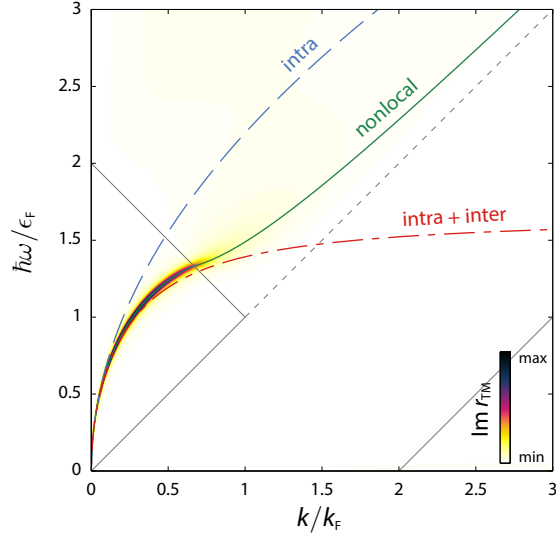


Figure 20: Dispersion of GPs in extended graphene in vacuum, calculated in the zero-loss zero-temperature limits. Three models are compared: an analytical local intraband description, Eq. (90), in dashed blue; the full local conductivity, accounting for both intra- and interband dispersion, Eqs. (81), in dash-dotted red; and finally the full nonlocal conductivity obtained from $\chi^0(k_{\parallel}, \omega)$, Eq. (77a), in green. The imaginary part of the TM Fresnel reflection coefficient, Eq. (85a), is superposed, calculated from the nonlocal description with Fermi energy $\epsilon_F = 0.4$ eV and finite Drude loss $\hbar\gamma = 12$ meV included through the Mermin-corrected relaxation-time approximation, Eq. (78). The regions of Landau damping are delimited by gray lines, with a dotted line separating the inter- and intraband regions.

damped. The feature can be appreciated also by the superposed colormap of r_{TM} in Figure 20, with loss included: the plasmon peak structure is absent in the Landau regions. Accordingly, it is questionable what physical content, if any, the calculated zero-loss dispersion retains in regions of Landau damping.

EXCITATION BY EELS As discussed at the outset of this section, GPs carry a momentum dramatically larger than the free-space wave vector k_0 ; in fact, k_0 is practically congruent with the vertical axis in Figure 20. Plasmons in extended graphene are consequently challenging to excite by optical means. Several approaches can circumvent this issue, e.g. by nanostructuring of graphene itself, which we return to in Section 5.3.2, or by the introduction of grating substrates [266]. EELS suggests an alternative excitation approach, which is appealing because it does not require structural modifications. The “missing” momentum is in EELS supplied directly by the traveling electron, see Section 2.3.2. Indeed, as we demonstrate in Figure 21, and as discussed previously by García de Abajo [267], even normally incident electrons can excite GPs. Moreover, as we derive in Appendix A and as evident from Figure 21, specific plasmon frequencies and momenta can be preferentially excited by tuning the electron’s total energy, or equivalently, its velocity v . Concretely, the EEL probability is peaked for matching electron and plasmon wave vectors: $\omega/v = k_{\parallel}$. For the intraband model (in which the low-loss EEL probability is analytically solvable, see Eq. (135), and where $k_{\parallel} = k_{\text{GP}}^{\text{intra}}$) this incurs v -dependent EELS peaks at $\hbar\omega_{\text{GP}}^{\text{intra}} = 2\alpha \frac{c}{v} \epsilon_F$. This condition highlights that GPs require rather slow electron velocities: for instance, to excite GPs with energies larger than, say, 10% of ϵ_F requires $v \lesssim 0.15c$ (equivalently, electron energies $\lesssim 5.5$ keV), contrasting the relativistic electrons, measured in the hundreds of keV, customarily employed in metal-plasmonic EELS.

See Appendix A for a treatment of the interaction between a traveling electron and graphene (normal incidence).

We note that the application of the linear EELS framework, Eq. (26), to such low acceleration energies carries an undeniable degree of unease: as the energy-loss eventually constitutes a sizable fraction of the total electron energy, the electron is deflected from its straight path, inducing further complication of a self-consistent kind that is not captured by Eq. (26).

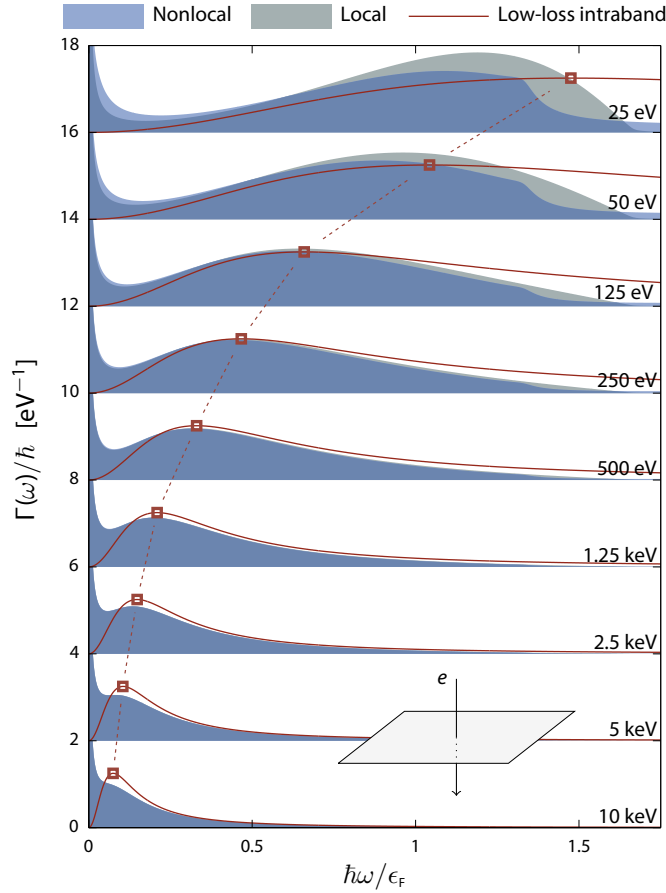


Figure 21: EEL probability $\Gamma(\omega)$ for an electron traversing an extended graphene sheet normally, as illustrated in inset, calculated from Eq. (134). A zero-temperature graphene sheet in vacuum is considered, doped to $\epsilon_F = 0.4$ eV and with Drude loss $\hbar\gamma = 0.4$ eV. Nonlocal, Eq. (77a), and local descriptions, Eqs. (81), are compared with a zero-loss intraband approximation, Eq. (135), in blue, gray, and red, respectively. Several electron energies are considered, with each spectrum offset by two units, corresponding to varying electron velocities. Red markers and dashed guide-lines indicate the intraband EEL peak condition $\omega/v = k_{\text{GP}}^{\text{intra}}(\omega)$, with associated energies $\hbar\omega_{\text{GP}}^{\text{intra}} = 2\alpha \frac{c}{v} \epsilon_F$.

Inspecting again Figure 21, now for the very slow electrons with energies below 50 eV (corresponding to $v \approx 0.014c \approx 4.6v_F$), we note that the effective GP momenta $\sim \omega/v$ sampled by the electron are increasingly large. Accordingly, the quality of the intraband approximation decreases at these velocities; eventually, discrepancies also develop between the full local and nonlocal treatments. These discrepancies are primarily due to nonvertical Landau transitions, which are included in the latter treatment, but not the former. Notably, a sharp “shoulder” develops in the nonlocal calculation, whose energy-loss matches the entry-point (located around $\hbar\omega/\epsilon_F \approx 1.33$) of the plasmon branch into the Landau region in Figure 20. This observation suggests an experimental route to study the limitations of the local description of graphene plasmonics through EELS with low-energy electrons. We note that the application of EELS in this energy regime is feasible [267]: electron energies below 100 eV are experimentally utilized in the field of low-energy EELS [268]. In fact, it has seen application to graphene already in angle-resolved studies employing 20 eV electrons [269].

5.2.3 TE graphene plasmon polaritons

The existence of a TE graphene plasmon polariton (GPP) represents something of a curiosity, given its nonexistence in metal-plasmonics. Nevertheless, it is straightforward to establish from r_{TE} , Eq. (85b), that a TE GPP could in principle exist with momentum $k_{\parallel}^{\text{TE}} > k_0$ provided that $\text{Im}[\sigma(k_{\parallel}^{\text{TE}}, \omega)] < 0$. In the special case of identical sub- and superstrates, i.e. $\varepsilon^{\pm} = \varepsilon$, a dispersion equation can be obtained by squaring [261]

$$k_{\parallel}^{\text{TE}} = k_0 \sqrt{\varepsilon - \left[\frac{\sigma(k_{\parallel}^{\text{TE}}, \omega)}{2\varepsilon_0 c} \right]^2}. \quad (91)$$

The TE GPP is also limited from above to energies $\hbar\omega/\epsilon_F < 2$ due to the onset of vertical Landau damping.

In a local description, including intra- and interband terms, the condition $\text{Im}[\sigma(\omega)] < 0$ is fulfilled for $\hbar\omega/\epsilon_F \gtrsim 1.6671$, see Figure 17, and TE GPPs may then in principle exist there: presenting us with an apparently remarkable flaunting of conventional wisdom. Partly motivated by the astounding interest this new mode has attracted following its prediction [261], we will use this opportunity to introduce a few tempering remarks.

Firstly, we note that the mode actually resembles an ordinary polariton much more than a plasmon. This is apparent if Eq. (91) is rewritten in terms of the characteristic conductivity $\sigma_0 \equiv e^2/\pi\hbar$, allowing

$$k_{\parallel}^{\text{TE}} = k_0 \sqrt{\varepsilon - \left[2\alpha \frac{\sigma(k_{\parallel}^{\text{TE}}, \omega)}{\sigma_0} \right]^2} \simeq k_0 \sqrt{\varepsilon} + \mathcal{O}(\alpha^2), \quad (92)$$

with the expansion following upon noting the smallness of $\alpha\sigma/\sigma_0$, given that $\sigma/\sigma_0 \sim 1$ while $\alpha \ll 1$. Equation (92) emphasizes that the TE SPP is only very slightly below the light line: accordingly, none of the usual plasmonic attributes transfer, e.g. its wavelength is only *very* slightly below the free-space value, and its lateral extension dramatically larger than its TM counterpart, extending hundreds of wavelengths away from the graphene sheet.

Secondly, we discuss the assumption of equal dielectric surroundings. Specifically, suppose the opposite, e.g. a vacuum superstrate and a dielectric substrate, such that $\varepsilon^+ = 1$ and $\varepsilon^- = \varepsilon$. The squaring procedure leading to Eq. (91) can again be applied to this case – however, the procedure must be explicitly checked by insertion of solutions into its non-squared counterpart, since the squaring procedure has the potential to

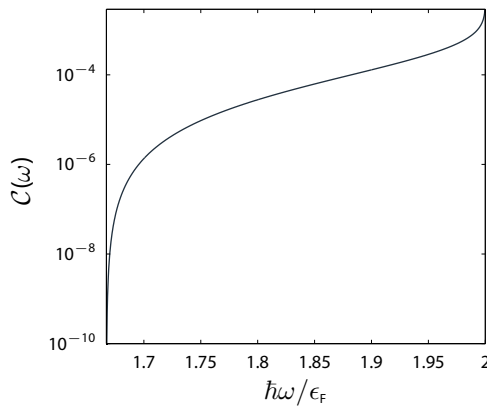


Figure 22: Energy-dependence of the condition function $\mathcal{C}(\omega)$, see Eq. (93), yielding the maximum allowable dielectric difference $\varepsilon - 1$ across a graphene sheet that can support TE GPPs. Here shown for the lossless zero-temperature local conductivity, Eqs. (81).

introduce fictitious solutions. Appendix B details this procedure, the end result of which is that TE GPPs exist *only* if the following condition holds true (ignoring loss)

$$\varepsilon - 1 < \left\{ 4\alpha \frac{\text{Im}[\sigma(\omega)]}{\sigma_0} \right\}^2 \equiv \mathcal{C}(\omega). \quad (93)$$

Figure 22 depicts the right-hand condition function, defined here as $\mathcal{C}(\omega)$: it indicates the maximum allowable difference in dielectric function of sub- and superstrate. Evidently, the allowable deviation in dielectric values is minuscule, excluding in practice any real substrate. Even for free-standing graphene in air the condition is a considerable worry; even minute pressure differences across the graphene sheet would effectively exclude the TE GPP solution. For these reasons, we suggest that the TE GPP, though at face-value a striking and oft-cited member of the graphene plasmonics zoo, has little or no physical value, especially considering that material loss is likely to only worsen the outlook.

It was recently suggested, perhaps optimistically, that this restriction on $\varepsilon - 1$ might be turned to functionality in the context of ultra-sensitive sensing of dielectric environments [270].

5.3 PLASMONS IN NANOSTRUCTURED GRAPHENE

The understanding of extended TM GPs of an infinite graphene sheet provide the conceptual underpinnings of graphene plasmonics. Nevertheless, their local counterparts in finite graphene samples are arguably of greater technological and practical interest, because of their ability to couple with propagating plane waves (i.e. for the same reasons that LSPs feature so prominently in metal plasmonics). In this section, we outline the formal framework for the understanding of GPs in graphene nanostructures within a local-conductivity description. Following a brief discussion of particulars related to the treatment of edges, we proceed to the concrete examination of various archetypal geometries, obtaining along the way an intuitive appreciation of the main features of localized GPs.

5.3.1 Nonretarded formulation

For a general 2D structure, positioned in the $z = 0$ plane and delimited in the xy -plane by a region $\mathbf{r}_{\parallel} \in \Omega$ with conductivity $\sigma(\mathbf{r}_{\parallel}, \omega)$, the nonretarded response can be deduced by self-consistently coupling the Coulomb law with the current-field relationship and the continuity equation. Specifically, considering a coordinate $\mathbf{r}_{\parallel} \in \Omega$, with associated induced surface charge density $\rho_{\parallel}(\mathbf{r}_{\parallel})$ and in-plane (total) potential $\phi(\mathbf{r}_{\parallel})$, the above noted elements provide the coupled equations

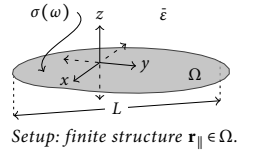
$$\phi(\mathbf{r}_{\parallel}) = \int_{\Omega} \frac{\rho_{\parallel}(\mathbf{r}'_{\parallel})}{4\pi\varepsilon_0\bar{\varepsilon}|\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}|} d^2\mathbf{r}'_{\parallel}, \quad (94a)$$

$$\rho_{\parallel}(\mathbf{r}_{\parallel}) = i\omega^{-1}\nabla_{\parallel} \cdot [\sigma(\mathbf{r}_{\parallel}, \omega)\nabla_{\parallel}\phi(\mathbf{r}_{\parallel})], \quad (94b)$$

with screening from sub- and superstrates accounted for via $\bar{\varepsilon}$ and with a 2D nabla operator $\nabla_{\parallel} \equiv \hat{\mathbf{x}}\partial_x + \hat{\mathbf{y}}\partial_y$. By combining Eqs. (94) one may evidently express either $\rho_{\parallel}(\mathbf{r}_{\parallel})$ or $\phi(\mathbf{r}_{\parallel})$ by an integro-differential equation. Choosing the potential, one finds an eigensystem with eigenpairs $\{\zeta_v, \phi_v\}$ defined by

$$\zeta_v\phi_v(\tilde{\mathbf{r}}_{\parallel}) = \frac{-1}{2\pi} \int_{\Omega} \frac{\tilde{\nabla}'_{\parallel} \cdot [f(\tilde{\mathbf{r}}'_{\parallel})\tilde{\nabla}'_{\parallel}\phi_v(\tilde{\mathbf{r}}'_{\parallel})]}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}'_{\parallel}|} d^2\tilde{\mathbf{r}}'_{\parallel}, \quad (95a)$$

$$\zeta_v \equiv \frac{2i\varepsilon_0\bar{\varepsilon}\omega_v L}{\sigma(\omega_v)}, \quad (95b)$$



The integration domain Ω is indicated in Eq. (95a) for mnemonic reasons only: in principle the integral extends over all of $\tilde{\mathbf{r}}'_{\parallel} \in \mathbb{R}^2$, but is limited in practice by the extent of $f(\tilde{\mathbf{r}}'_{\parallel})$, assumed bounded by Ω .

written in scale-independent form by the introduction of dimensionless coordinates $\tilde{\mathbf{r}}_{\parallel} = L^{-1}\mathbf{r}_{\parallel}$, $\tilde{\nabla}_{\parallel} = L\nabla_{\parallel}$, and $\tilde{\Omega} = L^{-1}\Omega$, all normalized to a characteristic length L . Furthermore, the conductivity has been separated into its frequency dependence, $\sigma(\omega)$, and its spatial dependence, $f(\mathbf{r}_{\parallel})$, such that $\sigma(\mathbf{r}_{\parallel}, \omega) \equiv \sigma(\omega)f(\mathbf{r}_{\parallel})$. In the special but very relevant case of uniform conductivity, the latter describes the geometric extent of Ω , expressible by the indicator function $f(\mathbf{r}_{\parallel}) = \mathbb{1}_{\Omega}(\mathbf{r}_{\parallel}) = \begin{cases} 1 & \text{if } \mathbf{r}_{\parallel} \in \Omega \\ 0 & \text{if } \mathbf{r}_{\parallel} \notin \Omega \end{cases}$. This particular approximation will occupy our full attention for the remainder of the chapter, with its qualities expected to mirror those of the uniform bulk LRA description applied in metal plasmonics.

The eigenvalues ζ_{ν} constitute the central object of interest, because they dictate the dispersion of the eigenfrequencies ω_{ν} through Eq. (95b) by means of the spatial scale L and the material response $\sigma(\omega)$. For instance, within a low-loss $\gamma \ll \omega$ intraband treatment, the plasmon dispersion is readily established as

$$\omega_{\nu}^{\text{intra}} \simeq \frac{1}{\hbar} \sqrt{\frac{e^2 \epsilon_{\text{F}}}{2\pi \epsilon_0 \tilde{\epsilon}} \frac{\zeta_{\nu}}{L} - \frac{i}{2} \gamma}, \quad (96)$$

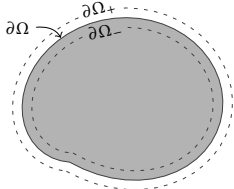
bearing an unmistakable resemblance with the extended-sheet result of Eq. (90) and confirming the expected $\propto \sqrt{1/L}$ argued for initially in Section 5.1.

We return to the concrete evaluation of ζ_{ν} for various geometries in Section 5.3.2. An evaluation of Eq. (95a), however, requires a definite treatment of the boundary $\partial\Omega$, made in practice by the inclusion of an edge charge contribution or by the application of a BC. Specifically, since the integration domain in Eq. (95a) extends an infinitesimal distance beyond Ω , the numerator of Eq. (95a) contributes a surficial Dirac delta term from the gradient of the indicator function, which in turn can be interpreted as a 1D (fictitious) edge charge density. If instead the integration domain is reduced by an infinitesimal amount, the singular contribution is sidestepped altogether – but the equation must then be supplemented by an appropriate BC, being in this case the vanishing of the normal current. Either approach has seen widespread, though disconnected use in the literature. The disconnect, perhaps, is reasonable given the qualitatively different perspectives, appearing to some extent diametrical in nature. Nevertheless, it is a simple matter to ascertain their equivalence by explicit calculation. To do so, we first define the circumstances of each approach: we take for the domain Ω with boundary $\partial\Omega$ the outward normal vector as $\hat{\mathbf{n}}$, and define their infinitesimal enlargement (+) and reduction (–) along $\hat{\mathbf{n}}$ as Ω_{\pm} and $\partial\Omega_{\pm}$, as pertinent to the edge charge and BC perspectives, respectively. With these definitions established, the equivalence follows by direct calculation:

The gradient of the indicator function is the analogue of the derivative of the Heaviside step function, in the sense that:

$$\int_{\Omega} [\nabla \mathbb{1}_{\Omega}(\mathbf{r})] \cdot \mathbf{g}(\mathbf{r}) \, d^n \mathbf{r} = - \oint_{\partial\Omega} \hat{\mathbf{n}} \cdot \mathbf{g}(\mathbf{r}) \, d^{n-1} \mathbf{r}$$

for $\Omega \in \mathbb{R}^n$ and outward normal vector $\hat{\mathbf{n}}$.



Domains Ω , Ω_- , and Ω_+ indicated by their associated boundaries.

EDGE CHARGE CONTRIBUTION Expanding the numerator of Eq. (95a) yields two terms $\mathbb{1}_{\Omega}(\tilde{\mathbf{r}}'_{\parallel}) \tilde{\nabla}_{\parallel}^{\prime 2} \phi_{\nu}(\tilde{\mathbf{r}}'_{\parallel})$ and $[\tilde{\nabla}'_{\parallel} \mathbb{1}_{\Omega}(\tilde{\mathbf{r}}'_{\parallel})] \cdot \tilde{\nabla}'_{\parallel} \phi_{\nu}(\tilde{\mathbf{r}}'_{\parallel})$, the latter of which contain a surficial Dirac delta function which evaluates to a line integral along $\partial\Omega$ since Ω_+ includes the boundary fully. Accordingly, Eq. (95a) expands into (omitting eigenindex ν for brevity)

$$\zeta \phi(\tilde{\mathbf{r}}_{\parallel}) = \frac{-1}{2\pi} \left\{ \int_{\tilde{\Omega}} \frac{\tilde{\nabla}_{\parallel}^{\prime 2} \phi(\tilde{\mathbf{r}}'_{\parallel})}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}'_{\parallel}|} \, d^2 \tilde{\mathbf{r}}'_{\parallel} - \oint_{\partial\tilde{\Omega}} \frac{\hat{\mathbf{n}}' \cdot \tilde{\nabla}'_{\parallel} \phi(\tilde{\mathbf{r}}'_{\parallel})}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}'_{\parallel}|} \, d\tilde{\mathbf{r}}'_{\parallel} \right\}. \quad (97)$$

This expression can be simplified by rewriting the surface integral term through initial use of the scalar-vector product rule $f \nabla \cdot \mathbf{g} = \nabla \cdot (f \mathbf{g}) - (\nabla f) \cdot \mathbf{g}$ and subsequent use of the divergence theorem, such that

$$\int_{\tilde{\Omega}} \frac{\tilde{\nabla}_{\parallel}^{\prime 2} \phi(\tilde{\mathbf{r}}'_{\parallel})}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}'_{\parallel}|} \, d^2 \tilde{\mathbf{r}}'_{\parallel} = \oint_{\partial\tilde{\Omega}} \frac{\hat{\mathbf{n}}' \cdot \tilde{\nabla}'_{\parallel} \phi(\tilde{\mathbf{r}}'_{\parallel})}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}'_{\parallel}|} \, d\tilde{\mathbf{r}}'_{\parallel} - \int_{\tilde{\Omega}} \left(\tilde{\nabla}'_{\parallel} \frac{1}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}'_{\parallel}|} \right) \cdot \tilde{\nabla}'_{\parallel} \phi(\tilde{\mathbf{r}}'_{\parallel}) \, d^2 \tilde{\mathbf{r}}'_{\parallel}. \quad (98)$$

The surface integral over Ω_+ reduces to an integral over Ω since $\mathbb{1}_{\Omega}(\mathbf{r}_{\parallel})$ vanishes for $\mathbf{r}_{\parallel} \notin \Omega$, i.e. there is no contribution from the small annulus $\Omega_+ \setminus \Omega$.

Evidently, the line integrals of Eqs. (97) and (98) cancel. As a consequence, Eq. (97) can be re-expressed in a form without edge singularities

$$\zeta\phi(\tilde{\mathbf{r}}_{\parallel}) = \frac{1}{2\pi} \int_{\tilde{\Omega}} \left(\tilde{\nabla}_{\parallel}' \frac{1}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}_{\parallel}'|} \right) \cdot \tilde{\nabla}_{\parallel}' \phi(\tilde{\mathbf{r}}_{\parallel}') d^2\tilde{\mathbf{r}}_{\parallel}'. \quad (99)$$

BOUNDARY CONDITION The BC of a vanishing normal current, i.e. $\hat{\mathbf{n}} \cdot \mathbf{K}(\mathbf{r}_{\parallel}) = 0$ for $\mathbf{r}_{\parallel} \in \partial\Omega_-$, is readily transferred to the potential in local models where $\mathbf{K}(\mathbf{r}_{\parallel}) = -\sigma(\omega)\nabla_{\parallel}\phi(\mathbf{r}_{\parallel})$, such that $\hat{\mathbf{n}} \cdot \nabla_{\parallel}\phi(\mathbf{r}_{\parallel}) = 0$ along $\mathbf{r}_{\parallel} \in \partial\Omega_-$. Moreover, since Ω_- excludes the boundary region $\partial\Omega$ the surficial Dirac delta term is never zeroed, and so does contribute. Accordingly, for the reduced integration domain Ω_- Eq. (95a) simplifies to

$$\zeta\phi(\tilde{\mathbf{r}}_{\parallel}) = \frac{-1}{2\pi} \int_{\tilde{\Omega}_-} \frac{\tilde{\nabla}_{\parallel}'^2 \phi(\tilde{\mathbf{r}}_{\parallel}')}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}_{\parallel}'|} d^2\tilde{\mathbf{r}}_{\parallel}'. \quad (100)$$

To make explicit use of the BC, the remaining term can be expanded by the same scheme that led to Eq. (98) – albeit with integration domains $\partial\Omega_-$ and Ω_- – in which case the BC forces the line integral term to vanish, revealing

$$\zeta\phi(\tilde{\mathbf{r}}_{\parallel}) = \frac{1}{2\pi} \int_{\tilde{\Omega}_-} \left(\tilde{\nabla}_{\parallel}' \frac{1}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}_{\parallel}'|} \right) \cdot \tilde{\nabla}_{\parallel}' \phi(\tilde{\mathbf{r}}_{\parallel}') d^2\tilde{\mathbf{r}}_{\parallel}'. \quad (101)$$

At this point it is observed that the end results of edge charge and BC approaches, Eqs. (99) and (101), differ only by an integral over an infinitesimal annulus $\Omega \setminus \Omega_-$ whose contribution vanishes as $\Omega_- \rightarrow \Omega$: accordingly, by the absence of any remaining terms with singularities specifically in the edge region, we conclude that the two approaches are equivalent.

5.3.2 Role of geometry

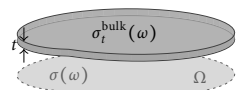
The primary value of the particular formulation of Eqs. (95) is arguably not the detailed form of integro-differential equation itself, but rather the conceptual understanding that the eigenfrequencies ω_v of a given setup depend only on three independent variables: namely, the eigenvalues ζ_v which are of a *purely geometric* origin, the length scale L , and the material response $\sigma(\omega)$. With this separation established, all outstanding questions pertain to the geometric dependence of ζ_v . These questions center on the actual values assumed by ζ_v , how they might be calculated in practice, and whether any can be obtained by analytical means. In practice, the latter question is answerable chiefly in the negative sense: only the extended sheet and the half-space have known exact analytical solutions – which, in the half-space, by no means feature simple or compact expressions [271].

Nevertheless, it is possible to compute ζ_v in geometries of symmetry, e.g. the ribbon and disk, by semi-analytical means; and in arbitrary geometries by numerical means. One particularly prevalent approach for numerical considerations has revolved around the application of standard 3D electromagnetic codes, e.g. BEM, FDTD, or FEM, applied to a slab of thickness t with lateral projection Ω and t -dependent *bulk* conductivity $\sigma_t^{\text{bulk}}(\omega) = \sigma(\omega)/t$. In the thin-slab limit $t \rightarrow 0$ the current is predominately in-plane and eventually reproduces the 2D case in the sense $\mathbf{J}(\mathbf{r}) \sim \mathbf{K}(\mathbf{r}_{\parallel})t^{-1}\theta(t/2 - |z|) \rightarrow \mathbf{K}(\mathbf{r}_{\parallel})\delta(z)$. Accordingly, the approach is rigorous in the $t \rightarrow 0$ limit: for instance, Eqs. (85) is derivable also from the $t \rightarrow 0$ limit of the reflection coefficient of an extended slab with dielectric function $\epsilon_t^{\text{bulk}} = 1 - i\sigma_t^{\text{bulk}}/\epsilon_0\omega$ [38]. In practice, the exact limit is attained only approximately by decreasing t until convergence, typically to ~ 1 nm. Though undeniably a productive approach, its computational efficiency is far from optimal – in turn, this in practice limits the achievable accuracy as well.

We emphasize that this step does not make use of the BC – it is merely a consequence of the assumed reduction of integration domain from Ω to Ω_- .

The response-magnitude at each eigenfrequency will, of course, depend on the associated eigenstates; but is crucially similarly separable in geometric, scale, and material dependencies.

We suggest that the comparative absence of analytical solutions relative to 3D metal plasmonics, can be traced to a dimensional mismatch between structure and space: a 2D graphene sample exists in a 3D electromagnetic reality; accordingly, severe restrictions on “good” symmetries are imposed from the outset.



Optical 3D approximation of a 2D graphene sample.

It is the purpose of this chapter to discuss strictly 2D approaches, with a focus mainly on results and interpretation rather than technical details. We will first briefly revisit the extended sheet, then nanoribbons and the half-sheet, then disks, and finally comment on arbitrary structures. We provide highly accurate values of ζ_ν for each; we hope this may have some use e.g. to the verification and benchmarking of new numerical methods for 2D materials, and perhaps also to the facile comparison of experiment and theory.

EXTENDED GRAPHENE The extended sheet solution of Eq. (95a) is straightforwardly obtained from an ansatz $\phi(\mathbf{r}_\parallel) = e^{ik_\parallel x}$. The momentum k_\parallel defines the only optical length scale, and the choice $L = k_\parallel^{-1}$ for the spatial normalization of ζ then follows naturally. The labor minimal, involving just the evaluation of integrals, feasible in either Cartesian or polar coordinates, with the end-result that $\zeta_{\text{surface}} = 2i\epsilon_0\bar{\epsilon}\omega k_\parallel^{-1}/\sigma(\omega) = 1$. Unsurprisingly, this result merely reestablishes Eq. (87).

In preparation for consideration of finite structures, it will benefit us to introduce a slight reformulation of Eq. (95b). Specifically, by introducing effective momenta $k_\nu^{\text{eff}} \equiv \zeta_\nu/L$ the dispersion can be recast to the form of the extended GP, Eq. (87), such that

$$\omega_\nu = \frac{\sigma(\omega_\nu)}{2i\epsilon_0\bar{\epsilon}} k_\nu^{\text{eff}}. \quad (102)$$

As we will see, the effective momentum k_ν^{eff} carries significant interpretational power since it allows the construction of simple standing wave arguments, furnishing us with a framework for an intuitive understanding of the relation between ζ_ν and Ω .

NANORIBBONS AND THE HALF-SHEET The simplest nontrivial modification of the extended sheet is that of a half-sheet, e.g. occupying the region $x < 0$. In turn, a ribbon of finite width W constitutes the generalization of the half-space. Given this hierarchal ordering, it is not surprising that several key features of the ribbon can be understood from its simpler variants. Let us therefore consider the half-sheet first.

Besides a continuum of quasi-bulk solutions which propagate away from $x = 0$ with near-unity eigenvalues $\zeta \geq 1$, the half-sheet also supports a single edge-localized plasmon propagating along y with momentum k_\parallel and associated phase $e^{ik_\parallel y}$. Very near the edge, i.e. for $|k_\parallel x| \ll 1$, the induced density associated with this edge plasmon diverges in the asymptotic sense $\sim |k_\parallel x|^{-1/2}$, while, appreciably away from the edge, $|k_\parallel x| \gg 1$, it decays exponentially as $\sim e^{\kappa_\perp x}$; see Figure 23. A closed-form solution of its properties can be deduced by means of the so-called Wiener–Hopf technique [271, 273]; though the solution is expansive and in (nested) integral-form. Alternatively, the eigenvalues can be computed semi-analytically by a polynomial expansion technique with analytical matrix-elements [274], although the approach suffers from comparatively poor eigendensity convergence. Regardless of approach, its eigenvalue is found at $\zeta_{\text{edge}} \approx 0.8216$. The asymptotic decay constant κ_\perp is derived in Ref. 271 as $\kappa_\perp = k_\parallel(1 - \zeta_{\text{edge}}^2)^{1/2}$, and describes the behavior of $\rho_\parallel(x)$ excellently cf. Figure 23 – moreover, it describes not only the induced density, but also the $x < 0$ decay of the associated in-plane potential. Given this exponential behavior, it is apparent that the half-sheet edge plasmon constitutes the lower-dimensional analogue of the familiar SP in metal-dielectric half-spaces.

As a consequence of translational invariance the nanoribbon case is – just as the half-sheet – parametrized by the momentum k_\parallel along the invariant direction, say y . In contrast, the ribbon width W , then along x , instates the first “real” geometric length scale encountered so far; accordingly it suggests the natural normalization choice of ζ_n , i.e. $L = W$ (here with parity-related eigenindex $\nu = n$). The eigenproblem can be solved efficiently either by direct discretization of the integro-differential equation of Eq. (95a), as discussed in [275], or, just as well, by the conversion to a strict integral

We note the necessary definite integrals [272]:
 $\int_0^{2\pi} e^{ir\cos\theta} d\theta = 2\pi J_0(r)$
 and $\int_0^\infty J_0(r) dr = 1$.

The analytical origin of the value $\zeta_{\text{edge}} \approx 0.8216$ is provided by the Wiener–Hopf technique: it is the solution to the integral equation [271]
 $\int_0^{\pi/2} \ln[(\zeta \sin x)^{-1} - 1] dx = 0$.

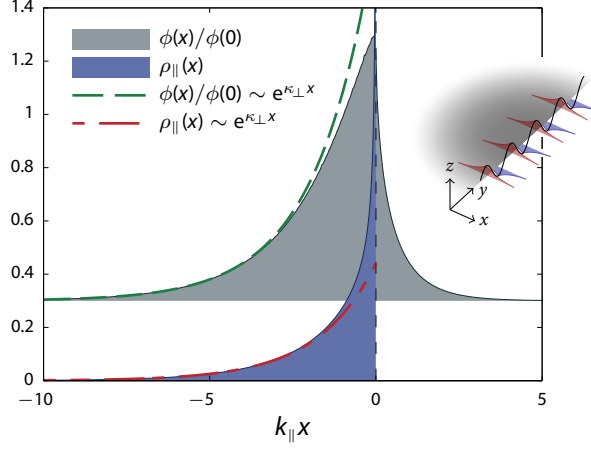


Figure 23: Induced potentials $\phi(x)$ and densities $\rho_{\parallel}(x)$ of an edge-plasmon propagating along y with momentum k_{\parallel} , i.e. with spatial dependence $\phi(\mathbf{r}_{\parallel}) = \phi(x)e^{ik_{\parallel}y}$ and $\rho_{\parallel}(\mathbf{r}_{\parallel}) = \rho_{\parallel}(x)e^{ik_{\parallel}y}$. Curves for the potential are offset by 0.3 units for intelligibility, and the normalization of $\rho_{\parallel}(x)$ is fixed by an arbitrary choice of integrated area. The asymptotic large $|k_{\parallel}x| \gg 1$ behavior of $\phi(x)$ and $\rho_{\parallel}(x)$ are indicated by dashed and dash-dotted lines. Inset illustrates the setup and the propagating nature of the edge plasmon.

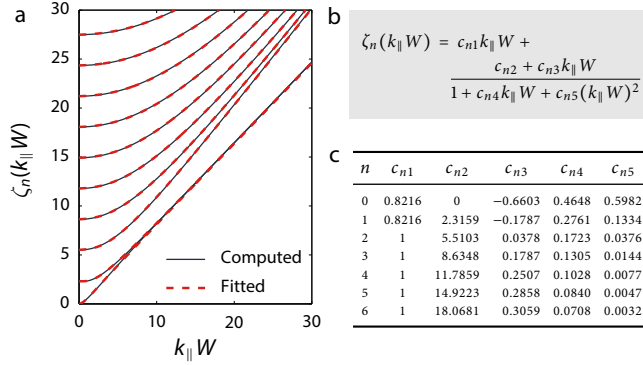


Figure 24: Eigenvalues ζ_n of a single nanoribbon, describing plasmons with transverse momentum k_{\parallel} and characteristic length L equaling the ribbon width W . (a) Computed eigenvalues of a single ribbon as a function of the normalized momentum $k_{\parallel}W$ and analytical fits. The method of Ref. 275 is employed with 2500 discretization points, i.e. convergence is ensured across all parameters. (b-c) Fitting formula and associated coefficients c_{nm} [as used in (a)]. The formula ensures the correct large momentum asymptotics; consequently, the fits are valid for *all* $k_{\parallel}W > 0$.

equation in $\rho_{\parallel}(x)$ via the introduction of a Green function [276] (being preferable in the context of hydrodynamic effects, a subject which we return to in Section 6.1). Using the former approach, we calculate and depict in Figures 24(a) the momentum dispersion of the eigenvalues for $n = 1, 2, \dots$, corresponding to monopoles, dipoles and so forth. We also provide in Figures 24(b-c) highly accurate analytical fits for eigenmodes up to $n = 7$ that are asymptotically valid; fitting formula restricted to $k_{\parallel}W \lesssim 10$ were previously provided in Ref. 278.

The coefficients c_{n1} and c_{n2} are of particular importance since they indicate, respectively, the large and small momentum limits of the dispersion. The large-momentum limits – namely, $\zeta_{n \geq 2} \rightarrow \zeta_{\text{surface}} \times k_{\parallel}W$ and $\zeta_{n \leq 1} \rightarrow \zeta_{\text{edge}} \times k_{\parallel}W$ – are especially illuminating, because they highlight a physical distinction between higher order modes $n \geq 2$ and the monopole and dipole modes $n \leq 1$. By virtue of this limiting behavior,

The fits exhibit a mean relative deviation of less than $\approx 2\%$ compared to the full numerical solution. The monopole fit, $n = 0$, however, is less accurate ($\approx 1\%$).

The monopole is physically distinct from the $n \neq 0$ modes for $k_{\parallel}W \ll 1$, exhibiting the distinctively 1D behavior (see Table 1) $\zeta_0 \propto (k_{\parallel}W)^2 \ln(k_{\parallel}W)$, with an associated frequency dispersion $\omega \propto k_{\parallel} \sqrt{\ln(k_{\parallel}W)}$; an analytical demonstration of this is provided by the variational treatment of Ref. 277.

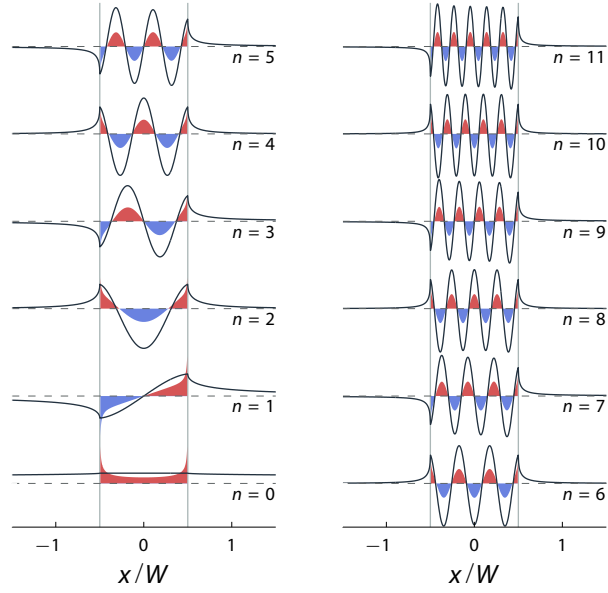
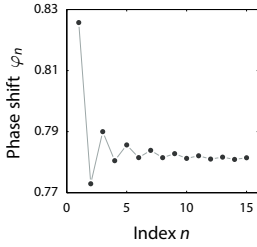


Figure 25: Eigenprofiles of a single nanoribbon, calculated at near-vanishing momentum ($k_{\parallel} W = 10^{-4}$). Induced charge densities $\rho_{\parallel}(x)$ indicated by filled areas, with associated sign (+/-) highlighted by color (red/blue), and in-plane potentials $\phi(x)$ in full lines. Gray vertical lines outline the nanoribbon's extent. Depicted quantities are normalized across all n with respect to their integrated absolute area, with distinct normalizations for $\rho_{\parallel}(x)$ and $\phi(x)$.

the monopole and dipole plasmons are naturally interpreted as symmetric and anti-symmetric hybridizations of two edge plasmons of the half-sheet, while higher order modes $n \geq 2$ essentially represent standing wave surface plasmons of the extended sheet. It is worthwhile to pursue the standing wave picture further also in the opposite limit, i.e. in the $k_{\parallel} W \rightarrow 0$ limit described by c_{n1} . Naïvely, one would expect a quantized behavior of the sort $2k_n^{\text{eff}} W = 2\pi n$, with $k_n^{\text{eff}} = \zeta_n(k_{\parallel} W = 0)/W$. It is clear, however, that the values of c_{n1} depart significantly from this simple picture. Nevertheless, a partial recovery of this intuition can be made by allowing a phase shift φ_n due to reflection at either edge, thereby generalizing the quantization condition to $2k_n^{\text{eff}} W + 2\varphi_n = 2\pi n$ [279]. Upon calculation, we find that the phase shift is nearly n -independent, converging rapidly to $\varphi_{n \gg 1} \approx 0.781$ suggestive of a physical affiliation with the edge plasmon at $\zeta_{\text{edge}} \approx 0.8216$. We summarize these considerations by visual means in Figure 25 by plotting the induced charge and potential of each eigenmode in the $k_{\parallel} W \rightarrow 0$ limit.



Reflection phase shifts in standing wave model of nanoribbon plasmons.

NANODISKS The circular truncation of radius R is the simplest strictly finite 2D structure. Perhaps as a result of this simplicity, it constitutes one of the most prominent members of the graphene plasmonics zoo, posing a key element in a range of theoretical [39, 265, 280–282] studies, as well as pioneering experimental demonstrations [239–241].

As a result of angular symmetry, the associated eigenindex ν is decomposable in azimuthal $l = 0, \pm 1, \dots$ and radial quantum numbers $n = 1, 2, \dots$. In practice, the eigensolutions can be computed with remarkable efficiency and accuracy by the polynomial expansion scheme introduced by Fetter [283] in the context of liquid helium 2D surfaces. A complete summary of the approach is supplied in Publication E; briefly, its key trait involves the expansion of the induced density (as opposed to the potential) in Jacobi polynomials, allowing an expression of overlap elements, and obtainment of expansion coefficients by inversion of an analytical matrix (in turn, the associated

We note an unfortunate (typographic) error in the Supplementary Material of Publication E: in Eq. (S13c), the term $\propto \delta_{j,k+1}$ is divided by $8 \prod_{p=1}^3 (l + 2j + p)$. This divisor should have read $8 \prod_{p=1}^3 (l + 2k + p)$ in order to provide a symmetric matrix.

	$n = 1$	$n = 2$	$n = 3$	$n = 4$
$l = 1$				
$\zeta_{1n}/2\pi$	1.0977	4.9140	8.1337	11.3079
$l = 2$				
$\zeta_{2n}/2\pi$	1.9942	6.2455	9.5455	12.7592
$l = 3$				
$\zeta_{3n}/2\pi$	2.8556	7.5124	10.8989	14.1596
$l = 4$				
$\zeta_{4n}/2\pi$	3.7032	8.7395	12.2117	15.5221

Table 2: Disk eigensolutions for different (non-monopolar) combinations of $\{ln\}$, here indicated by the associated in-plane eigenpotentials and eigenvalues. In each table entry, we provide in the upper panel a line plot of the in-plane potential $\text{Re } \phi_{ln}(x, y = 0)$ (zero-value indicated by dashed line), in the middle panel a contour plot of $\text{Re } \phi_{ln}(\mathbf{r}_{\parallel})$ (colormap ranging over blue, white, and red corresponding to maximally negative, zero, and maximally positive) – with disk boundary indicated in both cases by gray lines – and finally in the lower panel the eigenvalues ζ_{ln} divided by 2π and normalized by $L = 2\pi R$. Calculated by the method discussed in [Publication E](#): concretely by expansion in 50 and 250 Jacobi polynomials for ϕ_{ln} and ζ_{ln} , respectively.

potential follows from the same expansion coefficients, applied to hypergeometric functions rather than Jacobi polynomials). In [Table 2](#) we depict calculated in-plane eigenpotentials $\phi_{ln}(\mathbf{r}_{\parallel})$ and eigenvalues ζ_{ln} of the single nanodisk obtained by this method for the first few (non-monopolar) combinations of $\{ln\}$. The spatial dependence of $\phi_{ln}(\mathbf{r}_{\parallel})$, apart from the azimuthal component $\propto e^{il\theta}$, is qualitatively similar to that of the nanoribbon. The eigenvalues ζ_{ln} , here normalized by the disk circumference $L = 2\pi R$, also bear mention. In particular, although the phase shift argument applied to the nanoribbon cannot be transferred directly, it is evident that the $n = 1$ eigenvalues approximately fulfill a whispering gallery condition $2\pi R k_{l1}^{\text{eff}} \sim 2\pi l$ with the equivalent form $\zeta_{l1}/2\pi \sim l$. The $n \geq 2$ eigenvalues do not exhibit a similarly simple pattern, owing to additional quantization along the radial direction.

As for LSPs in metallic nanostructures, the plasmons of nanodisks are worthwhile primarily because of their large coupling with external plane waves, allowing e.g. dramatic absorption enhancement [[280](#)]. For this reason, let us here outline a pleasing

rigorous connection between the eigenmodes and absorption properties of *general* nanoscopic 2D structures – the derivation of which is given in [Appendix C](#). Concretely, the v th (composite index) eigenmodes of induced charge density $\rho_{\parallel v}(\tilde{\mathbf{r}}_{\parallel})$ and in-plane potential $\phi_v(\tilde{\mathbf{r}}_{\parallel})$ constitute a biorthogonal basis, whose normalization can be chosen conveniently as $\langle \rho_{\parallel v} | \phi_v \rangle = 2\varepsilon_0 \bar{\varepsilon} \zeta_v / L$ (integration over the scaled coordinates $\tilde{\mathbf{r}}_{\parallel} \in \tilde{\Omega}$). Considering excitation by a plane wave of amplitude E_0 , frequency ω , propagation direction z , and polarization along x , the induced dipole moment in the same direction follows from $p(\omega) = \varepsilon_0 \bar{\varepsilon} \alpha(\omega) E_0$ with the dipole polarizability

$$\alpha(\omega) = 2L^3 \sum_v \frac{|\langle \tilde{x} | \rho_{\parallel v} \rangle|^2}{\zeta_v - \zeta(\omega)}, \quad (103)$$

where we have re-instated specific declaration of frequency-dependence, and where the eigendensity $\rho_{\parallel v}$ denotes a dimensionless quantity courtesy of the normalization choice. We note that similar results have been discussed on several occasions, e.g. first in disks and rings [239], in nanoribbons [279], and recently in an electric-field formulation [282]; more generally, the statement is analogous to an earlier analysis within the BEM-formalism by Ouyang and Isaacson [284].

For an arbitrarily polarized excitation wave, e.g. along $\hat{\mathbf{n}}_{\mathbf{E}}$, the relevant polarizability is trivially obtained from Eq. (103) by the substitution $\tilde{x} \rightarrow \hat{\mathbf{n}}_{\mathbf{E}} \cdot \tilde{\mathbf{r}}_{\parallel}$.

n	$\zeta_{\pm 1, n}$	$ \langle \tilde{x} \rho_{\parallel n}^{\pm} \rangle ^2$	n	ζ_n	$ \langle \tilde{x} \rho_{\parallel n} \rangle ^2$
1	1.0977	2.8912	1	2.3159	0.8791
2	4.9140	0.1120	3	8.6348	0.0489
3	8.1337	0.0424	5	14.9223	0.0127
4	11.3079	0.0224	7	21.2024	0.0045
5	14.4675	0.0140	9	27.4843	0.0017
6	17.6205	0.0096	11	33.7656	0.0006

(a) Disk parameters for $l = \pm 1$ and normalization $L = R$.

(b) Ribbon parameters for $k_{\parallel} W = 0$, odd n , and normalization $L = W$.

Table 3: Eigendecomposition parameters associated with [Eq. \(103\)](#) in disk and ribbon geometries. (a) The disk oscillator strengths vanish for all azimuthal numbers l except for $l = \pm 1$. The listed oscillator strengths $|\langle \tilde{x} | \rho_{\parallel n}^{\pm} \rangle|^2$ sum the (equal) contributions of the $l = \pm 1$ terms such that $|\langle \tilde{x} | \rho_{\parallel n}^{\pm} \rangle|^2 \equiv \sum_{l=\pm 1} |\langle \tilde{x} | \rho_{\parallel, l n} \rangle|^2$, reducing the v -summation to one over $n = 1, 2, \dots$ with the noted oscillator strengths. (b) Similarly, the ribbon oscillator strengths vanish for all even n cf. parity considerations.

The interpretational and practical significance of [Eq. \(103\)](#) is readily appreciable. Specifically, only two geometric parameters are necessary to describe scattering properties of the v th resonance: namely, the eigenvalues ζ_v , describing the resonance position, and the (normalized) squared dipole moment $|\langle \tilde{x} | \rho_{\parallel v} \rangle|^2$, describing the resonance oscillator strength. As an example, we tabulate values of these oscillator parameters in [Table 3](#) for the nanodisk and nanoribbon, computed by means of the aforementioned highly accurate methods. As noted above, disk parameters were previously presented in [Ref. 239](#) for the dominant mode, computed by means of the finite-height approach combined with the BEM. Although no claim of absolute convergence was made, it is interesting to note the values attained thereby: in particular, for the important $\{l = \pm 1, n = 1\}$ dipole mode the stated values $\zeta_{\pm 1, 1}^{[239]} = 12.5/4\pi \approx 0.9947$ and $|\langle \tilde{x} | \rho_{\parallel 1}^{\pm [239]} \rangle|^2 = 2.6$ deviate by roughly 10% from the values obtained here by dedicated, specialized schemes. Though an acceptable deviation for many purposes, this comparison underscores the importance of dedicated two-dimensional approaches for accurate considerations of the interrelation between geometry and plasmonic properties.

Finally, we discuss a practical application of the formalism just discussed, namely calculation of absorption cross-sections via [Eqs. \(21\)](#) and [\(103\)](#). In [Figure 26](#) we depict

We remind that the nanoribbon's polarizability requires special interpretation cf. its semi-infinite extent; nevertheless the concept is still fruitful, e.g. in consideration of absorption per unit length.

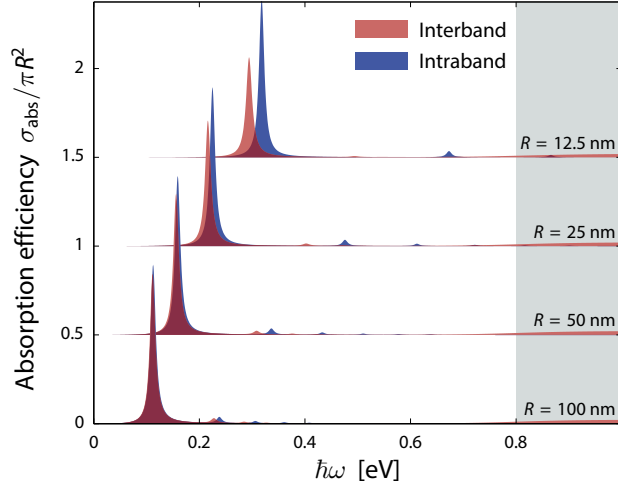


Figure 26: Absorption efficiencies of individual nanodisks in the radius range $R = 12.5, 25, 50,$ and 100 nm, as indicated. Calculated within local intraband (blue) as well as intra- and interband (red) conductivity models, for Fermi energy $\epsilon_F = 0.4$ eV, loss rate $\hbar\gamma = 12$ meV, and temperature $T = 300$ K. The region of vertical Landau transitions, $\hbar\omega \geq 2\epsilon_F$, is indicated by gray shading. Response computed via the eigendecomposition of [Table 3a](#) (being in *absolute* agreement with direct calculations that are not depicted here).

the absorption efficiency $\sigma_{\text{abs}}(\omega)/\pi R^2$ of individual graphene nanodisks for different radii. For realistic values of the Fermi energy, loss rate, and temperature the predicted absorption efficiencies of the $n = 1$ resonance approach unity in both intraband and full local treatments, while the $n > 2$ resonances only incur minor features. The radius influences the absorption spectra primarily by shifting the resonance, as appreciable from the intraband approximation of the ν th eigenfrequency $\omega_\nu^2 \propto \zeta_\nu/R$ cf. [Eq. \(96\)](#).

ARBITRARY STRUCTURES The cases so far considered lend themselves to semianalytic treatments due to their symmetries. In general structures without major symmetries, however, one must eventually retort to fully numerical mesh-based approaches. Nevertheless, this can still be achieved in a 2D formulation, using a method introduced in [Publication C \[285\]](#). Specifically, applying the ideas behind the FEM and working with a triangular mesh in barycentric coordinates, the governing equations [Eqs. \(94\)](#) can be discretized over the mesh vertices, yielding a method with computational scaling $\mathcal{O}(N^3)$ for a N -vertex mesh. Typically, excellently converged results are obtained with a few thousand vertices.

As an example of this method we tabulate in [Table 4](#) the eigenvalues $\zeta_\nu/2\pi$ in the four simplest regular polygons and in a disk with normalizing length L equaling the domain circumference C . Again, the whispering gallery argument, $k_\nu^{\text{eff}}C = 2\pi\nu \Leftrightarrow \zeta_\nu/2\pi = \nu$, provides a good qualitative description for the dipole ($\nu = 1$), quadrupole ($\nu = 2$), and hexapole ($\nu = 3$) plasmons. This is especially true for the smoother structures such as the pentagon, hexagon, and disk, although the argument's quality deteriorates for larger ν also there.

A last matter so far not touched upon is that of coupled structures. Several features are in principle immediately open to investigation: for instance, the method used to study ribbons is trivially generalizable to multiple ribbons, and similarly so for the mesh-based approach just discussed. The features fully follow the expectation: hybridization splits degeneracies with a strength proportional to the interstructural distance and large field enhancements arise in gap regions. We exemplify this statement in [Figure 27](#) by considering the energy-dispersion of in-plane coupled nanorib-

$N \times N$ matrix inversion exhibits computational complexity $\mathcal{O}(N^3)$ for the conventional Gauss-Jordan algorithm; improved scaling can be obtained for large N using iterative methods.

Graphene bowties, first treated in [Ref. 286](#), is discussed further in [Publication C](#).






ν					
1	0.864	1.069	1.106	1.110	1.098
2	2.496	2.900	1.811	1.931	1.994
3	3.387	4.407	3.191	3.798	2.856

Table 4: The three lowest eigenvalues $\zeta_\nu/2\pi$ of regular polygons, obtained from [Publication C](#), and a disk, see [Table 2](#) [63], with normalization length L chosen as the domain circumference C (monopolar modes excluded). Relative convergence of the polygonal eigenvalues is estimated as $\gtrsim 98.5\%$ by comparison of disk-eigenvalues obtained by the meshing-approach [285] and the semi-analytical polynomial expansion approach [63].

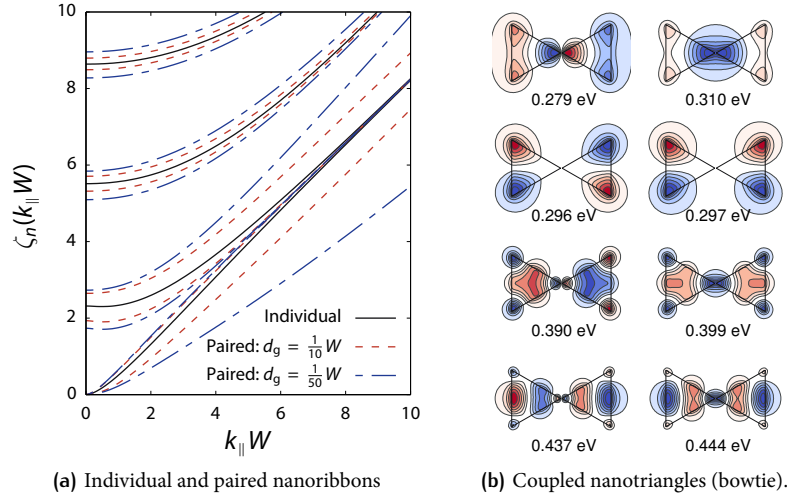


Figure 27: Plasmon features in coupled nanostructures. (a) Eigenvalue dispersion for individual (one) and paired (two) nanoribbons, the latter with in-plane gap d_g as indicated. (b) In-plane potential distribution for the eigenmodes (indicated by contours, ranging from minimum to maximum in blue to red) of a nanotriangle bowtie calculated for nanotriangles of sidelength L and intertip gap $d_g = L/40$ (reproduced from the data of [Publication C](#)). As a concrete measure of the splitting magnitude, resonance energies are indicated for the specific case of sidelength $L = 20$ nm, gap $d_g = 0.5$ nm, and Fermi energy $\epsilon_F = 0.4$ eV in a full local description. The corresponding individual nanotriangle dipole resonance is at ≈ 0.296 eV.

The otherwise interesting topic of lattice-coupling, i.e. nanostructures in periodic arrays, we leave entirely aside, although it is straightforwardly treatable by lattice-summation in the dipole limit [76].

bons and the spatial hybridization patterns of triangular bowtie structures. The key departures from the noncoupled scenario are indeed due to hybridization in even (antibonding) and odd (bonding) combinations, corresponding to blue- and redshifting resonances.

5.4 NONPLANAR GEOMETRIES: GRAPHENE-COATED SPHERE

The scenarios considered so far in this thesis – and indeed similarly so for the vast majority of efforts in graphene plasmonics in general – shares the common assumption of structural planarity. Recently, interest has emerged in exploring the properties of nonplanar geometries as well: e.g. in propagating modes in bent and corru-

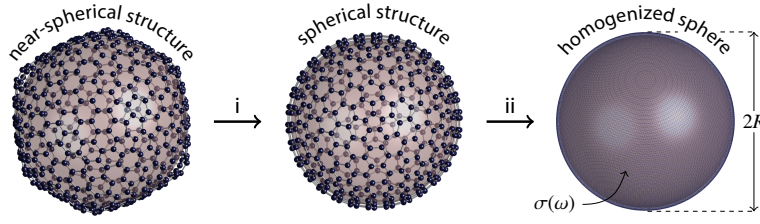


Figure 28: Illustration of the assumptions imposed on the treatment of graphene-coated nanospheres. Specifically, in step (i) all aspherical elements are neglected, while at step (ii) the microscopic details of the structure are replaced by the effective surface conductivity $\sigma(\omega)$. Depicted here for a C_{540} fullerene, for illustrative purposes only.

gated sheets [287, 288], for cloaking purposes [289, 290], in vibrating disks [285], and in particular in coated nanowire systems [291–296]. In [Publication D](#), motivated by the above-noted efforts, we examined the archetypal curved graphene structure: a graphene-coated nanosphere [71]. In this section, we review some of the results obtained therein.

The spherical geometry represents a worthwhile study for primarily two reasons. Firstly, given the sphere’s symmetry the electromagnetic analysis can be carried out analytically, and the results are straightforwardly interpretable. Accordingly, we suggest that the insights obtained here hold relevance to the interpretation and exploration of less symmetric curved configurations, such as bumps and corrugations. Secondly, the spherical geometry is of experimental relevance in its own right given recent demonstrations: in particular, we highlight the fabrication of reduced graphene-oxide hollow spheres [299], graphene encapsulation of hollow SnO_2 spheres with radii down to $R \approx 50$ nm [300], and porous multilayer graphene spheres with polystyrene interior [301]. Although each noted example exhibit several deviations from an idealized 2D spherical coating, they underscore the relevance of the spherical geometry beyond a theoretical perspective.

The simplifying assumptions imposed on the problem are summarized [Figure 28](#). Concretely, we assume that the coating is well-described by the local surface conductivity of extended graphene, i.e. we assume that the response function $\sigma(\omega)$ is approximately unmodified by the curvature. We expect this to be a valid assumption for sufficiently large spheres where $k_F R \gg 1$. Nevertheless, for smaller spheres the impact of quantization due to zone folding – familiar from carbon nanotubes [302] – as well as modified hopping amplitudes due to strain fields and orbital rotation [223] will undoubtedly weaken this assumption considerably. Neglecting these effects and proceeding with a surface conductivity description, the problem evidently poses a generalization of the Mie–Lorenz theory previously discussed in [Section 3.4.1](#); accordingly, obtainment of a set of modified Mie–Lorenz coefficients, t_l^{TE} and t_l^{TM} , constitute the natural starting point. By accounting for a nonzero tangential surface current $\mathbf{K}(\mathbf{r}, \omega) = \sigma(\omega)\mathbf{E}_\parallel(\mathbf{r}, \omega)$ the coefficients are readily obtained by the usual expansion in (solenoidal) vector waves. Specifically, denoting by $k_1 \equiv k_0\sqrt{\varepsilon_1}$ and $k_2 \equiv k_0\sqrt{\varepsilon_2}$ the wave numbers interior and exterior to the coating, respectively, the coefficients take the form

$$t_l^{\text{TE}} = \frac{-j_l(x_1)[x_2 j_l(x_2)]' + j_l(x_2)\{[x_1 j_l(x_1)]' - g(\omega)x_0^2 j_l(x_1)\}}{j_l(x_1)[x_2 h_l^{(1)}(x_2)]' - h_l^{(1)}(x_2)\{[x_1 j_l(x_1)]' - g(\omega)x_0^2 j_l(x_1)\}}, \quad (104a)$$

$$t_l^{\text{TM}} = \frac{-\varepsilon_2 j_l(x_2)[x_1 j_l(x_1)]' + [x_2 j_l(x_2)]'\{\varepsilon_1 j_l(x_1) + g(\omega)[x_1 j_l(x_1)]'\}}{\varepsilon_2 h_l^{(1)}(x_2)[x_1 j_l(x_1)]' - [x_2 h_l^{(1)}(x_2)]'\{\varepsilon_1 j_l(x_1) + g(\omega)[x_1 j_l(x_1)]'\}}, \quad (104b)$$

Although fullerenes represent a tempting small-scale analogy, they are likely not well-described by the theory developed here: in very small fullerenes quantum effects are important and an extended-graphene description accordingly poor, while the larger fullerenes exhibit icosahedral rather than spherical configurations [297, 298].

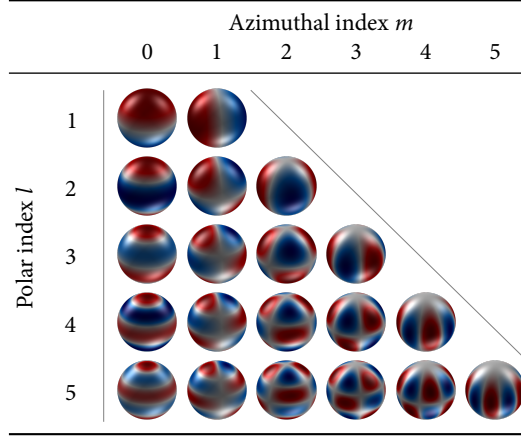


Table 5: The induced surface density in the spherical graphene coating (real part) of the lm th multipole plasmon. The density profile is $\propto P_l^m(\cos\theta)e^{im\varphi}$ and is depicted here in blue and red, corresponding to positive and negative. Azimuthal indices $m < 0$ differ from their positive counterparts only by rotational direction.

with dimensionless wave numbers $x_j \equiv k_j R$ for $j = \{0, 1, 2\}$, and with graphene's conductive properties included through the dimensionless parameter

$$g(\omega) \equiv \frac{i\sigma(\omega)}{\varepsilon_0 \omega R}. \quad (104c)$$

Unsurprisingly, we recover the standard Mie–Lorenz coefficients in the limit of vanishing $\sigma(\omega)$ where $g(\omega) \rightarrow 0$ [compare with Eq. (59) in the $\delta_l^{\text{HDM}} \rightarrow 0$ limit].

The coefficients in Eqs. (104) retain the full retarded response, and their poles, by extension, the retarded resonances. As established in Section 5.2.2, however, the properties of GPs are very well described by the nonretarded limit. Accordingly, we focus our efforts on the TM coefficients and derive the multipolar polarizability α_l from t_l^{TM} by the same $x_0 \rightarrow 0$ transitional relation previously used in Eq. (60), thereby obtaining

$$\alpha_l = 4\pi R^{2l+1} \frac{l[\varepsilon_1 - \varepsilon_2 + (l+1)g(\omega)]}{l\varepsilon_1 + (l+1)\varepsilon_2 + l(l+1)g(\omega)}. \quad (105)$$

Again, we obtain a result highly reminiscent of the noncoated counterpart, see Eq. (18), with the coating introducing only simple modifications through terms proportional to $g(\omega)$. The nonretarded resonances are immediately obtainable from the poles of Eq. (105), allowing the following generalized Fröhlich condition for the l th order multipole plasmon with frequency ω_l

$$l\varepsilon_1 + (l+1)\varepsilon_2 + l(l+1)g(\omega_l) = 0. \quad (106)$$

An interesting point of departure from the planar cases considered so far is the unequal and l -dependent weighting of the dielectric surroundings, contrasting the equal half-and-half weighting of the planar case via $\bar{\varepsilon} \equiv (\varepsilon_1 + \varepsilon_2)/2$. As a result, the eigenvalue ζ -approach, Eq. (95b), cannot be transferred completely, except in the limit of uniform dielectric surroundings $\bar{\varepsilon} = \varepsilon_1 = \varepsilon_2$ where one finds $\zeta_l^{\text{sphere}} = [1 + (2l+1)^{-1}]l$ for a normalization $L = R$, simplifying to $4/3$ for the optically important $l = 1$ dipole mode. The associated effective plasmon momenta $k_l^{\text{eff}} = \zeta_l^{\text{sphere}}/R$ approach l/R for large multipole orders indicating the asymptotic fulfillment of a whispering gallery condition. For the purpose of illustrating the spatial character of the modes, we depict in Table 5 the induced surface charge in the graphene coating for the first few multipoles, in their different azimuthal (i.e. m eigenindex) realizations.

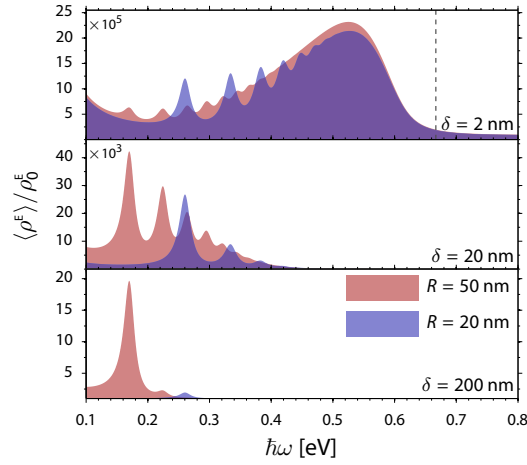


Figure 29: LDOS enhancement near graphene-coated spheres in vacuum surroundings $\epsilon_1 = \epsilon_2 = 1$ for different sphere-radii $R = 20$ and 50 nm, as indicated, and at three distinct distances $\delta = b - R$ from the sphere surface. Calculated by weighted summation of the nonretarded multipole polarizability [63, 303], using for $\sigma(\omega)$ the full local conductivity of graphene with Fermi energy $\epsilon_F = 0.4$ eV, loss rate $\hbar\gamma = 20$ meV, and temperature $T = 300$ K. In the top panel, the dashed line indicates the $l \rightarrow \infty$ multipole asymptote at $\hbar\omega/\epsilon_F \approx 1.6671$.

Although far-field measurements only significantly probe the dipole resonance, the higher multipoles play an important part in near-field measurements – just as they did for the LSPs of metallic nanospheres in Sections 3.4.3 and 3.5. To exemplify this, we consider in Figure 29 the spectral dependence of the (emitter-orientation-averaged) LDOS enhancement $\langle \rho^E \rangle / \rho_0^E$ for graphene-coated nanospheres in vacuum, at different radii and probe-to-surface separations δ . The interpretation of the results fully follows the intuition developed in Section 3.4.3: at large probe-to-surface separations the dipole mode is dominant, while, at increasingly shorter separations the higher multipoles contribute significant features. Eventually, at the shortest considered separation, a broad peak emerges, comprised of several spectrally overlapping multipoles which pile-up towards the local intra- and interband model’s asymptote at $\hbar\omega/\epsilon_F \approx 1.6671$, see Section 5.2.2, in a manner entirely analogous to the pile-up in metallic nanospheres near $\omega_p/\sqrt{2}$.

In looking forward, we comment that the formal study of more general curved structures presently remains largely unexplored territory. For instance, considering a “kink” or ridge, with characteristic radius of curvature L , inserted in an otherwise flat graphene sheet, one would naïvely expect the emergence of localized resonances comprised of wavepackets with center-momentum $\sim L^{-1}$. The rigorous study of these ideas require a transformation approach between “kinked” and flat scenarios; one which can be achieved using the tools of differential geometry by constructing an appropriate metric tensor g_{ij} for the deformation of the sheet. Though we shall not pursue this path further here, we note that a governing equation analogous to Eq. (95a) can be obtained with integrations extended over a strictly planar regime – the price paid, however, is the introduction of a metric-dependent occupation function and an effective Coulomb screening, $\mathbf{f}_g(\mathbf{r}_{\parallel})$ and $\epsilon_g(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel})$, respectively, exhibiting in the first case a spatially inhomogeneous and anisotropic nature, and in the second case kink-dependent nonlocality.

Several additional concrete examples and explicit calculations related to coated nanospheres are presented in Publication D – treating size-dispersion, far-field extinction, hybridization with an underlying Drude-sphere, and inclusion of hydrodynamic response – but are not repeated here for the sake of brevity.

5.5 SUMMARY

In this chapter we have introduced the main theoretical concepts of classical 2D plasmonics. In opening our considerations, we discussed first in [Section 5.1](#) the basic dimensional differences between plasmonics in 3D, 2D, and 1D by means of a straightforward analysis involving only the momentum transform of the Coulomb potential. Subsequently, specializing to our main focus – graphene plasmonics – we reviewed in [Section 5.2](#) the basic properties of plasmons in extended graphene. These considerations were expanded to nanostructured samples in [Section 5.3](#) by the introduction of a governing eigensystem taking the form of an integro-differential equation. The plasmon dispersion and its size-dependent scaling was shown to relate simply to the eigenvalues of this system, which introduced thereby a separation of geometry, scale, and material response. Using specialized methods, these eigenvalues were calculated for different geometries, specifically for an extended sheet, a half space, ribbons, disks, and regular polygons. In addition, a simple polarizability approach was established with tabulatable geometric elements computable from the eigendensities. Finally, in [Section 5.4](#) we discussed the existence of plasmons in nonplanar 2D configurations, specifically in graphene-coated nanospheres, finding simple analytical solutions by means of a generalized Mie–Lorenz approach.

6

NONCLASSICAL GRAPHENE PLASMONICS

In spirit, the preceding treatment of classical graphene plasmonics mirrors that of the conventional LRA approach from metal plasmonics: the linear local material response of the extended sheet, i.e. a bulk property, is applied to nanostructures with the justification that relevant electronic length scales remain small compared to the optical and geometric scales. We therefore expect the classical description of graphene plasmonics to fall short in the same circumstances as its metal correspondent, specifically in the circumstances discussed in [Section 3.1](#). However, graphene differs from noble metals on a number of counts: besides dimensionality and momentum-dispersion, the relevant Fermi wavelengths λ_F differ substantially. For instance, while $\lambda_F \approx 5 \text{ \AA}$ in gold and silver [56], graphene doped to, say, $\epsilon_F = 0.4 \text{ eV}$ exhibits $\lambda_F \approx 9 \text{ nm}$. Consequently, the effects of e.g. nonlocality enter comparatively sooner. In addition, graphene affords us with a unique opportunity to go beyond the jellium treatment with modest computational means, courtesy of the dual fortuity of a lowered dimensionality and an excellent single-particle electron description via affordable TB models. The accounting of atomistic features by a TB approach holds nontrivial surprises in particular relating to the existence of edge states along zigzag terminations.

It is the purpose of the present chapter to explore aspects of the above-noted departures from classicality. We start in the semi-classical domain, by introducing a hydrodynamic model (HDM) for graphene. Next, we introduce a quantum description of graphene plasmonics at the level of the RPA, accounting for features of level quantization, nonlocality, and edge configuration. In treating the single-particle properties, we consider both the atomistic TB model as well as the continuum Dirac model, focusing in both cases especially on the role of edge states. Finally, departing from the nonclassical theme of the chapter, we study the role of nonlinearity in graphene plasmonics.

6.1 HYDRODYNAMIC DESCRIPTION

Given our interpretation of hydrodynamics in the free-electron gas, established in [Section 3.2](#), as the real-space differential form of a low-momentum expansion of the Lindhard response, it is not a great surprise that we might establish a similar hydrodynamic description for graphene. In rigorous terms, the Boltzmann equation moment-approach applied in [Section 3.2.1](#) can be largely retraced for the intraband Dirac description, with one important difference. Concretely a 2D continuity equation follows from the zeroth moment, just as the for the free-electron gas. The first moment, however, requires changes because the quasiparticle velocity in graphene is $\mathbf{v}_{\mathbf{k}} = \hbar^{-1} \nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} = v_F \mathbf{k}/|\mathbf{k}|$, contrary to that of the parabolic gas where $\mathbf{v}_{\mathbf{k}} = \hbar \mathbf{k}/m_{\text{eff}}$. As a result, the Cauchy momentum equation is changed, specifically in the manner indicated by Müller *et al.* [304] in the context of transport considerations [304–306]. Briefly, in the linearized regime, a connection can be made with the free-electron gas result, [Eq. \(37b\)](#), by replacing the effective parabolic mass with graphene’s Drude mass, $m_{\text{eff}} \rightarrow m_{\text{eff}}^g \equiv \hbar k_F/v_F$, and by reducing the pressure contribution to that of a scalar Dirac degeneracy pressure, $\nabla \cdot \mathbf{P} \rightarrow \nabla_{\parallel} (u_{\omega} \epsilon_F n) \propto \nabla_{\parallel} n^{3/2}$ with $u_{\omega} \equiv \begin{cases} 1/3 & \text{for } \omega \ll \gamma \\ 1/2 & \text{for } \omega \gg \gamma \end{cases}$.

We refrain from further detailing the approach of Müller *et al.* [304], but emphasize that the above in principle provides a complete foundation for the construction of an intraband HDM of graphene. Instead we opt to study in detail a more pragmatic

The low-frequency degeneracy pressure of a Dirac system of density $n = N/A$ is derivable as the (negative) area (A) derivative of the total internal energy $U \propto N^{3/2} A^{-1/2}$ for fixed particle number N .

approach, explicated also in [Publication E](#), which elucidates the properties of not just intraband but also interband nonlocality. Specifically, we consider the 2D analogue of the HDM current-field relationship [Eq. \(41\)](#) (neglecting loss)

$$\mathbf{K}(\mathbf{r}_{\parallel}, \omega) + \frac{\beta^2}{\omega^2} \nabla_{\parallel} [\nabla_{\parallel} \cdot \mathbf{K}(\mathbf{r}_{\parallel}, \omega)] = \sigma_{\text{B}}(\omega) \mathbf{E}_{\parallel}(\mathbf{r}_{\parallel}, \omega), \quad (107)$$

and seek to determine the appropriate value of the pressure velocity β in graphene matching the intra- and interband LRA bulk-conductivity $\sigma_{\text{B}}(\omega)$. To do so, we first note the momentum-space hydrodynamic conductivity associated with [Eq. \(107\)](#) for longitudinal excitations $\mathbf{K}(k_{\parallel}, \omega) = \sigma_{\text{L}}^{\text{HDM}}(k_{\parallel}, \omega) \mathbf{E}_{\parallel}(k_{\parallel}, \omega)$ with $\mathbf{k}_{\parallel} \parallel \mathbf{E}_{\parallel}$

$$\sigma_{\text{L}}^{\text{HDM}}(k_{\parallel}, \omega) = \frac{\sigma_{\text{B}}(\omega)}{1 - \beta^2 \frac{k_{\parallel}^2}{\omega^2}} \simeq \sigma_{\text{B}}(\omega) \left(1 + \beta^2 \frac{k_{\parallel}^2}{\omega^2} \right). \quad (108)$$

Clearly, since the hydrodynamic approach must match the RPA result asymptotically, the value of β can be obtained by comparing [Eq. \(108\)](#) with the small-momentum expansion of the low-temperature nonlocal conductivity of graphene, obtained from [Eqs. \(77a\)](#) and [\(79\)](#), reading (in the region delimited by $\tilde{\omega} > \tilde{k}_{\parallel}$ and $\tilde{\omega} < 2 - \tilde{k}_{\parallel}$)

$$\sigma_{\text{L}}^{\text{RPA}}(k_{\parallel}, \omega) = \frac{2ie^2}{\pi\hbar} \frac{\tilde{\omega}}{\tilde{k}} \left\{ -1 + \frac{1}{8} F(\tilde{k}_{\parallel}, \tilde{\omega}) \left[\mathcal{W}\left(\frac{2+\tilde{\omega}}{\tilde{k}_{\parallel}}\right) - \mathcal{W}\left(\frac{2-\tilde{\omega}}{\tilde{k}_{\parallel}}\right) \right] \right\}, \quad (109)$$

with $\mathcal{W}(x)$ and $F(\tilde{k}_{\parallel}, \tilde{\omega}_{\parallel})$ defined in [Eq. \(77b\)](#), and with dimensionless quantities $\tilde{k}_{\parallel} \equiv k_{\parallel}/k_{\text{F}}$ and $\tilde{\omega} \equiv \hbar\omega/\epsilon_{\text{F}}$. The desired small-momentum expansion of $\sigma_{\text{L}}^{\text{RPA}}(k_{\parallel}, \omega)$ can be found in the region $\tilde{\omega} < 2$, allowing for $k_{\parallel} \ll k_{\text{F}}$

$$\sigma_{\text{L}}^{\text{RPA}}(k_{\parallel}, \omega) \simeq \sigma_{\text{intra}}(\omega) \left[1 + \beta_{\text{intra}}^2(\omega) \frac{k_{\parallel}^2}{\omega^2} \right] + \sigma_{\text{inter}}(\omega) \left[1 + \beta_{\text{inter}}^2 \frac{k_{\parallel}^2}{\omega^2} \right], \quad (110)$$

with *distinct* pressure velocities for the intra- and interband contributions, defined by

$$\beta_{\text{intra}}^2(\omega) \equiv \left[\frac{3}{4} - \left(\frac{\tilde{\omega}^2}{4 - \tilde{\omega}^2} \right)^2 \right] v_{\text{F}}^2 \stackrel{\tilde{\omega} \lesssim 1}{\simeq} \frac{3}{4} v_{\text{F}}^2, \quad \beta_{\text{inter}}^2 \equiv \frac{1}{2} v_{\text{F}}^2. \quad (111)$$

As a consequence of finding two distinct plasma velocities, it is clear that a one-fluid model, such as [Eq. \(107\)](#), strictly speaking does not suffice: indeed, a proper treatment would require a two-fluid model. Nonetheless, a reasonable compromise can be achieved in a one-fluid model by adopting the practical position that the pressure velocity of [Eq. \(107\)](#) equals $\beta^2 = \frac{3}{4} v_{\text{F}}^2$, thereby ignoring both the frequency-dependence of $\beta_{\text{intra}}(\omega)$ (reasonable for frequencies $\hbar\omega \lesssim \epsilon_{\text{F}}$), as well as the actual difference between inter- and intraband pressure velocities. We emphasize that the error incurred hereby is small because the interband term (and hence its nonlocal correction, upon which the error is inflicted) is small in the first place compared to the intraband term; roughly then, the error is second order. As the dust settles, we observe the remarkable fact that the β -factor of graphene ultimately mirrors that of the parabolic 2D gas, see [Eq. \(42\)](#), despite the marked qualitative difference in dispersion behavior.

The self-consistent interacting HDM treatment of graphene follows by a straightforward extension of the local constitutive relation [Eq. \(94b\)](#), namely $[1 + k_{\beta}^{-2} \nabla_{\parallel}^2] \rho_{\parallel} = i\omega^{-1} \nabla_{\parallel} \cdot (\sigma \nabla_{\parallel} \phi)$ with $k_{\beta}^2 \equiv \omega(\omega + i\gamma)/\beta^2$, obtained from the combination of [Eq. \(107\)](#) and the continuity equation. In turn, the scale invariance of the local governing eigen-system, [Eq. \(95a\)](#), is lost in a HDM treatment due to the introduction of a nonlocal length scale k_{β}^{-1} . The practical solution of this system can be achieved elegantly by the introduction of a Green function for the operator $1 + k_{\beta}^{-2} \nabla_{\parallel}^2$ consistent with a no-spill current BC.

The previously non-subscripted conductivity $\sigma = \sigma_{\text{intra}} + \sigma_{\text{inter}}$ is here appended with a B-subscript in anticipation of an impending need to differentiate several distinct local conductivity contributions.

We remind the low-loss low-temperature intra- and interband conductivities of [Eqs. \(81\)](#): $\sigma_{\text{intra}}(\omega) = \frac{ie^2}{\pi\hbar\omega}$ and $\sigma_{\text{inter}}(\omega) = \frac{ie^2}{4\pi\hbar} \ln \left| \frac{2-\tilde{\omega}}{2+\tilde{\omega}} \right|$.

The inclusion of loss can be achieved by straightforward modifications to $\sigma_{\text{B}}(\omega)$ and the substitution $\beta^2 \rightarrow \frac{\omega}{\omega+i\gamma} \beta^2$, see [Eq. \(41\)](#) and our earlier treatment of the parabolic HDM.

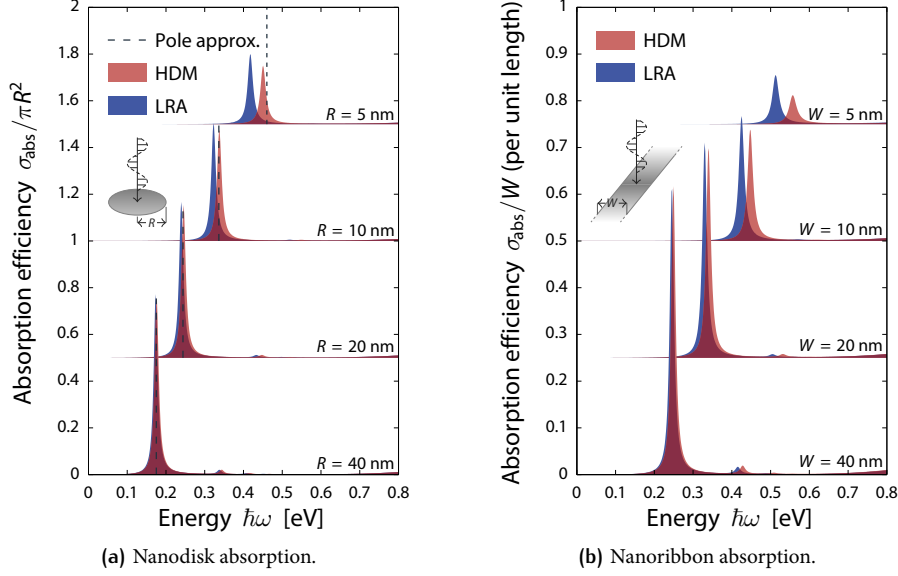


Figure 30: Absorption efficiency of (a) nanodisks and (b) nanoribbons, comparing the predictions of the classical LRA and semi-classical HDM models for varying geometry sizes, as indicated. In the nanodisk case, we include also the predictions of the pole-approximation discussed in the main text. In the nanoribbon case the exciting field is polarized along the ribbon width, i.e. along x , and the absorption efficiency is given per unit y -length, chosen as the ribbon width W . Material response includes the full intra- and interband contributions, calculated with $\epsilon_F = 0.4$ eV, $\hbar\gamma = 12$ meV, and $T = 300$ K.

For the disk and ribbon geometries this Green function can be constructed explicitly, allowing efficient solution procedures. We will not repeat the details here but refer instead to [Publication E](#) and [Ref. 283](#) for the disk and to [Ref. 276](#) for the ribbon. With these techniques, we can compute the HDM result for disks and nanoribbons: in [Figure 30](#) we depict the results of such calculations, exemplified by the absorption efficiency of graphene nanodisks and nanoribbons. The corresponding classical LRA predictions are also depicted. As we would expect on the basis of the intuition developed for the metal HDM, we observe a size-dependent blueshift of the main dipole resonance which increases with decreasing geometric size. The scaling of this shift can be discerned from a pole-approximation around the LRA resonances, as explained in [Publication E](#): for instance, for the disk dipole resonance the shift is approximately given by $\delta\omega \equiv \omega_{\text{dp}}^{\text{HDM}} - \omega_{\text{dp}}^{\text{LRA}} \simeq \Delta_{\text{dp}}\beta^2/2\omega_{\text{dp}}^{\text{LRA}}R^2$ with a pole-approximation eigenvalue $\Delta_{\text{dp}} \approx 3.39$. Accounting also for the (intraband) size-dispersion $\omega_{\text{dp}}^{\text{LRA}} \propto R^{-1/2}$ one ultimately obtains a scaling $\delta\omega \propto R^{-3/2}$. We illustrate the excellent quality of this approximation explicitly for the nanodisk in [Figure 30\(a\)](#). Notably, the scaling of the 2D HDM-correction contrasts that found in 3D metals, being there $\propto R^{-1}$ cf. [Eq. \(62\)](#). The difference, of course, arises simply because the corrections, of order $\sim R^{-1}$, apply relative to the LRA size-dispersion, being constant in 3D and $\propto R^{-1/2}$ in 2D.

This concludes our introduction to the graphene HDM for plasmonic purposes. In closing, we note that variants of this model, specifically the intraband kind, have seen use in other studies of graphene plasmonics [[307–309](#)]. We return to the use of the graphene HDM in [Section 6.2.3](#), where, in conjunction with a nonclassical edge-state correction, we invoke it to explain qualitatively the discrepancy between classical and quantum RPA calculations. Finally, we note that the graphene HDM – despite its approximate nature – obtains excellent agreement with full RPA calculations built from

While the ribbon treatment in [Ref. 276](#) seems mostly correct, we feel obliged to caution that the result for the Green function does not. For width-normalized momenta q and coordinates x , it should read

$$G(x, x') = \frac{1}{2q} \left\{ e^{-q|x-x'|} + \frac{e^{-q} \cosh[q(x-x')] + \cosh[q(x+x')]}{\sinh q} \right\},$$
as derivable from a minor extension of the half-sheet result [[274](#)].

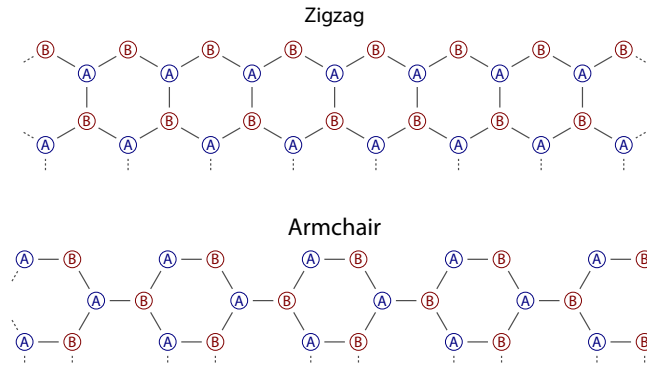


Figure 31: Illustration of the zigzag (ZZ) and armchair (AC) terminations of the honeycomb lattice. A- and B-sublattice atoms are indicated by lettering and color, blue and red, respectively. The extension of the lattice towards the bottom, left, and right is indicated by dashed lines. We highlight that the ZZ case terminates on a single sublattice, while the AC terminates on both.

the quantized states of a Dirac equation with a particularly “classical” electronic BC, known as the infinite mass BC: unfortunately, we will not find opportunity to discuss it further presently, but refer instead to the Supplementary Material of [Publication E](#).

6.2 QUANTUM GRAPHENE PLASMONICS

Our study of graphene plasmonics, and indeed also metal plasmonics, so far has labored under a continuum-assumption. This assumption – constituting essentially the ansatz that the continuous energy dispersion of a bulk material, and hence its response properties, remains representative also of nanostructures – was previously discussed in our overview of the shortcomings of the classical plasmonic approach, [Section 3.1](#), under the heading of quantum size effects. Similarly, the role of atomic configuration at the nanostructure boundary has been neglected. In this section we explore the consequences of including these previously omitted mechanisms; concretely by means of a quantized single-particle RPA treatment. The most dramatic consequence of such an expanded treatment arguably arises in the comparison of zigzag (ZZ) and armchair (AC) terminated nanostructures, see [Figure 31](#), with differences emerging due to the existence of so-called edge states in the ZZ configuration.

The section is organized as follows: first, we revisit the properties of the single-particle TB description of graphene with special emphasis on the emergence of edge states in nanostructures. Next, we discuss two computational approaches for the calculation of the RPA response in graphene, distinguished by the choice of frequency- or time-domain viewpoints, and subsequently discuss plasmonic properties in graphene nanostructures concretely in disks, triangles, and hexagons. Finally, by considering a Dirac description supplemented by termination-specific BCs, we discuss approaches to semi-classically account for edge states via local optical response functions, offering thereby a more direct understanding of their impact.

6.2.1 Edge states in graphene nanostructures

The existence of edge states in graphene can be explored in a multitude of manners and at several hierarchies of abstraction. To introduce the subject, we offer a simple argument applicable to arbitrary bipartite lattices, due to Inui *et al.* [[311](#)], which captures

Surprisingly, the earliest discussions of edge states in graphene precede its discovery by nearly two decades [[310](#)], studied then in the context of stacked benzene-chains.

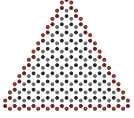
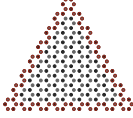
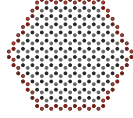
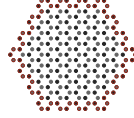
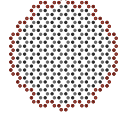
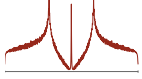
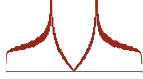
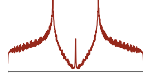
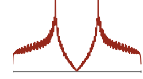
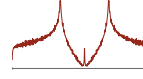
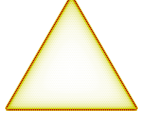
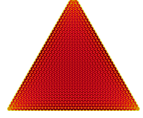
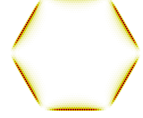
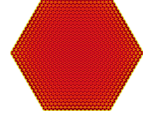
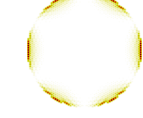
Configuration					
DOS(ϵ)					
LDOS($\mathbf{r}, \epsilon = 0$)					
N ($ N_A - N_B $)	5038 (68)	4920 (0)	5046 (0)	4902 (0)	5008 (0)
$\#\{\epsilon = 0 \text{ states}\}$	68	0	0	0	0
$\#\{\text{edge states}\}$	68	0	28	0	14

Table 6: Single-particle DOS and LDOS evaluated for five distinct graphene nanostructures of $N \approx 5 \times 10^3$ atoms (for clarity, the configuration illustrations include much fewer atoms) calculated by a nearest-neighbor TB model. The DOS is depicted as a function of energy, while the LDOS depicts the spatial variation at zero energy (for illustrative purposes, each site is represented by a short-range Gaussian function) from its maximal (red) to minimal (white) value: both are computed with broadening $\hbar\eta = 24$ meV. In addition to the exact number of atoms, N , and the sublattice asymmetry, $|N_B - N_A|$, the number of zero-energy states, $\#\{\epsilon = 0 \text{ states}\}$, and near-zero-energy edge states, $\#\{\text{edge states}\}$, are tabulated for each structure. The latter is determined by counting the number of states in the set $\{j \mid p_j < \text{tol} \wedge \epsilon_j \sim 0\}$, with a phenomenological participation tolerance $p_j < 0.03$ and an ad-hoc choice of the energy tolerance in each structure.

several characteristics of edge states. Concretely, we consider the usual TB Hamiltonian for a bipartite lattice with N sites

$$\epsilon_j \psi_{jl} = - \sum_{l' \in \text{nn}_l} t_{ll'} \psi_{jl'}, \quad (112)$$

for an eigenstate j of energy ϵ_j and ket $|\psi_j\rangle = \sum_l \psi_{jl} |l\rangle$, expanded in Wannier orbitals $|l\rangle$ centered at site l (of position \mathbf{r}_l) belonging to either the A- or B-sublattice, coupled to its nearest neighbors $l' \in \text{nn}_l$ belonging to the B- or A-sublattice, respectively. The coupling strength between sites l and l' is taken constant ($t_{ll'} = t_{AB} = 2.8$ eV) in graphene – but the following arguments are not restricted to this choice, being robust under e.g. coupling-perturbations near edges.

Consider then a finite restriction of the bipartite lattice with N_A and N_B sites on A- and B-sublattices, respectively, with $N_A \leq N_B$. Next, we seek zero-energy solutions where $\epsilon_j = 0$ and look for them in the special circumstance where $\psi_{j,l \in A} = 0$, i.e. when the wavefunction vanishes on all A sites. For $l \in B$ Eq. (112) is trivially satisfied, vanishing identically, while for $l \in A$ we must require

$$\sum_{l' \in \text{nn}_l} t_{ll'} \psi_{j,l' \in B} = 0, \quad (113)$$

yielding by iteration over $l \in A$ a set of N_A linear homogeneous equations with N_B unknowns. Standard linear algebra then guarantees at least $N_B - N_A$ (≥ 0 , by assumption) linearly independent nontrivial solutions, corresponding to fully lattice-polarized zero-energy states.

In general then, we expect at least $|N_B - N_A|$ zero-energy states in a finite bipartite lattice. It is instructive to compare this statement with explicit calculations in an

Lieb [312] offered a similar argument even earlier than Inui et al. [311], but did not comment directly on edge localization.

The variation of coupling strength for edge-atoms is ignored in our TB considerations – pragmatically, we assume the otherwise dangling bonds passivated by hydrogen atoms, and neglect p_z -orbital modifications arising therefrom.

$$\text{DOS}(\epsilon) = \frac{2}{\pi\mathcal{A}} \text{Im} \sum_j \frac{1}{\epsilon_j - \epsilon - i\hbar\eta}$$

for sample area \mathcal{A} and level-broadening η .

$$\text{LDOS}(\mathbf{r}_l, \epsilon) = \frac{2}{\pi\mathcal{A}} \text{Im} \sum_j \frac{|\psi_{jl}|^2}{\epsilon_j - \epsilon - i\hbar\eta}$$

assortment of graphene nanostructures; this is done in Table 6 for the ZZ- and AC-terminated triangles and hexagons, and for a bond-centered circular restriction with mixed terminations. For the triangles the above argument fully accounts for the outcome of the explicit calculation: exactly $|N_B - N_A|$ zero-energy states exist, with a nonzero count for the ZZ realization (which is terminated entirely on a single sublattice such that $|N_B - N_A| \neq 0$) and zero count for the AC realization. For the remaining structures in Table 6, however, the interpretation cannot be performed solely by the above arguments. Specifically, in the hexagon and disk cases, complete sublattice symmetry exists, i.e. $|N_B - N_A| = 0$, allowing no true zero-energy states. Nevertheless, the electronic single-particle DOS, display a sharp peak near zero-energy for the structures which terminate with ZZ regions, i.e. the ZZ hexagon and mixed boundary disk, mirroring features of the ZZ triangle. Still, this peak is not associated with true zero-energy states, but rather with near-zero-energy states. Both true and near-zero-energy states share a common feature: both are localized strongly to the edges of the structure, particularly to ZZ features (and are consequently also locally sublattice-polarized cf. Figure 31). This fact is appreciable from the electronic LDOS evaluated at $\epsilon = 0$, which is also depicted in Table 6.

A quantitative measure of the localization of a specific state j is given by the participation ratio [265, 313]

$$p_j = \frac{(\sum_l |\psi_{jl}|^2)^2}{N \sum_l |\psi_{jl}|^4}. \quad (114)$$

An additional class of localized states arise near the van Hove singularity at $\pm t_{AB}$ (but does not appear in the energy range of Figure 32).

For localized states with only a small fraction N_0/N of occupied sites, the participation ratio is $p_j \sim N_0/N$, while extended states with nearly constant occupancy have $p_j \sim 1$. Participation ratios of the triangle, hexagon, and disk configurations are shown in Figure 32. For the structures with ZZ terminations, a region of localized states exist that coincide with the near-zero-energy states, whereas for the strict AC case all considered states exhibit roughly similar participation ratio ~ 0.4 .

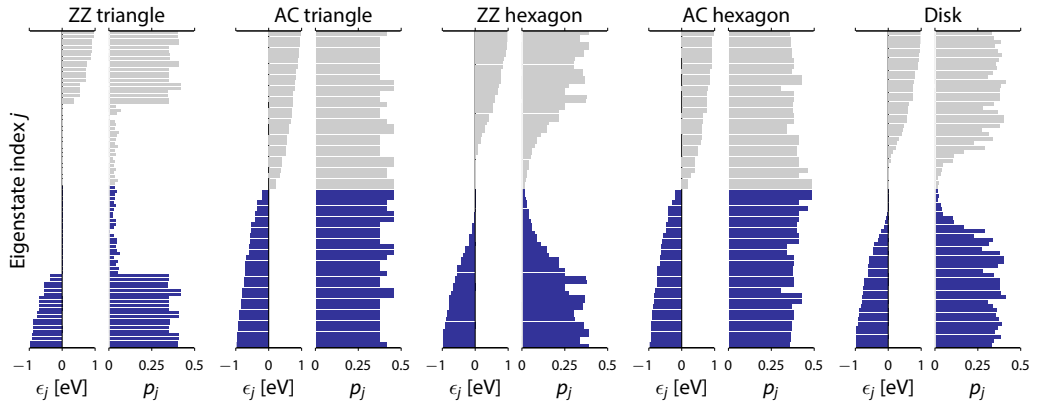


Figure 32: TB energies ϵ_j and participation ratios p_j as function of eigenstate index j , restricted to states with $|\epsilon_j| < 1$ eV. The configurations mirror those of Table 6 (indicated in titles) though with fewer atoms $N \sim 2400$.

Summarizing these introductory considerations, we suggest a phenomenological rule of thumb, which generalizes the argument of Inui *et al.* [311] on global sublattice asymmetry: specifically, near-zero-energy edge states arise whenever a *local* sublattice asymmetry exists, i.e. when the lattice terminates on a single sublattice across several connected sites; moreover, these states are localized near regions of local sublattice asymmetry. Rigorous topological arguments can be made for arbitrary edge terminations in the special case of ribbon restrictions [314, 315] (concretely, a Zak

phase of π designate edge-localized states), but are not generalizable to arbitrary restrictions. Nevertheless, the rule of thumb finds complementary and rigorous support in the work on BCs for the Dirac operator by Akhmerov and Beenakker [316], establishing that ZZ-type BCs result generically except at AC-terminations – or, in other words, that edge states analogous to those of the strict ZZ-termination associate with local sublattice asymmetry (since *only* the AC-termination exhibits complete sublattice symmetry).

In closing this introduction to edge states, we emphasize that the above theoretical considerations are well-established experimentally by a multitude of scanning tunneling microscopy and spectroscopy studies [317–321]. Furthermore, although challenges surely remain [322, 323], state-of-the-art nanofabrication techniques today allow the precise and direct control of the structural termination of graphene by a variety of methods [324–326]. With these encouraging facts in mind, we proceed to explore the role of edge states in graphene plasmonics.

6.2.2 Tight-binding and the RPA

Apart from an earlier study of thin nanoribbons [327], the full accounting of atomistic features in graphene plasmonics was approached first by Thongrattanasiri *et al.* [129] using the RPA with TB states, and subsequently explored further by the same authors in ensuing publications [286, 328–331]. Responding to these developments, we employed the same technique in Publications C and E. In this section, we detail the essential features of the frequency-domain RPA approach noted above. In addition, we introduce an equivalent method founded in a time-domain calculation, which allow complementary insights. Finally, we discuss the predictions of the TB-RPA, emphasizing especially the role played by ZZ- and AC-terminations as well as the deviations from classicality.

FREQUENCY-DOMAIN The central element of a frequency-domain RPA calculation involves the construction of the noninteracting density-density response function $\chi^0(\mathbf{r}, \mathbf{r}'; \omega)$ (occasionally referred to as the polarizability) in a real-space representation. For completeness we offer a derivation of $\chi^0(\mathbf{r}, \mathbf{r}'; \omega)$ in Appendix D from the perspective of the Kubo formalism (in addition, we generalize slightly previous results to the case of spinor wavefunctions, as needed for the Dirac treatment of Section 6.2.3). For a complete set of single-particle eigenstates $\{\epsilon_j, \psi_j(\mathbf{r})\}$ the result is

$$\chi^0(\mathbf{r}, \mathbf{r}'; \omega) = 2 \sum_{jj'} (f_j - f_{j'}) \frac{\psi_{j'}^\dagger(\mathbf{r}) \psi_j(\mathbf{r}) \psi_j^\dagger(\mathbf{r}') \psi_{j'}(\mathbf{r}')}{\epsilon_j - \epsilon_{j'} - \hbar(\omega + i\eta)}, \quad (115)$$

with Fermi–Dirac functions $f_j = \{1 + \exp[(\epsilon_j - \epsilon_F)/k_B T]\}^{-1}$, a multiplicative factor of 2 due to spin, and electronic decay rate η which, cf. Section 3.2.2, relates with the conventional optical rate through $\gamma = 2\eta$ in a (non-conserving) relaxation-time approximation. The RPA is instated in practice by allowing the induced charge density ρ to interact with the *total* self-consistent potential $\phi = \phi^{\text{ext}} + \phi^{\text{ind}}$ (rather than just the external potential ϕ^{ext}) through χ^0 ; leading to the self-consistent equations

$$\rho(\mathbf{r}, \omega) = e^2 \int \chi^0(\mathbf{r}, \mathbf{r}'; \omega) \phi(\mathbf{r}', \omega) d^2 \mathbf{r}', \quad (116a)$$

$$\phi(\mathbf{r}, \omega) = \phi^{\text{ext}}(\mathbf{r}, \omega) + \int V(\mathbf{r}, \mathbf{r}') \rho(\mathbf{r}') d^2 \mathbf{r}', \quad (116b)$$

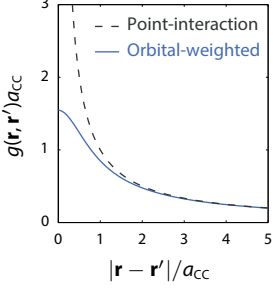
with the Coulomb interaction $V(\mathbf{r}, \mathbf{r}') = 1/4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|$. In the concrete case of a TB-treatment, the relevant coordinates naturally discretize in the site basis $\{|l\rangle\}$ with coordinates $\{\mathbf{r}_l\}$ and wavefunctions $\psi_j(\mathbf{r}_l) \rightarrow \psi_{jl}$. Consequently, all physical quantities

Further and much earlier support is found in the 1996 treatment of Ref. 233, in their (numerical) consideration of arbitrarily terminated nanoribbons.

For brevity, and since there is no chance of confusion, we here omit explicit declaration of in-plane quantities, i.e. we omit ||-subscripts.

The real-space form of χ^0 is central to the linear-response real-space and frequency-formulation of TDDFT, sometimes known as the Lehmann representation [135], which in addition includes the (dynamic) effects of xc-interaction.

Equivalently, the RPA density-density function $\chi^{\text{RPA}} = \chi^0 (1 - e^2 \mathbf{V} \chi^0)^{-1}$ relates the induced charge density to the external perturbation via $\rho = e^2 \chi^{\text{RPA}} \phi^{\text{ext}}$.



Scaled interaction $g(\mathbf{r}, \mathbf{r}')$ in point-like, $|\mathbf{r} - \mathbf{r}'|^{-1}$, and orbital-weighted accounts.

similarly find a discretized form, according to $\chi^0(\mathbf{r}_l, \mathbf{r}_{l'}) \equiv \chi_{ll'}^0$, $V(\mathbf{r}_l, \mathbf{r}_{l'}) \equiv V_{ll'}$, $\rho(\mathbf{r}_l) \equiv \rho_l$, and $\phi^{[\dots]}(\mathbf{r}_l) \equiv \phi_l^{[\dots]}$, while \mathbf{r}' -integrations reduce to l' -summations. In this case, the self-consistent system is reduced to a single matrix equation of the form $\rho = e^2 \chi^0 (\phi^{\text{ext}} + \mathbf{V} \rho)$ for each frequency. Following inversion of this system, the optical response to e.g. plane waves follows from the dipole moment of ρ_l .

The above introduction essentially outlines the entire working principle of the TB-RPA. Nevertheless, at least two further elements deserve mention:

COULOMB INTERACTION In the RPA the Coulomb interaction is strictly speaking a density-weighted Hartree interaction. Accordingly, although the interaction $V(\mathbf{r}_l, \mathbf{r}_{l'}) \equiv g(\mathbf{r}_l, \mathbf{r}_{l'})/4\pi\epsilon_0$ is largely point-like at large site-separations in the sense that $g(\mathbf{r}_l, \mathbf{r}_{l'}) \simeq |\mathbf{r}_l - \mathbf{r}_{l'}|^{-1}$ for $|\mathbf{r}_l - \mathbf{r}_{l'}| \gg a_{\text{CC}} = 1.42 \text{ \AA}$, it is not point-like at short separations. Specifically, the spatial extent of the $2p_z$ orbitals must then be accounted for explicitly. Thongrattanasiri *et al.* [129] performed this accounting by computing the orbital-weighted Hartree interaction using tabulated orbital data [332] – and have kindly communicated the distance-dependence of $g(\mathbf{r}, \mathbf{r}')$ to us. The most critical outcome of the orbital-accounting is a non-divergent self-interaction $g(\mathbf{r}_l, \mathbf{r}_l) \approx 1.55 a_{\text{CC}}^{-1}$ (self-energy of $\approx 15.7 \text{ eV}$).

SCALING AND FOURIER TECHNIQUES For an N -atom system, the direct construction of χ^0 requires a formidable $\mathcal{O}(N^4)$ operations (N^2 elements with N^2 summation terms each). Evidently, for configurations of interest this scaling is prohibitive (a disk, for instance, of radius 10 nm contains ≈ 3000 atoms). The scaling can be improved to $\mathcal{O}(N^3)$ by rewriting Eq. (115) as a frequency-convolution, which can be computed efficiently by the fast Fourier transform [129].

The above frequency-domain method was implemented by Weihua Wang, and is applied in Publications C and E. Subsequently, the author has explored and implemented a complementary but equivalent time-domain approach, explicated further in the following paragraph.

TIME-DOMAIN The RPA is interpretable from a multitude of perspectives. In the time-domain, it emerges as the linear limit of a dynamic Hartree treatment, historically known as time-dependent Hartree–Fock theory [333, 334]. As a consequence, time-propagated Hartree-interacting TB states *provide dynamics identical* to Eqs. (116). To be specific, we consider again a set of (occupied) TB orbitals $|\psi_j\rangle = \sum_l \psi_{jl} |l\rangle$ expressed in an (orthonormal) site-basis $\{|l\rangle\}$, representing Wannier states centered at $\{\mathbf{r}_l\}$. The time-independent ground state is described by Eq. (112), yielding in the present notation the single-particle TB Hamiltonian $\hat{H}^0 = -t_{\text{AB}} \sum_l \sum_{l' \in \text{nn}_l} \hat{c}_l^\dagger \hat{c}_{l'}$ with creation (annihilation) operators \hat{c}_l^\dagger ($\hat{c}_{l'}$) of $|l\rangle$ ($|l'\rangle$) orbitals. The time-dependent states $|\psi_j(t)\rangle = \sum_l \psi_{jl}(t) |l\rangle$ are governed by the self-consistent Hamiltonian

$$\hat{H}(t) = \hat{H}^0 + \hat{H}^{\text{ext}}(t) + \hat{H}^{\text{ind}}(t), \quad (117a)$$

with a perturbation $\hat{H}^l(t) = \hat{H}^{\text{ext}}(t) + \hat{H}^{\text{ind}}(t)$ with external and induced (Hartree) contributions

$$H^{\text{ext}}(\mathbf{r}, t) = -e\phi^{\text{ext}}(\mathbf{r}, t), \quad (117b)$$

$$H^{\text{ind}}(\mathbf{r}, t) = -e\phi^{\text{ind}}(\mathbf{r}, t) = -e \sum_{l'} V(\mathbf{r}, \mathbf{r}_{l'}) \rho_{l'}(t), \quad (117c)$$

with a diagonal TB representation $\hat{H}^{[\dots]}(t) = \sum_l H^{[\dots]}(\mathbf{r}_l, t) |l\rangle \langle l|$ and with induced charge density $\rho_l(t) \equiv -e[n_l(t) - n_l(0)]$ calculable from the time-dependent particle density $n_l(t) = 2 \sum_j^{\text{occ}} f_j |\psi_{jl}(t)|^2$ (summation restricted to occupied orbitals). The states evolve according to the time-dependent Schrödinger equation $i\hbar \partial_t |\psi_j(t)\rangle = \hat{H}(t) |\psi_j(t)\rangle$ [equivalently, as $i\hbar \partial_t \psi_{jl}(t) = \sum_{l'} H_{ll'}(t) \psi_{jl'}(t)$ in a site-representation

We assume a perturbation effective from $t = 0$, such that the ground state density is given by $n_l(0)$.

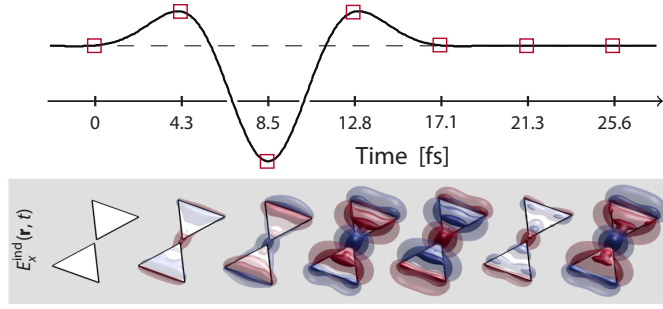


Figure 33: Excitation of an AC-terminated triangular bowtie (sidelength $L = 20$ nm and gap $d_g = 2$ nm) by a normally incident electric field (polarized along the gap-direction) with a fs temporal envelope. Top panel: Field amplitude of the incident excitation pulse as a function of time. Bottom panel: Snapshots of the x -component of the induced electric field at sample-times indicated in red squares. Isosurfaces drawn at 0.1, 0.3, and 0.9 times the incident pulse’s peak amplitude, with sign indicated in color (negative/positive in blue/red) and transparency scaled by magnitude.

with matrix-elements $H_{ll'}(t) \equiv \langle l | \hat{H}(t) | l' \rangle$. Formally, the problem can be inverted by introducing the unitary time-operator $\hat{U}(t, 0) = \exp[-i\hbar^{-1} \int_0^t \hat{H}(t') dt']$ which propagates the states from time $t' = 0$ to t in the manner $|\psi_j(t)\rangle = \hat{U}(t, 0)|\psi_j(0)\rangle$. From a numerical point-of-view, the primary challenge then becomes the faithful representation of \hat{U} in a time-discretized site-representation. Fortunately, this challenge is identical to the one faced in TDDFT, providing us immediately with a selection of well-developed schemes [135, 335]. For completeness, we summarize the technical elements of the scheme adopted in our calculations in Appendix E. From an overall perspective, the method allows the accurate computation of the (occupied) expansion coefficients $\psi_{jl}(t)$ on a discrete time-grid $t_{i+1} = t_i + \delta t$ for sufficiently small δt .

Ultimately, we remain interested predominately in the spectral response, in the form of e.g. absorption cross-sections. This, however, represents no real issue for a time-domain approach: indeed, the spectral response can be obtained by recording the temporal dynamics, say, of the dipole moment $\mathbf{p}(t) \equiv \sum_l \mathbf{r}_l \rho_l(t)$, due to excitation by a temporally short (spectrally broad) “kick”. The spectral response then follows subsequently from a frequency-transform. For illustration, we depict in Figure 33 the temporal response due to such a kick for the specific case of a triangular bowtie.

A single point of contention remains, namely the inclusion of loss through a decay rate $\gamma = 2\eta$. In fact, the above approach entirely neglects decay, which cannot be included in a time-propagated wavefunction approach. Nevertheless, in the linear limit the decay can be approximately accounted for by simply appending a factor $e^{-\eta(t-t_k)}\theta(t-t_k)$ to induced dynamic quantities with t_k indicating a time immediately after the passage of the kick, corresponding effectively to a relaxation-time substitution $\omega \rightarrow \omega + i\eta$. This approach is standard in TDDFT [336]. We caution, however, that while the time-domain scheme in principle is not limited to linear response, that this approximate treatment of loss certainly is. An accurate account of loss in the non-linear regime requires a density-matrix approach, such as the one recently employed in Ref. 337 (with which we have explicitly confirmed consensus in the linear regime).

RESULTS With the formal underpinnings of the TB-RPA explicated, we are now equipped to investigate its predictions. In this paragraph, we discuss results computed by means of the time-domain method; later, in Section 6.2.3 we utilize the frequency-domain method. Before discussing spectral properties, we briefly discuss one attractive opportunity posed by the time-domain method, related to a state-picture of plas-

While this heuristic approach to loss provides an excellent (TDDFT-)account of nonlinearities in e.g. molecules, it is not appropriate to transfer to plasmonic circumstances, where field-enhancement plays a major dynamic role. In our calculations, we use perturbations of sufficient weakness to ensure a linear response.

All time-domain calculations discussed here assume zero temperature, i.e. $T=0$.

mon excitations. Concretely, we track the time-dependent probability $P_j(t)$ which, by projection, gives the occupancy of the j th eigenstate at time t [128]

$$P_j(t) \equiv \sum_{j'} f_{j'} |\langle \psi_{j'}(t) | \psi_j(0) \rangle|^2. \quad (118)$$

In the $t = 0$ limit, P_j reduces to the ground state occupancy $P_j(0) = f_j = \theta(\epsilon_F - \epsilon_j)$. For nonzero times $P_j(t)$ yield the projection of the *total* state onto the j th noninteracting TB state, i.e. it gives the evolution of state occupancy. In Figure 34 we depict the temporal variation of the *change* of occupancy $\delta P_j(t) \equiv P_j(t) - f_j$: negative (positive) values of $\delta P_j(t)$ then indicate depleted (excited) regions relative to the ground state. We make several observations pertaining to Figure 34. First, we observe for both ZZ and AC triangles a periodic sloshing of population with a period ≈ 12 fs, corresponding to a resonance energy ≈ 3.4 eV. Moreover, the population exchange is strongest in the immediate neighborhood of the Fermi energy. These observations exemplify explicitly the interpretation of plasmons as collective and coherent electron-hole pair (EHP) excitations. Finally, we observe with interest that the zero-energy edge states of the ZZ triangle also partake appreciably, with significant excitation of edge state EHPs: they incur additional Landau damping of the plasmon and contribute, as we will see, to a comparatively smaller response in ZZ structures.

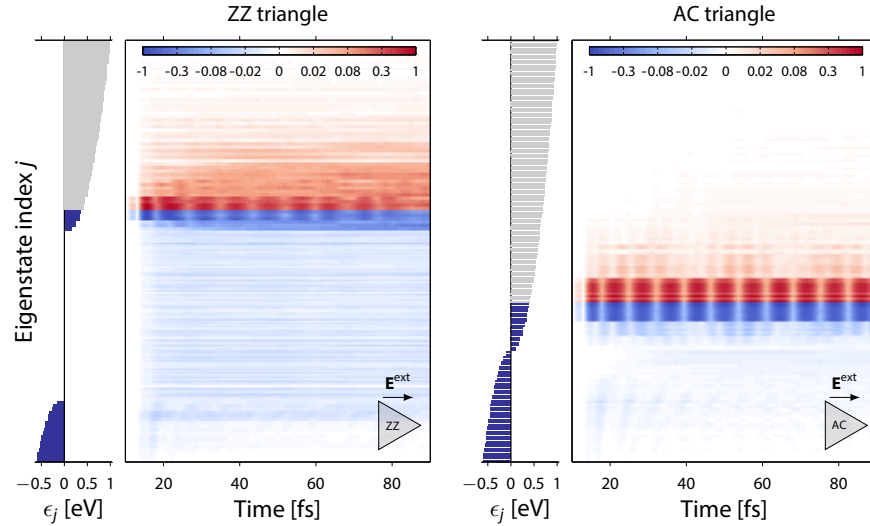


Figure 34: Temporal evolution of the state-overlap $\delta P_j(t) = P_j(t) - f_j$ relative to the ground state for ZZ and AC triangles of sidelength $L = 25$ nm. For each triangle we depict in the left-hand plot the TB equilibrium eigenenergies ϵ_j as a function of eigenstate index j for a selection of states near the Fermi energy $\epsilon_F = 0.4$ eV – equilibrium occupancy is indicated in color (filled/empty in blue/gray). Right-hand colormaps depict the temporal evolution of $\delta P_j(t)$ due to a sharp excitation kick around ~ 15 fs (polarization indicated in inset). The colormap depicts the normalized overlap $\delta P_j(t) / \max |P_j(t)|$ in a nonlinear intensity scale as indicated in the colorbar. Drude decay is not included, i.e. $\eta = 0$.

A meaningful assessment of nonclassical shifts require – in addition to precise quantum predictions – a highly accurate classical approach: crucially, the specialized methods discussed in Section 5.3.2 achieve this.

Next, we consider in Figure 35 the spectral response and size-dispersion of the absorption efficiency in distinct geometries, namely in ZZ and AC triangles and hexagons, and in a mixed-boundary (bond-centered) disk, and compare with the classical predictions. The size-dispersion of the primary peak is extracted in Figure 36. In comparing structures of strict ZZ- and AC-termination we observe, for moderately large sizes, a nonclassical red- and blueshift, respectively, and an absorption reduction in ZZ vis-à-vis AC circumstances. In both cases, the shift decreases with increasing size, tending asymptotically towards the classical predictions. For the mixed-boundary disk, we ob-

serve for both polarizations a response similar to that of the strict ZZ structures, i.e. redshifted and reduced. By our earlier considerations in Section 6.2.1, it is clear that the differences relate to the existence of zero-energy edge states in the ZZ and mixed-boundary structures. Similar observations were made previously in Refs. 129 and 286. In the following section, we pursue a semi-classical explanation of these observations, through a continuum Dirac description of an effectively ZZ-terminated disk.

In the opposite limit of very small sizes, $\sqrt{\text{area}} \lesssim 8$ nm, the plasmon peak is poorly developed, involving just a few single-particle transitions.

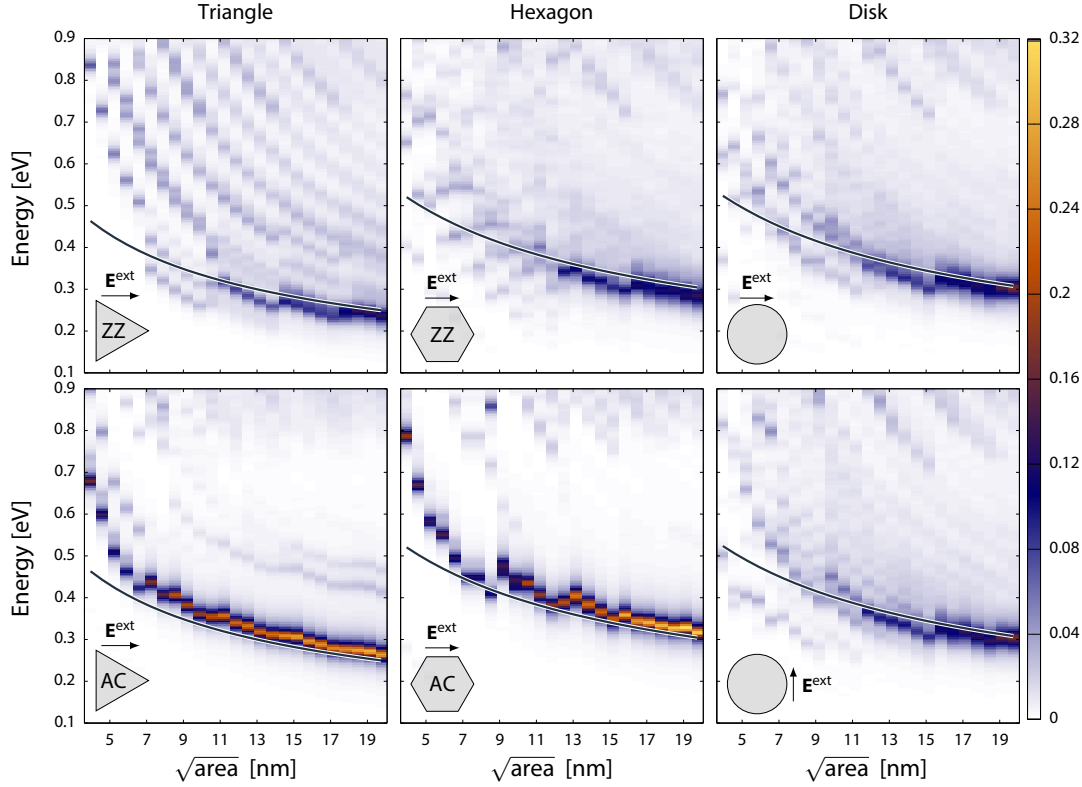


Figure 35: The absorption efficiency, σ_{abs} divided by geometric area, as a function of energy and (square-rooted) area for different geometries, as indicated in insets. In all cases, the structures are doped to $\epsilon_F = 0.4$ eV and a Drude loss $\hbar\gamma = 2\hbar\eta = 24$ meV is included. For the mixed-boundary (bond-centered) disk, two perpendicular polarization directions are examined. The classical intra- and interband dipole resonance is indicated in black lines, computed from the eigenvalues listed in Table 4.

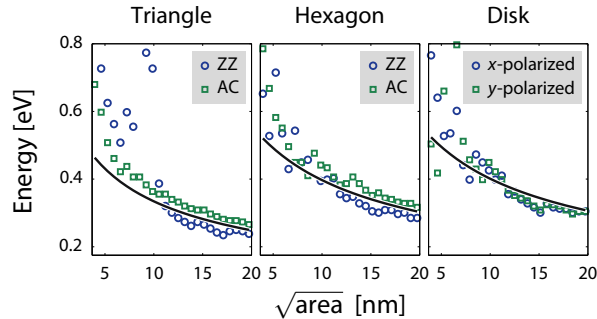


Figure 36: Resonance energies at peak response extracted from Figure 35, indicated in colored markers, compared with corresponding classical predictions in black lines.

Eq. (119) is strictly speaking a Dirac–Weyl equation, differentiated from the Dirac counterpart by the absence of a mass-term. Nevertheless, the distinction is seldom emphasized for graphene.

Kronecker products:

$$\tau_0 \otimes \sigma_x = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$\tau_z \otimes \sigma_y = \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{bmatrix}.$$

We note two additional BCs for general domains

Ω with boundary

normal $\hat{\mathbf{n}} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$.

AC terminations [338] admix

the valleys at $\mathbf{r} \in \partial\Omega$:

$$\psi_{A,B}^+(\mathbf{r})/\psi_{A,B}^-(\mathbf{r}) = -e^{i(\mathbf{K}^- - \mathbf{K}^+) \cdot \mathbf{r}}.$$

Infinite mass confinement

[339] – included by a

term $v_F^2 m(\mathbf{r}) \sigma_z$ with

$m(\mathbf{r}) = \begin{cases} 0, & \mathbf{r} \in \Omega \\ \infty, & \mathbf{r} \notin \Omega \end{cases}$ added

into \hat{H}_D^κ – decouples the

valleys but enforces a

sublattice phase relation

$$\phi_B^\kappa(\mathbf{r})/\phi_A^\kappa(\mathbf{r}) = ie^{i\mathbf{k} \cdot \theta} \text{ at } \mathbf{r} \in \partial\Omega.$$

6.2.3 Continuum Dirac equation approach

When expanded around the $\kappa = \pm 1$ valley points \mathbf{K}^κ the TB description of extended graphene takes the form of a Dirac equation. Depending on the lattice orientation, several notationally distinct – but ultimately equivalent – forms can be achieved. With the conventions of Ref. 223 one obtains a four-spinor Dirac equation, $\hat{H}_D \psi(\mathbf{r}) = \epsilon \psi(\mathbf{r})$, with Dirac Hamiltonian

$$\hat{H}_D = \hbar v_F (\tau_0 \otimes \sigma_x \hat{k}_x + \tau_z \otimes \sigma_y \hat{k}_y), \quad (119)$$

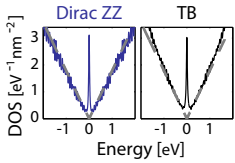
with momentum $\hat{\mathbf{k}} = -i\nabla$ and Pauli matrices τ_i and σ_i belonging to the valley and sublattice subspaces, respectively. The four-spinor $\psi(\mathbf{r}) \equiv [\psi_A^+(\mathbf{r}), \psi_B^+(\mathbf{r}), \psi_B^-(\mathbf{r}), \psi_A^-(\mathbf{r})]^T$ describes the state in sublattice (A and B) and valley (\pm) components, and is assigned the spinor-epithet courtesy of a characteristic π phase-accumulation under closed loops. In the absence of valley mixing the four-spinor equation decouples into a pair of two-spinor equations $\hat{H}_D^\kappa \psi^\kappa(\mathbf{r}) = \epsilon^\kappa \psi^\kappa(\mathbf{r})$ for $\psi^\kappa(\mathbf{r}) = [\psi_A^\kappa(\mathbf{r}), \psi_B^\kappa(\mathbf{r})]^T$ with valley-specific Hamiltonians $\hat{H}_D^+ = \hbar v_F \boldsymbol{\sigma} \cdot \hat{\mathbf{k}}$ and $\hat{H}_D^- = \hbar v_F \boldsymbol{\sigma}^* \cdot \hat{\mathbf{k}}$.

While structural termination is included straightforwardly in a direct TB treatment by the absence of “missing” neighbors, a Dirac treatment requires a more conscientious approach, concretely involving the application of BCs along a continuous domain boundary $\partial\Omega$. These BCs come in several variants depending on the exact type of termination considered [316, 340]. Here we will consider just a single BC, namely one of ZZ type – in Publication E we consider additionally an infinite mass BC. Specifically, imagining the lattice as terminating in ZZ fashion entirely on B-site atoms, we force the wave function to vanish at the “next” row of atoms, i.e. we require $\psi_A^\kappa(\mathbf{r} \in \partial\Omega) = 0$. The disk restriction of radius R is a particularly attractive geometry because it allows an analytical treatment [339, 341–343]. Concretely, as explicated in the Supplementary Material of Publication E, two classes of solutions emerge for the ZZ-BC Dirac disk, namely finite-energy bulk states, $\psi_{ln}^\kappa(\tilde{r}, \theta)$ for $l \in \mathbb{Z}$ and $n \in \mathbb{N}_+$, and zero-energy edge states, ϕ_ℓ^κ for $\ell = 0, 1, \dots, \ell_{\max}$, reading in polar coordinates as

$$\psi_{ln}^\kappa(\tilde{r}, \theta) = \frac{e^{il\theta}}{\sqrt{N_{ln}}} \begin{bmatrix} J_l(\beta_{ln} \tilde{r}) \\ i\kappa J_{l+\kappa}(\beta_{ln} \tilde{r}) e^{i\kappa\theta} \end{bmatrix}, \quad \phi_\ell^\kappa(\tilde{r}, \theta) = \frac{e^{-i\kappa\ell\theta}}{\sqrt{N_\ell}} \begin{bmatrix} 0 \\ \tilde{r}^\ell \end{bmatrix}, \quad (120)$$

with normalization constants $N_{ln} \equiv 2\pi R^2 J_{l+1}(\beta_{ln})$ and $N_\ell \equiv \pi R^2 / (\ell + 1)$, dimensionless radial coordinate $\tilde{r} \equiv r/R$, and with bulk energies $\epsilon_{ln} = \hbar \omega_R \beta_{ln}$ expressed through the characteristic frequency $\omega_R \equiv v_F/R$. The coefficients β_{ln} are dictated by the BC, given in the ZZ case by the zeros of the Bessel function $J_l(\beta_{ln}) \equiv 0$. Finally, a phenomenological angular momentum cutoff ℓ_{\max} is introduced for the edge states, motivated by the fact that TB predicts only a finite number of such states. The cutoff can be estimated from the density of edge states per unit length of zigzag boundary $\simeq 1/3\sqrt{3}a_{CC}$ as derived by Akhmerov and Beenakker [316], yielding an estimated total number of edge states $N_{\max}^{\text{edge}} = 2\pi R/3\sqrt{3}a_{CC}$. After accounting for degeneracies this fixes the angular cutoff by $4(\ell_{\max} + 1) = N_{\max}^{\text{edge}}$. With this adjustment, the DOS of the ZZ Dirac treatment qualitatively replicates that of the TB for equal size disks.

A full nonlocal RPA calculation can be performed on the basis of the spinor states in Eq. (120) by a slight generalization of Eq. (115) to account for the spinor’s non-scalar nature, see Appendix D. We will not discuss the technical details here – they are discussed in the Supplementary Material of Publication E – but instead proceed to discuss the promised semi-classical account of edge states. To do so, we construct a local-response conductivity $\sigma(\omega)$, reminiscent in spirit of a dipole-approximation,



DOS in $R=10$ nm disk, with 24 meV broadening, and low-energy bulk DOS in gray.

associated with the finite structure. Specifically, for an x -polarized perturbation the noted conductivity follows from a Fermi golden rule consideration [57]

$$\sigma(\omega) = \frac{2ie^2\omega}{\mathcal{A}} \sum_{jj'} (f_j - f_{j'}) \frac{|\langle \psi_j | x | \psi_{j'} \rangle|^2}{\epsilon_j - \epsilon_{j'} - \hbar(\omega + i\eta)}, \quad (121)$$

with the composite index j extending over both bulk and edge states, i.e. $\{j\} = \{l, n, \ell\}$. When expanded in summation, the conductivity naturally splits into two contribution: one due to bulk-bulk transitions and one due to bulk-edge transitions. The former necessarily tends to the conventional extended-graphene bulk conductivity $\sigma_b(\omega)$ in the large radius limit, while the latter, henceforth denoted $\sigma_E(\omega)$, contain the sought effects of edge states. This edge-state conductivity assumes the form of a summation over Bessel zeros

$$\sigma_E(\omega) = -\frac{16ie^2}{\pi\hbar} \frac{\omega}{\omega_R} \sum_{\ell=0}^{\ell_{\max}} \sum_{n}^{\hbar\omega_R \beta_{\ell n} > \epsilon_F} \frac{\ell+1}{\beta_{\ell n}^5 \left[1 - \left(\frac{\omega+i\eta}{\beta_{\ell n} \omega_R}\right)^2\right]}, \quad (122)$$

written here in the low-temperature limit for positive Fermi energies. Remarkably, from this formidable expression we can derive a simple asymptotic limit in the large radius $R \rightarrow \infty$ limit. We omit the somewhat cumbersome details here (though explicated in Publication E) and simply state the final result, denoted $\sigma_E^\infty(\omega)$, in the low-loss $\eta \rightarrow 0^+$ limit

$$\sigma_E^\infty(\omega) = \frac{8e^2}{3\pi^2\hbar} \frac{\omega_R}{\omega} \left[i \ln \left| \frac{\epsilon_F^2 - \hbar^2 \omega^2}{\epsilon_F^2} \right| + \pi \theta(\hbar\omega - \epsilon_F) \right]. \quad (123)$$

This result carries a substantial physical appeal, since it explicitly demonstrates that the accounting of edge states opens a nonclassical dispersive channel with a characteristic size-dependence $\omega_R \equiv v_F/R$. Moreover, Eq. (123) reveals the emergence, besides the usual vertical interband Landau transitions for $\hbar\omega \geq 2\epsilon_F$, of nonvertical edge-to-bulk transitions for $\hbar\omega \geq \epsilon_F$ with the necessary momentum $\sim k_F$ supplied by the structural truncation $\sim 1/R$.

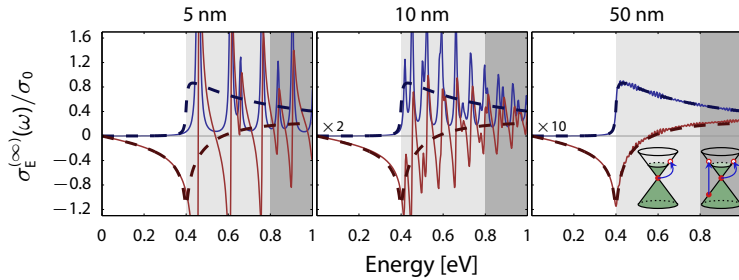


Figure 37: Full and large-radius asymptotic edge-state conductivities $\sigma_E^{(\infty)}(\omega)$ (normalized to $\sigma_0 \equiv e^2/4\hbar$) in full and bold dashed lines, respectively, with both real (blue) and imaginary (red) parts depicted. The nanodisk is assumed doped to $\epsilon_F = 0.4$ eV, with nonzero broadening $2\hbar\eta = 12$ meV, and with disk radii indicated in titles (note the scaling factor in the two right-hand panels). The region of edge-to-bulk transitions is indicated in light gray shading, while the region of concurrent edge-to-bulk and interband transitions feature in dark gray; both illustrated schematically by insets.

In Figure 37 we compare the full and asymptotic expressions for the edge-state conductivity $\sigma_E^{(\infty)}(\omega)$, Eqs. (122) and (123), for disks of different radii. A noticeable difference between the two is evident for small radii in the region $\hbar\omega > \epsilon_F$, where $\sigma_E(\omega)$

In words, this double series sums all dipole-allowed transitions between filled edge states and empty bulk states.

exhibits peaks at the discrete transitional energies $\hbar\omega \simeq \epsilon_{\ell n}$. Nevertheless, for increasing radii and concomitantly decreasing energy-spacing these discrete peaks eventually merge, reproducing asymptotically the smooth behavior predicted by $\sigma_E^\infty(\omega)$. Outside the region of direct transitions, i.e. for $\hbar\omega < \epsilon_F$, we observe a remarkably good agreement between full and asymptotic expressions: accordingly, for resonances below the Fermi energy, the asymptotic form can be taken as fully adequate. We highlight that the sign of $\text{Im} \sigma_E^\infty(\omega)$ is negative for $\hbar\omega < \frac{3}{2}\epsilon_F$, consequently incurring a resonance redshift cf. Eq. (95b). The magnitude of this shift is size-dependent, diminishing as the radius is increased since $\sigma_E^\infty \propto 1/R$. Nevertheless, even relatively large disks exhibit appreciable edge contributions to the conductivity; for instance, a maximal contribution $\approx 0.4\sigma_0$ is observed for the $R = 10$ nm disk, comparable to the contributions of interband bulk transitions, which amount roughly to σ_0 .

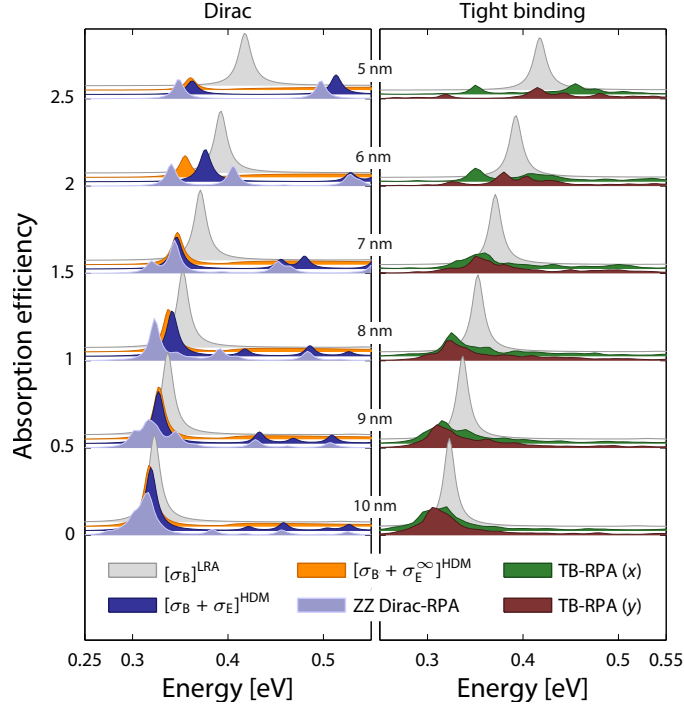


Figure 38: Absorption cross-sectional efficiency for nanodisks ($\epsilon_F = 0.4$ eV, $2\hbar\eta = 12$ meV, and $T = 300$ K) of different radii, as indicated, illuminated under normal incidence. Several distinct methods of calculation are compared: classical LRA bulk-response ($[\sigma_B]^{\text{LRA}}$), semi-classical edge-state conductivity corrected calculations including hydrodynamic response ($[\sigma_B + \sigma_E^{(\infty)}]^{\text{HDM}}$), frequency-domain RPA calculations built from Dirac states with a ZZ-BC (ZZ Dirac-RPA) as well as TB states (TB-RPA). TB calculations model a bond-centered disk, and examine also the variation with excitation polarization (along x and y). Spectra for different radii are offset by 0.5 units, while individual spectra at equal radii are offset by 0.025 units.

Finally, with all the tools now developed, we consider in Figure 38 a practical comparison of predictions in classical, semi-classical, and quantum frameworks. Specifically, we examine the absorption efficiency in nanodisks of varying radii. Besides a classical calculation, built from the bulk conductivity $\sigma_B(\omega)$, and quantum calculations with the frequency-domain RPA built from either Dirac ZZ or TB states, we include two semi-classical calculations. Specifically, in recognizing the equal footing of hydrodynamic and edge-state corrections, cf. an equal corrective behavior $\propto v_F/R$, we depict calculations including, in addition to $\sigma_B(\omega)$, both an edge-term $\sigma_E^{(\infty)}(\omega)$ and a hydrodynamic extension, see Eq. (107). We will not delve on the minute comparison of spectra, which is discussed in detail in Publication E, but pause to focus on

the interpretation of these semi-classical calculations. Firstly, we observe a qualitative agreement with the higher echelons of Dirac ZZ- and TB-RPA calculations for moderately large radii, expressed chiefly in terms of a resonance redshift and broadening relative to the classical predictions. Importantly, an overall redshift is replicated by the joint actions of $\sigma_E^{(\infty)}(\omega)$ and the HDM, even though they act to incur shifts in opposite directions, red and blue, respectively. The clear reduction in absorption efficiency displayed in TB-RPA – as well as the polarization dependence – is not reproduced in the continuum descriptions, being chiefly due to the neglect of atomic structure, which effectively facilitate a larger set of dipole-allowed Landau transitions.

In an attempt to distill the above considerations to their core essence, we conclude by the following retrospective observations, which qualitatively explain the features previously observed in [Figures 35](#) and [36](#):

ARMCHAIR TERMINATION With no edge states, the primary semi-classical features are due to nonlocal effects, well-modeled by a HDM approach, incurring resonance blueshifts.

ZIGZAG TERMINATION Edge states exist, which reduce the imaginary part of the total conductivity, thereby incurring a redshift. This redshift exceeds the blueshift due to nonlocality, yielding ultimately a net redshift.

MIXED TERMINATION Although fewer, edge states persist, localized near terminations of ZZ kind; accordingly, both nonlocal and edge-state effects apply. For the concrete case of a nanodisk the outcome is a net redshift.

In all cases, level-quantization also plays a role, particularly in the very small size limit, fracturing there the main plasmon peak into several subbands. In the semiclassical limit, it conceivably adds also to the nonlocal blueshift, though, we expect, to a lesser degree.

6.3 NONLINEAR RESPONSE

In this last section of the chapter we depart from the otherwise nonclassical theme, stepping instead into a realm which might be designated rather as nonconventional graphene plasmonics. Concretely, we here go beyond the conventional ansatz of linearity, and study the impact of a classical Kerr nonlinearity on the plasmonic properties of a graphene nanoribbon, as recently treated in [Publication A](#) [275]. Rather than retracing all elements of this publication, we offer here a condensed view, focusing primarily on the overall role of nonlinear corrections to plasmonic response in graphene.

Before doing so, let us briefly motivate our intentions. The topic of nonlinear plasmonics [344] constitutes an exciting subfield of optical nonlinearities [345, 346]. In many ways, the joining of plasmonics with nonlinear effects poses an obvious opportunity, because the extreme field enhancements of plasmons offer a natural approach to the attainment of large effective nonlinear interaction with modest input powers. Here, graphene offers an exceptionally interesting platform for two reasons: firstly, courtesy of its large electron mobilities [34, 248, 347], graphene plasmonics boasts an extraordinary potential for field enhancements compared with other plasmonic platforms. Secondly, graphene's nonlinearities are simply different from metals' [348–351]. Specifically, although a small intraband metal-nonlinearity exists due to nonlocal ponderomotive forces [352], allowing surface nonlinearities, the primary source of metal-nonlinearity are due to interband transitions [345]. Contrary to this, graphene exhibits a large intrinsic intraband nonlinearity due to its Dirac dispersion. Accordingly, both plasmonic and nonlinear properties of graphene originate in the same electronic features. With these matters in mind, it is not surprising that nonlinear graphene plasmonics has bloomed explosively in the last few years [337, 351, 353–361].

For the Kerr component, i.e. the third-order response oscillating at the perturbing frequency, the mentioned intraband nonlinearity can be derived from the Boltzmann equation in the manner lucidly outlined by Peres *et al.* [351], allowing an intensity-dependent conductivity (explicit declaration of frequency-dependence omitted)

The Kerr correction is of a self-focusing type, and accordingly must be augmented to include a saturating mechanism, or suffer runaway focusing (an issue familiar from optical waveguide modeling [362]). We achieve this in practice by using the two-level saturation model consistent with Eq. (12.4) plus a phenomenological accounting of two-photon absorption [360].

$$\sigma(\mathbf{r}) = \sigma_{(1)} \left[1 - \frac{|\mathbf{E}(\mathbf{r})|^2}{E_{(3)}^2} \right], \quad (124)$$

with linear response $\sigma_{(1)}(\omega) \equiv \sigma_{\text{intra}}(\omega) = ie^2 \epsilon_F / \pi \hbar^2 (\omega + i\gamma)$, approximated by the intraband response, and with a third-order characteristic field $E_{(3)} \equiv (8\tilde{\omega}_{(3)}^2) / (9\omega^2) E_{\text{sat}}^2$ linearly related to the saturation field $E_{\text{sat}} \equiv \epsilon_F \omega / ev_F$. We restrict ourselves here to the intraband approximation; the accounting of nonlinear interband effects is complicated by several singularities [350], the normalization of which require a band-opening at the Dirac point [363], i.e. a small mass-term in the Dirac Hamiltonian e.g. in the form of a spin-orbit interaction. Similarly, we restore a fully classical outlook, ignoring matters of nonlocality and lattice-termination – we note, however, that these aspects have recently been approached in a nonlinear context by a time-domain density-matrix formalism [337, 361], analogous to the time-domain TB-RPA discussed previously.

With the adoption of the nonlinear conductivity of Eq. (124), the governing nonretarded potential equations, Eqs. (95), become nonlinear through influence of the (inhomogeneous) normalized conductivity $f(\mathbf{r}) = \sigma(\mathbf{r})/\sigma_{(1)}$, which attains an explicit dependence on $\phi(\mathbf{r})$ via $\mathbf{E}(\mathbf{r})$. In practice, the resulting nonlinear problem can be solved by self-consistent iteration of an inhomogeneous linear equation, as described in detail in Publication A. In Figure 39 we examine the dependence of the ribbon-integrated

The third-order field is defined through a loss-modified frequency $\tilde{\omega}_{(3)}^2 \equiv (\omega + \frac{1}{2}\gamma)(\omega - i\gamma)$. We emphasize that this frequency is not just $(\omega + \frac{1}{2}\gamma)^2$ as a linear time-relaxation approximation would suggest: this underscores our previous warning that the time-domain TB-RPA approach fails in the nonlinear regime, since it does not provide a dynamic accounting of loss.

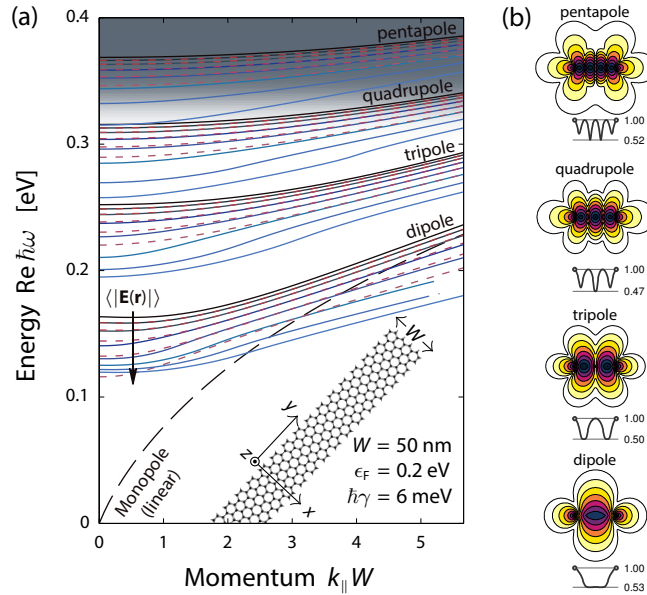


Figure 39: (a) Dispersion relation of an individual graphene nanoribbon (setup-parameters indicated in inset) under varying ribbon-averaged field strengths $\langle |\mathbf{E}(\mathbf{r})| \rangle$, ranging from negligible (black), i.e. linear, through 1×10^5 V/cm (darkest blue) to 4×10^5 V/cm (lightest blue) in steps of 0.5×10^5 V/cm increasing along the arrow. The perturbative result of Eq. (124) is indicated in dashed red lines for the first five field strengths. Mode titles (monopole, dipole, tripole, etc.) reference the nodal count of the induced charge; only the linear dispersion is shown for the monopole. The neglect of interband conductivity terms entails an overestimation of the eigenenergies; this is highlighted explicitly by the shaded region $\hbar\omega/\epsilon_F \gtrsim 1.6671$. (b) Field amplitude $|\mathbf{E}(x, z)|$ depicted by contours for the case $\langle |\mathbf{E}(\mathbf{r})| \rangle = 4 \times 10^5$ V/cm and $k_{\parallel} W = 0$. Contours are separated by factors 1.5, 1.75, 2, and 2.25, and colors range from minimum to maximum in light to dark. Sparklines below the maps display the variation of the normalized conductivity $f(\mathbf{r})$ along the ribbon width (and accordingly indicate the spatial scale of the contour maps).

field strength $\langle |\mathbf{E}(\mathbf{r})| \rangle \equiv W^{-1} \int_{-W/2}^{W/2} |\mathbf{E}(x)| dx$ on the eigenenergies $\hbar\omega_n(k_{\parallel})$ of an individual graphene nanoribbon. The primary impact of increased field strength, and hence mounting nonlinearity, is a redshift relative to the linear predictions. At a basic level, this is readily appreciable from the negative sign of the Kerr coefficient. At a more precise level, the shift can be estimated by a perturbative approach. Specifically, for an arbitrary geometry Ω , the nonlinear eigenenergies $\hbar\omega_v$ can be approximated from the *linear* eigenenergies and eigenfields, $\hbar\omega_v^{(0)}$ and $\mathbf{E}_v^{(0)}$, via

$$\hbar\omega_v \simeq \hbar\omega_v^{(0)} \sqrt{1 - \frac{9}{8} \frac{\langle |\mathbf{E}_v^{(0)}(\mathbf{r})|^4 \rangle}{\langle |\mathbf{E}_v^{(0)}(\mathbf{r})|^2 \rangle E_{\text{sat}}^2}}, \quad (125)$$

with averages performed over the domain Ω . The approximation is illustrated along the full self-consistent solutions in [Figure 39](#) in dashed red: evidently, for modest nonlinear strengths, the agreement is excellent – although the quality eventually deteriorates as the field strength is increased and the self-consistent nature of the interaction plays a larger role. The perturbative result also demonstrates in a clear manner the role played by field-inhomogeneity: by introducing a modal inhomogeneity parameter $\kappa_v \equiv \langle |\mathbf{E}_v^{(0)}(\mathbf{r})|^4 \rangle / \langle |\mathbf{E}_v^{(0)}(\mathbf{r})|^2 \rangle^2$, which has the property $\kappa_v \geq 1$ cf. the Cauchy–Schwarz inequality, we recognize that inhomogeneity *enhances* the nonlinear shift which is $\propto \kappa_v \langle |\mathbf{E}_v^{(0)}|^2 \rangle / E_{\text{sat}}^2$. This observation has bearing on recent dipole-considerations of nonlinear coupled nanodisk plasmons [356], which overlooked this effect. Additionally, a Taylor expansion of [Eq. \(125\)](#) provides a quantitative explanation to the recent numerical observations, made in the study of ribbon notch filters, of linear resonance shifts with intensity in the low-field regime [357].

With a mind to the overall focus of the chapter on *corrections* to the conventional behavior of graphene plasmonics, we choose to ultimately end our treatment of nonlinearities here. We emphasize, however, that our discussions in [Publication A](#) venture significantly beyond this; concretely by a treatment of plasmonic bistabilities under plane-wave excitation, as well as an examination of solitons propagation studied by the 1D nonlinear Schrödinger equation.

6.4 SUMMARY

In this chapter, we have discussed elements of graphene plasmonics which lie beyond the classical description of [Chapter 5](#). First, in [Section 6.1](#), we revisited the HDM previously studied for metals, and discussed its adaptation to graphene. The primary impact of the HDM in graphene plasmonics, namely a semiclassical blueshift, qualitatively mirrored the HDM correction in metals; apart from prefactors, the main distinction arises due to different dimensionalities, incurring a size-dependent shift $\delta\omega \propto R^{-3/2}$ and $\delta\omega \propto R^{-1/2}$ in the former and latter case, respectively. Embracing more fully a quantum perspective of plasmonics, we considered in [Section 6.2](#) an RPA approach built from a discrete electronic eigenspectrum. Following a brief account in [Section 6.2.1](#) of the TB single-particle properties of finite nanostructures, emphasizing particularly the emergence of edge states in regions of local sublattice asymmetry, we proceeded in [Section 6.2.2](#) to the explication and investigation of the TB-RPA approach, in both frequency- and time-domain formulations. In an effort to extract a clear physical understanding of observations of TB-RPA lattice-termination-dependent response, we closed our nonclassical considerations in [Section 6.2.3](#) by investigating the Dirac equation applied to nanodisks. By an extensive comparison of classical, semi-classical, and quantum approaches to response, we established that the primary quantum corrections arise from the competing interplay of nonlocality, edge-states, and level-quantization. In the two former cases, semi-classical accounts

Eq. (125) is a special case of a general result derived in [Publication A](#): for a nonperturbed setup $\{\Omega, f^{(0)}\}$ with solutions $\{\lambda_v^{(0)}, \mathbf{E}_v^{(0)}\}$, subjected to a small perturbation $f = f^{(0)} + f^{(1)}$, the perturbed eigenvalues $\lambda_v \simeq \lambda_v^{(0)} + \lambda_v^{(1)} + \dots$ attain a first-order shift $\lambda_v^{(1)} \simeq \lambda_v^{(0)} \frac{\langle \mathbf{E}_v^{(0)} | f^{(1)} | \mathbf{E}_v^{(0)} \rangle}{\langle \mathbf{E}_v^{(0)} | f^{(0)} | \mathbf{E}_v^{(0)} \rangle}$.

n	$\kappa_n(0)$	$\kappa_n(1)$
1	1.31	1.17
2	1.41	1.36
3	1.45	1.42
4	1.46	1.45

Nanoribbon inhomogeneity parameters $\kappa_n(k_{\parallel} W)$ for the dipole, tripole, etc. ($n=1, 2, \dots$).

can be obtained from the HDM and a dipole-approximated edge-state conductivity, respectively. Lastly, in [Section 6.3](#) we examined by numerical and analytical means the impact of a (classical) intraband Kerr nonlinearity, leading to substantial redshifts for field strengths comparable with a characteristic saturation field E_{sat} . Moreover, we unveiled the significant role played by field inhomogeneity for the total nonlinearity.

7

OUTLOOK AND CONCLUSIONS

This thesis has concerned itself with both classical, semi-classical, and quantum aspects of plasmonics in three- and two-dimensional nanostructures, in [Parts I and II](#), respectively. The thesis reflects the attempt at a cohesive and maximally clear exposition of the main results obtained in the course of this PhD project. Having already summarized the content of each chapter at their close, we discuss here only a subset of the matters treated. Our main focus, here at the curtain fall, is rather the interjection of a few last perspectives.

Setting the stage, we commenced our study in three dimensions in [Chapter 2](#) by introducing the key aspects of plasmonics. As we have had opportunity to observe on several occasions, the ancestry of plasmonics predate the present-day fascination with nanotechnology by several decades [3]. One may well wonder then, as others have done [364], where the present-day fascination stems from. In the author's view, contemporary interest in plasmonics is to a large extent fueled by the great leaps in fabrication and experimental capabilities gradually introduced by nanotechnology. From a theoretical perspective, the fundamental understanding of plasmonics today does not deviate far from its earliest accounts [7, 8]. The range and wealth of perspectives and applications developed since then, however, certainly does diverge from the initial path. This is especially true for the subbranch of localized surface plasmonics. Their features derive, as we have discussed, from the interplay between material properties and geometry, the latter of which is reducible into two parts: shape and scale.

This project has, as one of its core elements, pursued the exploration of the last element, scale. More specifically, we centered in [Chapter 3](#) our focus on the small-scale limit, with a particular emphasis on deviations from the classical picture which predicts simple scale-invariant behavior in the small-scale (nonretarded) limit. Explicitly, a break with classicality is expected on physical grounds as the optically relevant distances are reduced towards those of the plasmon-supporting electron gas. The exact character of this breakdown, its constituents, their interplay, and final net effect, remains a formidable theoretical challenge. In the extreme small-scale limit, i.e. the few-atom limit, the problem can be approached by *ab initio* methods, such as time-dependent density functional theory. In the intermediate mesoscopic range, i.e. the few- to several-nm range, we require, both for computational reasons and for the sake of physical understanding, the construction of semi-classical accounts.

Although, in reality, several nonclassical effects conspire in unison, we have here discussed just one: namely nonlocality, as described by a semi-classical hydrodynamic model. This model remedies a number of shortcomings of the conventional approach, and e.g. provides a qualitatively correct account of the nonclassical blueshifts observed in certain metals [25, 26, 174]. Nevertheless, because it accounts for only *one* aspect of nonclassicality, it is not a predictive model, as made obvious by measurements of nonclassical redshifts [175, 176]. Even so, the hydrodynamic description offers several insights, particularly in analytical aspects, whose consideration might in time suggest a more complete understanding. We highlight in this respect the considerations of [Section 3.6](#), wherein we proposed a generalization to arbitrary structures of the Feibelman *d*-parameter approach [54], which offers a versatile and nimble perspective on nonclassical effects. The salient appeal of this proposition, is the clear separation of scale, shape, and electronic properties. In looking forward, the finalization of these ideas constitutes one substantial desire of the author: as a teaser of further developments,

we note that the nonclassical geometric correction of Eqs. (68) finds an appealing interpretation as a ratio of classical energy terms.

In relinquishing a single dimension, and thus shifting to the topic of Part II – two-dimensional plasmonics – deviations from the three-dimensional understanding arise. The two-dimensional platform considered here, graphene, in addition hold novelties of its own, owed to its unique linear band-dispersion, which we discussed in Chapter 4. The classical plasmonic resonances of graphene, the topic of Chapter 5, deviate qualitatively from their three-dimensional metallic counterparts in two essential aspects: first, for dimensional reasons, by its momentum dependence $\omega \propto k_{\parallel}^{1/2}$, and second, due to graphene's band-structure, by its unique doping dependence $\omega \propto n_0^{1/4}$ and considerable tunability. This tunability – in conjunction with the quantitative differences of lower excitation energies, residing in the THz rather than optical domain, and potentially exceedingly low-loss attributes – does, in the author's view, justify in large parts the renewed interest in two-dimensional plasmonics brought about by graphene. Nevertheless, it seems to some extent that the great enthusiasm with which graphene plasmonics has been explored, occasionally has come at the cost of a historical blindness to earlier efforts in two-dimensional plasmonics. The benefits of looking back are significant; certainly, this thesis owes a debt of gratitude particularly to the earlier considerations of nanostructured two-dimensional plasmonics [274, 283]. We hope that our attempt at the explication of some of these ideas for nanostructures might highlight some of these similarities; and in addition underscore the comparative simplicity with which the classical properties of localized graphene plasmons can be understood.

Nevertheless, although graphene plasmonics in this sense exhibits strong similarities with other, and often older platforms, it does exhibit important unique traits of its own, as discussed in the above. Additional graphene-centric traits without counterpart are revealed in the nonclassical regime, as discussed in Chapter 6. For instance, the existence of edge states, owed to graphene's bipartite lattice, provide us with a captivating example of disparate response from macroscopically similar nanostructures, whose geometric distinction is only apparent at the microscopic level in the form of the termination configuration. Extracting a representative semi-classical account of such effects is a challenging and, fortunately, interesting problem. We have suggested one approach here, in the form of an edge-state conductivity, and shown that it, combined with a hydrodynamic consideration, yields predictions in qualitative agreement with full microscopic calculations. The rigorous extension of these considerations to arbitrary geometries remains an open problem. Speculating, we suggest tentatively that a general solution must extend beyond the kind of dipole-approximation applied here, e.g. by the introduction of a boundary-localized correction, taking, perhaps, the form of a dynamic optical boundary condition. Considering the similarities between this proposition and Feibelman's approach for the three-dimensional half-space, it appears that such ventures might find their resolution in the two-dimensional half-sheet.

The last element treated in this thesis, nonlinearity in graphene plasmonics, also in Chapter 6, constitutes a final example of (intensity-dependent) corrections to the conventional (linear) approach. The investigation of these corrections pose an exciting experimental challenge, which should be surmountable with current capabilities. At the undoubtedly more exacting and burdensome end of the spectrum, remains high-field questions related to the predicted emergence of plasmonic bistable behavior in nanoribbons [275].

With the end approaching, we offer an outlook on open problems pertaining to the subject matter discussed here. One critical omission in our treatment pertains to the coupling of nanostructures, particularly in scenarios with sub-nm gaps. The exact role of electron spill-out remains unsettled in the sense that experimental observa-

tions [365] find qualitative explanations in disparate models, which either include a classical account of spill-out overlap [83, 366], or include the type of Kreibig-related nonlocality discussed in Section 3.2.3 [86, 87]. It has not been the purpose of this thesis to explore this class of problems: going forward, however, we observe its relevance as a fundamental test-bed for the quality of general semi-classical corrective schemes.

More generally, the pursuit of ever smaller scales inevitably raises questions about the physical character of excitations in ultra-small structures [367]. For instance, is it sufficient for an excitation to be ‘plasmonic’ that its properties scale with the strength of the Coulomb interaction [328]? Or else, that an appreciable fraction of states participate significantly in the sloshing of induced charge [128]? What, if any, are the differences from conventional molecular excited states? The emerging topic of molecular plasmonics [368] accentuates the need for a decisive and quantitative answer to this fundamental question.

Ultimately, the worth of theoretical constructs can be gauged only by comparison with experiments. This is true for both classical and nonclassical plasmonics. The former, of course, is today well-established in essence. Nevertheless, despite its mature roots, surprises still remain, as exemplified e.g. by our observation of multipole plasmons in nanospheres [215] in Section 3.5. In the nonclassical case, experimental data is less abundant and certainly more taxing to obtain and interpret. Theoretical efforts, such as the present, are naturally left wanting. It is the hope of the author that the kind of investigations reported here, might contribute to the continued stimulation and fueling of experimental efforts in this direction.

Part III

APPENDICES

A

NONRETARDED INTERACTION OF FAST ELECTRONS WITH GRAPHENE

The calculation featured on the cover page considers the interaction of a fast electron with graphene, with the former normally incident upon the latter. This scenario was implicitly treated in the nonretarded limit by García de Abajo [22] for an electron traversing an interface between two bulk media. The extension to account for a graphene layer between the two media requires only modification of the appropriate reflection and transmission coefficients. For completeness, and because it is a simple and instructive example of interaction between graphene and an external, non-planar field, we here provide the details of the approach.

A.1 POTENTIAL AND SCREENED INTERACTION

The central component of the calculation relies on an appropriate expansion of the field associated with the fast electron. We consider the problem in the nonretarded limit, and treat the electron as a traveling point charge, moving normal to the interface along $\hat{\mathbf{z}}$ with (positive) velocity v and associated charge distribution $\rho(\mathbf{r}, t) = -e\delta(z - vt)\delta(\mathbf{r}_\parallel)$, where \mathbf{r}_\parallel denotes the in-plane coordinate along $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$. In a bulk medium of dielectric constant ε , i.e. in the absence of an interface, the associated field is obtained from the frequency-transform of $\rho(\mathbf{r}, t)$ via Coulomb's law:

$$\phi_0(\mathbf{r}, \omega) = \frac{-e}{4\pi\varepsilon_0 v} \int_{-\infty}^{\infty} \frac{e^{ik_v z'}}{\varepsilon|\mathbf{r} - z'\hat{\mathbf{z}}|} dz' = \frac{-e}{2\pi\varepsilon_0 \varepsilon v} K_0(k_v r_\parallel) e^{ik_v z}, \quad (126)$$

with momentum $k_v = \omega/v$ and modified Bessel function K_0 . The aforementioned expansion reformulates the above result, by expanding the Coulomb interaction $1/|\mathbf{r} - \mathbf{r}'|$ in the parallel momentum \mathbf{k}_\parallel , which is conserved under reflection due to translational symmetry. Indeed, for the bare field this can be accomplished by introducing the un-screened interaction

$$W_0(\mathbf{r}_\parallel, \mathbf{r}'_\parallel; z, z') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} = \int \frac{W_0(k_\parallel; z, z') e^{i\mathbf{k}_\parallel \cdot (\mathbf{r}_\parallel - \mathbf{r}'_\parallel)}}{(2\pi)^2}, \quad (127a)$$

$$W_0(k_\parallel; z, z') \equiv \frac{2\pi}{k_\parallel} e^{-k_\parallel |z - z'|}, \quad (127b)$$

such that

$$\phi_0(\mathbf{r}, \omega) = \frac{-e}{4\pi\varepsilon_0 v} \int_{-\infty}^{\infty} W(\mathbf{r}_\parallel, \mathbf{0}_\parallel; z, z') e^{ik_v z'} dz'. \quad (128)$$

Two integral relations are essential for the derivations included in this section, namely:

$$\begin{aligned} & \int_0^{2\pi} \exp(ik_\parallel r_\parallel \cos \phi) d\phi \\ &= 2\pi J_0(k_\parallel r_\parallel) \text{ and} \\ & \int_0^\infty \exp(-k_\parallel |z|) J_0(k_\parallel r_\parallel) dk_\parallel \\ &= (r_\parallel^2 + z^2)^{-1/2} [369]. \end{aligned}$$

By extension, the total field can be expressed by an analogous screened interaction, W , which accounts for reflection, transmission, and propagation of the k_{\parallel} -components

$$W(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; z, z') = \int W(k_{\parallel}; z, z') e^{i\mathbf{k}_{\parallel} \cdot (\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel})} \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2}, \quad (129a)$$

$$W(k_{\parallel}; z, z') \equiv \frac{2\pi}{k_{\parallel}} \begin{cases} \varepsilon_1^{-1} [e^{-k_{\parallel}|z-z'|} + r_{12} e^{-k_{\parallel}(z+z')}] & \text{for } z, z' > 0 \\ \varepsilon_2^{-1} [e^{-k_{\parallel}|z-z'|} + r_{21} e^{k_{\parallel}(z+z')}] & \text{for } z, z' < 0 \\ \varepsilon_2^{-1} t_{21} e^{-k_{\parallel}(z-z')} & \text{for } z > 0, z' < 0 \\ \varepsilon_1^{-1} t_{12} e^{-k_{\parallel}(-z+z')} & \text{for } z < 0, z' > 0 \end{cases}, \quad (129b)$$

with reflection and transmission coefficients r_{ij} and t_{ij} , respectively, for the potential (in the form $\phi_{\text{in}} = e^{ik_{\parallel}x} e^{k_{\parallel}z}$, $\phi_{\text{ref}} = r e^{ik_{\parallel}x} e^{-k_{\parallel}z}$, and $\phi_{\text{tra}} = t e^{ik_{\parallel}x} e^{k_{\parallel}z}$) under incidence from region i onto region j with associated dielectric functions ε_i and ε_j [with $\varepsilon(z) = \varepsilon_1 \theta(z) + \varepsilon_2 \theta(-z)$]. Accounting also for a graphene-sheet at $z = 0$ with in-plane conductivity σ , they are given and interrelated by

$$r_{ij} = -\frac{\varepsilon_j - \varepsilon_i + g}{\varepsilon_j + \varepsilon_i + g}, \quad t_{ij} = \frac{2\varepsilon_i}{\varepsilon_j + \varepsilon_i + g}, \quad 1 + r_{ij} = t_{ij}, \quad g \equiv \frac{i\sigma k_{\parallel}}{\varepsilon_0 \omega}, \quad (130)$$

which, we remind, are both momentum and frequency dependent. The total potential follows from the analogous form of Eq. (128), which can be reduced to

$$\begin{aligned} \phi(\mathbf{r}, \omega) &= \frac{-e}{4\pi\varepsilon_0 v} \int_{-\infty}^{\infty} W(\mathbf{r}_{\parallel}, \mathbf{0}_{\parallel}; z, z') e^{ik_{\parallel}z'} dz' \\ &= \frac{-e}{2(2\pi)^3 \varepsilon_0 v} \int_0^{\infty} k_{\parallel} \int_0^{2\pi} e^{ik_{\parallel}r_{\parallel} \cos \phi} d\phi \left[\int_{-\infty}^{\infty} W(k_{\parallel}; z, z') e^{ik_{\parallel}z'} dz' \right] dk_{\parallel} \\ &= \frac{-e}{4\pi\varepsilon_0 v} \int_0^{\infty} J_0(k_{\parallel}r_{\parallel}) F(k_{\parallel}, k_{\nu}; z) dk_{\parallel}, \end{aligned} \quad (131)$$

where $F(k_{\parallel}, k_{\nu}; z) \equiv \int_{-\infty}^{\infty} \frac{k_{\parallel}}{2\pi} W(k_{\parallel}; z, z') e^{ik_{\parallel}z'} dz'$ can be calculated analytically

$$\begin{aligned} F(k_{\parallel}, k_{\nu}; z) &= \frac{1}{\varepsilon(z)} \begin{cases} \left[r_{12} \frac{k_{\parallel} + ik_{\nu}}{k_{\parallel}^2 + k_{\nu}^2} + \left(\frac{\varepsilon_1}{\varepsilon_2} t_{21} - 1 \right) \frac{k_{\parallel} - ik_{\nu}}{k_{\parallel}^2 + k_{\nu}^2} \right] e^{-k_{\parallel}z} & \text{for } z > 0 \\ \left[r_{21} \frac{k_{\parallel} - ik_{\nu}}{k_{\parallel}^2 + k_{\nu}^2} + \left(\frac{\varepsilon_2}{\varepsilon_1} t_{12} - 1 \right) \frac{k_{\parallel} + ik_{\nu}}{k_{\parallel}^2 + k_{\nu}^2} \right] e^{k_{\parallel}z} & \text{for } z < 0 \end{cases} \\ &+ \frac{1}{\varepsilon(z)} F_0(k_{\parallel}, k_{\nu}; z), \end{aligned} \quad (132)$$

with $F_0(k_{\parallel}, k_{\nu}; z) = \frac{2k_{\parallel}}{k_{\parallel}^2 + k_{\nu}^2} e^{ik_{\parallel}z}$ yielding the vacuum-correspondent. The first term represents the induced potential [apart from bulk polarization due to $\varepsilon(z)$]. The total potential then follows from Eq. (131), with the first term amenable by numerical integration, while the second term follows directly from Eq. (126). The cover image depicts the induced potential obtained from such a calculation for a graphene-sheet in vacuum surroundings $\varepsilon_1 = \varepsilon_2 = 1$.

A.2 NONRETARDED ELECTRON ENERGY LOSS

With the induced potential established, we are in a position to also derive the associated energy loss, i.e. the EEL signal. Specifically, combining Eq. (131) and the nonretarded equivalent of Eq. (27), we derive the following EEL signal (measured relative to a scenario without reflections, i.e. we do not account for bulk loss in either medium)

$$\Gamma(\omega) = -\frac{2\alpha_{\nu}}{\pi\omega} \text{Im} \left[\frac{1}{\varepsilon_1 \varepsilon_2} \int_0^{\infty} \frac{k_{\parallel} k_{\nu}}{k_{\parallel}^2 + k_{\nu}^2} \left(\frac{\varepsilon_2 r_{12}}{k_{\parallel} - ik_{\nu}} + \frac{\varepsilon_1 r_{21}}{k_{\parallel} + ik_{\nu}} \right) \right] dk_{\parallel}, \quad (133)$$

Note that the reflection coefficients of the potential differ in sign from the corresponding nonretarded limit of the TM Fresnel coefficient; this is merely a consequence of the gradient interrelation between \mathbf{E} and ϕ .

Concretely, a bulk-term $\Gamma_{\text{bulk}}(\omega) = \frac{2\alpha_{\nu}}{\pi\omega} \text{Re} \left(\frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_1 \varepsilon_2} \int_0^{k_{\text{cut}}} \frac{k_{\parallel}}{k_{\parallel}^2 + k_{\nu}^2} dk_{\parallel} \right)$, which diverges for lossy surroundings, is omitted in Eq. (133).

where $\alpha_\nu = e^2/4\pi\epsilon_0\hbar\nu$ is a fine-structure constant at the effective propagation velocity ν . For graphene in vacuum, the expression is particularly simple, reducing there to the result previously reported in Ref. 267

$$\Gamma_{\text{vac}}(\omega) = \frac{-4\alpha_\nu}{\pi\omega} \int_0^\infty \frac{x^2}{(1+x^2)^2} \text{Im}[r(xk_\nu, \omega)] dx, \quad (134)$$

obtained by using $r_{12} = r_{21} \equiv r \equiv -g/(2+g)$, which retain frequency- and momentum-dependence through g , see Eq. (130). An analytical result can be obtained in the low-loss intraband limit, where g simplifies to $g = -2k_\parallel/k_{\text{sp}}^{\text{intra}}$ with $k_{\text{sp}}^{\text{intra}} = 2\pi(\hbar\omega)^2\epsilon_0/e^2\epsilon_F$, reducing in turn the reflection coefficient to $r = k_\parallel/(k_{\text{sp}}^{\text{intra}} - k_\parallel + i\eta)$ (where, for reasons of causality, an infinitesimal loss $\eta \rightarrow 0^+$ perseveres). By the Sokhotski–Plemelj theorem, $\lim_{\eta \rightarrow 0^+} (x + i\eta)^{-1} = \mathcal{P}x^{-1} - i\pi\delta(x)$, we then find $\text{Im} r = -\pi k_{\text{sp}}\delta(k_\parallel - k_{\text{sp}})$ and accordingly that

$$\Gamma_{\text{vac}}^{\text{intra}}(\omega) \simeq \frac{4\alpha_\nu}{\omega} k_{\text{sp}}^{\text{intra}} \int_0^\infty \frac{x^2}{(1+x^2)^2} \delta(xk_\nu - k_{\text{sp}}^{\text{intra}}) dx = \frac{2\hbar}{\epsilon_F} \frac{s^2}{(1+s^2)^2}, \quad (135)$$

with $s \equiv k_{\text{sp}}^{\text{intra}}/k_\nu = \alpha_\nu^{-1}\hbar\omega/2\epsilon_F$ and in agreement with Ref. 267. In principle, the result retains validity outside the intraband approximation as well, provided the low-loss assumption remains effective, requiring only a substitution of $k_{\text{sp}}^{\text{intra}}$ by the full intra- and interband dispersion k_{sp} , defined then by the implicit equation of Eq. (87). Equation (135) demonstrates that the EEL signal is maximal when $s = 1$, i.e. under conditions of phase matching $k_\nu = k_{\text{sp}}$: accordingly, the primary loss of energy is through the excitation of a plasmon of momentum k_ν . Nevertheless, as evident from the term $s^2/(1+s^2)^2$, a continuum of surface plasmons are excited around this momentum, all contributing to the total loss. We discuss concrete calculations of the EEL signal in graphene in Figure 21.

B

EXISTENCE CRITERIA: THE TRANSVERSE ELECTRIC PLASMON POLARITON

In this appendix we elucidate the analysis leading to the existence criteria [Eq. \(93\)](#) for TE GPPs in a scenario of vacuum superstrate $\varepsilon^+ = 1$ and arbitrary substrate $\varepsilon^- = \varepsilon$. We restrict ourselves to the lossless case, although generalization to the lossy case is straightforward.

The general (local) TE GPP dispersion condition, $k_{\perp}^+ + k_{\perp}^- + \mu_0 \omega \sigma(\omega) = 0$, following from the poles of [Eq. \(85b\)](#), is an example of a so-called radical equation. It can be converted to a secular equation by repeated squaring, revealing potential solutions (compactly expressed through the implicitly frequency-dependent $\Sigma \equiv 4ik_0 \alpha \sigma / \sigma_0$)

$$k_{\parallel}^2 = k_0^2 + \left[\frac{(\varepsilon - 1)k_0^2 + \Sigma^2}{2\Sigma} \right]^2. \quad (136)$$

We emphasize that [Eq. \(136\)](#) represents *potential* solutions, because the squaring procedure inadvertently may introduce fictitious, non-valid solutions. As a consequence, the validity of the solutions in [Eq. \(136\)](#) must be checked explicitly by insertion into the original radical equation. In doing so, the following equation is obtained

$$(\varepsilon - 1)k_0^2 - \Sigma^2 + \sqrt{[(\varepsilon - 1)k_0^2 - \Sigma^2]^2} = 0, \quad (137)$$

where, to avoid complications due to the branch cut of the complex square root, we have assumed negligible loss, i.e. purely imaginary σ (and hence real Σ), real ε , and positive $\varepsilon - 1$. The remaining square-root term can be straightforwardly evaluated by noting that $\sqrt{z^2} = \text{sgn}(\text{Re } z)z$, such that [Eq. \(137\)](#) reduces to the simpler condition

$$1 + \text{sgn}[(\varepsilon - 1)k_0^2 - \Sigma^2] = 0. \quad (138)$$

Finally, this condition is then trivially converted to

$$(\varepsilon - 1)k_0^2 < \Sigma^2, \quad (139)$$

which finally assumes the form discussed in [Eq. \(93\)](#) by the introduction of the condition function $\mathcal{C}(\omega) = \Sigma^2/k_0^2 = \{4\alpha \text{Im}[\sigma(\omega)]/\sigma_0\}^2$, with the latter equality following from the assumption of purely imaginary σ .

C

EIGENDECOMPOSITION IN 2D NANOSTRUCTURES: PLANE WAVE RESPONSE

In this appendix we outline the derivation leading to Eq. (103), i.e. we develop the necessary framework for an eigendecomposition of the response of a two-dimensional nanostructure.

Initially, let us recast Eqs. (94) in operator and bracket notation by introducing the kets $|\rho\rangle$ and $|\phi\rangle$ defining the real-space in-plane quantities $\rho_{\parallel}(\mathbf{r}_{\parallel}) \equiv \langle \mathbf{r}_{\parallel} | \rho \rangle$ and $\phi(\mathbf{r}_{\parallel}) \equiv \langle \mathbf{r}_{\parallel} | \phi \rangle$, as well as operators K and G^{-1} defined by

$$\langle \mathbf{r}_{\parallel} | K | \rho \rangle \equiv -\frac{1}{2\pi} \int_{\tilde{\Omega}} \frac{1}{|\tilde{\mathbf{r}}_{\parallel} - \tilde{\mathbf{r}}'_{\parallel}|} \rho_{\parallel}(\mathbf{r}'_{\parallel}), \quad (140a)$$

$$\langle \mathbf{r}_{\parallel} | G^{-1} | \phi \rangle \equiv \tilde{\nabla}_{\parallel} \cdot [f(\tilde{\mathbf{r}}_{\parallel}) \tilde{\nabla}_{\parallel} \phi(\tilde{\mathbf{r}}_{\parallel})], \quad (140b)$$

such that G constitutes the appropriate Green function for inversion of Eq. (94b) (respecting the applied BCs). With this notation, Eqs. (94) shorten considerably

$$|\phi\rangle = -Q^{-1}K|\rho\rangle, \quad (141a)$$

$$|\rho\rangle = -\frac{\sigma}{i\omega L^2} G^{-1}|\phi\rangle, \quad (141b)$$

with $Q \equiv 2\varepsilon_0 \bar{\varepsilon}/L$ introduced for the sake of brevity in ensuing formulas. Proceeding with this notation, we may also recast Eq. (95a) as an operator equation for $|\phi\rangle$ – or, alternatively, cast an equivalent equation for $|\rho\rangle$, such that

$$\zeta|\phi\rangle = KG^{-1}|\phi\rangle + \zeta|\phi^{\text{ext}}\rangle, \quad \zeta_v|\phi_v\rangle = KG^{-1}|\phi_v\rangle, \quad (142a)$$

$$\zeta G|\rho\rangle = K|\rho\rangle - Q|\phi^{\text{ext}}\rangle, \quad \zeta_v G|\rho_v\rangle = K|\rho_v\rangle, \quad (142b)$$

where in both cases, we have included also an external perturbing potential $|\phi^{\text{ext}}\rangle$ [included by addition into Eq. (141a)], such that $|\phi\rangle$ represents a *total* potential.

In the absence of an external potential, Eqs. (142) constitute two eigensystems [indicated by right-hand gray expressions in Eqs. (142)] with identical eigenvalues, such that $\{\zeta_v|\phi_v\rangle\}$ and $\{\zeta_v|\rho_v\rangle\}$ provide complementary but equally fundamental descriptions of the v th eigensolution. Moreover, since K and G are real and symmetric, they constitute a complete biorthogonal basis: $\langle \rho_v | \phi_v \rangle = \delta_{v\mu} Q^{-1} \zeta_v \langle \rho_v | G | \rho_v \rangle$ (in turn these properties of K and G also guarantee real, positive eigenvalues $\zeta_v > 0$). Thus, we may expand the induced charge density from Eq. (142b) in terms of the eigendensities according to

$$|\rho\rangle = \sum_v \xi_v |\rho_v\rangle. \quad (143)$$

The expansion coefficients ξ_v can be determined by combining Eqs. (142b) and (143) while utilizing the biorthogonal interrelation of $|\phi_v\rangle$ and $|\rho_v\rangle$, allowing

$$\xi_v = \frac{\langle \rho_v | \phi^{\text{ext}} \rangle}{(\zeta/\zeta_v - 1) \langle \rho_v | \phi_v \rangle} = Q \frac{\langle \rho_v | \phi^{\text{ext}} \rangle}{\zeta - \zeta_v}, \quad (144)$$

where we have introduced a normalization choice $\langle \rho_v | \phi_v \rangle = Q^{-1} \zeta_v$ at the second equality. Next we specialize to the case of illumination by a plane wave polarized along

We remind that quantities with tilde-notation are normalized by the characteristic length L , particularly $\tilde{\mathbf{r}}_{\parallel} \equiv L^{-1} \mathbf{r}_{\parallel}$, $\tilde{\nabla}_{\parallel} \equiv L \nabla_{\parallel}$, and $\tilde{\Omega} \equiv L^{-1} \Omega$.

The brackets $\langle \cdot | \cdot \rangle$ here implicitly denote integration over dimensionless coordinates $\tilde{\mathbf{r}}_{\parallel} \in \tilde{\Omega}$.

The normalization is equivalent with the statement $\langle \rho_v | G | \rho_v \rangle = 1$. Implicitly this choice introduces an assumption of dimensionless eigendensities.

x with amplitude E_0 , such that $\phi^{\text{ext}}(\mathbf{r}_{\parallel}) = -E_0 x = -E_0 \tilde{x} L$, in which case the expansion coefficients reduce to $\xi_v = E_0 Q L \langle \rho_v | \tilde{x} \rangle / (\zeta_v - \zeta)$. The dipole moment along the polarization direction, defined as $p \equiv \int_{\Omega} x \rho_{\parallel}(\mathbf{r}_{\parallel}) d^2 \mathbf{r}_{\parallel} = L^3 \langle \tilde{x} | \rho \rangle$, then follows immediately from ξ_v and Eq. (143)

$$p(\omega) = 2E_0 \varepsilon_0 \bar{\varepsilon} L^3 \sum_v \frac{|\langle \tilde{x} | \rho_v \rangle|^2}{\zeta_v - \zeta(\omega)}, \quad (145)$$

where parameters with frequency-dependence has had this dependence reinstated. The result Eq. (103) is then finally recovered after noting the definition of the polarizability: $\alpha(\omega) \equiv p(\omega) / \varepsilon_0 \bar{\varepsilon} E_0$.

D | NONINTERACTING DENSITY-DENSITY RESPONSE

In this appendix, we derive the noninteracting density-density response function χ^0 in a real-space basis for the case of a two-component spinor wavefunction, which, in its scalar limit reproduces Eq. (115), and in its full form allows response calculations using a Dirac spinor-description. The starting point for the derivation is the definition of the density-density response function, obtained from the Kubo formula [32]

$$\chi(\mathbf{r}, t; \mathbf{r}', t') = -\frac{i}{\hbar} \theta(t-t') \langle [\hat{\rho}(\mathbf{r}, t), \hat{\rho}(\mathbf{r}', t')] \rangle_0, \quad (146)$$

where $\hat{\rho}(\mathbf{r}, t)$ denotes the density operator at coordinate (\mathbf{r}, t) , and $\langle \dots \rangle_0$ denotes the thermal average with respect to the ground state (i.e. nonperturbed states).

To express the density operator we first remind the form of the creation and annihilation quantum field operators, $\hat{\Psi}^\dagger(\mathbf{r}, t)$ and $\hat{\Psi}(\mathbf{r}, t)$, for states $|v\rangle$ with real-space two-component spinor wavefunctions $\psi_v(\mathbf{r}) \equiv [\psi_v^A(\mathbf{r}), \psi_v^B(\mathbf{r})]^T$ (components associating with A- and B-sublattice occupancy in \mathbf{K} or \mathbf{K}' valleys)

$$\hat{\Psi}^\dagger(\mathbf{r}, t) = \sum_v \begin{bmatrix} \psi_v^{A*}(\mathbf{r}) \\ \psi_v^{B*}(\mathbf{r}) \end{bmatrix} \hat{c}_v^\dagger(t), \quad \hat{\Psi}(\mathbf{r}, t) = \sum_v \begin{bmatrix} \psi_v^A(\mathbf{r}) \\ \psi_v^B(\mathbf{r}) \end{bmatrix} \hat{c}_v(t), \quad (147)$$

where $\hat{c}_v^\dagger(t)$ and $\hat{c}_v(t)$ are time-dependent creation and annihilation operators of the two-component spinor $|v\rangle$, which we presently assume constitute a complete, normalized set; $\sum_v |v\rangle\langle v| = 1$. Next, we express the density operator $\hat{\rho}(\mathbf{r}, t)$ via $\hat{\Psi}^{(\dagger)}(\mathbf{r}, t)$

$$\begin{aligned} \hat{\rho}(\mathbf{r}, t) &= \hat{\Psi}^\dagger(\mathbf{r}, t) \hat{\Psi}(\mathbf{r}, t) \\ &= \sum_{v_1 v_2} [\psi_{v_1}^{A*}(\mathbf{r}) \psi_{v_2}^A(\mathbf{r}) + \psi_{v_1}^{B*}(\mathbf{r}) \psi_{v_2}^B(\mathbf{r})] \hat{c}_{v_1}^\dagger(t) \hat{c}_{v_2}(t) \\ &= \sum_{v_1 v_2} \sum_{\mu=A,B} \psi_{v_1}^{\mu*}(\mathbf{r}) \psi_{v_2}^\mu(\mathbf{r}) \hat{c}_{v_1}^\dagger(t) \hat{c}_{v_2}(t). \end{aligned} \quad (148)$$

In the noninteracting picture the time-dependence of the creation and annihilation operators is manageable, and is given by

$$c_v^\dagger(t) = c_v^\dagger e^{i\xi_v t/\hbar}, \quad c_v(t) = c_v e^{-i\xi_v t/\hbar}, \quad (149)$$

where ξ_v are the natural energies in the grand canonical ensemble $\xi_v \equiv \epsilon_v - \epsilon_F$, with ϵ_v denoting the single-particle eigenenergies and ϵ_F the Fermi level.

To derive the noninteracting density-density response $\chi^0(\mathbf{r}, \mathbf{r}'; \omega)$ – which Eq. (149) automatically restrict us to – in the frequency-domain, we employ the equations of motion technique. Inserting $\hat{\rho}(\mathbf{r}, t)$ from Eq. (148) into Eq. (146) we find

$$\chi^0(\mathbf{r}, \mathbf{r}'; t-t') = -\frac{i}{\hbar} \sum_{v_1 v_2 \mu} \psi_{v_1}^{\mu*}(\mathbf{r}) \psi_{v_2}^\mu(\mathbf{r}) \underbrace{\theta(t-t') e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \langle [\hat{c}_{v_1}^\dagger \hat{c}_{v_2}, \hat{\rho}(\mathbf{r}', t')] \rangle_0}_{\equiv \Xi_{v_1}^{v_2}(\mathbf{r}', t-t')} \quad (150)$$

Note that the creation (annihilation) operators here apply simultaneously to both A- and B-sublattices: in a Dirac spinor description, the spinors constitute the creatable (annihilable) entities – not their individual sublattice elements.

Time-dependence of the noninteracting operators can be derived from the commutation rules of the fermionic operators, and from the Heisenberg equation.

where the time-dependence has been reduced to the form $t - t'$ rather than (t, t') as a consequence of temporal invariance. To proceed with the equations of motion technique, we take the time-derivative of the auxiliary function $\Xi_{v_1}^{v_2}(\mathbf{r}', t - t')$

$$\begin{aligned} \partial_t \Xi_{v_1}^{v_2}(\mathbf{r}', t - t') &= \frac{i}{\hbar}(\epsilon_{v_1} - \epsilon_{v_2})\theta(t - t')e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \langle [\hat{c}_{v_1}^\dagger \hat{c}_{v_2}, \hat{\rho}(\mathbf{r}', t')] \rangle_0 \\ &\quad + \delta(t - t')e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \langle [\hat{c}_{v_1}^\dagger \hat{c}_{v_2}, \hat{\rho}(\mathbf{r}', t')] \rangle_0 \\ &= \frac{i}{\hbar}(\epsilon_{v_1} - \epsilon_{v_2})\Xi_{v_1}^{v_2}(\mathbf{r}', t - t') \\ &\quad + \delta(t - t') \underbrace{e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \langle [\hat{c}_{v_1}^\dagger \hat{c}_{v_2}, \hat{\rho}(\mathbf{r}', t')] \rangle_0}_{\equiv \Upsilon_{v_1}^{v_2}(\mathbf{r}', t - t')} \end{aligned} \quad (151)$$

To discern the explicit form of $\Upsilon_{v_1}^{v_2}(\mathbf{r}', t - t')$ we insert $\hat{\rho}(\mathbf{r}', t')$ from Eq. (148)

$$\Upsilon_{v_1}^{v_2}(\mathbf{r}', t - t') = e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \sum_{v_3 v_4 \gamma} \psi_{v_3}^{\gamma*}(\mathbf{r}') \psi_{v_4}^\gamma(\mathbf{r}') e^{i(\epsilon_{v_3} - \epsilon_{v_4})t'/\hbar} \langle [\hat{c}_{v_1}^\dagger \hat{c}_{v_2}, \hat{c}_{v_3}^\dagger \hat{c}_{v_4}] \rangle_0. \quad (152)$$

The four-element commutator can be expanded in two-element commutators using the fermionic rule $[\hat{c}_{v_1}^\dagger \hat{c}_{v_2}, \hat{c}_{v_3}^\dagger \hat{c}_{v_4}] = \hat{c}_{v_1}^\dagger \hat{c}_{v_4} \delta_{v_2 v_3} - \hat{c}_{v_3}^\dagger \hat{c}_{v_2} \delta_{v_1 v_4}$, allowing

$$\begin{aligned} \Upsilon_{v_1}^{v_2}(\mathbf{r}', t - t') &= e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \sum_{v_3 v_4 \gamma} \psi_{v_3}^{\gamma*}(\mathbf{r}') \psi_{v_4}^\gamma(\mathbf{r}') e^{i(\epsilon_{v_3} - \epsilon_{v_4})t'/\hbar} \\ &\quad \times [\delta_{v_2 v_3} \langle \hat{c}_{v_1}^\dagger \hat{c}_{v_4} \rangle_0 - \delta_{v_1 v_4} \langle \hat{c}_{v_3}^\dagger \hat{c}_{v_2} \rangle_0] \\ &= e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \sum_{v_4 \gamma} \psi_{v_2}^{\gamma*}(\mathbf{r}') \psi_{v_4}^\gamma(\mathbf{r}') e^{i(\epsilon_{v_2} - \epsilon_{v_4})t'/\hbar} \langle \hat{c}_{v_1}^\dagger \hat{c}_{v_4} \rangle_0 \\ &\quad - e^{i(\epsilon_{v_1} - \epsilon_{v_2})t/\hbar} \sum_{v_3 \gamma} \psi_{v_3}^{\gamma*}(\mathbf{r}') \psi_{v_1}^\gamma(\mathbf{r}') e^{i(\epsilon_{v_3} - \epsilon_{v_1})t'/\hbar} \langle \hat{c}_{v_3}^\dagger \hat{c}_{v_2} \rangle_0. \end{aligned} \quad (153)$$

The Fermi-Dirac distributions can be derived by noting that $\langle \hat{A} \rangle_0 = Z_G^{-1} \sum_v \langle v | \hat{A} | v \rangle e^{-\xi_v/k_B T}$ with grand canonical partition function Z_G .

The thermal averages yield Fermi-Dirac distributions since $\langle \hat{c}_{v'}^\dagger \hat{c}_{v'} \rangle_0 = f_v \delta_{v'v}$, with $f_v = [1 + \exp(\xi_v/k_B T)]^{-1}$. Accordingly, the v_4 - and v_3 -summations are reduced to

$$\Upsilon_{v_1}^{v_2}(\mathbf{r}', t - t') = \sum_{\gamma} \psi_{v_2}^{\gamma*}(\mathbf{r}') \psi_{v_1}^\gamma(\mathbf{r}') e^{i(\epsilon_{v_1} - \epsilon_{v_2})(t - t')/\hbar} (f_{v_1} - f_{v_2}). \quad (154)$$

Reinserting into Eq. (151) and Fourier transforming in time, $\partial_t \rightarrow -i\omega + \eta$, produces

$$\begin{aligned} -i(\omega + i\eta)\Xi_{v_1}^{v_2}(\mathbf{r}', \omega) &= \frac{i}{\hbar}(\epsilon_{v_1} - \epsilon_{v_2})\Xi_{v_1}^{v_2}(\mathbf{r}', \omega) + \sum_{\gamma} \psi_{v_2}^{\gamma*}(\mathbf{r}') \psi_{v_1}^\gamma(\mathbf{r}') (f_{v_1} - f_{v_2}) \\ \Leftrightarrow \Xi_{v_1}^{v_2}(\mathbf{r}', \omega) &= i\hbar \frac{f_{v_1} - f_{v_2}}{\epsilon_{v_1} - \epsilon_{v_2} + \hbar(\omega + i\eta)} \sum_{\gamma} \psi_{v_2}^{\gamma*}(\mathbf{r}') \psi_{v_1}^\gamma(\mathbf{r}'). \end{aligned} \quad (155)$$

Lastly, we Fourier transform Eq. (150) and subsequently insert $\Xi_{v_1}^{v_2}(\mathbf{r}', \omega)$ via Eq. (155), such that

$$\chi^0(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{v_1 v_2 \mu \gamma} (f_{v_1} - f_{v_2}) \frac{\psi_{v_1}^{\mu*}(\mathbf{r}) \psi_{v_2}^\mu(\mathbf{r}) \psi_{v_2}^{\gamma*}(\mathbf{r}') \psi_{v_1}^\gamma(\mathbf{r}')}{\epsilon_{v_1} - \epsilon_{v_2} + \hbar(\omega + i\eta)}. \quad (156)$$

Spin- and valley-degeneracies in addition yield multiplicative factors of 2 each, if not included explicitly in the summation.

By reintroducing vectorial notation the μ - and γ -summations can be removed, revealing the final result

$$\chi^0(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{v_1 v_2} (f_{v_1} - f_{v_2}) \frac{[\psi_{v_1}^\dagger(\mathbf{r}) \psi_{v_2}(\mathbf{r})][\psi_{v_2}^\dagger(\mathbf{r}') \psi_{v_1}(\mathbf{r}')] }{\epsilon_{v_1} - \epsilon_{v_2} + \hbar(\omega + i\eta)}. \quad (157)$$

E

DISCRETIZED UNITARY TIME-OPERATOR

The discretization of the unitary time-operator discussed in the context of Eqs. (117) is achieved by splitting the propagation over several steps. Noting the fundamental property $\hat{U}(t, t') = \hat{U}(t, t_1)\hat{U}(t_1, t')$ of the time-operator, we observe that evolution from $t' = 0$ to t can be decomposed into M steps according to

$$\hat{U}(t, 0) = \prod_{i=0}^{M-1} \hat{U}(t_{i+1}, t_i), \quad t_{i+1} = t_i + \delta t, \quad t_0 = 0, \quad t_M = t. \quad (158)$$

The remaining numerical challenge then revolves around the accurate approximation of $\hat{U}(t_i + \delta t, t_i)$ for small steps δt . A basic requirement for a successful methodology is a proper preservation of properties of time-ordering, unitarity, and time-reversal symmetry of \hat{U} at each time-step [335].

A particularly attractive approach is achieved by a combination of the exponential midpoint rule (a) for approximation of $\hat{U}(t_i + \delta t, t_i)$ and a splitting technique (b) for efficient computation of the relevant matrix-exponentials. In its essential form, these approximations, i.e. (a) and (b), approximate the exponential integral $\hat{U}(t_{i+1}, t_i) = \exp[-i\hbar^{-1} \int_{t_i}^{t_{i+1}} \hat{H}(t') dt']$ by

$$\begin{aligned} \hat{U}(t_{i+1}, t_i) &\stackrel{(a)}{\simeq} \exp\left[\frac{-i}{\hbar} \hat{H}(t_i + \delta t/2) \delta t\right] \\ &\stackrel{(b)}{\simeq} \exp\left(\frac{-i}{2\hbar} \hat{H}^0 \delta t\right) \exp\left[\frac{-i}{\hbar} \hat{H}'(t_i + \frac{1}{2} \delta t) \delta t\right] \exp\left(\frac{-i}{2\hbar} \hat{H}^0 \delta t\right) + \mathcal{O}(\delta t^3), \end{aligned} \quad (159)$$

with $\hat{H}'(t) = \hat{H}^{\text{ext}}(t) + \hat{H}^{\text{ind}}(t)$. The key advantage of the splitting technique is that computation of a full matrix-exponential, requiring $\mathcal{O}(N^3)$ operations, is required only *once* for the calculation of $\exp(-i\hat{H}^0 \delta t/2\hbar)$. In contrast, the perturbative term $\hat{H}'(t)$ is diagonal in the $\{|l\rangle\}$ basis, and so represents a very manageable numerical challenge. However, accurate determination of $\hat{H}'(t + \delta/2)$ is problematic, because the Hartree potential is unknown at time $t + \delta/2$. To retain an accuracy of $\mathcal{O}(\delta t^3)$ one must then perform an accurate extrapolation of \hat{H}^{ind} from t to $t + \delta/2$, involving in principle the need for a self-consistent cycle. Fortunately, a simpler and cost-effective alternative exists [which retains $\mathcal{O}(\delta t^3)$], suggested by Watanabe and Tsukada [370]

$$\hat{U}(t_{i+1}, t_i) \simeq \exp\left(\frac{-i}{2\hbar} \hat{H}^0 \delta t\right) \exp\left[\frac{-i}{\hbar} \hat{H}'_*(t_i) \delta t\right] \exp\left(\frac{-i}{2\hbar} \hat{H}^0 \delta t\right) + \mathcal{O}(\delta t^3), \quad (160)$$

with perturbation $\hat{H}'_*(t_i) \equiv \hat{H}^{\text{ext}}(t_i + \delta/2) + \hat{H}^{\text{ind}}[\rho_*]$, where the Hartree contribution is computed from an induced charge density ρ_* built from the half-evolved (occupied) states $\{\exp(-i\hat{H}^0 \delta t/2\hbar) |\psi_j(t_i)\rangle\}$ – in effect, the Hartree contribution is extrapolated from t_i to $t_i + \delta/2$ by a noninteracting, or bare, propagation of duration $\delta/2$. The added cost of this step is negligible (compared to the required matrix product operations), amounting simply to the calculation of ρ_* following the first application of $\exp(-i\hat{H}^0 \delta t/2\hbar)$ to the occupied states $\{|\psi_j(t)\rangle\}$.

Eqs. (158) and (160) allows the propagation of all occupied states $\{|\psi_j(t)\rangle\}$ across the time-series $\{t_i\}$. Relevant macroscopic parameters, $A(t)$, can be recorded at every time-step, and the spectral response computed subsequently by a Fourier transform

$$A(\omega) = \int_{-\infty}^{\infty} A(t) e^{i\omega t} e^{-\eta(t-t_k)\theta(t-t_k)} dt, \quad (161)$$

with decay accounted for approximately through $\eta = \gamma/2$ effective from time t_k . In practice, the transform is approximated by the discrete Fourier transform over the time-series $\{t_i\}$, conveniently computed by the fast Fourier transform.

In practice δt should be chosen much smaller than all relevant time-scales, including the perturbation's.

E.1 BARE EVOLUTION THRESHOLDING

Lastly, we mention that the computational cost of applying the half-step bare unitary time evolution $\hat{U}^0(\delta t/2) \equiv \exp(-i\hat{H}^0\delta t/2\hbar)$ – which constitutes the main computational effort, $\mathcal{O}(N^3)$, of the method – can be significantly reduced. In particular, in the $|l\rangle$ -basis the matrix representation is distinctly quasi-sparse, in the sense that the vast majority of elements are vanishingly small. These small entries coincide with transitions between site-orbitals $|l\rangle$ and $|l'\rangle$ of large spatial separation where $|\mathbf{r}_l - \mathbf{r}_{l'}|$ exceeds several lattice constants. In other words, for very short propagation times, the bare evolution only significantly connects nearby sites. Thus, by introducing either a tolerance for the values of matrix-elements or a cutoff separation $|\mathbf{r}_l - \mathbf{r}_{l'}| < r_{\text{cut}}$, the bare evolution operator can be significantly sparsified, especially in large systems. As a result, the number of elements in the matrix-representation of $\hat{U}_0(\delta t/2)$ can be brought to a scaling $\mathcal{O}(N)$ rather than $\mathcal{O}(N^2)$, since each site only connects to a finite number of nearby sites. As an example, for element-thresholds around 10^{-7} we obtain sparsity of $\sim 1\%$ for structures of characteristic size ~ 20 nm. This sparsity can be exploited to yield significant speedups by utilizing specialized sparse matrix product algorithms.

Without thresholding, a calculation with $N \approx 15 \times 10^3$ atoms and $M \approx 8.5 \times 10^3$ time-steps requires a run-time of ≈ 110 hours on a 20 core Intel Xeon E5-2660 processor. With element-thresholding at 10^{-10} , the run-time is reduced to ≈ 74 hours using just a single core.

Part IV

SELECTED PUBLICATIONS

A selection of the author's publications are included here, specifically, [Publications A, D, E, and F](#). Remaining publications are listed on [p. xiii](#).

Preceding each publication we include a short description of author contributions. The content of these descriptions is identical to those signed by the the coauthors for the mandatory author statements.

PHYS. REV. B **92**, 121407(R) (2015)

T. Christensen and W. Yan developed the initial idea and associated theory. T. Christensen implemented and performed all numerical calculations and wrote the manuscript with input from all authors.

PHYSICAL REVIEW B **92**, 121407(R) (2015)**Kerr nonlinearity and plasmonic bistability in graphene nanoribbons**Thomas Christensen,^{1,2} Wei Yan,^{1,2} Antti-Pekka Jauho,^{2,3} Martijn Wubs,^{1,2} and N. Asger Mortensen^{1,2,*}¹*Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*²*Center for Nanostructured Graphene, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*³*Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

(Received 8 June 2015; published 16 September 2015)

We theoretically examine the role of Kerr nonlinearities for graphene plasmonics in nanostructures, specifically in nanoribbons. The nonlinear Kerr interaction is included semiclassically in the intraband approximation. The resulting electromagnetic problem is solved numerically by self-consistent iteration with linear steps using a real-space discretization. We derive a simple approximation for the resonance shifts in general graphene nanostructures, and obtain excellent agreement with numerics for moderately high field strengths. Near plasmonic resonances the nonlinearities are strongly enhanced due to field enhancement, and the total nonlinearity is significantly affected by the field inhomogeneity of the plasmonic excitation. Finally, we discuss the emergence of a plasmonic bistability which exists for energies red-shifted relative to the linear resonance. Our results offer insights into the role of nonlinear interaction in nanostructured graphene and pave the way for experimental investigation.

DOI: [10.1103/PhysRevB.92.121407](https://doi.org/10.1103/PhysRevB.92.121407)

PACS number(s): 78.67.Wj, 73.20.Mf, 78.20.Ci, 78.20.Mg

Nonlinear optical effects [1,2], facilitated by strong light-matter interaction, are indispensable in modern photonics. Indeed, a host of phenomena and applications arise at sufficiently high field strengths, owing to superlinear photon-photon response mediated by strong light-matter interaction, ranging from frequency conversion through all-optical phase modulation to ultrafast switching, and is pursued in a broad range of platforms [3–5]. A perennial challenge in the discipline is to achieve significant nonlinear interaction at ever smaller excitation powers and interaction volumes, while maintaining *in situ* tunability and control. In achieving this goal, the field of plasmonics, describing the strong hybridization of the free electromagnetic field with collective oscillations of conduction electrons, suggests several promising avenues [6]. In particular, the extreme local field enhancements inherent to plasmonic excitations amplify intrinsic nonlinearities considerably, allowing large effective nonlinearities.

Nevertheless, plasmonic field enhancement is fundamentally limited by intrinsic Ohmic losses even in noble metals. The advent of the two-dimensional material graphene has garnered significant interest in the plasmonic community [7–10], in part due to extremely large electron mobilities [11–13] and concomitant extraordinary plasmonic field enhancements [14], exceeding even the very large enhancements known from metal plasmonics. Furthermore, graphene has attracted much interest also for its exceptional intrinsic nonlinear properties both theoretically [15–18] and experimentally [19–21]. Building on this compound fortuity, a body of research is rapidly emerging at the crossroad of nonlinear plasmonics and graphene [22–32].

Very recently, the role of Kerr nonlinearities in infinitely extended graphene has been studied, notably establishing the existence of bistable [22] and soliton solutions [23–25]. In this Rapid Communication, we study theoretically an analogous Kerr nonlinearity but in nanostructured graphene, specifically in nanoribbons, wherein plasmons, unlike in the extended

counterpart, are readily excited without momentum-matching concerns, e.g., by normally incident plane waves. We report an induced nonlinearity which is significantly affected by the degree of inhomogeneity of the electric fields of the plasmon, a feature which is absent in the corresponding extended system [22] or in coupled-dipole treatments [25]. Furthermore, we derive a simple perturbative expression for the nonlinear resonance shifts in general graphene nanostructures, which agrees excellently with full self-consistent calculations for moderately high field strengths and also explains recent numerical considerations of nonlinear ribbon notch filters [26]. Finally, we discuss the emergence of plasmonic bistability in nanoribbons under plane-wave excitation. First, however, we introduce the two components needed for a nonlinear treatment of graphene nanostructures, namely, a material response model and an exposition of the resulting electromagnetic problem.

Material response. For photon energies $\hbar\omega$ below the Fermi energy ϵ_F , the response of graphene is reasonably approximated by neglecting interband transitions. In this case, the intraband response can be derived from the Boltzmann equation. To third order in the perturbing field, the Kerr-corrected conductivity, i.e., the response oscillating at the perturbing frequency, is [22]

$$\sigma(\mathbf{r}) = \sigma_{(1)}[1 - |\mathbf{E}(\mathbf{r})|^2/E_{(3)}^2], \quad (1)$$

expressed in terms of the linear intraband conductivity $\sigma_{(1)} = ie^2\epsilon_F/\pi\hbar^2(\omega + i\gamma)$ with loss rate γ , and a third-order characteristic field $E_{(3)}^2 \equiv (8\omega_{(3)}^2)/(9\omega^2)E_{\text{sat}}^2$ linearly related to the saturation field $E_{\text{sat}} \equiv \epsilon_F\omega/ev_F$ through a loss-modified frequency $\omega_{(3)}^2 \equiv (\omega + \frac{1}{2}i\gamma)(\omega - i\gamma)$. Since the Kerr correction is of the self-focusing type [30], its usage in finite structures with inhomogeneous fields must be augmented to include a saturating mechanism, or else suffer nonphysical runaway self-focusing [33]. Here, we adopt the well-known two-level saturation model or, in other words, the [0/2] Padé approximant of $\sigma(\mathbf{r})$ consistent with Eq. (1):

$$\sigma(\mathbf{r}) \simeq \frac{\sigma_{(1)}(\mathbf{r})}{1 + |\mathbf{E}(\mathbf{r})|^2/E_{(3)}^2} + \sigma_{(3)2\gamma}(\mathbf{r}). \quad (2)$$

*asger@mailaps.org

CHRISTENSEN, YAN, JAUHO, WUBS, AND MORTENSEN

PHYSICAL REVIEW B 92, 121407(R) (2015)

This model reproduces the third-order result of Eq. (1) in the $|\mathbf{E}(\mathbf{r})|/E_{\text{sat}} \ll 1$ limit, while crucially exhibiting a sensible behavior beyond this limit as well [34]. Lastly, we include in Eq. (2) a term $\sigma_{(3)2\gamma}(\mathbf{r})$ to account for a high field loss mechanism through two-photon absorption via the phenomenological prescription suggested by Gorbach [30], via the dissipative correction $\sigma_{(3)2\gamma}(\mathbf{r}) = -i\alpha_{2\gamma}\sigma_{(1)}|\mathbf{E}(\mathbf{r})|^2/E_{\text{sat}}^2$ with $\alpha_{2\gamma} \approx 0.1$ estimated from measurements [21].

Before proceeding, we briefly discuss the limitations of the material response assumed in Eq. (2). First, the disregard of interband effects limits our consideration to energies sufficiently below $\sim 2\epsilon_F$. Second, nonlocality [35], potential edge states [36], and more generally atomistic features [31,32,37,38] are excluded, although they are important at small feature sizes. Consequently, we restrict our considerations to nanostructures of characteristic dimensions $\gtrsim 25$ nm where these effects only weakly perturb the intraband approximation.

Self-consistent response. In the quasistatic limit, the self-consistent response of graphene can be deduced from three elements: the Coulomb law, the continuity equation, and the current-field relationship as specified by a conductivity model. For a nanostructure defined by a two-dimensional domain Ω (e.g., at $z = 0$), these elements combine to form an integrodifferential equation for either the induced density or the total potential $\phi(\mathbf{r})$. Here, we choose the latter [8]:

$$\phi(\mathbf{r}) = \frac{i}{4\pi\epsilon_0\omega W} \int_{\Omega} d^2\mathbf{r}' V(\mathbf{r},\mathbf{r}') \nabla' \cdot [\sigma(\mathbf{r}') \nabla' \phi(\mathbf{r}')], \quad (3)$$

expressed in dimensionless coordinates $\mathbf{r}^{(i)} = [x^{(i)}, y^{(i)}, z]^\top$ normalized by a characteristic length W , with the Coulomb interaction $V(\mathbf{r},\mathbf{r}') = |\mathbf{r} - \mathbf{r}'|^{-1}$ [39], and with differential operators $\nabla' = [\partial_{x'}, \partial_{y'}]^\top$. The conductivity $\sigma(\mathbf{r})$ implicitly depends on frequency and, in a nonlinear treatment, also on the total field $\mathbf{E}(\mathbf{r})$. The spatial dependence of the conductivity can be conveniently expressed via a dimensionless occupation function $f(\mathbf{r}) \equiv \sigma(\mathbf{r})/\langle\sigma_{(1)}\rangle$ with $\langle\sigma_{(1)}\rangle$ denoting the average linear conductivity across Ω . Introducing operators $\mathbf{V}g(\mathbf{r}) \equiv \int d\mathbf{r}' V(\mathbf{r},\mathbf{r}')g(\mathbf{r}')$ and $\mathbf{D}g(\mathbf{r}') \equiv \nabla' \cdot [f(\mathbf{r}')\nabla'g(\mathbf{r}')] casts Eq. (3) as an eigenvalue problem for the composite operator \mathbf{VD} :$

$$\lambda\phi(\mathbf{r}) = \mathbf{VD}\phi(\mathbf{r}), \quad (4)$$

with eigenvalues $\lambda \equiv 4\pi\epsilon_0\omega W/i\langle\sigma_{(1)}\rangle$, dictating the permitted eigenenergies $\hbar\omega$. Operators \mathbf{V} and \mathbf{D} find simple matrix forms in a discretized real-space basis in both the general two-dimensional (2D) case as well as in the one-dimensional (1D) ribbon case [see Supplemental Material (SM) [40]]. A boundary condition of vanishing normal current at the boundary $\partial\Omega$ (or, equivalently, a discontinuous conductivity step) is incorporated explicitly in the construction of \mathbf{D} . In the presence of an external potential ϕ_{ext} , the eigenvalue problem in Eq. (4) becomes an inhomogeneous equation through the addition to the right-hand side of a source term $\lambda\phi_{\text{ext}}(\mathbf{r})$. To solve the nonlinear problem, with $\sigma(\mathbf{r})$, and hence $f(\mathbf{r})$ and \mathbf{D} , depending on the total electric field locally, we proceed iteratively until self-consistency is reached, exploiting at each iteration step the computational efficiency associated with linear systems [33] (see SM [40]).

With the formal premise established, we next specialize to the case of nanoribbons, translationally invariant along y

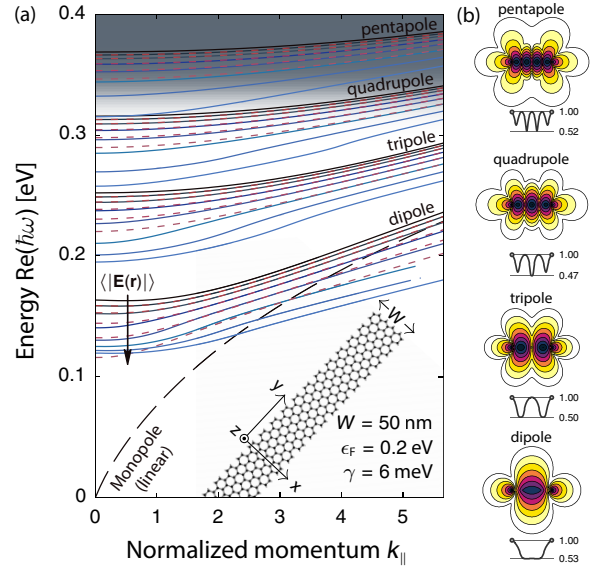


FIG. 1. (Color online) (a) Dispersion relation of a single nanoribbon. Ribbon-averaged field strength $\langle|\mathbf{E}(\mathbf{r})|\rangle$ ranges from negligible (black), i.e., linear, through 1×10^5 V/cm to 4×10^5 V/cm (lightest blue) in steps of 0.5×10^5 V/cm (increasing along arrow). For the first five $\langle|\mathbf{E}(\mathbf{r})|\rangle$, we indicate in dashed red the corresponding analytical estimate [see Eq. (5)]. For the monopole, only the linear calculation is shown. The region of significant interband modification is illustrated in shaded gray. Inset schematically depicts a single graphene nanoribbon. (b) Field amplitude $|\mathbf{E}(\mathbf{r})|$, contour maps for the case $\langle|\mathbf{E}(\mathbf{r})|\rangle = 4 \times 10^5$ V/cm, and $k_{||} = 0$. Color map ranges from maximal (dark) to minimal (light) logarithmically, with contours separated by factors of 1.5, 1.75, 2, and 2.25 for dipole, tripole, quadrupole, and pentapole cases, respectively. Sparklines below maps depict the variation of $|f(\mathbf{r})|$ along the ribbon, with maximal and minimal values indicated; their widths equal that of the ribbon, thus indicating the spatial scale.

and of finite extent W along x , a system which has already attracted much attention in the linear case [35,38,41,42]. As a consequence of translational symmetry, eigensolutions can be expanded in a momentum basis according to $\phi(\mathbf{r}) = \phi(x,z)\exp(ik_{||}y)$. Of central interest is the evolution of the eigenenergies with momentum $k_{||}$ (here dimensionless; conventional units via $k_{||}/W$), i.e. the dispersion relation $\hbar\omega_n(k_{||})$, and subsequently the response of the system to external fields.

Eigenmodes and nonlinear dispersion. For low field strengths, i.e., in the linear regime with $f(\mathbf{r})$ independent of $\mathbf{E}(\mathbf{r})$, the eigenmodes $\lambda_n(k_{||})$ of Eq. (4) are solely geometry dependent, but scale invariant, with associated eigenenergies $\hbar\omega_n(k_{||})$ dictated by $\lambda_n(k_{||}) = 4\pi\epsilon_0\omega_n(k_{||})W/i\langle\sigma_{(1)}\rangle$, allowing in the linear intraband approximation the simple scaling relation $\hbar\omega_n(k_{||}) \simeq (2\pi)^{-1}\sqrt{-\lambda_n(k_{||})e^2\epsilon_F/\epsilon_0}W$ [35,41]. Under significant nonlinear interaction, however, the eigenvalues $\lambda_n(k_{||})$ are field dependent and, by extension, scale dependent due to the self-consistent nature of the problem. In Fig. 1(a), we investigate the dispersion relation of the first few eigenmodes of a single $W = 50$ nm nanoribbon for different ribbon-averaged field strengths $\langle|\mathbf{E}(\mathbf{r})|\rangle \equiv W^{-1} \int_{\Omega} dx |\mathbf{E}(x)|$.

The most apparent impact of nonlinearity is a red-shift of all resonances. This is readily appreciated from the negativity of the Kerr correction. Indeed, the shift can be well approximated by perturbation theory for any general structure: denoting by $\hbar\omega_n^{(0)}$ and $\mathbf{E}_n^{(0)}$ the linear response eigenenergies and eigenfields [with $\langle |\mathbf{E}_n^{(0)}(\mathbf{r})| \rangle = \langle |\mathbf{E}(\mathbf{r})| \rangle$] the nonlinear eigenenergies are, to lowest order, approximately (see SM [40])

$$\omega_n \simeq \omega_n^{(0)} \sqrt{1 - \frac{9}{8} \frac{\langle |\mathbf{E}^{(0)}(\mathbf{r})|^4 \rangle}{\langle |\mathbf{E}^{(0)}(\mathbf{r})|^2 \rangle E_{\text{sat}}^2}}, \quad (5)$$

with averages taken over $\mathbf{r} \in \Omega$. The approximation is excellent for moderately high fields [see dashed red lines of Fig. 1(a)], although, naturally, inadequate for the largest fields due to the disregard of the self-consistent aspects of the nonlinearity. The approximately linear relation between resonance shift and intensity recently discussed for nonlinear ribbon filters [26] is similarly captured by Eq. (5), thus adding new physical insight. Moreover, by introducing the inhomogeneity parameter $\kappa \equiv \langle |\mathbf{E}^{(0)}(\mathbf{r})|^4 \rangle / \langle |\mathbf{E}^{(0)}(\mathbf{r})|^2 \rangle^2$ ($\kappa \geq 1$, cf. the Cauchy-Schwarz inequality), Eq. (5) reveals that inhomogeneity both modifies and enhances the nonlinearity which is $\propto \kappa \langle |\mathbf{E}^{(0)}(\mathbf{r})|^2 \rangle / E_{\text{sat}}^2$.

In Fig. 1(b), we explore this point further by depicting the modal character and inhomogeneous nature of the plasmonic modes. The modal labels are chosen from the perspective of the induced charge density $\rho(x)$ of the n th mode, with the monopole, dipole, tripole, quadrupole, and pentapole ($n = 0, 1, 2, 3$, and 4 , respectively) exhibiting n nodes of $\rho(x)$. Modes of even n are optically dark, owing to a vanishing dipole moment, and remain optically dark also under nonlinear perturbations (which preserves the system symmetry). The monopole violates charge conservation along x [but not along (x, y) for $k_i \neq 0$], is optically dark, and consistently does not converge at higher fields; as a consequence, we depict only its linear dispersion. The variation of the occupation function $f(\mathbf{r})$ under large fields is highlighted in the insets of Fig. 1(b). The pronounced spatial variation of $f(\mathbf{r})$, up to 50% for the considered $\langle |\mathbf{E}(\mathbf{r})| \rangle$, is a direct consequence of the strongly inhomogeneous nature of plasmons. Despite the significant variations of $f(\mathbf{r})$, the corresponding far-field mode profiles are highly similar in linear and nonlinear settings since they are dictated chiefly by the nodal character of $\rho(x)$.

Plane-wave excitation and bistability. Having considered the dispersion of eigenmodes, we next turn our attention to the response of the system due to a normally incident plane wave, polarized along x , i.e., $\mathbf{E}_{\text{ext}}(z=0) = E_0 \hat{\mathbf{x}}$ and $\phi_{\text{ext}}(z=0) = -E_0 x W$, corresponding to vanishing k_y . In addition to the power absorbed from the incident wave, the induced and total electric fields are of primary interest; here, we focus on the latter. For reasons of numerical efficiency and physical necessity, we compute for each separate energy the response by an initial linear calculation, followed by a ramping of the incident field strength first in upwards and then in downwards fashion, corresponding to a slow on-and-off turning of the maximum intensity; see SM [40] for implementational details.

In Fig. 2, we examine the spectral response of ribbons of widths $W = 25$ and 50 nm under different excitation strengths, i.e., under varying E_0 . For moderately high E_0 , the linear Lorentzian resonance is asymmetrically perturbed,

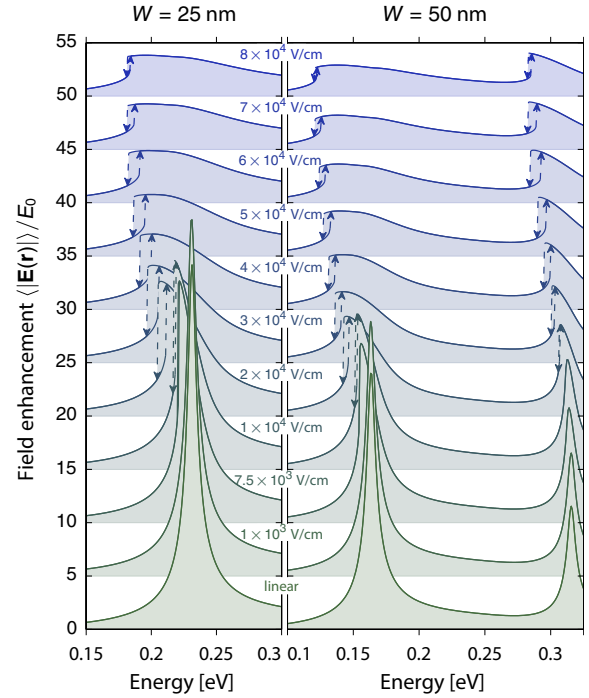


FIG. 2. (Color online) Field enhancement $\langle |\mathbf{E}(\mathbf{r})| \rangle / E_0$ as a function of energy $\hbar\omega$, for varying incident field strengths E_0 (as indicated above each spectrum). Each spectrum is offset vertically by five units. Two ribbon widths $W = 25$ and 50 nm are examined. Regions of bistability are delimited by dashed arrows which indicate the ramping direction. Material parameters are as in Fig. 1.

slightly broadened, and red-shifted. Furthermore, the upward and downward ramps to E_0 give identical spectra. As E_0 is increased further, these perturbations intensify. However, in certain energy ranges the response to upward and downward ramps toward E_0 differ (regions delimited by dashed arrows), a trademark of bistability. Similar features were found in Ref. [22] for extended graphene under normal incidence, in Ref. [25] for nanodisk chains in a coupled-dipole approximation, and in Ref. [31] for finite systems using a phenomenological anharmonic model. A key extension here is the full self-consistent accounting of the inhomogeneous nonlinear conductive profile arising in nanostructured systems. Also, coupling with the nanoribbon plasmons significantly expands the spectral region of bistability compared to the corresponding extended system under normal incidence (where plasmons are not excited), where it is restricted to $\hbar\omega < \sqrt{4/3} \alpha_{\text{fs}} \epsilon_{\text{F}}$ (with $\alpha_{\text{fs}} \equiv e^2 / 4\pi \epsilon_0 \hbar c$) [22]. Here, bistability is evident in the dipole mode for both $W = 25$ and 50 nm, but also in the quadrupole mode for $W = 50$ nm. In both cases, the area traced by the bistable region initially increases with E_0 and then decreases due to mounting saturation and absorption.

The history dependence of the response is further examined in Fig. 3(b), depicting hysteresis curves of E_0 versus $\langle |\mathbf{E}(\mathbf{r})| \rangle$ at a selection of fixed energies as indicated in the linear spectrum of Fig. 3(a). At energies far from the linear resonance at $\hbar\omega^{(0)}$ the response $\langle |\mathbf{E}(\mathbf{r})| \rangle$ relates linearly with E_0 . As the energy is

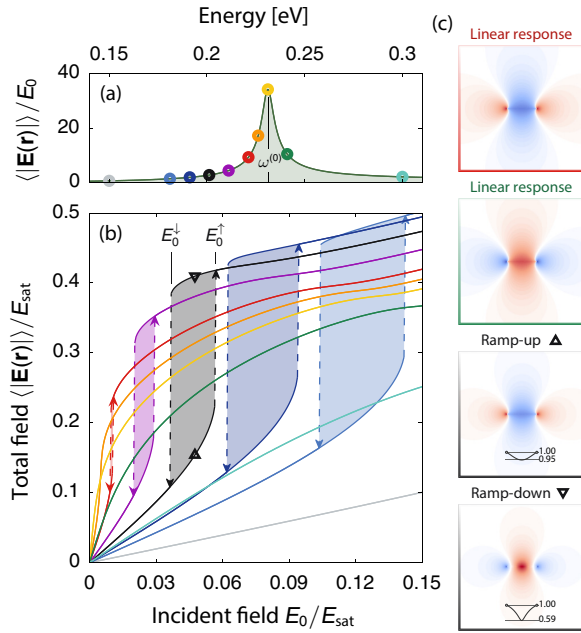


FIG. 3. (Color online) Hysteresis arising from bistable behavior in a $W = 25$ nm nanoribbon excited by a plane wave $E_0\hat{x}$ (material parameters are as in Fig. 1). (a) Linear response field-enhancement spectrum. Selected energies are highlighted by colored markers, and the linear resonance energy $\hbar\omega^{(0)}$ is labeled. (b) Hysteresis curves at fixed energy [corresponding colorwise to those in (a)] for total field $\langle |\mathbf{E}(\mathbf{r})| \rangle$ versus incident field E_0 . Bistable regions are indicated by shading and delimited by energy-dependent low- and high-point field strengths E_0^{\downarrow} and E_0^{\uparrow} . (c) Intensity maps of the induced electric field $\text{Re}[E_x^{\text{ind}}(x, z)]$ in a 50×50 nm² cross-sectional region. Color scale is identical across the four maps, ranging from positive (red) through zero (white) to negative (blue) in a symmetric range. Absolute magnitudes are scaled logarithmically for intelligibility. Frame color indicates association with energies in (a). Field strengths in the high-field maps are specified by corresponding triangles in (b). Sparklines, defined as in Fig. 1(b), indicate the range and variation of $|f(\mathbf{r})|$.

increased towards $\hbar\omega^{(0)}$, a nonlinear discrepancy develops with increasing E_0 which eventually gives way to a discontinuous jump at a critical field strength E_0^{\uparrow} , indicated for a selected energy in Fig. 3(b). As E_0 is reduced on the downward ramp, its response initially traces out that of the upward ramp, but departs from its upward counterpart after E_0^{\uparrow} and eventually undergoes a discontinuous jump at E_0^{\downarrow} after which the initial path is retraced. The hysteresis area, indicated by shaded areas, increases with positive $\omega^{(0)} - \omega$ (although E_0^{\uparrow} similarly increases, delaying the onset of hysteresis), but vanishes for $\omega \gtrsim \omega^{(0)}$ due to the red-shifting of the resonance with E_0 . Due to plasmonic field enhancement of the total field, the onset of bistability is reached for incident field strengths considerably below E_{sat} .

Lastly, we comment on the field profiles of the excitations. Specifically, we highlight the π phase shift that arises between the two bistable solutions in the black-framed maps. The phase

shift can be appreciated from a simple anharmonic oscillator model [31], in partial analogy to the shift exhibited by the red- and green-framed maps of the linear resonance (see SM [40]).

Summary and discussion. In this Rapid Communication, we have analyzed the impact of Kerr nonlinearity on the plasmonic response of graphene nanostructures, specifically for nanoribbons. The key distinction of nanostructures compared to the corresponding extended system arises from the strongly inhomogeneous fields of localized plasmonic excitations, which in turn incur an inhomogeneous conductive profile. We have derived a simple analytic expression (5), which approximates the nonlinear resonance shifts, while accounting for both inhomogeneity and overall amplitude of the nonlinear perturbation. The characteristic field of the Kerr nonlinearity in graphene is the saturation field E_{sat} . However, significant nonlinear interaction can be achieved near plasmonic resonances even for much weaker incident fields owing to plasmonic field enhancement. Finally, we discussed the existence of a plasmonic bistability in nanoribbons under normal incidence.

The applications of optical bistabilities are well known and long pursued [1,2], with implications particularly in optical switching. Indeed, a range of platforms have been scrutinized for this purpose, in recent years, e.g., in photonic crystal cavities (PCC) where nonlinearities are enhanced by large Q -factors and light slowdown [3]. Whether graphene can further the state of the art in this mature field remains to be seen [43]. We expect, however, that a very profitable avenue for progress exists in hybrid approaches, utilizing, e.g., PCC and graphene in unison, as has in fact been explored experimentally [21], albeit without taking advantage of the resonant plasmonic nonlinearity described herein. Advances in this direction require improved understanding of nonlinearities in nanostructures; this work constitutes one such effort. Several features, however, remain unexplored, underscoring the fertility and richness of the field. For example, from a semi-classical perspective, barring atomistic approaches [31,32], questions remain relating to the role of interband nonlinearities [18], nonlocality, and the effective role of edge states. The impact of substrate interactions, which affords, e.g., plasmon-phonon coupling [44,45], is similarly unexplored in nonlinear settings.

Penultimately, we highlight that graphene solitons, sustained by Kerr nonlinearities, also pose a number of intriguing opportunities [23–25]. Our present considerations actually generalize readily to treat temporal solitons through the 1D nonlinear Schrödinger equation. For instance, the group velocity dispersion and self-phase modulation parameters can be derived analytically in terms of $\lambda_n(k_i)$ (see SM [40]). There, too, inhomogeneity has a pronounced impact.

In closing, we mention a final question of singular practical relevance, namely, damage thresholds. So far, to the best of our knowledge, measurements do not exist in the infrared, but in the optical domain [46–48] the reported thresholds fall in the rather broad range from $\sim 10^6$ V/cm in fs-pulsed operation [47] to just $\sim 10^4$ V/cm for hour-long continuous-wave operation [46], while technologically important substrates such as SiO₂ exhibit thresholds up to $\sim 10^8$ V/cm [49]. For comparison, the saturation field at $\hbar\omega = \epsilon_F = 0.2$ eV is $E_{\text{sat}} \approx 6.7 \times 10^5$ V/cm. Although direct comparison is impossible,

in part due to frequency range, pulse conditions, and the uncertain impact of field enhancement, this highlights that even resonantly enhanced nonlinearities in graphene walk a narrow road, not unlike previous contenders for large nonlinearities. Given the promising results presented herein, however, we believe the journey will be worth the effort.

The Center for Nanostructured Graphene is sponsored by the Danish National Research Foundation, Project No. DNRF58. This work was also supported by the Danish Council for Independent Research, Project No. 1323-00087. W.Y. acknowledges support from the Lundbeck Foundation, Grant No. 70802.

- [1] R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic, Burlington, MA, 2008).
- [2] H. M. Gibbs, *Optical Instability: Controlling Light with Light* (Academic, New York, 1985).
- [3] M. Soljačić and J. D. Joannopoulos, *Nat. Mater.* **3**, 211 (2004).
- [4] J. M. Dudley and J. R. Taylor, *Nat. Photonics* **3**, 85 (2009).
- [5] J. Leuthold, C. Koos, and W. Freude, *Nat. Photonics* **4**, 535 (2010).
- [6] M. Kauranen and A. V. Zayats, *Nat. Photonics* **6**, 737 (2012).
- [7] Y. V. Bludov, A. Ferreira, N. M. R. Peres, and M. I. Vasilevskiy, *Int. J. Mod. Phys. B* **27**, 1341001 (2013).
- [8] F. J. García de Abajo, *ACS Photonics* **1**, 135 (2014).
- [9] T. Stauber, *J. Phys.: Condens. Matter* **26**, 123201 (2014).
- [10] T. Low and P. Avouris, *ACS Nano* **8**, 1086 (2014).
- [11] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, *Science* **306**, 666 (2004).
- [12] K. I. Bolotin, K. J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. L. Stormer, *Solid State Commun.* **146**, 351 (2008).
- [13] P. Tassin, T. Koschny, and C. M. Soukoulis, *Science* **341**, 620 (2013).
- [14] S. Thongrattanasiri and F. J. García de Abajo, *Phys. Rev. Lett.* **110**, 187401 (2013).
- [15] S. A. Mikhailov and K. Ziegler, *J. Phys.: Condens. Matter* **20**, 384204 (2008).
- [16] S. A. Mikhailov, *Physica E (Amsterdam)* **44**, 924 (2012).
- [17] X. Yao and A. Belyanin, *Phys. Rev. Lett.* **108**, 255503 (2012).
- [18] J. L. Cheng, N. Vermeulen, and J. E. Sipe, *New J. Phys.* **16**, 053014 (2014).
- [19] E. Hendry, P. J. Hale, J. Moger, A. K. Savchenko, and S. A. Mikhailov, *Phys. Rev. Lett.* **105**, 097401 (2010).
- [20] H. Zhang, S. Virally, Q. Bao, L. K. Ping, S. Massar, N. Godbout, and P. Kockaert, *Opt. Lett.* **37**, 1856 (2012).
- [21] T. Gu, N. Petrone, J. F. McMillan, A. van der Zande, M. Yu, G. Q. Lo, D. L. Kwong, J. Hone, and C. W. Wong, *Nat. Photonics* **6**, 554 (2012).
- [22] N. M. R. Peres, Y. V. Bludov, J. E. Santos, A.-P. Jauho, and M. I. Vasilevskiy, *Phys. Rev. B* **90**, 125425 (2014).
- [23] M. L. Nesterov, J. Bravo-Abad, A. Y. Nikitin, F. J. García-Vidal, and L. Martín-Moreno, *Laser Photon. Rev.* **7**, L7 (2013).
- [24] D. A. Smirnova, I. V. Shadrivov, A. I. Smirnov, and Y. S. Kivshar, *Laser Photon. Rev.* **8**, 291 (2014).
- [25] D. A. Smirnova, R. E. Noskov, L. A. Smirnov, and Y. S. Kivshar, *Phys. Rev. B* **91**, 075409 (2015).
- [26] H. Nasari and M. S. Abrishamian, *J. Opt. Soc. Am. B* **31**, 1691 (2014).
- [27] M. Gullans, D. E. Chang, F. H. L. Koppens, F. J. García de Abajo, and M. D. Lukin, *Phys. Rev. Lett.* **111**, 247401 (2013).
- [28] X. Yao, M. Tokman, and A. Belyanin, *Phys. Rev. Lett.* **112**, 055501 (2014).
- [29] D. A. Smirnova, I. V. Shadrivov, A. E. Miroshnichenko, A. I. Smirnov, and Y. S. Kivshar, *Phys. Rev. B* **90**, 035412 (2014).
- [30] A. V. Gorbach, *Phys. Rev. A* **87**, 013830 (2013).
- [31] J. D. Cox and F. J. García de Abajo, *Nat. Commun.* **5**, 5725 (2014).
- [32] J. D. Cox and F. J. García de Abajo, *ACS Photonics* **2**, 306 (2015).
- [33] X.-H. Wang, *Finite Element Methods for Nonlinear Optical Waveguides*, Advances in Nonlinear Optics Vol. 2 (Gordon and Breach, Amsterdam, 1995).
- [34] Although the third-order term is recovered, the saturation model is in contradiction at the fifth order, whose Kerr contribution is $\sigma_{(5)} = -\sigma_{(1)}|E|^4/E_{(5)}^4$ with $E_{(5)}^4 = (16\varpi_{(5)}^4)/(25\omega^4)E_{\text{sat}}^4$ and $\varpi_{(5)}^4 \equiv \varpi_{(3)}^2(\omega + \frac{1}{3}i\gamma)(\omega - \frac{1}{2}i\gamma)$. This, however, can in principle be easily remedied by going to the [0/4] Padé approximant.
- [35] W. Wang and J. M. Kinaret, *Phys. Rev. B* **87**, 195424 (2013).
- [36] T. Christensen, W. Wang, A.-P. Jauho, M. Wubs, and N. A. Mortensen, *Phys. Rev. B* **90**, 241414(R) (2014).
- [37] W. Wang, T. Christensen, A.-P. Jauho, K. S. Thygesen, M. Wubs, and N. A. Mortensen, *Sci. Rep.* **5**, 9535 (2015).
- [38] I. Silveiro, J. M. P. Ortega, and F. J. García de Abajo, *Light: Sci. Appl.* **4**, e241 (2015).
- [39] Rescaling $V(\mathbf{r}, \mathbf{r}') \rightarrow \varepsilon_{\text{B}}^{-1} V(\mathbf{r}, \mathbf{r}')$ with $\varepsilon_{\text{B}} \equiv \frac{1}{2}(\varepsilon_{-} + \varepsilon_{+})$ accounts for nonunity and/or dispersive sub- and superstrates $\varepsilon_{\pm}(\omega)$.
- [40] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.92.121407> for a description of the iterative procedure applied the nonlinear problem, matrix representations of V and D for the nanoribbon, derivation of the perturbation approach, discussion of the qualitative anharmonic oscillator model, and a semianalytical treatment of plasmonic solitons.
- [41] J. Christensen, A. Manjavacas, S. Thongrattanasiri, F. H. L. Koppens, and F. J. García de Abajo, *ACS Nano* **6**, 431 (2012).
- [42] K. A. Velizhanin, *Phys. Rev. B* **91**, 125429 (2015).
- [43] J. B. Khurgin, *Appl. Phys. Lett.* **104**, 161116 (2014).
- [44] M. Freitag, T. Low, W. Zhu, H. Yan, F. Xia, and P. Avouris, *Nat. Commun.* **4**, 1951 (2013).
- [45] X. Zhu, W. Wang, W. Yan, M. B. Larsen, P. Bøggild, T. G. Pedersen, S. Xiao, J. Zi, and N. A. Mortensen, *Nano Lett.* **14**, 2907 (2014).
- [46] B. Krauss, T. Lohmann, D.-H. Chae, M. Haluska, K. von Klitzing, and J. H. Smet, *Phys. Rev. B* **79**, 165428 (2009).
- [47] A. Roberts, D. Cormode, C. Reynolds, T. Newhouse-Illige, B. J. LeRoy, and A. S. Sandhu, *Appl. Phys. Lett.* **99**, 051912 (2011).
- [48] M. Currie, J. D. Caldwell, F. J. Bezares, J. Robinson, T. Anderson, H. Chun, and M. Tadjer, *Appl. Phys. Lett.* **99**, 211909 (2011).
- [49] W. H. Lowdermilk and D. Milam, *IEEE J. Quantum Electron.* **17**, 1888 (1981).

SUPPLEMENTAL MATERIAL

Kerr nonlinearity and plasmonic bistability in graphene nanoribbons

Thomas Christensen,^{1,2} Wei Yan,^{1,2} Antti-Pekka Jauho,^{2,3} Martijn Wubs,^{1,2} and N. Asger Mortensen^{1,2}

¹Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

²Center for Nanostructured Graphene, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

³Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

I. ITERATIVE PROCEDURE FOR NONLINEAR PROBLEM

We here discuss an iterative approach to solving the nonlinear equation

$$\lambda\phi(\mathbf{r}) = \lambda\phi_{\text{ext}}(\mathbf{r}) + \text{VD}[f[\phi]]\phi(\mathbf{r}), \quad (\text{S1})$$

which is essentially just the driven correspondent of Eq. (4), and where we have emphasized the dependence of D on $\phi(\mathbf{r})$ through $f(\mathbf{r})$. The problem is evidently nonlinear, but can be solved efficiently by iteration with only linear algebra at each step. We follow the usual iteration scheme, as e.g. also used previously in the studies of bistability in dielectric waveguides [S1].

1. Compute a linear solution based on an initial guess of $f = f_{\text{ini}}$, i.e. solve Eq. (S1) with $D[f[\phi]] \rightarrow D[f = f_{\text{ini}}]$. Denote the obtained solution as $\phi^{[0]}$. Set the iteration step $m = 0$.
2. Calculate the m th guess at the occupation function $f^{[m]}$ from the potential $\phi^{[m]}$.
3. Compute the $(m + 1)$ th iteration by solving the linear system $\lambda\phi^{[m+1]} = \lambda\phi_{\text{ext}} + \text{VD}[f^{[m]}\phi^{[m+1]}}$.
4. Iterate steps 2 and 3 until convergence, otherwise update iteration step $m \rightarrow m + 1$.

We impose convergence criteria corresponding to the simultaneous fulfillment of (with $\text{tol} = 10^{-5}$)

$$\max_{\mathbf{r} \in \Omega} |\phi^{[m+1]}(\mathbf{r}) - \phi^{[m]}(\mathbf{r})| / \max_{\mathbf{r} \in \Omega} |\phi^{[m]}(\mathbf{r})| < \text{tol}, \quad (\text{S2a})$$

$$\max_{\mathbf{r} \in \Omega} |f^{[m+1]}(\mathbf{r}) - f^{[m]}(\mathbf{r})| / \max_{\mathbf{r} \in \Omega} |f^{[m]}(\mathbf{r})| < \text{tol}, \quad (\text{S2b})$$

being of standard type for iterative approaches to nonlinearity [S1]. In all considered cases the iterative procedure converged after at most several hundred iterations. One exception should be mentioned however; the dipolar eigenmodes at field strengths 3×10^5 V/cm and 3.5×10^5 V/cm failed to converge after 1250 iterations for $k_{\parallel} \gtrsim 5$ and are consequently absent in Fig. 1 for these momenta. This could likely be remedied by a more elaborate stepping procedure, though such investigations have not been pursued further in this work.

Two additional extensions of the simple iterative scheme described above are employed. Firstly, for numerical stability we apply a linear mixing scheme for updating guesses on f , specifically we use $D[f_{\text{mix}}^{[m]}]$ with $f_{\text{mix}}^{[m]} = (1 - \xi_{\text{mix}})f^{[m-1]} + \xi_{\text{mix}}f^{[m]}$ in step 2 (mixing parameter $\xi_{\text{mix}} = 0.275$) rather than the unmixed $D[f^{[m]}]$. Secondly, the initial guess f_{ini} is always taken from the previous field strength in ramping scenarios. This provides a significant numerical speed-up and, crucially, allows us to investigate hysteresis and bistability. The initial guess at the first field strength is naturally $f_{\text{ini}} = 1$.

For eigenmodal calculations where $\phi_{\text{ext}} = 0$, we normalize ϕ_n at each iteration to impose the desired ribbon-averaged field strength $\langle |\mathbf{E}(\mathbf{r})| \rangle$, and in addition determine ω_n from $\lambda_n(\omega_n)$ by numerically solving the equation in the complex frequency-plane.

A. Ramping details

As discussed in the main text, we compute, for each fixed energy $\hbar\omega$, solutions to Eq. (S1) for a ramp-array of many field-strengths, going both up and down. We here explicate the rather straightforward details

of this array: consider for each $\hbar\omega$ a ramp-array $\{E_{0,n}\}_{n=1}^N$ with $E_{0,n+1} > E_{0,n}$ and with $E_{0,1}$ sufficiently small to be considered a linear perturbation. Starting from $E_{0,1}$ we compute associated solutions and proceed, generally, to field strength $E_{0,n+1}$ with initial guesses on f and ϕ (provided to step 1 of the iterative procedure discussed in Section I) obtained from the n th solution. This defines the upward ramp, corresponding to slowly turning the incident intensity up. Upon reaching $n = N$ we invert the procedure and follow a downward ramp, in the pattern $E_{0,n} \rightarrow E_{0,n-1}$, corresponding to slowly turning the intensity down.

II. MATRIX REPRESENTATION OF V AND D IN A DISCRETIZED BASIS

We here elaborate the reduction of the differential and integral operators D and V to matrix representations \mathbf{D} and \mathbf{V} using an equidistant discrete basis. Specifically, we discuss the 1D ribbon case, although the generalization to general 2D restrictions is straightforward. Specifically, we imagine a system in the xy -plane, translationally invariant along y and with finite extent along x . For simplicity, we assume just a single ribbon, such that x is limited to the simple domain $x \in [0, 1]$. Furthermore, as the operators necessarily act on a potential $\phi(\mathbf{r})$, we impose translational invariance along y by the decomposition $\phi(\mathbf{r}) = \phi(x)e^{ik_{\parallel}y}$.

Starting with the differential operator D , we consider its operation onto $\phi(\mathbf{r})$, which takes the form $D\phi(\mathbf{r}) = \partial_x[f(x)\partial_x\phi(x)]e^{ik_{\parallel}y} - k_{\parallel}^2 f(x)\phi(x)e^{ik_{\parallel}y}$. By extension, we define the operation of D onto the single-variable function $\phi(x)$ through $D\phi(x) \equiv \partial_x[f(x)\partial_x\phi(x)] - k_{\parallel}^2 f(x)\phi(x)$. To proceed, we introduce a discretization of the x -coordinates as $\{x_j\}_{j=1}^N$ with associated values $\phi_j \equiv \phi(x_j)$ and $f_j \equiv f(x_j)$ (we take $N = 150$, being well-converged in all considered cases). Though not strictly necessary, we assume equidistant x_j with constant spacing $x_{j+1} - x_j = a$, see Fig. S1.

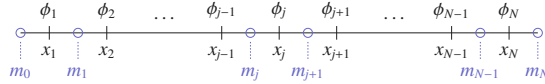


FIG. S1 Sketch of the discretization approach applied to a single ribbon.

The matrix elements D_{jl} of the finite-element representation of D is then defined by $D\phi_j = \sum_l D_{jl}\phi_l$. The elements can be deduced using finite differences at the midpoints. Specifically, using central differences $\partial_x[f_j\partial_x\phi_j] \simeq a^{-1}(m_j - m_{j-1})$ where m_j defines midpoint-values of the function $m(x) \equiv f(x)\partial_x\phi(x)$ such that $m_j \simeq (2a)^{-1}(f_{j+1} + f_j)(\phi_{j+1} - \phi_j)$, see Fig. S1. For all interior points, $j \in [2, N - 1]$, this then allows a decomposition of D_{jl} as the tridiagonal matrix

$$D_{jl} = \frac{1}{2a^2} [\delta_{j-1,l}(f_{j-1} + f_j) - \delta_{jl}(f_{j-1} + 2f_j + f_{j+1}) + \delta_{j+1,l}(f_j + f_{j+1})] - \delta_{jl}k_{\parallel}^2 f_j. \quad (\text{S3a})$$

At the end-points $j = 1$ and $j = N$ we explicitly account for boundary conditions. Specifically, we ensure a vanishing of normal current, equivalent to the condition $\partial_x\phi(x) = 0$ for $x = 0$ and $x = 1$. In turn, this forces $m_0 = m_N = 0$, allowing

$$D_{1l} = \frac{1}{2a^2}(f_1 + f_2)(-\delta_{1,l} + \delta_{2,l}) - \delta_{1,l}k_{\parallel}^2 f_1, \quad (\text{S3b})$$

$$D_{Nl} = \frac{1}{2a^2}(f_{N-1} + f_N)(\delta_{N-1,l} - \delta_{N,l}) - \delta_{N,l}k_{\parallel}^2 f_N. \quad (\text{S3c})$$

As an alternative to taking explicit account of the boundary condition, one can allow a slightly larger x -range, and explicitly include points with $f(\mathbf{r}) = 0$ outside $\mathbf{r} \in \Omega$ – the step in $f(\mathbf{r})$ at $\mathbf{r} \in \partial\Omega$ then mimics an edge charge and accounts numerically for the boundary condition; such a procedure may be preferable in finite structures without any geometric symmetries compatible with a square grid.

The integral operator V is similarly amenable to explicit expression on the equidistant grid. Specifically, letting V operate on a function $g(\mathbf{r}) = g(x)e^{ik_{\parallel}y}$ one finds [S2; S3]

$$Vg(\mathbf{r}) = e^{ik_{\parallel}y} \int dx' 2K_0(k_{\parallel}|x - x'|)g(x'), \quad (\text{S4})$$

where $k_{\parallel} > 0$ is assumed and with K_0 denoting the zeroth order modified Bessel function of the second kind. Assuming a slowly varying $g(x)$ and an equidistant $\{x_j\}$ then allows a matrix decomposition of \mathbf{V} via $Vg_j = \sum_l V_{jl}g_l$ where [S4]

$$V_{jl} = 2 \int_{x_l-a/2}^{x_l+a/2} dx' K_0(k_{\parallel}|x_j - x'|) = \pi \sum_{\tilde{x}=x_{jl}\pm a/2} \tilde{x} \{K_0(k_{\parallel}|\tilde{x}|) [\mathbf{L}_1(k_{\parallel}|\tilde{x}|) + \frac{2}{\pi}] + K_1(k_{\parallel}|\tilde{x}|) \mathbf{L}_0(k_{\parallel}|\tilde{x}|)\}, \quad (\text{S5})$$

with $x_{jl} \equiv x_j - x_l$ and $\mathbf{L}_{0,1}$ denoting modified Struve functions of zeroth and first order.

A final detail which should be discussed is the special case $k_{\parallel} = 0$, where the kernel $K_0(k_{\parallel}|x - x'|)$ in Eq. (S5) diverges. Despite this divergence, finite and meaningful matrix elements can be retrieved by invoking charge conservation. Specifically, we note the small argument expansion $K_0(k_{\parallel}|x - x'|) \sim -\ln(|x - x'|) - \ln(k_{\parallel}) + \alpha$ where $\alpha = \ln(2) - \gamma_{\text{EM}}$ (γ_{EM} is the Euler–Mascheroni constant) [S4]. The x' -independent term $-\ln(k_{\parallel}) + \alpha$ gives a contribution $[-\ln(k_{\parallel}) + \alpha] \int dx' g(x')$ to Eq. (S4) and appears divergent as $k_{\parallel} \rightarrow 0$. Nevertheless, this contribution vanishes for the functions $g(\mathbf{r}')$ of relevance since they always represent induced charges [as evident from Eq. (3)] and obey charge conservation $\int dx' g(x') = 0$. As such, the $k_{\parallel} = 0$ case can be calculated by simply letting $K_0(k_{\parallel}|x - x'|) \rightarrow -\ln(|x - x'|)$ in Eq. (S5) [S3], yielding $V_{jl} = -2 \sum_{s=\pm} s(x_{jl} + s\frac{a}{2}) \ln(|x_{jl} + s\frac{a}{2}|)$ for $k_{\parallel} = 0$.

This concludes the real-space discretization approach for reduction of the abstract operator equation of Eq. (4) into a matrix equation $\lambda\phi = \mathbf{VD}\phi$ with ϕ denoting the vector form of ϕ_j .

III. PERTURBATION ESTIMATE OF THE NONLINEAR SHIFT OF EIGENFREQUENCIES

We here provide the derivations that allow the approximate result of Eq. (5). As we explain below, the approach relies on the formulation of a Hermitian eigenproblem followed by application of standard perturbation theory to a spatially inhomogeneous problem.

The compound operator \mathbf{VD} defined in Eq. (4) is – though numerically practical – inconvenient for analytical considerations, because it is not symmetric. However, the problem can (of course) be cast as a Hermitian eigenproblem with eigenvalues λ_n [though, strictly speaking, only for real, positive occupation functions $f(\mathbf{r})$, which we restrict our analysis to here], as also noted recently in Refs. [S5; S6]. Specifically, consider the application of the scaled gradient operation $-\sqrt{f(\mathbf{r})}\nabla$ onto Eq. (3):

$$-\lambda \sqrt{f(\mathbf{r})}\nabla\phi(\mathbf{r}) = \sqrt{f(\mathbf{r})}\nabla \int_{\Omega} d^2\mathbf{r}' V(\mathbf{r}, \mathbf{r}')\nabla' \cdot \{ \sqrt{f(\mathbf{r}')} [-\sqrt{f(\mathbf{r}')} \nabla' \phi(\mathbf{r}')]\}. \quad (\text{S6})$$

Defining the scaled in-plane field $\xi(\mathbf{r}) \equiv -\sqrt{f(\mathbf{r})}\nabla\phi(\mathbf{r})$ and manipulating further allows

$$\begin{aligned} \lambda\xi(\mathbf{r}) &= \sqrt{f(\mathbf{r})}\nabla \int_{\Omega} d^2\mathbf{r}' V(\mathbf{r}, \mathbf{r}')\nabla' \cdot [\sqrt{f(\mathbf{r}')} \xi(\mathbf{r}')] \\ &\stackrel{a}{=} \sqrt{f(\mathbf{r})}\nabla \left\{ \int_{\Omega} d^2\mathbf{r}' \nabla' \cdot [V(\mathbf{r}, \mathbf{r}') \sqrt{f(\mathbf{r}')} \xi(\mathbf{r}')] - \int_{\Omega} d^2\mathbf{r}' [\nabla' V(\mathbf{r}, \mathbf{r}')] \cdot [\sqrt{f(\mathbf{r}')} \xi(\mathbf{r}')] \right\} \\ &\stackrel{b}{=} -\sqrt{f(\mathbf{r})}\nabla \int_{\Omega} d^2\mathbf{r}' \sqrt{f(\mathbf{r}')} [\nabla' V(\mathbf{r}, \mathbf{r}')] \cdot \xi(\mathbf{r}') \\ &\stackrel{c}{=} - \int_{\Omega} d^2\mathbf{r}' \sqrt{f(\mathbf{r})f(\mathbf{r}')} [\nabla \otimes \nabla' V(\mathbf{r}, \mathbf{r}')] \xi(\mathbf{r}') \end{aligned} \quad (\text{S7})$$

with associated steps $a - c$ explicated below for convenience:

- a. Application of chain rule to expand integrand.
- b. The first integral term in step a vanishes, as can be deduced by application of the divergence theorem which transforms the term to $\oint_{\partial\Omega} V(\mathbf{r}, \mathbf{r}') \sqrt{f(\mathbf{r}')} [\xi(\mathbf{r}') \cdot \mathbf{n}']$. The integrand vanishes for all $\mathbf{r}' \in \partial\Omega$ due to the no-spill boundary condition on the induced current which forces $\xi(\mathbf{r}') \cdot \mathbf{n}' = 0$ on $\mathbf{r}' \in \partial\Omega$.
- c. The term $\sqrt{f(\mathbf{r})}\nabla$ is taken under the integral sign. ∇ operates on \mathbf{r} and hence only on $V(\mathbf{r}, \mathbf{r}')$. The operation $\nabla \cdot [\nabla' V(\mathbf{r}, \mathbf{r}') \cdot \nu(\mathbf{r}')]$ is rewritten in the equivalent outer-product form $[\nabla \otimes \nabla' V(\mathbf{r}, \mathbf{r}')] \nu(\mathbf{r}')$ with elements $[\nabla \otimes \nabla']_{ij} = \partial_r \partial_{r'_j}$.

We then define the operator M by its action on a field-ket $|\xi\rangle$ [where, as usual, $\langle \mathbf{r}|\xi\rangle \equiv \xi(\mathbf{r})$]

$$\langle \mathbf{r}|M|\xi\rangle \equiv \int_{\Omega} d^2\mathbf{r}' \sqrt{f(\mathbf{r})f(\mathbf{r}')} [\nabla \otimes \nabla' V(\mathbf{r}, \mathbf{r}')] \xi(\mathbf{r}'), \quad (\text{S8})$$

with associated eigenspectrum $(-\lambda_n)$ and $|\xi_n\rangle$:

$$(-\lambda_n)|\xi_n\rangle = M|\xi_n\rangle. \quad (\text{S9})$$

The operator M is evidently symmetric, positive semi-definite, and thus Hermitian. Accordingly, the eigenspectrum $\{-\lambda_n\}$ is non-negative and real; and the eigenkets $|\xi_n\rangle$ are orthogonal $\langle \xi_n|\xi_{n'}\rangle = \delta_{nn'}$ and span the solution space for $\mathbf{r} \in \Omega$.

With these facts established, we can now discuss a perturbation treatment. Specifically, we consider the simple case where $f(\mathbf{r}) = f^{(0)} + \delta f^{(1)}(\mathbf{r})$ for $\mathbf{r} \in \Omega$ with ‘‘groundstate’’ $f^{(0)} = 1$ and perturbation $f^{(1)}$ with strength δ . The corresponding expansion of $M = M^{(0)} + \delta M^{(1)} + \mathcal{O}(\delta^2)$ is found by expansion of Eq. (S8), yielding

$$\langle \mathbf{r}|M^{(0)}|\xi\rangle = \int_{\Omega} d^2\mathbf{r}' [\nabla \otimes \nabla' V(\mathbf{r}, \mathbf{r}')] \xi(\mathbf{r}'), \quad (\text{S10a})$$

$$\langle \mathbf{r}|M^{(1)}|\xi\rangle = \frac{1}{2} \int_{\Omega} d^2\mathbf{r}' [f^{(1)}(\mathbf{r}) + f^{(1)}(\mathbf{r}')] [\nabla \otimes \nabla' V(\mathbf{r}, \mathbf{r}')] \xi(\mathbf{r}'). \quad (\text{S10b})$$

Since M is a Hermitian operator usual perturbation theory applies [S7]. Specifically, for a ‘‘groundstate’’ eigenspectrum $\{-\lambda_n^{(0)}, |\xi_n^{(0)}\rangle\}$ the leading-order correction to the perturbed eigenvalue $\lambda_n = \lambda_n^{(0)} + \delta\lambda_n^{(1)} + \mathcal{O}(\delta^2)$ is derivable by application of Eqs. (S10) [by using the $(\mathbf{r}, \mathbf{r}')$ -symmetry of the resulting equation]

$$\lambda_n^{(1)} = -\frac{\langle \xi_n^{(0)}|M_1|\xi_n^{(0)}\rangle}{\langle \xi_n^{(0)}|\xi_n^{(0)}\rangle} = \lambda_n^{(0)} \frac{\langle \xi_n^{(0)}|f^{(1)}|\xi_n^{(0)}\rangle}{\langle \xi_n^{(0)}|\xi_n^{(0)}\rangle}.$$

For nonlinear purposes, we unfortunately do not know the exact perturbation $f^{(1)}$ as it should be determined self-consistently with the total field $|\xi_n\rangle$. However, for low field-strengths this self-consistency can be neglected and we can approximate $f[|\xi_n\rangle] \simeq f[|\xi_n^{(0)}\rangle]$ with $|\xi_n^{(0)}\rangle$ referring to the electric field predicted by a *linear* calculation (at the desired field strength). For the Kerr-type nonlinearity of Eq. (1) the resulting correction is therefore [assuming vanishingly small loss and noting $\xi^{(0)}(\mathbf{r}) = \mathbf{E}^{(0)}(\mathbf{r})$ for $f^{(0)} = 1$]

$$\lambda_n^{(1)} \simeq -\lambda_n^{(0)} \frac{9}{8} \frac{\int_{\Omega} d^2\mathbf{r} |\mathbf{E}^{(0)}(\mathbf{r})|^4}{E_{\text{sat}}^2 \int_{\Omega} d^2\mathbf{r} |\mathbf{E}^{(0)}(\mathbf{r})|^2} = -\lambda_n^{(0)} \frac{9}{8} \frac{\langle |\mathbf{E}^{(0)}(\mathbf{r})|^4 \rangle}{E_{\text{sat}}^2 \langle |\mathbf{E}^{(0)}(\mathbf{r})|^2 \rangle}, \quad (\text{S11})$$

with E_{sat} similarly evaluated at the linear resonance frequency $\omega_n^{(0)}$ associated with $\lambda_n^{(0)}$. Finally, the result of the main text, Eq. (5), is obtained by invoking the relation between eigenvalues λ_n and eigenfrequencies ω_n together with the lossless intraband conductivity $\sigma_{(1)}(\omega) \simeq ie^2\epsilon_f/\pi\hbar^2\omega$.

IV. QUALITATIVE ANHARMONIC OSCILLATOR MODEL

We review the basics of the simple anharmonic oscillator model [S8; S9], and discuss how it – in connection with a polarizability consideration – explains the π phase-shift observed for the bistable solutions in Fig. 3(c).

Before considering the nonlinear problem, we note first that the field profiles depicted in the red- and green-framed maps of Fig. 3(c), corresponding to energies just below and above the resonance energy $\hbar\omega^{(0)}$, exhibit the well-known π phase shift between each other. The phase-shift can be appreciated e.g. by inspection of the linear harmonic-oscillator polarizability $\alpha(\omega) \propto [(\omega^{(0)})^2 - \omega(\omega + i\gamma)]^{-1}$ which exhibits a sign-change of its real part as ω traverses the resonance at $\omega^{(0)}$: as a result, the induced dipole $p(\omega) = \alpha(\omega)E_0$ changes sign for $\omega \lesssim \omega^{(0)}$, and correspondingly so for the induced fields. As noted in the main text, a similar phase-shift is observed in the bistable comparison, see black-framed modes in Fig. 3(c). Again, the origin of the sign change can be appreciated from a polarizability consideration by including a third-order anharmonic term to the harmonic oscillator model [S9]; we do this below.

In this qualitative model, we represent the induced dipole by a single (time-dependent) coordinate x , which obeys the simple equation of motion

$$m\ddot{x} + m\gamma\dot{x} = -efE_0(t) - \partial_x U(x), \quad (\text{S12})$$

with an effective anharmonic restoring potential $U(x) = \frac{1}{2}m(\omega^{(0)})^2x^2 - \frac{1}{4}max^4$, effective oscillator mass m , linear resonance $\omega^{(0)}$, anharmonic parameter a (note that $a > 0$ in our case cf. sign of Kerr conductivity), and coupling factor f . We seek the solution that oscillates at $e^{-i\omega t}$ in response to a perturbation $E_0(t) = E_0(\omega)e^{-i\omega t}$, i.e. the Kerr response; we denote this term by $x^{(1\omega)}(\omega)e^{-i\omega t}$. Working with Eq. (S12) one finds (omitting declaration of ω -dependence)

$$m[(\omega^{(0)})^2 - \omega(\omega + i\gamma) - 3a|x^{(1\omega)}|^2]x^{(1\omega)} = -efE_0. \quad (\text{S13})$$

The polarizability $\alpha^{(1)}$ is linked to $x^{(1\omega)}$ via the induced dipole $p^{(1\omega)} = -ex^{(1\omega)} = \alpha^{(1\omega)}E_0$, allowing (ignoring loss, being nonessential for the present considerations)

$$[(\omega^{(0)})^2 - \omega^2 - 3ae^{-2}|\alpha^{(1\omega)}|^2E_0^2]\alpha^{(1\omega)} = e^2f/m. \quad (\text{S14})$$

For the bistable scenarios the term $(\omega^{(0)})^2 - \omega^2$ is always positive, see e.g. Figs. 2 and 3. Depending on the magnitude of $3ae^{-2}\alpha^{(1\omega)}E_0^2$ relative to $(\omega^{(0)})^2 - \omega^2$ it is then clear that polarizability-solutions of opposing sign can arise, depending on the sign of the terms bracketed on the left-hand side of Eq. (S14). Furthermore, if we denote the positive and negative solutions $\alpha_+^{(1\omega)}$ and $\alpha_-^{(1\omega)}$, respectively, it can then be deduced by direct inspection of Eq. (S14) that $|\alpha_+^{(1\omega)}| < |\alpha_-^{(1\omega)}|$. In other words, the induced dipole – and hence the induced fields – of the positive solution should be lower than its negative counterpart; upon identifying the lower branches of Fig. 3(b) with $\alpha_+^{(1)}$ and vice versa for the upper branch, we see that this is exactly the case. As such, the anharmonic model describes not only the phase-shift, but also the magnitude interrelationship. Lastly, we mention for completeness that the anharmonic model describes also a third solution, which, however, is physically irrelevant as it is unstable (and correspondingly is not found in the iterative procedure employed in this study, nor in experimental investigation).

V. PLASMONIC SOLITONS AND THE NONLINEAR SCHRÖDINGER EQUATION

As a simple, practical example of our general considerations, we here discuss how our results can be applied to study 1D plasmonic solitons within the framework of the 1D nonlinear Schrödinger equation (NLSE). Specifically, pulse propagation along the ribbon's y -direction (i.e. along k_{\parallel}) can be well-described by the NLSE under the same assumptions underlying its use in nonlinear fiber optics [S10]. For a slowly varying pulse (in y) with center frequency ω_0 and associated center momentum k_{\parallel}^0 , and under the assumption of negligible propagation loss the NLSE reads [S8; S10]

$$i\frac{\partial \tilde{A}(y, \tau)}{\partial y} - \frac{\beta_2}{2}\frac{\partial^2 \tilde{A}(y, \tau)}{\partial \tau^2} + \gamma_{\text{nl}}|\tilde{A}(y, \tau)|^2\tilde{A}(y, \tau) = 0, \quad (\text{S15})$$

expressed in the retarded time-frame $\tau = t - \frac{\partial k_{\parallel}}{\partial \omega}|_0 y$ with $|_0$ indicating evaluation in the low-field limit at center frequency and momentum ω_0 and k_{\parallel}^0 with all complementary parameters held fixed. Here $\tilde{A}(y, \tau) = A(y, t)$ is the y -dependent envelope function of the field amplitude $E(y, t) = A(y)e^{ik_{\parallel}^0 y - i\omega_0 t}$. In the 1D treatment, $E(y, t)$ corresponds physically to the y -dependence of the x -averaged amplitude. Finally, parameters β_2 and γ_{nl} give the group velocity dispersion $\beta_2 \equiv \frac{\partial^2 k_{\parallel}}{\partial \omega^2}|_0$ and the nonlinear parameter $\gamma_{\text{nl}} \equiv \frac{\partial k_{\parallel}}{\partial |A|^2}|_0$.

Our main point here is to highlight that the coefficients β_2 and γ_{nl} can be analytically expressed for moderate field strengths in terms of eigenvalues $\lambda_n(k_{\parallel})$, inhomogeneity parameter κ , saturation field E_{sat} , and ribbon setup ϵ_f and W . In particular, working in the intraband approximation where $\hbar\omega(k_{\parallel}) = \hbar\Omega\sqrt{-\lambda(k_{\parallel})}$, with $\hbar\Omega \equiv (2\pi)^{-1}\sqrt{e^2\epsilon_f/\epsilon_0 W}$, allows obtention of

$$\beta_2 = -\frac{4}{\Omega^2}\left(\frac{\partial \lambda(k_{\parallel})}{\partial k_{\parallel}}\Big|_0\right)^{-1}\left[\frac{1}{2} - \lambda(k_{\parallel}^0)\frac{\partial^2 \lambda(k_{\parallel})}{\partial k_{\parallel}^2}\Big|_0\left(\frac{\partial \lambda(k_{\parallel})}{\partial k_{\parallel}}\Big|_0\right)^{-2}\right]. \quad (\text{S16})$$

Similarly, using $d\omega = \frac{\partial\omega}{\partial k_{\parallel}} dk_{\parallel} + \frac{\partial\omega}{\partial |\tilde{A}|^2} d|\tilde{A}|^2 = 0$ and Eq. (5), allows expression of γ_{nl}

$$\begin{aligned} \gamma_{nl} &= \frac{\partial k_{\parallel}}{\partial |\tilde{A}|^2} \Big|_0 = - \frac{\partial\omega}{\partial |\tilde{A}|^2} \Big|_0 \left(\frac{\partial\omega}{\partial k_{\parallel}} \Big|_0 \right)^{-1} \quad \text{where} \quad \frac{\partial\omega}{\partial |\tilde{A}|^2} \Big|_0 = \frac{-9 \kappa \omega_0}{16 E_{\text{sat}}^2} \quad \text{and} \quad \frac{\partial\omega}{\partial k_{\parallel}} \Big|_0 = \frac{-\Omega}{2[-\lambda(k_{\parallel}^0)]^{1/2}} \frac{\partial\lambda(k_{\parallel})}{\partial k_{\parallel}} \Big|_0, \\ &= \frac{9}{8} \frac{\kappa}{E_{\text{sat}}^2} \lambda(k_{\parallel}^0) \left(\frac{\partial\lambda(k_{\parallel})}{\partial k_{\parallel}} \Big|_0 \right)^{-1}, \end{aligned} \quad (\text{S17})$$

where we have used that $\langle |\mathbf{E}^{(0)}|^2 \rangle \simeq |\tilde{A}|^2$ to linear order, and where κ denotes the inhomogeneity parameter defined by $\langle |\mathbf{E}^{(0)}|^4 \rangle \equiv \kappa \langle |\mathbf{E}^{(0)}|^2 \rangle^2$. Note that κ is also (weakly) momentum-dependent and hence evaluated at k_{\parallel}^0 ; higher-order corrections $\propto \frac{\partial\kappa}{\partial k_{\parallel}} \Big|_0$ are neglected.

Crucially, for negative β_2 Eq. (S15) exhibits an analytical (bright) soliton solution, with spatial phase factor $\delta k_{\parallel} \equiv -\beta_2/2\tau_0^2$ [S8]

$$\tilde{A}(y, \tau) = \tilde{A}^0 \text{sech}(\tau/\tau_0) e^{i\delta k_{\parallel} y}, \quad (\text{S18})$$

provided the pulse width τ_0 and maximum amplitude \tilde{A}^0 are interrelated by

$$N^2 \equiv \frac{\gamma_{nl} |\tilde{A}^0|^2 \tau_0^2}{|\beta_2|} = 1. \quad (\text{S19})$$

In Fig. S2 we plot $\beta_2 \Omega^2$ and $\gamma_{nl} E_{\text{sat}}^2$ – which are dimensionless, universal functions of the ribbon, independent of setup parameters ϵ_r and W – as functions of k_{\parallel}^0 for the first few eigenmodes of the nanoribbon ($n = 0, 1, \dots$, corresponding to monopole, dipole, etc.). Apart from the monopole and dipole, the modes exhibit simultaneously positive γ_{nl} and negative β_2 in the entire or most of the considered k_{\parallel}^0 -region [restricted to small k_{\parallel}^0 to ensure validity of Eq. (S15)]. Accordingly, fundamental soliton solutions, of the type in Eq. (S18), are allowed for these modes for appropriate values of $|\tilde{A}^0|^2 \tau_0^2$. The dipole mode exhibits a small region of feasible soliton parameters for larger $k_{\parallel}^0 \gtrsim 0.49$.

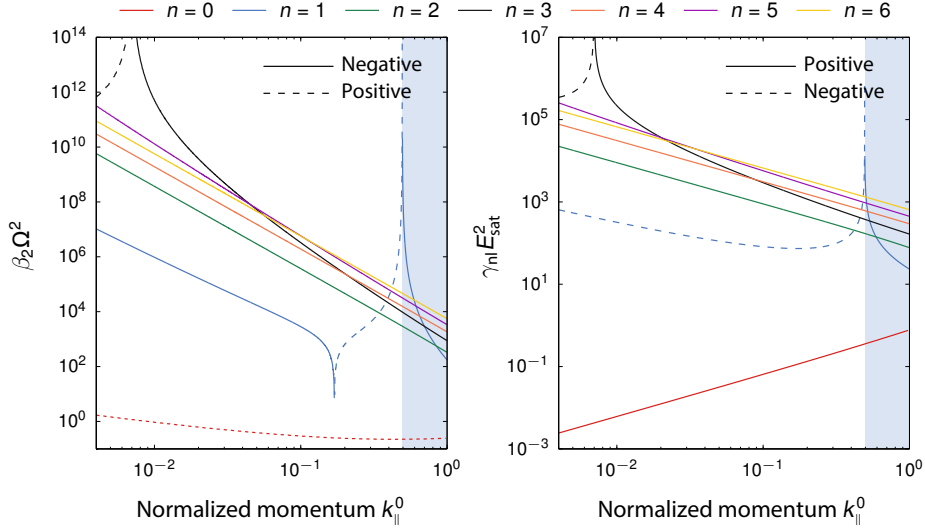


FIG. S2 Universal parameters $\beta_2 \Omega^2$ and $\gamma_{nl} E_{\text{sat}}^2$ for the n th eigenmodes of the nanoribbon. Sign is indicated by line style. Momentum regions where full lines coexist, i.e. where $\beta_2 < 0$ and $\gamma_{nl} > 0$, indicate feasible soliton regimes. This region is highlighted for the dipole in gray shading.

References

- [S1] X.-H. Wang, *Finite Element Methods for Nonlinear Optical Waveguides*, Advances in Nonlinear Optics, Vol. 2 (Gordan and Breach Publishers, 1995).
- [S2] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. (Elsevier Academic Press, 2007).
- [S3] W. Wang and J.M. Kinaret, *Phys. Rev. B* **87**, 195424 (2013).
- [S4] National Institute of Standards and Technology, "Digital library of mathematical functions," Release 1.0.9 of 2014-08-29.
- [S5] K.A. Velizhanin, *Phys. Rev. B* **91**, 125429 (2015).
- [S6] F.J. García de Abajo and A. Manjavacas, *Faraday Discuss.* **178**, 87 (2015).
- [S7] S. Gasiorowicz, *Quantum physics*, 3rd ed. (John Wiley & Sons, 2003).
- [S8] R.W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic Press, 2008).
- [S9] J.D. Cox and F.J. García de Abajo, *Nat. Commun.* **5**, 5725 (2014).
- [S10] G.P. Agrawal, *J. Opt. Soc. Am. B* **28**, A1 (2011).

PHYS. REV. B **91**, 125414 (2015)

T. Christensen developed the initial idea, performed all calculations, and wrote the manuscript with input from all authors.

Localized plasmons in graphene-coated nanospheres

Thomas Christensen,^{1,2} Antti-Pekka Jauho,^{2,3} Martijn Wubs,^{1,2} and N. Asger Mortensen^{1,2,*}

¹*Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*

²*Center for Nanostructured Graphene, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*

³*Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*

(Received 22 December 2014; revised manuscript received 12 February 2015; published 10 March 2015)

We present an analytical derivation of the electromagnetic response of a spherical object coated by a conductive film, here exemplified by a graphene coating. Applying the framework of Mie-Lorenz theory augmented to account for a conductive boundary condition, we derive the multipole scattering coefficients, modified essentially through the inclusion of an additive correction in numerator and denominator. By reductionist means, starting from the retarded response, we offer simple results in the quasistatic regime by analyzing the multipolar polarizability and associated dispersion equation for the localized plasmons. We consider graphene coatings of both dielectric and conducting spheres, where the graphene coating in the former case introduces the plasmons and in the latter case modifies in interesting ways the existing ones. Finally, we discuss our analytical results in the context of extinction cross section and local density of states. Recent demonstrations of fabricated spherical graphene nanostructures make our study directly relevant to experiments.

DOI: [10.1103/PhysRevB.91.125414](https://doi.org/10.1103/PhysRevB.91.125414)

PACS number(s): 78.20.Bh, 78.67.Wj, 73.20.Mf, 78.20.Ci

I. INTRODUCTION

The study of interaction between electromagnetic fields and graphene has seen a riveting development in recent years. In particular, efforts have centered on oscillation energies near the experimentally achievable Fermi level of graphene, typically in the sub-eV range, where the principal features are due to the excitation of either propagating or localized two-dimensional plasmons [1–3]. A large variety of structural configurations has been investigated, ranging from, e.g., (semi-)finite structures [4–7] to periodic arrays [8–10]. Of these studies, the overwhelming majority exhibit the common assumption of structural planarity. Recently, interest has emerged also in exploring the properties of plasmonic interaction in curved configurations, e.g., propagating modes in bent and corrugated sheets [11], in the context of cloaking [12,13], and in various coated nanowire systems [14–17].

In this paper, we examine the archetypal curved graphene structure: a sphere, highly amenable to analytics and exhibiting the key features necessary to gain clear physical insight in the role of curvature. The spherical geometry is also of experimental relevance, given recent fabrication demonstrations. Notably, demonstrations include reduced graphene-oxide hollow spheres [18], graphene encapsulation of hollow SnO₂ spheres of radii down to ~50 nm [18], and porous multilayer graphene nanospheres supported by a polystyrene interior [19]. Though the graphene in these recent demonstrations exhibits several deviations from an idealized two-dimensional spherical coating, it underscores the relevance of the geometry beyond a theoretical perspective. At the opposite end of the size spectrum, the fullerenes represent a tempting analogy. However, it is now well established that larger fullerenes, such as C₃₂₀ and beyond, prefer faceted, predominately icosahedral rather than spherical configurations [20,21]. Additionally, the use of semiclassical, local response functions in graphene is reasonable only for structures in excess of ~10⁴ carbon

atoms [22]. In this paper, we take a classical, local surface conductivity approach, and as such we expect our predictions to be reliable chiefly for radii larger than ~5 nm (a graphene sphere of N atoms has a radius of $R \sim \sqrt{N} \times 0.457 \text{ \AA}$), significantly beyond the range of fullerenes.

The paper is structured as follows: for graphene described by a local surface conductivity, we show in Sec. II A that its electromagnetic response can be understood through a modified Mie-Lorenz theory. From the asymptotic limit of these results to the quasistatic realm, we derive the multipolar polarizability in Sec. II B and identify the multipole plasmon conditions in the quasistatic regime, which are particularly transparent. Specifically, we show that an intuitive, effective momentum mapping connects the sphere resonances with those of an infinite plane. In Sec. III we present calculations for graphene-coated dielectric spheres, considering first the size dispersion of the multipole resonances. Next, comparing two observables, the extinction cross section and the local density of states (LDOS), we highlight the physical significance of the dipole and higher-order multipoles. In Sec. IV we study the interaction between a localized plasmon supported by a Drude sphere, e.g., a doped semiconductor, and the plasmon supported by the graphene coating. We close our treatment of coated Drude spheres by discussing a corollary of the formalism related to surficial damping in metal plasmonics. Finally, we summarize the results in Sec. V.

II. THEORETICAL DESCRIPTION

Within the local-response approximation (LRA), the electric field $\mathbf{E}(\mathbf{r}, \omega)$ in a homogeneous region \mathcal{V}_j with dielectric constant $\varepsilon_j(\omega)$ satisfies the homogeneous Helmholtz wave equation:

$$\nabla^2 \mathbf{E}(\mathbf{r}, \omega) - k_0^2 \varepsilon_j(\omega) \mathbf{E}(\mathbf{r}, \omega) = 0, \quad (1)$$

where $k_0 = \omega/c$ denotes the free-space wave number and where $\varepsilon_j(\omega)$ denotes the effective LRA dielectric constant, potentially exhibiting a frequency dependence.

* asger@mailaps.org

A. Retarded solution by expansion in vector waves

We solve Eq. (1) by expansion in vector wave functions, as befitting for structures with curvilinear symmetries [23]. In particular, within the LRA the electric field is divergence free, or solenoidal, in which case the monochromatic solutions of the electric field in a homogeneous region \mathcal{V}_j can be expanded in the basis of the solenoidal vector wave functions $\mathbf{M}_\nu^{[i]}(\mathbf{r})$ and $\mathbf{N}_\nu^{[i]}(\mathbf{r})$ of that region:

$$\mathbf{E}(\mathbf{r}) = \sum_{i\nu} a_\nu \mathbf{M}_\nu^{[i]}(\mathbf{r}) + b_\nu \mathbf{N}_\nu^{[i]}(\mathbf{r}), \quad (2)$$

with ν denoting a geometry-dependent expansion index, while i denotes expansion over in- and outgoing waves, and, finally, with $a_\nu^{[i]}$ and $b_\nu^{[i]}$ denoting associated expansion coefficients. The functions $\mathbf{M}_\nu^{[i]}(\mathbf{r})$ and $\mathbf{N}_\nu^{[i]}(\mathbf{r})$ describe the TE and TM parts, respectively, of the electric field, and describe the propagation of transverse modes (cf. their solenoidality). In spherical coordinates $\mathbf{r} = (r, \theta, \varphi)$ the index ν partitions into polar and azimuthal quantum numbers, $l \in [1, \infty[$ and $m \in [-l, l]$, with associated vector waves (usually referred to as multipoles) [23]:

$$\mathbf{M}_{lm}^{[i]}(\mathbf{r}) = \nabla \times \mathbf{r} \psi_{lm}^{[i]}(\mathbf{r}), \quad (3a)$$

$$\mathbf{N}_{lm}^{[i]}(\mathbf{r}) = \frac{1}{k} \nabla \times \nabla \times \mathbf{r} \psi_{lm}^{[i]}(\mathbf{r}), \quad (3b)$$

defined in terms of the scalar generating functions $\psi_{lm}^{[i]}(r, \theta, \varphi) = z_l^{[i]}(kr) P_l^m(\cos \theta) e^{im\varphi}$, where $z_l^{[i]}$ denotes spherical Bessel or Hankel functions (of the first kind), j_l and $h_l^{(i)}$, for $i = 1$ and 2 , respectively, corresponding to in- and outgoing waves. The wave number $k \equiv k_0 \sqrt{\varepsilon_j}$ relates the dimensionless argument kr with the material properties. An additional class of vector wave functions exists, denoted $\mathbf{L}_\nu(\mathbf{r})$, which are irrotational. These vector waves are needed, e.g., in the description of longitudinal modes arising in nonlocal response or in the presence of sources, but are irrelevant in homogeneous media described by the LRA [24–26].

Here we consider the specific case of a two-component spherically symmetric system, centered at origo, coated by a conductive film at the bulk-component interface at radius R , as indicated in Fig. 1. We assume that the system is illuminated from a source in the external region, denoted \mathcal{V}_2 , by the ingoing field $\mathbf{E}^{\text{inc}}(\mathbf{r})$. The incident field excites an outgoing scattered field, $\mathbf{E}^{\text{sca}}(\mathbf{r})$, in \mathcal{V}_2 , and an ingoing transmitted field, $\mathbf{E}^{\text{tra}}(\mathbf{r})$, in the interior region, denoted \mathcal{V}_1 . Explicitly, the field inside and outside the sphere is expanded via

$$\mathbf{E}_{\mathcal{V}_1}(\mathbf{r}) = \sum_{lm} a_{lm}^{\text{tra}} \mathbf{M}_{lm}^{[1]}(\mathbf{r}) + b_{lm}^{\text{tra}} \mathbf{N}_{lm}^{[1]}(\mathbf{r}), \quad r < R, \quad (4a)$$

$$\begin{aligned} \mathbf{E}_{\mathcal{V}_2}(\mathbf{r}) = & \sum_{lm} a_{lm}^{\text{inc}} \mathbf{M}_{lm}^{[1]}(\mathbf{r}) + b_{lm}^{\text{inc}} \mathbf{N}_{lm}^{[1]}(\mathbf{r}) \\ & + \sum_{lm} a_{lm}^{\text{sca}} \mathbf{M}_{lm}^{[2]}(\mathbf{r}) + b_{lm}^{\text{sca}} \mathbf{N}_{lm}^{[2]}(\mathbf{r}), \quad r > R, \end{aligned} \quad (4b)$$

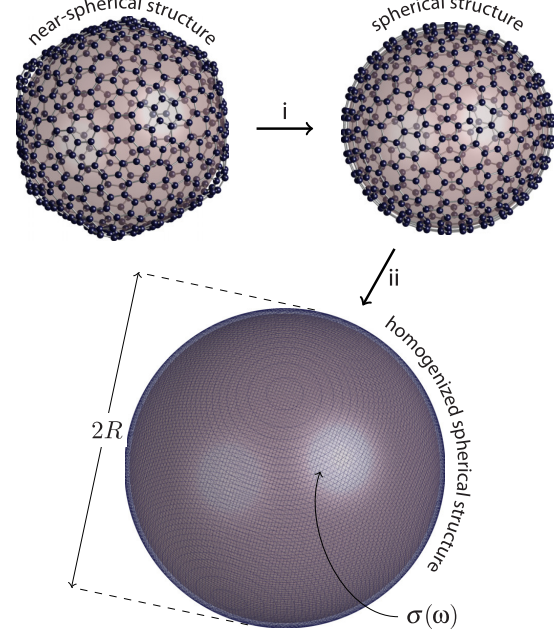


FIG. 1. (Color online) Schematic illustration of the introduced conceptual simplifications in the treatment of optical response of graphene nanospheres via a surface conductivity approach. Specifically at step (i) any aspherical elements are neglected, while at step (ii) the microscopic details of the structure are replaced by the bulk response function $\sigma(\omega)$. Here depicted for a C_{540} fullerene, for illustrative purposes solely.

where regions \mathcal{V}_j are implicitly associated with wave numbers $k_j = k_0 \sqrt{\varepsilon_j}$.

The transmitted and scattered amplitudes, $\{a_{lm}^{\text{tra}}, b_{lm}^{\text{tra}}\}$ and $\{a_{lm}^{\text{sca}}, b_{lm}^{\text{sca}}\}$, are linearly proportional to the incident amplitudes, $\{a_{lm}^{\text{inc}}, b_{lm}^{\text{inc}}\}$. Their interrelation is dictated by the boundary conditions (BCs) at the domain interface at $r = R$. In the presence of a conductive surface at $r = R$ a finite surface current \mathbf{K} is introduced, in which case the BCs read as $\hat{\mathbf{n}} \times (\mathbf{E}_{\mathcal{V}_2} - \mathbf{E}_{\mathcal{V}_1}) = 0$ and $\hat{\mathbf{n}} \times (\mathbf{H}_{\mathcal{V}_2} - \mathbf{H}_{\mathcal{V}_1}) = \mathbf{K}$ at all surficial points. We take the induced surface current at a surficial point \mathbf{r} , with associated normal $\hat{\mathbf{n}}$, as linearly related to the tangential field $\mathbf{E}_\parallel(\mathbf{r})$, constructed such that $\mathbf{E}_\parallel(\mathbf{r}) \cdot \hat{\mathbf{n}} = 0$, via an LRA surface conductivity $\sigma(\omega)$, such that $\mathbf{K}(\mathbf{r}) = \sigma(\omega) \mathbf{E}_\parallel(\mathbf{r})$.

Enforcing these BCs translates into local, linear relations between the scattered and incident amplitudes:

$$a_{lm}^{\text{sca}} = t_l^{\text{TE}} a_{lm}^{\text{inc}} \delta_{ll'} \delta_{mm'}, \quad b_{lm}^{\text{sca}} = t_l^{\text{TM}} b_{lm}^{\text{inc}} \delta_{ll'} \delta_{mm'}, \quad (5)$$

where the proportionality constants, often referred to as Mie-Lorenz scattering coefficients, are given by

$$t_l^{\text{TE}} = \frac{-j_l(x_1)[x_2 j_l(x_2)]' + j_l(x_2)[x_1 j_l(x_1)]' - g(\omega)x_0^2 j_l(x_1) j_l(x_2)}{j_l(x_1)[x_2 h_l^{(1)}(x_2)]' - h_l^{(1)}(x_2)[x_1 j_l(x_1)]' + g(\omega)x_0^2 j_l(x_1) h_l^{(1)}(x_2)}, \quad (6a)$$

$$t_l^{\text{TM}} = \frac{-x_2^2 j_l(x_2)[x_1 j_l(x_1)]' + x_1^2 j_l(x_1)[x_2 j_l(x_2)]' + g(\omega)x_0^2 [x_1 j_l(x_1)]' [x_2 j_l(x_2)]'}{x_2^2 h_l^{(1)}(x_2)[x_1 j_l(x_1)]' - x_1^2 j_l(x_1)[x_2 h_l^{(1)}(x_2)]' - g(\omega)x_0^2 [x_1 j_l(x_1)]' [x_2 h_l^{(1)}(x_2)]'}, \quad (6b)$$

written in terms of the dimensionless argument $x_j \equiv k_j R$ for $j = \{0, 1, 2\}$, and where the influence of the conductive surface is included via the dimensionless parameter:

$$g(\omega) \equiv \frac{i\sigma(\omega)}{\epsilon_0 \omega R}. \quad (7)$$

Naturally, for vanishing surface conductivity $g(\omega) \rightarrow 0$, whereby the solution reduces to the standard Mie-Lorenz coefficients [27].

B. The multipolar polarizability and the nonretarded plasmon dispersion

The amplitudes in Eq. (6) give the fully retarded response. However, as is well known, the quasistatic limit is excellent in the context of plasmonic excitations in graphene when $\hbar\omega/\epsilon_F \gg \alpha_{fs} \approx 1/137$ [28]. For optical interactions in the quasistatic regime, the multipolar polarizability, α_l , constitutes the central object, and can be derived from the TM Mie-Lorenz coefficients via [25]

$$\alpha_l = -4\pi i \frac{l![(2l+1)!!]^2}{(l+1)(2l+1)} \lim_{x_0 \rightarrow 0} \left[\frac{t_l^{\text{TM}}}{k_2^{2l+1}} \right], \quad (8)$$

with !! denoting the double factorial. From this we can derive (using the small-argument asymptotic expansions of the spherical Bessel functions) the multipolar polarizability in the quasistatic limit:

$$\alpha_l = 4\pi R^{2l+1} \frac{l[\epsilon_1 - \epsilon_2 + (l+1)g(\omega)]}{l\epsilon_1 + (l+1)\epsilon_2 + l(l+1)g(\omega)}. \quad (9)$$

This expression is naturally highly reminiscent of the well-known result for the polarizability of a two-component spherical system [29,30], but extended via $g(\omega)$ to account for the presence of a conductive surface.

The plasmonic resonances of the system are obtained from the poles of the Mie-Lorenz coefficients of Eq. (8), or, in the quasistatic regime, from the poles of the polarizability of Eq. (9). In the latter case, we can derive an uncomplicated resonance condition for the l -order multipolar plasmon at frequency ω_l , extending the Fröhlich condition to account for a conductive surface contribution:

$$l\epsilon_1 + (l+1)\epsilon_2 + l(l+1)g(\omega_l) = 0. \quad (10)$$

Though usually—in the absence of a coating—the existence of a plasmon requires $\epsilon_1\epsilon_2 < 0$, it is evident that plasmons may exist here even when $\epsilon_1, \epsilon_2 > 0$ provided that $g(\omega_l)$ is sufficiently negative, achievable for a surface conductivity with $\text{Im}(\sigma) < 0$.

For a uniform background, $\epsilon_1 = \epsilon_2 = \epsilon$, the condition is particularly elucidating, reading

$$\frac{2i\epsilon_0\epsilon\omega_l}{\sigma(\omega_l)} = \left(1 + \frac{1}{2l+1}\right) \frac{l}{R} \equiv q_l^{\text{eff}}. \quad (11)$$

Here we have cast the condition in the equivalent form as that of the infinite sheet plasmon condition [31], whereby we are able to identify an effective momentum q_l^{eff} , which, rather suggestively, approaches l/R asymptotically as $l \rightarrow \infty$, as a consequence of the modes perceiving the curving surface as increasingly flat with higher l and concomitant

shorter wavelengths [26,30]. For the optically important dipole resonance, we find $q_1^{\text{eff}} = \frac{4}{3}R^{-1}$.

The identification of an effective momentum suggests a phenomenological approach to incorporate the effects of nonlocal response (momentum dispersion), by substituting $\sigma(\omega) \rightarrow \sigma(q_l^{\text{eff}}, \omega)$, with the latter expression obtainable, e.g., from the noninteracting polarizability [32,33]. However, though such a speculative approach certainly is alluring, it would constitute an overextension of the momentum analogy. Indeed, upon including nonlocal response through its hydrodynamic approximation one finds that the correct effective momentum takes a form $q_l^{\text{eff,H}} \equiv \sqrt{l(l+1)}/R$, clearly distinct from q_l^{eff} . For completeness we discuss the inclusion of hydrodynamic response in the conductive coating in the Appendix, whose contribution can be accounted for by a straightforward rescaling of the local-response conductivity.

III. GRAPHENE COATING OF DIELECTRIC SPHERES

The analysis so far is valid for any spectral dependence of $\sigma(\omega)$ (or, indeed, of ϵ_1 or ϵ_2). For a graphene-coated system, we take $\sigma(\omega)$ as graphene's bulk LRA conductivity, which for a Fermi level ϵ_F at finite temperature T reads as $\sigma(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega)$ [34,35]:

$$\frac{\sigma_{\text{intra}}(\omega)}{\sigma_0} = \frac{2ik_B T}{\hbar\tilde{\omega}} \ln \left[2 \cosh \left(\frac{\epsilon_F}{2k_B T} \right) \right], \quad (12a)$$

$$\frac{\sigma_{\text{inter}}(\omega)}{\sigma_0} = \frac{\pi}{4} H \left(\frac{1}{2} \hbar\tilde{\omega} \right) + i\hbar\tilde{\omega} \int_0^\infty d\epsilon \frac{H(\epsilon) - H(\frac{1}{2}\hbar\tilde{\omega})}{\hbar^2\tilde{\omega}^2 - 4\epsilon^2}, \quad (12b)$$

with definitions $\tilde{\omega} \equiv \omega + i\gamma_g$ where γ_g denotes the optical loss rate of graphene [38,39], $\sigma_0 \equiv e^2/\pi\hbar$ is the quantum of conductance, and $H(\epsilon)$ is the population difference between energies $\mp\epsilon$:

$$H(\epsilon) = \frac{\sinh(\epsilon/k_B T)}{\cosh(\epsilon_F/k_B T) + \cosh(\epsilon/k_B T)}. \quad (12c)$$

In the ensuing subsections we consider nondispersive bulk media, i.e., spectrally constant ϵ_1 and ϵ_2 , that is, dielectrics. In this case, the existence of localized plasmons is strictly due to the graphene coating. In Sec. IV we explore a dispersive interior, concretized by a graphene-coated Drude sphere, with the accompanying plasmons emerging from the interaction of the plasmon branches of each bare subsystem.

A. Size dispersion and modal profile

In Fig. 2(a) we investigate the size dispersion of the plasmonic modes of graphene spheres in vacuum in the low-temperature, low-loss limit, by solving Eq. (10) numerically. It is evident that for large spheres and sufficiently low l the intraband, low-loss approximation, $\sigma(\omega) \simeq \sigma_{\text{intra}}(\omega)|_{T=0}^{\gamma_g=0} = ie^2\epsilon_F/\pi\hbar^2\omega$, is a good approximation, yielding the dispersion $\hbar\omega_l \simeq [e^2\epsilon_F/\pi\epsilon_0\epsilon_l^B R]^2$ with $\epsilon_l^B \equiv \epsilon_1/(l+1) + \epsilon_2/l$. For smaller spheres, and concomitant larger resonance frequencies, the interband term redshifts the resonances significantly [40]. Furthermore, since $\text{Im}(\sigma_{\text{intra}} + \sigma_{\text{inter}})$ changes sign from positive to negative at $\hbar\omega/\epsilon_F \approx 1.6671$ the LRA predicts plasmon resonances restricted to the range below this

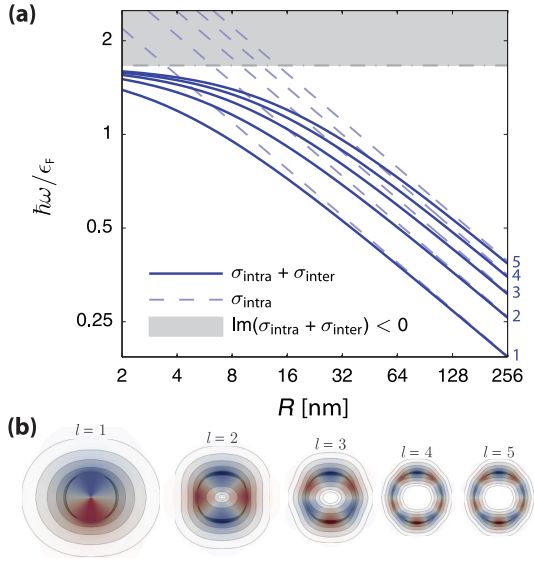


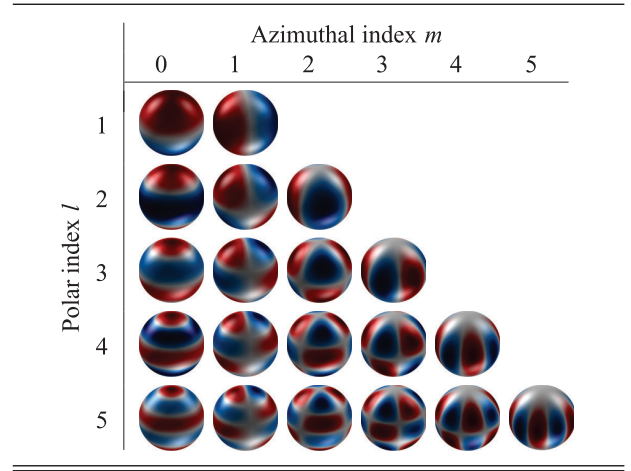
FIG. 2. (Color online) (a) Dispersion of the plasmon resonance frequency as a function of sphere radius for the first five multipole plasmons (l indicated in blue text) depicted in a doubly logarithmic plot. Resonances are calculated for lossless graphene spheres with Fermi level $\epsilon_F = 0.4$ eV in vacuum ($\epsilon_1 = \epsilon_2 = 1$) at zero temperature. The intraband approximation (discussed in the text) is indicated in dashed blue lines, while the region of negative imaginary part of the conductivity, $\hbar\omega/\epsilon_F \gtrsim 1.6671$, is indicated in gray shading. (b) Induced multipole modal profiles in the xz plane, calculated for $R = 20$ nm and $m = 1$ at resonance. Indicated is $|\mathbf{E}|$ in contours (separated by factors of 2) and $\text{Re}(E_\theta)$ in blue and red, corresponding to positive and negative, respectively.

frequency—though the inclusion of nonlocal response relaxes this restriction [32,33]. As seen from Fig. 2(b) the electric fields associated with each multipole plasmon are increasingly confined to the surface region with increasing l , in analogy with the increasing confinement experienced by a surface-plasmon polariton with increasing momentum. In line with this analogy, the plasmon modes also exhibit a monotonically decreasing circumferential wavelength, displaying l nodal lines of E_θ in the xz plane for the l th mode. In general, as shown in Table I, the induced charge exhibits exactly l nodal lines, regardless of the value of m .

B. Extinction and dipolar response

For direct optical measurements the quantity of interest is typically the extinction cross section, which is obtainable from either the Mie-Lorenz coefficients or, approximately, from the dipolar polarizability via [27] $\sigma_{\text{ext}} = 2\pi k_2^{-2} \sum_{l=1}^{\infty} (2l+1) \text{Re}(t_l^{\text{TE}} + t_l^{\text{TM}}) \simeq k_2 \text{Im}(\alpha_1) + (6\pi)^{-1} k_2^4 |\alpha_1|^2$. As is evident from the quasistatic approximation of σ_{ext} only the dipole plasmon influences the cross section in small spheres. In Fig. 3 we show the extinction cross-sectional efficiency of graphene-coated spheres of vacuum and polystyrene, surrounded externally by vacuum. Indeed, it is evident that the dipole approximation is excellent, even for graphene-coated spheres of several hundred nanometers. As already

TABLE I. (Color online) Representation of the induced charge profiles of the lm -multipole plasmons. The charge profile is $\rho_{lm}(\theta, \varphi) \propto P_l^m(\cos \theta) e^{im\varphi}$, of which we here depict the real part of the latter. Negative m values differ from their positive counterparts only by rotational direction.



observed in Fig. 2(a), the resonance position is redshifted with increasing radius, leading to a size-dependent extinction cross section. This stands in contrast to the resonances of metallic nanospheres which, in the classic quasistatic picture, exhibit size-independent resonances (though the inclusion of nonlocal response introduces a size dependence [41]). The inclusion of a nonunity dielectric as the spherical substrate redshifts and lowers the overall response efficiency as seen from the calculations for coated polystyrene spheres ($\epsilon_1 = 2.4$). This is consistent with the redshift generally arising from a reduction

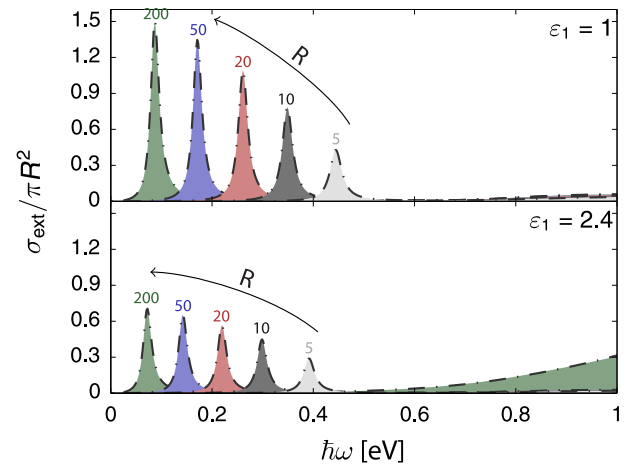


FIG. 3. (Color online) Extinction cross-sectional efficiency for graphene spheres in vacuum of varying radii $R = 5, 10, 20, 50,$ and 200 nm (indicated in color), with Fermi level $\epsilon_F = 0.4$ eV, loss rate $\hbar\gamma_g = 20$ meV, and temperature $T = 300$ K. Top and bottom panels consider the interior spheres consisting of vacuum and polystyrene ($\epsilon_1 = 1$ and 2.4), respectively. The quasistatic dipole approximation as well as fully retarded calculations are presented, here shown in dashed black lines and colored areas, respectively.

of the effective Coulomb interaction $1/\epsilon_l^B R$ (since ϵ_l^B increases with ϵ_2).

Lastly, we comment on the maximum values of the extinction cross section, which, for the nanoscopic spheres considered here, is chiefly due to absorption with only negligible scattering contributions. Though the achieved maximum cross section of the dipole is far below the single-channel limit of absorption (at $3\lambda^2/8\pi$), as relevant, e.g., in the pursuit of superscattering [42], a substantial resonant absorption is evident in Fig. 3. In particular, above- or near-unity cross-sectional efficiencies are achieved, constituting a noteworthy surpassing of the comparatively low $\sim 2.3\%$ absorptance of pristine graphene [3]. The size dependence of the maximum cross-sectional efficiency (attained at the dipole resonance), equaling approximately $\sigma_{\text{ext}}/\pi R^2 \simeq \frac{24}{(2+\epsilon_1)^2} \frac{\text{Im}(\sigma)^2}{\epsilon_0 c \text{Re}(\sigma)}$ in the low-loss regime, is strictly due to interband dispersion: upon neglecting interband terms and temperature dependence, the maximum attains the size-independent value of $\sigma_{\text{ext}}/\pi R^2 \simeq \alpha \frac{96}{(2+\epsilon_1)^2} \frac{\epsilon_F}{\hbar \gamma_g}$, with $\alpha \approx 1/137$ denoting the fine-structure constant.

C. LDOS and multipolar response

Exploring the properties of plasmons beyond the dipole resonance is best facilitated by near-field measurements whose exciting fields (and sampling profile) are not plane waves [26]. Such nonplanar exciting fields are naturally associated with nearby dipole emitters, such as dyes. The interplay between emitter and plasmonic system, leading, e.g., to decay enhancement [28], is then governed by the electric LDOS [43]. As for the cross section, the LDOS enhancement, i.e., the LDOS near the nanosphere, ρ^E , relative to the LDOS in free space, ρ_0^E , can be obtained from the Mie-Lorenz coefficients—or, more simply, from the multipolar polarizability. In particular, in the quasistatic limit, the emitter-orientation-averaged LDOS enhancement reads as [26,44]

$$\frac{\langle \rho^E \rangle}{\rho_0^E} = 1 + \frac{1}{8\pi k_2^3} \sum_{l=1}^{\infty} (l+1)(2l+1) \frac{\text{Im}(\alpha_l)}{(R+\delta)^{2(l+2)}}, \quad (13)$$

evaluated at radial distance δ from the sphere surface.

In Fig. 4 we consider the spectral dependence of the orientation averaged LDOS at varying distances from the spherical coating. At large surface-to-probe separations the LDOS is dominated by the dipolar plasmon, whilst at increasingly shorter separations the high-order multipoles appear as significant features. The LDOS evaluated at just 2 nm above the surface reveals a broad multiple multipole feature. This broad feature is comprised of several spectrally overlapping multipole plasmons, that are increasingly congested as the energies approach $\approx 1.6671\epsilon_F$. Once more, this effect has a close analog in metallic nanospheres where local response incurs a pile up of multipole modes near the flat-interface surface-plasmon resonance at $\omega_p/\sqrt{2}$, with ω_p denoting the metallic plasma frequency [26,30].

IV. GRAPHENE COATING OF DRUDE SPHERES

Proceeding from the study of a nondispersive interior, we consider next a graphene-coated Drude sphere, wherein we

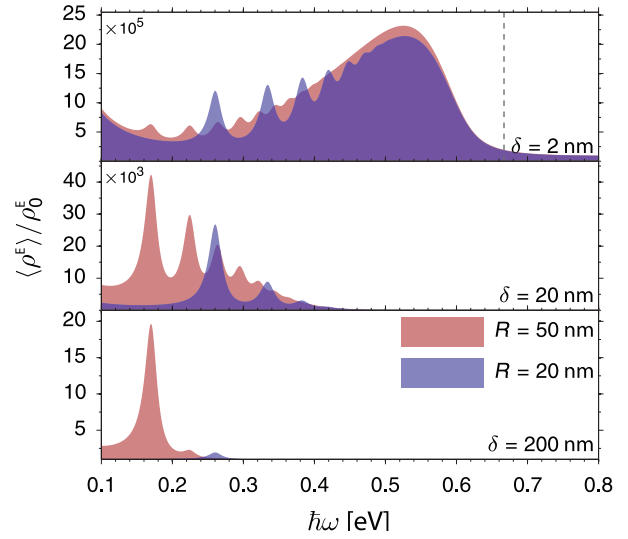


FIG. 4. (Color online) LDOS enhancement for graphene spheres in vacuum of varying radii $R = 20$ and 50 nm, with setup otherwise as in Fig. 3. The surface-to-observation distance δ is indicated in each panel. In the top panel, the $l \rightarrow \infty$ multipole asymptote at $\hbar\omega/\epsilon_F \approx 1.6671$ is indicated by the dashed line. For each panel, the y axis ranges from unity and upwards.

assign the interior dielectric function a Drude form:

$$\epsilon_1(\omega) = \epsilon_\infty - \frac{\omega_p^2}{\omega(\omega + i\gamma_D)}, \quad (14)$$

where ϵ_∞ gives the residual high-frequency response of the material, and γ_D gives the optical loss rate of the Drude material. The Drude dispersion is traditionally applied to metals, but also reliably describes strongly doped semiconductors much larger in extent than the Fermi wavelength [45,46]. The case of doped semiconductor spheres is significantly more interesting from the perspective of mode hybridization, as the range of plasma frequencies of doped semiconductors overlaps the realizable Fermi energies of graphene.

In this case where the interior is dispersive and well described by Eq. (14) the resonances of the coated system then follow directly from Eq. (10). If we include only the low-temperature intraband response of graphene, via $\sigma_{\text{intra}}(\omega) = ie^2\epsilon_F/\pi\hbar^2(\omega + i\gamma_g)$, the dipole resonance condition is particularly simple, reading as

$$\frac{\omega_p^2}{\omega(\omega + i\gamma_D)} + \frac{\omega_{gR}^2}{\omega(\omega + i\gamma_g)} = \epsilon_\infty + 2\epsilon_2, \quad (15a)$$

where

$$\omega_{gR}^2 \equiv \frac{2e^2\epsilon_F}{\epsilon_0\pi\hbar^2 R} \quad (15b)$$

plays the role of an effective plasma frequency of the graphene coating. To first order in the loss rates (γ_D, γ_g) we find a single positive resonance frequency $\omega \equiv \omega_R - i\omega_I$ with [47]

$$\omega_R \simeq \sqrt{\frac{\omega_p^2 + \omega_{gR}^2}{\epsilon_\infty + 2\epsilon_2}}, \quad \omega_I \simeq \frac{1}{2} \frac{\omega_{gR}^2\gamma_g + \omega_p^2\gamma_D}{\omega_{gR}^2 + \omega_p^2}. \quad (16)$$

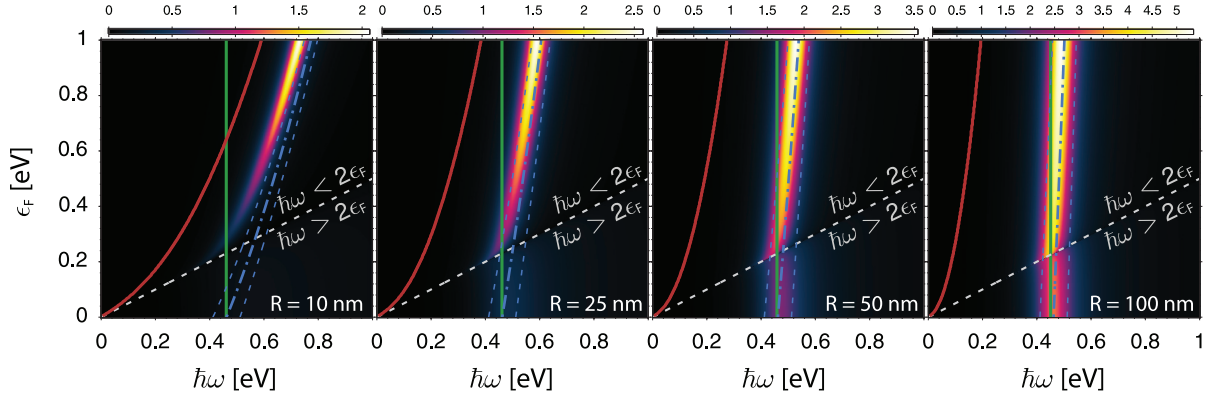


FIG. 5. (Color online) Extinction cross-sectional efficiency in a graphene-coated Drude sphere in vacuum, explored as a function of graphene Fermi energy and frequency, for several fixed values of the radius (as indicated). The Drude material is characterized by its plasma frequency $\hbar\omega_p = 0.8$ eV and loss rate $\hbar\gamma_D = 0.1$ eV, while graphene parameters are evaluated with $\hbar\gamma_g = 20$ meV at $T = 0$ K. The color coded response indicates the extinction efficiency, $\sigma_{\text{ext}}/\pi R^2$, calculated from the fully retarded expression. The green and red lines indicate the retarded bare resonance positions of the Drude sphere and graphene-coated vacuum sphere, respectively. In blue is given the quasistatic intraband approximation, Eq. (16), with the dash-dotted line indicating the resonance position and the dashed lines indicating the resonance width via $\omega_R \pm \omega_I$. Finally, the dashed white line separates the regions of zero and nonzero Landau damping.

The result bears a high resemblance with the standard quasistatic dipole resonance of a Drude sphere, but here lucidly adapted to account for the graphene coating through ω_{gR}^2 . Considering ω_I it is interesting to note that for $\gamma_D \neq \gamma_g$ it is possible to achieve an effectively reduced plasmon decay rate compared, e.g., to the uncoated Drude sphere by appropriate scaling of ω_{gR} relative to ω_p : specifically, if $\omega_{gR}^2/\omega_p^2 \equiv a$ then $\omega_I \simeq \frac{1}{2(a+1)}(a\gamma_g + \gamma_D)$, illustrating that the decay can be tuned from predominately graphene- to Drude-like by varying the ratio a . Evidently, the comparative importance of graphene's response relative to the bulk Drude material's is indicated by the a , with the former dominant when $a > 1$.

In Fig. 5 we explore the response of a graphene-coated Drude sphere, conceptually representative of a doped semiconductor, by considering the extinction cross-sectional efficiency. The plasma frequency considered, $\hbar\omega_p = 0.8$ eV, overlaps with the considered graphene Fermi energy range. The intraband approximation in Eq. (16) plainly provides an excellent account of both the position and width of the dipole resonance in the region $\hbar\omega < 2\epsilon_F$, i.e., outside the region of interband Landau damping, that is, for sufficiently large spheres at sufficiently high graphene Fermi energies. Notably, the hybridization between the bare resonances of the Drude sphere and a graphene-coated vacuum sphere leads to just a single hybridized mode, rather than the familiar construction of a red- and blueshifted bonding and antibonding mode. In the joint Drude-graphene case, due to the absence of spatial separation between the induced charge regions in the two materials—both residing at the sphere surface—just a single hybridized plasmon is formed. As is evident from Eq. (16) the hybridized resonance is blueshifted compared to the bare resonances. Significant tunability is achievable by varying either the sphere radius or, for dynamical purposes, graphene's doping level. The latter scenario could be achieved, e.g., by application of an external gating field, with significant

retainment of tunability expected [48], even in the presence of a fixed substrate charge-transfer contribution [49].

A. Relation to surface scattering

At this point we digress briefly from considerations of graphene coatings, to consider an endearing corollary of Eq. (10) in the dipole case related to surface scattering and Kreibig's size-dependent damping model. Specifically, suppose that a metallic particle, well described by Eq. (14), exhibits a slightly increased damping rate $\tilde{\gamma}_D = \gamma_D + \delta\gamma$ near the surface, e.g., due to roughness. We assume that this region is thin; specifically, it is reasonable to take its width as a single plasma wavelength $w = v_F/\omega_p$. We include this thin region approximately via a surface conductivity $\sigma = \sigma_{\text{bulk}}w$, where σ_{bulk} denotes the bulk Drude conductivity of loss rate $\tilde{\gamma}_D$. In this case, working from Eq. (10), one finds to first order in the loss rates and in the ratio w/R that the resonance broadens as $\text{Im}(\omega) \simeq -\frac{1}{2}[\gamma_D + (2\delta\gamma/\omega_p)(v_F/R)]$, which follows exactly the Kreibig form [50], $\gamma \rightarrow \gamma + Av_F/R$, with dimensionless damping parameter $A = 2\delta\gamma/\omega_p$. From experimental studies, it is well known that A is on the order of unity [50]—with this in mind, we recognize that surface scattering due to a spatial dependence of γ is only a minor contributor to the experimentally measured A , since $\delta\gamma/\omega_p \ll 1$ for any reasonable imagined $\delta\gamma$. Indeed, it was established by Apell and Penn, using density functional theory, that the primary contributor to A arises from density inhomogeneity in the surface region [51].

V. SUMMARY

In this paper we have examined the electromagnetic response—and, as a key element, the plasmonic properties—of a two-component spherical structure, coated at the interface by a conductive film, exemplified here by a graphene coating.

Within the naturally suited framework of vector waves we derived a corresponding set of generalized Mie-Lorenz scattering coefficients. Restricting our considerations to the quasistatic regime, we derived mathematically uncomplicated dispersion equations for the multipole plasmons. Considering the modest number of established analytical dispersion equations in graphene plasmonics, we believe that the additional member introduced here offers a complementing view, particularly in the emerging subfield of plasmonic interaction in nonplanar two-dimensional structures. Finally, as useful applications of the theory developed herein, we considered two distinct types of spherical substrates for the coating: specifically, dielectric and Drude substrates. In the former case, this allowed us to explore the localized plasmons arising strictly from the charge carriers in the graphene coating, while, in the latter case, we explored the interplay between plasmons supported independently by the bulk and the coating.

ACKNOWLEDGMENTS

T.C. expresses his gratitude to Weihua Wang for surfacing the idea of considering a spherical structure, and to Wei Yan for encouraging and stimulating discussions. The Center for Nanostructured Graphene is sponsored by the Danish National Research Foundation, Project No. DNRF58. This work was also supported by the Danish Council for Independent Research, Project No. 1323-00087.

APPENDIX: INCLUDING HYDRODYNAMIC NONLOCALITY IN GRAPHENE'S RESPONSE IN SPHERICAL GEOMETRIES

We here discuss how to appropriately account for nonlocal response acting in the graphene coating through a hydrodynamic model. For simplicity—and because it is justified in the size regime relevant for nonlocal response in finite structures—we work in the quasistatic regime.

Hydrodynamic response is included by a modest generalization of the usual Ohm's law $\mathbf{K}(\mathbf{r}) = \sigma(\omega)\mathbf{E}_\parallel(\mathbf{r})$ by appending to the left-hand side a term $\frac{\beta^2}{\omega^2}\nabla_\parallel[\nabla_\parallel \cdot \mathbf{K}(\mathbf{r})]$, which in turn, combined with the continuity equation, enforces a relationship between the induced charge density ρ and the potential ϕ [40,52]:

$$\rho(\mathbf{r}) + \frac{\beta^2}{\omega^2}\nabla_\parallel^2\rho(\mathbf{r}) = \frac{i\sigma(\omega)}{\omega}\nabla_\parallel^2\phi(\mathbf{r}), \quad (\text{A1})$$

with plasma velocity β proportional to the Fermi velocity v_F , interrelated approximately [52] by $\beta^2 = \frac{3}{4}v_F^2$ in graphene. For the potential, we expand it in the exterior and interior regions, \mathcal{V}_2 and \mathcal{V}_1 , by making use of axial symmetry to freely choose the azimuthal $m = 0$ component of a complete set of bounded, nonconstant solutions of the Laplace

equation [53]:

$$\phi_{\nu_1}(\mathbf{r}) = \sum_{l=1} c_l^{\text{tra}} r^l P_l(\cos\theta), \quad r < R, \quad (\text{A2a})$$

$$\phi_{\nu_2}(\mathbf{r}) = \sum_{l=1} [c_l^{\text{inc}} r^l + c_l^{\text{sca}} r^{-(l+1)}] P_l(\cos\theta), \quad r > R, \quad (\text{A2b})$$

with associated incident, scattered, and transmitted multipole coefficients c_l^{inc} , c_l^{sca} , and c_l^{tra} , respectively. Due to spherical symmetry, the coefficients can be matched multipole by multipole, i.e., separately for each l . The matching is governed by the BCs $\hat{\mathbf{n}} \times (\mathbf{E}_{\nu_2} - \mathbf{E}_{\nu_1}) = 0$ and $\hat{\mathbf{n}} \times (\mathbf{D}_{\nu_2} - \mathbf{D}_{\nu_1}) = \rho$, which translate into BCs for the potential ϕ , reading as $\partial_\theta\phi_{\nu_2} = \partial_\theta\phi_{\nu_1}$ and $\varepsilon_1\partial_r\phi_{\nu_1} - \varepsilon_2\partial_r\phi_{\nu_2} = \rho$ at all surficial points. The induced charge density associated with a potential ϕ_{ν_1} of multipole order l is denoted ρ_l and is obtained by solving Eq. (A1) subject to Eq. (A2a) for fixed l yielding

$$\rho_l = -c_l^{\text{tra}} \frac{i\sigma_l^{\text{H}}(\omega)}{\omega} R^{l-2} l(l+1) P_l(\cos\theta), \quad (\text{A3a})$$

expressed via a hydrodynamically corrected conductivity:

$$\sigma_l^{\text{H}}(\omega) \equiv \frac{\sigma(\omega)}{1 - \frac{\beta^2}{\omega^2} \frac{l(l+1)}{R^2}}. \quad (\text{A3b})$$

Applying the BCs to Eqs. (A2) and (A3a) then yields a direct relation between the scattered and incident multipole coefficients:

$$c_l^{\text{sca}} = -(4\pi)^{-1} \alpha_l^{\text{H}} c_l^{\text{inc}}, \quad (\text{A4})$$

expressed in terms of a hydrodynamic multipole polarizability α_l^{H} . Importantly, as is evident from Eq. (A3a), the inclusion of hydrodynamic response acts only to introduce an effective conductivity σ_l^{H} . As such, the hydrodynamic multipole polarizability α_l^{H} differs only from its LRA counterpart α_l of Eq. (9) by the substitution $\sigma \rightarrow \sigma_l^{\text{H}}$.

Interestingly, in momentum space the hydrodynamic conductivity of a planar sheet takes the form $\sigma(q, \omega) = \sigma(\omega)[1 - \frac{\beta^2}{\omega^2} q^2]^{-1}$. Clearly, a mapping between the planar case and Eq. (A3b) can be achieved by introducing an effective momentum $q_l^{\text{eff,H}} \equiv \sqrt{l(l+1)}/R$. Notably, this differs from the optically relevant effective momentum q_l^{eff} at order $\mathcal{O}(l^{-1})$.

Concluding our considerations of hydrodynamics, we comment that the effective nonlocal interaction range β/ω is ~ 1 nm for graphene (for a resonance, e.g., at $\hbar\omega = 0.5$ eV), which is relatively large compared to noble metals [26], e.g., $\beta/\omega \sim 3$ Å in Ag at relevant plasmonic frequencies. As such, the onset of significant hydrodynamic perturbations to the LRA predictions in graphene can be expected to occur comparatively earlier when decreasing the characteristic feature length. In particular, we should expect notable nonlocal corrections throughout the few-nanometer domain.

[1] A. N. Grigorenko, M. Polini, and K. S. Novoselov, *Nat. Photonics* **6**, 749 (2012).

[2] Y. V. Bludov, A. Ferreira, N. M. R. Peres, and M. I. Vasilevskiy, *Int. J. Mod. Phys. B* **27**, 1341001 (2013).

[3] F. J. García de Abajo, *ACS Photonics* **1**, 135 (2014).

[4] W. Wang, P. Apell, and J. Kinaret, *Phys. Rev. B* **84**, 085423 (2011).

- [5] J. Chen, M. Badioli, P. Alonso-González, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenović, A. Centeno, A. Pesquera, P. Godignon, A. Z. Elorza, N. Camara, F. J. García de Abajo, R. Hillenbrand, and F. H. L. Koppens, *Nature (London)* **487**, 77 (2012).
- [6] Z. Fei, A. S. Rodin, G. O. Andreev, W. Bao, A. S. McLeod, M. Wagner, L. M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, M. M. Fogler, A. H. Castro Neto, C. N. Lau, F. Keilmann, and D. N. Basov, *Nature (London)* **487**, 82 (2012).
- [7] S. Thongrattanasiri and F. J. García de Abajo, *Phys. Rev. Lett.* **110**, 187401 (2013).
- [8] S. Thongrattanasiri, F. H. L. Koppens, and F. J. García de Abajo, *Phys. Rev. Lett.* **108**, 047401 (2012).
- [9] Z. Fang, S. Thongrattanasiri, A. Schlather, Z. Liu, L. Ma, Y. Wang, P. M. Ajayan, P. Nordlander, N. J. Halas, and F. J. García de Abajo, *ACS Nano* **7**, 2388 (2013).
- [10] X. Zhu, W. Wang, W. Yan, M. B. Larsen, P. Bøggild, T. G. Pedersen, S. Xiao, J. Zi, and N. A. Mortensen, *Nano Lett.* **14**, 2907 (2014).
- [11] W. B. Lu, W. Zhu, H. J. Xu, Z. H. Ni, Z. G. Dong, and T. J. Cui, *Opt. Express* **21**, 10475 (2013).
- [12] P.-Y. Chen, J. Soric, and A. Alù, *Adv. Mater.* **24**, OP281 (2012).
- [13] P. Chen, J. Soric, Y. R. Padooru, H. M. Bernety, A. B. Yakovlev, and A. Alù, *New J. Phys.* **15**, 123029 (2013).
- [14] B. Zhu, G. Ren, Y. Gao, Y. Yang, Y. Lian, and S. Jian, *Opt. Express* **22**, 24096 (2014).
- [15] Y. Gao, G. Ren, B. Zhu, H. Liu, Y. Lian, and S. Jian, *Opt. Express* **22**, 24322 (2014).
- [16] Y. Gao, G. Ren, B. Zhu, J. Wang, and S. Jian, *Opt. Lett.* **39**, 5909 (2014).
- [17] Z.-R. Huang, L.-L. Wang, B. Sun, M.-D. He, J.-Q. Liu, H.-J. Li, and X. Zhai, *J. Opt.* **16**, 105004 (2014).
- [18] H. Yang, Z. Hou, N. Zhou, B. He, J. Cao, and Y. Kuang, *Ceramics International* **40**, 13903 (2014).
- [19] J.-S. Lee, S.-I. Kim, J.-C. Yoon, and J.-H. Jang, *ACS Nano* **7**, 6047 (2013).
- [20] D. Bakowies, M. Buehl, and W. Thiel, *J. Am. Chem. Soc.* **117**, 10113 (1995).
- [21] P. Calaminici, G. Geudtner, and A. M. Köster, *J. Chem. Theory Comput.* **5**, 29 (2009).
- [22] S. Thongrattanasiri, A. Manjavacas, and F. J. García de Abajo, *ACS Nano* **6**, 1766 (2012).
- [23] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941).
- [24] R. Ruppin, *Phys. Rev. Lett.* **31**, 1434 (1973).
- [25] C. David and F. J. García de Abajo, *J. Phys. Chem. C* **115**, 19470 (2012).
- [26] T. Christensen, W. Yan, S. Raza, A.-P. Jauho, N. A. Mortensen, and M. Wubs, *ACS Nano* **8**, 1745 (2014).
- [27] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983).
- [28] F. H. L. Koppens, D. E. Chang, and F. J. García de Abajo, *Nano Lett.* **11**, 3370 (2011).
- [29] R. Fuchs and F. Claro, *Phys. Rev. B* **35**, 3722 (1987).
- [30] V. Myroshnychenko, J. Rodríguez-Fernández, I. Pastoriza-Santos, A. M. Funston, C. Novo, P. Mulvaney, L. M. Liz-Marzán, and F. J. García de Abajo, *Chem. Soc. Rev.* **37**, 1792 (2008).
- [31] T. Low and P. Avouris, *ACS Nano* **8**, 1086 (2014).
- [32] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, *New J. Phys.* **8**, 318 (2006).
- [33] E. H. Hwang and S. Das Sarma, *Phys. Rev. B* **75**, 205418 (2007).
- [34] L. A. Falkovsky and A. A. Varlamov, *Eur. Phys. J. B* **56**, 281 (2007).
- [35] Here we assume that the conductive response of the spherical geometry is well approximated by its planar equivalent. However, as is well known from carbon nanotubes [36], zone folding quantizes the allowable electronic momenta and so modifies the linear energy dispersion of planar graphene. Nevertheless, these perturbations incur only negligible changes to the conductive response provided the inverse circumference remains small relative to the Fermi momentum k_F , i.e., provided $k_F R \gg 1$. Lastly, the influence of curvature and strain on the effective hopping amplitude [37] is similarly neglected in this approximation.
- [36] J.-C. Charlier, X. Blase, and S. Roche, *Rev. Mod. Phys.* **79**, 677 (2007).
- [37] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009).
- [38] V. P. Gusynin, S. G. Sharapov, and J. P. Carbotte, *J. Phys.: Condens. Matter* **19**, 026222 (2007).
- [39] G. W. Hanson, *J. Appl. Phys.* **104**, 084314 (2008).
- [40] W. Wang and J. M. Kinaret, *Phys. Rev. B* **87**, 195424 (2013).
- [41] N. A. Mortensen, S. Raza, M. Wubs, T. Søndergaard, and S. I. Bozhevolnyi, *Nat. Commun.* **5**, 3809 (2014).
- [42] Z. Ruan and S. Fan, *Appl. Phys. Lett.* **98**, 043101 (2014).
- [43] K. Joulain, R. Carminati, J.-P. Mulet, and J.-J. Greffet, *Phys. Rev. B* **68**, 245405 (2003).
- [44] J. Vielma and P. Leung, *J. Chem. Phys.* **126**, 194704 (2007).
- [45] A. M. Schimpf, N. Thakkar, C. E. Gunthardt, D. J. Masiello, and D. R. Gamelin, *ACS Nano* **8**, 1065 (2014).
- [46] H. Zhang, V. Kulkarni, E. Prodan, P. Nordlander, and A.O. Govorov, *J. Phys. Chem. C* **118**, 16035 (2014).
- [47] For completeness, we note that for a general multipole order l the results of Eq. (16) generalize according to the substitutions $\omega_{gR}^2 \rightarrow \frac{l+1}{2} \omega_{gR}^2$ and $\varepsilon_2 \rightarrow \frac{l+1}{l} \varepsilon_2$.
- [48] C. Gong, G. Lee, B. Shan, E. M. Vogel, R. M. Wallace, and K. Cho, *J. Appl. Phys.* **108**, 123711 (2010).
- [49] L. Kong, C. Bjelkevig, S. Gaddam, Y. H. Lee, G. H. Han, H. K. Jeong, N. Wu, Z. Zhang, J. Xiao, P. A. Dowben, and J. A. Kelber, *J. Phys. Chem. C* **114**, 21618 (2010).
- [50] U. Kreibig and L. Genzel, *Surf. Sci.* **156**, 678 (1985).
- [51] P. Apell and D. R. Penn, *Phys. Rev. Lett.* **50**, 1316 (1983).
- [52] T. Christensen, W. Wang, A.-P. Jauho, M. Wubs, and N. A. Mortensen, *Phys. Rev. B* **90**, 241414(R) (2014).
- [53] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

PHYS. REV. B 90, 241414(R) (2014)

The Dirac approach was initially suggested by W. Wang. T. Christensen subsequently derived and performed all classical, hydrodynamic, and Dirac calculations, including the LRA edge-state conductivity. W. Wang implemented and performed the frequency-domain TB-RPA calculations. T. Christensen wrote the manuscript with critical input and revisions from all authors.

PHYSICAL REVIEW B **90**, 241414(R) (2014)**Classical and quantum plasmonics in graphene nanodisks: Role of edge states**Thomas Christensen,^{1,2} Weihua Wang,^{1,2} Antti-Pekka Jauho,^{2,3} Martijn Wubs,^{1,2} and N. Asger Mortensen^{1,2,*}¹*Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*²*Center for Nanostructured Graphene, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*³*Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

(Received 15 July 2014; revised manuscript received 4 November 2014; published 16 December 2014)

Edge states are ubiquitous for many condensed matter systems with multicomponent wave functions. For example, edge states play a crucial role in transport in zigzag graphene nanoribbons. Here, we report microscopic calculations of quantum plasmonics in doped graphene nanodisks with zigzag edges. We express the nanodisk conductivity $\sigma(\omega)$ as a sum of the conventional bulk conductivity $\sigma_B(\omega)$, and a novel term $\sigma_E(\omega)$, corresponding to a coupling between the edge and bulk states. We show that the edge states give rise to a redshift and broadening of the plasmon resonance, and that they often significantly impact the absorption efficiency. We further develop simplified models, incorporating nonlocal response within a hydrodynamical approach, which allow a semiquantitative description of plasmonics in the ultrasmall size regime. Furthermore, we show that the effect of hydrodynamic and edge-conductivity corrections scale identically, approximately with the inverse of the disk radius, highlighting their equatable importance. However, the polarization dependence is only given by fully microscopic models. The approach developed here should have many applications in other systems supporting edge states.

DOI: [10.1103/PhysRevB.90.241414](https://doi.org/10.1103/PhysRevB.90.241414)

PACS number(s): 73.22.Pr, 73.20.Mf, 78.20.Ci, 78.67.Wj

Introduction. Plasmonics at the nanoscale introduces a host of novel phenomena, both in terms of improved efficacy of certain classical phenomena, e.g., extreme field enhancements, but also conceptually by offering a tunable transition from the classical to the quantum regime [1]. Probing and understanding this transition in detail, and in particular the breakdown of classical predictions, is an important task in view of the progress in nanofabrication [2,3]. With the emergence of low-dimensional materials such as graphene, new avenues develop, both experimentally and theoretically. Graphene, and several other low-dimensional systems, exhibits an approximately linear, gapless, two-band energy dispersion $\epsilon = \pm\hbar v_F k$, with Fermi velocity v_F . The plasmonic consequences of this non-standard dispersion and dimensionality have been investigated vigorously in recent years [4–10].

The accurate description of low-energy excitations in graphene by simple tight-binding (TB) Hamiltonians allows investigations of nonclassical plasmonic features of relatively large graphene structures [11,12]. It has recently been theoretically demonstrated that the optical excitations of few-atom graphene nanostructures involve multiple individual electron-hole pairs (EHPs) strongly modified by the Coulomb interaction, occasionally referred to as molecular plasmons [13]. Conversely, experimental measurements on ensembles of larger disks, of radii $R \gtrsim 50$ nm, have exhibited distinctly classical features [14,15]. Improving our understanding of this transition between opposing realms is underscored by the many advances in fabrication of graphene nanostructures [16–18].

In this Rapid Communication we show that for smaller graphene disks, though larger than $R \gtrsim 7$ nm, two essential modifications of the classical single-disk response arise, due to edge states and to nonlocal response, producing an overall

redshift and broadening of the dipole resonance. In particular, we show that the existence of edge states due to zigzag (ZZ) features can be accounted for via an edge-state conductivity, while the impact of nonlocal response can be accounted for effectively within a hydrodynamic model. This affirms and extends the supposition regarding the crucial role of edge states in prior numerical work [11]. In Fig. 1 we outline and summarize the different computational approaches considered in this Rapid Communication.

Electronic states. The simplest atomistic description of the conduction electrons of graphene, without explicit treatment of spin, is given by the p_z -orbital nearest-neighbor TB Hamiltonian with hopping energy $t_{AB} = 2.8$ eV:

$$\hat{H}_{\text{TB}} = -t_{AB} \sum_{\langle j,j' \rangle} \hat{a}_j^\dagger \hat{b}_{j'} + \hat{b}_j^\dagger \hat{a}_{j'}, \quad (1)$$

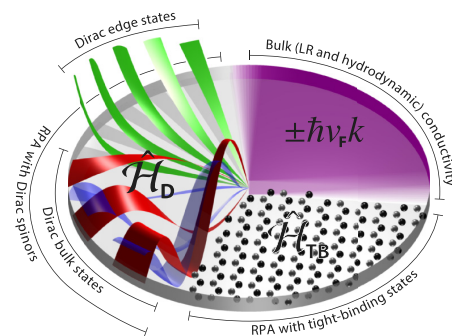


FIG. 1. (Color online) Illustration of considered levels of approximation for a graphene nanodisk. Angular slices of Dirac ZZ bulk state spinor components are indicated in red and blue, and edge state nonzero components in green.

*asger@mailaps.com

CHRISTENSEN, WANG, JAUHO, WUBS, AND MORTENSEN

PHYSICAL REVIEW B **90**, 241414(R) (2014)

with A - and B -sublattice annihilation (creation) operators $\hat{a}_j^{(\dagger)}$ and $\hat{b}_j^{(\dagger)}$ at sites j , and with $\langle j, j' \rangle$ indicating summation over nearest neighbors. In the low-energy limit, for extended graphene, the characteristics of the TB approach are asymptotically reproduced by the four-spinor Dirac equation, $\hat{H}_D \psi(\mathbf{r}) = \epsilon \psi(\mathbf{r})$, with the Hamiltonian [19]

$$\hat{H}_D = v_F(\tau_0 \otimes \sigma_x \hat{p}_x + \tau_z \otimes \sigma_y \hat{p}_y), \quad (2)$$

where $\hat{p} = -i\hbar\nabla$ denotes momentum, and with Pauli matrices τ_i and σ_i belonging to valley and sublattice subspaces, respectively. In the absence of valley mixing, the four-spinor equation for $\psi(\mathbf{r}) = [\psi_A^+(\mathbf{r}), \psi_B^+(\mathbf{r}), \psi_A^-(\mathbf{r}), \psi_B^-(\mathbf{r})]^T$ decouples into a pair of two-spinor equations for $\psi^\kappa(\mathbf{r}) = [\psi_A^\kappa(\mathbf{r}), \psi_B^\kappa(\mathbf{r})]^T$ associated with valley indices $\kappa = \pm 1$ pertaining to Dirac valleys $\mathbf{K}^\kappa = [\sqrt{3}, \kappa]^T 2\pi/3a_{LC}$, with lattice constant $a_{LC} = 2.46 \text{ \AA}$ [20].

Finite graphene structures are easily modeled with Eq. (1) by omitting the absent neighbors in the matrix representation of \hat{H}_{TB} , whose dimension equals the number of constituent carbon atoms. For the continuum Dirac equation, Eq. (2), suitable boundary conditions (BCs) are needed. General considerations, enforcing no-spill current conditions, Hermiticity, and unitarity, lead to a rather broad family of allowable BCs [22,23], which can be made explicit by using the atomistic details of the structural termination. In the present work we consider ZZ lattice termination (which can be considered appropriate, in general, for nonarmchair minimal lattice terminations as argued in Ref. [23]) forcing a single sublattice component to vanish, e.g., forcing $\psi_A^\kappa(\mathbf{r}) = 0$ on the boundary if the ZZ edge belongs to the B -sublattice. For comparison we also consider the infinite mass (IM) BC [24], corresponding microscopically to confinement due to an atomically staggered potential [23], which enforces an intersublattice phase relationship $\psi_B^\kappa(\mathbf{r})/\psi_A^\kappa(\mathbf{r}) = i e^{i\kappa\theta}$, with θ denoting the tangential boundary angle [25].

Upon application of BCs, the otherwise linear Dirac dispersion $\epsilon = \pm\hbar v_F k$ is transformed into a discrete set of energies and associated spinors. For the case of a homogeneous disk of radius R , the nonzero-energy spinors are quantized in angular and radial quantum numbers $l = 0, \pm 1, \pm 2, \dots$ and $n = 1, 2, \dots$ [26,27]:

$$\psi_{ln}^\kappa(\tilde{r}, \theta) = \frac{e^{il\theta}}{\sqrt{N_{ln}^\kappa}} \begin{bmatrix} J_l(\beta_{ln}\tilde{r}) \\ i\kappa J_{l+\kappa}(\beta_{ln}\tilde{r}) e^{i\kappa\theta} \end{bmatrix}, \quad (3)$$

with normalization N_{ln}^κ [see Supplemental Material (SM) [28]] expressed through the dimensionless radial coordinate $\tilde{r} = r/R$ and momenta $\beta_{ln} = \epsilon_{ln}/\hbar\omega_R$ with circumferential fermion frequency $\omega_R = v_F/R$. The ZZ BC energies are valley independent and correspond to zeros of the Bessel function, i.e., β_{ln} fulfills $J_l(\beta_{ln}) = 0$, while the IM BC yields valley-dependent energies, given by $\kappa J_{l+\kappa}(\beta_{ln}^\kappa) = J_l(\beta_{ln}^\kappa)$. Additionally, for the ZZ BC a set of zero-energy spinors exist, here denoted by ϕ_ℓ^κ , discretized in angular quantum numbers $\ell = 0, 1, \dots, \ell_{\max}$ [26,27]:

$$\phi_\ell^\kappa(\tilde{r}, \theta) = \frac{e^{-i\kappa\ell\theta}}{\sqrt{N_\ell}} \begin{bmatrix} 0 \\ \tilde{r}^\ell \end{bmatrix}, \quad (4)$$

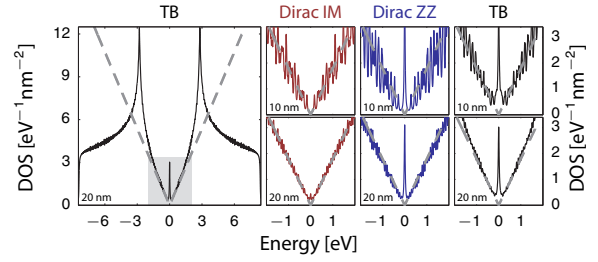


FIG. 2. (Color online) DOS for graphene nanodisk in Dirac ZZ and IM, and TB treatments, broadened by a loss of $\hbar\eta = 24 \text{ meV}$, with disk diameter indicated. The asymptotic low-energy, bulk graphene DOS is indicated by dashed gray. The region of approximately linear DOS is indicated by gray shading.

with normalization N_ℓ (see SM). The phenomenologically introduced cutoff angular quantum number ℓ_{\max} is required to avoid a divergence of the density of states at zero energy, and is chosen to ensure a total number of zero-energy states (including spin and valley degeneracy) $\sim 2\pi R/3a_{LC}$ [23] (see SM). Except for the $\ell = 0$ case, the zero-energy states are predominately localized at the disk edge, with a characteristic localization length $R[1 - e^{-1/2(\ell+1)}]$ (see SM).

In Fig. 2 we show the resulting noninteracting density of states (DOS), phenomenologically broadened by an electron collision rate η , computed as $\text{DOS}(\epsilon) = \frac{2}{\pi\mathcal{A}} \sum_\nu \text{Im}[(\epsilon_\nu - \epsilon - i\hbar\eta)^{-1}]$ with \mathcal{A} denoting the sample area and with \sum_ν denoting summation over all states ν (excluding spin, which contributes a factor 2). Also, the DOS for a TB model is shown, for a bond-centered disk. A key feature of both TB and Dirac ZZ treatments is a prominent peak at zero energy associated with edge states, which is not reproduced in either Dirac IM or in bulk approximations. Additionally, due to breaking of valley and azimuthal symmetry in TB the interstate energy-level spacing is overestimated in Dirac treatments relative to TB. Nevertheless, the total number of edge and bulk states in Dirac ZZ and TB is in good agreement (see SM). Due to the absence of edge states in Dirac IM vis-à-vis its presence in TB, we focus in the following on Dirac ZZ. Finally, we note the complete absence of nonconical dispersion effects, e.g., trigonal warping [19] and van Hove quasisingularities at $\pm t_{AB}$ [29], in the Dirac treatment, whose exemption, however, is expected to be unimportant in low-energy plasmonics.

Random-phase approximation. To compute the optical response of graphene disks in both TB and Dirac approaches, the first step is to evaluate the noninteracting polarizability [11,30]:

$$\chi^0(\mathbf{r}, \mathbf{r}'; \omega) = 2 \sum_{\nu\nu'} (f_\nu - f_{\nu'}) \frac{\psi_\nu^\dagger(\mathbf{r}) \psi_\nu(\mathbf{r}) \psi_{\nu'}^\dagger(\mathbf{r}') \psi_{\nu'}(\mathbf{r}')}{\epsilon_\nu - \epsilon_{\nu'} - \hbar(\omega + i\eta)}, \quad (5)$$

where f_ν denotes Fermi-Dirac equilibrium functions evaluated at energy ϵ_ν , and electron relaxation is included phenomenologically through a finite rate η . We give explicit expressions for the Dirac-disk polarizability in the SM.

The random-phase approximation (RPA) is instated by coupling the induced charge density $\rho(\mathbf{r})$ to the *total* field via

χ^0 , leading to a self-consistent integral equation, reading, in operator notation, as $\rho = e^2 \chi^0 (\phi^{\text{ext}} + \mathbf{V}\rho)$, with \mathbf{V} denoting the Coulomb interaction and ϕ^{ext} an external potential [31]. Henceforth, depending on the choice for single-particle states used in constructing χ^0 , we distinguish between approaches by the self-evident notation RPA@Dirac and RPA@TB. In the SM we elucidate the technical details for efficiently computing RPA@Dirac via an angular momentum decomposition, and follow the scheme introduced in Ref. [11] for RPA@TB. The computational complexity of these approaches is discussed and compared in the SM. The optical absorption cross section, i.e., the absorbed power relative to the intensity of an incident plane wave, relates to the induced charge density via $\sim \omega \text{Im}[p(\omega)]$, with $p(\omega)$ denoting the dipole moment obtained from $\rho(\mathbf{r})$.

Local response. For comparison with the two quantum approaches described above, we also consider the traditional, classical approach, wherein the induced charge density in graphene is determined from the well-known bulk local-response (LR) conductivity with intra- and interband terms $\sigma_{\text{B}}(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega)$ [32]. The interband term induces a redshift [33] of the dipolar plasmon resonance with decreasing radius, but not to the extent observed in TB-RPA calculations [11]. For the electrostatic disk, the LR problem is solved most elegantly by using a polynomial expansion technique, as explicated by Fetter [34], and summarized for completeness in the SM, allowing a semianalytical solution requiring only a numerical matrix inversion. Applying this technique, we find that the singly radially quantized dipole plasmon resonance, ω_{dp} , being the resonance of primary relevance in nanoscopic disks, relates to the total LR conductivity $\sigma(\omega)$ via $\omega_{\text{dp}}/\sigma(\omega_{\text{dp}}) = \zeta/2i\epsilon_0\epsilon_{\text{B}}R$, with ϵ_{B} denoting the background dielectric constant and $\zeta \approx 1.0977$ accounting for the disk geometry [35]. An intraband approximation then entails the scaling $\omega_{\text{dp}} \propto 1/R^{1/2}$.

Although the bulk LR conductivity $\sigma_{\text{B}}(\omega)$ is usually derived from a starting point of a continuum of bulk graphene Dirac states, it may as well be derived from the large-radius limit of the finite sample's conductivity using the states $\psi_{l_n}^{\kappa}$ from Eq. (3). Specifically, in the LR limit, the current response due to an x -polarized field is obtained from the conductivity [31]:

$$\sigma(\omega) = \frac{2ie^2\omega}{\mathcal{A}} \sum_{vv'} (f_v - f_{v'}) \frac{|\langle \psi_v^{\kappa} | x | \psi_{v'}^{\kappa} \rangle|^2}{\epsilon_v - \epsilon_{v'} - \hbar(\omega + i\eta)}. \quad (6)$$

Considering the Dirac ZZ states in Eqs. (3) and (4) this gives rise to two distinct terms, one tending asymptotically to $\sigma_{\text{B}}(\omega)$ with increasing radius, originating from bulk-to-bulk transitions $|\langle \psi_{l_n}^{\kappa} | x | \psi_{l_n'}^{\kappa} \rangle|^2$, and one novel term originating from edge-to-bulk transitions:

$$\sigma_{\text{E}}(\omega) = \frac{4ie^2\omega}{\mathcal{A}} \sum_{\kappa l n l'} (f_{l_n} - f_0) \frac{|\langle \psi_{l_n}^{\kappa} | x | \phi_{l'}^{\kappa} \rangle|^2}{\epsilon_{l_n}^2 - \hbar^2(\omega + i\eta)^2}, \quad (7)$$

with f_0 denoting the Fermi-Dirac function at zero energy, and ϵ_{l_n} denoting the Dirac ZZ energies corresponding to $\psi_{l_n}^{\kappa}$. This edge contribution, physically representing all interactions between occupied zero-energy edge states and unoccupied nonzero-energy bulk states, can be worked out explicitly as

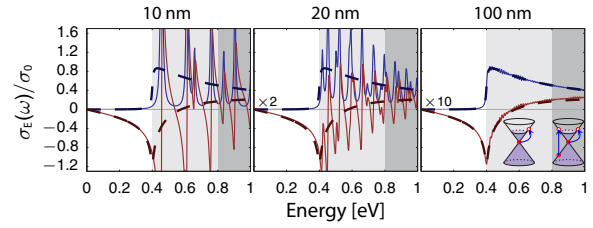


FIG. 3. (Color online) Edge-state conductivity in graphene nanodisks (doped to 0.4 eV) normalized to $\sigma_0 = e^2/4\hbar$ calculated with finite damping $\hbar\eta = 6$ meV. Disk diameter indicated in titles; note the scaling factor in the center and right-hand graphs. Explicit evaluation of Eq. (8) given in full, and large-radius limit, Eq. (9), in bold dashed; real and imaginary parts in blue and red, respectively. The region of edge-to-bulk transitions is indicated in gray, while the region of concurrent edge-to-bulk and interband transitions is dark-gray, and illustrated schematically.

a summation over the Bessel function zeros β_{l_n} (see SM for details). In the low-temperature limit, assuming positive ϵ_{F} , the edge-state conductivity becomes

$$\sigma_{\text{E}}(\omega) = \frac{-16ie^2}{\pi\hbar} \frac{\omega}{\omega_R} \sum_{\ell=0}^{\ell_{\text{max}}} \sum_n^{\hbar\omega_R\beta_{l_n} > \epsilon_{\text{F}}} \frac{\ell+1}{\beta_{l_n}^5 \left[1 - \left(\frac{\omega+i\eta}{\beta_{l_n}\omega_R}\right)^2\right]}. \quad (8)$$

Remarkably, the above expression allows a simple asymptotic form in the large-radius limit $R \rightarrow \infty$. Replacing the angular momenta $\ell+1$ by their average at fixed energy $\ell+1 \rightarrow \langle \ell+1 \rangle_{\epsilon} \simeq \xi\epsilon/\hbar\omega_R$, with proportionality constant $\xi = 4/3\pi$ (see SM), introducing the bulk-energy substitution $\beta_{l_n} \rightarrow \epsilon/\hbar\omega_R$, and transforming the l_n summations into integrals over $\epsilon_{\text{F}} \leq \epsilon < \infty$, we find

$$\sigma_{\text{E}}^{\infty}(\omega) = \xi \frac{2e^2}{\pi\hbar} \frac{v_{\text{F}}}{\omega R} \left[i \ln \left| \frac{\epsilon_{\text{F}}^2 - \hbar^2\omega^2}{\epsilon_{\text{F}}^2} \right| + \pi\theta(\hbar\omega - \epsilon_{\text{F}}) \right], \quad (9)$$

shown here, for simplicity, in the low-loss limit $\eta \rightarrow 0^+$ [36]. Interestingly, Eq. (9) shows that the inclusion of edge states opens a dispersive channel scaling with $\omega_R = v_{\text{F}}/R$, mathematically reminiscent of, but physically distinct from, the scaling phenomenologically introduced in Kreibitz damping [37] and recently derived from the viewpoint of nonlocal diffusion dynamics [38]. In addition to Landau damping due to vertical transitions, as included in $\sigma_{\text{inter}}(\omega)$ for $\hbar\omega \geq 2\epsilon_{\text{F}}$, edge-to-bulk transitions allow nonvertical transitions at $\hbar\omega \geq \epsilon_{\text{F}}$, with the necessary momentum supplied by the structural truncation with a strength proportional to $1/R$. In Fig. 3 we consider $\sigma_{\text{E}}(\omega)$ and compare with $\sigma_{\text{E}}^{\infty}(\omega)$ for three disk diameters. At smaller diameters $\sigma_{\text{E}}(\omega)$ and $\sigma_{\text{E}}^{\infty}(\omega)$ differ substantially in the region $\hbar\omega > \epsilon_{\text{F}}$ with σ_{E} exhibiting peaks at discrete transitional energies $\hbar\omega \simeq \epsilon_{l_n}$; as the diameter is increased, and the energy difference between distinct transitional energies decreases accordingly, σ_{E} approaches $\sigma_{\text{E}}^{\infty}$ asymptotically, as anticipated. We note that a generally good agreement is apparent, even for small disks, when $\hbar\omega < \epsilon_{\text{F}}$. The importance of the edge-state conductivity vis-à-vis the bulk conductivity diminishes with increasing diameter due to the $1/R$ scaling of $\sigma_{\text{E}}(\omega)$. Nonetheless, even at large disk diameters, e.g., 20 nm, the maximal edge-state conductivity is

CHRISTENSEN, WANG, JAUHO, WUBS, AND MORTENSEN

PHYSICAL REVIEW B **90**, 241414(R) (2014)

still on the order of $\sim 0.4\sigma_0$, while the magnitude of vertical interband transitions roughly amounts to σ_0 .

Hydrodynamic response. The noninteracting polarizability, the key constituent of the RPA, as considered in Eq. (5), accounts not only for the discretized and individual nature of the allowable states, but also for the nonlocal nature of the electromagnetic response, manifest in the finitude of $\chi^0(\mathbf{r}, \mathbf{r}'; \omega)$ for $\mathbf{r} \neq \mathbf{r}'$. An approximate accounting of nonlocal response can be facilitated by a hydrodynamic model [33]:

$$\left(1 + \frac{\beta^2}{\omega^2} \nabla_{\parallel}^2\right) \rho(\mathbf{r}) = \frac{i\sigma(\omega)}{\omega} \nabla_{\parallel}^2 \phi(\mathbf{r}), \quad (10)$$

with ∇_{\parallel}^2 being the two-dimensional Laplacian, and with plasma velocity denoted by $\beta^2 = \frac{3}{4}v_F^2$ (see SM). For brevity, we will denote hydrodynamic calculations with a backbone conductivity $\sigma(\omega)$ by $[\sigma]^H(\omega)$. The primary effect of the hydrodynamic model is to introduce a blueshift, which, in $[\sigma_B]^H(\omega)$, approximately amounts to a shift $\delta\omega_{dp} \simeq 1.27\omega_R^2/\omega_{dp}$. Predictions of the hydrodynamic model at the level $[\sigma_B]^H(\omega)$ agree excellently with predictions of RPA@Dirac IM as we show in the SM. This underscores the accuracy of a hydrodynamic description, since RPA@Dirac IM neglects the existence of edge states, and thus, at large radii, is modified primarily by nonlocal effects. Moreover, through this, we qualitatively explain the blueshift predicted by RPA@TB in armchair nanostructures [39,40] as a hydrodynamic shift.

Results and discussion. Figure 4 depicts the absorption cross-sectional efficiency, i.e., cross section normalized to disk area, for a normally incident excitation field, i.e., propagating along z , of graphene nanodisks for different diameters, contrasting results obtained by LR with and without hydrodynamic and edge-state conductivity, RPA@TB, and RPA@Dirac ZZ. A fundamental feature of RPA@TB, not captured by any of the continuum models, is a polarization dependence, considered in Fig. 4 for x - and y -polarized incident fields, of the optical response due to the discrete nature of the description. For smaller disks, only few EHPs contribute, leading to a strong polarization dependence. For larger disks, as the number of contributing EHPs increases, and the collective nature of the plasmon emerges, this dependence diminishes rapidly.

The primary feature of both RPA@TB and RPA@Dirac ZZ for disk diameters larger than approximately 14 nm, is the emergence of a broad dominant plasmonic resonance redshifted with respect to the LR bulk predictions. Comparison with $[\sigma_B + \sigma_E]^H(\omega)$ and $[\sigma_B + \sigma_E^{\infty}]^H(\omega)$ agrees qualitatively. A similar redshift is reproduced, $\propto \omega_R^2/\omega_{dp} \propto 1/R^{3/2}$ (see SM), but slightly underestimated in magnitude due to the assumption of a constant total field, inherent to the dipole approximation in Eq. (6), contrasting the actual electric field distribution of the plasmon, which is significantly concentrated near the edge [41]. Furthermore, the dipole resonance in RPA@TB is damped and broadened to a larger degree than both RPA@Dirac ZZ and $[\sigma_B + \sigma_E]^H(\omega)$ as a result of the explicit breaking of azimuthal and valley symmetry in the discrete treatment, permitting additional dipole-allowed transitions.

In conclusion, the redshift observed between predictions of RPA@TB and bulk LR calculations arises from the competing

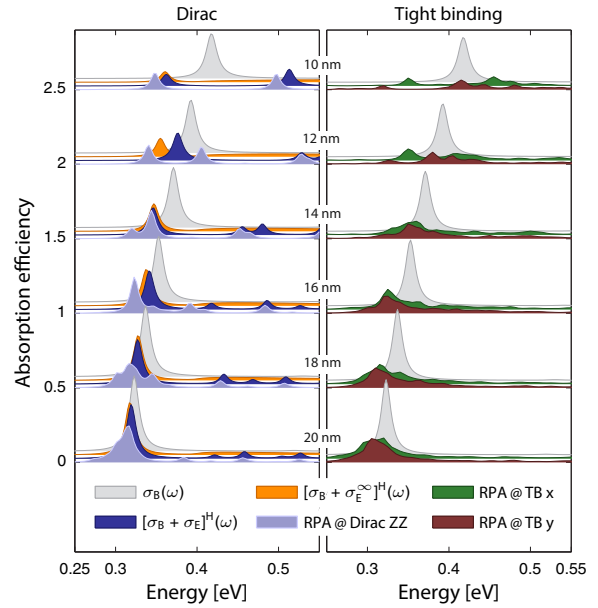


FIG. 4. (Color online) Absorption cross-sectional efficiency at normal incidence of graphene nanodisks calculated via LR, hydrodynamics with bulk and edge-state conductivities, RPA@Dirac ZZ, and RPA@TB (for x - and y -polarized light) for disks of increasing diameter. The sample is considered doped to $\epsilon_F = 0.4$ eV, with electron relaxation-rate $\hbar\eta = 6$ meV, and at a temperature $T = 300$ K. Spectra for different diameters are offset by 0.5, while individual spectra at identical diameters are offset by 0.025. Spectra for intermediate diameters available in SM.

effects of edge conductivity and nonlocal response, with the former prevailing, shifting the dipole resonance ω_{dp} to the red and blue, respectively, with a strength $\propto \omega_R^2/\omega_{dp}$ in both cases. The simultaneous accounting of both effects is thus of paramount importance in semiclassically reproducing the key plasmonic features of full RPA@TB predictions, with significant corrections even at relatively large diameters ~ 20 nm. Our identification and effective description of this additional dispersive channel via edge states illustrates an important difference between graphene and metal plasmonics. Additionally, the equivalent size-dependent scaling with, e.g., nonlocal corrections, accentuates its high-ranking position in the hierarchy of nonclassical corrections for plasmonics at the nanoscale. Encouragingly, the salient features of Eq. (9) are geometry independent, and we accordingly predict that the simple analytical term can be qualitatively extended to other graphene nanostructures, e.g., by substituting $R \rightarrow 2A/C$ with A and C denoting system area and circumference, respectively, reasonable for sufficiently smooth boundaries. Finally, generalizations to a much wider class of systems supporting edge or surface states appear feasible, e.g., in topological insulators such as bismuth bilayers [42] or silicene [43], MoS₂ nanotriangles [44], nanostructures with Ag(111) facets [45], or indeed in *any* finite bipartite system which generally supports zero-energy localized states [46].

Acknowledgments. We thank Wei Yan for many fruitful discussions, helpful suggestions, and continued interest. The Center for Nanostructured Graphene is sponsored by the

Danish National Research Foundation, Project DNRF58. This work was also supported by the Danish Council for Independent Research, Project 1323-00087.

- [1] M. S. Tame, K. R. McEnery, Ş. K. Özdemir, J. Lee, S. A. Maier, and M. S. Kim, *Nat. Phys.* **9**, 329 (2013).
- [2] K. Ostrikov, U. Cvelbar, and B. Murphy, *J. Phys. D: Appl. Phys.* **44**, 174001 (2011).
- [3] J. Henzie, J. Lee, M.H. Lee, W. Hasan, and T. W. Odom, *Annu. Rev. Phys. Chem.* **60**, 147 (2009).
- [4] A. N. Grigorenko, M. Polini, and K. S. Novoselov, *Nat. Photonics* **6**, 749 (2012).
- [5] Y. V. Bludov, A. Ferreira, N. M. R. Peres, and M. I. Vasilevskiy, *Int. J. Mod. Phys. B* **27**, 1341001 (2013).
- [6] F. J. García de Abajo, *ACS Photonics* **1**, 135 (2014).
- [7] M. Jablan, H. Buljan, and M. Soljačić, *Phys. Rev. B* **80**, 245435 (2009).
- [8] F. H. L. Koppens, D. E. Chang, and F. J. García de Abajo, *Nano Lett.* **11**, 3370 (2011).
- [9] V. W. Brar, M. S. Jang, M. Sherrott, J. J. Lopez, and H. A. Atwater, *Nano Lett.* **13**, 2541 (2013).
- [10] H. Yan, T. Low, W. Zhu, Y. Wu, M. Freitag, X. Li, F. Guinea, P. Avouris, and F. Xia, *Nat. Photonics* **7**, 394 (2013).
- [11] S. Thongrattanasiri, A. Manjavacas, and F. J. García de Abajo, *ACS Nano* **6**, 1766 (2012).
- [12] S. Thongrattanasiri, A. Manjavacas, P. Nordlander, and F. J. García de Abajo, *Laser Photon. Rev.* **7**, 297 (2013).
- [13] A. Manjavacas, F. Marchesin, S. Thongrattanasiri, P. Koval, P. Nordlander, D. Sánchez-Portal, and F. J. García de Abajo, *ACS Nano* **7**, 3635 (2013).
- [14] Z. Fang, S. Thongrattanasiri, A. Schlather, Z. Liu, L. Ma, Y. Wang, P. M. Ajayan, P. Nordlander, N. J. Halas, and F. J. García de Abajo, *ACS Nano* **7**, 2388 (2013).
- [15] X. Zhu, W. Wang, W. Yan, M. B. Larsen, P. Bøggild, T. G. Pedersen, S. Xiao, J. Zi, and N. A. Mortensen, *Nano Lett.* **14**, 2907 (2014).
- [16] K. Kim, S. Coh, C. Kisielowski, M. F. Crommie, S. G. Louie, M. L. Cohen, and A. Zettl, *Nat. Commun.* **4**, 2723 (2013).
- [17] S. K. Hämäläinen, Z. Sun, M. P. Boneschanscher, A. Uppstu, M. Ijäs, A. Harju, D. Vanmaekelbergh, and P. Liljeroth, *Phys. Rev. Lett.* **107**, 236803 (2011).
- [18] X. Jia, M. Hofmann, V. Meunier, B. G. Sumpter, J. Campos-Delgado, J. M. Romo-Herrera, H. Son, Y.-P. Hsieh, A. Reina, J. Kong, M. Terrones, and M. S. Dresselhaus, *Science* **323**, 1701 (2009).
- [19] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009).
- [20] In the present study we forfeit consideration of armchair lattice termination, which induces valley admixing [21], and thus examine the two-spinor Dirac equations with $\hat{\mathcal{H}}_D^+ = v_F \sigma \cdot \hat{p}$ and $\hat{\mathcal{H}}_D^- = v_F \sigma^* \cdot \hat{p}$.
- [21] L. Brey and H. A. Fertig, *Phys. Rev. B* **73**, 195408 (2006).
- [22] E. McCann and V. I. Fal'ko, *J. Phys.: Condens. Matter* **16**, 2371 (2004).
- [23] A. R. Akhmerov and C. W. J. Beenakker, *Phys. Rev. B* **77**, 085423 (2008).
- [24] M. V. Berry and R. J. Mondragon, *Proc. R. Soc. London, Ser. A* **412**, 53 (1987).
- [25] N. M. R. Peres, J. N. B. Rodrigues, T. Stauber, and J. M. B. Lopes dos Santos, *J. Phys.: Condens. Matter* **21**, 344202 (2009).
- [26] B. Wunsch, T. Stauber, and F. Guinea, *Phys. Rev. B* **77**, 035316 (2008).
- [27] M. Grujić, M. Zarenia, A. Chaves, M. Tadić, G. A. Farias, and F. M. Peeters, *Phys. Rev. B* **84**, 205441 (2011).
- [28] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.90.241414> for additional details, including derivation of results, supporting simulations, and description of methodology.
- [29] A. Zhou and W. Sheng, *J. Appl. Phys.* **112**, 094313 (2012).
- [30] A. Manjavacas and F. J. García de Abajo, *Nat. Commun.* **5**, 3548 (2014).
- [31] G. Grosso and G. P. Parravicini, *Solid State Physics*, 2nd ed. (Academic, New York, 2014).
- [32] L. A. Falkovsky and A. A. Varlamov, *Eur. Phys. J. B* **56**, 281 (2007).
- [33] W. Wang and J. M. Kinaret, *Phys. Rev. B* **87**, 195424 (2013).
- [34] A. L. Fetter, *Phys. Rev. B* **33**, 5221 (1986).
- [35] Similar results were found in Ref. [14], albeit with the less precise value of $\zeta \approx 0.995$ likely due to use of a general, but less tailored, numerical approach. Precise evaluation of ζ is crucial to accurately assess the actual resonance shifts of RPA@TBA relative to LR bulk calculations.
- [36] Finite loss, $\eta > 0$, is included through the rigorous substitutions $\omega_R/\omega \rightarrow \omega_R\omega/(\omega + i\eta)$, $\pi\theta(\omega - \bar{\epsilon}_F) \rightarrow \arctan[(\bar{\epsilon}_F + \omega)/\eta] - \arctan[(\bar{\epsilon}_F - \omega)/\eta]$, and $\ln|(\bar{\epsilon}_F^2 - \omega^2)/\bar{\epsilon}_F^2| \rightarrow \frac{1}{2}\ln\{[(\bar{\epsilon}_F - \omega)^2 + \eta^2]/\bar{\epsilon}_F^2\} + \frac{1}{2}\ln\{[(\bar{\epsilon}_F + \omega)^2 + \eta^2]/\bar{\epsilon}_F^2\}$, where $\bar{\epsilon}_F = \epsilon_F/\hbar$.
- [37] U. Kreibitz and C. Fragstein, *Z. Phys.* **224**, 307 (1969).
- [38] N. A. Mortensen, S. Raza, M. Wubs, T. Søndergaard, and S. I. Bozhevolnyi, *Nat. Commun.* **5**, 3809 (2014).
- [39] S. Thongrattanasiri and F. J. García de Abajo, *Phys. Rev. Lett.* **110**, 187401 (2013).
- [40] W. Wang, T. Christensen, A.-P. Jauho, K. S. Thygesen, M. Wubs, and N. A. Mortensen, [arXiv:1410.0537](https://arxiv.org/abs/1410.0537).
- [41] W. Wang, P. Apell, and J. Kinaret, *Phys. Rev. B* **84**, 085423 (2011).
- [42] I. K. Drozdov, A. Alexandradinata, S. Jeon, S. Nadj-Perge, H. Ji, R. J. Cava, B. Andrei Bernevig, and A. Yazdani, *Nat. Phys.* **10**, 664 (2014).
- [43] K. Kikutake, M. Ezawa, and N. Nagaosa, *Phys. Rev. B* **88**, 115432 (2013).
- [44] M. V. Bollinger, K. W. Jacobsen, and J. K. Nørskov, *Phys. Rev. B* **67**, 085410 (2003).
- [45] T. C. Hsieh, T. Miller, and T.-C. Chiang, *Phys. Rev. Lett.* **55**, 2483 (1985).
- [46] M. Inui, S. A. Trugman, and E. Abrahams, *Phys. Rev. B* **49**, 3190 (1994).

SUPPLEMENTARY MATERIAL

Classical and Quantum Plasmonics in Graphene Nanodisks: the Role of Edge States

Thomas Christensen,^{1,2} Weihua Wang,^{1,2} Antti-Pekka Jauho,^{2,3} Martijn Wubs,^{1,2} and N. Asger Mortensen^{1,2}

¹Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

²Center for Nanostructured Graphene, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

³Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

CONTENTS

I. Tight-binding calculations	1	A. Formulation of model and comparison with bulk RPA	9
A. Atomic configuration	1	B. Comparison of hydrodynamics and RPA with Dirac IM states	9
B. Numerical values of physical constants	1		
II. Dirac states in disks	2	VII. Electrostatic interaction in local-response nanodisk	10
A. Nonzero-energy states	2	A. Reduction to matrix eigenvalue problem by polynomial expansion	11
B. Zero-energy states	3	1. Eigenmode resonance frequencies in local-response	12
1. Estimate of the number of zero-energy states in a zigzag-like Dirac disk	3	2. Pole approximations for hydrodynamic and edge-state conductivity contributions	12
2. Localization length of zero-energy states in Dirac disk	3	B. Interaction with external potentials and the absorption cross-section	13
3. Identification of edge states in tight-binding calculations	4	VIII. Electrostatic interaction in Dirac-response nanodisk	13
III. Non-interacting density response of Dirac disk	4	A. Computational complexity and comparison with RPA at tight-binding level	14
IV. Edge-state conductivity	5	IX. Absorption spectra at intermediate and smaller diameters	15
A. Large-radius asymptotic form	6	References	16
1. Evaluation of proportionality constant ξ	7		
V. Local-response bulk conductivity	8		
VI. Linearized hydrodynamic model for graphene	8		

I. TIGHT-BINDING CALCULATIONS

A. Atomic configuration

For definiteness, we explicate the atomic configuration considered in our Communication for tight-binding (TB) calculations. In all cases we consider the so-called carbon-carbon (CC) centered configuration [S1], wherein a carbon-carbon bond is centered at origo, see Fig. S1. Note the existence of a mixture of both zigzag- and armchair-like edge regions. In our Dirac approach we approximate this nontrivial boundary configuration by a simple zigzag boundary. Such a treatment is justified in light of the good qualitative agreement in the DOS of TB and Dirac ZZ treatments, see Fig. 2 of our Communication. Additionally, as derived by Akhmerov and Beenakker, zigzag boundary conditions result generically for non-armchair minimal lattice terminations [S2].

B. Numerical values of physical constants

In all calculations we employ a nearest-neighbor hopping energy of $t_{AB} \approx 2.8$ eV. For correspondence with results using a Dirac description we use a Fermi velocity $v_F = \sqrt{3}a_{LC}t_{AB}/2\hbar \approx 0.91 \times 10^6$ m/s, where $a_{LC} = 2.46$ Å denotes the lattice constant. The resulting Fermi velocity slightly overestimates the measured electronic-band velocity, $v_F^{\text{exp}} \approx 1.0 \times 10^6$ m/s, deviating from the simple linear-dispersion prediction due to the effects of many-body interactions [S3]. Nevertheless, for formal consistency within our present

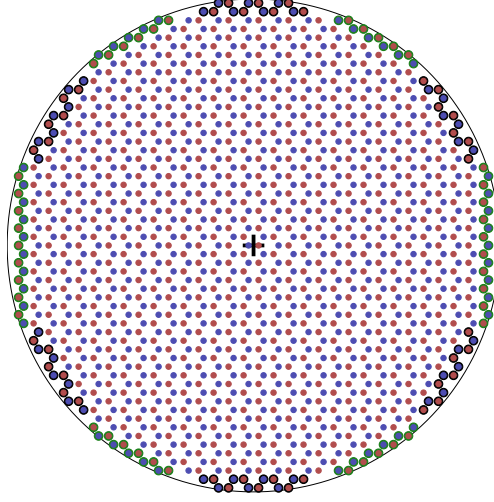


FIG. S1 CC-centered graphene disk ($R = 3.5$ nm, containing 1456 sites) with dangling bonds eliminated. A and B sublattice atoms, indicated in blue and red, respectively, are positioned symmetrically to the left and right of origo. The presence of zigzag- and armchair-like edge regions is indicated by green and black shading, respectively.

framework, and because we choose to focus on the features associated with the Dirac Hamiltonian, we maintain $v_F \approx 0.91 \times 10^6$ m/s throughout.

The phenomenologically introduced loss rate is chosen as $\hbar\gamma = 2\hbar\eta = 12$ meV in all calculations except in depiction of the DOS (broadened therein for clarity and to facilitate direct comparison with bulk DOS). This value is chosen to be in agreement with experimentally attained mobilities in graphene at a Fermi level near 0.4 eV, see data collected in Ref. [S4].

II. DIRAC STATES IN DISKS

The Dirac equation, see Eq. (2), for uncoupled valleys $\kappa = \pm 1$ can be written as a two-spinor equation $\hat{\mathcal{H}}_D^\kappa \psi^\kappa(\mathbf{r}) = \epsilon \psi^\kappa(\mathbf{r})$ which, in a polar coordinate-system (r, θ) , reads as

$$\hat{\mathcal{H}}_D^\kappa = -i\hbar v_F \begin{bmatrix} 0 & \hat{L}_{-\kappa} \\ \hat{L}_\kappa & 0 \end{bmatrix}, \quad \text{with } \hat{L}_\pm = e^{\pm i\theta} \left(\partial_r \pm \frac{i}{r} \partial_\theta \right). \quad (\text{S1})$$

The corresponding solutions for systems of azimuthal symmetry take the general form [S5]:

$$\psi_l^\kappa(r, \theta) = e^{il\theta} \begin{bmatrix} f_l^A(r) \\ e^{ik\theta} f_l^B(r) \end{bmatrix}, \quad \text{with } l \in \mathbb{Z}. \quad (\text{S2})$$

A. Nonzero-energy states

The explicit form of $f_l^A(r)$ and $f_l^B(r)$ is readily determined after simple manipulations of Eqs. (S1) and (S2), while enforcing regularity at the origin $r = 0$, yielding the spinor in Eq. (3), reproduced here for convenience:

$$\psi_{ln}^\kappa(r, \theta) = \frac{e^{il\theta}}{\sqrt{N_{ln}^\kappa}} \begin{bmatrix} J_l(k_{ln}^\kappa r) \\ ik J_{l+\kappa}(k_{ln}^\kappa r) e^{ik\theta} \end{bmatrix}, \quad (\text{S3})$$

expressed via effective momenta $k_{ln}^k = \beta_{ln}^k/R$ associated with nonzero energies $\epsilon_{ln}^k = \hbar\omega_R\beta_{ln}^k$, where $\omega_R = v_F/R$. The normalization N_{ln}^k ensures that $\langle \psi_{ln}^k | \psi_{ln}^k \rangle = 1$, and is generally expressible as [S6, (5.54.2)]:

$$N_{ln}^k = \pi R^2 \sum_{\zeta \in \{0, \kappa\}} \left[J_{l+\zeta}^2(\beta_{ln}^k) - J_{l-1+\zeta}(\beta_{ln}^k) J_{l+1+\zeta}(\beta_{ln}^k) \right]. \quad (\text{S4})$$

For the zigzag boundary condition (ZZ BC), which enforces $J_l(\beta_{ln}) = 0$, its form is particularly simple since Bessel zeros obey $J_{l-1}(\beta_{ln}) = -J_{l+1}(\beta_{ln})$ cf. the Bessel recursion relation $J_l(x) = \frac{x}{2l} [J_{l-1}(x) + J_{l+1}(x)]$, such that the zigzag normalization reduces to a valley-invariant form $N_{ln}^{ZZ} = 2\pi R^2 J_{l-1}^2(\beta_{ln}) = 2\pi R^2 J_{l+1}^2(\beta_{ln})$.

B. Zero-energy states

For zero-energy states the two-spinor equation decouples into two separate homogeneous equations $\hat{L}_{\kappa} \psi_{\kappa}^k(r, \theta) = 0$ and $\hat{L}_{-\kappa} \psi_{\kappa}^k(r, \theta) = 0$. Inserting the azimuthal form Eq. (S2) into these equations and solving produces solutions of the form $f_l^{\kappa}(r) = c^{\kappa} r^{\kappa l}$ and $f_l^{\kappa}(r) = c^{\kappa} r^{-\kappa(l+1)}$. Evidently, it is impossible to find normalizable solutions to the infinite mass boundary condition (IM BC), $\psi_{\kappa}^k(R, \theta)/\psi_{\kappa}^k(R, \theta) = ie^{i\kappa\theta}$, since its enforcement would require simultaneously nonzero coefficients $c^{\kappa, \text{B}}$ [S5].

For the ZZ BC, however, we require only $\psi_{\kappa}^k(R, \theta) = 0$, i.e. $c^{\kappa} = 0$. Regularity at the origin imposes conditions on the allowable l -values, which when relabeled to an integer variable ℓ yields the zero-energy states ϕ_{ℓ}^k from Eq. (4), reproduced here for convenience [S7]:

$$\phi_{\ell}^k(r, \theta) = \frac{e^{-i\kappa\ell\theta}}{\sqrt{\mathcal{N}_{\ell}}} \begin{bmatrix} 0 \\ r^{\ell} \end{bmatrix}, \quad \text{with } \ell = 0, 1, 2, \dots, \ell_{\max}, \quad (\text{S5})$$

with normalization $\mathcal{N}_{\ell} = \pi R^{2(\ell+1)}/(\ell+1)$ [in Eq. (4) of our Communication ϕ_{ℓ}^k is written in terms of the dimensionless coordinate $\tilde{r} = r/R$ – as a consequence, the normalization constant associated with that form is simply $\mathcal{N}_{\ell} = \pi R^2/(\ell+1)$]. Note that, apart from the case $\ell = 0$, the states ϕ_{ℓ}^k have an edge-like quality in the sense that they are localized near $r = R$, with the localization increasing with ℓ .

1. Estimate of the number of zero-energy states in a zigzag-like Dirac disk

The Dirac equation, Eq. (S1), itself offers no bound on the allowable upper value of ℓ . Nevertheless, such bounds can easily be introduced. For instance, the discreteness of the carbon lattice imposes a natural upper bound on the allowable angular momentum on the order of $\ell_{\max} \sim R/a_{\text{LC}}$ [S7]. The introduction of a more precise bound, which agrees qualitatively with the above, can be facilitated by consideration of a result by [Akhmerov and Beenakker](#). In particular, Ref. S2 demonstrates that the density of edge states per unit length of zigzag edge equals approximately $1/3a_{\text{LC}}$ (including spin- and valley-degeneracies), which for a disk of circumference $2\pi R$ yields a maximal number of edge states equal to $N_{\max}^{\text{edge}} = 2\pi R/3a_{\text{LC}}$. Considering the range of ℓ in Eq. (S5) and accounting for degeneracies, it is then apparent that ℓ_{\max} approximately fulfills $4(\ell_{\max} + 1) = N_{\max}^{\text{edge}}$ for a zigzag disk.

Finally, we mention a heuristic modification applied in our calculations to the expression $N_{\max}^{\text{edge}} = 2\pi R/3a_{\text{LC}}$. In particular, we introduce a radial offset in the estimate of N_{\max}^{edge} to account for finite-size effects and the absence of edge states in TB calculations at very small radii. Specifically we find that $N_{\max}^{\text{edge}} \simeq 2\pi(R - R_0)/3a_{\text{LC}}$ with $R_0 = 1.5$ nm agrees well with TB calculations, see the following section, and use this estimate in calculation of ℓ_{\max} .

2. Localization length of zero-energy states in Dirac disk

The degree of localization of the zero-energy states of Eq. (S5) can be investigated by assigning to each edge state of angular momentum ℓ a localization length δR_{ℓ} , defined by the criterion:

$$\int_{R-\delta R_{\ell} \leq r \leq R} d^2 \mathbf{r} |\phi_{\ell}^k(\mathbf{r})|^2 = \text{threshold}, \quad (\text{S6})$$

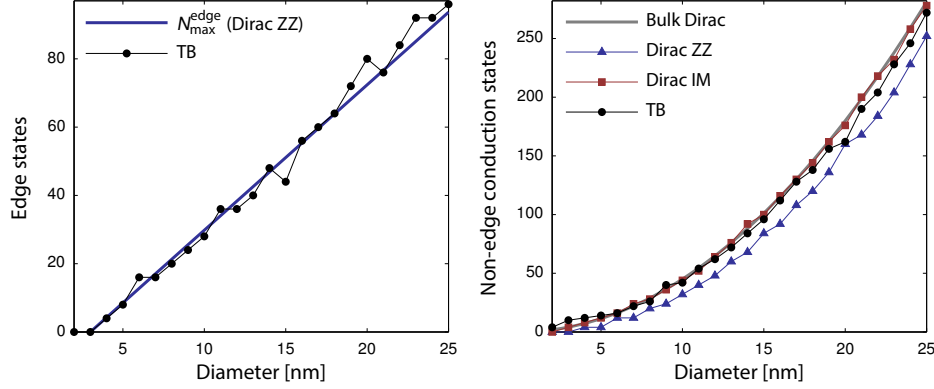


FIG. S2 Total number of edge states (left), as a function of diameter, and total number of non-edge conduction-states for a Fermi level of $\epsilon_f = 0.8$ eV (right) for a CC-centered graphene disk. The calculation of N_{\max}^{edge} includes the heuristic offset R_0 discussed in Section II.B.1.

involving a principally arbitrary threshold, which we choose as $\text{threshold} = 1 - e^{-1}$. With this definition δR_ℓ denotes the width of a two-dimensional annulus with outer radius R containing precisely 50% of the mode-density. Combining Eqs. (S5) and (S6) one finds that $\delta R_\ell/R = 1 - e^{-1/2(\ell+1)}$, i.e. that the characteristic localization length depends exponentially on the angular momentum ℓ .

3. Identification of edge states in tight-binding calculations

To assess the spatial character of the eigenstates of a TB calculation we introduce the participation ratio p of a TB-state ψ [S8]

$$p[\psi] = \frac{\left(\sum_n |\psi_n|^2\right)^2}{N \sum_n |\psi_n|^4}, \quad (\text{S7})$$

where N denotes the number of carbon atoms in the structure and ψ_n denotes a normalized TB-state evaluated at carbon site n . For localized states, where only a small fraction, N_0/N , of carbon sites are occupied by the electron, we have $p \sim N_0/N$, while for extended states with nearly equal site occupancy we have $p \sim 1$. In practice we categorize a given state ψ as a localized edge state if $p[\psi] < p_{\text{cut}}$ with $p_{\text{cut}} = 0.1$ (the total number of edge states being rather insensitive to this particular choice of cutoff). For the optically relevant energy-range $|\epsilon| \lesssim 2$ eV all edge states are found near the Dirac point at zero energy, albeit with a finite energy-spread. Additionally, a cluster of edge states is found near the van Hove singularity at $t_{\text{Ab}} \approx 2.8$ eV which is not included in the count.

We summarize the considerations of Sections II.B.1 and II.B.3 in Fig. S2, wherein we plot the number of edge states in a zigzag Dirac treatment with cutoff ℓ_{\max} and as estimated from a TB-treatment using the participation ratio. Additionally, we also plot the number of non-edge conduction-states, i.e. states with energy $0 < \epsilon \leq \epsilon_f$, and compare with ZZ and IM treatments, and also a bulk Dirac treatment with conduction electron density $n_e = \pi^{-1}(\epsilon_f/\hbar v_F)^2$. A good agreement is observed across all approaches, although the zigzag Dirac treatment slightly underestimates the number of conduction electrons.

III. NON-INTERACTING DENSITY RESPONSE OF DIRAC DISK

In this section, we present analytical expressions for the non-interacting density response function for a Dirac disk. Quite generally, we allow for the existence of nonzero-energy states ψ , see Eq. (S3), and zero-energy states ϕ , see Eq. (S5). Expanding Eq. (5) yields two terms, one due to transitions between nonzero-energy states, $\chi_{\text{bulk-bulk}}^0$, and one due to transitions between zero- and nonzero-energy states,

$\chi_{\text{edge-bulk}}^0$ (but none from transitions between distinct zero-energy states due to Pauli blocking):

$$\chi_0(\mathbf{r}, \mathbf{r}'; \omega) = \chi_{\text{bulk-bulk}}^0(\mathbf{r}, \mathbf{r}'; \omega) + \chi_{\text{edge-bulk}}^0(\mathbf{r}, \mathbf{r}'; \omega), \quad (\text{S8a})$$

$$\chi_{\text{bulk-bulk}}^0(\mathbf{r}, \mathbf{r}'; \omega) = 2 \sum_{\kappa l n l' n'} \frac{f_{l n, l' n'}^{\kappa}}{\epsilon_{l n, l' n'}^{\kappa} - \hbar(\omega + i\eta)} \psi_{l' n'}^{\kappa \dagger}(\mathbf{r}) \psi_{l n}^{\kappa}(\mathbf{r}) \psi_{l n}^{\kappa \dagger}(\mathbf{r}') \psi_{l' n'}^{\kappa}(\mathbf{r}'), \quad (\text{S8b})$$

$$\begin{aligned} \chi_{\text{edge-bulk}}^0(\mathbf{r}, \mathbf{r}'; \omega) &= 2 \sum_{\kappa \ell l n} \frac{f_{0, l n}^{\kappa}}{-\epsilon_{l n}^{\kappa} - \hbar(\omega + i\eta)} \psi_{l n}^{\kappa \dagger}(\mathbf{r}) \phi_{\ell}^{\kappa}(\mathbf{r}) \phi_{\ell}^{\kappa \dagger}(\mathbf{r}') \psi_{l' n'}^{\kappa}(\mathbf{r}') \\ &+ 2 \sum_{\kappa \ell l n} \frac{f_{l n, 0}^{\kappa}}{\epsilon_{l n}^{\kappa} - \hbar(\omega + i\eta)} \phi_{\ell}^{\kappa \dagger}(\mathbf{r}) \psi_{l n}^{\kappa}(\mathbf{r}) \psi_{l' n'}^{\kappa \dagger}(\mathbf{r}') \phi_{\ell}^{\kappa}(\mathbf{r}'), \end{aligned} \quad (\text{S8c})$$

where $\epsilon_{l n, l' n'}^{\kappa} = \epsilon_{l n}^{\kappa} - \epsilon_{l' n'}^{\kappa}$ denotes an energy difference and $f_{l n, l' n'}^{\kappa} = f_{l n}^{\kappa} - f_{l' n'}^{\kappa}$ a population difference between states $\psi_{l n}^{\kappa}$ and $\psi_{l' n'}^{\kappa}$, while f_0 denotes a Fermi function evaluated at zero energy. The inclusion of an explicit valley summation (κ) is unnecessary for the ZZ BC, wherein the valley-degeneracy can be accounted for by a simple factor 2, but essential for the IM BC, where the valley-symmetry is broken. In the following we continue to include this explicit sum for generality.

The bulk-bulk and bulk-edge expressions can be explicated by inserting the spinors from Eqs. (S3) and Eq. (S5):

$$\chi_{\text{edge-bulk}}^0(\mathbf{r}, \mathbf{r}'; \omega) = 2 \sum_{\kappa l n l' n'} \frac{f_{l n, l' n'}^{\kappa}}{\epsilon_{l n, l' n'}^{\kappa} - \hbar(\omega + i\eta)} \frac{e^{i(l-l')(\theta-\theta')}}{N_{l n}^{\kappa} N_{l' n'}^{\kappa}} \sum_{\xi \mu}^{\{0, \kappa\}} \mathbb{J}_{l'+\xi}^{l'+\mu}(r, r'; k_{l' n'}^{\kappa}) \mathbb{J}_{l+\xi}^{l+\mu}(r, r'; k_{l n}^{\kappa}), \quad (\text{S9a})$$

$$\chi_{\text{edge-bulk}}^0(\mathbf{r}, \mathbf{r}'; \omega) = 2 \sum_{s=\pm 1} \sum_{\kappa \ell l n} \frac{s f_{l n, 0}^{\kappa}}{s \epsilon_{l n}^{\kappa} - \hbar(\omega + i\eta)} \frac{e^{is(\kappa \ell + l + \kappa)(\theta - \theta')}}{N_{l n}^{\kappa} N_{\ell}} (r r')^{\ell} \mathbb{J}_{l+\kappa}^{l+\kappa}(r, r'; k_{l n}^{\kappa}), \quad (\text{S9b})$$

where we have introduced short-hand notation for the double Bessel function:

$$\mathbb{J}_q^l(r, r'; k) = J_q(kr) J_l(kr'). \quad (\text{S9c})$$

Taking advantage of the azimuthal symmetry we decompose $\chi^0(\mathbf{r}, \mathbf{r}'; \omega)$ in angular components via

$$\chi^0(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{m=-\infty}^{\infty} \chi_m^0(r, r'; \omega) e^{im(\theta-\theta')}, \quad (\text{S10})$$

with associated parts $\chi_m^0 = \chi_{m|\text{bulk-bulk}}^0 + \chi_{m|\text{edge-bulk}}^0$. After some algebra [which explicitly yield selection rules $m = l - l'$ for the bulk-bulk contribution, and $m = s(\kappa \ell + l + \kappa)$ for the edge-bulk contribution] this allows us to identify the angular components as:

$$\chi_{m|\text{bulk-bulk}}^0(r, r'; \omega) = 2 \sum_{\kappa l n l' n'} \frac{f_{l n, l' n'}^{\kappa}}{\epsilon_{l n, l' n'}^{\kappa} - \hbar(\omega + i\eta)} \frac{1}{N_{l n}^{\kappa} N_{l' n'}^{\kappa}} \sum_{\xi \mu}^{\{0, \kappa\}} \mathbb{J}_{l'-m+\xi}^{l'-m+\mu}(r, r'; k_{l' n'}^{\kappa}) \mathbb{J}_{l+\xi}^{l+\mu}(r, r'; k_{l n}^{\kappa}), \quad (\text{S11a})$$

$$\chi_{m|\text{edge-bulk}}^0(r, r'; \omega) = 2 \sum_{s=\pm 1} \sum_{\kappa \ell l n} \frac{s f_{l n, 0}^{\kappa}}{s \epsilon_{l n}^{\kappa} - \hbar(\omega + i\eta)} \frac{1}{N_{l n}^{\kappa} N_{\ell}} (r r')^{\ell} \mathbb{J}_{L+\kappa}^{L+\kappa}(r, r'; k_{l n}^{\kappa}), \quad (\text{S11b})$$

where short-hand notation $L = sm - \kappa(\ell + 1)$ is understood in the last equation.

In practical calculations we increase the maximal values of l , n , and n' (up to an appropriate energy cutoff) until convergence is reached. Conversely, ℓ is always limited by $\ell_{\text{max}} - \text{incidentally, the resulting optical properties are relatively insensitive to small variations of } \ell_{\text{max}}$.

IV. EDGE-STATE CONDUCTIVITY

In this section we consider the derivation of Eqs. (8) and (9) from (6) assuming throughout a Dirac ZZ treatment (and as such, we do not allow for valley dependence in energies or occupation functions). For convenience, we reiterate the form of the local-response (LR) conductivity $\sigma(\omega)$ due to an x -polarized

incident field:

$$\sigma(\omega) = \frac{2ie^2\omega}{\mathcal{A}} \sum_{\nu\nu'} (f_\nu - f_{\nu'}) \frac{|\langle \psi_{\nu}|x|\psi_{\nu'} \rangle|^2}{\epsilon_\nu - \epsilon_{\nu'} - \hbar(\omega + i\eta)}. \quad (\text{S12})$$

Inserting spinors, Eq. (S3) and (S5), we find, in a similarity with the nonlocal treatment in Section III, a bulk and an edge term:

$$\sigma(\omega) = \tilde{\sigma}_B(\omega) + \sigma_E(\omega), \quad (\text{S13a})$$

$$\tilde{\sigma}_B(\omega) = \frac{2ie^2\omega}{\mathcal{A}} \sum_{\kappa l n l' n'} \frac{f_{l n, l' n'}}{\epsilon_{l n, l' n'} - \hbar(\omega + i\eta)} |\langle \psi_{l n}^\kappa | x | \psi_{l' n'}^\kappa \rangle|^2, \quad (\text{S13b})$$

$$\sigma_E(\omega) = \frac{4ie^2\omega}{\mathcal{A}} \sum_{\kappa l n l} \frac{f_{l n, 0} \epsilon_{l n}}{\epsilon_{l n}^2 - \hbar^2(\omega + i\eta)^2} |\langle \psi_{l n}^\kappa | x | \phi_\ell^\kappa \rangle|^2. \quad (\text{S13c})$$

Again, no term representing transitions between distinct edge-states arise, since their distribution functions are identical and so cancel via $f_\nu - f_{\nu'} = f_0 - f_0 = 0$.

We consider in the following just the edge term $\sigma_E(\omega)$, knowing that the term $\tilde{\sigma}_B(\omega)$ will tend asymptotically towards the bulk infinite-extent graphene conductivity $\sigma_B(\omega)$, see Section V. The transition matrix element can be evaluated analytically and yields:

$$\begin{aligned} |\langle \psi_{l n}^\kappa | x | \phi_\ell^\kappa \rangle|^2 &\stackrel{\text{a}}{=} \frac{\pi^2}{N_{l n} \mathcal{N}_\ell} \left[\sum_{s=\pm 1} \delta_{l, -\kappa(\ell+1+s)} \right] \left| \int_0^R dr r^{\ell+2} J_{l+\kappa}(k_{l n} r) \right|^2 \\ &\stackrel{\text{b}}{=} \sum_{s=\pm 1} \delta_{l, -\kappa(\ell+1+s)} \frac{\pi^2}{N_{\ell+1+s, n} \mathcal{N}_\ell} \left| \int_0^R dr r^{\ell+2} J_{\ell+s}(k_{\ell+1+s, n} r) \right|^2 \\ &\stackrel{\text{c}}{=} \sum_{s=\pm 1} \delta_{l, -\kappa(\ell+1+s)} \frac{\pi^2 R^{2(\ell+3)}}{N_{\ell+1+s, n} \mathcal{N}_\ell} 4\delta_{s, -1} \beta_{\ell n}^{-4} J_{\ell-1}^2(\beta_{\ell n}) \\ &\stackrel{\text{d}}{=} 2\delta_{l, -\kappa\ell} (\ell+1) R^2 \beta_{\ell n}^{-4}, \end{aligned} \quad (\text{S14})$$

where we highlight relevant steps in the derivation in the following (which all rely on the assumption of application of the ZZ BC):

- Orthogonality of azimuthal components. Additionally, normalization is valley-independent.
- Generally $J_l(x) = (-1)^l J_{-l}(x)$, such that $k_{l n} = k_{-l n}$ and $N_{l n} = N_{-l n}$.
- Use of the integral-identity $\int_0^1 d\tilde{r} \tilde{r}^{\ell+2} J_{\ell+s}(\beta_{\ell+1+s, n} \tilde{r}) = 2\delta_{s, -1} \beta_{\ell n}^{-2} J_{\ell-1}(\beta_{\ell n})$, valid for $\beta_{\ell n}$ denoting the n th zero of the Bessel function $J_\ell(x)$. Identity derivable from recurrence relation for the Bessel function and the standard integral identity $\int dx x^{p+1} J_p(x) = x^{p+1} J_{p+1}(x)$.
- Inserting normalization constants $N_{\ell n} = 2\pi R^2 J_{\ell-1}^2(\beta_{\ell n})$ and $\mathcal{N}_\ell = \pi R^{2(\ell+1)} / (\ell+1)$.

Upon combining Eqs. (S13c) and (S14), evaluating the κ -sum (yielding a simple factor 2), and expressing all energies through $\epsilon_{\ell n} = \hbar\omega_R \beta_{\ell n}$:

$$\sigma_E(\omega) = \frac{16ie^2}{\pi\hbar} \frac{\omega}{\omega_R} \sum_{\ell n} \frac{\ell+1}{\beta_{\ell n}^5} \frac{f_{\ell n, 0}}{1 - \left(\frac{\omega+i\eta}{\beta_{\ell n} \omega_R}\right)^2}. \quad (\text{S15})$$

In the low-temperature limit, $T \rightarrow 0$, the occupation term reduces to $f_{\ell n, 0} = \theta(\hbar\omega_R \beta_{\ell n} - \epsilon_F) - 1$, assuming $\epsilon_F > 0$. With this simplification we readily find the result in Eq. (8) of our Communication.

A. Large-radius asymptotic form

Rather than deal explicitly with the details of the zeros of the Bessel function, $\beta_{\ell n}$, it is convenient to note that the summation in (8) simply runs over all energies above the Fermi-level for states with angular momentum $0 \leq \ell \leq \ell_{\max}$. In the large radius limit, the energies tend to $\epsilon \approx \hbar v_F k$, with k denoting a continuous momentum, and an associated density of states $D(\epsilon) = g_d |\epsilon| / 2\pi \hbar^2 v_F^2$ (where $g_d = 1$ since we

have already accounted for spin- and valley-degeneracies explicitly). Ignoring the upper limit, ℓ_{\max} , on ℓ , which is reasonable in the large-disk case, this allows us to transform the sum via:

$$\frac{1}{\mathcal{A}} \sum_{\ell n}^{\hbar\omega_R\beta_{\ell n} \geq \epsilon_F} \rightarrow \frac{1}{2} \int_{\epsilon_F}^{\infty} d\epsilon D(\epsilon), \quad (\text{S16})$$

where the factor 1/2 accounts for the fact that we only include terms $\ell \geq 0$ in the sum rather than all $\ell \in \mathbb{Z}$. The Bessel zeros are naturally converted via $\beta_{\ell n} \rightarrow \epsilon/\hbar\omega_R$. The remaining difficulty lies with the factor $\ell + 1$. We replace it by its average value at fixed energy, i.e. $\ell + 1 \rightarrow \langle \ell + 1 \rangle_\epsilon$, which in turn can be approximated by $\langle \ell + 1 \rangle_\epsilon \simeq \xi\epsilon/\hbar\omega_R$, where ξ denotes a proportionality constant $\xi = 4/3\pi$, see Section IV.A.1.

Introducing these considerations into Eq. (8) then allows us to derive (9):

$$\begin{aligned} \sigma_E^\infty(\omega) &= \frac{-16ie^2}{\pi\hbar} \frac{\omega}{\omega_R} \frac{\xi(\hbar\omega_R)^2}{4} \int_{\epsilon_F}^{\infty} d\epsilon \frac{1}{\epsilon^3} \frac{1}{1 - \left[\frac{\hbar(\omega+i\eta)}{\epsilon}\right]^2} \\ &= -\xi \frac{4ie^2}{\pi} \hbar\omega\omega_R \int_{\epsilon_F}^{\infty} d\epsilon \frac{1}{2\epsilon^2} \left[\frac{1}{\epsilon - \hbar(\omega + i\eta)} + \frac{1}{\epsilon + \hbar(\omega + i\eta)} \right] \\ &= -\xi \frac{2ie^2}{\pi} \hbar\omega\omega_R \int_{\epsilon_F}^{\infty} d\epsilon \frac{1}{\epsilon^2} \left[\mathcal{P}\left(\frac{1}{\epsilon - \hbar\omega}\right) + \mathcal{P}\left(\frac{1}{\epsilon + \hbar\omega}\right) + i\pi\delta(\epsilon - \hbar\omega) - i\pi\delta(\epsilon + \hbar\omega) \right] \\ &= \xi \frac{2e^2}{\pi\hbar} \frac{\omega_R}{\omega} \left[i \ln \left| \frac{\epsilon_F^2 - \hbar^2\omega^2}{\epsilon_F^2} \right| + \pi\theta(\hbar\omega - \epsilon_F) \right], \end{aligned} \quad (\text{S17})$$

where we have used the identity $\frac{1}{x \pm i\eta} = \mathcal{P}\frac{1}{x} \mp i\pi\delta(x)$ valid for $\eta \rightarrow 0^+$, and that $\{\epsilon_F, \omega\} > 0$.

1. Evaluation of proportionality constant ξ

The value of ξ can be determined by considering initially the average value of ℓ . Suppose that the probability of finding a Bessel zero $J_\ell(\beta_{\ell n}) = 0$ near some value $\epsilon/\hbar\omega_R$ is given by $P_\ell(\epsilon)$, and that the maximal ℓ -value (after which all $P_\ell(\epsilon) = 0$) is L , then:

$$\langle \ell \rangle_\epsilon = \frac{\sum_{\ell=0}^L P_\ell(\epsilon)\ell}{\sum_{\ell=0}^L P_\ell(\epsilon)}. \quad (\text{S18})$$

Evaluating this expression is straight-forward numerically. Specifically, using the following recipe one can estimate the probability $P_\ell(\epsilon)$:

1. For fixed energy ϵ , and small energy-interval $\Delta\epsilon$, find all zeros, b_j , of the Bessel function $J_\ell(x)$ in the interval $x \in [\epsilon - \frac{1}{2}\Delta\epsilon/2, \epsilon + \frac{1}{2}\Delta\epsilon]/\hbar\omega_R$ for all $\ell = 0, 1, \dots, L$. Denote this set as $\mathcal{B} = \{b_1, b_2, \dots, b_M\}$, with M denoting the cardinality of the set.
2. Let $\#\ell(\mathcal{B})$ denote the number of zeros in \mathcal{B} of order ℓ . The probability is then $P_\ell(\epsilon) = \#\ell(\mathcal{B})/M$.

Having computed $P_\ell(\epsilon)$ and thereby $\langle \ell \rangle_\epsilon$ the value of ξ follows upon comparison with the ansatz $\langle \ell + 1 \rangle_\epsilon = \langle \ell \rangle_\epsilon + 1 \simeq \xi\epsilon/\hbar\omega_R$. In Fig. S3 we show computed values of $P_\ell(\epsilon)$, averaged across bins. A clear trend is evident and indicated in dashed blue; the probability is approximately of the form $P_\ell(\epsilon) \simeq A\sqrt{1 - \ell^2/L^2}$, with A being an undetermined scaling. With this form of the probability, we can approximate the expression in Eq. (S18) through the continuum limit via $\langle \ell \rangle_\epsilon \simeq \left[\int_0^L d\ell P_\ell(\epsilon)\ell \right] / \left[\int_0^L d\ell P_\ell(\epsilon) \right]$, where we treat ℓ now as a quasi-continuous variable. Evaluating this expression yields $\langle \ell + 1 \rangle_\epsilon \simeq \langle \ell \rangle_\epsilon \simeq \frac{4}{3\pi}L$ (where the approximate equivalence of $\langle \ell + 1 \rangle_\epsilon$ and $\langle \ell \rangle_\epsilon$ follows from our assumption of large $\epsilon/\hbar\omega_R$). Finally, we note that $L \simeq \epsilon/\hbar\omega_R$, as is also evident from Fig. S3 (this is asymptotically true for large L , since the first zero, $\beta_{\ell,1}$, of $J_\ell(x)$, goes like $\beta_{\ell,1} \sim \ell + O(\ell^{1/3})$ for large ℓ , while all zeros of differing orders are bounded by $\beta_{\ell,1} < \beta_{\ell+1,1} < \dots$ [S9, (10.21.2) & (10.21.40)]). Thus, we find analytically that $\langle \ell + 1 \rangle_\epsilon \simeq \xi\epsilon/\hbar\omega_R$ with $\xi = 4/3\pi$. We compare the numerical and analytical estimates in Fig. S3 and find excellent agreement across a wide range of $\epsilon/\hbar\omega_R$, improving with larger effective energies.

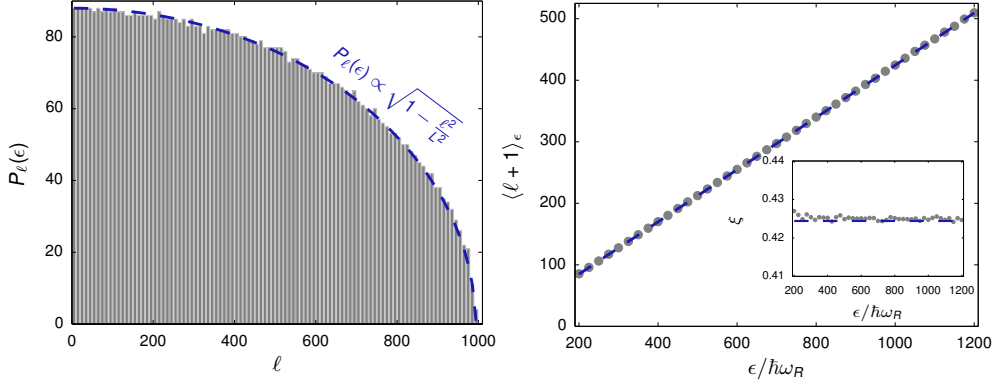


FIG. S3 Numerical and analytical considerations of the probability $P_\ell(\epsilon)$ and the proportionality constant ξ . Left: Probability $P_\ell(\epsilon)$ calculated numerically for $\epsilon/\hbar\omega_R = 1000$ and $\Delta\epsilon/\hbar\omega_R = 25$ with adjoining ℓ -values collected in bins of size 10. Indicated in dashed blue is a fit $P_\ell(\epsilon) = A\sqrt{1 - \ell^2/L^2}$, with L denoting the maximal ℓ -value of the computed Bessel zeros. Right: The average value $\langle \ell + 1 \rangle_\epsilon$ as a function of $\epsilon/\hbar\omega_R$, computed numerically using Eq. (S18) with $\Delta\epsilon/\hbar\omega_R = 25$ and analytically via $\langle \ell + 1 \rangle_\epsilon \simeq \xi\epsilon/\hbar\omega_R$ with $\xi = 4/3\pi$, in gray markers and dashed blue, respectively. Inset shows computed values of ξ for different effective energies compared with the value $4/3\pi$.

V. LOCAL-RESPONSE BULK CONDUCTIVITY

For completeness, and for comparison with the result for the edge-state conductivity, we here also give the local-response bulk conductivity, $\sigma_B(\omega)$, utilized in Fig. 4 of our Communication. At finite temperature, T , the local-response intra- and interband conductivity terms in $\sigma_B(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega)$ are [S10]:

$$\sigma_{\text{intra}}(\omega) = \sigma_0 \frac{8ik_B T}{\pi\hbar(\omega + i\gamma)} \ln \left[2 \cosh \left(\frac{\epsilon_F}{2k_B T} \right) \right], \quad (\text{S19a})$$

$$\sigma_{\text{inter}}(\omega) = \sigma_0 \left[H(\hbar\omega/2) + \frac{4i\hbar(\omega + i\gamma)}{\pi} \int_0^\infty d\epsilon \frac{H(\epsilon) - H(\hbar\omega/2)}{\hbar^2(\omega + i\gamma)^2 - 4\epsilon^2} \right], \quad (\text{S19b})$$

where $\gamma = 2\eta$ denotes the optical loss-rate, $\sigma_0 \equiv e^2/4\hbar$ is a characteristic conductive magnitude, and

$$H(\epsilon) = f(-\epsilon) - f(\epsilon) = \frac{\sinh(\epsilon/k_B T)}{\cosh(\epsilon_F/k_B T) + \cosh(\epsilon/k_B T)}, \quad (\text{S19c})$$

denotes a population difference. Following the prescription in Refs. S11 and S12 we have included phenomenological loss in both intra- and interband terms via γ . In the low-loss, low-temperature limit the terms are particularly simple, and especially the interband term is mathematically quite evocative of $\sigma_E^\infty(\omega)$ from Eq. (S17):

$$\sigma_{\text{intra}}(\omega) \stackrel{\gamma, T=0}{=} \sigma_0 \frac{4i\epsilon_F}{\pi\hbar\omega}, \quad \sigma_{\text{inter}}(\omega) \stackrel{\gamma, T=0}{=} \sigma_0 \left[\frac{i}{\pi} \ln \left| \frac{2\epsilon_F - \hbar\omega}{2\epsilon_F + \hbar\omega} \right| + \theta(\hbar\omega - 2\epsilon_F) \right]. \quad (\text{S20})$$

VI. LINEARIZED HYDRODYNAMIC MODEL FOR GRAPHENE

In this section we discuss the use of an approximate hydrodynamic model for graphene, and compare with the results of a small-momentum expansion of the bulk graphene density response from RPA. We also show numerically that absorption spectra calculated using Eq. (S11a) with Dirac IM states agree excellently with predictions of a hydrodynamic description in the large radius limit.

A. Formulation of model and comparison with bulk RPA

Linearized hydrodynamic descriptions, recently widely applied to metallic systems, relate the induced current \mathbf{J} and the electric field \mathbf{E} . For graphene, due to its two-dimensional nature, only the in-plane components of the electric field, \mathbf{E}_\parallel , and the surface-current \mathbf{K} are of relevance. In this case, the hydrodynamic equivalent of Ohm's law (neglecting loss for the nonce) reads as [S13]:

$$\mathbf{K}(\mathbf{r}, \omega) + \frac{\beta^2}{\omega^2} \nabla_\parallel [\nabla_\parallel \cdot \mathbf{K}(\mathbf{r}, \omega)] = \sigma(\omega) \mathbf{E}_\parallel(\mathbf{r}, \omega), \quad (\text{S21})$$

where $\sigma(\omega)$ is a local conductivity, and β a plasma velocity.

It is instructive to consider the resulting hydrodynamic conductivity in momentum space for longitudinal excitations $\mathbf{K}(k, \omega) = \sigma_\parallel^L(k, \omega) \mathbf{E}_\parallel(k, \omega)$ where $\mathbf{k} \parallel \mathbf{E}_\parallel$:

$$\sigma_\parallel^L(k, \omega) = \sigma(\omega) \left(1 - \beta^2 \frac{k^2}{\omega^2} \right) \simeq \sigma(\omega) \left(1 + \beta^2 \frac{k^2}{\omega^2} \right), \quad (\text{S22})$$

assuming $k \ll \omega/\beta$. The values of β and $\sigma(\omega)$ appropriate for graphene, can be discerned by comparison with the small-momentum expansion of the low-temperature nonlocal conductivity for bulk graphene $\sigma^0(k, \omega)$, whose full expression reads as [S14; S15]:

$$\sigma^0(k, \omega) = \frac{ie^2}{\pi\hbar} \frac{\tilde{\omega}}{\tilde{k}^2} \left\{ -2 + \frac{1}{4} F(\tilde{k}, \tilde{\omega}) \left[\mathcal{W}\left(\frac{2+\tilde{\omega}}{\tilde{k}}\right) - \mathcal{W}\left(\frac{2-\tilde{\omega}}{\tilde{k}}\right) \right] \right\}, \quad (\text{S23a})$$

$$F(\tilde{k}, \tilde{\omega}) = \frac{\tilde{k}^2}{\sqrt{\tilde{\omega}^2 - \tilde{k}^2}}, \quad \mathcal{W}(x) = x \sqrt{x^2 - 1} - \ln(x + \sqrt{x^2 - 1}), \quad (\text{S23b})$$

expressed in terms of normalized momenta and frequencies $\tilde{k} = k/k_F$ and $\tilde{\omega} = \hbar\omega/\epsilon_F$. A series expansion of this expression yields (assuming $\tilde{\omega} < 2$ to restrict the study to regions without Landau damping):

$$\sigma^0(k, \omega) = \underbrace{\frac{ie^2}{\pi\hbar} \frac{\epsilon_F}{\hbar\omega}}_{\sigma_{\text{intra}}(\omega)} \left[1 + \beta_{\text{intra}}^2(\omega) \frac{k^2}{\omega^2} \right] + \underbrace{\frac{ie^2}{4\pi\hbar} \ln\left(\frac{2\epsilon_F - \hbar\omega}{2\epsilon_F + \hbar\omega}\right)}_{\sigma_{\text{inter}}(\omega)} \left(1 + \beta_{\text{inter}}^2 \frac{k^2}{\omega^2} \right), \quad (\text{S24a})$$

where we have indicated association with the usual local-response conductivities. The plasma velocities for the intra- and interband terms are given by:

$$\beta_{\text{intra}}^2(\omega) = \left\{ \frac{3}{4} - \left[\frac{1}{4(\epsilon_F/\hbar\omega)^2 - 1} \right]^2 \right\} v_F^2 \stackrel{\hbar\omega \leq \epsilon_F}{\simeq} \frac{3}{4} v_F^2, \quad \beta_{\text{inter}}^2 = \frac{1}{2} v_F^2. \quad (\text{S24b})$$

Evidently, an appropriate hydrodynamic treatment would require a two-fluid model, accounting for the difference in plasma velocities for the intra- and interband terms. Nevertheless, a one-fluid model is reasonable, since the interband-term - and hence its nonlocal correction - is small compared to the intraband term. Hence, upon also ignoring the frequency-dispersion of β_{intra} (reasonable for disks in the large radius limit where the plasmonic resonances lie below the Fermi level), we find that Eq. (S21) with $\beta^2 = \frac{3}{4} v_F^2$ and $\sigma(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega)$ provides a reasonable accounting for the lowest order nonlocal corrections of the electronic plasma response.

In calculations we also account approximately for relaxation in the nonlocal part of Eq. (S21), with a rate γ , by introducing the substitutions $\beta^2 \rightarrow \frac{\omega}{\omega+i\gamma} \beta^2$ [S16; S17].

B. Comparison of hydrodynamics and RPA with Dirac IM states

In Fig. S4 we show computational results for graphene disks of varying diameter. Very good agreement between RPA with Dirac IM states and a hydrodynamic description based on the bulk conductivity $\sigma_B(\omega)$ is obtained, apart from the smallest considered diameter. Since RPA with Dirac IM states fully accounts for the effects of nonlocality, but completely neglects the presence of edge states, this underlines the excellent description of nonlocality by a simple hydrodynamic model with $\beta^2 = \frac{3}{4} v_F^2$.

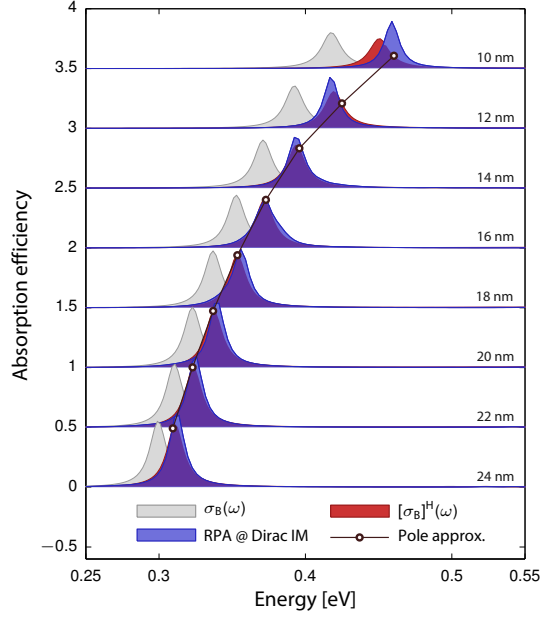


FIG. S4 Absorption cross-sectional efficiency in graphene disks, with $\epsilon_F = 0.4$ eV, $\hbar\gamma = 12$ meV, and $T = 300$ K and varying diameter (indicated above each spectra). Spectra calculated from local-response with bulk conductivity $\sigma_B(\omega)$, from hydrodynamic response with bulk conductivity (indicated in legend as $[\sigma_B]^{H}(\omega)$), and from RPA with Dirac IM states. A pole approximation for the spectral position of the hydrodynamic resonances is also indicated, see Eq. (S35). Spectra from different disk diameters are offset by 0.5.

VII. ELECTROSTATIC INTERACTION IN LOCAL-RESPONSE NANODISK

A complete solution of the electrostatic problem in a two-dimensional conductive disk, accounting also for a hydrodynamic interaction in the form of Eq. (S21) and screening due to nearby grounded planes, was offered by Fetter in Ref. S18. For convenience, we repeat here the most important steps of the derivation as relevant to our case (i.e. in the absence of grounded planes). Throughout we suppress explicit declaration of frequency dependence.

The electrostatic potential, $\phi(\mathbf{r})$, is governed by the Poisson equation with a charge number density $\rho(\mathbf{r}) = \delta(z)\rho_{\parallel}(\mathbf{r}_{\parallel})$, where $\rho_{\parallel}(\mathbf{r}_{\parallel})$ indicates the induced in-plane charge density in the graphene sample. For a disk $\rho_{\parallel}(\mathbf{r}_{\parallel}) = \rho_{\parallel}(r, \theta)$ which vanishes for $r > R$. Thus;

$$\nabla^2 \phi(\mathbf{r}) = -\frac{1}{\epsilon_0 \epsilon(\mathbf{r})} \delta(z) \rho_{\parallel}(r, \theta), \quad (\text{S25})$$

where we allow for a dielectric background above and below the disk via $\epsilon(\mathbf{r}) = \epsilon_+ \theta(z) + \epsilon_- \theta(-z)$ [and $\epsilon(\mathbf{r}_{\parallel}, z=0) = \frac{1}{2}(\epsilon_+ + \epsilon_-) \equiv \epsilon_B$].

Due to the azimuthal symmetry of the disk, the potential can be decomposed in cylindrical coordinates via $\phi(\mathbf{r}) = \phi_{\parallel}(r)\phi_{\perp}(z)e^{i\theta}$ [with $\phi_{\perp}(0) \equiv 1$], and similarly for the charge density $\rho_{\parallel}(\mathbf{r}_{\parallel}) = \rho_{\parallel}(r)e^{i\theta}$ [we suppress explicit indication of the l -dependence of $\phi_{\parallel}(r)$, $\phi_{\perp}(z)$, and $\rho_{\parallel}(r)$ in the following, and assume it implicitly understood]. Carrying out an l -order Hankel transform of Eq. (S25) in coordinate r allows solving for $\phi_{\perp}(z)$, which, through use of the boundary conditions for the potential, allows a relation between the Hankel transformed in-plane potential and charge density. By an inverse l -order Hankel transform of this relation one arrives at:

$$\phi_{\parallel}(\tilde{r}) = \frac{R}{2\epsilon_0 \epsilon_B} \int_0^1 d\tilde{r}' K_l(\tilde{r}, \tilde{r}') \rho_{\parallel}(\tilde{r}') \tilde{r}', \quad (\text{S26})$$

with kernel $K_l(\tilde{r}, \tilde{r}') \equiv \int_0^{\infty} dp J_l(p\tilde{r}) J_l(p\tilde{r}')$, written in terms of normalized radial coordinates $\tilde{r} = r/R$.

Similarly, we can derive an additional relation between ϕ_{\parallel} and ρ_{\parallel} by considering the hydrodynamic constitutive equation of Eq. (S21), which in the electrostatic limit is equivalent to

$$(1 + k_{\text{NL}}^{-2} \nabla_{\parallel}^2) \rho_{\parallel}(\mathbf{r}_{\parallel}) = i\omega^{-1} \sigma \nabla_{\parallel}^2 \phi_{\parallel}(\mathbf{r}_{\parallel}), \quad (\text{S27})$$

with $k_{\text{NL}} = \omega/\beta$, valid for $r < R$. The point at $r = R$ requires special attention: it can either be accounted for by including a Dirac delta term, $\sim \delta(r - R)$, in the constitutive equation, *or* by imposing a boundary condition at the edge. The boundary conditions of relevance is the no-spill current condition $\hat{\mathbf{r}} \cdot \mathbf{K}(r = R, \theta) = 0$ and the boundedness of $\phi_{\parallel}(r)$ and $\rho_{\parallel}(r)$. The no-spill condition can be converted to a condition on $\phi_{\parallel}(r = R)$ and $\rho_{\parallel}(r = R)$ by projecting on Eq. (S21).

Following this approach we can solve Eq. (S27) for $\phi_{\parallel}(\tilde{r})$ by using a Green function to account for the boundary condition, resulting in:

$$\phi_{\parallel}(\tilde{r}) = -i\omega\sigma^{-1} k_{\text{NL}}^{-2} \rho_{\parallel}(\tilde{r}) + i\omega\sigma^{-1} R^2 \int_0^1 d\tilde{r}' G_l(\tilde{r}, \tilde{r}') \rho_{\parallel}(\tilde{r}') \tilde{r}', \quad (\text{S28})$$

where $G_l(\tilde{r}, \tilde{r}') = (2l)^{-1} [(\tilde{r}\tilde{r}')^l + (\tilde{r}_</\tilde{r}_>)^l]$ [with $\tilde{r}_< = \min(\tilde{r}, \tilde{r}')$ and $\tilde{r}_> = \max(\tilde{r}, \tilde{r}')$] is a Green function defined by $\Delta_{\text{B}}^l G(\tilde{r}, \tilde{r}') = -\tilde{r}^{-1} \delta(\tilde{r} - \tilde{r}')$ subject to $\partial_{\tilde{r}} G(1, \tilde{r}') = 0$ and boundedness for $\{\tilde{r}, \tilde{r}'\} \in [0, 1]$ (where Δ_{B}^l denotes the Bessel differential operator of order l acting on \tilde{r}) [S19, Section 43]. Here we have implicitly assumed that $l \neq 0$: the case of axisymmetric modes requires a separate treatment, due to the necessity of introducing an additional boundary condition which ensures that the total induced charge vanishes, see Ref. S18.

By combining Eqs. (S26) and (S28) self-consistent equations for either the in-plane potential or charge density can be found - which, in the latter case, offers:

$$\frac{\beta^2}{R^2} \rho_{\parallel}(\tilde{r}) - \omega^2 \int_0^1 G_l(\tilde{r}, \tilde{r}') \rho_{\parallel}(\tilde{r}') \tilde{r}' + \Omega_0^2(\omega) \int_0^1 K_l(\tilde{r}, \tilde{r}') \rho_{\parallel}(\tilde{r}') \tilde{r}' = 0, \quad (\text{S29})$$

where $\Omega_0^2(\omega) \equiv -i\omega\sigma(\omega)/2\varepsilon_0\varepsilon_{\text{B}}R$ is a characteristic round trip frequency associated with traversal of the disk at a velocity defined by $\sigma(\omega)$ and ε_{B} . The solution of this integral equation gives the eigenmodes, ω_n and $\rho_{\parallel n}(r)$, of the disk for modes of angular momentum l .

A. Reduction to matrix eigenvalue problem by polynomial expansion

Rather astonishingly, the complicated integral equation in Eq. (S29) can be reduced to a matrix eigenvalue problem *with analytical matrix elements*. This fact was demonstrated in Ref. S18, by use of an expansion in a complete set of Jacobi polynomials $\{P_j^{l,0}(1 - 2\tilde{r}^2)\}_{j=0}^{\infty}$:

$$\rho_{\parallel}(\tilde{r}) = \tilde{r}^l \sum_{j=0}^{\infty} c_j P_j^{l,0}(1 - 2\tilde{r}^2). \quad (\text{S30})$$

The reduction to a matrix equation is achieved by juxtaposing $x^{l+1} P_k^{l,0}(1 - 2x^2)$ onto Eq. (S29) and integrating over \tilde{r} , which, after utilizing the orthogonality of Jacobi polynomials, yields an algebraic equation $\sum_{j=0}^{\infty} [\frac{\beta^2}{R^2} \mathbf{D}_{jk} - \omega^2 \mathbf{G}_{jk} + \Omega_0^2(\omega) \mathbf{K}_{jk}] c_j = 0$. Truncating to some cutoff $\{j, k\} = 0, 1, \dots, J$ allows reformulating this as a $J \times J$ matrix equation:

$$\left[\frac{\beta^2}{R^2} \mathbf{D} - \omega^2 \mathbf{G} + \Omega_0^2(\omega) \mathbf{K} \right] \mathbf{c} = 0, \quad (\text{S31a})$$

where the matrix elements can be evaluated exactly and take the form:

$$D_{jk} = \frac{\delta_{jk}}{2(l+2j+1)}, \quad (\text{S31b})$$

$$G_{jk} = \frac{\delta_{j0}\delta_{k0}}{8l(l+1)^2} + \frac{\delta_{jk}}{4(l+2j)(l+2j+1)(l+2j+2)} + \frac{\delta_{j+1,k} + \delta_{j,k+1}}{8(l+2j+1)(l+2j+2)(l+2j+3)}, \quad (\text{S31c})$$

$$K_{jk} = \frac{(-1)^{k-j+1}}{\pi[4(k-j)^2 - 1](l+k+j+\frac{1}{2})(l+k+j+\frac{3}{2})}. \quad (\text{S31d})$$

1. Eigenmode resonance frequencies in local-response

In the absence of nonlocality the matrix equation reduces to a generalized eigenvalue problem $\mathbf{K}\mathbf{c}_n = \zeta_n \mathbf{G}\mathbf{c}_n$ with eigenvector \mathbf{c}_n and eigenvalues $\zeta_n = \omega_n^2 / \Omega_0^2(\omega_n)$. The resonances can thus be categorized by the value of ζ_n . In particular, the electrostatic resonances are thus governed by:

$$\frac{\omega_n}{\sigma(\omega_n)} = \frac{\zeta_n}{2i\epsilon_0\epsilon_n R}, \quad (\text{S32})$$

where the value of l used in construction of \mathbf{K} and \mathbf{G} dictates the angular momentum of the mode, while n gives its radial quantization ($n = 1, 2, \dots$ for singly, doubly, etc.). In Table I we summarize values of ζ_n for different angular momenta and radial quantization. In the electrostatic regime, the eigenmode relevant for plane-wave interaction in nanodisks is naturally the $l = n = 1$ resonance.

n	$l = 1$	$l = 2$	$l = 3$	$l = 4$
1	1.0977	1.9942	2.8556	3.7032
2	4.9140	6.2455	7.5124	8.7395
3	8.1337	9.5455	10.8989	12.2117
4	11.3079	12.7592	14.1596	15.5221

TABLE I Tabulated values of ζ_n for eigenmodes of angular momentum l and radial quantization n . Calculated with a cutoff of $J = 250$, giving fully converged results accurate to last indicated decimal.

2. Pole approximations for hydrodynamic and edge-state conductivity contributions

Assuming that the hydrodynamic response is a small perturbation we can justify making a pole approximation to assess the impact of nonlocality. Specifically, if we indicate the local-response eigenmodes by ω_n^0 and \mathbf{c}_n^0 the pole approximation of Eq. (S31a) becomes:

$$\left[\frac{\beta^2}{R^2} \mathbf{D} - \omega^2 \mathbf{G} + \Omega_0^2(\omega_n^0) \mathbf{K} \right] \mathbf{c}_n = 0 \quad \Leftrightarrow \quad \frac{\beta^2}{R^2} \mathbf{D} \mathbf{c}_n = \omega^2 \mathbf{G} \mathbf{c}_n - (\omega_n^0)^2 \mathbf{G} \mathbf{c}_n^0 - \Omega_0^2(\omega_n^0) \mathbf{K} \delta \mathbf{c}_n, \quad (\text{S33})$$

where we have used that $\Omega_0^2(\omega_n^0) \mathbf{K} \mathbf{c}_n^0 = \omega_n^0 \mathbf{G} \mathbf{c}_n^0$ and also introduced the small difference $\delta \mathbf{c}_n = \mathbf{c}_n - \mathbf{c}_n^0$. Ignoring this small difference, thereby letting $\mathbf{c}_n^0 \simeq \mathbf{c}_n$, this equation can be recast as a generalized eigenvalue problem with eigenvalues Δ_n :

$$\mathbf{D} \mathbf{c}_n \simeq \frac{R^2}{\beta^2} [\omega_n^2 - (\omega_n^0)^2] \mathbf{G} \mathbf{c}_n \equiv \Delta_n \mathbf{G} \mathbf{c}_n. \quad (\text{S34})$$

Thus, within the pole approximation, the resonance frequency is approximately:

$$\omega_n = \sqrt{(\omega_n^0)^2 + \Delta_n \frac{\beta^2}{R^2}} \simeq \omega_n^0 + \frac{\Delta_n \beta^2}{2\omega_n^0 R^2}, \quad (\text{S35})$$

where all the eigenvalues Δ_n are positive, leading invariably to a blueshift of the resonance, with a strength $\sim \omega_R^2 / \omega_n^0$. For the optically relevant $l = n = 1$ mode we find numerically that $\Delta_1 \approx 3.39$. See Fig. S4 for

an assessment of the accuracy of the approximation.

By a similar reasoning, the approximate impact of an edge-state conductivity $\sigma_E(\omega)$ [see Eq. (S17)] can also be accounted for. Introducing again a pole approximation by letting $\sigma(\omega_n) = \sigma_B(\omega_n) + \sigma_E(\omega_n) \rightarrow \sigma_B(\omega_n^0) + \sigma_E(\omega_n^0)$ in Eq. (S32), and using Eq. (S17), one finds:

$$\omega_n \simeq \omega_n^0 + \frac{\zeta_n \sigma_E(\omega_n^0)}{2i\epsilon_0 \epsilon_B R} \simeq \omega_n^0 + \frac{4\zeta_n \xi \alpha_g}{\epsilon_B} \frac{\omega_R^2}{\omega_n^0} \ln \left| \frac{\epsilon_F^2 - (\hbar\omega_n^0)^2}{\epsilon_F^2} \right|, \quad (\text{S36})$$

where we have introduced the effective fine-structure constant for graphene $\alpha_g = e^2/4\pi\epsilon_0\hbar v_F \approx 2.40$. It is evident that the edge-contribution leads to a redshift since $\text{Im}[\sigma_E(\omega_n^0)] < 0$ for $\hbar\omega_n^0 < \sqrt{2}\epsilon_F$. Note that the strength of the hydrodynamic and edge-contribution corrections scale similarly, i.e. with ω_R^2/ω_n^0 .

B. Interaction with external potentials and the absorption cross-section

Including interaction with external potentials of definite angular momentum l is straightforward. In particular, if $\phi_{\parallel}^{\text{ext}}(\tilde{r}) = \tilde{r}^l \sum_j^{\infty} d_j P_j^{l,0}(1 - 2\tilde{r}^2)$ then the matrix system in Eq. (S31a) evolves into the inhomogeneous system:

$$\left[\frac{\beta^2}{R^2} \mathbf{D} - \omega^2 \mathbf{G} + \Omega_0^2(\omega) \mathbf{K} \right] \mathbf{c} = \frac{i\omega\sigma(\omega)}{R^2} \mathbf{D} \mathbf{d}, \quad (\text{S37})$$

The important case of the absorption cross-section, that is, the absorbed power relative to the incident intensity of a plane wave, is similarly straightforward. In particular, considering an x -polarized plane wave normally incident on the disk, $\mathbf{E}^{\text{ext}}(\mathbf{r}_{\parallel}, z = 0) = E_0 \hat{\mathbf{x}}$, the associated electrostatic potential $\phi_{\parallel}^{\text{ext}}(\mathbf{r}_{\parallel}) = -E_0 x = -\frac{1}{2} E_0 r (e^{i\theta} + e^{-i\theta})$, and thereby also the induced charge density, is dipolar, i.e. $l = \pm 1$. The dipole moment $\mathbf{p}(\omega)$ of the disk is just (restoring explicit frequency dependence for clarity)

$$\mathbf{p}(\omega) = \int_{r < R} d^2 \mathbf{r}_{\parallel} \mathbf{r}_{\parallel} \rho_{\parallel}(\mathbf{r}_{\parallel}, \omega) = \hat{\mathbf{x}} 2\pi R^3 \int_0^1 d\tilde{r} \tilde{r}^2 \rho_{\parallel}(\tilde{r}, \omega), \quad (\text{S38})$$

where we have used that the radial charge density, $\rho_{\parallel}(\tilde{r}, \omega)$, is identical for $l = 1$ and $l = -1$. Using the polynomial expansion of Eq. (S30) then yields a simple connection between $\mathbf{p}(\omega) = p(\omega) \hat{\mathbf{x}}$ and the expansion coefficients $\mathbf{c}(\omega)$ [S6, 7.391.3]:

$$p(\omega) = 2\pi R^3 \sum_{j=0}^{\infty} c_j(\omega) \underbrace{\int_0^1 \tilde{r}^3 P_j^{1,0}(1 - 2\tilde{r}^2)}_{= \delta_{j0}/4} = \frac{1}{2} \pi R^3 c_0(\omega), \quad (\text{S39})$$

with $c_0(\omega)$ obtained via Eq. (S37) with $d_j = -\frac{1}{2} E_0 R \delta_{j0}$ and $l = \pm 1$.

Finally, we note that the quasistatic polarizability relates to the dipole moment via $\alpha_0(\omega) = p(\omega)/\epsilon_0 \epsilon_B E_0$, and that the absorption cross-section, neglecting retardation-corrections, relates to the quasistatic polarizability via $\sigma_{\text{abs}}(\omega) = \frac{\omega}{c} \text{Im}[\alpha_0(\omega)]$.

VIII. ELECTROSTATIC INTERACTION IN DIRAC-RESPONSE NANODISK

The starting point for our treatment is a self-consistent equation for the induced density. In the RPA this self-consistent equation is derived by coupling the *total* potential, with contributions from both external and induced potentials $\phi = \phi^{\text{ext}} + \phi^{\text{ind}}$, and the induced charge density ρ through the non-interacting polarizability and a Hartree interaction:

$$\rho(\mathbf{r}) = e^2 \int d\mathbf{r}' \chi^0(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}') \quad (\text{S40a})$$

$$\phi(\mathbf{r}) = \phi^{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}' V(\mathbf{r}, \mathbf{r}') \rho(\mathbf{r}'), \quad (\text{S40b})$$

with $V(\mathbf{r}, \mathbf{r}') = 1/4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|$ denoting the Coulomb interaction. Unlike in Section VII, all coordinates and functions refer to surface quantities, since the explicit accounting for the z -dimension is unnecessary. Combining these equations one readily finds integral equations for either the potential or the induced charge density. For the induced charge density, as considered also in our Communication, one finds:

$$\rho(\mathbf{r}) = e^2 \int d\mathbf{r}' \chi^0(\mathbf{r}, \mathbf{r}') \left[\phi^{\text{ext}}(\mathbf{r}') + \int d\mathbf{r}'' V(\mathbf{r}', \mathbf{r}'') \rho(\mathbf{r}'') \right]. \quad (\text{S41})$$

By expanding all quantities in angular momenta,

$$\rho(r, \theta) = \sum_{m=-\infty}^{\infty} \rho_m(r) e^{im\theta}, \quad \phi^{\text{ext}}(r, \theta) = \sum_{m=-\infty}^{\infty} \phi_m^{\text{ext}}(r) e^{im\theta}, \quad V(r, r'; \theta - \theta') = \sum_{m=-\infty}^{\infty} V_m(r, r') e^{im(\theta - \theta')}, \quad (\text{S42})$$

with $\chi^0(\mathbf{r}, \mathbf{r}')$ expanded already in Eq. (S10), we can decouple Eq. (S41) into separate equations for each angular momentum:

$$\rho_m(\tilde{r}) = 2\pi e^2 R^2 \int_0^1 d\tilde{r}' \tilde{r}' \chi_m^0(\tilde{r}, \tilde{r}') \left[\phi_m^{\text{ext}}(\tilde{r}') + 2\pi R^2 \int_0^1 d\tilde{r}'' \tilde{r}'' V_m(\tilde{r}', \tilde{r}'') \rho_m(\tilde{r}'') \right], \quad (\text{S43})$$

expressed again in dimensionless coordinates $\tilde{r} = r/R$. The Coulomb angular elements can be read off by comparison with Eq. (S26) yielding $V_m(\tilde{r}, \tilde{r}') = K_m(\tilde{r}, \tilde{r}')/4\pi\epsilon_0\epsilon_b R$.

The solution of this equation can be attempted e.g. via polynomial expansion - but analytical expressions for the matrix elements cannot be attained. As a consequence it is desirable to pursue a simple discretization-solution: we discretize \tilde{r} by the set $\{\tilde{r}_n\}_{n=1}^{N_d} \in [0, 1]$ and find good convergence for $N_d = 100$ in all considered cases, including in our computations also the weakly singular diagonal elements.

A. Computational complexity and comparison with RPA at tight-binding level

The primary computational hurdle in applying RPA, both in the Dirac and the TB approaches, involves computation of $\chi^0(\mathbf{r}, \mathbf{r}')$ on the set of all relevant positions $\{\mathbf{r}, \mathbf{r}'\}$. We compare the complexity below:

Tight binding – The set $\{\mathbf{r}, \mathbf{r}'\}$ is predetermined as all carbon-atom locations of which there are N , such that the matrix representation of χ^0 has dimensions $N \times N$. Computation of the density-response at each point, i.e. of $\chi^0(\mathbf{r}_n, \mathbf{r}_{n'})$, requires $\mathcal{O}(N^2)$ operations [cf. the double sum $\sum_{\nu\nu'}$ in Eq. (5) of our Communication]. As such, direct construction of χ^0 in TB requires $\mathcal{O}(N^4)$ operations – which, however, can be reduced to $\mathcal{O}(N^3)$ operations by using the fast Fourier transform following the scheme suggested in Ref. [S1].

Dirac equation – As discussed in the previous section, the electrostatic problem can be decoupled into multipolar components χ_m^0 (cf. the continuum-assumption of the Dirac equation) and the radial coordinates (r, r') can subsequently be discretized e.g. on a regular grid with N_d points. Since only a single multipolar component, namely the $m = 1$ ($m = -1$) component is necessary for studying interaction with plane-waves, the dimensions of the matrix representation of χ^0 are just $N_d \times N_d$. The evaluation of $\chi_0^0(r_n, r_{n'})$ requires $\mathcal{O}(N^\beta)$ operations, with $1 < \beta < 2$ [cf. the reduction of state-summations in Eq. (5) of our Communication due to the selection rules discussed in Section III], with N here indicating the number of Dirac-states in the considered energy range (similar scaling as in TB). Thus, direct evaluation of χ_m^0 requires $\mathcal{O}(N_d^2 N^\beta)$.

Self-consistent electrostatic problem – In both cases, the self-consistent problem finally requires solving a matrix equation ($N \times N$ in TB and $N_d \times N_d$ in Dirac) scaling with the third power of the dimensionality by direct Gauss-Jordan elimination.

Since N is on the order of several thousands, while we use $N_d = 100$ this illustrates that application of RPA@Dirac requires significantly less computational effort compared with RPA@TB.

Of course, solving the electrostatic problem with a local conductivity $\sigma(\omega)$ is vastly simpler complexity-wise compared to RPA at any level – in particular, that problem is scale-invariant (assuming a scale-invariant conductivity). Furthermore, applying the semi-analytical solution for nanodisks discussed in Section VII requires only solving a $J \times J$ matrix equation with $J = 250$ much more than sufficient.

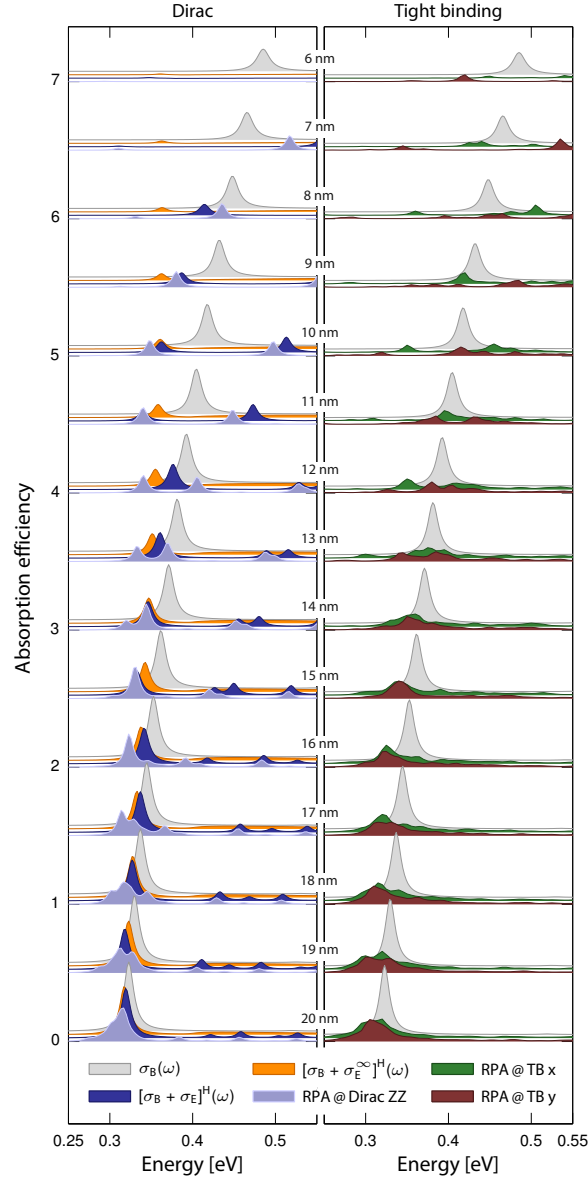


FIG. S5 Identical setup as in Fig. 4 of our Communication, but with an extended range of diameters considered. As in Fig. 4, disk-diameters are indicated above each spectra, with different diameters offset by 0.5, while individual spectra at identical diameters are offset by 0.025.

IX. ABSORPTION SPECTRA AT INTERMEDIATE AND SMALLER DIAMETERS

In support of Fig. 4 of our Communication, we here offer additional data in Fig. S5, adding to Fig. 4 absorption spectra at intermediate diameters, as well as considering also smaller diameters. The spectra at intermediate diameters follow the trends also observed in Fig. 4. At very small diameters, predictions of the spectral position of resonances in both local and Dirac approaches are unable to match those of TB – and the polarization-dependence becomes even more pronounced – highlighting the necessity for atomistic treatments in this size-range. Nevertheless, one feature qualitatively reproduced by both edge-corrected LR and Dirac approaches at very small diameters is a dramatically reduced absorption efficiency, not captured

by a bulk LR approach. This reduction sets in as the resonances move into the region of edge-to-bulk EHP transitions. In fact, the overall reduction is slightly overestimated in Dirac and edge-corrected LR since the Dirac approach overestimates the energy-level spacing due to assumed azimuthal and valley symmetry which is not present in a TB treatment.

References

- [S1] S. Thongrattanasiri, A. Manjavacas, and F.J. García de Abajo, *ACS Nano* **6**, 1766 (2012).
- [S2] A.R. Akhmerov and C.W.J. Beenakker, *Phys. Rev. B* **77**, 085423 (2008).
- [S3] A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, and A.K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009).
- [S4] P. Tassin, T. Koschny, and C.M. Soukoulis, *Science* **341**, 620 (2013).
- [S5] M. Grujić, M. Zarenia, A. Chaves, M. Tadić, G.A. Farias, and F.M. Peeters, *Phys. Rev. B* **84**, 205441 (2011).
- [S6] I.S. Gradshteyn and I.M. Ryzhik, *Table of integrals, series, and products*, 7th ed. (Academic Press, 2007).
- [S7] B. Wunsch, T. Stauber, and F. Guinea, *Phys. Rev. B* **77**, 035316 (2008).
- [S8] M. Wimmer, A.R. Akhmerov, and F. Guinea, *Phys. Rev. B* **82**, 045409 (2010).
- [S9] National Institute of Standards and Technology, “[Digital library of mathematical functions](#),” Release 1.0.8 of 2014-04-25.
- [S10] L.A. Falkovsky and A.A. Varlamov, *Eur. Phys. J. B* **56**, 281 (2007).
- [S11] V.P. Gusynin, S.G. Sharapov, and J.P. Carbotte, *J. Phys.: Condens. Matter* **19**, 026222 (2007).
- [S12] G.W. Hanson, *J. Appl. Phys.* **104**, 084314 (2008).
- [S13] W. Wang and J.M. Kinaret, *Phys. Rev. B* **87**, 195424 (2013).
- [S14] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, *New J. Phys.* **8**, 318 (2006).
- [S15] E.H. Hwang and S. Das Sarma, *Phys. Rev. B* **75**, 205418 (2007).
- [S16] P. Halevi, *Phys. Rev. B* **51**, 7497 (1995).
- [S17] S. Raza, G. Toscano, A.-P. Jauho, M. Wubs, and N.A. Mortensen, *Phys. Rev. B* **84**, 121412(R) (2011).
- [S18] A.L. Fetter, *Phys. Rev. B* **33**, 5221 (1986).
- [S19] A.L. Fetter and J.D. Walecka, *Theoretical Mechanics of Particles and Continua* (Dover Publications, 1980).

ACS NANO 8, 1745 (2014)

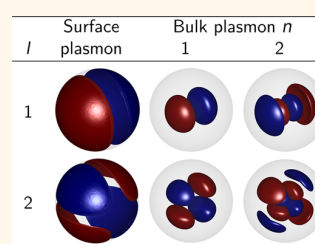
The topic was initially conceived by T. Christensen, S. Raza, N.A. Mortensen, and M. Wubs. W. Yan and S. Raza introduced T. Christensen to Mie-Lorenz theory. T. Christensen implemented and performed all calculations. T. Christensen and M. Wubs co-wrote the manuscript with input from all authors.

Nonlocal Response of Metallic Nanospheres Probed by Light, Electrons, and Atoms

Thomas Christensen,^{†,‡} Wei Yan,^{†,‡} Søren Raza,^{†,§} Antti-Pekka Jauho,^{†,⊥} N. Asger Mortensen,^{†,‡} and Martijn Wubs^{†,‡,*}

[†]Department of Photonics Engineering, [‡]Center for Nanostructured Graphene, [§]Center for Electron Nanoscopy, and [⊥]Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

ABSTRACT Inspired by recent measurements on individual metallic nanospheres that cannot be explained with traditional classical electrodynamics, we theoretically investigate the effects of nonlocal response by metallic nanospheres in three distinct settings: atomic spontaneous emission, electron energy loss spectroscopy, and light scattering. These constitute two near-field and one far-field measurements, with zero-, one-, and two-dimensional excitation sources, respectively. We search for the clearest signatures of hydrodynamic pressure waves in nanospheres. We employ a linearized hydrodynamic model, and Mie–Lorenz theory is applied for each case. Nonlocal response shows its mark in all three configurations, but for the two near-field measurements, we predict especially pronounced nonlocal effects that are not exhibited in far-field measurements. Associated with every multipole order is not only a single blueshifted surface plasmon but also an infinite series of bulk plasmons that have no counterpart in a local-response approximation. We show that these increasingly blueshifted multipole plasmons become spectrally more prominent at shorter probe-to-surface separations and for decreasing nanosphere radii. For selected metals, we predict hydrodynamic multipolar plasmons to be measurable on single nanospheres.



KEYWORDS: nonlocal response · nanoplasmonics · EELS · extinction · LDOS · spontaneous emission · multipole plasmons

A plethora of effects arises in structured metals due to collective excitations of conduction electrons and their interaction with the electromagnetic field. This constitutes plasmonics, a research field with mature roots^{1,2} that is continuing to develop strongly.³ Notably, applications for plasmonics are found in the biochemistry and biomedical fields, such as in surface-enhanced Raman spectroscopy (SERS),⁴ biosensing⁵ and biomedical imaging,⁶ drug delivery,⁷ and phototherapy of cancer cells.⁸ Purely photonic applications are also emerging, for example, in plasmonic waveguiding,⁹ optical nano-antennas,^{10,11} and photovoltaics.¹²

Recent years' advances in fabrication, synthesis, and characterization techniques have allowed well-controlled experimental investigations of plasmonics even at the nanoscale. Yet in this growing field of nanoplasmonics,^{3,13} the commonly employed theory for light–matter interaction is still traditional classical electrodynamics, where the response of the material constituents to light is described collectively in terms of local,

bulk material response functions. Indeed, this approach usually remains very accurate, even for subwavelength phenomena.

Interestingly, recent measurements on individual few-nanometer plasmonic particles have shown phenomena that are clearly beyond classical electrodynamics. Electron energy loss spectroscopy (EELS) of Ag spheres resting on dielectric substrates showed surface plasmon resonance blueshifts up to 0.5 eV as compared to classical theory.^{14,15} Earlier similar measurements were performed on ensembles of nanoparticles.¹⁶ Classical electrodynamics was also shown to fail in experiments involving (sub)nanometer-sized gaps between dimers^{17–19} or between nanoparticles and a substrate.²⁰

To explain these features arising beyond the validity of classical electrodynamics, various physical mechanisms are invoked. First, classical electrodynamics assumes a step-function profile of the free-electron density at a metal–dielectric interface. The finite quantum mechanical spill-out²¹ of the

* Address correspondence to mwubs@fotonik.dtu.dk.

Received for review November 29, 2013 and accepted January 17, 2014.

Published online January 17, 2014
10.1021/nn406153k

© 2014 American Chemical Society

electron density redshifts the surface plasmon resonance,^{22,23} may give rise to nonresonant field enhancement,²⁴ and may enable charge transfer between nontouching plasmonic dimers.^{18,19,25} Second, a stronger confinement of the free electrons gives rise to blueshifts. In cluster physics, it is single-particle excitations that are blueshifted due to quantum confinement,²⁶ while confinement in nanoplasmonics blueshifts collective resonances and gives rise to Friedel oscillations in the electron density.^{27,28} A third, semiclassical physical mechanism beyond classical electrodynamics is nonlocal response, discussed in more detail below, which becomes important when reducing the particle size or gap size of a dimer down to the range of the nonlocality²⁹ (ξ_{NL} denoting the spatial extent of significant nonlocal interaction, to be introduced shortly), and blueshifts surface plasmon resonance frequencies.

Large experimental blueshifts of the localized surface plasmon (LSP) dipole resonance seem to indicate that several physical mechanisms add up.^{15,30} Certainly, in experiments, all of these physical mechanisms beyond traditional classical electrodynamics are at work simultaneously, thus emphasizing the importance of microscopic theories³¹ (e.g., density functional theory, DFT) or effective models³² that incorporate multiple mechanisms. Yet at the same time, it is important to ascertain the relative strength and compatibility of the various mechanisms. Indeed, it is paramount to know—and to measure—the unique characteristics of each mechanism, that is to say, find their individual “smoking guns”, in order to appreciate the dominant physical mechanisms under different nonstandard circumstances. We foresee an increasing number of such decisive experiments on individual nanoparticles in the near future.

The boundary between cluster physics and nanoplasmonics is an interesting one. Metal clusters require a quantum description of interacting electron states, often studied with DFT. In contrast, nanoplasmonics could be defined to start for nanoparticle sizes that allow an effective quantum description in terms of noninteracting plasmons.²⁸ A current interesting issue is where to place the origin of the observed blueshift of the surface plasmon resonance of individual nanospheres: is it primarily due to quantum confinement of single-particle states,^{14,16} or due to confinement of collective modes?^{15,27,28,32} In this article, we assume the latter and identify new observable consequences. We focus on nanoparticles that are considered large enough ($2R \geq 3$ nm) that so-called core plasmons, although collective in nature, can be neglected according to DFT calculations.²⁸

Nonlocal response is a semiclassical effect which emerges in nanoplasmonics at few-nanometer length scales. The general nonlocal relation between the displacement and electric fields, $\mathbf{D}(\mathbf{r},\omega) = \varepsilon_0 \int \varepsilon(\mathbf{r},\mathbf{r}';\omega) \mathbf{E}(\mathbf{r}',\omega) d\mathbf{r}'$,

becomes simpler and more familiar in the local-response approximation (LRA), that is, $\varepsilon(\mathbf{r},\mathbf{r}';\omega) \approx \varepsilon_{\text{LRA}}(\mathbf{r},\omega) \delta(\mathbf{r} - \mathbf{r}')$. In many cases, this approximation provides an excellent effective description due to the short-range nature of the nonlocal interaction. However, the LRA is not justifiable when the nonlocal interaction length, ξ_{NL} , becomes comparable with characteristic feature sizes of structural or optical kind.²⁹

Here we consider inclusion of the classically neglected Fermi–Dirac pressure of the electron gas. Its associated pressure waves give rise to a nonlocal optical response. The simplest way to study the effects of Fermi pressure in nanoplasmonics is by assuming a hydrodynamic model,^{15,33–39} which neglects the aforementioned spill-out and confinement effects on the static electron density. In hydrodynamics, the nonlocal interaction length becomes $\xi_{\text{NL}} = v_{\text{F}}/\omega$, with $v_{\text{F}} = \hbar\sqrt{3\pi^2 n_0}/m$ denoting the Fermi velocity, defined through the effective mass m and free-electron density n_0 . This corresponds to ξ_{NL} -values in the range of 2–5 Å for typical plasmonic metals at optical frequencies; see Table S1 in the Supporting Information (SI). We will focus on the linearized hydrodynamic model here but would like to emphasize that the full hydrodynamic model involves both nonlocality and nonlinearity, predicting nonlinear effects such as second-harmonic generation at the surface of metal nanoparticles for larger field strengths.^{40–42}

The strongest evidence of hydrodynamic behavior in metals originates from experiments on thin metal films, where resonances due to standing waves of confined bulk plasmons have been identified, in silver by Lindau and Nilsson,⁴³ in potassium by Anderegg *et al.*,⁴⁴ in magnesium by Chen,⁴⁵ and very recently by Özer *et al.*⁴⁶ Rather surprisingly, Özer *et al.*⁴⁶ could measure confined bulk plasmon resonances (*i.e.*, standing Fermi pressure waves) even for ultrathin magnesium films of only three atomic monolayers and found qualitative agreement with theory even when neglecting electronic spill-out. For nanospheres on the other hand, the observations of blueshifted dipole resonances of localized surface plasmons (LSPs) in individual nanospheres^{14,15,30} and of broad resonance features above the plasma frequency in ensembles,⁴⁷ tentatively suggested as associated with confined bulk plasmons,³³ are perhaps less conclusive evidence of hydrodynamic behavior. This may in part be due to a line of reasoning which addresses just a single resonance, namely, the dipole.

Our aim in this article is then to examine theoretically which phenomena constitute the clearest evidence of hydrodynamic pressure waves in plasmonic nanospheres and how best to observe them. Powerful measurement techniques include scattering measurements, as realized for example, in the infrared regime by Fourier transform infrared spectroscopy (FTIR), scanning near-field optical microscopy (SNOM),⁴⁸ EELS,^{49,50}

and fluorescence microscopy techniques, utilizing decay enhancement of emitters near plasmonic resonances.^{51,52} In this theoretical article, we systematically explore three prominent measurement techniques, each with different excitation sources, namely, the extinction cross section, the EELS probability, and the electric local density of states (LDOS). The excitation sources are, respectively, a linearly polarized plane wave, a traveling electron with kinetic energy in the kiloelectronvolt range, and an electric dipole emitter, corresponding to a two-, one-, and zero-dimensional source. The three measurement principles represent both far- and near-field types, and we show their spectra to be qualitatively different.

We investigate not only the strongest (dipolar) LSP resonance of nanospheres but also higher-order multipole LSPs, as well as bulk plasmons, for all three measurements considered. We show that hydrodynamic response leads to a significant spectral separation of the sphere's multipole plasmons at small radii, allowing them to extend above the LRA asymptotic limit at $\omega_p/\sqrt{2}$. Resonance features above this limit have already been observed in polydisperse ensembles of nanospheres and previously been interpreted instead in terms of single-particle confinement.¹⁶ We find significant qualitative disparity between properties measurable in the far-field (*i.e.*, via extinction) and in the near-field (*i.e.*, via EELS or LDOS). Our findings result in concrete suggestions to experimentally observe hydrodynamic nonlocal phenomena in the near-field, by identifying the multipolar plasmon resonances of individual nanospheres of selected metals.

RESULTS AND DISCUSSION

Theoretical Framework. In a linearized hydrodynamic description, the current density $\mathbf{J}(\mathbf{r}, \omega)$ and the electric field $\mathbf{E}(\mathbf{r}, \omega)$ are inter-related by the nonlocal relation:^{53,54}

$$\frac{\beta_F^2}{\omega(\omega + i\eta)} \nabla[\nabla \cdot \mathbf{J}(\mathbf{r}, \omega)] + \mathbf{J}(\mathbf{r}, \omega) = \sigma(\omega)\mathbf{E}(\mathbf{r}, \omega) \quad (1a)$$

where $\sigma(\omega) = i\varepsilon_0\omega_p^2/(\omega + i\eta)$ is the usual Drude conductivity of a free-electron gas with plasma frequency ω_p , including a phenomenological loss rate η , and $\beta_F^2 = (3/5)v_F^2$ is the hydrodynamical velocity of plasma pressure waves in the metal. The hydrodynamic model can be classified as 'semiclassical' because eq 1a relates the classical fields \mathbf{J} and \mathbf{E} via the parameter $\beta_F \propto v_F$, which is proportional to \hbar . Hydrodynamic response appears as a lowest-order spatially nonlocal correction to the local Ohm's law, with a strength proportional to $\xi_{NL}^{-2}k^2$ in momentum k -space.

In addition to eq 1a, the electric field must satisfy the Maxwell wave equation

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}, \omega) - k_0^2 \varepsilon_\infty(\omega)\mathbf{E}(\mathbf{r}, \omega) = i\omega\mu_0\mathbf{J}(\mathbf{r}, \omega) \quad (1b)$$

with $k_0 = \omega/c$ denoting the usual free-space wavenumber and $\varepsilon_\infty(\omega)$ the dielectric response of the bound

charges, that is, the response not due to the free-electron plasma. The sum of the bound- and free-electron response gives the transverse response of the metal $\varepsilon_M(\omega) = \varepsilon_\infty(\omega) + \sigma(\omega)/i\varepsilon_0\omega$, familiar from the LRA. For calculations involving a measured transverse metal response $\varepsilon_M(\omega)$, the bound response $\varepsilon_\infty(\omega)$ is determined by fixing $\omega_p = (n_0e^2/\varepsilon_0m)^{1/2}$, that is, through the free-electron density n_0 and effective mass m , thus determining the free response $\sigma(\omega)$ and allowing $\varepsilon_\infty(\omega)$ to be determined by subtraction.³⁴

The practical solution of eqs 1 in structures with curvilinear symmetries can be aided significantly by expansion in the so-called vector wave functions. Concretely, a monochromatic electromagnetic field in a region of uniform dielectric function can be expanded in the basis composed of the solenoidal, $\mathbf{M}_\nu(\mathbf{r})$ and $\mathbf{N}_\nu(\mathbf{r})$, and irrotational, $\mathbf{L}_\nu(\mathbf{r})$, vector wave functions:^{55,56}

$$\mathbf{E}(\mathbf{r}) = \sum_\nu a_\nu \mathbf{M}_\nu(\mathbf{r}) + b_\nu \mathbf{N}_\nu(\mathbf{r}) + c_\nu \mathbf{L}_\nu(\mathbf{r}) \quad (2)$$

where ν denotes a composite expansion index with a_ν , b_ν , and c_ν being associated expansion coefficients. The functions $\mathbf{M}_\nu(\mathbf{r})$ and $\mathbf{N}_\nu(\mathbf{r})$ describe the TE and TM parts, respectively, of the electric field and describe the propagation of transverse, or divergence-free, modes.⁵⁶ The functions $\mathbf{L}_\nu(\mathbf{r})$ are irrotational and, as such, are irrelevant in media described by the LRA. However, their inclusion is indispensable for the treatment of plasmonic nanoparticles by hydrodynamic response, in order to account for the inclusion of longitudinal modes.

Next, we consider the case of an arbitrary external exciting field \mathbf{E}^{ex} that originates in an outer dielectric region and scatters upon a spherical metallic particle of radius R that is centered at the origin. This induces scattered fields \mathbf{E}^{sc} outside the particle and transmitted fields \mathbf{E}^{tr} inside (see Figure 1). For spherical nanoparticles, the choice of multipolar vector wave functions separates the composite expansion index ν into the angular momentum quantum numbers l and m ; for details, see the Methods section.

Outside the nanosphere ($r > R$), the fields \mathbf{E}^{ex} and \mathbf{E}^{sc} can be expanded solely in terms of the in- and outgoing transverse multipoles $\{\mathbf{M}_{lm}^{\text{ex}}, \mathbf{N}_{lm}^{\text{ex}}\}$ and $\{\mathbf{M}_{lm}^{\text{sc}}, \mathbf{N}_{lm}^{\text{sc}}\}$, respectively, since the dielectric region does not support longitudinal waves. The corresponding expansion coefficients are $\{a_{lm}^{\text{ex}}, b_{lm}^{\text{ex}}\}$ and $\{a_{lm}^{\text{sc}}, b_{lm}^{\text{sc}}\}$. The transmitted field \mathbf{E}^{tr} inside the nanosphere ($r < R$) requires ingoing transverse multipoles, $\{\mathbf{M}_{lm}^{\text{tr}}, \mathbf{N}_{lm}^{\text{tr}}\}$, and also ingoing longitudinal modes, $\mathbf{L}_{lm}^{\text{tr}}$, which correspondingly necessitates three sets of expansion coefficients $\{a_{lm}^{\text{tr}}, b_{lm}^{\text{tr}}, c_{lm}^{\text{tr}}\}$.

The fields inside and outside the nanosphere are related by boundary conditions (BCs); see the Methods section. This translates into linear relations between the expansion coefficients of the exciting and

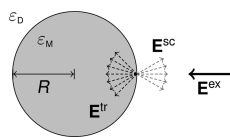


Figure 1. Sketch of an exciting wave E^{ex} interacting with a metallic sphere embedded in a dielectric background, giving rise to scattered and transmitted fields E^{sc} and E^{tr} , respectively.

scattered fields^{56,57}

$$a_{lm}^{\text{sc}} = t_l^{\text{TE}} a_{lm}^{\text{ex}} \delta_{ll'} \delta_{mm'}, \quad b_{lm}^{\text{sc}} = t_l^{\text{TM}} b_{lm}^{\text{ex}} \delta_{ll'} \delta_{mm'} \quad (3)$$

where δ_{jk} is the Kronecker delta. The proportionality constants t_l^{TE} and t_l^{TM} are known as the Mie–Lorenz coefficients.⁵⁸ For nanospheres with nonlocal response, they are given by^{33,34}

$$t_l^{\text{TE}} = \frac{-j_l(x_M)[x_D j_l(x_D)]' + j_l(x_D)[x_M j_l(x_M)]'}{j_l(x_M)[x_D h_l^{(1)}(x_D)]' - h_l^{(1)}(x_D)[x_M j_l(x_M)]'} \quad (4a)$$

$$t_l^{\text{TM}} = \frac{-\varepsilon_M j_l(x_M)[x_D j_l(x_D)]' + \varepsilon_D j_l(x_D)\{[x_M j_l(x_M)]' + \Delta_l\}}{\varepsilon_M j_l(x_M)[x_D h_l^{(1)}(x_D)]' - \varepsilon_D h_l^{(1)}(x_D)\{[x_M j_l(x_M)]' + \Delta_l\}} \quad (4b)$$

where $x_D = k_D R$ and $x_M = k_M R$ are dimensionless parameters in terms of the dielectric and transverse metal wavenumbers (see Methods) and the radius R of the nanosphere. The primes denote the derivatives with respect to $x_{D,M}$. As for the usual Mie–Lorenz coefficients in the LRA, these hydrodynamic Mie–Lorenz coefficients are independent of the multipole label m , due to the spherical geometry of the scatterer. Spatial nonlocality influences the Mie–Lorenz coefficients through the hydrodynamic term^{33,34}

$$\Delta_l = l(l+1)j_l(x_M) \frac{\varepsilon_M - \varepsilon_\infty}{\varepsilon_\infty} \frac{j_l(x_{\text{NL}})}{x_{\text{NL}} j_l'(x_{\text{NL}})} \quad (4c)$$

with $x_{\text{NL}} = k_{\text{NL}} R$ introducing the longitudinal metal wavenumber (see Methods). As expected, the correction Δ_l vanishes in the LRA limit since $|x_{\text{NL}}| \rightarrow \infty$ as $\beta_F \rightarrow 0$. Note that only the scattering of TM waves is affected by the inclusion of spatial nonlocality. There are no contributions to the magnetic field from the longitudinal multipoles \mathbf{L}_m^{tr} (cf. the Maxwell–Faraday equation), thus leaving the TE waves, sometimes called the magnetic waves, unaffected.

The significance of the Mie–Lorenz coefficients is that they specify the scattering laws outside the sphere; that is, they determine the outcome of external measurements. In particular, a general linear measurement \mathcal{O} on a nanosphere can be expressed as a linear combination of them. As discussed in more detail below, all three measurements that we consider can be expressed in the general form

$$\mathcal{O} = \sum_{lm} \mathcal{O}_{lm}^{\text{TE}} \text{Re}(t_l^{\text{TE}}) + \mathcal{O}_{lm}^{\text{TM}} \text{Re}(t_l^{\text{TM}}) \quad (5)$$

where the coefficients $\mathcal{O}_{lm}^{\text{TE,TM}}$ contain all information regarding the measurement (e.g., type and position), while

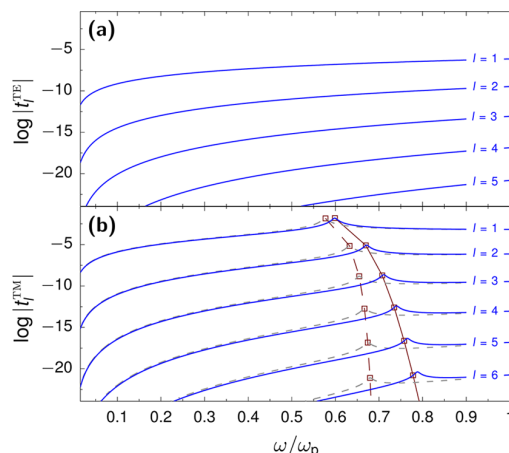


Figure 2. Absolute value of the Mie–Lorenz coefficients (a) t_l^{TE} and (b) t_l^{TM} on a logarithmic scale as a function of frequency for the first few values of l . Considered is a $R = 2.5$ nm sphere with Drude metal parameters $\omega_p = 10$ eV, $\eta = 0.1$ eV, and $\varepsilon_\infty = 1$ embedded in vacuum, $\varepsilon_D = 1$. For comparison, the LRA TM Mie–Lorenz coefficients are illustrated as gray dashed lines. Approximate resonance predictions for LRA and hydrodynamics, as predicted by eqs 7 and 8, are given as dashed and full red lines, respectively.

$t_l^{\text{TE,TM}}$ contain all information regarding the scattering geometry (e.g., dielectric composition and size). Crucially, the inclusion of hydrodynamic nonlocality modifies only the Mie–Lorenz coefficients t_l^{TM} but not the measurement coefficients $\mathcal{O}_{lm}^{\text{TE,TM}}$.

For this reason, we can first focus on the Mie–Lorenz coefficients and look for the local and nonlocal plasmonic resonances that in principle affect all measurements. After that, we will identify the measurements in which these resonances make a prominent appearance and where the impact of hydrodynamic dispersion is especially strong.

Multipole Plasmon Resonances. Figure 2 depicts the frequency dependence of the first few Mie–Lorenz coefficients $t_l^{\text{TE,TM}}$ of a free-electron $R = 2.5$ nm nanosphere. Clearly, large l multipoles in general scatter significantly weaker than small l multipoles (notice the log scale). In addition, the t_l^{TM} coefficients exhibit a series of resonances, corresponding to poles of the coefficient, associated with excitation of LSPs of dipole, quadrupole, hexapole (and so on) character, for $l = 1, 2, 3, \dots$, respectively. By contrast, the t_l^{TE} coefficients exhibit no such resonances. Moreover, they are several orders of magnitude smaller than their equal momenta TM correspondents. As a result, the TM interaction dominates the response of plasmonic nanospheres. It is this dominant TM interaction which is modified by nonlocal response.

Surface Plasmon Resonance Conditions. A trademark of hydrodynamic response is its blueshift of resonances as compared to local response. Figure 2 illustrates that for nanospheres these blueshifts show up in the TM Mie–Lorenz coefficients and are

increasingly shifted for larger l .⁵⁹ We study this quantitatively and find the multipole plasmon resonances of order l from the pole of the t_l^{TM} coefficient. The non-retarded limit can be applied to the small spheres under consideration, leading to the plasmon condition⁶⁰

$$l\varepsilon_M + (l+1)(1+\delta_l)\varepsilon_D = 0 \quad (6)$$

where $\delta_l = \Delta_l/[j_l(x_{\text{NL}})(l+1)]$ accounts for the hydrodynamic correction; see SI for additional details. (A similar multipole plasmon condition was derived in ref 60 for metallic spheres in vacuum, but with a missing factor of i/x_{NL} in their equivalent definition of δ_l .) Evidently, non-locality can be interpreted as modifying the dielectric surrounding, by introducing an effective l -dependent dielectric constant $\varepsilon_{lD}^{\text{eff}} = (1+\delta_l)\varepsilon_D$. Since δ_l is a function of frequency and angular momentum, eq 6 defines plasmon resonances only implicitly. Nevertheless, their spectral location can be determined by approximation while retaining the essential physics, as we shall show below.

In the LRA limit $\delta_l \rightarrow 0$ and upon neglecting dispersion of the bound response and damping, that is, taking $\varepsilon_M(\omega) = \varepsilon_\infty - \omega_p^2/\omega^2$, the well-known local electrostatic plasmon resonances are immediately recovered from eq 6 as

$$\omega_l^{\text{L}} = \frac{\omega_p}{\sqrt{\varepsilon_\infty + \frac{l+1}{l}\varepsilon_D}} \quad (7)$$

Thus, in local theory, for $l=1$, we find the well-known (dipolar) LSP resonance $\omega_1^{\text{L}} = \omega_p/(\varepsilon_\infty + 2\varepsilon_D)^{1/2}$, which reduces to $\omega_p/\sqrt{3}$ for a free Drude metal sphere in vacuum. The high-order multipole plasmons tend asymptotically from below toward the local planar interface surface plasmon $\omega_p/(\varepsilon_\infty + \varepsilon_D)^{1/2}$ for $l \rightarrow \infty$, reducing to $\omega_p/\sqrt{2}$ for a free Drude metal sphere in vacuum. The l -dependence of ω_l^{L} as described by eq 7 is depicted by the red dashed line in Figure 2, clearly showing the asymptotic behavior for large l .

Turning now from local to nonlocal response, let us assume that δ_l in eq 6 is a small perturbation, which is valid for small l and for $R \gg \beta_F/\omega_p$. We circumvent the implicitness of the resonance condition by making a pole approximation, replacing the dispersive function $\delta_l(\omega)$ by its value $\delta_l^{\text{L}} = \delta_l(\omega_l^{\text{L}})$ in the local resonance frequency ω_l^{L} , the latter given by eq 7. The hydrodynamically corrected resonances ω_l^{NL} then occur at approximately⁵⁹

$$\omega_l^{\text{NL}} \approx \frac{\omega_p}{\sqrt{\varepsilon_\infty + \frac{l+1}{l}(1+\delta_l^{\text{L}})\varepsilon_D}} \approx \omega_l^{\text{L}} + \frac{\beta_F}{R} \sqrt{\frac{l(l+1)\varepsilon_D}{4\varepsilon_\infty}} \quad (8)$$

where, at the last step, in addition to a Taylor expansion of the square-root term, we have utilized the large imaginary x_{NL} limit of the hydrodynamic correction, $\delta_l = l[(\varepsilon_M - \varepsilon_\infty)/\varepsilon_\infty]i/x_{\text{NL}}$, which is applicable at

frequencies below the screened plasma frequency $\omega_p^\infty \equiv \omega_p/\varepsilon_\infty$. These approximate nonlocal surface plasmon resonance frequencies are illustrated by the solid red line in Figure 2. The approximation captures the exact nonlocal blueshift well but is less accurate for larger l , as expected. By implication of these nonlocal blueshifts, excitations appear between the LRA $l = \infty$ mode (the planar surface plasmon) and the volume plasmon at ω_p , classically a resonance-free frequency interval.¹⁶

Bulk Plasmon Resonance Condition. Besides blueshifting the multipolar LSP resonances that already exist in the LRA, hydrodynamical theory also predicts the appearance of additional resonances due to confined bulk plasmons for which no LRA counterparts exist.^{33,54} More microscopic theories have also predicted the emergence of such bulk plasmons.^{28,31} These bulk plasmons emerge due to the presence of propagating longitudinal pressure waves above the plasma frequency. In hydrodynamics, the confined bulk plasmons are then easily interpreted as the standing wave resonances of longitudinal waves. Table 1 depicts isosurfaces of the induced charge density for LSPs and bulk plasmons for comparison.

An approximation for these bulk resonances can be found by neglecting the coupling of the pressure waves to light, that is, by searching for standing wave solutions of $\mathbf{L}_{lm}^{\text{tr}}$, thus neglecting the transverse components. For nanospheres, this gives radially quantized confined bulk plasmons resonating at the frequencies ω_n^{bulk} (see SI for details):

$$\omega_n^{\text{bulk}}(\omega_n^{\text{bulk}} + i\eta) = \frac{\omega_p^2}{\varepsilon_\infty} + w_{ln}^2 \left(\frac{\beta_F}{R}\right)^2 \quad (9)$$

where w_{ln} is the n th positive root of $j_l'(w)$, the derivative of the l th-order spherical Bessel function (see refs 33 and 61 for lengthier, more accurate approximations). Modes associated with the first root at $n=0$ are in fact not resonant but are artifacts of the approximation that arise due to having neglected the transverse field components. Regardless, for every multipole order l , there is an infinite number of confined bulk plasmons associated with $n=1, 2, \text{etc.}$

As for the LSP resonances, we first illustrate the signature of these bulk plasmons in the Mie–Lorenz coefficients before considering the experiments in which their presence is most pronounced. In Figure 3, we depict the frequency dependence of the first few Mie–Lorenz transmission coefficients q_l^{L} near and above ω_p . These coefficients give the transmission amplitude to a longitudinal mode due to excitation by an incident TM mode and are defined analogously to the scattering coefficients $t_l^{\text{TE, TM}}$ of eq 3 through $c_{lm}^{\text{tr}} = q_l^{\text{L}} b_{lm}^{\text{ex}} \delta_{ll'} \delta_{mm'}$ (see SI for their explicit form). The first dipolar and quadrupolar bulk plasmon resonances of a nanosphere clearly show up as Lorentzian resonances, and the bulk plasmon approximation (eq 9) is quite accurate.

TABLE 1. Charge Densities of Multipole Surface and Bulk Plasmons^a

<i>l</i>	Surface plasmon	Bulk plasmon quantization number <i>n</i>		
		1	2	3
1				
2				
3				

^a Isosurfaces are drawn for the real part of the charge density, calculated in a hydrodynamic treatment, at isovalues equal to plus/minus (red/blue) twice the mean of the absolute value of the charge density in the sphere. The nanosphere outline is indicated in shaded gray.

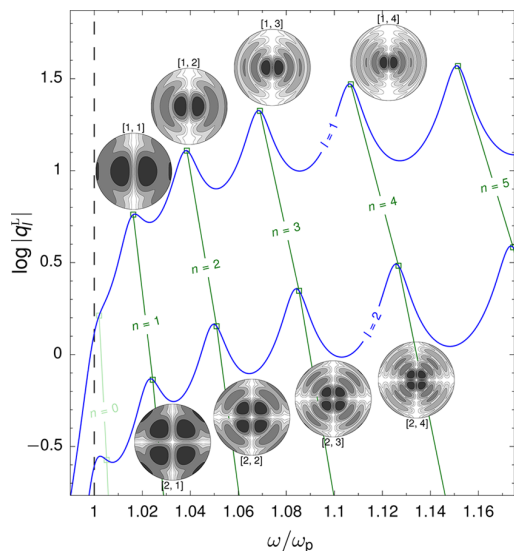


Figure 3. Absolute value of the Mie–Lorenz transmission coefficients q_l^T on a logarithmic scale as a function of frequency. The coefficients give the coupling amplitude between transmitted longitudinal multipoles and incident TM multipoles. Setup parameters are identical to those in Figure 2. Shown are the dipolar, q_1^T , and the quadrupolar, q_2^T , coefficients in blue. Both exhibit peaks above ω_p , corresponding to a series of confined bulk plasmons labeled by $n = 0, 1, 2, \dots$. Green curves show approximate resonance positions; see eq 9. The absence of an $n = 0$ resonance is apparent. Insets depict logarithmic scale contour plots, with contours separated by factors of 2, of the absolute value of the induced charge density of the bulk resonances, with $[l, n]$ indices labeled, in the xz plane.

The resonant charge distributions in the insets illustrate the radial quantization of the confined bulk plasmons. To the best of our knowledge, only the dipole ($l = 1$) confined bulk plasmons have been considered previously, for example, in relation with extinction features above the plasma frequency in nanospheres.^{33,54} In our investigation

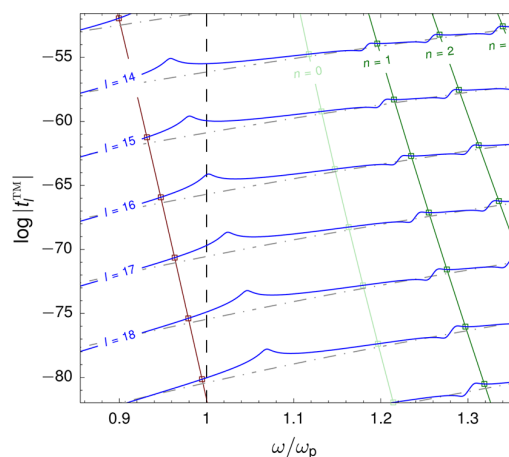


Figure 4. Absolute value of the TM Mie–Lorenz coefficients, t_l^{TM} , on a logarithmic scale, as a function of frequency for high-angular momenta. Setup parameters are identical to those in Figure 2. Hydrodynamic results are illustrated as blue solid lines, while LRA results are illustrated as gray dashed lines for comparison. The transition across the plasma frequency is marked by the black dashed line. The red line depicts the approximate LSP resonance of eq 8; the green lines show eq 9 and approximate the first few confined bulk plasmon resonances. The bulk plasmons show up as Fano-like resonances in $|t_l^{TM}|$.⁶²

of EELS and LDOS below, we consider additionally if these higher- l bulk plasmons may influence the spectral response in the near-field. First, however, we discuss the properties of higher-order LSP multipoles.

Large- l Plasmonic Resonances. We have seen in Figure 2 that multipolar hydrodynamic LSP modes blueshift away from the classical limit, the LRA planar surface plasmon at $\omega_p/\sqrt{2}$. What is more, Figure 4 illustrates that high-multipole nonlocal LSP resonances can even appear above the plasma frequency ω_p . There is no indication that the plasma frequency would mark a qualitative transition. This is despite the change from predominantly imaginary metal wavenumbers (k_M and k_{NL}) for frequencies $\omega < \omega_p^\infty$ to predominantly real metal wavenumbers for $\omega > \omega_p^\infty$. In particular, the transition from predominantly imaginary to real wavenumbers does not carry with it a transition from predominantly bound surface modes to volume-like modes as assumed in the past.³⁶ [Such a transition does not emerge since $|x_n|$ remains comparative with $(l + 1)^{1/2}$, which, cf. eqs 11 and the small-argument asymptotic form $j_l(x) \approx x^l/(2l + 1)!!$ valid for $|x| \ll (l + 1)^{1/2}$, implies that $|j_l(x)| \sim |j_l(ix)|$ for $|x| > |x_{NL}|$, whereby the charge density is left qualitatively unchanged and surface-bound.] Hydrodynamic surface plasmons above the plasma frequency have also been found theoretically for a planar metal–dielectric interface, for a thin metal slab, and for planar metamaterials.^{63,64}

It is fruitful to pursue further the analogy between the LSPs of our nanospheres and of planar structures. The analogy is well-known for local response, but the hydrodynamic version holds a surprise. The large- l LSP

resonances below and above the plasma frequency can both be characterized by wave propagation along the surface of the nanosphere. The l th surface mode accommodates exactly l oscillation periods along the periphery of the sphere. One can therefore ascribe an effective surface wavelength $\lambda_l^s = 2\pi R/l$ and an effective surface wavenumber $k_l^s = l/R$ to the l th mode. For larger l , the effective wavelength becomes shorter and the modes perceive the curving surface of the sphere as increasingly flat. For that reason, the dispersion would mimic that of a planar metal–dielectric interface for large l .

To test this prediction from the analogy, we compute the exact plasmon resonances from eq 6 and show them in a pseudodispersion plot in Figure 5. For local response, Figure 5 indeed shows the well-known result that for larger l the dispersion of the nanosphere LSPs approaches more and more that of a flat interface. For nonlocal response, also shown in Figure 5, we first note that the LSP dispersion indeed does not show a transition at the plasma frequency, as we already guessed from Figure 4. Second, there is satisfactory agreement of the hydrodynamic dispersion of LSPs for a nanosphere and for the flat interface, so the analogy is also meaningful for hydrodynamic response. However, and this is the surprising third point, unlike for local response, the agreement does *not* converge toward a complete agreement as l increases: a discrepancy develops for large l . The discrepancy is larger in Figure 5a for $R = 2.5$ nm spheres than for the twice larger spheres in Figure 5b.

This can be explained by noting that in the LRA all the induced free charge resides only *on* the surface of the sphere, whereas it is distributed *close to* this surface in the hydrodynamic description. The latter is illustrated in Figure 5c. Note the surficial standing wave quantization of the LSPs in Figure 5 and also the absence of radial quantization, being associated only with the bulk plasmons as shown in Figure 3. For large l , the neighboring hydrodynamic charge patterns in Figure 5 get squeezed into each other due to the finite curvature, producing the discrepancy with the planar interface. An alternative explanation of the discrepancy as due to interaction across antipodal surface points can be ruled out since the insets of Figure 5a,b show that the electric fields corresponding to high- l modes are well-localized near the surface of the nanosphere, even those above the plasma frequency (in contrast to predictions of ref 36), so that fields on opposite angular regions of the sphere are spatially well-separated. This agrees with recent findings for hydrodynamic LSP modes in a planar thin metal slab, which do not show finite size effects either for sufficiently large wavevectors. Rather, since the slab has no curvature, the large- k dispersion of its LSP modes does indeed agree with that of the single interface.⁶³

Extinction, EELS, and LDOS. Having discussed the characteristics of the multipole plasmons, and in particular

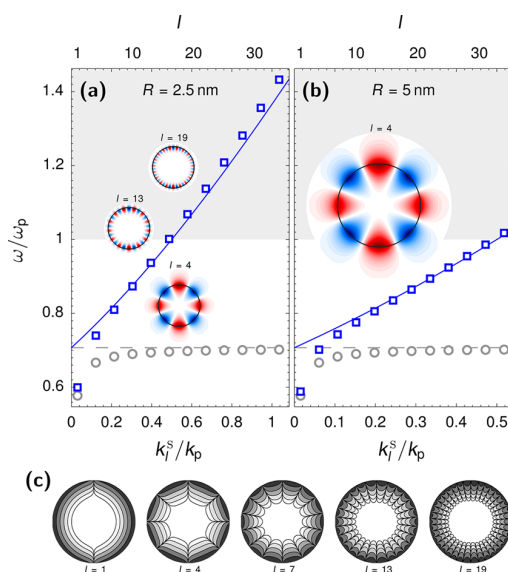


Figure 5. Dispersion of the nonretarded surface plasmon resonances of nanospheres. Material parameters as in Figure 2 but with $\eta = 0$. Wavenumbers are normalized to the plasma wavenumber $k_p = \omega_p/\beta_F$. The hydrodynamic model is shown in blue and the LRA in gray. The $l = 1, 4, 7, \dots, 34$ multipole LSP resonances are indicated by squares and circles; nonretarded dispersion relations⁶⁴ for a planar interface are shown as solid lines. Insets in panels (a) and (b) show the real parts of the electric field of selected LSP modes in the xz plane along θ polarization (on separate color scales). (c) Contour plots of the absolute value of the hydrodynamic charge density of selected LSP modes in the same nanosphere and in the same plane (contours separated by factors of 10 with separate, and logarithmic color scales).

the modifications due to hydrodynamic response, we will now consider three distinct measurements, each with a different sensitivity to the various surface and bulk plasmons:

- 1. Light scattering.** This measurement gives the extinction cross section $\sigma_{\text{ext}}(\omega)$, yielding the ratio of power dissipated due to scattering and absorption of a plane wave relative to incident intensity.
- 2. Electron energy-loss spectroscopy.** EELS gives information on the electron loss function $\Gamma(\omega)$ that expresses the probability that a relativistic electron will lose an energy $\hbar\omega$ due to interaction with the particle. We consider electrons traveling with velocity $v\hat{z}$ and impact parameter \mathbf{b} in the xy plane outside the sphere ($|\mathbf{b}| = b > R$).
- 3. Atomic spontaneous emission.** A dipole orientation-averaged measurement of local spontaneous emission rates relates linearly to the electric local density of states (or LDOS) $\rho^E(\omega)$. We consider emitter positions \mathbf{b} outside the nanosphere ($b > R$).

These three measurements constitute examples of illumination of the sphere by plane-, cylinder-, and spherical-like waves. Extinction is measured in the

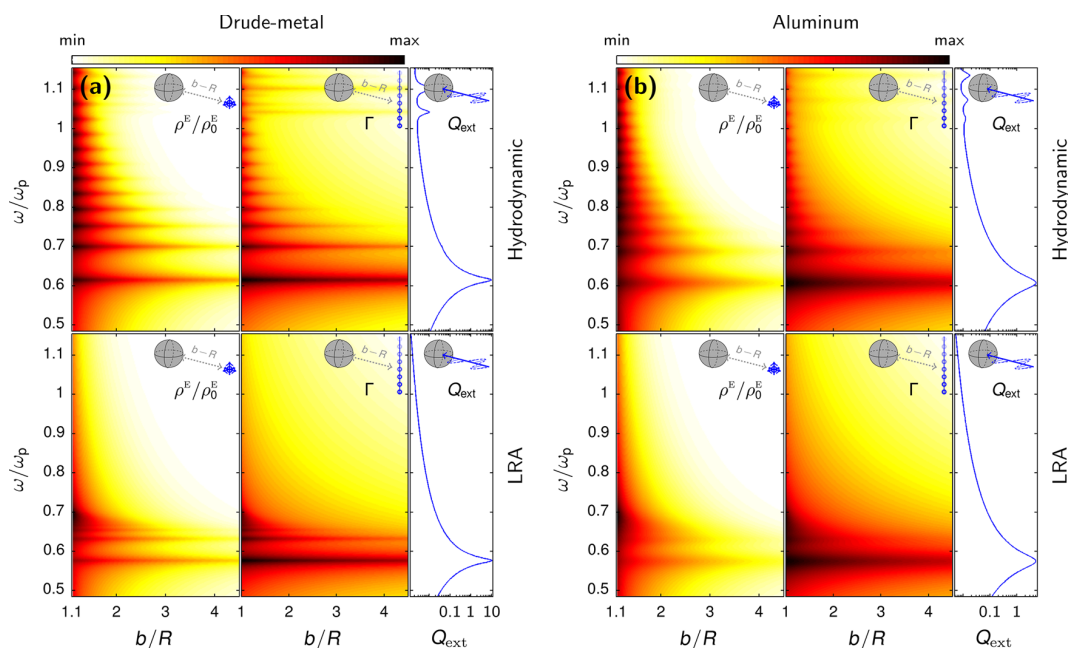


Figure 6. Normalized LDOS, ρ^E/ρ_0^E , EELS probability, Γ , and extinction efficiency, Q_{ext} , in left, center, and right panels, respectively. LDOS and EELS calculations are illustrated on independent logarithmic color scales. An $R = 1.5$ nm sphere in vacuum is considered. Electron energy in EELS calculations is $E_e = 200$ keV. (a) Drude metal with $\omega_p = 10$ eV, $\eta = 0.1$ eV, and $\epsilon_\infty = 1$. (b) Aluminum with bound response included from measured data from ref 74 via $\epsilon_\infty(\omega)$, with $\omega_p = 14.94$ eV and $\eta = 0.075$ eV.

archetypical far-field scattering setup, while the EELS probability and LDOS can be measured locally in the near-field. Subnanometer control of the probe–surface separation is routinely achieved in EELS⁴⁹ and also demonstrated in fluorescence measurements,^{65,66} permitting experimental investigation of the various calculated spectra that we will show below.

Let us briefly discuss the computation of these measurements in the multipole basis. The arbitrary exciting field can be decomposed into the multipole basis; that is, the coefficients $\{a_{lm}^{\text{ex}}, b_{lm}^{\text{ex}}\}$ can be determined. The scattered field is then obtained through the Mie–Lorenz coefficients using eq 3. A general linear measurement \mathcal{Q} may involve components of the scattered field at a single location, as for the LDOS, or a continuous weighting of different spatial components of the field, as for the extinction cross section or the EELS probability. In any case, the measurements can be expressed through a weighted lm summation of the scattering amplitudes $t_l^{\text{TE}} a_l^{\text{ex}}$ and $t_l^{\text{TM}} b_l^{\text{ex}}$. As stated above, for the extinction cross section,⁶⁷ EELS probability,^{68,49} and LDOS,^{69–73} the measurements \mathcal{Q} can all be expressed in terms of the Mie–Lorenz coefficients in the general form of eq 5. For the specific forms that eq 5 takes for each of the three measurements, we refer the reader to eqs S3, S6, and S9 of the SI.

In the following, we normalize the extinction cross section to the geometric cross section, πR^2 , yielding the extinction efficiency $Q_{\text{ext}}(\omega) \equiv \sigma_{\text{ext}}(\omega)/\pi R^2$, and similarly normalize the LDOS to the free-space LDOS $\rho_0^E(\omega)$, yielding the LDOS enhancement $[\rho^E/\rho_0^E](\omega)$.

Near-Field versus Far-Field. Figure 6a depicts the probe-to-surface separation dependence of the LDOS and EELS spectra in a Drude metal nanosphere of $R = 1.5$ nm and for comparison also depicts the extinction resonances. Hydrodynamic and LRA calculations are shown to be distinctly different. Most conspicuous in Figure 6a is perhaps that many new resonances appear in the nonlocal EELS and LDOS spectra, many more than in extinction, and that drastic changes occur when we vary b/R from the contact scenario $b/R = 1$ to $b/R = 4.5$. When fixing $b/R = 2$, we obtain the spectra of Figure 7a. Below we discuss both figures in more detail, but before that, Figure 6a already makes clear that only a rudimentary understanding of EELS measurements can be obtained by comparing them with calculated extinction or absorption spectra. Such comparisons have nevertheless been quite common.^{14,15}

Let us interpret Figures 6a and 7a in more detail by first discussing the region below the plasma frequency, where both in local and nonlocal response the extinction efficiency exhibits just the single dipolar ($l = 1$) surface plasmon resonance. Higher-order multipole plasmons do not contribute since the sphere size is much smaller than the wavelength of the incident plane wave.³³

In stark contrast to these known extinction spectra, several additional multipole LSP resonances are observable in the EELS and LDOS spectra and better so for smaller probe-to-surface separations. Notice that higher-order LSP modes do exist in the LRA, as we have seen in the analysis of the Mie–Lorenz coefficients, but these

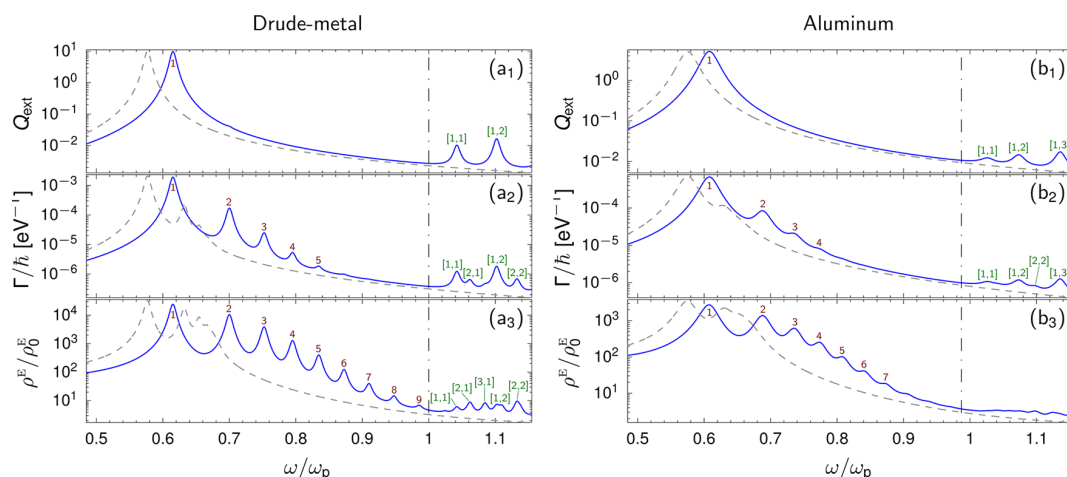


Figure 7. Extinction efficiency, Q_{ext} , EELS probability, Γ , and normalized LDOS, ρ^E/ρ_0^E . Hydrodynamics in solid blue and LRA in dashed gray. The screened plasma frequency is indicated in dashed-dotted black. When distinguishable, the LSP multipole order l is noted in red, while bulk plasmon $[l,n]$ orders are noted in green. Parameters in (a) are as in Figure 6a; in (b) as in Figure 6b. The EELS probability and LDOS are computed for $b/R = 2$ in all three cases.

additional LSP resonances converge toward the $l = \infty$ limit at $\omega_p/\sqrt{2}$ and rapidly become indistinguishable due to losses. By contrast, the higher-order LSP resonances are much more clearly visible in the hydrodynamic spectra because of the l -dependent nonlocal blueshift of eq 8, which pushes the multipole resonances in the EELS and LDOS spectra beyond the LRA $l = \infty$ limit and moreover separates them despite the loss-induced broadening.^{37,38}

Observation of a multipolar resonance above the $l = \infty$ limit was reported by vom Felde *et al.*¹⁶ in EELS measurements on ensembles of potassium clusters of radius 1–2 nm embedded in magnesium oxide. Vom Felde *et al.* attributed this blueshift into the classically quiet region to quantum size effects. Here we show that there is a good alternative explanation, namely, collective hydrodynamic multipolar LSP resonances. Thus the ongoing discussion how to interpret the blueshift of the main dipolar LSP resonance as seen in EELS^{14,15,32} can now be extended to higher-order LSP resonances, observable in both EELS and LDOS measurements. This improves the outlook of obtaining conclusive evidence for hydrodynamic behavior in plasmonic nanospheres.

Importantly, our calculations performed for aluminum ($\omega_p = 14.94$ eV) in Figures 6b and 7b, using measured data from ref 74, confirm the feasibility of measuring multipole resonances beyond the $l = \infty$ limit for realistic (*i.e.*, non-Drude) metals: at least 4 orders of surface plasmons besides both dipole and quadrupole bulk plasmons are discernible. The nanosphere radius considered in Figures 6 and 7 is, however, relatively small at $R = 1.5$ nm. While consideration of such small nanospheres eases interpretation and labeling, it also approaches the emergence of the realm of cluster physics. Nevertheless, similar spectral features persist for larger spheres, upholding the pertinence of the

analysis. Supporting calculations for $R = 3$ nm nanospheres are presented in the SI.

We emphasize that one should not view the results in Figures 6b and 7b as being fully representative of experiments: the semiclassical plasma-in-a-box hydrodynamic model necessarily cannot contain all relevant physics. In particular, it is known that the nonlocal blueshift of the *dipolar* SPP for aluminum spheres in vacuum will be more than fully compensated by a redshift due to electronic spill-out.⁷⁵

However, for higher-order multipoles, we expect that the centroid of the induced charge will be pushed inward at larger multipole orders, and that nonlocality will come to dominate the effects of spill-out. These considerations are supported by calculations in ref 22 on planar simple metal surfaces, which show that the induced charge recedes to the interior of the metal at large momentum transfers, equivalent to high-multipole order. This indicates that spill-out does not undo our prediction that higher-order SPP resonances will be well-separated due to nonlocal response and thus suggests a novel direction for identification of hydrodynamic behavior in nanospheres. The key features of our theoretical near-field spectra for aluminum are encouraging in this respect. Accordingly, experimental investigation and further theoretical study with more microscopic models is highly desirable.

Additionally, we note that electronic spill-out is not a property of the metal nanoparticle alone but also of its surrounding dielectric, in a similar way that the atomic spontaneous emission rate is not a property of the atom alone but also of its electromagnetic environment. This gives additional experimental freedom: by embedding metal spheres into a solid matrix, electronic spill-out can be controlled and the associated redshift suppressed.¹⁶ A high-index dielectric surrounding can significantly reduce the electronic

spill-out, even in simple metals. Thus with high-index background dielectrics, our plasma-in-a-box model is expected to be more accurate. The key effects of a nonunity background dielectric function on the SPP and bulk plasmon resonances of Figures 6 and 7 can be readily discerned from eqs 7–9.

As further promising experiments, we propose to use the same materials as in ref 16, namely, potassium (or Na or Rb) nanospheres in an MgO matrix, but now for doing EELS on an individual nanosphere, so that inhomogeneous broadening would no longer obscure individual multipolar peaks. Similarly, rather than utilizing a continuous embedding matrix, it may be feasible to suppress the electronic spill-out just by coating the nanospheres with a suitable dielectric, thereby also providing protection from oxidation.

The higher-order LSPs that we propose to observe were not seen in the recent EELS measurements on silver nanospheres of refs 14 and 15. This agrees with calculations performed by us for silver, which are detailed in the SI: due to strong interband effects, higher-order multipole LSP resonances are obscured even in individual Ag nanospheres.

Above the plasma frequency, two hydrodynamic peaks can be seen in the (identical) extinction spectra of Figures 6a and 7a. They clearly have no analogue in the LRA and correspond to the first two dipolar confined bulk plasmon resonances, with labels $[l,n] = [1,1]$ and $[1,2]$ that we also identified in the hydrodynamic Mie–Lorenz coefficients in Figure 3. They have first been predicted by Ruppin to exist in the extinction spectrum.³³ Interestingly, in the EELS and LDOS spectra of Figures 6 and 7, we see more resonances above the plasma frequency than the two dipolar bulk plasmons of the extinction spectrum. According to our investigations of the Mie–Lorenz coefficients in Figures 3 and 5, these additional resonances in principle could be either high- l LSP resonances or quadrupolar and higher-order bulk plasmon resonances. They all turn out to be bulk plasmons and are therefore labeled accordingly; the high- l LSP resonances are much weaker and absent in the spectrum.

Better than observing shifts in LSP peaks, observing the confined bulk plasmon peaks would constitute a unique identification of hydrodynamic pressure waves in nanospheres. However, since we find them to be three orders of magnitude weaker than the dipolar LSP resonance, actually the same order of magnitude weaker as found in recent density functional calculations,²⁸ they are difficult to measure in nanospheres. To our knowledge, they have not yet been observed (unlike their counterparts in thin films), so to date, bulk plasmons are “non-smoking guns” of hydrodynamic pressure waves in nanospheres.

Overall, Figures 6 and 7 illustrate the importance of the dimensionality of the excitation source. As is well-known, the plane wave used in extinction measurements

only excites dipole resonances in deeply subwavelength spheres. As to the EELS spectra, the one-dimensional source of a traveling electron excites a cylinder-like field, which for short probe-to-surface separations is sufficiently inhomogeneous to excite higher-order ($l > 1$) plasmons, as well. Lastly, the LDOS spectra illustrate the largest sensitivity to the multipole modes, with all LSPs discernible and significant response from several bulk plasmon orders. The spherical-like field of the zero-dimensional dipole induces locally a more inhomogeneous excitation field than the traveling electron, thus accounting for the increased multipole sensitivity in LDOS compared to EELS. At large probe–surface separations shown in Figure 6, the exciting fields in both EELS and LDOS are almost homogeneous near the sphere, and the response due to higher-order multipoles is diminished. As a consequence, for large probe-to-surface separations, the spectral response in extinction, EELS, and LDOS is qualitatively the same. See SI for analytical considerations of this latter point, regarding the asymptotics of the LDOS and EELS spectra.

Distance Dependence of LDOS. In the preceding sections, we established that the response of high-order plasmons is significantly enhanced with probes of low-dimensionality when examined in the near-field, where the observability of multipolar LSPs is enhanced by hydrodynamics. Let us therefore finally focus solely on the LDOS spectra, where the response of these high-order multipoles is most pronounced. In Figure 8, we display the variation of the LDOS spectra as a function of the probe-to-surface separation, varying from $b/R = 1$ (i.e., source on surface) to $b/R = 5$ (10 nm separation). For the panels with $b/R < 2$, contributions from high-order multipoles are increasingly important, as the excitation of multiple LSP orders contribute to the spectrum.⁷⁶ Consequently, in the LRA, the largest LDOS occurs at $\omega_p/\sqrt{2}$, the limiting frequency of the high-order LSPs, coinciding with the pile-up of LRA multipoles. By contrast, the hydrodynamically blueshifted LSPs do not have a finite limiting frequency or an associated similar pile-up of modes, but instead exhibit distinguishable peaks associated with excitation of different multipoles. The qualitative discrepancy between local and nonlocal spectra is even substantial. For larger spheres, the multipole peaks merge and instead give rise to a broad-band enhancement above $\omega_p/\sqrt{2}$, even extending beyond the plasma frequency; see SI for supporting calculations on an $R = 10$ nm sphere. This suggests a hitherto largely unexplored regime of studying nonlocal response in comparatively large nanostructures but at short surface-to-probe separations.

As is well-known, in the extreme limit $b = R$, the LRA LDOS diverges (hence not shown) due to the $1/(b - R)^3$ scaling of the nonradiative decay rate. For the $b/R = 1$ panel of Figure 8, we obtain convergent results for the hydrodynamic response and associated finite LDOS

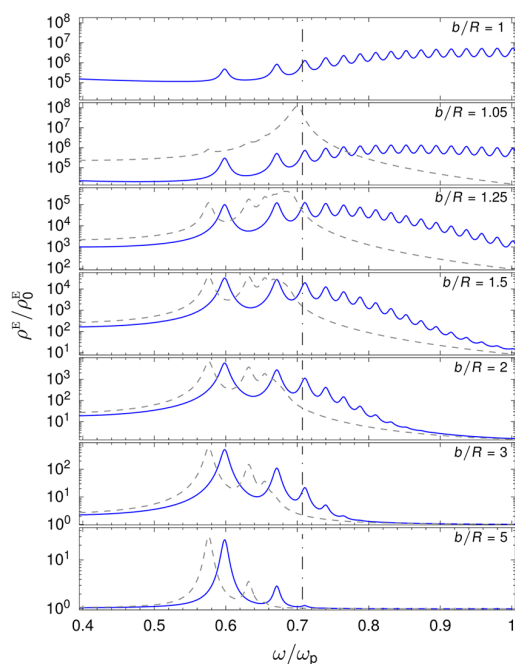


Figure 8. Normalized LDOS for different probe-to-surface separations in hydrodynamic and LRA treatments, in solid blue and dashed gray, respectively, for a Drude metal with material parameters as in Figure 6a, for a $R = 2.5$ nm sphere.

spectra. The convergence, however, hinges upon the choice of a simple Drude metal with real-valued ϵ_{∞} , as discussed in ref 77. As such, hydrodynamic response does not fully regularize the divergence of the LDOS for real metals with dissipative bound response. Complete regularization in real metals would likely necessitate an appropriate nonlocal treatment of not only the free response but also the bound response. In addition, for these very close proximities between source and nanosphere, the effect of high-order moments—beyond the dipole—of the source itself, due to the finite size of the source, would certainly modify the decay rates, as well.⁷⁸ For emitters at the larger separations (*e.g.*, in the panel with $b/R = 5$), the dipole mode of the nanosphere is again the primary feature but with the quadrupolar LSP still imposing a significant spectral feature.

CONCLUSIONS

In this paper, we have aimed to identify indisputable signatures of hydrodynamic response in plasmonic nanospheres. The corresponding evidence for layered systems is the observation, found both with light^{43,44} and with electrons,⁴⁶ of confined bulk plasmons in thin films. Employing the hydrodynamic Drude model, we predict the existence of confined bulk plasmons also in nanospheres. An important question then is whether such excitations would be observable. A series of confined bulk plasmons of dipolar character have been predicted before to show up in extinction spectra.³³

Here we additionally found that besides the dipole series, also series of quadrupolar and higher-order bulk plasmons emerge in near-field EELS and LDOS spectra. However, we find the strength of these bulk plasmon resonances in nanospheres to be about three orders of magnitude weaker than the dominant LSP peak. Their experimental observation in nanospheres, for example with EELS or LDOS, remains an open challenge. Another promising technique is core-level photoemission.⁴⁶

Of a more immediate, accessible nature experimentally is our prediction that, in the near-field EELS and LDOS spectra, also quadrupolar and higher-order LSPs appear, besides the well-known dominant dipolar LSPs. In itself, it is no surprise that higher-order LSPs show up in near-field spectra because the LRA predicts them already.⁷⁶ The salient point here is that LRA LSPs exhibit the surface plasmon $\omega_p/\sqrt{2}$ of a planar interface as a limiting upper frequency, while we predict hydrodynamic LSPs to be observable also above $\omega_p/\sqrt{2}$. This follows from our prediction that higher- l plasmons exhibit a larger nonlocal blueshift. Indeed, we found that high- l LSPs in principle can occur above the plasma frequency in few-nanometer spheres, with their mode profiles still well-bound to the surface. An upper limiting frequency for LSPs actually does not exist in the hydrodynamic model.

Not all multipolar LSPs will be observable, though. For silver, we predict all LSPs besides the dipolar one to be suppressed due to interband effects. On the other hand, we predict that for aluminum nanospheres several higher-order LSPs should be observable in near-field EELS and LDOS spectra. In ensembles of alkali metal (Na, K, Rb) nanospheres in a MgO matrix, resonances above the LRA limit $\omega_p/\sqrt{2}$ have actually already been observed, but individual resonance peaks could not be resolved due to ensemble averaging.¹⁶ We propose to do these measurements on individual alkali metal nanospheres, something that has already been achieved with silver nanospheres.^{14,15}

Would such measurements constitute the unequivocal evidence, the “smoking gun”, of hydrodynamic nonlocal response in nanospheres that we set out to identify? We can only suggest ‘perhaps’ at this stage because alternative explanations for resonances above $\omega_p/\sqrt{2}$ do exist. In particular, vom Felde *et al.* invoke quantum confinement (cluster physics) rather than hydrodynamics (nanoplasmonics) to explain their intriguing observation of resonances above the LRA limit.¹⁶ It is safe to assume, however, that fitting the two distinct models to a measured series of LSP resonances will be more conclusive than fitting only the dominant dipolar LSP, which remains state-of-the-art.^{14,15,32} We therefore suggest to measure near-field EELS and LDOS spectra of nanospheres of aluminum and alkali metals embedded in a solid dielectric environment.

The plasmonic resonances emerge with strikingly different weights in the three types of spectra that we

calculated, so that for example the state-of-the-art comparison of EELS experiments with theoretical absorption cross sections¹⁴ or extinction cross sections¹⁵ can be of limited use. The comparison happened to be useful for silver nanospheres,^{14,15} where interband effects suppress the beyond-dipole LSP resonances that otherwise would show up in near-field EELS and LDOS experiments.

METHODS

Hydrodynamics and Multipole Basis. By eliminating the current density in eqs 1, the hydrodynamic equations can be recast solely in terms of the electric field:

$$(\nabla^2 + k_M^2)\nabla \times \mathbf{E}(\mathbf{r}, \omega) = 0 \quad (10a)$$

$$(\nabla^2 + k_{NL}^2)\nabla \times \mathbf{E}(\mathbf{r}, \omega) = 0 \quad (10b)$$

where $k_M^2 = k_0^2 \epsilon_M$ and $k_{NL}^2 = (\omega_p/\beta_F)^2 \epsilon_M / [\epsilon_\infty(\epsilon_\infty - \epsilon_M)]$ denote the transverse and longitudinal wavenumbers in the metal, respectively. The transverse response of the metal is governed by $\epsilon_M(\omega) = \epsilon_\infty(\omega) - \sigma(\omega)/i\epsilon_0\omega$.

The vector wave functions, $\mathbf{M}_l(\mathbf{r})$, $\mathbf{N}_l(\mathbf{r})$, and $\mathbf{L}_l(\mathbf{r})$, are defined in terms of a pilot vector \mathbf{c} and a generating scalar function $\psi_l(\mathbf{r})$, satisfying the Helmholtz equation $\nabla^2 \psi_l(\mathbf{r}) + k^2 \psi_l(\mathbf{r}) = 0$. In spherically symmetric structures, it is natural to express the generating functions in spherical coordinates $\mathbf{r} = (r, \theta, \phi)$ and to choose the pilot vector as the (nonconstant) outward radial vector $\mathbf{c} = \mathbf{r}$. In this case, the degeneracy label ν separates into the angular momentum quantum numbers l and m , and the vector wave functions read as

$$\mathbf{M}_{lm}(\mathbf{r}) = \nabla \times \mathbf{r}\psi_{lm}(\mathbf{r}) \quad (11a)$$

$$\mathbf{N}_{lm}(\mathbf{r}) = \frac{1}{k} \nabla \times \nabla \times \mathbf{r}\psi_{lm}(\mathbf{r}) \quad (11b)$$

$$\mathbf{L}_{lm}(\mathbf{r}) = \frac{1}{k} \nabla \psi_{lm}(\mathbf{r}) \quad (11c)$$

with $\psi_{lm}(r, \theta, \phi) = z_l(kr)P_l^m(\cos \theta)e^{im\phi}$, where z_l denotes spherical Bessel or Hankel functions of the first kind, j_l or $h_l^{(1)}$, for in- and outgoing waves, respectively. Finally, P_l^m denotes the associated Legendre polynomials. In addition, by requirements of continuity along ϕ and boundedness at the polar extremes, the angular momentum quantum numbers are restricted to integer values in the ranges $l \in [1, \infty[$ and $m \in [-l, l]$. This particular basis is usually referred to as the multipole basis.

The k -dependence of the vector wave functions used in the field expansions varies inside and outside the sphere. By insertion of the external field into the vector Helmholtz equation, $\nabla^2 \mathbf{E} + k_D^2 \mathbf{E} = 0$, which is valid outside the sphere, it is clear that the appropriate choice of wavenumber is $k_D = \sqrt{\epsilon_D}k_0$ outside the sphere. Similarly, by insertion of the internal field into eqs 10, it is clear that the solenoidal vector waves \mathbf{M}_{lm}^i and \mathbf{N}_{lm}^i inside the sphere are associated with the transverse wavenumber k_M , while the irrotational vector wave \mathbf{L}_{lm}^i is associated with the longitudinal wavenumber k_{NL} .

Finally, the matching of internal and external expansions is facilitated by application of BCs. The usual BCs for the electromagnetic field requires the continuity of the tangential components of the electric and magnetic field at $r = R$, that is, $\mathbf{E}_{||}^{i,c} + \mathbf{E}_{||}^{e,c} = \mathbf{E}_{||}^i$ and $\mathbf{H}_{||}^{i,c} + \mathbf{H}_{||}^{e,c} = \mathbf{H}_{||}^i$. Furthermore, an additional BC is required to account for the presence of the longitudinal waves inside the metal, which, in the case of an abrupt dielectric boundary, is unambiguously chosen as the continuity of the normal component of the induced current, equivalent to the continuity of the normal component of the bound charge depolarization at $r = R$, corresponding to $\epsilon_D \mathbf{E}_{\perp}^{i,c} + \epsilon_D \mathbf{E}_{\perp}^{e,c} = \epsilon_\infty \mathbf{E}_{\perp}^i$.^{53,63}

Even for the relatively simple hydrodynamic theory that we used here, the near-field spectra of nanospheres become rather elaborate and rich—but they can be understood rigorously. We therefore expect that our results could also assist in the interpretation of near-field spectra calculated with more microscopic calculations, with some features attributable to hydrodynamic nonlocal response.

Conflict of Interest: The authors declare no competing financial interest.

Acknowledgment. The Center for Nanostructured Graphene is sponsored by the Danish National Research Foundation, Project DNRF58. This work was also supported by the Danish Council for Independent Research - Natural Sciences, Project 1323-00087.

Supporting Information Available: Additional information regarding computation of extinction cross section, EELS probability, and LDOS in the multipole basis along with asymptotics of EELS and LDOS in near-extinction limits. Also provides derivation of the quasistatic multipolar polarizability, associated LSP resonance conditions, Mie–Lorenz transmission coefficients, approximate longitudinal resonance conditions for bulk plasmons, quasistatic LDOS expressions and associated calculations for a $R = 10$ nm sphere, and supplementary calculations for silver and larger nanospheres. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Ritchie, R. H. Plasma Losses by Fast Electrons in Thin Films. *Phys. Rev.* **1957**, *106*, 874–881.
- Pines, D.; Bohm, D. A Collective Description of Electron Interactions: II. Collective vs Individual Particle Aspects of the Interactions. *Phys. Rev.* **1952**, *85*, 338–353.
- Stockman, M. I. Nanoplasmonics: The Physics Behind the Applications. *Phys. Today* **2011**, *64* (2), 39–44.
- Campion, A.; Kambhampati, P. Surface-Enhanced Raman Scattering. *Chem. Soc. Rev.* **1998**, *27*, 241–250.
- Vazquez-Mena, O.; Sannomiya, T.; Villanueva, L.; Voros, J.; Brugger, J. Metallic Nanodot Arrays by Stencil Lithography for Plasmonic Biosensing Applications. *ACS Nano* **2011**, *5*, 844–853.
- Khlebtsov, N.; Dykman, L. A. Optical Properties and Biomedical Applications of Plasmonic Nanoparticles. *J. Quant. Spectrosc. Radiat. Transfer* **2010**, *111*, 1–35.
- Kyrsting, A.; Bendix, P. M.; Stamou, D. G.; Oddershede, L. B. Heat Profiling of Three-Dimensionally Optically Trapped Gold Nanoparticles Using Vesicle Cargo Release. *Nano Lett.* **2011**, *11*, 888–892.
- Lal, S.; Clare, S. E.; Halas, J. N. Nanoshell-Enabled Photothermal Cancer Therapy: Impending Clinical Impact. *Acc. Chem. Res.* **2008**, *41*, 1842–1851.
- Bozhevolnyi, S. I.; Volkov, V. S.; Devaux, E.; Laluet, J.-Y.; Ebbesen, T. W. Channel Plasmon Subwavelength Waveguide Components Including Interferometers and Ring Resonators. *Nature* **2006**, *440*, 508–511.
- Novotny, L.; Van Hulst, N. Antennas for Light. *Nat. Photonics* **2011**, *5*, 83–90.
- Muskens, O. L.; Giannini, V.; Sánchez-Gil, J. A.; Gómez Rivaz, J. Strong Enhancement of the Radiative Decay Rate of Emitters by Single Plasmonic Nanoantennas. *Nano Lett.* **2007**, *7*, 2871–2875.
- Wu, J.-L.; Chen, F.-C.; Hsiao, Y.-S.; Chien, F.-C.; Chen, P.; Kuo, C.-H.; Huang, M.; Hsu, C.-S. Surface Plasmonic Effects of Metallic Nanoparticles on the Performance of Polymer Bulk Heterojunction Solar Cells. *ACS Nano* **2011**, *5*, 959–967.

13. Stockman, M. I. Nanoplasmonics: Past, Present, and Glimpse into Future. *Opt. Express* **2011**, *19*, 22029–22106.
14. Scholl, J. A.; Koh, A. L.; Dionne, J. A. Quantum Plasmon Resonances of Individual Metallic Nanoparticles. *Nature* **2012**, *483*, 421–427.
15. Raza, S.; Stenger, N.; Kadkhodazadeh, S.; Fischer, S. V.; Kostesha, N.; Jauho, A.-P.; Burrows, A.; Wubs, M.; Mortensen, N. A. Blueshift of the Surface Plasmon Resonance in Silver Nanoparticles Studied with EELS. *Nanophotonics* **2013**, *2*, 131–138.
16. vom Felde, A.; Fink, J.; Ekardt, W. Quantum Size Effects in Excitations of Potassium Clusters. *Phys. Rev. Lett.* **1988**, *61*, 2249–2252.
17. Kern, J.; Grossmann, S.; Tarakina, T.; Hackel, N. V.; Emmerling, M.; Kamp, M.; Huang, J.-S.; Biagioni, P.; Prangma, J. C.; Hecht, B. Atomic-Scale Confinement of Resonant Optical Fields. *Nano Lett.* **2012**, *12*, 5504–5509.
18. Savage, K. J.; Hawkeye, M. M.; Esteban, R.; Borisov, A. G.; Aizpurua, J.; Baumberg, J. J. Revealing the Quantum Regime in Tunnelling Plasmonics. *Nature* **2012**, *491*, 574–577.
19. Scholl, J.; Garcia-Etxarri, A.; Koh, A. L.; Dionne, J. A. Observation of Quantum Tunneling between Two Plasmonic Nanoparticles. *Nano Lett.* **2013**, *13*, 564–569.
20. Ciraci, C.; Hill, R. T.; Mock, J. J.; Urzhumov, Y.; Fernández-Domínguez, A. I.; Maier, S. A.; Pendry, J. B.; Chilkoti, A.; Smith, D. R. Probing the Ultimate Limits of Plasmonic Enhancement. *Science* **2012**, *337*, 1072–1074.
21. Lang, N. D.; Kohn, W. Theory of Metal Surfaces: Charge Density and Surface Energy. *Phys. Rev. B* **1970**, *1*, 4555–4568.
22. Liebsch, A. Surface-Plasmon Dispersion and Size Dependence of Mie Resonance: Silver versus Simple Metals. *Phys. Rev. B* **1993**, *48*, 11317–11328.
23. Teperik, T. V.; Nordlander, P.; Aizpurua, J.; Borisov, A. G. Robust Subnanometric Plasmon Ruler by Rescaling of the Nonlocal Optical Response. *Phys. Rev. Lett.* **2013**, *110*, 263901.
24. Öztürk, Z. F.; Xiao, S.; Yan, M.; Wubs, M.; Jauho, A.-P.; Mortensen, N. A. Field Enhancement at Metallic Interfaces Due to Quantum Confinement. *J. Nanophotonics* **2011**, *5*, 051602.
25. Esteban, R.; Borisov, A. G.; Nordlander, P.; Aizpurua, J. Bridging Quantum and Classical Plasmonics with a Quantum-Corrected Model. *Nat. Commun.* **2012**, *3*, 825.
26. de Heer, W. A. The Physics of Simple Metal Clusters: Experimental Aspects and Simple Models. *Rev. Mod. Phys.* **1993**, *65*, 611–676.
27. Keller, O.; Xiao, M.; Bozhevolnyi, S. Optical Diamagnetic Polarizability of a Mesoscopic Metallic Sphere: Transverse Self-Field Approach. *Opt. Commun.* **1993**, *102*, 238–244.
28. Townsend, E.; Bryant, G. W. Plasmonic Properties of Metallic Nanoparticles: The Effects of Size Quantization. *Nano Lett.* **2011**, *12*, 429–434.
29. Ginzburg, P.; Zayats, A. V. Localized Surface Plasmon Resonances in Spatially Dispersive Nano-objects: Phenomenological Treatise. *ACS Nano* **2013**, *7*, 4334–4342.
30. Raza, S.; Yan, W.; Stenger, N.; Wubs, M.; Mortensen, N. A. Blueshift of the Surface Plasmon Resonance in Silver Nanoparticles: Substrate Effects. *Opt. Express* **2013**, *21*, 27344–27355.
31. Stella, L.; Zhang, P.; García-Vidal, F. J.; Rubio, A.; García-González, P. Performance of Nonlocal Optics When Applied to Plasmonic Nanostructures. *J. Phys. Chem. C* **2013**, *117*, 8941–8949.
32. Carmina Monreal, R.; Antosiewicz, T. J.; Apell, S. P. Competition between Surface Screening and Size Quantization for Surface Plasmons in Nanoparticles. *New J. Phys.* **2013**, *15*, 083044.
33. Ruppín, R. Optical Properties of a Plasma Sphere. *Phys. Rev. Lett.* **1973**, *31*, 1434–1437.
34. David, C.; García de Abajo, F. J. Spatial Nonlocality in the Optical Response of Metal Nanoparticles. *J. Phys. Chem. C* **2012**, *115*, 19470–19475.
35. Fuchs, R.; Claro, F. Multipolar Response of Small Metallic Spheres: Nonlocal Theory. *Phys. Rev. B* **1987**, *35*, 3722–3726.
36. Baltz, R. v.; Mensch, M.; Zohm, H. Longitudinal Electric Response and Loss-Function of Metallic Microspheres and Voids. *Z. Phys. B* **1995**, *98*, 151–161.
37. Tran Thoai, D. Surface Excitations on Small Metal Particles by Fast Electrons. *Phys. Status Solidi B* **1986**, *136*, 291–298.
38. Tran Thoai, D.; Zeitler, E. Multipole Surface Excitations on Small Oxide-Covered Metal Particles by Fast Electrons. *Appl. Phys. A: Mater. Sci. Process.* **1988**, *45*, 249–253.
39. Mortensen, N. A. Nonlocal Formalism for Nanoplasmonics: Phenomenological and Semi-classical Considerations. *Photonics Nanostruct.: Fundamental Appl.* **2013**, *11*, 302–308.
40. Sipe, J. E.; So, V. C. Y.; Fukui, M.; Stegeman, G. I. Analysis of Second-Harmonic Generation at Metal Surfaces. *Phys. Rev. B* **1980**, *21*, 4389–4402.
41. Ginzburg, P.; Krasavin, A.; Sonnefraud, Y.; Murphy, A.; Pollard, R. J.; Maier, S. A.; Zayats, A. V. Nonlinearly Coupled Localized Plasmon Resonances: Resonant Second-Harmonic Generation. *Phys. Rev. B* **2012**, *86*, 085422.
42. Ciraci, C.; Poutrina, E.; Scalora, M.; Smith, D. R. Second-Harmonic Generation in Metallic Nanoparticles: Clarification of the Role of the Surface. *Phys. Rev. B* **2012**, *86*, 115451.
43. Lindau, I.; Nilsson, P. O. Experimental Verification of Optically Excited Longitudinal Plasmons. *Phys. Scr.* **1971**, *3*, 87–92.
44. Anderegg, M.; Feuerbacher, B.; Fitton, B. Optically Excited Longitudinal Plasmons in Potassium. *Phys. Rev. Lett.* **1971**, *27*, 1565–1568.
45. Chen, C. H. Plasmon Dispersion in Single-Crystal Magnesium. *J. Phys. C: Solid State Phys.* **1976**, *9*, L321–L324.
46. Özer, M. M.; Moon, E. J.; Eguluz, A. G.; Weitering, H. H. Plasmon Response of a Quantum-Confined Electron Gas Probed by Core-Level Photoemission. *Phys. Rev. Lett.* **2011**, *106*, 197601.
47. Duthler, C. J.; Johnson, S. E.; Broida, H. P. Plasma-Resonance Scattering from Small Sodium Particles Formed in a Flowing Gas Stream. *Phys. Rev. Lett.* **1971**, *26*, 1236–1239.
48. Greffet, J.-J.; Carminati, R. Image Formation in Near-Field Optics. *Prog. Surf. Sci.* **1997**, *56*, 133–237.
49. García de Abajo, F. J. Optical Excitations in Electron Microscopy. *Rev. Mod. Phys.* **2010**, *82*, 209–275.
50. Egerton, R. F. Electron Energy-Loss Spectroscopy in the TEM. *Rep. Prog. Phys.* **2009**, *72*, 016502.
51. Schmelzeisen, M.; Zhao, Y.; Klapper, M.; Müllen, K.; Kreiter, M. Fluorescence Enhancement from Individual Plasmonic Gap Resonances. *ACS Nano* **2010**, *4*, 3309–3317.
52. Willets, K. Super-resolution Imaging of Interactions between Molecules and Plasmonic Nanostructures. *Phys. Chem. Chem. Phys.* **2013**, *15*, 5345–5354.
53. Boardman, A. D. *Electromagnetic Surface Modes. Hydrodynamic Theory of Plasmon-Polaritons on Plane Surfaces*; John Wiley and Sons: Chichester, UK, 1982.
54. Raza, S.; Toscano, G.; Jauho, A.-P.; Wubs, M.; Mortensen, N. A. Unusual Resonances in Nanoplasmonic Structures Due to Nonlocal Response. *Phys. Rev. B* **2011**, *84*, 121412(R).
55. Chew, W. C. *Waves and Fields in Inhomogeneous Media*; IEEE Press: New York, 1994.
56. Stratton, J. A. *Electromagnetic Theory*; McGraw-Hill Book Company: New York, 1941.
57. Quinten, M. *Optical Properties of Nanoparticle Systems: Mie and Beyond*; Wiley-VCH: Weinheim, Germany, 2011.
58. Mie, G. Articles on the Optical Characteristics of Turbid Tubes, Especially Colloidal Metal Solutions. *Ann. Phys.* **1908**, *25*, 377–445.
59. Yan, W.; Mortensen, N. A.; Wubs, M. Green's Function Surface-Integral Method for Nonlocal Response of Plasmonic Nanowires in Arbitrary Dielectric Environments. *Phys. Rev. B* **2013**, *88*, 155414.
60. Boardman, A. D.; Paranjape, B. V. The Optical Surface Modes of Metal Spheres. *J. Phys. F: Metal Phys.* **1977**, *7*, 1935–1945.
61. Gildenburg, V. B.; Kostin, V. A.; Pavlichenko, I. A. Resonances of Surface and Volume Plasmons in Atomic Clusters. *Phys. Plasmas* **2011**, *18*, 092101.

62. Tribelsky, M.; Miroshnichenko, A.; Kivshar, Y. Unconventional Fano Resonances in Light Scattering by Small Particles. *Europhys. Lett.* **2012**, *97*, 44005.
63. Yan, W.; Wubs, M.; Mortensen, N. A. Hyperbolic Metamaterials: Nonlocal Response Regularizes Broadband Super-singularity. *Phys. Rev. B* **2012**, *86*, 205429.
64. Raza, S.; Christensen, T.; Wubs, M.; Bozhevolnyi, S.; Mortensen, N. A. Nonlocal Response in Thin-Film Waveguides: Loss versus Nonlocality and Breaking of Complementarity. *Phys. Rev. B* **2013**, *88*, 115401.
65. Anger, P.; Bharadwaj, P.; Novotny, L. Enhancement and Quenching of Single-Molecule Fluorescence. *Phys. Rev. Lett.* **2006**, *96*, 113002.
66. Dulkeith, E.; Ringler, M.; Klar, T. A.; Feldmann, J.; Javier Muñoz, A.; Parak, W. J. Gold Nanoparticles Quench Fluorescence by Phase Induced Radiative Rate Suppression. *Nano Lett.* **2005**, *5*, 585–589.
67. Bohren, C. F.; Huffman, D. R. *Absorption and Scattering of Light by Small Particles*; John Wiley & Sons: New York, 1983.
68. García de Abajo, F. J. Relativistic Energy Loss and Induced Photon Emission in the Interaction of a Dielectric Sphere with an External Electron Beam. *Phys. Rev. B* **1999**, *59*, 3095–3107.
69. Kerker, M.; Wang, D.-S.; Chew, H. Surface Enhanced Raman Scattering (SERS) by Molecules Adsorbed at Spherical Particles: Errata. *Appl. Opt.* **1980**, *19*, 4159–4174.
70. Ruppin, R. Decay of an Excited Molecule near a Small Sphere. *J. Chem. Phys.* **1982**, *76*, 1681–1684.
71. Chew, H. Transition Rates of Atoms Near Spherical Surfaces. *J. Chem. Phys.* **1987**, *87*, 1355–1360.
72. Dung, H. T.; Knöll, L.; Welsch, D.-G. Decay of an Excited Atom near an Absorbing Microsphere. *Phys. Rev. A* **2001**, *64*, 013804.
73. Vos, W. L.; Koenderink, A. F.; Nikolaev, I. S. Orientation-Dependent Spontaneous Emission Rates of a Two-Level Quantum Emitter in Any Nanophotonic Environment. *Phys. Rev. A* **2009**, *80*, 053802.
74. Rakić, A. D. Algorithm for the Determination of Intrinsic Optical Constants of Metal Films: Application to Aluminum. *Appl. Opt.* **1995**, *34*, 4755–4767.
75. Mandal, S.; Wang, J.; Winans, R. E.; Jensen, L.; Sen, A. Quantum Size Effects in the Optical Properties of Ligand Stabilized Aluminum Nanoclusters. *J. Phys. Chem. C* **2013**, *117*, 6741–6746.
76. Moroz, A. Non-radiative Decay of a Dipole Emitter Close to a Metallic Nanoparticle: Importance of Higher-Order Multipole Contributions. *Opt. Commun.* **2010**, *283*, 2277–2287.
77. Datsyuk, V. V.; Tovkach, O. Optical Properties of a Metal Nanosphere with Spatially Dispersive Permittivity. *J. Opt. Soc. Am. B* **2011**, *28*, 1224–1230.
78. Andersen, M. L.; Stobbe, S.; Sørensen, A. S.; Lodahl, P. Strongly Modified Plasmon–Matter Interaction with Mesoscopic Quantum Emitters. *Nat. Phys.* **2011**, *7*, 215–218.

Supporting Information for:
Nonlocal Response of Metallic Nanospheres
Probed by Light, Electrons, and Atoms

Thomas Christensen,^{†,‡} Wei Yan,^{†,‡} Søren Raza,^{†,¶} Antti-Pekka Jauho,^{§,‡}
N. Asger Mortensen,^{†,‡} and Martijn Wubs^{*,†,‡}

*DTU Fotonik, Center for Nanostructured Graphene, Center for Electron Nanoscopy, and
DTU Nanotech*

E-mail: mwubs@fotonik.dtu.dk

*To whom correspondence should be addressed

[†]Department of Photonics Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[‡]Center for Nanostructured Graphene, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[¶]Center for Electron Nanoscopy, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

[§]Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

Hydrodynamic nonlocal interaction range ξ_{NL}

Table S1 provides a listing of typical values for the nonlocal interaction range $\xi_{\text{NL}} = v_F/\omega$ considered in the optical domain at a wavelength of 500 nm, computed from tabulated plasma frequencies for various relevant plasmonic metals.

Table S1: Table of values for the plasma frequency, ω_p , Fermi velocity, v_F , and nonlocal interaction range, ξ_{NL} (at a wavelength of 500 nm), for a selection of plasmonic metals.

Metal	$\hbar\omega_p$ [eV]	v_F [10^6 m/s]	ξ_{NL} [\AA]	Reference
Li	8.05	1.29	3.4	S1
Na	6.04	1.07	2.8	S1
Al	14.94	1.95	5.2	S2
K	4.39	0.86	2.3	S1
Cu	10.83	1.58	4.2	S3
Pd	9.72	1.47	3.9	S3
Ag	9.01	1.39	3.7	S3
Pt	9.59	1.45	3.9	S3
Au	9.03	1.40	3.7	S3

Measurement coefficients

In this section, we review the particulars of the Mie–Lorenz coefficient expansion of the extinction cross-section, σ_{ext} , the EELS probability, Γ , and the free-space normalized local density of states (LDOS), ρ^E/ρ_0^E . First, we briefly remind how a given exciting field can be decomposed into multipoles.

Multipole expansion of exciting field. With the relationship between exciting and scattered fields established, Eq. (3), the problem of deducing the scattered field due to some exciting field is reduced to expanding the exciting field in the multipole basis. As a consequence of the orthogonality of the vector wave functions on the surface of a sphere, the expansion coefficients can principally be obtained from:^{S4}

$$a_{lm}^{\text{ex}} = \frac{\int_0^{2\pi} \int_0^\pi \mathbf{E}^{\text{ex}} \cdot \bar{\mathbf{M}}_{lm}^{[1]} \sin \theta \, d\theta d\phi}{\int_0^{2\pi} \int_0^\pi |\mathbf{M}_{lm}^{[1]}|^2 \sin \theta \, d\theta d\phi}, \quad (\text{S1})$$

with $\bar{\mathbf{F}}$ denoting the complex conjugate of \mathbf{F} . An identical equation for b_{lm}^{ex} exists with $\mathbf{M}_{lm}^{[1]}$ replaced by $\mathbf{N}_{lm}^{[1]}$.

At this point, it is worthwhile noting that the inclusion of hydrodynamics incurs no additional analytical difficulties in the external region, compared to local theory; all results from local theory remain valid for $r > R$, provided the local TM Mie–Lorenz coefficients are supplemented by the hydrodynamic correction Δ_l from Eq. (4c). As such, for measurements restricted to the external region, hydrodynamics can be immediately included using well-established results from local theory. In the internal region, the additional wave components due to the longitudinal multipoles, $\mathbf{L}_{lm}^{[1]}$, break this convenient correspondence. In the following we consider evaluation of the extinction cross-section, the EELS probability and the electric LDOS in the external region.

Extinction cross-section. In the case of an incident plane wave, propagating along the z -direction and polarized along the x -direction, $\mathbf{E}^{\text{ex}}(\mathbf{r}) = e^{ik_D z} \hat{\mathbf{e}}_x$, the exciting field can be decomposed in a multipole basis with $m = \pm 1$, leading to expansion coefficients:^{S4}

$$a_{lm}^{\text{ex}} = E_l m \delta_{|m|1}, \quad b_{lm}^{\text{ex}} = E_l \delta_{|m|1}, \quad (\text{S2})$$

with $E_l = -i^{l+1}(2l+1)/[2l(l+1)]$.

The extinction cross-section, σ_{ext} , which measures the ratio of power dissipated due to both scattering and absorption by the sphere, W_{ext} , to the incident intensity, I_0 , can be obtained by application of the optical theorem,^{S5} giving;

$$\sigma_{\text{ext}} = \frac{W_{\text{ext}}}{I_0} = \frac{2\pi}{k_D^2} \sum_{l=1}^{\infty} (2l+1) \text{Re}(t_l^{\text{TE}} + t_l^{\text{TM}}). \quad (\text{S3})$$

For discussion of actual results, we prefer the dimensionless extinction efficiency $Q_{\text{ext}} = \sigma_{\text{ext}}/\pi R^2$ rather than the cross-section. Regardless of the choice of efficiency or cross-section, the characteristics of the extinction closely mirrors those of standard experimental transmission measurements on widely separated particle arrays.

Electron energy loss probability with aloof electron. The case of the EELS probability for aloof electron trajectories is also approachable by expansion in the multipole basis. In particular, an electron traveling at constant velocity $\mathbf{v} = v\hat{\mathbf{z}}$ with $t = 0$ impact parameter \mathbf{b} in the xy -plane, emanates a cylinder-like wave from the electron trajectory $\mathbf{r}_e(t) = \mathbf{b} + \mathbf{v}t$. Specifically, if $\mathbf{b} = 0$ the traveling charge density is $\rho(\mathbf{r}, t) = -e\delta(\mathbf{r} - \mathbf{v}t)$ which excites an electric field $\mathbf{E}^{\text{ex}}(\mathbf{r}, \omega) = \frac{e\omega}{2\pi\epsilon_0 v^2 \gamma \epsilon} e^{i\omega z/v} \left[\frac{i}{\gamma} K_0\left(\frac{\omega r_{\parallel}}{v\gamma}\right) \hat{\mathbf{z}} - K_1\left(\frac{\omega r_{\parallel}}{v\gamma}\right) \hat{\mathbf{r}}_{\parallel} \right]$ with $\gamma = 1/\sqrt{1 - (v/c)^2}$ denoting the Lorentz contraction factor.^{S6,S7} This incident field scatters off the metallic sphere, and the scattered field,

working back on the electron, incurs a measurable energy loss, ΔE , for the total electron energy

$$\Delta E = e \int_{-\infty}^{\infty} \mathbf{E}^{\text{sc}}[\mathbf{r}_e(t), t] \cdot \mathbf{v} dt. \quad (\text{S4})$$

The loss can also be expressed in terms of a frequency-decomposition through $\Delta E \equiv \int_0^{\infty} \hbar\omega\Gamma(\omega) d\omega$, with $\Gamma(\omega)$ denoting the electron loss probability. The time-domain scattered field in Eq. (S4) can be represented by its frequency-domain Fourier components, which, together with the definition of $\Gamma(\omega)$ and the property $\mathbf{E}(\mathbf{r}, \omega) = \bar{\mathbf{E}}(\mathbf{r}, -\omega)$, allows expression of $\Gamma(\omega)$ in terms of the scattered field:

$$\Gamma(\omega) = \frac{e\mathbf{v}}{\pi\hbar\omega} \int_{-\infty}^{\infty} \hat{\mathbf{z}} \cdot \text{Re} \left\{ \mathbf{E}^{\text{sc}}[\mathbf{r}_e(t), \omega] e^{-i\omega t} \right\} dt. \quad (\text{S5})$$

The problem of determining the appropriate multipole expansion of the exciting field due to the traversing electron, and the subsequent integration of the induced field as required to obtain $\Gamma(\omega)$ through Eq. (S5), was solved by F. J. García de Abajo in Ref. S8 for the case of a sphere embedded in vacuum, $\varepsilon_D = 1$, and for aloof electron trajectories, $|\mathbf{b}| = b > R$. The resulting expression is:^{S7,S8}

$$\Gamma(\omega) = \frac{\alpha}{\omega} \sum_{l=1}^{\infty} \sum_{m=-l}^l K_m^2 \left(\frac{\omega b}{v\gamma} \right) [C_{lm}^{\text{TE}} \text{Re}(t_l^{\text{TE}}) + C_{lm}^{\text{TM}} \text{Re}(t_l^{\text{TM}})], \quad (\text{S6})$$

where $\alpha = \frac{e^2 \varepsilon_0}{\hbar c 4\pi}$ is the fine-structure constant, K_m denotes the modified Bessel function of the second kind of order m , and C_{lm}^{TE} and C_{lm}^{TM} are functions of l , m , and v/c given by:^{S7}

$$C_{lm}^{\text{TE}} = \frac{1}{l(l+1)} |2m\Pi_{lm}|^2, \quad C_{lm}^{\text{TM}} = \frac{1}{l(l+1)} \left| \frac{c}{v\gamma} \Xi_{lm} \right|^2, \quad (\text{S7a})$$

with

$$\Pi_{lm} = \sqrt{\frac{(2l+1)(l-|m|)!(2|m|-1)!!}{\pi(l+|m|)! (v\gamma/c)^{|m|}}} C_{l-|m|}^{(|m|+1/2)} \left(\frac{c}{v} \right), \quad (\text{S7b})$$

$$\Xi_{lm} = \Pi_{l,m+1} \sqrt{(l+m+1)(l-m)} + \Pi_{l,m-1} \sqrt{(l-m+1)(l+m)}, \quad (\text{S7c})$$

where $C_n^{(\nu)}(x)$ denotes the n th Gegenbauer polynomial of order ν .^{S9}

Lastly, we note that the relativistic kinetic energy of the electron, E_e , relates to its velocity, v , through $E_e = m_e\gamma c^2 - m_e c^2$, where m_e denotes the electron mass. Consequently, a given

kinetic electron energy E_e corresponds to the velocity

$$\frac{v}{c} = \sqrt{1 - \left(\frac{m_e c^2}{E_e + m_e c^2} \right)^2}. \quad (\text{S8})$$

Local density of states. The third and final relevant excitation field and measurement to be considered here, is that of an electric dipole field and a measurement of the LDOS, relevant *e.g.*, for the spontaneous decay of an electric dipole emitter near the sphere. The problem of dipole radiation outside a sphere was first considered by M. Kerker *et al.* in Ref. S10 using the multipole basis, in the context of surface enhanced Raman scattering (SERS). Subsequently, the problem of decay rates of emitters near metallic and dielectric spheres was treated, relating the Mie–Lorenz coefficients to the decay rate enhancement.^{S11,S12}

The partial electric LDOS experienced by an emitter of transition frequency ω with its dipole-moment oriented along the radial and tangential directions, ρ_{\perp}^E and ρ_{\parallel}^E , respectively, at a distance $b > R$ from origo is given by:^{S12,S13}

$$\begin{aligned} \frac{\rho_{\perp}^E}{\rho_0^E} &= 1 + \frac{3}{2} \frac{1}{y^2} \sum_{l=1}^{\infty} (2l+1)l(l+1) \text{Re} \left[t_l^{\text{TM}} h_l^{(1)}(y)^2 \right], \\ \frac{\rho_{\parallel}^E}{\rho_0^E} &= 1 + \frac{3}{4} \frac{1}{y^2} \sum_{l=1}^{\infty} (2l+1) \text{Re} \left[t_l^{\text{TE}} \xi_l(y)^2 + t_l^{\text{TM}} \xi_l'(y)^2 \right], \end{aligned} \quad (\text{S9})$$

where ρ_0^E denotes the LDOS in the absence of the sphere $\rho_0^E = \omega^2/\pi^2 c^3$, $y = k_0 b$ denotes phase-accumulation from dipole to sphere, and finally ξ_l denotes the Riccati-Bessel function $\xi_l(y) = y h_l^{(1)}(y)$ introduced for brevity of notation. The orientation-averaged LDOS, ρ^E , can be obtained from the partial LDOS through summation^{S14,S15} $\rho^E = \frac{1}{3} \rho_{\perp}^E + \frac{2}{3} \rho_{\parallel}^E$.

We note that the radially oriented emitter couples solely with TM-polarized fields, while the tangentially oriented emitter couples to both TM and TE polarizations. However, for small spheres and probe distances, $y \ll 1$, the TM contribution dominates the TE contribution as can be verified from the $0 < x \ll \sqrt{l+1}$ asymptotic behavior of $\xi_l(x) \simeq N_{l+1} x^{l+1} - i N_l^{-1} x^{-l}$ and $\xi_l'(x) \simeq (l+1) N_{l+1} x^l + i N_l^{-1} x^{-l-1}$ with $N_l = 2^l l! / (2l)!$. As a consequence, we expect strong enhancement of either radial or tangential LDOS to arise primarily due to TM polarized interaction.

Asymptotics of LDOS and EELS - similarities with extinction

It is instructive to consider the limits in which the LDOS and EELS spectra are qualitatively similar to the extinction spectra, in other words, to study the regimes wherein extinction measurements gives information directly comparative to EELS or LDOS measurements on the same system. In this section, we derive asymptotic expressions for the normalized orientation-averaged LDOS in the large-separation range, $k_0 b \gg 1$, and similarly for the EELS signal in the ultra-relativistic limit, $v/c \rightarrow 1$. We show that these asymptotics display the same qualitative behavior as the full extinction spectra, apart from minor probe-related differences.

LDOS at large surface-to-probe separations. We consider the large- y limit of Eqs. (S9), corresponding to probe-to-surface separations exceeding the wavelength in the surrounding dielectric. From the large-argument asymptotics^{S16} of the spherical Hankel and Ricatti-Bessel functions [$h_l^{(1)}(y) \simeq i^{-l-1} y^{-1} e^{iy}$, $\xi_l(y) \simeq i^{-l-1} e^{iy}$, and $\xi_l'(y) \simeq i^{-l} e^{iy}$ valid for $y \gg 1$] we find that Eq. (S9) reduces to:

$$\frac{\rho_{\perp}^E}{\rho_0^E} = 1 + \frac{3}{2} \frac{1}{y^4} \sum_{l=1}^{\infty} (2l+1)l(l+1) \text{Re} \left[t_l^{\text{TM}} e^{2iy} \right] (-1)^{l+1}, \quad (\text{S10a})$$

$$\frac{\rho_{\parallel}^E}{\rho_0^E} = 1 + \frac{3}{4} \frac{1}{y^2} \sum_{l=1}^{\infty} (2l+1) \text{Re} \left[(t_l^{\text{TE}} - t_l^{\text{TM}}) e^{2iy} \right] (-1)^{l+1}, \quad (\text{S10b})$$

which, for the orientation-averaged normalized LDOS, to lowest order in y^{-2} , gives:

$$\frac{\rho^E}{\rho_0^E} = 1 + \frac{1}{2} \frac{1}{y^2} \sum_{l=1}^{\infty} (2l+1) \text{Re} \left[(t_l^{\text{TE}} - t_l^{\text{TM}}) e^{2iy} \right] (-1)^{l+1}, \quad (\text{S11})$$

valid for $y \gg 1$. Apart from scaling, the factor e^{2iy} , which is due to interference of emitted and reflected waves, and some signs, this form is qualitatively similar to the form of the extinction spectra in (S3). In particular, the same $(2l+1)$ magnitude of the weighting is present. This illustrates why just the dipolar term is significant in the large separation limit for small spheres, where t_1^{TM} is logarithmically dominant compared to the remaining Mie–Lorenz coefficients.

EELS signal in ultra-relativistic limit. In the $v/c \rightarrow 1$ limit we can evaluate the Gegenbauer polynomials at unity argument using^{S17} $C_n^{(\nu)}(1) = (2\nu + n - 1)! / [(2\nu - 1)! n!]$ - this value is approached linearly or quadratically as a function of the argument, depending on the parity of n . The limit $v/c \rightarrow 1$ also sends $\gamma \rightarrow \infty$, but a sub-linear rate of divergence. As such, we evaluate

the limit $v/c \rightarrow 1$ in Eq. (S6), but retain γ as finite. To lowest order in γ^{-2} we find after some manipulations that $C_{lm}^{\text{TE, TM}} \simeq \pi^{-1} \gamma^{-2} (2l+1) \delta_{|m|,1}$, leading to:

$$\Gamma(\omega) = \frac{\alpha}{\omega} \frac{2}{\pi \gamma^2} K_1^2 \left(\frac{\omega b}{v \gamma} \right) \sum_{l=1}^{\infty} (2l+1) \text{Re} [t_l^{\text{TE}} + t_l^{\text{TM}}], \quad (\text{S12})$$

valid for v/c near unity, *i.e.*, for ultra-relativistic electron velocities. Evidently, apart from the two frequency-terms outside the sum, the spectral response is identical to that obtained for extinction in Eq. (S3), and as such provides an identical weighting to the various multipoles. Effectively, then, the dipole is the prominent peak due to the dominance of t_l^{TM} relative to the remaining Mie–Lorenz coefficients, at least for small spheres.

Finally, from the $x \gg m$ asymptotic form of the modified Bessel function,⁵⁹ $K_m^2(x) \simeq \frac{\pi}{2x} e^{-2x}$, it is evident that the loss probability decays approximately exponentially in the far-probe region. Due to Lorenz contraction, the transition to the far-probe region is postponed until the contracted distance b/γ is comparable with the electron wavenumber ω/v .

Multipolar polarizability and nonretarded plasmon resonances

We can derive the nonretarded multipolar polarizability, α_l , giving the response to incident potentials of pole order $l+1$, from the Mie–Lorenz coefficients, by noting the interrelationship⁵¹⁸

$$\alpha_l = \lim_{c/\omega R \rightarrow \infty} \left[- \frac{4\pi i \mathcal{N}_l}{k_D^{2l+1}} t_l^{\text{TM}} \right], \quad (\text{S13})$$

with $\mathcal{N}_l = \frac{[(2l+1)!!]^2}{(l+1)(2l+1)}$. This allows determination of the hydrodynamic multipolar polarizability, using Eqs. (4), and yields [upon using the small-argument limiting forms of the spherical Bessel and Hankel functions, namely $j_l(x) \simeq x^l/(2l+1)!!$ and $h_l^{(1)}(x) \simeq x^l/(2l+1)!! - i(2l-1)!!/x^{l+1}$ valid for $x \ll \sqrt{l+1}$]:

$$\alpha_l = 4\pi R^{2l+1} \frac{l [\varepsilon_M - (1 + \delta_l) \varepsilon_D]}{l \varepsilon_M + (l+1)(1 + \delta_l) \varepsilon_D}, \quad (\text{S14})$$

with $\delta_l = \Delta_l/[j_l(x_M)(l+1)]$, see Eq. (4c). The hydrodynamic correction δ_l vanishes in the local limit, such that α_l appropriately reduces to the local response approximation (LRA) multipolar polarizability $\alpha_l^l = 4\pi R^{2l+1} \frac{l(\varepsilon_M - \varepsilon_D)}{l \varepsilon_M + (l+1) \varepsilon_D}$. The hydrodynamic multipolar polarizability was also considered by Fuchs and Claro in Ref. S19, through their more general consideration of the multipolar polarizability of a sphere with dielectric constant $\varepsilon(k, \omega)$. The form of Eq. (S14), however, is more elucidating in the direct comparison of hydrodynamic and local models.

Notably, the above form reinforces an idea of hydrodynamics as acting to effectively modify the neighboring dielectric surrounding,^{S20} rendering it weakly frequency- and l -dependent.

The nonretarded plasmon condition, Eq. (6), is obtained immediately from the poles of the polarizability.

Mie–Lorenz transmission coefficients

In analogy with the *scattering* coefficients $t_l^{\text{TE, TM}}$ discussed in Eq. (3), it is natural to introduce *transmission* coefficients $q_l^{\text{TE, TM, L}}$ for the multipoles transmitted into the interior of the sphere:

$$a_{lm}^{\text{tr}} = q_l^{\text{TE}} a_{l'm'}^{\text{ex}} \delta_{ll'} \delta_{mm'}, \quad b_{lm}^{\text{tr}} = q_l^{\text{TM}} b_{l'm'}^{\text{ex}} \delta_{ll'} \delta_{mm'}, \quad c_{lm}^{\text{tr}} = q_l^{\text{L}} b_{l'm'}^{\text{ex}} \delta_{ll'} \delta_{mm'}. \quad (\text{S15})$$

Note that the longitudinal multipoles, $\mathbf{L}_{lm}^{\text{tr}}$, are excitable only by the ingoing TM multipoles, $\mathbf{N}_{lm}^{\text{ex}}$, but not by ingoing TE multipoles, $\mathbf{M}_{lm}^{\text{ex}}$ – a fact that is directly tied with the invariance of the TE scattering coefficients, t_l^{TE} , under inclusion of longitudinal waves.

Matching of the fields and currents at the boundary of the sphere yields expressions for the transmission coefficients, here given in terms of the scattering coefficients from Eqs. (4):

$$q_l^{\text{TE}} = \frac{t_l^{\text{TE}} h_l^{(1)}(x_D) + j_l(x_D)}{j_l(x_M)}, \quad (\text{S16a})$$

$$q_l^{\text{TM}} = \frac{t_l^{\text{TM}} h_l^{(1)}(x_D) + j_l(x_D)}{j_l(x_M)} \sqrt{\frac{\epsilon_D}{\epsilon_M}}, \quad (\text{S16b})$$

$$q_l^{\text{L}} = l(l+1) \frac{t_l^{\text{TM}} h_l^{(1)}(x_D) + j_l(x_D)}{j_l'(x_{\text{NL}})} \left(\frac{\epsilon_D}{\epsilon_\infty} - \frac{\epsilon_D}{\epsilon_M} \right). \quad (\text{S16c})$$

Longitudinal modes and approximate bulk plasmons

In this section, we consider an approximate criterion for the existence of bulk plasmons in finite or semi-finite metallic structures, and in particular derive Eq. (9).

In analogy with Eqs. (10), the governing equations, Eqs. (1), can be recast solely in terms of the current density (omitting explicit declaration of frequency dependence):

$$(\nabla^2 + k_M^2) \nabla \times \mathbf{J}(\mathbf{r}) = 0, \quad (\text{S17a})$$

$$(\nabla^2 + k_{\text{NL}}^2) \nabla \cdot \mathbf{J}(\mathbf{r}) = 0, \quad (\text{S17b})$$

By performing a spatial Fourier transform, and thus letting $\nabla \rightarrow i\mathbf{q}$, it is clear that Eq. (S17a)

describes the transverse (divergence-free or solenoidal) part of the current, while Eq. (S17b) describes the longitudinal (curl-free or irrotational) part of the current. In extended, infinite structures it is well-known that plasmon resonances above the plasma-frequency are inherently longitudinal. By extension, we introduce the assumption that bulk plasmons in finite structures are also solely longitudinal – thus effectively neglecting interaction with transverse light, thereby constituting the approximation in our following considerations.

To find longitudinal current density solutions, we examine the existence of solutions to Eqs. (S17) which are curl-free (*i.e.*, $\nabla \times \mathbf{J} = 0$). This condition can be satisfied by expressing the current density *via* a scalar velocity potential $\psi(\mathbf{r})$ through $\mathbf{J}(\mathbf{r}) \equiv \nabla\psi(\mathbf{r})$, since the curl of a gradient is always zero. With these assumptions for $\mathbf{J}(\mathbf{r})$ Eq. (S17a) is automatically fulfilled, and Eq. (S17b) reduces to a scalar equation for $\psi(\mathbf{r})$:

$$(\nabla^2 + k_{\text{NL}}^2) \nabla^2 \psi(\mathbf{r}) = 0. \quad (\text{S18})$$

The potential $\psi(\mathbf{r})$ may be expanded in any complete, orthogonal set of functions, say $\{\psi_\nu(\mathbf{r})\}_{\nu=1}^\infty$, such that

$$\psi(\mathbf{r}) = \sum_{\nu} d_{\nu} \psi_{\nu}(\mathbf{r}), \quad (\text{S19})$$

For the solution of Eqs. (S17) in a bounded domain Ω we only need to impose the additional BC that the normal component of the current density vanishes at the boundary $\partial\Omega$, *i.e.*, that $\hat{\mathbf{n}} \cdot \mathbf{J}(\mathbf{r}) = 0$ for $\mathbf{r} \in \partial\Omega$ with $\hat{\mathbf{n}}$ denoting the normal unit vector to $\partial\Omega$. By requiring that each component of the set $\{\psi_\nu(\mathbf{r})\}_{\nu=1}^\infty$ respects this BC, we impose the homogeneous Neumann BC that $\hat{\mathbf{n}} \cdot \nabla\psi_\nu(\mathbf{r})$ vanishes on the boundary $\partial\Omega$. A particularly convenient choice for the generating potentials can be constructed from the solutions of the Helmholtz equation,

$$(\nabla^2 + \kappa_\nu^2) \psi_\nu(\mathbf{r}) = 0, \quad (\text{S20})$$

which constitute a complete, orthogonal set.

Applying the expansion in Eq. (S19), along with property of Eq. (S20), to Eq. (S18) determines the allowed values of the nonlocal propagation constants as $k_{\text{NL}}^2 = \kappa_\nu^2$. Using the definition for k_{NL}^2 we finally obtain the longitudinal resonance frequencies in terms of the eigenvalues, κ_ν^2 , of the Neumann BC Helmholtz equation

$$\omega_\nu(\omega_\nu + i\eta) = \frac{\omega_p^2}{\varepsilon_\infty(\omega_\nu)} + \kappa_\nu^2 \beta_F^2, \quad (\text{S21})$$

which describes the dispersion of hydrodynamic longitudinal modes in an arbitrary geometry.

The problem of determining the dispersion or resonance frequencies of bulk plasmons in any metallic structure is thus simplified to finding the eigenvalues of the Neumann BC Helmholtz equation. A similar result was obtained in Ref. S21 for the case of a metallic sphere.

By rescaling the Helmholtz equation through $\mathbf{x} = \mathbf{r}/R$, where R denotes some geometrically characteristic distance (such as the radius in a spherical geometry), the size-dependence of the resonances can be discerned. Specifically, by this rescaling, the eigenvalues of the dimensionless Helmholtz equation $(\nabla_{\mathbf{x}}^2 + w_{\nu}^2)\psi(\mathbf{x})$ occur at $w_{\nu} = \kappa_{\nu}R$. Consequently, the resonance condition, Eq. (S21), can be cast in terms of the dimensionless eigenvalues w_{ν} :

$$w_{\nu}(\omega_{\nu} + i\eta) = \frac{\omega_{\text{p}}^2}{\varepsilon_{\infty}(\omega_{\nu})} + \frac{w_{\nu}^2 \beta_{\text{F}}^2}{R^2}, \quad (\text{S22})$$

highlighting the size-dependent R^2 -scaling of the bulk plasmons.

For a spherical geometry, the bounded solutions of the Helmholtz equation are $\psi_{lm}(x, \theta, \phi) = j_l(w_{\nu}x)P_l^m(\cos \theta)e^{im\phi}$ as discussed also in the Methods section. Upon application of the Neumann BC at $x = 1$ (corresponding to $r = R$), the eigenvalues are seen to be solutions of the transcendental equation $j_l'(w_n) = 0$.

Quasistatic LDOS and $R = 10$ nm Drude-metal sphere

In this section we provide supporting calculations for the discussions related to Figure 8 of the main text, for a larger sphere of $R = 10$ nm. For a larger sphere inclusion of additional multipoles are required, compared to the 50 multipoles included for $R = 2.5$ nm, to achieve computational convergence (since the multipole resonances are positioned more densely). For computational purposes, it is convenient to work in the electrostatic regime for $l_{\text{max}} > 50$ to avoid numerical instability associated with evaluation of high-order spherical Bessel functions. In this section, we therefore briefly discuss how to derive the electrostatic limit of the electric LDOS enhancement - an approximation which remains representative of the essential physics for an $R = 10$ nm sphere, provided b/R is not too large.

The electrostatic form for the LDOS enhancement can be recovered by considering the $c/\omega R \sim \lambda/R \rightarrow \infty$ asymptotics of the retarded expressions in Eqs. (S9). For this purpose we note the small-argument asymptotic forms $h_l^{(1)}(y) \simeq -i(2l-1)!!y^{-l-1}$ and $\xi_l'(y) \simeq il(2l-1)!!y^{-l-1}$, the connection between the TM Mie-Lorenz coefficients and the multipole polarizability in

Eq. (S13), and finally that the contribution from TE components vanish in the considered limit:^{S22}

$$\frac{\rho_{\perp}^E}{\rho_0^E} = 1 + \frac{3}{2} \frac{1}{k_D^3} \sum_{l=1}^{\infty} (l+1)^2 \frac{1}{b^{2(l+2)}} \operatorname{Im} \left(\frac{\alpha_l}{4\pi} \right), \quad (\text{S23a})$$

$$\frac{\rho_{\parallel}^E}{\rho_0^E} = 1 + \frac{3}{4} \frac{1}{k_D^3} \sum_{l=1}^{\infty} l(l+1) \frac{1}{b^{2(l+2)}} \operatorname{Im} \left(\frac{\alpha_l}{4\pi} \right), \quad (\text{S23b})$$

which, for the orientation-averaged LDOS, $\rho^E = \frac{1}{3}\rho_{\perp}^E + \frac{2}{3}\rho_{\parallel}^E$, yields:

$$\frac{\rho^E}{\rho_0^E} = 1 + \frac{1}{2} \frac{1}{k_D^3} \sum_{l=1}^{\infty} (2l+1)(l+1) \frac{1}{b^{2(l+2)}} \operatorname{Im} \left(\frac{\alpha_l}{4\pi} \right). \quad (\text{S23c})$$

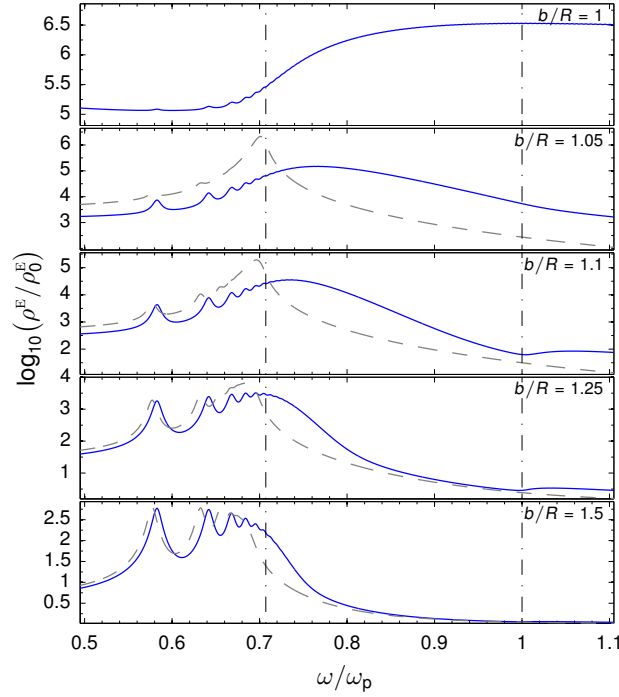


Figure S1: Logarithm of the normalized LDOS in hydrodynamic and LRA treatments, in full blue and dashed gray, respectively, for a $R = 10$ nm sphere with Drude-metal composition ($\omega_p = 10$ eV, $\eta = 0.1$ eV, and $\varepsilon_{\infty} = \varepsilon_D = 1$). Calculated in the electrostatic approximation *via* Eq. (S23c) with 250 multipoles, *i.e.*, with $l_{\max} = 250$.

In Figure S1 we depict the results of an electrostatic calculation of the LDOS for an $R = 10$ nm Drude-metal sphere. The electrostatic calculation gives representative results in the considered parameter-space, apart from a missing redshift (~ 19 meV for the dipole resonance) due to radiation reaction for both LRA and hydrodynamic treatments. Interestingly, the results show

that hydrodynamics predicts distinct differences from the LRA not only for small spheres, but also for larger spheres, provided the probe-to-surface separation, $b - R$, is sufficiently small. The disparity arises due to the lifting of the singular pile-up of modes predicted by the LRA near the planar-interface resonance at $\omega_p/\sqrt{2}$. As the probe approaches the surface additional multipoles are excited, until, at the surface, all multipoles in the spectral vicinity contribute. This suggests an alternate approach for examining the presence of nonlocal effects, even in large structures: consider the LDOS enhancement spectrally for short probe-to-surface distances. Both a significant broadening and a spectral shift of the peak LDOS enhancement is predicted by hydrodynamics compared to the LRA.

Finally, for intermediate separations, $b/R = 1.1$ and $b/R = 1.25$, the existence of low-order bulk plasmons produce a shoulder above ω_p – for shorter probe-to-surface separations the excitation of high-order multipole LSPs, existing above ω_p , overshadow this effect.

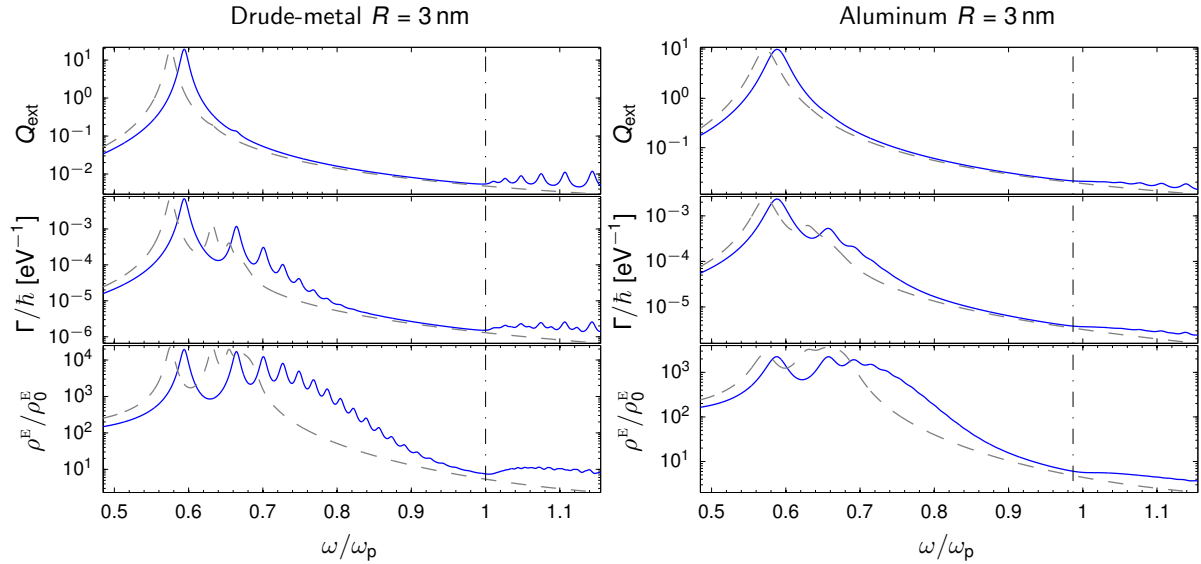


Figure S2: Extinction efficiency, EELS probability, and normalized LDOS. Setup parameters are identical to those in Figure 7, but with radius $R = 3$ nm and $b/R = 1.5$ (corresponding to maintaining the same probe-to-surface separation, 1.5 nm, as in Figure 7). Hydrodynamics and LRA in full blue and dashed gray, respectively, as before. Screened plasma frequency in dash-dotted black.

Additional spectra for larger nanospheres and for silver

To support the discussion related to Figure 7 of the main text, we here give additional calculations for a Drude-metal and aluminum in nanospheres with $R = 3$ nm in Figure S2. Clearly, to observe

the multipole features discussed in the main text the particles under consideration should be rather small – however, even at $R = 3$ nm at least 3 multipole resonances are discernible for aluminum.

Lastly, we give calculations for silver in $R = 1.5$ nm and $R = 3$ nm nanospheres in Figure S3. For the case of silver the effects of higher-order multipoles are entirely suppressed by the strongly dispersive background due to the bound response.

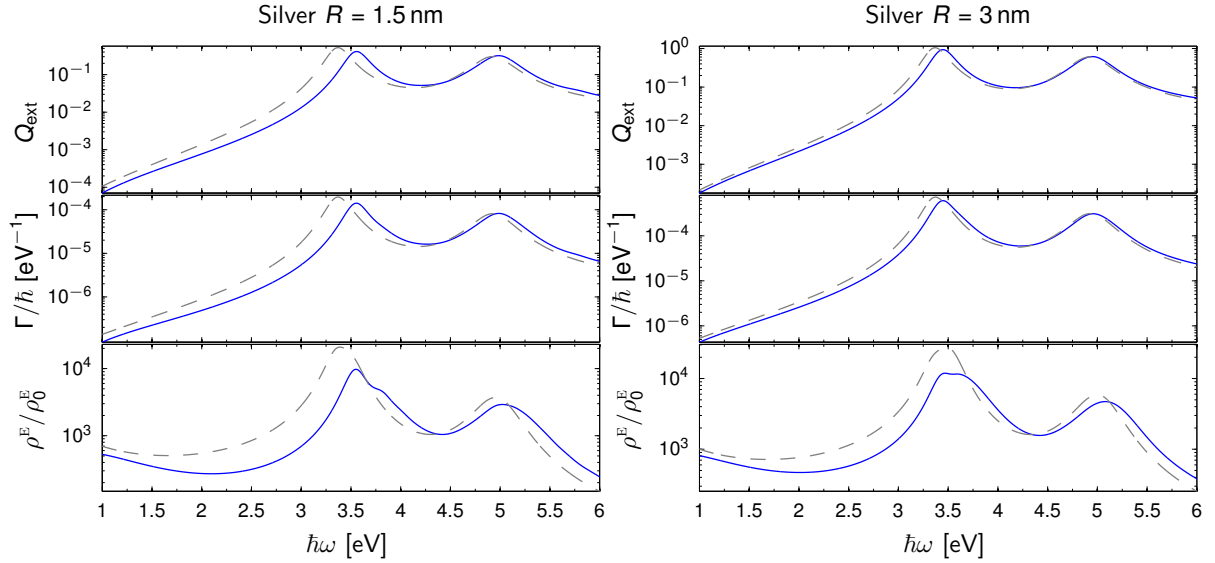


Figure S3: Extinction efficiency, EELS probability, and normalized LDOS. Setup parameters are identical to those in Figure S2, but for silver of two different radius. Probe-to-surface separation is $b - R = 1.5$ nm in both cases. Hydrodynamics and LRA in full blue and dashed gray, respectively, as before. Experimental data for silver taken from Ref. S3, with free-electron parameters $\omega_p = 9.01$ eV and $\eta = 48$ meV.

References

- (S1) Kittel, C. *Introduction to Solid State Physics*; Jon Wiley & Sons, 2005.
- (S2) Rakić, A. D. Algorithm for the Determination of Intrinsic Optical Constants of Metal Films: Application to Aluminum. *Appl. Opt.* **1995**, *34*, 4755–4767.
- (S3) Rakić, A. D.; Djurišić, A. B.; Elazar, J. M.; Majewski, M. L. Optical Properties of Metallic Films for Vertical-Cavity Optoelectronic Devices. *Appl. Optics* **1998**, *37*, 5271.
- (S4) Stratton, J. A. *Electromagnetic Theory*; McGraw-Hill Book Company: New York, 1941.

- (S5) Bohren, C. F.; Huffman, D. R. *Absorption and Scattering of Light by Small Particles*; John Wiley & Sons: New York, 1983.
- (S6) Jackson, J. D. *Classical Electrodynamics*, 3rd ed.; John Wiley & Sons: New York, 1999.
- (S7) García de Abajo, F. J. Optical Excitations in Electron Microscopy. *Rev. Mod. Phys.* **2010**, *82*, 209–275.
- (S8) García de Abajo, F. J. Relativistic Energy Loss and Induced Photon Emission in the Interaction of a Dielectric Sphere with an External Electron Beam. *Phys. Rev. B* **1999**, *59*, 3095–3107.
- (S9) Abramowitz, M.; Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*; Dover: New York, 1972.
- (S10) Kerker, M.; Wang, D.-S.; Chew, H. Surface Enhanced Raman Scattering (SERS) by Molecules Adsorbed at Spherical Particles: Errata. *Appl. Opt.* **1980**, *19*, 4159–4174.
- (S11) Ruppin, R. Decay of an Excited Molecule near a Small Sphere. *J. Chem. Phys.* **1982**, *76*, 1681–1684.
- (S12) Chew, H. Transition Rates of Atoms Near Spherical Surfaces. *J. Chem. Phys.* **1987**, *87*, 1355–1360.
- (S13) Dung, H. T.; Knöll, L.; Welsch, D.-G. Decay of an Excited Atom near an Absorbing Microsphere. *Phys. Rev. A* **2001**, *64*, 013804.
- (S14) Vos, W. L.; Koenderink, A. F.; Nikolaev, I. S. Orientation-Dependent Spontaneous Emission Rates of a Two-Level Quantum Emitter in any Nanophotonic Environment. *Phys. Rev. A* **2009**, *80*, 053802.
- (S15) Datsyuk, V. V. Ultimate Enhancement of the Local Density of Electromagnetic States Outside an Absorbing Sphere. *Phys. Rev. A* **2007**, *75*, 043820.
- (S16) Olver, F.; Lozier, D.; Boisvert, R.; Clark, C. *NIST Handbook of Mathematical Functions*; National Institute of Standards and Technology, 2010.
- (S17) Gradshteyn, I.; Ryzhik, I. *Table of Integrals, Series, and Products*, 7th ed.; Elsevier Academic Press, 2007.

- (S18) David, C.; García de Abajo, F. J. Spatial Nonlocality in the Optical Response of Metal Nanoparticles. *J. Phys. Chem. C* **2012**, *115*, 19470–19475.
- (S19) Fuchs, R.; Claro, F. Multipolar Response of Small Metallic Spheres: Nonlocal Theory. *Phys. Rev. B* **1987**, *35*, 3722–3726.
- (S20) Raza, S.; Yan, W.; Stenger, N.; Wubs, M.; Mortensen, N. A. Blueshift of the Surface Plasmon Resonance in Silver Nanoparticles: Substrate Effects. *Opt. Express* **2013**, *21*, 27344–27355.
- (S21) Barberán, N.; Bausells, J. Plasmon Excitation in Metallic Spheres. *Phys. Rev. B* **1985**, *31*, 6354–6359.
- (S22) Vielma, J.; Leung, P. Nonlocal Optical Effects on the Fluorescence and Decay Rates for Admolecules at a Metallic Nanoparticle. *J. Chem. Phys.* **2007**, *126*, 194704.

BIBLIOGRAPHY

All journal entries are abbreviated according to CASSI conventions.

- [1] A. Ruivo, C. Gomes, A. Lima, M.L. Botelho, R. Melo, A. Belchior, and A.P. de Matos, *Gold nanoparticles in ancient and contemporary ruby glass*, *C. R. Phys.* **9**, 649 (2008).
- [2] P. Colomban, *The use of metal nanoparticles to produce yellow, red and iridescent colour, from bronze age to present times in lustre pottery and glass: solid state chemistry, spectroscopy and nanostructure*, *J. Nano Res.* **8**, 109 (2009).
- [3] G. Mie, *Articles on the optical characteristics of turbid tubes, especially colloidal metal solutions*. *Ann. Phys.* **25**, 377 (1908), originally 'Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen'.
- [4] H. Siedentopf and R. Zsigmondy, *On visualization and size-determination of ultra-microscopic particles, with particular applications to gold ruby glasses*. *Ann. Phys.* **315**, 1 (1902), originally 'Über Sichtbarmachung und Größenbestimmung ultramikroskopischer Teilchen, mit besonderer Anwendung auf Goldrubingläser'.
- [5] T. Mappes, N. Jahr, A. Csaki, N. Vogler, J. Popp, and W. Fritzsche, *The invention of immersion ultramicroscopy in 1912 – the birth of nanotechnology?* *Angew. Chem.* **51**, 11208 (2012).
- [6] L. Tonks and I. Langmuir, *Oscillations in ionized gases*, *Phys. Rev.* **33**, 195 (1929).
- [7] D. Pines and D. Bohm, *A collective description of electron interactions: II. collective vs individual particle aspects of the interactions*, *Phys. Rev.* **85**, 338 (1952).
- [8] R.H. Ritchie, *Plasma losses by fast electrons in thin films*, *Phys. Rev.* **106**, 874 (1957).
- [9] J.M. Pitarke, V.M. Silkin, E.V. Chulkov, and P.M. Echenique, *Theory of surface plasmons and surface-plasmon polaritons*, *Rep. Prog. Phys.* **70**, 1 (2007).
- [10] S.A. Maier, *Plasmonics: Fundamentals and Applications* (Springer, 2007).
- [11] A. Baev, P.N. Prasad, H. Ågren, M. Samoć, and M. Wegener, *Metaphotonics: an emerging field with opportunities and challenges*, *Phys. Rep.* **594**, 1 (2015).
- [12] M.I. Stockman, *Nanoplasmonics: the physics behind the applications*, *Phys. Today* **64**, 39 (2011).
- [13] M.I. Stockman, *Nanoplasmonics: past, present, and glimpse into future*, *Opt. Express* **19**, 22029 (2011).
- [14] J.A. Schuller, E.S. Barnard, W. Cai, Y.C. Jun, J.S. White, and M.L. Brongersma, *Plasmonics for extreme light concentration and manipulation*, *Nat. Mater.* **9**, 193 (2010).
- [15] D.K. Gramotnev and S.I. Bozhevolnyi, *Plasmonics beyond the diffraction limit*, *Nat. Photonics* **4**, 83 (2010).
- [16] C.F. Bohren, *How can a particle absorb more than the light incident on it?* *Am. J. Phys.* **51**, 323 (1983).

- [17] K. Kneipp, Y. Wang, H. Kneipp, L.T. Perelman, I. Itzkan, R.R. Dasari, and M.S. Feld, *Single molecule detection using surface-enhanced raman scattering (SERS)*, *Phys. Rev. Lett.* **78**, 1667 (1997).
- [18] N. Engheta and R.W. Ziolkowski, eds., *Metamaterials: Physics and Engineering Explorations* (Wiley–IEEE Press, 2006).
- [19] L. Novotny and N. van Hulst, *Antennas for light*, *Nat. Photonics* **5**, 83 (2011).
- [20] T. Ebbesen, H.J. Lezec, H.F. Ghaemi, T. Thio, and P.A. Wolff, *Extraordinary optical transmission through sub-wavelength hole arrays*, *Nature* **391**, 667 (1998).
- [21] J. Henzie, J. Lee, M.H. Lee, W. Hasan, and T.W. Odom, *Nanofabrication of plasmonic structures*, *Annu. Rev. Phys. Chem.* **60**, 147 (2009).
- [22] F.J. García de Abajo, *Optical excitations in electron microscopy*, *Rev. Mod. Phys.* **82**, 209 (2010).
- [23] X. Chen, H.-R. Park, M. Pelton, X. Piao, N.C. Lindquist, H. Im, Y.J. Kim, J.S. Ahn, N. Park, D.-S. Kim, and S.-H. Oh, *Atomic layer lithography of wafer-scale nanogap arrays for extreme confinement of electromagnetic waves*, *Nat. Commun.* **4**, 2361 (2013).
- [24] J. Kern, S. Grossmann, T. Tarakina, N.V Hackel, M. Emmerling, M. Kamp, J.-S. Huang, P. Biagioni, J.C. Prangsma, and B. Hecht, *Atomic-scale confinement of resonant optical fields*, *Nano Lett.* **12**, 5504 (2012).
- [25] J.A. Scholl, A.L. Koh, and J.A. Dionne, *Quantum plasmon resonances of individual metallic nanoparticles*, *Nature* **483**, 421 (2012).
- [26] S. Raza, N. Stenger, S. Kadkhodazadeh, S.V. Fischer, N. Kostesha, A.-P. Jauho, A. Burrows, M. Wubs, and N.A. Mortensen, *Blueshift of the surface plasmon resonance in silver nanoparticles studied with EELS*, *Nanophotonics* **2**, 131 (2013).
- [27] I.V. Tokatly, *Time-dependent density functional theory for many-electron systems interacting with cavity photons*, *Phys. Rev. Lett.* **110**, 233001 (2013).
- [28] C. Pellegrini, J. Flick, I.V. Tokatly, H. Appel, and A. Rubio, *Optimized effective potential for quantum electrodynamical time-dependent density functional theory*, *Phys. Rev. Lett.* **115**, 093001 (2015).
- [29] M.S. Tame, K.R. McEnery, S.K. Özdemir, J. Lee, S.A. Maier, and M.S. Kim, *Quantum plasmonics*, *Nature Phys.* **9**, 329 (2013).
- [30] S. Kumar, N.I. Møller, A. Huck, and U.L. Andersen, *Generation and controlled routing of single plasmons on a chip*, *Nano Lett.* **14**, 663 (2014).
- [31] C.C. Gerry and P.L. Knight, *Introductory Quantum Optics*, 1st ed. (Cambridge University Press, 2005).
- [32] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics* (Oxford University Press, 2004).
- [33] T. Ando, A.B. Fowler, and F. Stern, *Electronic properties of two-dimensional systems*, *Rev. Mod. Phys.* **54**, 437 (1982).
- [34] K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, and A.A. Firsov, *Electric field effect in atomically thin carbon films*, *Science* **306**, 666 (2004).

- [35] K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, M.I. Katsnelson, I.V. Grigorieva, S.V. Dubonos, and A.A. Firsov, *Two-dimensional gas of massless dirac fermions in graphene*, *Nature* **438**, 197 (2005).
- [36] K.S. Novoselov, D. Jiang, F. Schedin, T.J. Booth, V.V. Khotkevich, S.V. Morozov, and A.K. Geim, *Two-dimensional atomic crystals*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10451 (2005).
- [37] M. Jablan, H. Buljan, and M. Soljačić, *Plasmonics in graphene at infrared frequencies*, *Phys. Rev. B* **80**, 245435 (2009).
- [38] A. Vakil and N. Engheta, *Transformation optics using graphene*, *Science* **332**, 1291 (2011).
- [39] F.H.L. Koppens, D.E. Chang, and F.J. García de Abajo, *Graphene plasmonics: a platform for strong light–matter interactions*, *Nano Lett.* **11**, 3370 (2011).
- [40] A.N. Grigorenko, M. Polini, and K.S. Novoselov, *Graphene plasmonics*, *Nat. Photonics* **6**, 749 (2012).
- [41] M. Jablan, M. Soljačić, and H. Buljan, *Plasmons in graphene: fundamental properties and potential applications*, *Proc. IEEE* **101**, 1689 (2013).
- [42] Y.V. Bludov, A. Ferreira, N.M.R. Peres, and M.I. Vasilevskiy, *A primer on surface plasmon-polaritons in graphene*, *Int. J. Mod. Phys. B* **27**, 1341001 (2013).
- [43] F. Xia, H. Wang, D. Xiao, M. Dubey, and A. Ramasubramaniam, *Two-dimensional material nanophotonics*, *Nat. Photonics* **8**, 899 (2014).
- [44] F.J. García de Abajo, *Graphene plasmonics: challenges and opportunities*, *ACS Photonics* **1**, 135 (2014).
- [45] T. Low and P. Avouris, *Graphene plasmonics for terahertz to mid-infrared applications*, *ACS Nano* **8**, 1086 (2014).
- [46] T. Stauber, *Plasmonics in Dirac systems: from graphene to topological insulators*, *J. Phys.: Condens. Matter* **26**, 123201 (2014).
- [47] J.D. Jackson, *Classical Electrodynamics*, 3rd ed. (John Wiley & Sons, 1999).
- [48] C.-T. Tai, *Dyadic Green Functions in Electromagnetic Theory*, 2nd ed., Series on Electromagnetic Waves (IEEE Press, 1993).
- [49] M. Wubs, L.G. Suttorp, and A. Lagendijk, *Spontaneous-emission rates in finite photonic crystals of plane scatterers*, *Phys. Rev. E* **69**, 016616 (2004).
- [50] J.D. Joannopoulos, S.E. Johnson, J.N. Winn, and R.D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University Press, 2008).
- [51] F. Wijnands, J.B. Pendry, F.J. García-Vidal, P.M. Bell, P.J. Roberts, and L. Martín Moreno, *Green's functions for Maxwell's equations: application to spontaneous emission*, *Opt. Quantum Electron.* **29**, 199 (1997).
- [52] K.M. Lee, P.T. Leung, and K.M. Pang, *Dyadic formulation of morphology-dependent resonances. I. Completeness relation*, *J. Opt. Soc. Am. B* **16**, 1409 (1999).
- [53] R.-G. Ge, P.T. Kristensen, J.F. Young, and S. Hughes, *Quasinormal mode approach to modelling light emission and propagation in nanoplasmonics*, *New J. Phys.* **16**, 113048 (2014).

- [54] P.J. Feibelman, *Surface electromagnetic fields*, *Prog. Surf. Sci.* **12**, 287 (1982).
- [55] L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, 2012).
- [56] N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Saunders College Publishing, 1976).
- [57] G. Grosso and G.P. Parravicini, *Solid State Physics*, 2nd ed. (Academic Press, 2014).
- [58] P.B. Johnson and R.W. Christy, *Optical constants of the noble metals*, *Phys. Rev. B* **6**, 4370 (1972).
- [59] A.D. Rakić, *Algorithm for the determination of intrinsic optical constants of metal films: Application to aluminum*, *Appl. Opt.* **34**, 4755 (1995).
- [60] A.D. Rakić, A.B. Djurišić, J.M. Elazar, and M.L. Majewski, *Optical properties of metallic films for vertical-cavity optoelectronic devices*. *Appl. Optics* **37**, 5271 (1998).
- [61] J.M. Luther, P.K. Jain, T. Ewers, and A.P. Alivisatos, *Localized surface plasmon resonances arising from free carriers in doped quantum dots*, *Nat. Mater.* **10**, 361 (2011).
- [62] A.M. Schimpf, N. Thakkar, C.E. Gunthardt, D.J. Masiello, and D.R. Gamelin, *Charge-tunable quantum plasmons in colloidal semiconductor nanocrystals*, *ACS Nano* **8**, 1065 (2014).
- [63] T. Christensen, W. Yan, S. Raza, A.-P. Jauho, N.A. Mortensen, and M. Wubs, *Nonlocal response of metallic nanospheres probed by light, electrons, and atoms*, *ACS Nano* **8**, 1745 (2014).
- [64] G. Ruthemann, *Discrete energy loss of fast electrons in passage through thin-films*, *Ann. Phys.* **437**, 113 (1948), originally 'Diskrete Energieverluste mittelschneller Elektronen beim Durchgang durch dünne Folien'.
- [65] L. Marton, J.A. Simpson, H.A. Fowler, and N. Swanson, *Plural scattering of 20-keV electrons in aluminum*, *Phys. Rev.* **126**, 182 (1962).
- [66] C.H. Chen, *Plasmon dispersion in single-crystal magnesium*, *J. Phys. C: Solid State Phys.* **9**, L321 (1976).
- [67] R.W. Alexander, G.S. Kovener, and R.J. Bell, *Dispersion curves for surface electromagnetic waves with damping*, *Phys. Rev. Lett.* **32**, 154 (1974).
- [68] A. Archambault, A. Teperik, F. Marquier, and J.-J. Greffet, *Surface plasmon Fourier optics*, *Phys. Rev. B* **79**, 195414 (2009).
- [69] R. Fuchs and F. Claro, *Multipolar response of small metallic spheres: Nonlocal theory*, *Phys. Rev. B* **35**, 3722 (1987).
- [70] V. Myroshnychenko, J. Rodríguez-Fernández, I. Pastoriza-Santos, A.M. Funston, C. Novo, P. Mulvaney, L.M. Liz-Marzán, and F.J. García de Abajo, *Modelling the optical response of gold nanoparticles*, *Chem. Soc. Rev.* **37**, 1792 (2008).
- [71] T. Christensen, A.-P. Jauho, M. Wubs, and N.A. Mortensen, *Localized plasmons in graphene-coated nanospheres*, *Phys. Rev. B* **91**, 125414 (2015).

- [72] R.M. Cole, J.J. Baumberg, F.J. García de Abajo, S. Mahajan, M. Abdelsalam, and P.N. Bartlett, *Understanding plasmons in nanoscale voids*, **Nano Lett.** **7**, 2094 (2007).
- [73] S.P. Apell, P.M. Echenique, and R.H. Ritchie, *Sum rules for surface plasmon frequencies*, **Ultramicroscopy** **65**, 53 (1996).
- [74] Z. Ruan and S. Fan, *Superscattering of light from subwavelength nanostructures*, **Phys. Rev. Lett.** **105**, 013901 (2010).
- [75] C.L.C. Smith, N. Stenger, A. Kristensen, N.A. Mortensen, and S.I. Bozhevolnyi, *Gap and channeled plasmons in tapered grooves: a review*, **Nanoscale** **7**, 9355 (2015).
- [76] F.J. García de Abajo, *Colloquium: Light scattering by particle and hole arrays*, **Rev. Mod. Phys.** **79**, 1267 (2007).
- [77] A. Rusina, M. Durach, and M.I. Stockman, *Theory of spoof plasmons in real metals*, **Appl. Phys. A** **100**, 375 (2010).
- [78] R. Ruppin, *Surface modes of two spheres*, **Phys. Rev. B** **26**, 3440 (3441).
- [79] P. Nordlander, C. Oubre, E. Prodan, K. Li, and M.I. Stockman, *Plasmon hybridization in nanoparticle dimers*, **Nano Lett.** **4**, 899 (2004).
- [80] I. Romero, J. Aizpurua, G.W. Bryant, and F.J. García de Abajo, *Plasmons in nearly touching metallic nanoparticles: singular response in the limit of touching dimers*, **Opt. Express** **14**, 9988 (2006).
- [81] A.I. Fernández-Domínguez, A. Wiener, F.J. García-Vidal, S.A. Maier, and J.B. Pendry, *Transformation-optics description of nonlocal effects in plasmonic nanostructures*, **Phys. Rev. Lett.** **108**, 106802 (2012).
- [82] G. Toscano, S. Raza, A.-P. Jauho, N.A. Mortensen, and M. Wubs, *Modified field enhancement and extinction in plasmonic nanowire dimers due to nonlocal response*, **Opt. Express** **20**, 4176 (2012).
- [83] R. Esteban, A.G. Borisov, P. Nordlander, and J. Aizpurua, *Bridging quantum and classical plasmonics with a quantum-corrected model*, **Nat. Commun.** **3**, 825 (2012).
- [84] 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).
- [85] K. Andersen, K. Jensen, N.A. Mortensen, and K.S. Thygesen, *Visualizing hybridized quantum plasmons in coupled nanowires: from classical to tunneling regime*, **Phys. Rev. B** **87**, 235433 (2013).
- [86] N.A. Mortensen, S. Raza, M. Wubs, and S.I. Bozhevolnyi, *A generalized nonlocal optical response theory for plasmonic nanostructures*, **Nat. Commun.** **5**, 3809 (2014).
- [87] S. Raza, M. Wubs, S.I. Bozhevolnyi, and N.A. Mortensen, *Nonlocal study of ultimate plasmon hybridization*, **Opt. Lett.** **40**, 839 (2015).
- [88] U. Leonhardt and T.G. Philbin, *Transformation optics and the geometry of light*, **Prog. Opt.** **53**, 69 (2009).

- [89] A.I. Fernández-Domínguez, S.A. Maier, and J.B. Pendry, *Collection and concentration of light by touching spheres: a transformation optics approach*, *Phys. Rev. Lett.* **105**, 266807 (2010).
- [90] A.V. Lavrinenko, J. Lægsgaard, N. Gregersen, F. Schmidt, and T. Søndergaard, *Numerical Methods in Photonics*, 1st ed. (CRC Press, 2014).
- [91] A. Taflove and S.C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. (Artech House Inc., 2005).
- [92] J.-M. Jin, *The Finite Element Method in Electromagnetics*, 3rd ed. (Wiley-IEEE Press, 2014).
- [93] T. Søndergaard, *Modeling of plasmonic nanostructures: Green's function integral equation methods*, *Phys. Status Solidi B* **244**, 3448 (2007).
- [94] B.T. Draine and P.J. Flatau, *Discrete-dipole approximation for scattering calculations*, *JOSA A* **11**, 1491 (1994).
- [95] F.J. García de Abajo and A. Howie, *Retarded field calculation of electron energy loss in inhomogeneous dielectrics*, *Phys. Rev. B* **65**, 115418 (2002).
- [96] U. Hohenester and A. Trügler, *MNPBEM – a Matlab toolbox for the simulation of plasmonic nanoparticles*, *Comput. Phys. Commun.* **183**, 370 (2012).
- [97] U. Hohenester, *Simulating electron energy loss spectroscopy with the MNPBEM toolbox*, *Comput. Phys. Commun.* **185**, 1177 (2014).
- [98] C.F. Bohren and D.R. Huffman, *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, New York, 1983).
- [99] J.W. Strutt (Lord Rayleigh), *On the light from the sky, its polarization and colour*, *Philos. Mag.* **37**, 388 (1871).
- [100] R. Carminati, J.-J. Greffet, C. Henkel, and J.M. Vigoureux, *Radiative and non-radiative decay of a single molecule close to a metallic nanoparticle*, *Opt. Commun.* **261**, 368 (2006).
- [101] H.C. van de Hulst, *Light scattering by small particles* (Dover Publications, 1981).
- [102] J.-J. Greffet and R. Carminati, *Image formation in near-field optics*, *Prog. Surf. Sci.* **56**, 133 (1997).
- [103] J.-C. Weeber, J.R. Krenn, A. Dereux, B. Lamprecht, Y. Lacroute, and J.P. Goudonnet, *Near-field observation of surface plasmon polariton propagation on thin metal stripes*, *Phys. Rev. B* **64**, 045411 (2001).
- [104] A.V. Zayats, I.I. Smolyaninov, and A.A. Maradudin, *Nano-optics of surface plasmon polaritons*, *Phys. Rep.* **408**, 131 (2005).
- [105] S.I. Bozhevolnyi, V.S. Volkov, E. Devaux, J.-Y. Laluet, and T.W. Ebbesen, *Channel plasmon subwavelength waveguide components including interferometers and ring resonators*, *Nature* **440**, 508 (2006).
- [106] J. Chen, M. Badioli, P. Alonso-González, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenović, A. Centeno, A. Pesquera, P. Godignon, A.Z. Elorza, N. Camara, F.J. García de Abajo, R. Hillenbrand, and F.H.L. Koppens, *Optical nano-imaging of gate-tunable graphene plasmons*, *Nature* **487**, 77 (2012).

- [107] Z. Fei, A.S. Rodin, G.O. Andreev, W. Bao, A.S. McLeod, M. Wagner, L.M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, M.M. Fogler, A.H. Castro Neto, C.N. Lau, F. Keilmann, and D.N. Basov, *Gate-tuning of graphene plasmons revealed by infrared nano-imaging*, *Nature* **487**, 82 (2012).
- [108] J. Chen, M.L. Nesterov, A.Y. Nikitin, S. Thongrattanasiri, P. Alonso-González, T.M. Slipchenko, F. Speck, O. Markus, T. Seyller, I. Crassee, F.H.L. Koppens, L. Martin-Moreno, F.J. García de Abajo, A. Kuzmenko, and R. Hillenbrand, *Strong plasmon reflection at nanometer-size gaps in monolayer graphene on SiC*, *Nano Lett.* **13**, 6210 (2003).
- [109] P. Alonso-González, A.Y. Nikitin, F. Golmar, A. Centeno, A. Pesquera, S. Vélez, J. Chen, G. Navickaite, F.H.L. Koppens, A. Zurutuza, F. Casanova, L.E. Hueso, and R. Hillenbrand, *Controlling graphene plasmons with resonant metal antennas and spatial conductivity patterns*, *Science* **344**, 1369 (2014).
- [110] K.A. Willets, *Super-resolution imaging of SERS hot spots*, *Chem. Soc. Rev.* **43**, 3854 (2014).
- [111] K. Joulain, R. Carminati, J.-P. Mulet, and J.-J. Greffet, *Definition and measurement of the local density of electromagnetic states close to an interface*, *Phys. Rev. B* **68**, 245405 (2003).
- [112] V. Weisskopf and E. Wigner, *Calculation of the natural linewidth in the Dirac theory of light*, *Z. Phys.* **63**, 54 (1930), originally 'Berechnung der natürlichen Linienbreite auf Grund der Diracschen Lichttheorie'.
- [113] K.H. Drexhage, *Influence of a dielectric interface on fluorescence decay time*, *J. Lumin.* **1-2**, 693 (1970).
- [114] E. Dulkeith, M. Ringle, T.A. Klar, J. Feldmann, A. Javier Muñoz, and W.J. Parak, *Gold nanoparticles quench fluorescence by phase induced radiative rate suppression*, *Nano Lett.* **5**, 585 (2005).
- [115] P. Anger, P. Bharadwaj, and L. Novotny, *Enhancement and quenching of single-molecule fluorescence*, *Phys. Rev. Lett.* **96**, 113002 (2006).
- [116] R.F. Egerton, *Electron energy-loss spectroscopy in the TEM*, *Rep. Prog. Phys.* **72**, 016502 (2009).
- [117] M. Kociak and F.J. García de Abajo, *Nanoscale mapping of plasmons, photons, and excitons*, *MRS Bull.* **37**, 39 (2012).
- [118] F.J. García de Abajo and M. Kociak, *Probing the photonic local density of states with electron energy loss spectroscopy*, *Phys. Rev. Lett.* **100**, 106804 (2008).
- [119] U. Hohenester, H. Ditlbacher, and J. R. Krenn, *Electron-energy-loss spectra of plasmonic nanoparticles*, *Phys. Rev. Lett.* **103**, 106801 (2009).
- [120] N.D. Lang and W. Kohn, *Theory of metal surfaces: charge density and surface energy*, *Phys. Rev. B* **1**, 4555 (1970).
- [121] Z.F. Öztürk, S. Xiao, M. Yan, M. Wubs, A.-P. Jauho, and N.A. Mortensen, *Field enhancement at metallic interfaces due to quantum confinement*, *J. Nanophotonics* **5**, 051602 (2011).
- [122] M.L. Brongersma, N.J. Halas, and P. Nordlander, *Plasmon-induced hot carrier science and technology*, *Nat. Nanotechnol.* **10**, 25 (2015).

- [123] U. Kreibig and L. Genzel, *Optical absorption of small metallic particles*, *Surf. Sci.* **156**, 678 (1985).
- [124] J.B. Khurgin, *Graphene – a rather ordinary nonlinear optical material*, *Appl. Phys. Lett.* **104**, 161116 (2014).
- [125] W.P. Halperin, *Quantum size effects in metal particles*, *Rev. Mod. Phys.* **58**, 533 (1986).
- [126] F.J. García de Abajo, *Microscopy: plasmons go quantum*, *Nature* **483**, 417 (2012).
- [127] E. Townsend and G.W. Bryant, *Plasmonic properties of metallic nanoparticles: the effects of size quantization*, *Nano Lett.* **12**, 429 (2011).
- [128] E. Townsend and G.W. Bryant, *Which resonances in small metallic nanoparticles are plasmonic?* *J. Opt.* **16**, 114022 (2014).
- [129] S. Thongrattanasiri, A. Manjavacas, and F.J. García de Abajo, *Quantum finite-size effects in graphene plasmons*, *ACS Nano* **6**, 1766 (2012).
- [130] P. Zhang, J. Feist, A. Rubio, P. García-González, and F.J. García-Vidal, *Ab initio nanoplasmonics: The impact of atomic structure*, *Phys. Rev. B* **90**, 161407 (2014).
- [131] O. Keller, M. Xiao, and S.I. Bozhevolnyi, *Optical diamagnetic polarizability of a mesoscopic metallic sphere: transverse self-field approach*, *Opt. Commun.* **102**, 238 (1993).
- [132] P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, *Phys. Rev.* **136**, 864 (1964).
- [133] E. Runge and E.K.U. Gross, *Density-functional theory for time-dependent systems*, *Phys. Rev. Lett.* **52**, 997 (1984).
- [134] R.O. Jones, *Density functional theory: its origins, rise to prominence, and future*, *Rev. Mod. Phys.* **87**, 897 (2015).
- [135] M.A.L. Marques, C.A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E.K.U. Gross, eds., *Time-Dependent Density Functional Theory*, Lecture Notes in Physics (Springer, 2006).
- [136] W. Yan, M. Wubs, and N.A. Mortensen, *Projected dipole model for quantum plasmonics*, *Phys. Rev. Lett.* **115**, 137403 (2015).
- [137] K.-D. Tsuei, E.W. Plummer, A. Liebsch, E. Pehlke, K. Kempa, and P. Bakshi, *The normal modes at the surface of simple metals*, *Surf. Sci.* **247**, 302 (1991).
- [138] P. Apell and A. Ljungbert, *A general non-local theory for the electromagnetic response of a small metal particle*, *Physica Scripta* **26**, 113 (1982).
- [139] P. Apell and A. Ljungbert, *Red shift of surface plasmons in small metal particles*, *Solid State Commun.* **44**, 1367 (1982).
- [140] K. Andersen, K.W. Jacobsen, and K.S. Thygesen, *Spatially resolved quantum plasmon modes in metallic nano-films from first-principles*, *Phys. Rev. B* **86**, 245129 (2012).
- [141] F. Bloch, *Stopping power of atoms with several electrons*, *Z. Phys. A* **81**, 363 (1933), originally ‘Bremsvermögen von Atomen mit mehreren Elektronen’.
- [142] A.D. Boardman, *Electromagnetic Surface Modes* (John Wiley & Sons, 1982) Chap. Hydrodynamic Theory of Plasmon-Polaritons on Plane Surfaces.

- [143] A. Eguiluz and J.J. Quinn, *Hydrodynamic model for surface plasmons in metals and degenerate semiconductors*, *Phys. Rev. B* **14**, 1347 (1976).
- [144] V.L. Lignères and E.A. Carter, *Handbook of Materials Modeling* (Springer, 2005) Chap. An Introduction to Orbital-Free Density Functional Theory.
- [145] R. Balescu, *Equilibrium and nonequilibrium statistical mechanics* (John Wiley & Sons, 1975).
- [146] I. Tokatly and O. Pankratov, *Hydrodynamic theory of an electron gas*, *Phys. Rev. B* **60**, 15550 (1999).
- [147] J.E. Sipe, V.C.Y. So, M. Fukui, and G.I. Stegeman, *Analysis of second-harmonic generation at metal surfaces*, *Phys. Rev. B* **21**, 4389 (1980).
- [148] P. Ginzburg, A. Krasavin, Y. Sonnefraud, A. Murphy, R.J. Pollard, S.A. Maier, and A.V. Zayats, *Nonlinearly coupled localized plasmon resonances: resonant second-harmonic generation*, *Phys. Rev. B* **86**, 085422 (2012).
- [149] C. Ciraci, E. Poutrina, M. Scalora, and D.R. Smith, *Second-harmonic generation in metallic nanoparticles: clarification of the role of the surface*, *Phys. Rev. B* **86**, 115451 (2012).
- [150] J.D. Jackson, *Classical Electrodynamics*, 1st ed. (John Wiley & Sons, 1962), see Section 10.9. ‘High-Frequency Plasma Oscillations’ [unavailable in later editions].
- [151] A.L. Fetter, *Electrodynamics of a layered electron gas. I. Single layer*, *Ann. Phys.* **81**, 367 (1973).
- [152] P. Halevi, *Hydrodynamic model for the degenerate free-electron gas: generalization to arbitrary frequencies*, *Phys. Rev. B* **51**, 7497 (1995).
- [153] N. D. Mermin, *Lindhard dielectric function in the relaxation-time approximation*, *Phys. Rev. B* **1**, 2362 (1970).
- [154] M. del Castillo-Mussot, G.J. Vázquez Fonseca, and G.H. Coccoletzi, *Hydrodynamic model for the 2D degenerate free-electron gas for arbitrary frequencies*, *Rev. Mex. Fis.* **49**, 231 (2003).
- [155] S. Raza, T. Christensen, M. Wubs, S.I. Bozhevolnyi, and N.A. Mortensen, *Non-local response in thin-film waveguides: loss versus nonlocality and breaking of complementarity*, *Phys. Rev. B* **88**, 115401 (2013).
- [156] F. Forstmann and H. Stenschke, *Electrodynamics at metal boundaries with inclusion of plasma waves*, *Phys. Rev. Lett.* **38**, 1365 (1977).
- [157] A.D. Boardman and R. Ruppin, *The boundary conditions between spatially dispersive media*, *Surf. Sci.* **112**, 153 (1981).
- [158] P. Jewsbury, *Electrodynamic boundary conditions at metal interfaces*, *J. Phys. F: Met. Phys.* **11**, 195 (1981).
- [159] 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).
- [160] J.M. McMahon, S.K. Gray, and G.C. Schatz, *Calculating nonlocal optical properties of structures with arbitrary shape*, *Phys. Rev. B* **82**, 035423 (2010).

- [161] W.L. Mochán, M. del Castillo-Mussot, and R.G. Barrera, *Effect of plasma waves on the optical properties of metal-insulator superlattices*, *Phys. Rev. B* **35**, 1088 (1987).
- [162] A. Wiener, A. I. Fernández-Domínguez, A. P. Horsfield, J. B. Pendry, and S. A. Maier, *Nonlocal effects in the nanofocusing performance of plasmonic tips*, *Nano Lett.* **12**, 3308 (2012).
- [163] C. Ciraci, R.T. Hill, J.J. Mock, Y. Urzhumov, A.I. Fernández-Domínguez, S.A. Maier, J.B. Pendry, A. Chilkoti, and D.R. Smith, *Probing the ultimate limits of plasmonic enhancement*, *Science* **337**, 1072 (2012).
- [164] W. Yan, M. Wubs, and N.A. Mortensen, *Hyperbolic metamaterials: nonlocal response regularizes broadband supersingularity*, *Phys. Rev. B* **86**, 205429 (2012).
- [165] S. Raza, S.I. Bozhevolnyi, M. Wubs, and N.A. Mortensen, *Nonlocal optical response in metallic nanostructures*, *J. Phys.: Condens. Matter* **27**, 183204 (2015).
- [166] J. Lindhard, *On the properties of a gas of charged particles*, *Dan. Mat. Fys. Medd.* **28**, 8 (1954).
- [167] M. Ichikawa, *Theory of localized plasmons for metal nanostructures in random-phase approximation*, *J. Phys. Soc. Jpn.* **80**, 044606 (2011).
- [168] G.A. Baker, *Essentials of Padé Approximants* (Academic Press, 1975).
- [169] U. Kreibig and C. Fragstein, *The limitation of electron mean free path in small silver particles*, *Z. Physik* **224**, 307 (1969).
- [170] P. Apell and D.R. Penn, *Optical properties of small metal spheres: Surface effects*, *Phys. Rev. Lett.* **50**, 1316 (1983).
- [171] L. Genzel, T.P. Martin, and U. Kreibig, *Dielectric function and plasma resonances of small metal particles*, *Z. Phys. B: Condens. Matter* **21**, 339 (1975).
- [172] F. Ouyang, P.E. Batson, and M. Isaacson, *Quantum size effects in the surface-plasmon excitation of small metallic particles by electron-energy-loss spectroscopy*, *Phys. Rev. B* **46**, 15421 (1992).
- [173] K.-P. Charlé, L. König, S. Nijpko, I. Rabin, and W. Schulze, *The surface plasmon resonance of free and embedded Ag-clusters in size range $1.5 \text{ nm} < D < 30 \text{ nm}$* , *Cryst. Res. Technol.* **33**, 1085 (1998).
- [174] J. Tiggesbäumker, L. Köller, K.-H. Meiwes-Broer, and A. Liebsch, *Blue shift of the Mie plasma frequency in Ag clusters and particles*, *Phys. Rev. A* **48**, R1749 (1993).
- [175] S. Mandal, J. Wang, R.E. Winans, L. Jensen, and A. Sen, *Quantum size effects in the optical properties of ligand stabilized aluminum nanoclusters*, *J. Phys. Chem. C* **117**, 6741 (2013).
- [176] r. Tsuei, E.W. Plummer, and P.J. Feibelman, *Surface-plasmon dispersion in simple metals*, *Phys. Rev. Lett.* **63**, 2256 (1989).
- [177] A. Liebsch, *Dynamical screening at simple-metal surfaces*, *Phys. Rev. B* **36**, 7378 (1987).
- [178] A. Liebsch, *Surface-plasmon dispersion and size dependence of Mie resonance: silver versus simple metals*, *Phys. Rev. B* **48**, 11317 (1993).

- [179] C. David, N.A. Mortensen, and J. Christensen, *Perfect imaging, epsilon-near zero phenomena and waveguiding in the scope of nonlocal effects*, *Sci. Rep.* **3**, 2526 (2013).
- [180] A. Moreau, C. Ciraci, and D.R. Smith, *Impact of nonlocal response on metallodielectric multilayers and optical patch antennas*, *Phys. Rev. B* **87**, 045401 (2013).
- [181] W. Yan, N.A. Mortensen, and M. Wubs, *Hyperbolic metamaterial lens with hydrodynamic nonlocal response*, *Opt. Express* **21**, 15026 (2013).
- [182] J. Benedicto, R. Pollès, C. Ciraci, E. Centeno, D.R. Smith, and A. Moreau, *Numerical tool to take nonlocal effects into account in metallo-dielectric multilayers*, *J. Opt. Soc. Am. A* **32**, 1581 (2015).
- [183] F. Intravaia and K. Busch, *Fluorescence in nonlocal dissipative periodic structures*, *Phys. Rev. A* **91**, 053836 (2015).
- [184] I. Lindau and P.O. Nilsson, *Experimental verification of optically excited longitudinal plasmons*, *Phys. Scr.* **3**, 87 (1971).
- [185] M. Anderegg, B. Feuerbacher, and B. Fitton, *Optically excited longitudinal plasmons in potassium*, *Phys. Rev. Lett.* **27**, 1565 (1971).
- [186] M.M. Özer, E.J. Moon, A.G. Eguiluz, and H.H. Weitering, *Plasmon response of a quantum-confined electron gas probed by core-level photoemission*, *Phys. Rev. Lett.* **106**, 197601 (2011).
- [187] J.C. Maxwell Garnett, *Colours in metal glasses and in metallic films*, *Philos. Trans. R. Soc., A* **203**, 385 (1904).
- [188] M. Quinten, *Optical Properties of Nanoparticle Systems – Mie and beyond* (Wiley-VCH, 2011).
- [189] M. Hu, C. Novo, A. Funston, H. Wang, H. Staleva, S. Zou, P. Mulvaney, Y. Xia, and G.V. Hartland, *Dark-field microscopy studies of single metal nanoparticles: understanding the factors that influence the linewidth of the localized surface plasmon resonance*, *J. Mater. Chem.* **18**, 1949 (2008).
- [190] M.W. Knight, J. Fan, F. Capasso, and N.J. Halas, *Influence of excitation and collection geometry on the dark field spectra of individual plasmonic nanostructures*, *Opt. Express* **18**, 2579 (2010).
- [191] X. Tian, Y. Zhou, S. Thota, S. Zou, and J. Zhao, *Plasmonic coupling in single silver nanosphere assemblies by polarization-dependent dark-field scattering spectroscopy*, *J. Phys. Chem. C* **118**, 13801 (2014).
- [192] M. Schmelzeisen, Y. Zhao, M. Klapper, K. Müllen, and M. Kreiter, *Fluorescence enhancement from individual plasmonic gap resonances*, *ACS Nano* **4**, 3309 (2010).
- [193] G.M. Akselrod, C. Argyropoulos, T.G. Hoang, C. Ciraci, C. Fang, J. Huang, D.R. Smith, and M.H. Mikkelsen, *Probing the mechanisms of large Purcell enhancement in plasmonic nanoantennas*, *Nat. Photonics* **8**, 835 (2014).
- [194] J. Nelayah, M. Kociak, O. Stéphan, F.J. García de Abajo, M. Tencé, L. Henrard, D. Taverna, I. Pastoriza-Santos, L.M. Liz-Marzán, and C. Colliex, *Mapping surface plasmons on a single metallic nanoparticle*, *Nat. Physics* **3**, 348 (2007).

- [195] A.L. Koh, A.I. Fernández-Domínguez, D.V. McComb, S.A. Maier, and J.K.W. Yang, *High-resolution mapping of electron-beam-excited plasmon modes in lithographically defined gold nanostructures*, *Nano Lett.* **11**, 1323 (2011).
- [196] C. Diaz-Egea, R. Abargues, J.P. Martínez-Pastor, W. Sigle, P.A. van Aken, and S.I. Molina, *High spatial resolution mapping of individual and collective localized surface plasmon resonance modes of silver nanoparticle aggregates: correlation to optical measurements*, *Nanoscale Res. Lett.* **10**, 310 (2015).
- [197] N. Yamamoto, K. Araya, and F.J. García de Abajo, *Photon emission from silver particles induced by a high-energy electron beam*, *Phys. Rev. B* **64**, 205419 (2001).
- [198] S. Kadkhodazadeh, J.R. de Lasson, M. Beleggia, H. Kneipp, J.B. Wagner, and K. Kneipp, *Scaling of the surface plasmon resonance in gold and silver dimers probed by EELS*, *J. Phys. Chem. C* **118**, 5478 (2014).
- [199] L. Lorenz, *Light propagation in and outside a sphere illuminated by plane waves*, *K. Dan. Vidensk. Selsk. Skrift.* **6**, 6 (1890), originally 'Lysbevægelsen i og uden for en af plane Lysbølger belyst Kugle'.
- [200] P. Debye, *The light pressure on spheres of arbitrary material*, *Ann. Phys.* **335**, 57 (1909), originally 'Der Lichtdruck auf Kugeln von beliebigem Material'.
- [201] W. Hergert and T. Wriedt, eds., *The Mie Theory* (Springer, 2012).
- [202] J.A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, 1941).
- [203] W.C. Chew, *Waves and Fields in Inhomogeneous Media* (IEEE Press, 1994).
- [204] R. Ruppín, *Optical properties of a plasma sphere*, *Phys. Rev. Lett.* **31**, 1434 (1973).
- [205] C. David and F. J. García de Abajo, *Spatial nonlocality in the optical response of metal nanoparticles*, *J. Phys. Chem. C* **115**, 19470 (2012).
- [206] W. Yan, N.A. Mortensen, and M. Wubs, *Green function surface-integral method for nonlocal response of plasmonic nanowires in arbitrary dielectric environments*, *Phys. Rev. B* **88**, 155414 (2013).
- [207] A. Delga, J. Feist, J. Bravo-Abad, and F.J. García-Vidal, *Quantum emitters near a metal nanoparticle: strong coupling and quenching*, *Phys. Rev. Lett.* **112**, 253601 (2014).
- [208] A. Delga, J. Feist, J. Bravo-Abad, and F.J. García-Vidal, *Theory of strong coupling between quantum emitters and localized surface plasmons*, *J. Opt.* **16**, 114018 (2014).
- [209] F.J. García de Abajo, *Relativistic energy loss and induced photon emission in the interaction of a dielectric sphere with an external electron beam*, *Phys. Rev. B* **59**, 3095 (1999).
- [210] M. Kerker, D.-S. Wang, and H. Chew, *Surface enhanced Raman scattering (SERS) by molecules adsorbed at spherical particles: errata*, *Appl. Opt.* **19**, 4159 (1980).
- [211] R. Ruppín, *Decay of an excited molecule near a small sphere*, *J. Chem. Phys.* **76**, 1681 (1982).
- [212] H. Chew, *Transition rates of atoms near spherical surfaces*, *J. Chem. Phys.* **87**, 1355 (1987).

- [213] H.T. Dung, L. Knöll, and D.-G. Welsch, *Decay of an excited atom near an absorbing microsphere*, *Phys. Rev. A* **64**, 013804 (2001).
- [214] A. vom Felde, J. Fink, and W. Ekardt, *Quantum size effects in excitations of potassium clusters*, *Phys. Rev. Lett.* **61**, 2249 (1988).
- [215] S. Raza, S. Kadhodazadeh, T. Christensen, M. Di Vece, M. Wubs, N.A. Mortensen, and N. Stenger, *Multipole plasmons and their disappearance in few-nanometre silver nanoparticles*, *Nat. Commun.* **6**, 8788 (2015).
- [216] T.L. Ferrell, R.J. Warmack, V.E. Anderson, and P.M. Echenique, *Analytical calculation of the stopping power for isolated small spheres*, *Phys. Rev. B* **35**, 7365 (1987).
- [217] W. Yan, *Hydrodynamic theory for quantum plasmonics: linear-response dynamics of the inhomogeneous electron gas*, *Phys. Rev. B* **91**, 115416 (2015).
- [218] G. Toscano, J. Straubel, A. Kwiatkowski, C. Rockstuhl, F. Evers, H. Xu, N.A. Mortensen, and M. Wubs, *Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics*, *Nat. Commun.* **6**, 7132 (2015).
- [219] C. David and F.J. García de Abajo, *Surface plasmon dependence on the electron density profile at metal surfaces*, *ACS Nano* **8**, 9558 (2014).
- [220] W. Yang, *Gradient correction in Thomas–Fermi theory*, *Phys. Rev. A* **34**, 4575 (1986).
- [221] F.J. García de Abajo and J. Aizpurua, *Numerical simulation of electron energy loss near inhomogeneous dielectrics*, *Phys. Rev. B* **56**, 15873 (1997).
- [222] D.C. Brody, *Biorthogonal quantum mechanics*, *J. Phys. A: Math. Theor.* **47**, 035305 (2014).
- [223] A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, and A.K. Geim, *The electronic properties of graphene*, *Rev. Mod. Phys.* **81**, 109 (2009).
- [224] H.P. Boehm, A. Clauss, G.O. Fischer, and U. Hofmann, *The absorption properties of very thin carbon films*, *Z. Anorg. Allg. Chem.* **316**, 119 (1962), originally ‘Das Adsorptionsverhalten sehr dünner Kohlenstoff-Folien’.
- [225] N.R. Gall, E.V. Rut’kov, and A.Y. Tontegode, *Two dimensional graphite films on metal and their intercalation*, *Int. J. Mod. Phys. B* **11**, 1865 (1997).
- [226] E.S. Reich, *Nobel prize committee under fire*, Nature News (2010).
- [227] W. de Heer, *Letter to the 2010 Nobel committee*, online (2010).
- [228] P.R. Wallace, *The band theory of graphite*, *Phys. Rev.* **71**, 622 (1947).
- [229] J.C. Slonczewski and P.R. Weiss, *Band structure of graphite*, *Phys. Rev.* **109**, 272 (1958).
- [230] G.W. Semenoff, *Condensed-matter simulation of a three-dimensional anomaly*, *Phys. Rev. Lett.* **53**, 2449 (1984).
- [231] K.W.-K. Shung, *Dielectric function and plasmon structure of stage-1 intercalated graphite*, *Phys. Rev. B* **34**, 979 (1986).
- [232] R. Saito, M. Fujita, G. Dresselhaus, and M.S. Dresselhaus, *Electronic structure of chiral graphene tubules*, *Appl. Phys. Lett.* **60**, 2204 (1992).

- [233] K. Nakada, M. Fujita, G. Dresselhaus, and M.S. Dresselhaus, *Edge state in graphene ribbons: nanometer size effect and edge shape dependence*, *Phys. Rev. B* **54**, 17954 (1996).
- [234] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejón, *Tight-binding description of graphene*, *Phys. Rev. B* **66**, 035412 (2002).
- [235] Y. Yao, F. Ye, X.L. Qi, S.C. Zhang, and Z. Fang, *Spin-orbit gap of graphene: first-principles calculations*, *Phys. Rev. B* **75**, 041401 (2007).
- [236] H. Liu, Y. Liu, and D. Zhu, *Chemical doping of graphene*, *J. Mater. Chem.* **21**, 3335 (2011).
- [237] C.-F. Chen, C.-H. Park, B.W. Boudouris, J. Horng, B. Geng, C. Girit, A. Zettl, M.F. Crommie, R.A. Segalman, S.G. Louie, and F. Wang, *Controlling inelastic light scattering quantum pathways in graphene*, *Nature* **471**, 617 (2011).
- [238] D.K. Efetov and P. Kim, *Controlling electron-phonon interactions in graphene at ultrahigh carrier densities*, *Phys. Rev. Lett.* **105**, 256805 (2010).
- [239] Z. Fang, S. Thongrattanasiri, A. Schlather, Z. Liu, L. Ma, Y. Wang, P.M. Ajayan, P. Nordlander, N.J. Halas, and F.J. García de Abajo, *Gated tunability and hybridization of localized plasmons in nanostructured graphene*, *ACS Nano* **7**, 2388 (2013).
- [240] X. Zhu, W. Wang, W. Yan, M.B. Larsen, P. Bøggild, T.G. Pedersen, S. Xiao, J. Zi, and N.A. Mortensen, *Plasmon-phonon coupling in large-area graphene dot and antidot arrays fabricated by nanosphere lithography*, *Nano Lett.* **14**, 2907 (2014).
- [241] Z. Fang, Y. Wang, A.E. Schlather, Z. Liu, P.M. Ajayan, F.J. García de Abajo, P. Nordlander, X. Zhu, and N.J. Halas, *Active tunable absorption enhancement with graphene nanodisk arrays*, *Nano Letters* **14**, 299 (2014).
- [242] H. Ehrenreich and M.H. Cohen, *Self-consistent field approach to the many-electron problem*, *Phys. Rev.* **115**, 786 (1959).
- [243] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, *Dynamical polarization of graphene at finite doping*, *New J. Phys.* **8**, 318 (2006).
- [244] E.H. Hwang and S. Das Sarma, *Dielectric function, screening, and plasmons in two-dimensional graphene*, *Phys. Rev. B* **75**, 205418 (2007).
- [245] B.E. Sernelius, *Retarded interactions in graphene systems*, *Phys. Rev. B* **85**, 195427 (2012).
- [246] L.A. Falkovsky and A.A. Varlamov, *Space-time dispersion of graphene conductivity*, *Eur. Phys. J. B* **56**, 281 (2007).
- [247] G.W. Hanson, *Quasi-transverse electromagnetic modes supported by a graphene parallel-plate waveguide*, *J. Appl. Phys.* **104**, 084314 (2008).
- [248] P. Tassin, T. Koschny, and C.M. Soukoulis, *Graphene for terahertz applications*, *Science* **341**, 620 (2013).
- [249] Z. Fei, G.O. Andreev, W. Bao, L.M. Zhang, A.S. McLeod, C. Wang, M.K. Stewart, Z. Zhao, G. Dominguez, M. Thieme, M.M. Fogler, M.J. Tauber, A.H. Castro Neto, C.N. Lau, F. Keilmann, and D.N. Basov, *Infrared nanoscopy of Dirac plasmons at the graphene SiO₂ interface*, *Nano Lett.* **11**, 4701 (2011).

- [250] H. Yan, T. Low, W. Zhu, Y. Wu, M. Freitag, X. Li, F. Guinea, P. Avouris, and F. Xia, *Damping pathways of mid-infrared plasmons in graphene nanostructures*, *Nat. Photonics* **7**, 394 (2013).
- [251] R.R. Nair, P. Blake, A.N. Grigorenko, K.S. Novoselov, T.J. Booth, T. Stauber, N.M.R. Peres, and A.K. Geim, *Fine structure constant defines visual transparency of graphene*, *Science* **320**, 1308 (2008).
- [252] F. Stern, *Polarizability of a two-dimensional electron gas*, *Phys. Rev. Lett.* **18**, 546 (1967).
- [253] C.C. Grimes and G. Adams, *Observation of two-dimensional plasmons and electron-ripplon scattering in a sheet of electrons on liquid helium*, *Phys. Rev. Lett.* **36**, 145 (1976).
- [254] T.N. Theis, *Plasmons in inversion layers*, *Surf. Sci.* **98**, 515 (1980).
- [255] S.J. Allen Jr., D.C. Tsui, and R.A. Logan, *Observation of the two-dimensional plasmon in silicon inversion layers*, *Phys. Rev. Lett.* **38**, 980 (1977).
- [256] T.N. Theis, J.P. Kotthaus, and P.J. Stiles, *Two-dimensional magnetoplasmon in the silicon inversion layer*, *Solid State Commun.* **24**, 273 (1977).
- [257] D.B. Mast, A.J. Dahm, and A.L. Fetter, *Observation of bulk and edge magnetoplasmons in a two-dimensional electron fluid*, *Phys. Rev. Lett.* **54**, 1706 (1985).
- [258] D.C. Glattli, E.Y. Andrei, G. Deville, and F.I.B. Williams, *1D perimeter waves in a classical 2D electron system*, *Surf. Sci.* **170**, 70 (1986).
- [259] B.M. Santoyo and M. del Castillo-Mussot, *Plasmons in three, two and one dimension*, *Rev. Mex. Fis.* **39**, 640 (1993).
- [260] K. Linghua, Y. Baorong, and H. Xiwei, *Dispersion relations of longitudinal plasmons in one, two and three dimensional electron gas of metals*, *Plasma Sci. Technol.* **9**, 519 (2007).
- [261] S.A. Mikhailov and K. Ziegler, *New electromagnetic mode in graphene*, *Phys. Rev. Lett.* **99**, 016803 (2007).
- [262] A. Bondi, *Van der Waals volumes and radii*, *J. Phys. Chem.* **68**, 441 (1964).
- [263] R. Messina, J.-R. Hugonin, J.-J. Greffet, F. Marquier, Y. De Wilde, A. Belarouci, L. Frechette, Y. Cordier, and P. Ben-Abdallah, *Tuning the electromagnetic local density of states in graphene-covered systems via strong coupling with graphene plasmons*, *Phys. Rev. B* **87**, 085421 (2013).
- [264] V. Ariel and A. Natan, *Electron effective mass in graphene*, in *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA)* (2013) pp. 696–698.
- [265] T. Christensen, W. Wang, A.-P. Jauho, M. Wubs, and N.A. Mortensen, *Classical and quantum plasmonics in graphene nanodisks: role of edge states*, *Phys. Rev. B* **90**, 241414(R) (2014).
- [266] X. Zhu, W. Yan, P.U. Jepsen, O. Hansen, N.A. Mortensen, and S. Xiao, *Experimental observation of plasmons in a graphene monolayer resting on a two-dimensional subwavelength silicon grating*, *Appl. Phys. Lett.* **102**, 131101 (2013).
- [267] F.J. García de Abajo, *Multiple excitation of confined graphene plasmons by single free electrons*, *ACS Nano* **7**, 11409 (2013).

- [268] M. Rocca, *Low-energy EELS investigation of surface electronic excitations on metals*, *Surf. Sci. Rep.* **22**, 1 (1995).
- [269] Y. Liu, R.F. Willis, K.V. Emtsev, and T. Seyller, *Plasmon dispersion and damping in electrically isolated two-dimensional charge sheets*, *Phys. Rev. B* **78**, 201403 (2008).
- [270] O.V. Kotov, M.A. Kol'chenko, and Y.E. Lozovik, *Ultrahigh refractive index sensitivity of TE-polarized electromagnetic waves in graphene at the interface between two dielectric media*, *Opt. Express* **21**, 13533 (2013).
- [271] V.A. Volkov and S.A. Mikhailov, *Edge magnetoplasmons: low frequency weakly damped excitations in inhomogeneous two-dimensional electron systems*, *Sov. Phys. JETP* **67**, 1639 (1988).
- [272] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. (Elsevier Academic Press, 2007).
- [273] V.A. Volkov and S.A. Mikhailov, *Theory of edge magnetoplasmons in a two-dimensional electron gas*, *JETP Lett.* **42**, 556 (1985).
- [274] A.L. Fetter, *Edge magnetoplasmons in a two-dimensional electron fluid confined to a half-space*, *Phys. Rev. B* **33**, 3717 (1986).
- [275] T. Christensen, W. Yan, A.-P. Jauho, M. Wubs, and N.A. Mortensen, *Kerr non-linearity and plasmonic bistability in graphene nanoribbons*, *Phys. Rev. B* **92**, 121407(R) (2015).
- [276] V. Cataudella and G. Iadonisi, *Magnetoplasmons in a two-dimensional electron gas: strip geometry*, *Phys. Rev. B* **35**, 7443 (1987).
- [277] S. Rudin and M. Dyakonov, *Edge and strip plasmons in a two-dimensional electron fluid*, *Phys. Rev. B* **55**, 4684 (1997).
- [278] J. Christensen, A. Manjavacas, S. Thongrattanasiri, F.H.L. Koppens, and F.J. García de Abajo, *Graphene plasmon waveguiding and hybridization in individual and paired nanoribbons*, *ACS Nano* **6**, 431 (2012).
- [279] K.A. Velizhanin, *Geometric universality of plasmon modes in graphene nanoribbon arrays*, *Phys. Rev. B* **91**, 125429 (2015).
- [280] S. Thongrattanasiri, F.H.L. Koppens, and F.J. García de Abajo, *Complete optical absorption in periodically patterned graphene*, *Phys. Rev. Lett.* **108**, 047401 (2012).
- [281] W. Wang, P.S. Apell, and J.M. Kinaret, *Edge magnetoplasmons and the optical excitations in graphene disks*, *Phys. Rev. B* **86**, 125450 (2012).
- [282] F.J. García de Abajo and A. Manjavacas, *Plasmonics in atomically thin materials*, *Faraday Discuss.* **178**, 87 (2015).
- [283] A.L. Fetter, *Magnetoplasmons in a two-dimensional electron fluid: disk geometry*, *Phys. Rev. B* **33**, 5221 (1986).
- [284] F. Ouyang and M. Isaacson, *Surface plasmon excitation of objects with arbitrary shape and dielectric constant*, *Philos. Mag. B* **60**, 481 (1989).
- [285] W. Wang, T. Christensen, A.-P. Jauho, K.S. Thygesen, M. Wubs, and N.A. Mortensen, *Plasmonic eigenmodes in individual and bow-tie graphene nanotriangles*, *Sci. Rep.* **5**, 9535 (2015).

- [286] S. Thongrattanasiri and F.J. García de Abajo, *Optical field enhancement by strong plasmon interaction in graphene nanostructures*, *Phys. Rev. Lett.* **110**, 187401 (2013).
- [287] W.B. Lu, W. Zhu, H.J. Xu, Z.H. Ni, Z.G. Dong, and T.J. Cui, *Flexible transformation plasmonics using graphene*, *Opt. Express* **21**, 10475 (2013).
- [288] D.A. Smirnova, S.H. Mousavi, Z. Wang, Y.S. Kivshar, and B. Khanikaev, *Trapping and guiding surface plasmons in curved graphene landscapes*, *arXiv:1508.02729* (2015).
- [289] P.-Y. Chen, J. Soric, and A. Alú, *Invisibility and cloaking based on scattering cancellation*, *Adv. Mater.* **24**, OP281 (2012).
- [290] P. Chen, J. Soric, Y.R. Padooru, H.M. Bernety, A.B. Yakovlev, and A. Alú, *Nanostructured graphene metasurface for tunable terahertz cloaking*, *New J. Phys.* **15**, 123029 (2013).
- [291] B. Zhu, G. Ren, Y. Gao, Y. Yang, Y. Lian, and S. Jian, *Graphene-coated tapered nanowire infrared probe: a comparison with metal-coated probes*, *Opt. Express* **22**, 24096 (2014).
- [292] Y. Gao, G. Ren, B. Zhu, H. Liu, Y. Lian, and S. Jian, *Analytical model for plasmon modes in graphene-coated nanowire*, *Opt. Express* **22**, 24322 (2014).
- [293] Y. Gao, G. Ren, B. Zhu, J. Wang, and S. Jian, *Single-mode graphene-coated nanowire plasmonic waveguide*, *Opt. Lett.* **39**, 5909 (2014).
- [294] Z.-R. Huang, L.L. Wang, B. Sun, M.-D. He, J.-Q. Liu, H.-J. Li, and X. Zhai, *A mid-infrared fast-tunable graphene ring resonator based on guided-plasmonic wave resonance on a curved graphene surface*, *J. Opt.* **16**, 105004 (2014).
- [295] L. Martin-Moreno, F.J. García de Abajo, and F.J. García-Vidal, *Ultra-efficient coupling of a quantum emitter to the tunable guided plasmons of a carbon nanotube*, *arXiv:1502.02488* (2015).
- [296] M. Riso, M. Cuevas, and R.A. Depine, *Tunable plasmonic enhancement of light scattering and absorption in graphene-coated subwavelength wires*, *J. Opt.* **17**, 075001 (2015).
- [297] D. Bakowies, M. Buehl, and W. Thiel, *Can large fullerenes be spherical?* *J. Am. Chem. Soc.* **117**, 10113 (1995).
- [298] P. Calaminici, G. Geudtner, and A. M. Köster, *First-principle calculations of large fullerenes*, *J. Chem. Theory Comput.* **5**, 29 (2009).
- [299] L. Wu, H. Feng, M. Liu, K. Zhang, and J. Li, *Graphene-based hollow spheres as efficient electrocatalysts for oxygen reduction*, *Nanoscale* **5**, 10839 (2013).
- [300] H. Yang, Z. Hou, N. Zhou, B. He, J. Cao, and Y. Kuang, *Graphene-encapsulated SnO₂ hollow spheres as high-performance anode materials for lithium ion batteries*, *Ceram. Int.* **40**, 13903 (2014).
- [301] J.-S. Lee, S.-I. Kim, J.-C. Yoon, and J.-H. Jang, *Chemical vapor deposition of mesoporous graphene nanoballs for supercapacitor*, *ACS Nano* **7**, 6047 (2013).
- [302] J.-C. Charlier, X. Blase, and S. Roche, *Electronic and transport properties of nanotubes*, *Rev. Mod. Phys.* **79**, 677 (2007).

- [303] J. Vielma and P. Leung, *Nonlocal optical effects on the fluorescence and decay rates for admolecules at a metallic nanoparticle*, *J. Chem. Phys.* **126**, 194704 (2007).
- [304] M. Müller, J. Schmalian, and L. Fritz, *Graphene: a nearly perfect fluid*, *Phys. Rev. Lett.* **103**, 025301 (2009).
- [305] M. Mendoza, H.J. Herrmann, and S. Succi, *Preturbulent regimes in graphene flow*, *Phys. Rev. Lett.* **106**, 156601 (2011).
- [306] M. Mendoza, H.J. Herrmann, and S. Succi, *Hydrodynamic model for conductivity in graphene*, *Sci. Rep.* **3**, 1052 (2013).
- [307] G.W. Hanson, *Dyadic Green's functions for an anisotropic, non-local model of biased graphene*, *IEEE Trans. Antennas Propag.* **56**, 747 (2008).
- [308] W. Wang and J.M. Kinaret, *Plasmons in graphene nanoribbons: interband transitions and nonlocal effects*, *Phys. Rev. B* **87**, 195424 (2013).
- [309] A. Fallahi, T. Low, M. Tamagnone, and J. Perruisseau-Carrier, *Nonlocal electromagnetic response of graphene nanostructures*, *Phys. Rev. B* **91**, 121405 (2015).
- [310] K. Tanaka, S. Yamashita, H. Yamabe, and T. Yamabe, *Electronic properties of one-dimensional graphite family*, *Synth. Met.* **17**, 143 (1987).
- [311] M. Inui, S.A. Trugman, and E. Abrahams, *Unusual properties of midband states in systems with off-diagonal disorder*, *Phys. Rev. B* **49**, 3190 (1994).
- [312] E.H. Lieb, *Two theorems on the Hubbard model*, *Phys. Rev. Lett.* **62**, 1201 (1989).
- [313] M. Wimmer, A.R. Akhmerov, and F. Guinea, *Robustness of edge states in graphene quantum dots*, *Phys. Rev. B* **82**, 045409 (2010).
- [314] P. Delplace, D. Ullmo, and G. Montambaux, *Zak phase and the existence of edge states in graphene*, *Phys. Rev. B* **84**, 195452 (2011).
- [315] M. Bellec, U. Kuhl, G. Montambaux, and F. Mortessagne, *Manipulation of edge states in microwave artificial graphene*, *New J. Phys.* **16**, 113023 (2014).
- [316] A.R. Akhmerov and C.W.J. Beenakker, *Boundary conditions for Dirac fermions on a terminated honeycomb lattice*, *Phys. Rev. B* **77**, 085423 (2008).
- [317] Z. Klusek, Z. Waqar, E.A. Denisov, T.N. Kompaniets, I.V. Makarenko, A.N. Titkov, and A.S. Bhatti, *Observations of local electron states on the edges of the circular pits on hydrogen-etched graphite surface by scanning tunneling spectroscopy*, *Appl. Surf. Sci.* **161**, 508 (2000).
- [318] Y. Kobayashi, K. Fukui, T. Enoki, K. Kusakabe, and Y. Kaburagi, *Observation of zigzag and armchair edges of graphite using scanning tunneling microscopy and spectroscopy*, *Phys. Rev. B* **71**, 193406 (2005).
- [319] Y. Kobayashi, K. Fukui, T. Enoki, and K. Kusakabe, *Edge state on hydrogen-terminated graphite edges investigated by scanning tunneling microscopy*, *Phys. Rev. B* **73**, 125415 (2006).
- [320] K.A. Ritter and J.W. Lyding, *The influence of edge structure on the electronic properties of graphene quantum dots and nanoribbons*, *Nat. Mater.* **8**, 235 (2009).
- [321] C. Tao, L. Jiao, O.V. Yazyev, Y.-C. Chen, J. Feng, X. Zhang, R.B. Capaz, J.M. Tour, A. Zettl, S.G. Louie, H. Dai, and M.F. Crommie, *Spatially resolving edge states of chiral graphene nanoribbons*, *Nat. Physics* **7**, 616 (2011).

- [322] X. Jia, J. Campos-Delgado, M. Terrones, V. Meunier, and M.S. Dresselhaus, *Graphene edges: a review of their fabrication and characterization*, *Nanoscale* **3**, 86 (2011).
- [323] M. Acik and Y.J. Chabal, *Nature of graphene edges: a review*, *Jpn. J. Appl. Phys.* **50**, 070101 (2011).
- [324] S.K. Hämäläinen, Z. Sun, M.P. Boneschanscher, A. Uppstu, M. Ijäs, A. Harju, D. Vanmaekelbergh, and P. Liljeroth, *Quantum-confined electronic states in atomically well-defined graphene nanostructures*, *Phys. Rev. Lett.* **107**, 236803 (2011).
- [325] X. Jia, M. Hofmann, V. Meunier, B.G. Sumpter, J. Campos-Delgado, J.M. Romo-Herrera, H. Son, Y.-P. Hsieh, A. Reina, J. Kong, M. Terrones, and M.S. Dresselhaus, *Controlled formation of sharp zigzag and armchair edges in graphitic nanoribbons*, *Science* **323**, 1701 (2009).
- [326] K. Kim, S. Coh, C. Kisielowski, M.F. Crommie, S.G. Louie, M.L. Cohen, and A. Zettl, *Atomically perfect torn graphene edges and their reversible reconstruction*, *Nat. Commun.* **4**, 2723 (2013).
- [327] L. Brey and H.A. Fertig, *Elementary electronic excitations in graphene nanoribbons*, *Phys. Rev. B* **75**, 125434 (2007).
- [328] A. Manjavacas, F. Marchesin, S. Thongrattanasiri, P. Koval, P. Nordlander, D. Sánchez-Portal, and F.J. García de Abajo, *Tunable molecular plasmons in polycyclic aromatic hydrocarbons*, *ACS Nano* **7**, 3635 (2013).
- [329] A. Manjavacas, S. Thongrattanasiri, and F.J. García de Abajo, *Plasmons driven by single electrons in graphene nanoislands*, *Nanophotonics* **2**, 139 (2013).
- [330] S. Thongrattanasiri, A. Manjavacas, P. Nordlander, and F.J. García de Abajo, *Quantum junction plasmons in graphene dimers*, *Laser Photon. Rev.* **7**, 297 (2013).
- [331] I. Silveiro, J.M.P. Ortega, and F.J. García de Abajo, *Quantum nonlocal effects in individual and interacting nanoribbons*, *Light: Sci. Appl.* **4**, e241 (2015).
- [332] E. Clementi and C. Roetti, *Roothan–Hartree–Fock atomic wave functions*, *At. Data Nucl. Data Tables* **14**, 177 (1974).
- [333] D.J. Rowe, *An interpretation of time-dependent Hartree–Fock theory*, *Nucl. Phys.* **80**, 209 (1966).
- [334] D.J. Rowe, *Nuclear collective motion* (World Scientific, 1970).
- [335] A. Castro, *A first-principles time-dependent density functional theory scheme for the computation of the electromagnetic response of nanostructures*, Ph.D. thesis, Universidad de Valladolid (2004).
- [336] D.C. Marinica, A.K. Kazansky, P. Nordlander, J. Aizpurura, and A.G. Borisov, *Quantum plasmonics: nonlinear effects in the field enhancement of a plasmonic nanoparticle dimer*, *Nano Lett.* **12**, 1333 (2012).
- [337] J.D. Cox and F.J. García de Abajo, *Electrically tunable nonlinear plasmonics in graphene nanoislands*, *Nat. Commun.* **5**, 2014 (2014).
- [338] M. Zarenia, A. Chaves, G. Farias, and F.M. Peeters, *Energy levels of triangular and hexagonal graphene quantum dots: a comparative study between the tight-binding and Dirac equation approach*, *Phys. Rev. B* **84**, 245403 (2011).

- [339] M.V. Berry and R.J. Mondragon, *Neutrino billiards: time-reversal symmetry-breaking without magnetic fields*, *Proc. R. Soc. Lond. A* **412**, 53 (1987).
- [340] E. McCann and V.I. Fal'ko, *Symmetry of boundary conditions of the Dirac equation for electrons in carbon nanotubes*, *J. Phys.: Condens. Matter* **16**, 2371 (2004).
- [341] N.M.R. Peres, J.N.B. Rodrigues, T. Stauber, and J.M.B. Lopes dos Santos, *Dirac electrons in graphene-based quantum wires and quantum dots*, *J. Phys.: Condens. Matter* **21**, 344202 (2009).
- [342] B. Wunsch, T. Stauber, and F. Guinea, *Electron-electron interactions and charging effects in graphene quantum dots*, *Phys. Rev. B* **77**, 035316 (2008).
- [343] M. Grujić, M. Zarenia, A. Chaves, M. Tadić, G.A. Farias, and F.M. Peeters, *Electronic and optical properties of a circular graphene quantum dot in a magnetic field: influence of the boundary conditions*, *Phys. Rev. B* **84**, 205441 (2011).
- [344] M. Kauranen and A.V. Zayats, *Nonlinear plasmonics*, *Nature Photon.* **6**, 737 (2012).
- [345] R.W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic Press, 2008).
- [346] H.M. Gibbs, *Optical bistability: Controlling light with light* (Academic Press, 1985).
- [347] K.I. Bolotin, K.J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H.L. Stormer, *Ultrahigh electron mobility in suspended graphene*, *Solid State Commun.* **146**, 351 (2008).
- [348] S.A. Mikhailov and K. Ziegler, *Nonlinear electromagnetic response of graphene: frequency multiplication and the self-consistent-field effects*, *Jour. Phys.: Cond. Matt.* **20**, 384204 (2008).
- [349] S.A. Mikhailov, *Theory of the nonlinear optical frequency mixing effect in graphene*, *Physica E* **44**, 924 (2012).
- [350] J.L. Cheng, N. Vermeulen, and J.E. Sipe, *Third order optical nonlinearity of graphene*, *New J. Phys.* **16**, 053014 (2014).
- [351] N.M.R. Peres, Y.V. Bludov, J.E. Santos, A.-P. Jauho, and M.I. Vasilevskiy, *Optical bistability of graphene in the terahertz range*, *Phys. Rev. B* **90**, 125425 (2014).
- [352] P. Ginzburg, A. Hayat, N. Berkovitch, and M. Orenstein, *Nonlocal ponderomotive nonlinearity in plasmonics*, *Opt. Lett.* **35**, 1551 (2010).
- [353] M.L. Nesterov, J. Bravo-Abad, A.Y. Nikitin, F.J. García Vidal, and L. Martín-Moreno, *Graphene supports the propagation of subwavelength optical solitons*, *Laser Photon. Rev.* **7**, L7 (2013).
- [354] D.A. Smirnova, A.V. Gorbach, I.V. Iorsh, I.V. Shadrivov, and Y.S. Kivshar, *Nonlinear switching with a graphene coupler*, *Phys. Rev. B* **88**, 045443 (2013).
- [355] D.A. Smirnova, I.V. Shadrivov, A.I. Smirnov, and Y.S. Kivshar, *Dissipative plasmon-solitons in multilayer graphene*, *Laser Photon. Rev.* **8**, 291 (2014).
- [356] D.A. Smirnova, R.E. Noskov, L.A. Smirnov, and Y.S. Kivshar, *Dissipative plasmon solitons in graphene nanodisk arrays*, *Phys. Rev. B* **91**, 075409 (2015).
- [357] H. Nasari and M.S. Abrishamian, *All-optical tunable notch filter by use of Kerr nonlinearity in the graphene microribbon array*, *J. Opt. Soc. Am. B* **31**, 1691 (2014).

- [358] M. Gullans, D.E. Chang, F.H.L. Koppens, F.J. García de Abajo, and M.D. Lukin, *Single-photon nonlinear optics with graphene plasmons*, *Phys. Rev. Lett.* **111**, 247401 (2013).
- [359] X. Yao, M. Tokman, and A. Belyanin, *Efficient nonlinear generation of THz plasmons in graphene and topological insulators*, *Phys. Rev. Lett.* **112**, 055501 (2014).
- [360] A.V. Gorbach, *Nonlinear graphene plasmonics: amplitude equation for surface plasmons*, *Phys. Rev. A* **87**, 013830 (2013).
- [361] J.D. Cox and F.J. García de Abajo, *Plasmon-enhanced nonlinear wave mixing in nanostructured graphene*, *ACS Photonics* **2**, 306 (2015).
- [362] X.-H. Wang, *Finite Element Methods for Nonlinear Optical Waveguides*, *Advances in Nonlinear Optics*, Vol. 2 (Gordan and Breach Publishers, 1995).
- [363] B. Semnani, A.H. Majedi, and S. Safavi-Naeini, *Nonlinear quantum optical properties of graphene: the role of chirality and symmetry*, *arXiv:1502.02203* (2015).
- [364] J. Krenn and D. Pile, *Perspective on plasmonics*, *Nat. Photonics* **6**, 714 (2012).
- [365] K. J. Savage, M. M. Hawkeye, R. Esteban, A. G. Borisov, J. Aizpurua, and J. J. Baumberg, *Revealing the quantum regime in tunnelling plasmonics*, *Nature* **491**, 574 (2012).
- [366] R. Esteban, A. Zugarramurdi, P. Zhang, P. Nordlander, F.J. García-Vidal, A.G. Borisov, and J. Aizpurua, *A classical treatment of optical tunneling in plasmonic gaps: extending the quantum corrected model to practical situations*, *Faraday Discuss.* **178**, 151 (2015).
- [367] F.J. García de Abajo, R. Sapienza, M. Noginov, F. Benz, J. Baumberg, S. Maier, D. Graham, J. Aizpurua, T. Ebbesen, A. Pinchuk, J. Khurgin, K. Matczyszyn, J.T. Hugal, N. van Hulst, P. Dawson, C. Roberts, M. Nielsen, L. Bursi, M. Flatté, J. Yi, O. Hess, N. Engheta, M. Brongersma, V. Podolskiy, V. Shalaev, E. Narimanov, and A. Zayats, *Plasmonic and new plasmonic materials: general discussion*, *Faraday Discuss.* **178**, 123 (2015).
- [368] A. Lauchner, A.E. Schlather, A. Manjavacas, Y. Cui, M.J. McClain, G.J. Stec, F.J. García de Abajo, P. Nordlander, and N.J. Halas, *Molecular plasmonics*, *Nano Lett.* **15**, 6208 (2015).
- [369] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, 1972).
- [370] N. Watanabe and M. Tsukada, *Efficient method for simulating quantum electron dynamics under the time-dependent Kohn–Sham equation*, *Phys. Rev. E* **65**, 036705 (2002).