

**Dielectric Response and Interactions in
Low-dimensional Carbon Materials from First
Principles Calculations.**

by

Boris Kozinsky

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

© Massachusetts Institute of Technology 2007. All rights reserved.

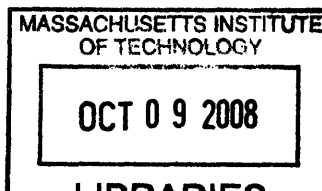
Author
Department of Physics

1100 MA May 11, 2007

Certified by
Nicola Marzari
Associate Professor of Materials Science and Engineering
Thesis Supervisor

Certified by
John D. Joannopoulos
Francis Wright Davis Professor of Physics
Thesis Supervisor

Accepted by
Thomas J. Greytak
Associate Department Head for Education



ARCHIVES

Dielectric Response and Interactions in Low-dimensional Carbon Materials from First Principles Calculations

by
Boris Kozinsky

Submitted to the Department of Physics
on May 11, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This work examines the dielectric response of nanoscale one- and two-dimensional systems using first-principles calculations. In the first part we discuss the peculiarities of the long-wavelength in-plane dielectric response of graphene and boron-nitride sheets. Graphene is shown to possess features that classify it as an intermediate between a two-dimensional metal and an insulator, unlike the simple insulating case of boron-nitride. Carbon nanotubes inherit the unusual scale-invariance of graphene's response, which manifests itself as a high transverse polarizability and a radius-insensitive screening. Boron nitride nanotubes behave more like hollow dielectric cylinders. For a single-wall nanotube of each type, we find that the longitudinal response is controlled by the band gap, while the transverse response is sensitive only to the effective radius. To study the static response of multiwall nanotubes, we construct a simple yet accurate electrostatic model that predicts qualitative differences between the transverse response of carbon nanotubes and that of boron-nitride nanotubes. The second part of this work is concerned with the Luttinger liquid description of interactions in nanoscale one-dimensional conductors, that is believed to explain the power laws observed in experiments in the photoemission spectra and the dependence of the conductance on voltage and temperature. We describe a method for ab-initio calculations of the parameters of the Luttinger liquid model and apply it to metallic carbon nanotubes, finding good agreement with experiments. The third part is devoted to an analysis of issues that arise in ab-initio calculations of low-dimensional systems using fully periodic boundary conditions. We present a comprehensive review of the difficulties arising for different dimensionalities and discuss efficient ways of using electrostatics to dramatically improve the speed and accuracy of computations.

Thesis Supervisor: Nicola Marzari
Title: Associate Professor of Materials Science and Engineering

Thesis Supervisor: John D. Joannopoulos
Title: Francis Wright Davis Professor of Physics

Acknowledgments

I am very grateful to my thesis advisor Prof. Nicola Marzari for his inspiring creative suggestions, unwavering support and infectious enthusiasm for cool new problems. I greatly enjoyed the warm and productive atmosphere of the Quasiamore group, where I worked for the past three years, and I would like to thank all of the members with whom I had the pleasure to interact and collaborate: Mayeul D’Avezac, Francesca Baletto, Nicola Bonini, Matteo Cococcioni, Ismaila Dabo, Oswaldo Dieguez, Jivtesh Garg, Heather Kulik, Young-Su Lee, Elise Li, Arash Mostofi, Nicolas Mounet, Nicholas Singh-Miller, Nicolas Poilvert, Damian Scherlis, Kathryn Simons, Patrick Sit, Timo Thonhauser, Paolo Umari, and Brandon Wood.

I cannot thank enough Prof. Leonid Levitov, whom I had the fortune to learn from and collaborate with since my undergraduate days. His help and sage advice were essential during my years at MIT, and his profound insights have inspired much of the work presented in this thesis. I am thankful to Prof. Mildred Dresselhaus for important comments on the thesis draft and wonderful discussions on the history of nanotubes.

I owe a great deal to my parents for their love and encouragement of my early scientific pursuits and to my sister Inna for being my role model throughout much of my life. Finally, my wonderful wife Katya has my inexpressible gratitude for her devotion, patience and persistent encouragement that made this work possible. She also deserves credit for some of the best-looking illustrations in this thesis.

Contents

1	Introduction	17
1.1	Nanotubes	18
1.2	Graphene	20
1.3	Outline	21
2	Basics of ab-initio calculations	23
2.1	Many-electron problem	23
2.2	Hohenberg-Kohn theorems	25
2.3	Kohn-Sham equations	27
2.4	Exchange-correlation functionals	29
2.5	Planewave basis calculations	30
2.6	Density functional perturbation theory	33
3	Structure of graphene and carbon nanotubes	37
3.1	Electronic structure of graphene	37
3.2	Geometric and electronic structure of carbon nanotubes	44
4	Static dielectric response of sheets and tubes	51
4.1	Dielectric linear response formalism	52
4.1.1	Definitions and identities	52
4.1.2	Random-phase approximation	59
4.1.3	Polarizabilities in low dimensions	62
4.2	2D: Graphene and boron-nitride	63

4.2.1	Wannier function analysis	70
4.3	1D: Nanotubes	72
4.3.1	Longitudinal polarizabilities	72
4.3.2	Transverse polarizabilities and screening factors	75
4.3.3	Electrostatic models of nanotubes	82
4.3.4	Nanotube bundles	89
4.3.5	Linearity of the response	90
4.4	Separation and alignment of nanotubes	91
5	Calculation of Luttinger liquid parameters in carbon nanotubes	95
5.1	Introduction	95
5.2	The Luttinger liquid	99
5.3	DFT approach to parameter calculation	102
6	Electrostatic corrections in periodic boundary conditions	113
6.1	Charged Systems	115
6.1.1	Point countercharge method	115
6.2	Polarized Systems	120
6.2.1	Polarizability corrections	123
6.2.2	Depolarization shape effects in finite samples	127
6.3	Inhomogeneous Systems	128
6.4	Density countercharge method	133
6.5	Defects in bulk medium	135
7	Summary and future directions	141
A	Electrostatics of dielectric cylinders	145
B	Matlab code for MWNT models	149
C	Polarizability tensor of a metallic ellipsoid	153

List of Figures

3-1	Honeycomb lattice consists of two sublattices A(\bullet) and B(\circ).	38
3-2	Bands of graphene as obtained from the tight binding calculation. Courtesy of Jonas Hauptmann.	42
3-3	Brillouin zone of graphene.	43
3-4	Chiral and translational vectors of single-wall nanotubes. Structures on the right are examples of zigzag, general chiral, and armchair nan- otubes. Reproduced from http://www.nanoelectronics.jp	46
3-5	(a),(b) The transverse k-vector component k_{\perp} is quantized when peri- odic boundary conditions are applied. (c) The values of k_{\perp} allowed by quantization slice through the bands of graphene. Pictures courtesy of Paul McEuen's group.	47
3-6	Quantization of the k -vectors in the Brillouin zone and the resulting bands.	48
3-7	Band structure of armchair and zigzag carbon SWNTs obtained from DFT calculations. Courtesy of Young-Su Lee.	48
4-1	Charge density oscillations in a graphene sheet induced by a perturbing sinusoidal potential with a wavelength equal to 60 unit cells. Rapid oscillations are caused by charge transfer between the two carbon sub- lattices.	65

4-2	Scaling of the self-consistent response matrix element $\chi_{0,0}(q)$ of a graphene and a BN sheet as a function of the wavelength of the perturbing sinusoidal potential. The doping level for graphene is 0.005 electrons per unit cell ($\varepsilon_F=0.01$ Ha). At lower doping levels the curve exhibits a sharper variation at smaller q	67
4-3	Scaling of the full response matrix element $\chi_{0,0}(q)$ of the BN sheet. Extrapolation to $q = 0$ is used to compute the in-plane polarizability.	68
4-4	Inverse screened compressibility of a doped graphene sheet versus the doping density. The dashed curve is computed from the independent-particle linear band model via (4.53).	69
4-5	Maximally localized Wannier functions of the occupied states in graphene. (a) σ -bonding MLWF; (b) π -bonding MLWF. Courtesy of Young-Su Lee.	70
4-6	Percent relative contribution of π -electrons to the in-plane response function of the undoped graphene sheet.	71
4-7	Log scale plot of α_{\parallel} of zigzag nanotubes as a function of band gap. The dashed line has slope -2. The inset shows the values for large-gap SWNTs as a function of R_0/Δ_g^2	73
4-8	Longitudinal polarizability of boron-nitride nanotubes as a function of radius.	73
4-9	Transverse polarizabilities α_{\perp} of armchair and zigzag nanotubes as a function of \tilde{R}^2 . The dashed line is the best-fit result of our semi-metallic shell model; the solid line $\alpha_{\perp} = \frac{1}{2}\tilde{R}^2$ corresponds to an ideal metallic cylinder.	76
4-10	Transverse polarizabilities of boron-nitride nanotubes as a function of the radius. The dashed line is the best-fit result of our thick dielectric shell model.	76
4-11	Electrostatic potential for a (10,10) SWNT in an applied homogeneous transverse field E_{out} . The electric field through the center slice is shown in the inset.	77

4-12	Screening factors of carbon and boron-nitride nanotubes versus the radius.	80
4-13	Transverse response of a tube to a uniform field can be viewed as the in-plane response of a sheet to a sinusoidal potential.	82
4-14	Electrostatic models for transverse response of single-wall nanotubes. The first line has the correct radius dependence for CNTs but is obviously wrong.	82
4-15	Transverse polarizability of a carbon MWNT of radius 100\AA versus the number of inner layers, as predicted by our electrostatic model. The outer few layers strongly dominate the transverse response.	87
4-16	Transverse polarizability of a boron-nitride MWNT of radius 100\AA versus the number of inner layers, as predicted by our electrostatic model.	87
4-17	Screening factors in a carbon MWNT of outer radius 100\AA versus the number of nanotube layers, as predicted by our electrostatic model. The vertical axis is on the log scale, so screening is exponential in the number of layers.	88
4-18	Screening factors of a boron-nitride MWNT of outer radius 100\AA versus the number of inner layers, as predicted by our electrostatic model. The innermost layers contribute more to the transverse response.	88
4-19	Electrostatic potential for a (10,10) SWNT in an applied homogeneous transverse field E_{out} . The electric field through the center slice is shown in the inset.	91
5-1	Momentum distribution function. Dotted line represents free electron gas. Red line represents the Fermi liquid, essentially similar to the free gas: the distribution still has a discontinuity at k_F but with a reduced amplitude $Z < 1$. Blue line represents the Luttinger liquid, where quasiparticles are unstable excitations.	96

5-2	Single particle excitation in 1D induced by a photon (a) splits into an excitation containing only a charge degree of freedom and another one with spin (b).	100
5-3	Typical experimental setup with a nanotube of radius R suspended above an infinite metallic gate distance h away.	106
5-4	The inverse capacitance of a periodic array of (10,10) nanotubes as a function of the lattice spacing L on a log scale. The trend deviates slightly from logarithmic at small separations due to proximity effects.	108
5-5	Microscopic length scale λ as determined from our calculations for nanotubes of different sizes.	110
5-6	Luttinger liquid interaction parameter g as a function of the dimensionless ratio of two length scales. Here we set $v_F = 0.36$ a.u. = 8×10^5 m/s.	111
6-1	Left: potential of a single 2D charge in OBC. Right: potential of a 5x5 square array of 2D charges in OBC. Note the overall parabolic shape of the potential.	115
6-2	Left: potential of a 5x5 square array of 2D charges plotted together with the parabolic potential due to a uniform jellium of the same average density. Right: subtracting the jellium potential from the system yields precisely the same potential as the PBC system of point charge.	116
6-3	Left: Potential of images ϕ^{img} in an OBC array (i.e. one charge removed). Right: PBC potential with an isolated charge potential removed to show the parabolic correction.	117
6-4	Change in the force between ions of a H_2^+ molecule due to periodic images as a function of lattice spacing.	120
6-5	Left: because potentials of point charges are linear in 1D, the contribution of images vanishes. Lower curve shows $\phi^{img} = 0$. Right: Potentials of PBC and OBC 1D systems differ by exactly the parabolic potential ϕ^{jet} due to the fictitious jellium.	121

6-6	Left: Potential of a square array of 2D dipoles in OBC with one dipole removed. The average slope corresponds to the macroscopic electric field. Right: Zoom-in on the origin shows that the field there due to all other images vanishes.	122
6-7	Convergence of α_{\perp} and α_{\perp}^b with respect to L for a (5,5) SWNT. The point at L=10.6Å corresponds to a typical tube-tube separation in a bundle.	125
6-8	Convergence of the screening factor with respect to L for a (10,10) SWNT.	125
6-9	Response function versus the periodic inter-plane distance before and after the inhomogeneous electrostatic correction.	132
6-10	Schematic representation of the energy contribution from an isolated defect in bulk. Left: E_{def}^0 . Right: E_{def}^{∞}	136
6-11	The symmetry duality that allows us to calculate E_{def}^{∞}	137

List of Tables

3.1	Structural parameters for (n,n) armchair and $(n,0)$ zigzag nanotubes: translational period, diameter, and the number of carbon atoms in a unit cell.	46
4.1	Summary of $q \rightarrow 0$ scaling of dielectric response quantities of metals and insulators of different dimensionalities, as derived within RPA. ν is the density of states at the Fermi energy.	62
4.2	Radius, band gap, screening ratio, longitudinal and transverse polarizabilities (per unit length) of carbon nanotubes as a function of the chiral vector (n, m) . R_0 is the radius of the carbon backbone. The numbers in parentheses have been obtained with the electrostatic model for multi-wall CNTs.	78
4.3	Screening ratios, transverse polarizabilities (per unit length), longitudinal polarizabilities and radii of boron-nitride nanotubes as a function of the chiral vector (n, m) . R_0 is the radius of the B-N backbone and α_{\perp}^r is the full static polarizability that includes ionic response. The numbers in parentheses have been obtained with the electrostatic model for multi-wall BNNTs.	79

Chapter 1

Introduction

Recent sweeping advances in nanotechnology have elevated the importance of low-dimensional nanoscale structures, and a good understanding of their physical properties became essential. It is well-known that the dimensionality of a system has pronounced effects on the electron-electron interactions. In conventional three-dimensional metals elementary excitations form a weakly interacting Fermi liquid of quasiparticles, reminiscent of the independent electron picture. In two dimensions, electron-electron interactions in metals can produce exotic behaviors, such as the fractional quantum Hall effect, where excitations are very different from the bare electrons and can even carry fractional charge. Finally, in one-dimensional metals the conventional quasiparticles are not stable excitations at all even for weak interactions, and arising collective excitations determine thermodynamic and transport properties, as described within the Luttinger liquid paradigm. In semiconductors, static screening properties and the scaling of excitonic energies also depend critically on the dimensionality.

Even in the cases where the basic theory of electronic interactions is believed to be well understood, it is still a challenge to interpret results of experiments within the theoretical framework because nanoscale devices are extremely sensitive to their microscopic environment. Various environmental factors that can influence the outcome of measurements often cannot be incorporated into simple models and are still not well understood at the nano-scale. This important niche is filled by first principles calculations that bridge the gap between experiment and basic theory in that

they offer much more control over the realistic details of the system of interest and at the same time provide access to measurable quantities with close to experimental accuracy. Quantum-mechanical modeling is particularly valuable in situations where experimental control is difficult. Low-dimensional nanoscale structures, particularly nanotubes, are good examples of this situation – even though much is known about their properties, there is neither a way to grow a nanotube of a specified size and chirality, nor to fully characterize an already grown multi-wall nanotube.

The goal of this thesis is to combine *ab-initio* and analytical methods to develop a detailed understanding of the effects of electron-electron interactions on the properties of low-dimensional structures. Our study spans dielectric response properties of one- and two- dimensional insulators, metals, and zero-gap semiconductors, concentrating on carbon-based structures and comparing them to similar ones based on boron nitride. We study primarily nanotubes and single-layer sheets, materials that possess many fascinating and potentially useful properties, and whose theoretical and experimental characterization is far from complete. A complementary goal is to present applications of established and recently developed *ab-initio* methods to low-dimensional structures and highlight issues that arise in studying these materials from first principles.

1.1 Nanotubes

Carbon nanotubes have been first observed and reported in 1952 by Radushkevich and Lukyanovich [1] and have been rediscovered several times over the years, culminating in the famous papers by Iijima on single-wall carbon nanotubes in the early 1990's [2] that brought them to the awareness of the wide scientific community and generated a lot of excitement. Subsequently, the physics of carbon nanotubes has rapidly evolved into a research field of its own and significant progress has been made in understanding and characterizing their physical properties [3, 4]. They have attracted much attention for their exceptional properties, such as extremely high tensile strength, elasticity, excellent thermal conductivity, and ballistic conductance. One of

the most unusual properties of carbon nanotubes is their electronic structure which depends only on their geometry, i.e. chirality and diameter, and is unique in solid state physics. This property stems from the electronic properties of the graphene sheet and the constraints of the geometrical shape of a nanotube, resulting in either semiconducting or metallic tubes. Metallic carbon nanotubes are almost ideal quasi-one-dimensional quantum wires that can sustain extremely high current densities. Apart from the obvious applications as transistors and interconnects in electronics, carbon nanotubes have been envisioned to penetrate into various other areas of technology. The high surface-to-volume ratio, combined with structure-sensitive electronic properties, makes them a good candidate for chemical sensor applications. The high aspect-ratio and mechanical stability qualifies them for electron emitter and super-capacitor applications.

Unfortunately, synthesis and separation of specific nanotubes remains a central challenge that impedes further progress in scientific study and commercial applications. One approach may be to separate a random mixture of nanotubes after growth, based on their distinct characteristics. Variations in chirality and size influence dielectric properties, which in turn can be exploited for separation; e.g. electric fields have been used to align nanotubes during CVD synthesis [5, 6] and to separate different tubes in solutions [7, 8]. Another motivation for understanding of their dielectric response is the need to characterize optical excitations, screening at contacts, plasmons in nanotube arrays, and the degree of control achievable in endohedral fillings. In recent years the response of carbon single-wall nanotubes (SWNTs) has been studied with tight-binding [9, 10, 11] and first-principles approaches [12, 13], while multi-wall nanotubes (MWNTs) – a more common product of synthesis – have received much less attention due to their complexity. In this work we will examine the static dielectric properties of different types of nanotubes with the goals of engineering better ways of controlling their structure experimentally and, in general, of advancing the understanding of the electronic response of these unique one-dimensional systems.

By analogy with carbon nanotubes, boron-nitride nanotubes (BNNTs) were predicted to exist using ab-initio methods in 1993 [14], and were discovered shortly

afterwards [15]. They share good mechanical and thermal qualities with carbon nanotubes, while due to the insulating nature of the boron-nitride sheet, the electronic properties of these nanotubes are much more uniform and do not depend so strongly on the structure.

1.2 Graphene

Graphene is a relative newcomer in carbon physics, attracting a rapidly growing research interest, despite the fact that it has been known and studied for decades as an integral part of graphite and graphite-intercalated compounds. In the last decade, graphene has been viewed as a hypothetical building block whose understanding was a prerequisite to the study of fullerenes and nanotubes. By itself, the free standing single-layer graphene was thought not to even exist in a stable form, after Landau and Peierls argued that strictly two-dimensional crystals were thermodynamically unstable due to divergences of thermal fluctuations [16, 17]. The argument was later extended by Mermin [18] and is strongly supported by experimental observations that the melting temperature of many thin films rapidly decreases with decreasing thickness, and they become unstable. Atomic monolayers have so far been known only as an integral part of larger three-dimensional structures, usually grown epitaxially on top of substrates. Without being embedded in a 3D structure, 2D materials were presumed not to exist until the experimental discovery in 2004 of the graphene monolayer and other free-standing 2D atomic crystals, including boron-nitride [19, 20].

Among the remarkable properties of free-standing graphene are the extremely high carrier mobility, which only weakly depends on temperature, and the presence of the anomalous quantum Hall effect, observed even at room temperature. One indicator of the high crystal quality of graphene is that charge carriers can travel thousands of interatomic distances without scattering. Graphene is also exceptional from a theoretical point of view: its charge carriers mimic relativistic particles and are naturally described by the Dirac equation rather than the Schrodinger equation, although there is nothing particularly relativistic about electrons moving around carbon atoms.

This unusual electronic structure is responsible for a number of unique features in the static dielectric response of graphene, also inherited by carbon nanotubes, and we will describe them in the following chapters.

1.3 Outline

The thesis is organized as follows. In Chapter 2 we review the basic theory underlying the methods of first principles calculations that we use in this work, in particular density functional theory and density-functional perturbation theory. Chapter 3 contains a brief description of the intimate connection between of the geometric and electronic structures of graphene, boron-nitride and the corresponding nanotubes. Chapter 4 deals with calculations of dielectric linear response of these 2D materials, highlighting the unusual response of graphene. These results are then used to quantitatively understand the response of nanotubes to electric fields and to construct an accurate classical electrostatic model that captures the response of multi-wall nanotubes. In Chapter 5 we describe the effects of very strong correlations in metallic single-wall carbon nanotubes within the Luttinger liquid picture and present a method of calculating from first-principles the parameters of this model. Finally, in Chapter 6 we present methods, that we developed and used throughout this work, for increasing efficiency of ab-initio calculations of dielectric response of low-dimensional structures.

Chapter 2

Basics of ab-initio calculations

Quantum mechanics is extremely successful in describing microscopic properties of matter and has provided fundamental understanding of an enormous range of phenomena in physics, chemistry and biology. Despite the ambiguities in logical interpretation of the underlying equations, the incredible accuracy of quantitative predictions on the microscopic level make it such a powerful tool that hardly any branch of physical science today remains unaffected by quantum methods and concepts. Nanotechnology in particular owes its very existence to the possibility of calculating with high precision material quantities like the equilibrium lattice constants, vibrational spectra, dielectric tensors and even phase diagrams. The total energy of the ground state is a central quantity in quantum theory, and a vast majority of material properties are related to the total energy of the electron-ion system and its variation with various parameters or external perturbations. This explains the amount of work that has been done in developing efficient methods of total-energy calculations in recent years [21]. We present here briefly the fundamentals of density functional theory - the most widely studied and used method of first-principles calculations.

2.1 Many-electron problem

At the heart of non-relativistic quantum mechanics lies the many-body time-independent Schrödinger equation, which describes the system of ions at positions $\{\mathbf{R}\}$ and elec-

trons at $\{\mathbf{r}\}$

$$\hat{H}\Psi_i(\mathbf{r}, \mathbf{R}) = E_i\Psi_i(\mathbf{r}, \mathbf{R}) . \quad (2.1)$$

Solutions of this equation provide complete information about an ensemble of particles interacting with each other and with external fields. In practice we are only interested in the ground state solution; however, even this problem is impossible to treat in an exact way, and only a few cases are solvable analytically. Numerical solutions are also limited to very few particles, primarily due to the fact that there is no general way to decouple the equation for the $3N$ degrees of freedom into a set of independent equations. One is forced to resort to approximations, of which the adiabatic (Born-Oppenheimer) approximation is the most commonly used. It amounts to assuming that the motions of ions and electrons occur on such drastically different time scales that they can be decoupled. This means that in studying the electronic degrees of freedom the ions can be considered at rest, so that the total wavefunction can be written as a product of two functions: one describing the ions and another describing electrons while treating ionic positions as parameters

$$\Psi(\mathbf{r}, \mathbf{R}) = \Phi(\mathbf{R})\Psi_{\mathbf{R}}(\mathbf{r}) . \quad (2.2)$$

We are now left with a slightly easier problem corresponding to the electronic system with the nuclei fixed in the configuration $\{\mathbf{R}\}$. Electrons are fermions, and the electronic wavefunction is antisymmetric with respect to exchange of two electrons. The Hamiltonian for the electrons is given by

$$\hat{H} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} = \sum_i^N \left(-\frac{1}{2}\nabla_i^2 \right) + \sum_i^N v(\mathbf{r}_i) + \sum_{i<j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} , \quad (2.3)$$

$$v(\mathbf{r}_i) = -\sum_l \frac{Z_l}{|\mathbf{r}_i - \mathbf{R}_l|} , \quad (2.4)$$

where E is the electronic ground state energy, \mathbf{r}_i the coordinate of electron i , \mathbf{R}_I the coordinate of nucleus I , and Z_I the charge of nucleus I .¹ The total energy W , including the ion-ion interactions V_{nn} is then given by

$$W = E + V_{nn} = E + \sum_{I < J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (2.5)$$

Nearly all physical properties of a system can be calculated once the ground state wavefunction is known. This is the fundamental idea behind *first-principles* or *ab-initio* approaches: material properties are obtained from the solution of the Schrodinger equation without any experimental input. But we still have an interacting many-body quantum problem which is far too complicated to be solved exactly in a practical way. Further approximations are required to perform ab-initio calculations for real materials, and Density Functional theory (DFT) provides a framework for making progress.

2.2 Hohenberg-Kohn theorems

The first prescription for constructing the total energy in terms of only the electronic density was given by Thomas and Fermi [22]. They used the expression for the kinetic, exchange and correlation energies of the homogeneous electron gas to construct the same quantities for the inhomogeneous system in the local density approximation. In 1964 Hohenberg and Kohn formulated and proved two theorems that put those early ideas on solid mathematical ground. The first theorem states that no two different potentials acting on the electrons of a given system can give rise to the same ground-state electronic charge density, i.e. the external potential $v(\mathbf{r})$ is uniquely determined by the electronic density $n(\mathbf{r})$, up to a constant [23]. The external potential $v(\mathbf{r})$ then fixes the Hamiltonian and thus the ground state wavefunction Ψ_0 ; the ground state of a many-body system is therefore a unique functional of $n(\mathbf{r})$. Their second theorem introduces a formal variational principle on the charge density itself. Since

¹This equation and all others in this work, unless otherwise noted, are written in atomic units: $\hbar = m_e = 4\pi\epsilon_0 = 1$. Energies are in the units of Hartree [Ha] and distances are in [bohr].

Ψ_0 is determined by $n(\mathbf{r})$, one can define the functional

$$F[n(\mathbf{r})] \equiv \langle \Psi_0 | \hat{T}_e + \hat{V}_{ee} | \Psi_0 \rangle . \quad (2.6)$$

It is worth noting that $F[n(\mathbf{r})]$ is a universal functional independent of $v(\mathbf{r})$ since the kinetic energy T_e and the electron-electron interaction energy V_{ee} are functionals only of $n(\mathbf{r})$. Then the ground state energy functional is defined as

$$E_v[n(\mathbf{r})] \equiv F[n(\mathbf{r})] + \int v(\mathbf{r})n(\mathbf{r}) d\mathbf{r} . \quad (2.7)$$

The theorem states that for a given external potential $v(\mathbf{r})$, ground state density $n_0(\mathbf{r})$ and any trial electron density $\tilde{n}(\mathbf{r})$, the energy functional is minimized by the electron charge density of the ground state

$$E_0 = E_v[n_0] \leq E_v[\tilde{n}] . \quad (2.8)$$

Furthermore, the value of the minimum coincides with the ground-state energy E_0 . Therefore, the ground state density must satisfy the equation

$$\left[\frac{\delta F[n(\mathbf{r})]}{\delta n(\mathbf{r})} \right]_{n_0} = -v(\mathbf{r}) . \quad (2.9)$$

These theorems provide the foundation of what is currently known as *density-functional theory* [24]. The Hohenberg-Kohn approach represents a great conceptual simplification of the quantum-mechanical problem of the search of the ground-state properties of a system of interacting electrons, for it replaces the traditional description based on wave-functions with a much more tractable description in terms of the charge density. If the explicit form of the universal functional $F[n(\mathbf{r})]$ were known, the problem of determining the ground state would reduce to the problem of the minimization of a functional of the 3-dimensional electron density, i.e., a function of 3 coordinates, instead of $3N$ (N being the number of electrons). The Hohenberg and Kohn theorem implies that all physical properties of a system of interacting electrons are uniquely

determined by its ground-state charge-density distribution. The major problem that hampers a straightforward application of this remarkable result is that the form of the universal $F[n(\mathbf{r})]$ functional is of course unknown. This problem can be approached by mapping the system onto an auxiliary system of non interacting electrons [25] and by making appropriate approximations along the lines described in the next section.

2.3 Kohn-Sham equations

The Hohenberg-Kohn theorems hold independently of the precise form of the electron-electron interaction. In particular, in a non-interacting system, $F[n(\mathbf{r})]$ is simply the kinetic energy as a functional of the ground-state charge-density distribution, $T_0[n(\mathbf{r})]$. This fact has been used by Kohn and Sham to map the problem of a system of interacting electrons onto an equivalent non-interacting problem [25]. The key is to be able to reproduce the exact result for the non-interacting case, regardless of any approximations introduced to treat interactions. In this regard, Kohn and Sham argued that the ground state density of the interacting system can be represented as the ground state density of a *non-interacting* system in some local external potential $V_{KS}(\mathbf{r})$. Then according to the general density-functional formalism, the ground state density can be obtained by minimizing the non-interacting energy functional

$$E_{V_{KS}}^0[n(\mathbf{r})] \equiv T_0[n(\mathbf{r})] + \int V_{KS}(\mathbf{r})n(\mathbf{r}) d\mathbf{r} . \quad (2.10)$$

The unknown functional $F[n(\mathbf{r})]$ is written as

$$F[n(\mathbf{r})] = T_0[n(\mathbf{r})] + \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}[n(\mathbf{r})], \quad (2.11)$$

where the second term is the classical electrostatic self-interaction of the electron charge-density distribution, and the unknown last term is the so-called *exchange-correlation energy* is actually defined by Eq. (2.11). $E_{xc}[n]$ captures all remaining contributions including the complex many-body effects of exchange and correlation beyond the Hartree mean-field. Variation of the energy functional with respect to

$n(\mathbf{r})$ with the constraint that the number of electrons is kept fixed leads formally to the same equation that would hold for a system of non-interacting electrons subject to an effective self-consistent Kohn-Sham potential:

$$V_{KS}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}), \quad (2.12)$$

where

$$v_{xc}(\mathbf{r}) \equiv \frac{\delta E_{xc}}{\delta n(\mathbf{r})} \quad (2.13)$$

is the functional derivative of the exchange-correlation energy, also called the *exchange-correlation potential*. Although $v_{xc}(\mathbf{r})$ is a local potential, its functional dependence on the density is nonlocal. The resulting effective Hamiltonian for the system is the one describing a non-interacting gas feeling the effective potential V_{KS} in which all the interparticle interactions of the real system are contained

$$\hat{H}_{KS}\psi_i(\mathbf{r}) = \left(-\frac{1}{2}\nabla^2 + V_{KS}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}). \quad (2.14)$$

The ground-state charge-density distribution and non-interacting kinetic energy functional would then be given in terms of the *auxiliary* Kohn-Sham (KS) orbitals, $\psi_i(\mathbf{r})$:

$$n(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2 \quad (2.15)$$

$$T_0[n(\mathbf{r})] = -\frac{1}{2} \sum_{i=1}^N \int \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) d\mathbf{r}, \quad (2.16)$$

where N is the number of electrons. The ground-state energy given by Eqs. (2.11) can be equivalently expressed in terms of the KS eigenvalues:

$$E[n(\mathbf{r})] = \sum_{n=1}^N \epsilon_n - \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}[n(\mathbf{r})] - \int n(\mathbf{r})v_{xc}(\mathbf{r})d\mathbf{r}. \quad (2.17)$$

Eq. (2.14) has the form of a non-linear Schrödinger equation whose potential depends on its own eigenfunctions through the electron charge-density distribution. Once

an explicit form for the exchange-correlation energy is available, this equation can be solved in an iterative self-consistent way using a variety of methods. It should be noted that the wavefunctions ψ_i and eigenvalues ϵ_i appearing in the Kohn-Sham equations have no direct physical meaning: they are the eigenstates and eigenvalues of the auxiliary non-interacting system that are used in an intermediate step of the method, but cannot be considered the wave functions and energy levels for the electrons of the real system. Nevertheless, for lack of better and equally general methods, KS eigenvalues are often used to estimate excitation energies. One important exception is the highest occupied KS eigenvalue, which can be shown to correspond exactly to the true Fermi energy of an extended gapless system, through a theorem by Janak [26]. The general features of the low-lying energy bands in solids are generally considered to be at least qualitatively correct, in spite of the fact that oftentimes optical gaps in insulators are substantially underestimated. Though the Kohn-Sham approach involves solving for N one-electron wavefunctions instead of the charge density, the explicit treatment of wavefunctions allows very accurate approximations of the exact kinetic energy. This, combined with reasonable approximations for the exchange-correlation functional has led to remarkable predictive accuracy and has made first-principles calculations very successful. It is expected that this approach is very efficient in situations where the energy is dominated by the kinetic and electrostatic terms.

2.4 Exchange-correlation functionals

The Kohn-Sham scheme constitutes a practical way to implement DFT, provided an accurate and reasonably easy-to-use approximation is available for the exchange-correlation energy, $E_{xc}[n(\mathbf{r})]$. In their original paper [25], they used the idea that each small volume of the system can be considered locally uniform so that one can use the same exchange-correlation energy as an equal volume of a homogeneous electron gas at the same density. With this assumption, the exchange-correlation energy functional

and potential are:

$$E_{xc}^{LDA}[n(\mathbf{r})] = \int \epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r}, \quad (2.18)$$

$$v_{xc}^{LDA}(n(\mathbf{r})) = \frac{\delta E_{xc}^{LDA}[n(\mathbf{r})]}{\delta n(\mathbf{r})} = \left(\epsilon_{xc}(n) + n \frac{d\epsilon_{xc}(n)}{dn} \right)_{n=n(\mathbf{r})}, \quad (2.19)$$

where $\epsilon_{xc}(n)$ is the exchange-correlation energy per particle in a homogeneous electron gas at density n . This approximation is known as the *local density approximation* (LDA). It is designed to work with systems in which the electronic charge density is smooth, like in simple metals and semiconductors, but in practice it gives surprisingly good results for non-homogeneous systems, like covalently bonded materials. LDA typically gives good agreement with experiment for equilibrium structures and vibrational properties but tends to underestimate bond lengths. To improve on LDA, some extensions of the method were introduced, of which the Generalized Gradient Approximation (GGA) is one of the most popular [27]:

$$E_{xc}^{GGA}[n(\mathbf{r})] = \int f(n(\mathbf{r}), \nabla n(\mathbf{r}))d\mathbf{r}. \quad (2.20)$$

GGA gives more accurate results than LDA in many cases, but it is nonetheless not a systematic improvement. That can be improved by further powers of the gradient expansions.

2.5 Planewave basis calculations

For periodic solids, one needs to deal with infinitely many electrons and ions, and thus an infinite basis set to represent each particle's wavefunctions. This formidable problem can be surmounted by performing calculations on periodic systems and applying Bloch's theorem to the electronic wave functions [21]. According to this theorem, when the Hamiltonian operator \hat{H} displays translational symmetry,

$$\hat{H}\psi_i(\mathbf{r}) = \left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \quad (2.21)$$

where

$$v(\mathbf{r}) = v(\mathbf{r} + \mathbf{R}) \quad (\text{for all } \mathbf{R} \text{ in a Bravais lattice}), \quad (2.22)$$

the eigenfunction ψ_i can be chosen of the following form:

$$\psi_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}\psi_{n\mathbf{k}}(\mathbf{R}) \quad (2.23)$$

or equivalently,

$$\begin{aligned} \psi_{n\mathbf{k}}(\mathbf{r}) &= e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}) \\ u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) &= u_{n\mathbf{k}}(\mathbf{r}). \end{aligned} \quad (2.24)$$

which essentially expresses an eigenfunction as a lattice-periodic part multiplied by a plane-wave envelope. Each wavefunction ψ is thus labeled by a wave vector \mathbf{k} , the crystal momentum, and a discrete band index n which classifies states corresponding to the same \mathbf{k} . Due to translational invariance, different \mathbf{k} -points can be treated independently and the Schrödinger equation becomes decoupled:

$$\left[-\frac{1}{2}(\nabla + i\mathbf{k})^2 + v(\mathbf{r}) \right] u_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r}), \quad (2.25)$$

where $\{\psi_{n\mathbf{k}}\}$ are orthonormal wavefunctions, i.e. $\langle \psi_{n\mathbf{k}} | \psi_{n'\mathbf{k}'} \rangle = \delta_{nn'}\delta_{\mathbf{k}\mathbf{k}'}$. Eq. 2.25 must be solved for all \mathbf{k} inside the Brillouin zone (BZ) of the reciprocal space, and physical quantities are integrals over the BZ, e.g. the charge density is given by

$$n(\mathbf{r}) = \sum_{\mathbf{k}} \sum_n |u_{n\mathbf{k}}(\mathbf{r})|^2 \quad (2.26)$$

where the n -sum runs over the occupied bands at \mathbf{k} . While, in principle, an infinite number of \mathbf{k} vectors must be sampled to obtain the exact result (equivalent to simulating a crystal of infinite dimensions), in practice, the number of \mathbf{k} -vectors can be systematically increased until the physical quantities of interest converge within the desired accuracy. The required number of \mathbf{k} -points can be significantly reduced by using symmetries of the crystal of interest, via the so-called special point technique [28]. A practical way of solving the KS equation (2.25) is achieved by expanding the

electronic wavefunctions in a finite basis of plane waves. This is a natural choice of basis functions for studying periodic crystalline systems, and it allows to take advantage of efficient algorithms, such as the Fast Fourier Transform (FFT), to go between the real and reciprocal spaces. A wavefunction in a plane-wave basis is expressed as

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}}^{N_{pw}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} , \quad (2.27)$$

or

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}}^{N_{pw}} c_{n\mathbf{k}}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} , \quad (2.28)$$

where \mathbf{G} is a reciprocal lattice vector. The sum is customarily taken over the set of \mathbf{G} vectors satisfying $\frac{1}{2}|\mathbf{k} + \mathbf{G}|^2 < E_{cut}$, where E_{cut} is the plane-wave cutoff energy. Using this expansion, the KS equation can be written as an algebraic equation in reciprocal space:

$$\sum_{\mathbf{G}'} \left[\frac{1}{2} |\mathbf{k} + \mathbf{G}|^2 + v_H(\mathbf{G} - \mathbf{G}') + v_{ext}(\mathbf{G} - \mathbf{G}') + v_{xc}(\mathbf{G} - \mathbf{G}') \right] c_n(\mathbf{k} + \mathbf{G}') = \epsilon_{kn} c_n(\mathbf{k} + \mathbf{G}) \quad (2.29)$$

The Hamiltonian in this representation is block-diagonal with respect to \mathbf{k} -vectors and diagonalization can be performed separately on each block. A great advantage of this basis set is that it treats all parts of a general inhomogeneous system on an equal footing, unlike a basis of localized orbitals. The accuracy of the calculation is directly controlled by the parameter E_{cut} , which determines the size of the basis set to represent all wavefunctions. Once the atomic coordinates, ionic potentials and total number of electrons are supplied, the Fourier components of the wavefunctions are computed iteratively without any further assumptions regarding the system structure. Thus all the symmetries and degeneracies of the electronic states (such as the K-point of graphene) are correctly captured. A disadvantage of the planewave basis is that the number of basis functions is usually large compared to a localized basis, especially for the case of isolated systems. We will further address this concern in Chapter 6 and introduce an efficient technique to deal with such scaling issues. Another disadvantage

of the plane-wave basis is the fact that all regions of space are resolved on the same uniform grid, whereas it is clear that close to ionic cores the charge density varies vary rapidly and in principle requires a very large number of \mathbf{G} vectors to be accurately described. One way around this difficulty is the *pseudopotential* method, which is based on the assumption that the most relevant physical and chemical properties of a system are due to its valence electrons only, while the ionic cores with the inner electrons can be considered frozen. The valence electrons thus feel an effective field produced by these inert ionic cores and a pseudopotential tries to reproduce this field on the outer valence electrons without explicit calculation of the core states. The total number of electrons is thus reduced to the number of valence electrons, and a much smaller number of basis functions is required when only the smoother valence electron wavefunctions are involved. A typical norm-conserving pseudopotential, where the valence pseudo-wavefunctions satisfy the orthonormality condition $\langle \psi_i^{PS} | \psi_j^{PS} \rangle = \delta_{ij}$, is constructed following the prescriptions of Ref. [29]. Norm-conserving pseudopotentials require a relatively high E_{cut} and since the computational cost grows as the planewave cutoff increases, a smoother pseudo wavefunction would be desirable. Most of our calculations are done using ultrasoft pseudopotentials, that generate optimally smooth pseudo wavefunctions by relaxing the norm-conserving condition [30].

2.6 Density functional perturbation theory

Once the approximate ground state charge density is known, it is possible to conveniently calculate linear response quantities within the framework of density-functional perturbation theory (DFPT), which we now briefly discuss. Let us add a perturbation to the original external potential $v_\lambda(\mathbf{r}) = v(\mathbf{r}) + \lambda\Delta v(\mathbf{r})$ parametrized by λ . Then, according to the Hellman-Feynman theorem, the first and second derivatives of the

ground-state energy are:

$$\frac{\partial E}{\partial \lambda} = \sum_n \langle \psi_n | \frac{\partial v_\lambda(\mathbf{r})}{\partial \lambda} | \psi_n \rangle = \int \Delta v(\mathbf{r}) n_\lambda(\mathbf{r}) d\mathbf{r}, \quad (2.30)$$

$$\frac{\partial^2 E}{\partial \lambda^2} = \int \frac{\partial n_\lambda(\mathbf{r})}{\partial \lambda} \Delta v(\mathbf{r}) d\mathbf{r} \quad (2.31)$$

The electron-density response, $\partial n_\lambda(\mathbf{r})/\partial \lambda$, appearing in (2.31) can be evaluated by linearizing Eqs. (2.15), (2.14), and (2.12) with respect to wave-function, density, and potential variations. A general feature of linear response is that only the first-order density response is needed to compute the second derivative of the total energy with respect to the perturbation (e.g. phonon frequency or dielectric susceptibility). Linearization of the density variation yields

$$\Delta n(\mathbf{r}) = 2 \sum_{n=1}^N \psi_n^*(\mathbf{r}) \Delta \psi_n(\mathbf{r}). \quad (2.32)$$

The variation of the KS orbitals, $\Delta \psi_n(\mathbf{r})$, is obtained by standard first-order perturbation theory:

$$(H_{KS} - \epsilon_n) |\Delta \psi_n\rangle = -(\Delta V_{KS} - \Delta \epsilon_n) |\psi_n\rangle, \quad (2.33)$$

where

$$H_{KS} = -\frac{1}{2} \nabla_{\mathbf{r}}^2 + V_{KS}(\mathbf{r}) \quad (2.34)$$

is the unperturbed Kohn-Sham Hamiltonian, and

$$\Delta V_{KS}(\mathbf{r}) = \Delta V(\mathbf{r}) + \int \frac{\Delta n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \left. \frac{dv_{xc}(n)}{dn} \right|_{n=n(\mathbf{r})} \Delta n(\mathbf{r}), \quad (2.35)$$

is the first-order correction to the self-consistent potential, and $\Delta \epsilon_n = \langle \psi_n | \Delta V_{KS} | \psi_n \rangle$ is the first order variation of the KS eigenvalue, ϵ_n . Equations (2.32)–(2.35) form a set of self-consistent equations for the perturbed system completely analogous to the KS equations in the unperturbed case—Eqs. (2.12), (2.14), and (2.15)—with the KS eigenvalue equation, Eq. (2.14), being replaced by the solution of a linear system, Eq. (2.33). A great benefit of this method is that, once reformulated in Fourier

space, all perturbing frequencies decouple to linear order, and response to a periodic perturbation of a given wave-vector q can be solved in a separate calculation (see (4.12) below). For a detailed review of DPFT and its applications see Ref. [31].

Chapter 3

Structure of graphene and carbon nanotubes

3.1 Electronic structure of graphene

Graphene is a single sheet of carbon atoms forming a hexagonal grid. Every atom has four valence electrons - one $2s$ and three $2p$, and electronic wave functions of neighboring atoms overlap. While the s , p_x , and p_y orbitals are symmetric with respect to reflections across the graphene plane, the p_z orbital changes sign. Therefore, the overlaps of p_x and p_y with p_z orbitals vanish and p_z orbitals can be treated independently of the other valence electrons. The hybridization of the in-plane orbitals produces a hexagonal network of σ bonds, while p_z orbitals form π bonds that point perpendicular to the plane. The energies of σ bands lie far away from the Fermi level, so that they contribute very little to the electronic and chemical properties of graphene and nanotubes (apart from curvature effects in narrow tubes). As we will see shortly, the bonding and anti-bonding π bands actually cross at the Fermi level and they are responsible for the unique semi-metallic nature of graphene. In the description of the electronic excitations we will rely on a tight-binding formulation, which translates the problem into that of finding the spectrum of the hopping Hamiltonian for the fermion

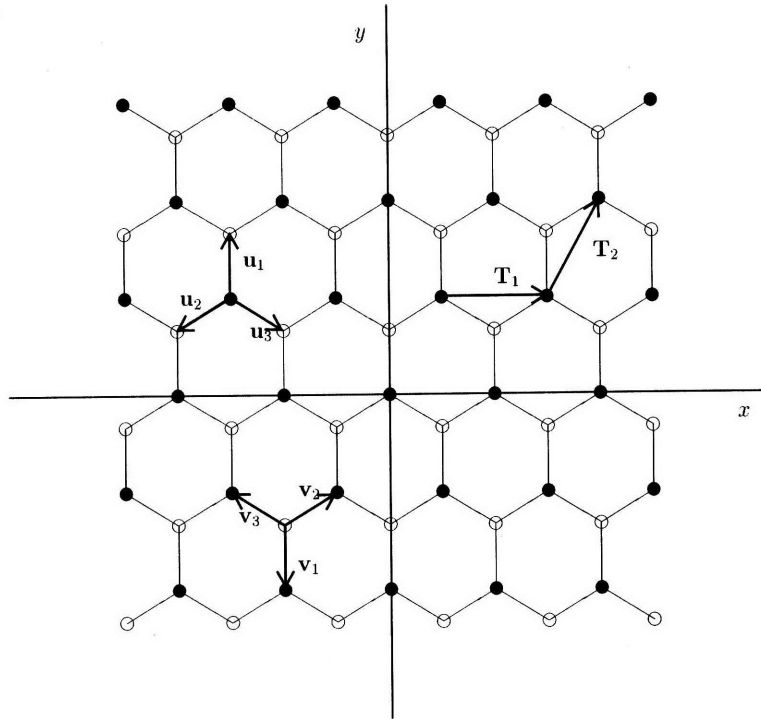


Figure 3-1: Honeycomb lattice consists of two sublattices A(●) and B(○).

operators a_i, a_j^\dagger

$$H = \gamma \sum_{\langle i,j \rangle} a_i^\dagger a_j \quad (3.1)$$

with the sum running over nearest neighbors i, j in the lattice and a_i, a_j^\dagger are the canonically anticommuting operators

$$\{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0 \quad \{a_i, a_j^\dagger\} = \delta_{ij} \quad (3.2)$$

When the single hopping parameter γ is appropriately adjusted based on experimental data, the tight-binding model works exceptionally well in reproducing the features of the conduction and valence π bands. It is also important to stress that, in spite of neglecting the coulombic interaction between electrons, the tight-binding approximation predicts energy levels which are in good quantitative agreement with existing experimental results. The set of discrete translations under which the lattice potential

$V(\mathbf{r})$ is invariant can be generated by 2D vectors \mathbf{T}_1 and \mathbf{T}_2 so that

$$V(\mathbf{r} + p\mathbf{T}_1 + q\mathbf{T}_2) = V(\mathbf{r}) \quad \forall p, q \in Z \quad (3.3)$$

According to Bloch theorem 2.24 the energy eigenfunctions of the quantum problem can be expressed as

$$\Psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \quad (3.4)$$

with $u_{n\mathbf{k}}(\mathbf{r})$ obeying

$$u_{n\mathbf{k}}(\mathbf{r} + p\mathbf{T}_1 + q\mathbf{T}_2) = u_{n\mathbf{k}}(\mathbf{r}) \quad \forall p, q \in Z \quad (3.5)$$

In the above expression n stands for the band label, while \mathbf{k} labels different states in a band within the Brillouin zone. We can write the energy eigenfunctions (3.4) as a linear combination of the normalized atomic orbitals ϕ_n

$$\Psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i} \phi_n(\mathbf{r} - \mathbf{r}_i) \quad (3.6)$$

where N is the number of unit cells in the crystal and the sum runs over all the lattice points

$$\mathbf{r}_i = p_i\mathbf{T}_1 + q_i\mathbf{T}_2 \quad p_i, q_i \in Z \quad (3.7)$$

It can be shown that (3.6) is a Bloch wavefunction and, vice versa, that an eigenfunction with the property (3.4) can always be cast in the form (3.6). One can then insert the linear combination (3.6) in the computation of the energy eigenvalues. When the orbitals ϕ are localized around each lattice site, we can approximate the numerator by integrals involving only nearest neighbor orbitals, which is the essence of the tight-binding method. This gives for the energy levels:

$$E_{\mathbf{k}} = \int d^2r \phi^*(\mathbf{r}) H \phi(\mathbf{r}) + \sum_{\substack{\text{nearest} \\ \text{neighbors}}} e^{i\mathbf{k}\cdot\mathbf{r}_i} \int d^2r \phi^*(\mathbf{r}) H \phi(\mathbf{r} - \mathbf{r}_i) \quad (3.8)$$

where the sum runs over all nearest neighbors i of the origin. We can generalize the above procedure to two atoms per unit cell, labeled A and B , and take as generators of the lattice the vectors

$$\mathbf{T}_1 = \sqrt{3}\mathbf{e}_x \quad \mathbf{T}_2 = \frac{\sqrt{3}}{2}\mathbf{e}_x + \frac{3}{2}\mathbf{e}_y \quad (3.9)$$

One atom is placed at the origin of the primitive cell and another is at $\mathbf{d} = \mathbf{e}_y$, so that the honeycomb lattice is obtained by repeated application of the generators. The wavefunctions ϕ may be thought of being composed of two orbitals ϕ_A and ϕ_B localized respectively around each of the two atoms in the unit cell. We make here our first approximation by writing a convenient linear combination of these two orbitals as

$$\phi(\mathbf{r}) = c_A\phi_A(\mathbf{r}) + c_B e^{i\mathbf{k}\cdot\mathbf{d}}\phi_B(\mathbf{r} - \mathbf{d}) \quad (3.10)$$

By introducing this expression into the energy functional we get, in the tight-binding approximation,

$$\begin{aligned} E_{\mathbf{k}} = & c_A^*c_A \int d^2r \phi_A^*(\mathbf{r})H\phi_A(\mathbf{r}) + c_B^*c_B \int d^2r \phi_B^*(\mathbf{r} - \mathbf{d})H\phi_B(\mathbf{r} - \mathbf{d}) \\ & + c_A^*c_B \sum_j e^{i\mathbf{k}\cdot\mathbf{u}_j} \int d^2r \phi_A^*(\mathbf{r})H\phi_B(\mathbf{r} - \mathbf{u}_j) \\ & + c_B^*c_A \sum_j e^{i\mathbf{k}\cdot\mathbf{v}_j} \int d^2r \phi_B^*(\mathbf{r} - \mathbf{d})H\phi_A(\mathbf{r} - \mathbf{d} - \mathbf{v}_j) \end{aligned} \quad (3.11)$$

where $\{\mathbf{u}_j\}$ is a triad of vectors of unit length pointing respectively in the direction of the nearest neighbors of a A point, and $\{\mathbf{v}_j\}$ the triad made of their respective opposites (see Fig. 3-1). The symmetry of the problem imposes that

$$\int d^2r \phi_A^*(\mathbf{r})H\phi_A(\mathbf{r}) = \int d^2r \phi_B^*(\mathbf{r} - \mathbf{d})H\phi_B(\mathbf{r} - \mathbf{d}) = \beta \quad (3.12)$$

$$\int d^2r \phi_A^*(\mathbf{r})H\phi_B(\mathbf{r} - \mathbf{u}_i) = \int d^2r \phi_B^*(\mathbf{r} - \mathbf{d})H\phi_A(\mathbf{r} - \mathbf{d} - \mathbf{v}_i) = \gamma \quad (3.13)$$

In the basis $\{\phi_A, \phi_B\}$ the energy functional can now be written as a quadratic form

$$E_{\mathbf{k}} = \begin{pmatrix} c_A^* & c_B^* \end{pmatrix} \begin{pmatrix} \beta & \gamma \sum_j e^{i\mathbf{k}\cdot\mathbf{u}_j} \\ \gamma \sum_j e^{i\mathbf{k}\cdot\mathbf{v}_j} & \beta \end{pmatrix} \begin{pmatrix} c_A \\ c_B \end{pmatrix} \quad (3.14)$$

which can be viewed as a variational expression for the energy. This is equivalent to a problem of diagonalization of the above quadratic form. For simplicity we disregard the diagonal contribution in (3.14) and set $\beta = 0$ because it is independent of \mathbf{k} and will lead to minor changes of the spectrum in the end (see Ref. [3] for details).

Equivalently, we can see that what we are doing actually is diagonalizing the hopping hamiltonian (3.1). To solve this problem, we note that an eigenstate of the Hamiltonian must also be an eigenstate of the translations \mathbf{T}_1 and \mathbf{T}_2

$$\Psi = \sum_{i \in A} c_A e^{i\mathbf{k}\cdot\mathbf{r}_i} a_i^\dagger |O\rangle + \sum_{i \in B} c_B e^{i\mathbf{k}\cdot\mathbf{r}_i} a_i^\dagger |O\rangle \quad (3.15)$$

assigning different coefficients c_A and c_B to respective sublattices, as depicted in Fig. 3-1. The two sublattices are mapped into each other by the action of the Hamiltonian and we have

$$\begin{aligned} H\Psi &= \gamma \sum_{i \in A} \sum_{\langle i,j \rangle} c_B e^{i\mathbf{k}\cdot\mathbf{r}_j} a_i^\dagger |O\rangle + \gamma \sum_{i \in B} \sum_{\langle i,j \rangle} c_A e^{i\mathbf{k}\cdot\mathbf{r}_j} a_i^\dagger |O\rangle \\ &= \gamma \sum_j e^{i\mathbf{k}\cdot\mathbf{u}_j} \sum_{i \in A} c_B e^{i\mathbf{k}\cdot\mathbf{r}_i} a_i^\dagger |O\rangle + \gamma \sum_j e^{i\mathbf{k}\cdot\mathbf{v}_j} \sum_{i \in B} c_A e^{i\mathbf{k}\cdot\mathbf{r}_i} a_i^\dagger |O\rangle \end{aligned} \quad (3.16)$$

It is clear that the state (3.15) is an eigenvector of H provided that the coefficients c_A and c_B are solutions of the eigenvalue problem

$$\begin{pmatrix} 0 & \gamma \sum_j e^{i\mathbf{k}\cdot\mathbf{u}_j} \\ \gamma \sum_j e^{i\mathbf{k}\cdot\mathbf{v}_j} & 0 \end{pmatrix} \begin{pmatrix} c_A \\ c_B \end{pmatrix} = E_{\mathbf{k}} \begin{pmatrix} c_A \\ c_B \end{pmatrix} \quad (3.17)$$

This is just another way of writing the variational problem (3.14). From (3.17) after

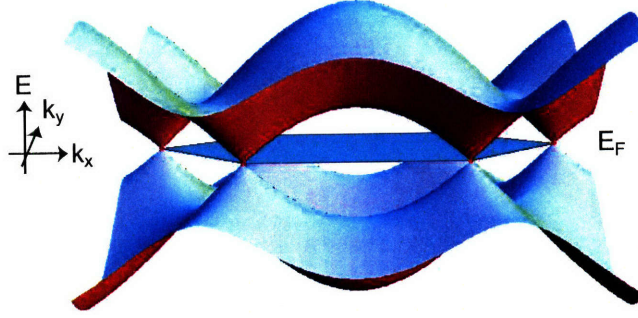


Figure 3-2: Bands of graphene as obtained from the tight binding calculation. Courtesy of Jonas Hauptmann.

a bit of algebra we derive the dispersion relation

$$E_{\mathbf{k}} = \pm \gamma \sqrt{1 + 4 \cos^2 \frac{\sqrt{3}}{2} k_x + 4 \cos \frac{\sqrt{3}}{2} k_x \cos \frac{3}{2} k_y} \quad (3.18)$$

The structure of this band has very striking properties when considered at half-filling. This situation physically corresponds to graphene, where each site of the honeycomb lattice is occupied by one electron, while each level of the band may accommodate two states due to the spin degeneracy, and the Fermi level turns out to be at the midpoint of the band, $E_{\mathbf{k}} = 0$. The honeycomb lattice at half-filling has, therefore, six isolated Fermi points, instead of a whole Fermi line. The reciprocal lattice is generated by

$$\mathbf{K}_1 = \frac{2\pi}{\sqrt{3}} \mathbf{e}_x - \frac{2\pi}{3} \mathbf{e}_y \quad \mathbf{K}_2 = \frac{4\pi}{3} \mathbf{e}_y \quad (3.19)$$

and the Brillouin zone is a hexagon, as shown in Fig. 3.1. The only points which reach the Fermi level are the six vertices of the hexagon

$$\begin{aligned} k_x &= \pm \frac{4\pi}{3\sqrt{3}} & k_y &= 0 \\ k_x &= \pm \frac{2\pi}{3\sqrt{3}} & k_y &= \pm \frac{2\pi}{3} \end{aligned} \quad (3.20)$$

It can be checked that these are the only roots of $E_{\mathbf{k}} = 0$. But due to translation symmetry in the reciprocal space, only two of the Fermi points are independent, and

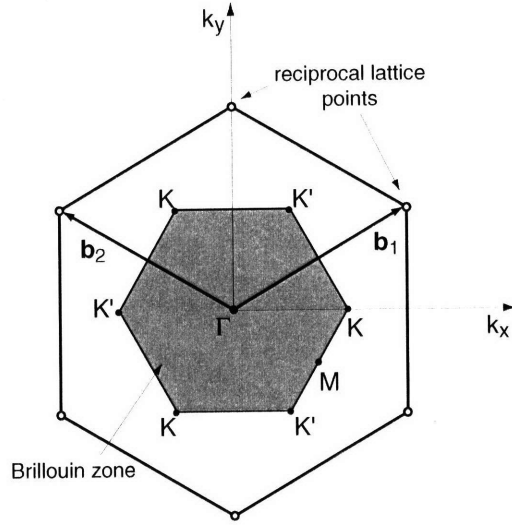


Figure 3-3: Brillouin zone of graphene.

are commonly referred to as K- and K'-points. The existence of a finite number of Fermi points at half-filling has important consequences in the description of the spectrum about the Fermi level. The low energy excitations can be studied by taking the continuum limit at any two independent Fermi points. For this purpose we introduce a parameter a measuring the carbon-carbon bond length and expand the 2×2 operator in (3.17) around any of two independent Fermi points. At the first Fermi point in (3.20), for instance, we have

$$\mathbf{k} = \frac{4\pi}{3\sqrt{3}}\mathbf{e}_x + \delta\mathbf{k} \quad (3.21)$$

and

$$\mathcal{H} \equiv \begin{pmatrix} 0 & \gamma \sum_j e^{ia\mathbf{k} \cdot \mathbf{u}_j} \\ \gamma \sum_j e^{ia\mathbf{k} \cdot \mathbf{v}_j} & 0 \end{pmatrix} \approx -\frac{3}{2}\gamma a \begin{pmatrix} 0 & \delta k_x + i\delta k_y \\ \delta k_x - i\delta k_y & 0 \end{pmatrix} + O((a \delta k)^2) \quad (3.22)$$

which can be written in terms of the Pauli matrices

$$\mathcal{H} \approx -\frac{3}{2}a\gamma \boldsymbol{\sigma}^T \cdot \delta\mathbf{k} \quad (3.23)$$

The effective Hamiltonian in the continuum limit turns out to be the Dirac operator in two dimensions, which means that the low energy excitations of the honeycomb lattice at half-filling are described by an effective theory of two two-dimensional Dirac spinors. This is very different from the usual effective mass continuum approximations, where a quadratic dispersion relation at high symmetry points of the Brillouin zone gives rise to an effective Schrodinger equation, with one parameter (the mass) chosen to reproduce the band curvature. We should note that the conical shape of the bands of graphene is a robust feature that is not restricted to nearest-neighbor tight-binding models. Rather it is due entirely to the group symmetry of the lattice [32] and is also observed in DFT band structure calculations [33].

For the purposes of a comparison, we mention that sheets of hexagonal boron nitride possess a closely related sp^2 -bonded structure, where A sites are replaced by boron and those at B sites by nitrogen. However, analogous tight binding calculations (e.g. [14]) show that in this case the inequivalence of the two sublattices results in lower symmetry and prevents the two bands from touching. The two bands are separated by a wide gap of about 5.8 eV, and the usual effective mass theory is the appropriate description in this case. As we will later see, this results in dramatic qualitative differences in the low energy and long-wavelength properties of the two 2D materials, as well as of the nanotubes formed by rolling up the corresponding sheets.

3.2 Geometric and electronic structure of carbon nanotubes

A tube made of a single graphene layer rolled up into a hollow cylinder is called a single-wall carbon nanotube (SWCNT); a tube consisting of several concentric cylinders is called a multi-wall nanotube (MWCNT). This rolling happens under the right growth conditions, since it is energetically favorable to avoid the dangling bonds that are on the edges of a graphite sheet, since these dangling bonds corresponds to very

high energy states. To avoid these high energy states the atoms cluster together into structures such as smaller fullerenes or nanotubes. This is energetically favorable, in spite of the cost in energy due to the strain that this roll-up puts on the bonds. In experiments, diameters of SWCNTs are observed to be in the range 1 to 2 nm, while diameters of MWCNTs vary from 5 to 200 nm. For a qualitative understanding of the electronic structure of single-wall carbon nanotubes, it is sufficient to start with the band structure of graphene and use the zone folding approach. This method allows an easy calculation of band structures of nanotubes of all sizes and chiralities, something more elaborate methods do not. We begin by describing the geometry of a nanotube using the graphene lattice. The unit cell of a carbon nanotube is defined by the chiral vector \mathbf{C} and the translational vector \mathbf{T} . Here \mathbf{C} , which is the roll-up direction and defines the circumference of the tube, can be expressed as a linear combination of the two lattice vectors of graphene, \mathbf{a}_1 and \mathbf{a}_2 (whose length is $|\mathbf{a}_1| = |\mathbf{a}_2| = a_0 = 2.46 \text{ \AA}$) as in Fig. 3-4:

$$\mathbf{C} = n\mathbf{a}_1 + m\mathbf{a}_2 \equiv (n, m) \quad (n, m \text{ are integers, } 0 \leq |m| \leq n). \quad (3.24)$$

Thus each nanotube is characterized by an index (n, m) that is derived from that of the chiral vector. The translational vector \mathbf{T} defines the translational period along the tube axis and is expressed as

$$\mathbf{T} = \frac{2m+n}{\mathcal{R}}\mathbf{a}_1 - \frac{2n+m}{\mathcal{R}}\mathbf{a}_2, \quad (3.25)$$

where \mathcal{R} is the greatest common divisor of $(2m+n)$ and $(2n+m)$ [3]. The highest symmetry cases, containing mirror symmetries, are the so-called achiral nanotubes: (n, n) armchair and $(n, 0)$ zigzag tubes. They have $4n$ atoms per unit cell and are convenient to study within periodic boundary conditions along the axis. Chiral tubes, on the other hand, generally have very long unit cells, containing many more atoms, and are thus less amenable to computational studies. A chiral tube and two types of achiral tubes are shown on the right of Fig. 3-4. The structural parameters for (n, n) armchair and $(n, 0)$ zigzag tubes are summarized in Table 3.1.

	$ \mathbf{T} $	$2R$	N_A
armchair (n, n)	a_0	$\sqrt{3}na_0/\pi$	$4n$
zigzag $(n, 0)$	$\sqrt{3}a_0$	na_0/π	$4n$

Table 3.1: Structural parameters for (n,n) armchair and $(n,0)$ zigzag nanotubes: translational period, diameter, and the number of carbon atoms in a unit cell.

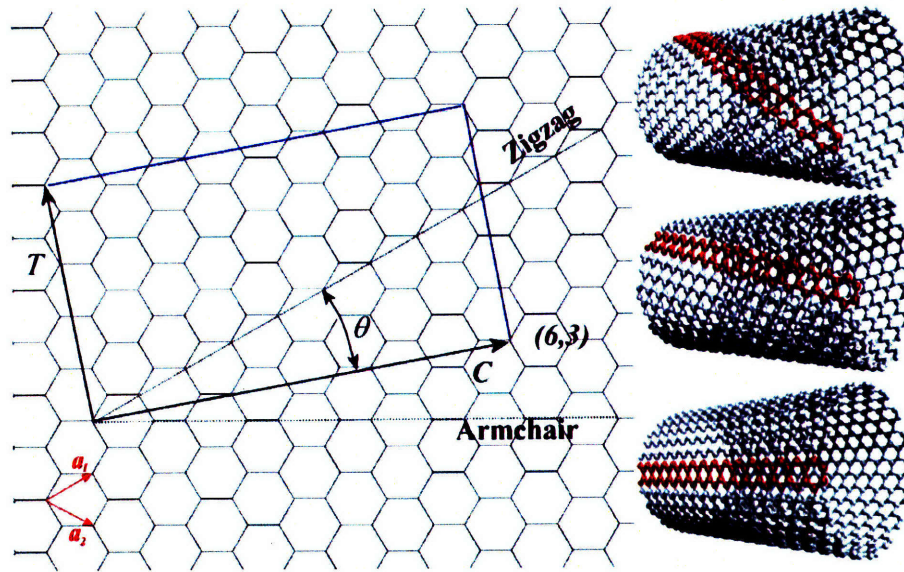


Figure 3-4: Chiral and translational vectors of single-wall nanotubes. Structures on the right are examples of zigzag, general chiral, and armchair nanotubes. Reproduced from <http://www.nanoelectronics.jp>

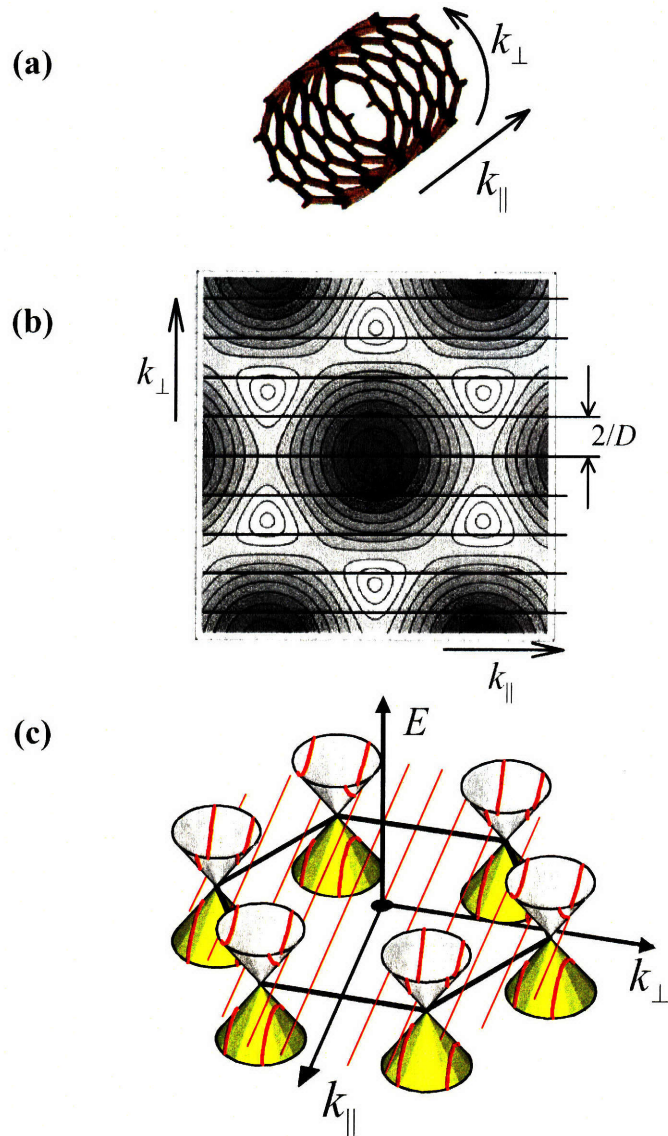


Figure 3-5: (a),(b) The transverse k-vector component k_{\perp} is quantized when periodic boundary conditions are applied. (c) The values of k_{\perp} allowed by quantization slice through the bands of graphene. Pictures courtesy of Paul McEuen's group.

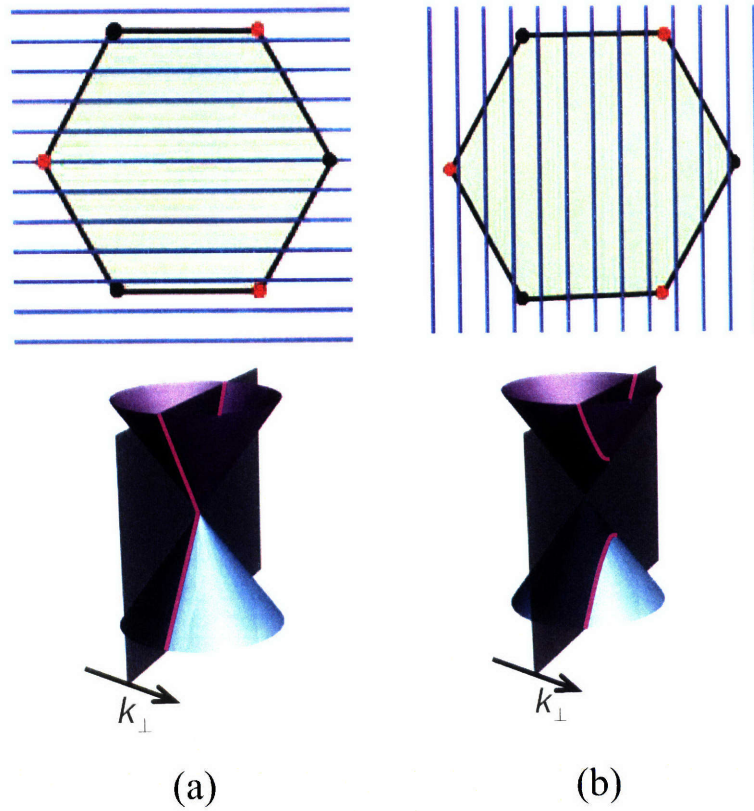


Figure 3-6: Quantization of the k -vectors in the Brillouin zone and the resulting bands.

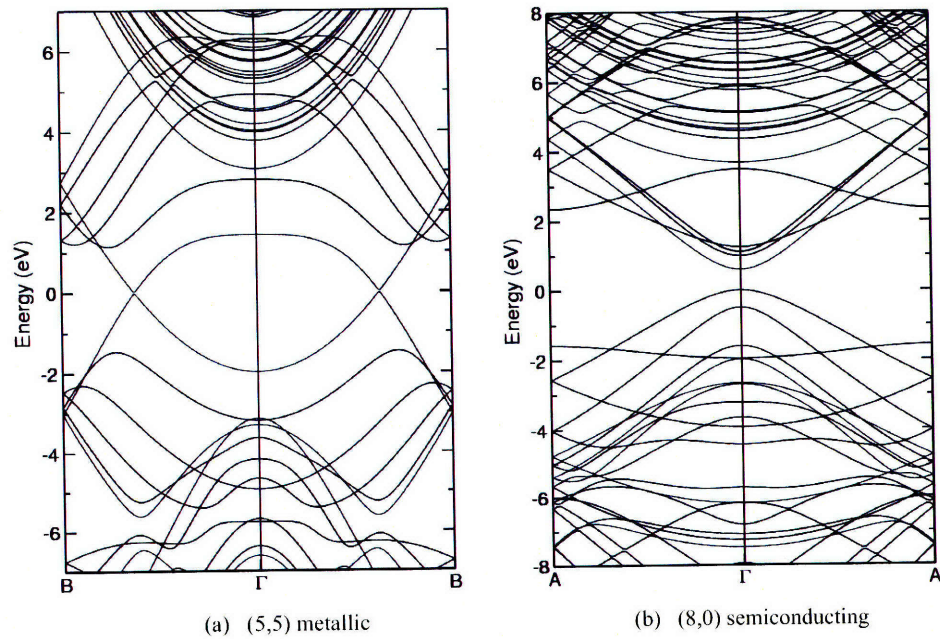


Figure 3-7: Band structure of armchair and zigzag carbon SWNTs obtained from DFT calculations. Courtesy of Young-Su Lee.

Like their crystalline structure, the electronic structure of a CNT can be derived from that of two-dimensional graphene by imposing the geometrical constraints from the nanotube on the band structure of graphene. In the case of carbon nanotubes, confinement of the electron wavefunction along the circumference only allows discrete wave vectors in the direction of \mathbf{C} . Along the axis, the electronic states are Bloch functions as in usual extended systems. Thus, the allowed wave-vectors appear as a series of parallel lines in the Brillouin zone of graphene, as illustrated in Fig. 3-5. The most striking consequence of this quantization is that single-wall carbon nanotubes can be metallic or semiconducting. Indeed, as we saw earlier graphene is a zero-gap semiconductor: conduction and valence band touch each other at the corners of the Brillouin zone (K -points). When the lines allowed by the quantization pass through the K -point, a CNT is metallic; otherwise it is semiconducting, as shown in Fig. 3-7. In terms of the (n,m) index, this zone folding can be summarized as following:

$$\text{mod}(n - m, 3) = \begin{cases} 0 & : \text{ metallic} \\ 1, 2 & : \text{ semiconducting} \end{cases} . \quad (3.26)$$

Thus a (5,5) CNT is metallic while an (8,0) CNT is semiconducting. Since the zone folding approach ignores curvature and the accompanying orbital re-hybridizations, exceptions exist in small-diameter tubes. However, armchair CNTs are always metallic and have two degenerate eigenstates at the Fermi energy, regardless of diameter. This dependence of electronic character only on geometry is unlike anything else in condensed matter theory, and this is partly the reason why carbon nanotubes captured so much interest, particularly in the possible electronics applications. A more detailed discussion can be found in Refs. [3] and [4], and references therein.

Chapter 4

Static dielectric response of sheets and tubes

This chapter begins with a comprehensive picture of the static dielectric response of graphene and boron-nitride sheets, computed with several ab-initio methods. We draw a comparison between the two structures in order to illustrate the unusual screening properties of graphene, such as the scale-invariance of the response of π -electrons, which dominates at long wavelengths. We find that the in-plane response of graphene is intermediate between a 2D insulator and a metal, while the response of the boron-nitride sheet is that of an insulator. In the second half of the chapter we use these results to explain quantitatively the differences of the calculated dielectric response of carbon and boron-nitride single-wall nanotubes, and generalize to the case of multi-wall tubes. We find that the longitudinal polarizability of SWNTs scales as the inverse square of the band gap, while in MWNTs and bundles it is given by the sum of the polarizabilities of the constituent tubes. The transverse polarizabilities and screening factors of SWNTs of both types are insensitive to band gaps and chiralities but exhibit qualitatively different radius dependence. As expected, for carbon nanotubes the transverse response is intermediate between metallic and insulating, and for boron-nitride it is insulating. We generalize the single-wall results with the construction of a simple and accurate electrostatic model of multi-wall nanotubes that captures the dependence of the static transverse dielectric response on radii and

number of layers for both CNTs and BNNTs. We conclude this chapter by highlighting the relevance of our calculations to methods of separation of various nanotubes by chiralities and radii.

4.1 Dielectric linear response formalism

4.1.1 Definitions and identities

Any experiment conducted on a system involves observation and measurement of changes induced by some external probe. If the interaction between the probe and the system is weak, the response is linear: it is determined by the properties of the system in the absence of the perturbation, i.e. it is a ground state property. The system's response is closely linked to the correlations between positions and momenta of different particles, so the inter-particle interaction plays a central role. Since we want to study the effects of interactions in a system consisting of charged particles, we are interested in the density response to external electric fields. It is worth noting that density-functional theory is particularly well suited to such calculations - it is in principle an exact theory of the ground state of an interacting system, and density is the fundamental parameter. As such, calculations of zero-temperature *static* linear response quantities related to the ground state energy and density contain no uncontrolled approximations or possible misinterpretations (unlike calculations of excited-state and non-equilibrium properties).

In this section we review the standard expressions for the linear-response dielectric susceptibilities of periodic systems of different dimensionalities. We will consider only static response and drop the time/frequency dependence of all quantities, e.g. $\chi(\mathbf{q}, \omega = 0) = \chi(\mathbf{q})$. For completeness, we will briefly discuss the frequency-dependence in the end of this subsection and show that we are justified in assuming that all quantities of interest below are purely real in the static regime. Let us begin by considering an external perturbing potential $V_{ext}(\mathbf{r})$ that couples linearly to the

number density operator

$$\hat{n}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i), \quad (4.1)$$

so that the perturbed Hamiltonian is

$$\hat{H}_V = \hat{H} + \int V_{ext}(\mathbf{r}') \hat{n}(\mathbf{r}') d\mathbf{r}'. \quad (4.2)$$

The static density-density response function (susceptibility) $\chi(\mathbf{r}, \mathbf{r}')$ describes the response of the expectation value of the number density operator (induced density) at point \mathbf{r} to the perturbation:

$$\delta n(\mathbf{r}) = \int \chi(\mathbf{r}, \mathbf{r}') V_{ext}(\mathbf{r}') d\mathbf{r}', \quad (4.3)$$

which in Fourier space becomes

$$\delta n(\mathbf{q}) = \int \chi(\mathbf{q}, \mathbf{q}') V_{ext}(\mathbf{q}') d\mathbf{q}'. \quad (4.4)$$

In general, a perturbation with wave vector \mathbf{q}' will induce density modulations at all the wave vectors \mathbf{q} for which the response function $\chi(\mathbf{q}, \mathbf{q}')$ is non-zero. For systems with continuous translational invariance a great simplification takes place - the response function depends only on the difference $\mathbf{r} - \mathbf{r}'$ and so $\delta n(\mathbf{q}) = \chi(\mathbf{q}) V_{ext}(\mathbf{q})$. We can view $\chi(\mathbf{q})$ as a measure of stiffness of the system against a field that attempts to change the value of $n(\mathbf{q})$. If we denote the unperturbed ground state energy by $E(0)$, we have that the perturbed energy is given in the linear regime by the stiffness theorem [34]

$$E(\delta n) = E(0) - \frac{1}{2} \frac{\delta n^2}{\chi} \geq E(0) \quad (4.5)$$

that necessitates that $\chi \leq 0$.

Now because electrons are charged, the change in density $\delta n(\mathbf{r})$ creates an additional coulomb potential $V_{ind}(\mathbf{r})$ which is superimposed on the external field. So a

test charge sees a total resulting potential (called the screened potential)

$$V_{sc}(\mathbf{r}) = V_{ext}(\mathbf{r}) + V_{ind}(\mathbf{r}), \quad (4.6)$$

where

$$V_{ind}(\mathbf{r}) = \int d\mathbf{r}' \frac{\delta n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (4.7)$$

Using here (4.3) we obtain for the screened potential

$$V_{sc}(\mathbf{r}) = \int d\mathbf{r}' \epsilon^{-1}(\mathbf{r}, \mathbf{r}') V_{ext}(\mathbf{r}'), \quad (4.8)$$

where we define the inverse (in the matrix sense) dielectric function as

$$\epsilon^{-1}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') + \int d\mathbf{r}'' \frac{1}{|\mathbf{r} - \mathbf{r}''|} \chi(\mathbf{r}'', \mathbf{r}). \quad (4.9)$$

Again, in the simple case of homogeneous systems, we get the algebraic relation for the Fourier components

$$V_{sc}(q) = \frac{V_{ext}(q)}{\epsilon(q)}, \quad (4.10)$$

where

$$\frac{1}{\epsilon(q)} = 1 + v_q \chi(q), \quad (4.11)$$

with v_q being the Fourier transform of the coulomb interaction. In the inhomogeneous case of a periodic crystal we have $\chi(\mathbf{r} + L, \mathbf{r}' + L) = \chi(\mathbf{r}, \mathbf{r}')$ and in Fourier space

$$\delta n(\mathbf{q} + \mathbf{G}) = \sum_{\mathbf{G}'} \chi(\mathbf{q} + \mathbf{G}, \mathbf{q} + \mathbf{G}') V_{ext}(\mathbf{q} + \mathbf{G}'), \quad (4.12)$$

where \mathbf{G} is a vector in the reciprocal lattice and \mathbf{q} is restricted to the first Brillouin zone. Using a more compact notation $\chi_{\mathbf{G}, \mathbf{G}'}(\mathbf{q}) = \chi(\mathbf{q} + \mathbf{G}, \mathbf{q} + \mathbf{G}')$ we can rewrite (4.9) as

$$[\epsilon^{-1}]_{\mathbf{G}, \mathbf{G}'}(\mathbf{q}) = \delta_{\mathbf{G}, \mathbf{G}'} + v_{\mathbf{G}}(\mathbf{q}) \chi_{\mathbf{G}, \mathbf{G}'}(\mathbf{q}). \quad (4.13)$$

The inverse dielectric matrix has generally non-zero off-diagonal elements, due to

crystal inhomogeneity, which generate umklapp effects in the response called *local-field effects*. In particular, $\epsilon(0,0)$ is in general different from $1/\epsilon^{-1}(0,0)$. It can be shown that in crystals with high symmetry (e.g. 3D cubic, 2D hexagonal) the local-field effects simplify in the long-wavelength limit ($q \rightarrow 0$) [35], and the macroscopic dielectric constant ϵ_m is the inverse of the $\mathbf{G} = \mathbf{G}' = 0$ component of the inverse dielectric matrix:

$$\frac{1}{\epsilon_m} = [\epsilon^{-1}]_{0,0}(0). \quad (4.14)$$

For simplicity, we will be interested in the response of our systems to a monochromatic sinusoidal potential $V_{ext}(\mathbf{q} + \mathbf{G}) = V_0 \delta(\mathbf{q} - \mathbf{q}_0)$, i.e. we will consider only the matrix elements with $\mathbf{G} = \mathbf{G}' = 0$. We will often drop these subscripts to simplify notation, but it is important to remember, however, that within DFPT calculations all details of the local-field effects enter into the response matrices, whose particular elements we ultimately use.

It is also convenient to define a *screened* density-density response function $\tilde{\chi}(\mathbf{r}, \mathbf{r}')$ that gives the response of the density $\delta n(\mathbf{r})$ to the screened potential $V_{sc}(\mathbf{r}')$

$$\delta n(\mathbf{r}) = \int \tilde{\chi}(\mathbf{r}, \mathbf{r}') V_{sc}(\mathbf{r}') d\mathbf{r}'. \quad (4.15)$$

It is easy to see that the relationship between the screened response function and the full one is

$$\tilde{\chi}^{-1}(\mathbf{r}, \mathbf{r}') = \chi^{-1}(\mathbf{r}, \mathbf{r}') + \frac{1}{|\mathbf{r} - \mathbf{r}'|} \quad (4.16)$$

$$\tilde{\chi}^{-1}(\mathbf{q}) = \chi^{-1}(\mathbf{q}) + v_q. \quad (4.17)$$

Writing now the screened potential as

$$V_{sc}(\mathbf{r}) = V_{ext}(\mathbf{r}) + \int d\mathbf{r}' \int d\mathbf{r}'' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \tilde{\chi}(\mathbf{r}', \mathbf{r}'') V_{sc}(\mathbf{r}'') \quad (4.18)$$

and comparing with (4.8) we get a direct relationship between the dielectric function

and the screened response function:

$$\epsilon(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') - \int d^3\mathbf{r}'' \frac{1}{|\mathbf{r} - \mathbf{r}''|} \tilde{\chi}(\mathbf{r}'', \mathbf{r}) \quad (4.19)$$

$$\epsilon(\mathbf{q}) = 1 - v_q \tilde{\chi}(\mathbf{q}). \quad (4.20)$$

If we turn off coulomb interactions, the two response functions χ and $\tilde{\chi}$ are the same and equal to the independent particle susceptibility given by

$$\chi_{0,0}^0(\mathbf{q}) = \sum_{\mathbf{k}, n_1, n_2} \frac{|\langle \mathbf{k}, n_1 | e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}} | \mathbf{k} + \mathbf{q}, n_2 \rangle|^2}{E_{n_1}(\mathbf{k}) - E_{n_2}(\mathbf{k} + \mathbf{q})} [f_{n_1}(\mathbf{k}) - f_{n_2}^0(\mathbf{k} + \mathbf{q})], \quad (4.21)$$

where \mathbf{q} and \mathbf{k} are restricted to the first Brillouin zone, and $f_n(\mathbf{k})$ are the temperature-dependent Fermi occupation functions of bands n_1, n_2 [35]. This expression reduces to the familiar Lindhard function in the homogeneous case.

The presence of long-range coulomb interactions causes certain physical quantities, (e.g. bare compressibility) that are finite in neutral liquids, to diverge in the long wave-length limit, due to the infinite self-energy. For this reason the linear density response of charged liquids is usually discussed within the *jellium* model, in which a system of negatively charged electrons is embedded in a uniform background of positive charge density so as to impose overall charge neutrality. For instance, in the Hartree approximation the coulomb self-energy of the positive background plus that of the electron gas exactly cancels the interaction energy of the electrons with the positive background. The screened response function $\tilde{\chi}$ is essentially a construct designed to facilitate dealing with these divergences. In effect, $\tilde{\chi}$ corresponds to the density response function for a fictitious neutral system with the law of interaction given by the (short-range) screened coulomb interaction.

We now mention an important result of the linear response theory, called the compressibility sum rule. Taking the long wavelength limit of (4.16)

$$\lim_{q \rightarrow 0} [\chi^{-1}(q) + v_q] = \lim_{q \rightarrow 0} \tilde{\chi}^{-1}(q) = -\frac{1}{n^2 \tilde{K}} \quad (4.22)$$

This sum rule tells us that the full density stiffness

$$\lim_{q \rightarrow 0} \chi^{-1}(q) = -\frac{\partial^2 E}{\partial n^2} = -\frac{1}{n^2 K} \quad (4.23)$$

consists of two parts: (i) an electrostatic part v_q which diverges for $q \rightarrow 0$ due to the infinite coulomb energy cost of the uniform change of the electronic charge density, with the positive background remaining fixed; and (ii) a finite additional contribution known as screened or “proper stiffness” arising from the change in kinetic and exchange-correlation energies in the overall neutral system. In other words, the screened compressibility \tilde{K} of the system is related to the long-wavelength limit of the screened response function $\tilde{\chi}(\mathbf{q})$.

We now connect the above discussion to the general dynamic case by re-introducing time-dependence of the response. The frequency-dependent density response function of the free electron gas is given by the well-known Lindhard function

$$\chi^0(\mathbf{q}, \omega) = \sum_{\mathbf{k}, n_1, n_2} \frac{f_{n_1}(\mathbf{k}) - f_{n_2}(\mathbf{k} + \mathbf{q})}{E_{n_1}(\mathbf{k}) - E_{n_2}(\mathbf{k} + \mathbf{q}) + \omega + i\eta}, \quad (4.24)$$

where η is an infinitesimal positive number that regulates the divergences at resonant frequencies of the system and the limit $\eta \rightarrow 0^+$ is assumed. In general the response function is complex, with the imaginary part linked to the dissipation of the energy supplied by the perturbation. We can separate the real and imaginary components of (4.24) using the formula

$$\lim_{\eta \rightarrow 0^+} \frac{1}{\omega - y + i\eta} = \mathcal{P} \frac{1}{\omega - y} - i\pi \delta(\omega - y). \quad (4.25)$$

We see that at $\omega = 0$ only the principle part of (4.24) survives because we have $(f_1 - f_2) \delta(E_1 - E_2) = 0$, i.e. for a given energy eigenvalue all states are either occupied or not. This is an implicit assumption of *equilibrium*, or net zero current. Indeed, a situation with a net current corresponds to only some states of a given energy being occupied, which results in an overall non-zero momentum. This is only possible in a system with partially occupied bands at the Fermi energy - a metal.

In an insulator, no current can flow because all bands are completely filled. In the classical harmonic resonance picture, an insulator has the lowest resonance frequency different from zero, while a conductor has free electrons whose response is singular at zero frequency [36]. When an electric field is applied to a conductor, the effective electric field that drives the current is the gradient of the external potential V plus the gradient of the local chemical potential associated with the inhomogeneous induced density $\delta n(\mathbf{q}, \omega) = \chi(\mathbf{q}, \omega) V(\mathbf{q}, \omega)$. The chemical potential is given by

$$\mu(\mathbf{q}, \omega) = -\frac{\delta n(\mathbf{q}, \omega)}{\chi(\mathbf{q}, \omega)}. \quad (4.26)$$

which is just the potential that would produce the density δn if the system is allowed to reach equilibrium. The effective electric field is then

$$\mathbf{E}_{eff}(\mathbf{q}, \omega) = -iq(\mu(\mathbf{q}, \omega) + V(\mathbf{q}, \omega)). \quad (4.27)$$

Note that the field vanishes for $\omega = 0$ because $\delta n(\mathbf{q}, 0) = \chi(\mathbf{q}, 0) V(\mathbf{q}, 0)$ since no current flows at equilibrium. At non-zero frequency the current density is

$$\mathbf{j}(\mathbf{q}, \omega) = \sigma(\mathbf{q}, \omega) \mathbf{E}_{eff}(\mathbf{q}, \omega) \quad (4.28)$$

which defines the conductivity σ , a quantity that contains information about the dissipation of the energy through electron-phonon scattering or impurities. With the help of equations (4.26)-(4.28) and the continuity equation $\delta n(\mathbf{q}, \omega) = \mathbf{q} \cdot \mathbf{j}(\mathbf{q}, \omega)/\omega$ we arrive at the relation

$$\frac{1}{\chi(\mathbf{q}, \omega)} = \frac{1}{\chi(\mathbf{q}, 0)} + i \frac{\omega}{q^2 \sigma(\mathbf{q}, \omega)}. \quad (4.29)$$

For normal metals the conductivity is finite and real at zero frequency, which implies that the imaginary part of the response vanishes at $\omega = 0$, once equilibrium is assumed. Thus, for insulating systems, finite metallic systems and extended metallic systems at equilibrium the zero-frequency density response is purely real, and in

this work we will be dealing only with such systems. However, if a steady current is allowed to flow, an imaginary component of the static dielectric constant will be non-zero, corresponding to some sort of dissipation. Frequency-dependent response and non-equilibrium transport phenomena require more elaborate methods (such as time-dependent DFT and non-equilibrium Green's functions) and lie outside the scope of this thesis.

4.1.2 Random-phase approximation

So far we have been discussing an exact formalism, but in order to compute the response of actual systems, it is useful to make approximations that would allow us to input some details of a model and study the result analytically. The simplest analytical approximation one can make to describe the response of a charged system is termed the random-phase approximation (RPA), also known as the self-consistent field approximation. It consists of replacing the screened response function $\tilde{\chi}$ by that of the independent electron gas χ^0 so that (4.19) becomes

$$\epsilon^{RPA}(q) = 1 - v_q \chi^0(q) \quad (4.30)$$

and the full response function is given within RPA by

$$\chi^{RPA}(q) = \frac{\chi^0(q)}{1 - v_q \chi^0(q)}. \quad (4.31)$$

The RPA is effectively a single-particle approximation of the full many-electron problem using the self-consistent potential which incorporates many-body effects in an average sense. It is a very convenient and powerful analysis tool for studying scaling of various quantities sensitive to screening. In particular, it is generally very accurate in describing the long-wavelength dielectric response in the high density limit. However, it neglects the short-range exchange and correlation effects altogether, and the response at finite q is less reliable.

We now consider the zero-temperature static dielectric response at small q of

simple metals and insulators within RPA to get a sense of scaling for these quantities before we present the results of full ab-initio calculations. In this limit the independent particle susceptibility of a metal is a non-zero constant, which can be seen from the following argument. Considering only one partially filled band in (4.21) we have (taking $e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}} \sim 1 + O(q)$)

$$\chi^0(\mathbf{q}) = \sum_{\mathbf{k}} \frac{f^0(\mathbf{k}) - f^0(\mathbf{k} + \mathbf{q})}{E(\mathbf{k}) - E(\mathbf{k} + \mathbf{q})} \quad (4.32)$$

$$\rightarrow \int \frac{\mathbf{q} \cdot \frac{\partial f^0}{\partial E} \nabla_{\mathbf{k}} E(\mathbf{k})}{\mathbf{q} \cdot \nabla_{\mathbf{k}} E(\mathbf{k})} d\mathbf{k} \quad (4.33)$$

$$= \int \frac{\partial f^0}{\partial E} \nu(E) dE = -\nu(0), \quad (4.34)$$

where $\nu(E)$ is the density of states and the Fermi energy is taken to be zero. This result is clearly independent of the dimensionality of the system.

To find an approximation to the equation (4.21) in the case of insulators, we can use the sum rule for oscillator strengths [37]

$$\sum_n (E_n - E_s) |\langle n | e^{i\mathbf{q}\cdot\hat{\mathbf{r}}} | s \rangle|^2 = \frac{q^2}{2}. \quad (4.35)$$

Taking the smallest value of the energy difference (the band gap) to be the most important term in the summation,

$$E_{n_1}(\mathbf{k}) - E_{n_2}(\mathbf{k} + \mathbf{q}) \approx \Delta_{gap} \quad (4.36)$$

we have then

$$|\langle \mathbf{k}, n_1 | e^{i\mathbf{q}\cdot\hat{\mathbf{r}}} | \mathbf{k} + \mathbf{q}, n_2 \rangle|^2 \approx \frac{1}{\Delta_g} \frac{q^2}{2}. \quad (4.37)$$

Only matrix elements between valence and conduction states contribute to the sum in (4.21), because the occupation factor $f_{n_2}^0(\mathbf{k} + \mathbf{q}) - f_{n_1}^0(\mathbf{k}) \neq 0$ only if $n_1 \neq n_2$. Putting (4.37) into (4.21) and counting the contributions from each of N electrons in

the filled bands in the unit cell of volume Ω , we get

$$\chi^0(q) \approx -\frac{N}{\Omega\Delta_g^2}q^2 \quad (4.38)$$

which is valid for an insulator of any dimensionality. We now use the two expressions (4.32) and (4.38) to study the $q \rightarrow 0$ limit of the dielectric response of both types of extended systems in all three dimensionalities. We will need explicit expressions for the Fourier transforms of the Coulomb potentials. They are

$$v_q^{3D} = \frac{4\pi}{q^2} \quad (4.39)$$

$$v_q^{2D} = \frac{2\pi}{q} \quad (4.40)$$

$$v_q^{1D} = -e^{q^2 R^2} \text{Ei}(-q^2 R^2), \quad (4.41)$$

where Ei is the exponential integral function. The one-dimensional case is more delicate, as one also needs a short-distance cutoff R to avoid the non-integrable divergence of the interaction at the origin. This can be done by modeling a 1D system as a 3D cylinder of very small radius R , so that the transverse motion of electrons is essentially frozen in the lowest energy localized state. For $qR \ll 1$ we have $v_q^{1D} \approx -2 \ln(qR)$, and we will use this approximation from now on.

In 3D bulk metals $\chi_0(q \rightarrow 0)$ is nonzero according to (4.32), so the dielectric constant $\epsilon_{RPA}^{3D} = 1 + \frac{4\pi}{q}\nu(0)$ tends to infinity, resulting in the familiar complete screening of the electric field. The same situation is observed in perfect metals in both the 2D and 1D cases, resulting in a diverging Coulomb interaction and consequently an infinite dielectric constant $\epsilon_{RPA}^{1D,2D} \rightarrow +\infty$. In 3D insulators the q -dependence in (4.30) cancels out and the dielectric constant has a scale-invariant nonzero gap-dependent value

$$\epsilon_{RPA}^{3D}(q) \approx 1 + \frac{4\pi n}{\Delta_g^2}, \quad (4.42)$$

where n is the average electronic density. In low-dimensional insulators, on the other hand, the q -dependent factors do not cancel, and in the small- q limit the $\chi^0(q)$ factor

	1D: $v_q = -2 \ln qR$	2D: $v_q = 2\pi/q$	3D: $v_q = 4\pi/q^2$
metal $\tilde{\chi} \sim -\nu$	$\epsilon \sim -2\nu \ln qR$ $\chi \sim 1/\ln qR$	$\epsilon \sim 2\pi\nu/q$ $\chi \sim -q/2\pi$	$\epsilon \sim 4\pi\nu/q^2$ $\chi \sim -q^2/4\pi$
insulator $\tilde{\chi} \sim -Cq^2$	$\epsilon \rightarrow 1$ $\chi \sim -Cq^2$	$\epsilon \rightarrow 1$ $\chi \sim -Cq^2$	$\epsilon \rightarrow \epsilon_\infty$ $\chi \sim -\frac{C}{\epsilon_\infty}q^2$

Table 4.1: Summary of $q \rightarrow 0$ scaling of dielectric response quantities of metals and insulators of different dimensionalities, as derived within RPA. ν is the density of states at the Fermi energy.

vanishes quicker than v_q so we have $\epsilon_{RPA}^{1D,2D} \rightarrow 1$. In other words, 2D and 1D semiconducting systems do not screen electric fields at all, because their reduced dimensions do not allow a polarization field to develop. These results are summarized in Table 4.1.

4.1.3 Polarizabilities in low dimensions

It is useful to understand the electrostatics of low dimensional objects considered as embedded in a 3D periodic bulk system. This is a practical issue that arises when calculating the dielectric response of a low-dimensional object in periodic boundary conditions, where we are actually calculating the dielectric constant of the 3D bulk system consisting of a periodic array of these objects separated by vacuum. Let us take, for example, a cubic unit cell of size L containing an atom (a 0D object) of polarizability α^{0D} . The dielectric constant of this material is given by

$$\epsilon = \frac{\mathbf{D}}{\mathbf{E}} = 1 + 4\pi \frac{\mathbf{P}}{\mathbf{E}} \quad (4.43)$$

where $\mathbf{P} = \mathbf{E}\alpha^{0D}/L^3$ is the polarization density. For a solid made up of a square array of tubes or an array of sheets the corresponding relationships are

$$\epsilon = 1 + \frac{4\pi}{L^2}\alpha^{1D} \quad \epsilon = 1 + \frac{4\pi}{L}\alpha^{2D} \quad (4.44)$$

where α^{1D} is the polarizability per unit length and α^{2D} is the polarizability per unit area. The dielectric constant of these systems tends to 1 when the cell size is taken to

infinity, so bulk ϵ is not a useful quantity to describe the response of a low-D system, and one needs to extract the intrinsic values of α 's. Let us now make the connection with the response function formalism. For simplicity we specialize to the case of a 2D sheet and denote quantities computed for periodic (bulk) systems by superscript b :

$$\epsilon^b = \lim_{q \rightarrow 0} 1 - \frac{4\pi}{q^2} \tilde{\chi}^b(q) = 1 + \frac{4\pi}{L} \alpha_{\parallel}^{2D}. \quad (4.45)$$

At the same time we have

$$\tilde{\chi}^b(q) = \frac{\delta n^b(q)}{V_{sc}(q)} = \frac{\delta n^{2D}(q)/L}{V_{sc}(q)} = \frac{\tilde{\chi}_{\parallel}^{2D}(q)}{L} \quad (4.46)$$

so that we get the identity

$$\alpha_{\parallel}^{1D,2D} = -\lim_{q \rightarrow 0} \frac{\tilde{\chi}_{\parallel}(q)}{q^2}, \quad (4.47)$$

which relates the 2D in-plane response function to the polarizability per unit area, while in 1D this expression related the longitudinal response function to the longitudinal polarizability per unit length. Clearly for insulators, where $\tilde{\chi}_{\parallel} \sim -q^2$, the polarizability α_{\parallel} is a finite quantity, while for metals (and semi-metals) it is divergent. It should be noted that the out-of-plane (transverse) polarizability α_{\perp} is not accessible within the low-dimensional response function formalism because it is a 3D quantity describing the response to transverse fields of the low-dimensional system embedded in 3D space.

4.2 2D: Graphene and boron-nitride

In this section we study the in-plane static response of two-dimensional sheets of graphene and boron-nitride [38]. It is useful in studying the effects of charged impurities and in analyzing the electrostatic environment of functionalizations. Electron-phonon coupling is also very sensitive to the way the low-frequency lattice perturbations are screened by the electrons. As we shall see later in Sec. 4.3 the in-plane response of a sheet also helps us understand the transverse response of single- and

multi-wall nanotubes. The in-plane response function of graphene at zero doping was first considered by Gonzalez et al. [39] and derived for independent particles, starting with the homogeneous model with a linear band dispersion. Motivated by the expected strong influence of doping on the nature of interactions, Ando [40] recently generalized this result to non-zero doping for $q < 2k_F$ at $T=0$:

$$\chi^0(q) = -\frac{q}{4v_F} + \frac{2\varepsilon_F}{\pi v_F^2} \quad (4.48)$$

where the Fermi energy is related to the doping density n by

$$n = \frac{\varepsilon_F^2}{\pi v_F^2}. \quad (4.49)$$

Note that at zero doping ($\varepsilon_F = 0$) the free response function is *linear* in q . Remembering the 2D Coulomb factor (4.39), we can use the RPA expression (4.30) for the in-plane dielectric constant at zero doping

$$\epsilon \approx 1 + \frac{2\pi}{q} \frac{q}{4v_F}. \quad (4.50)$$

Using the DFT result of $v_F = 0.38 \text{ bohr}\cdot\text{Ha} = 8.31 \times 10^5 \text{ m/s}$ we get $\epsilon = 5.13$. Because of the cancelation of the q -dependent factors, we obtain that the in-plane dielectric constant $\epsilon(q = 0)$ has a non-trivial value, resembling the situation in a 3D bulk dielectric. In other words, the in-plane dielectric response is q -invariant. It follows that graphene cannot be classified as either a metal or a insulator in light of the discussion in Sec. 4.1.2 and Table 4.1, but rather it is intermediate – a *zero-gap semiconductor*. This anomalous screening property of graphene is due to the unique combination of its two-dimensional nature with linear electronic spectrum and resulting vanishing density of states at the Fermi K-points. This corresponds to the unusual linear screening noted previously in graphite- intercalated compounds [41]. On the other hand, any finite amount of doping or a finite temperature will create free carriers and we should expect to see the signature of metallic response. We should mention also that the temperature dependence of the static screening is

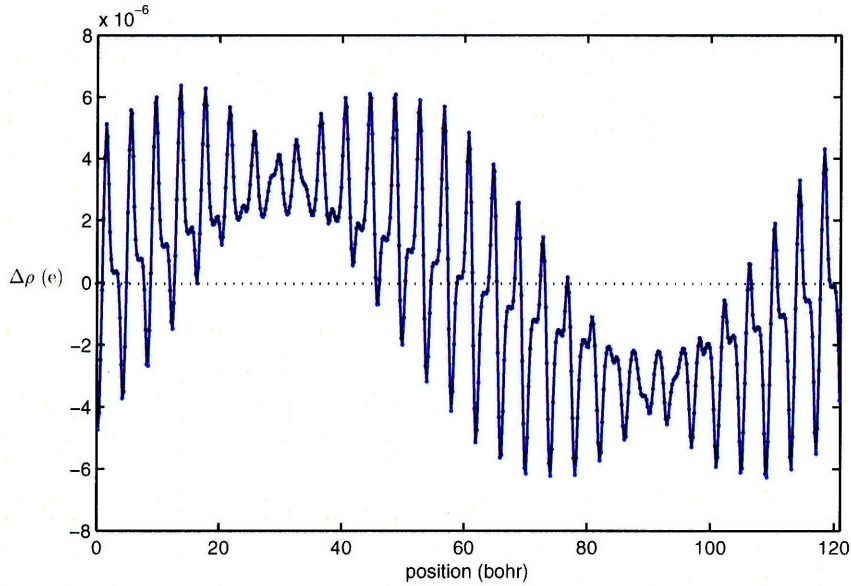


Figure 4-1: Charge density oscillations in a graphene sheet induced by a perturbing sinusoidal potential with a wavelength equal to 60 unit cells. Rapid oscillations are caused by charge transfer between the two carbon sub-lattices.

predicted within RPA to be also qualitatively different from that in conventional metals [40].

We will now investigate the $q \rightarrow 0$ scaling of the in-plane response of graphene and boron-nitride using density functional theory. In our calculations we apply an external sinusoidal perturbing potential $V_{ext}(q)$ of wavelength $2\pi/q$ and calculate the Fourier component of the resulting self-consistent charge density response δn_q . By taking the ratio of the Fourier components we obtain the full response function matrix element $\mathbf{G} = \mathbf{G}' = 0$

$$\chi_{0,0}(q) = \frac{\delta n_q}{V_{ext}(q)}. \quad (4.51)$$

The calculation is done in two different ways. First, we set up a long orthorhombic super-cell containing up to 60 repetitions of the basic unit cell, and add a small-amplitude sinusoidal potential with a period spanning the length of the super-cell. The induced charge density response is calculated using the self-consistent iterative DFT procedure as implemented in the Quantum-ESPRESSO package [42]. Alternatively, we have implemented the linear dielectric response within the framework of DFPT, which allows much more efficient calculations of response matrix elements at

arbitrary q without the need for super-cells. Both approaches yield identical results as shown in Fig. 4-2. We set the graphene lattice parameter at $|\mathbf{T}_1| = 4.654$ bohr and for boron nitride at 4.73 bohr. We use the Perdew-Burke-Ernzerhof generalized gradient approximation with ultra-soft (in the super-cell method) and norm-conserving (in the DFPT calculation) pseudo-potentials in a plane-wave basis with a cutoff energy of 25 and 40 Ry, respectively. These parameters have been shown to accurately predict the crystalline structures and the physical properties of diamond and graphite [43]. The Brillouin zone was sampled with a k -point grid of at least 64x64x1 points. Periodic boundary conditions are imposed in all three dimensions, and special measures are taken to remove the periodic image interactions that renormalize the density response, especially at small q . In particular, we use the correction method described in Sec. 6.3 with the inter-plane distance of about 45 bohr being sufficient for convergence at the smallest value of q .

Fig. 4-2 shows the values of the matrix element $\frac{2\pi}{q}\chi_{0,0}(q)$ of the full response function as a function of q for both 2D sheets. It is essential to note the drastically different small- q behavior of the susceptibilities of the two materials. The curve for boron nitride is seen to approach a linear trend near the origin, confirming the expected dependence $\chi(q) \sim q^2$, which characterizes boron nitride as a 2D insulator. The same quantity for undoped graphene, on the other hand, approaches a nontrivial constant ≈ 0.8 as $q \rightarrow 0$, which implies $\chi(q) \sim q$ to leading order, confirming the RPA result. With the help of expressions (4.13) and (4.14), we see that $\epsilon_m \approx 1/(1-0.8) = 5$ in excellent agreement with the RPA estimate above. Finally, for doped graphene the curve is seen to approach 1 for small q , i.e. $\chi_{0,0} \rightarrow -\frac{q}{2\pi}$, and using (4.30) this results in an infinite $\epsilon(q = 0)$, as expected of a truly metallic 2D system. The $q \rightarrow 0$ in-plane response for graphene is thus very sensitive to doping with either electrons or holes. This can be very relevant for the screening of impurities and the strength of electron-phonon coupling.

We note that one should be careful in interpreting the calculated results plotted in Figs. 4-2 and 4-3, as they show the variation of just a single matrix element of the full response function. For instance, these values are not sufficient to obtain $\tilde{\chi}(q)$ and

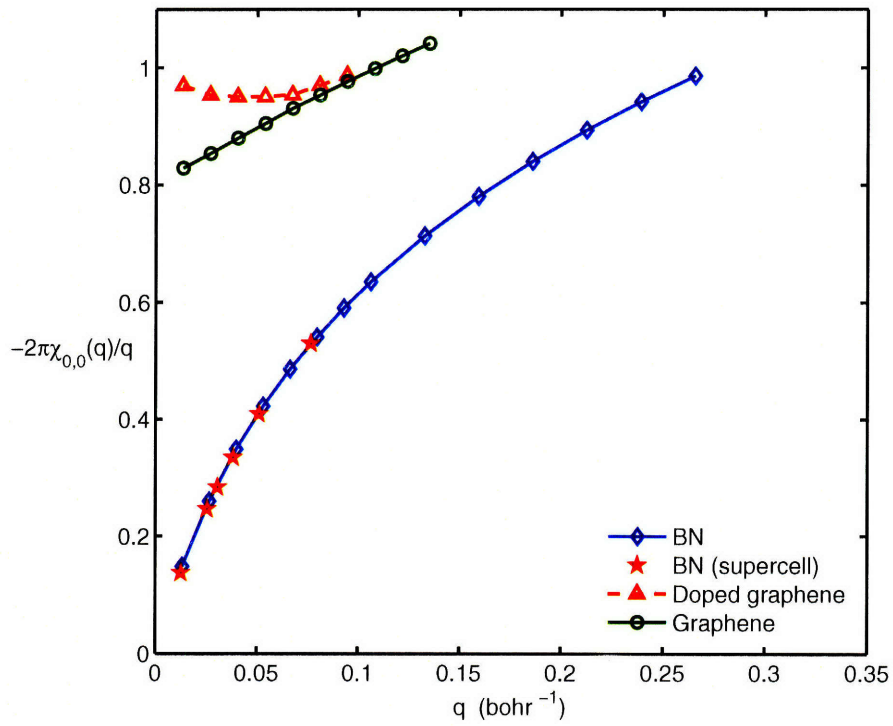


Figure 4-2: Scaling of the self-consistent response matrix element $\chi_{0,0}(q)$ of a graphene and a BN sheet as a function of the wavelength of the perturbing sinusoidal potential. The doping level for graphene is 0.005 electrons per unit cell ($\varepsilon_F=0.01$ Ha). At lower doping levels the curve exhibits a sharper variation at smaller q .

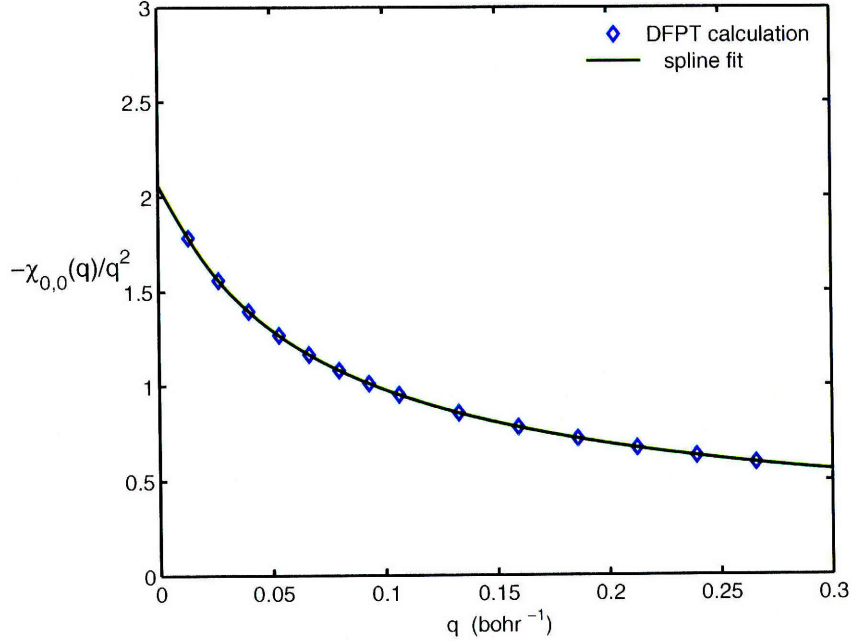


Figure 4-3: Scaling of the full response matrix element $\chi_{0,0}(q)$ of the BN sheet. Extrapolation to $q = 0$ is used to compute the in-plane polarizability.

$\epsilon(q)$ at finite q because the off-diagonal matrix elements are nonzero due to local field effects. However, due to (4.14) the $q \rightarrow 0$ limit is well-defined and we are allowed to go between these quantities as in a homogeneous system. For example, we can calculate the in-plane polarizability per unit area of the BN sheet in two ways: directly from the dielectric tensor we get $\alpha_{\parallel}^{2D} = 2.04$ bohr and $\alpha_{\perp}^{2D} = 0.34$ bohr. Alternatively from the small- q limit of the response function using (4.47)

$$\alpha_{\parallel}^{2D} = -\lim_{q \rightarrow 0} \frac{\tilde{\chi}_{0,0}(q)}{q^2} = -\lim_{q \rightarrow 0} \frac{\chi_{0,0}(q)}{q^2} \quad (4.52)$$

we get by extrapolation $\alpha_{\parallel}^{2D} = 2.06$ bohr (see Fig. 4-3). In order to find $\tilde{\chi}_{0,0}(q)$ at finite q one needs in general to invert the entire full response matrix $\chi(q)$, the calculation of which is more challenging within DFPT.

We investigate now the long-wavelength limit of the response of the graphene sheet at finite doping. Recalling the compressibility sum rule (4.22) and applying the RPA

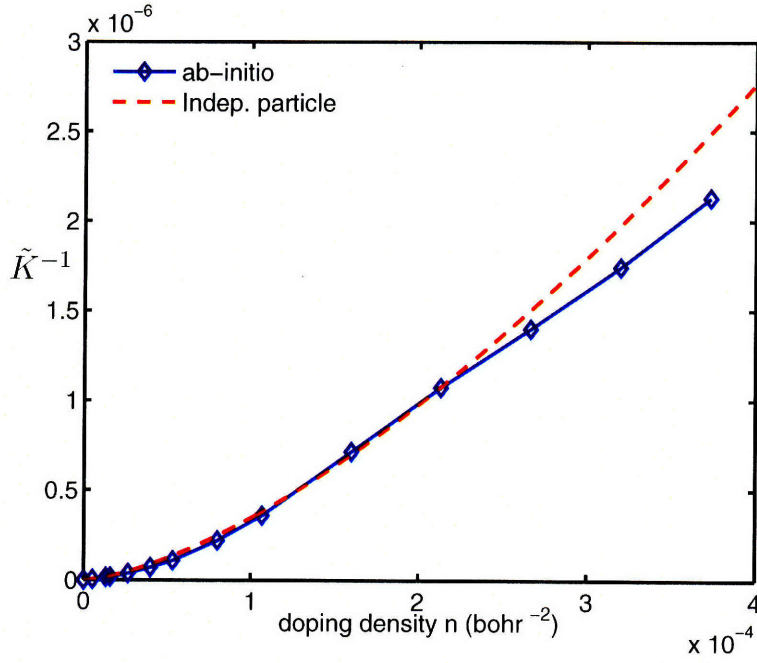


Figure 4-4: Inverse screened compressibility of a doped graphene sheet versus the doping density. The dashed curve is computed from the independent-particle linear band model via (4.53).

estimate of the response function (4.49), we can write the screened compressibility as

$$-\frac{1}{n^2 \tilde{K}} = \lim_{q \rightarrow 0} \tilde{\chi}^{-1}(q) \approx \frac{\pi v_F^2}{2\varepsilon_F} \quad (4.53)$$

The free inverse compressibility is seen to vanish at zero doping density, which follows directly from the vanishing density of states, i.e. the Fermi energy as a function of doping density has an infinite slope at zero doping. We calculate the screened compressibility in DFT by taking the second derivative of the total energy with respect to the average charge density and appropriately removing the electrostatic terms using methods described in Chapter 6. Alternatively, we can proceed to calculate the Fermi energy as a function of the doping level and take the derivative with respect to the doping density. The two calculations yield identical results, as guaranteed by Janak's theorem [26], and the results are plotted in Fig. 4-4. The calculated compressibility agrees well with the free linear-band model result (4.53) at low doping densities, yielding credibility to the simple model used in RPA calculations. Once again, this

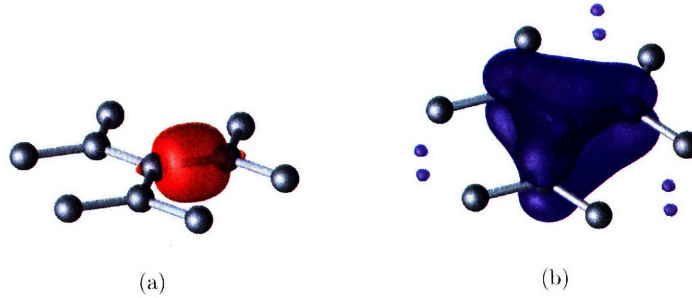


Figure 4-5: Maximally localized Wannier functions of the occupied states in graphene. (a) σ -bonding MLWF; (b) π -bonding MLWF. Courtesy of Young-Su Lee.

follows from the fact that π -electrons dominate the response at long wavelengths. At higher doping levels we see a discrepancy that results from the nonlinearities of the DFT band dispersions and exchange-correlation contributions.

4.2.1 Wannier function analysis

It is usually tacitly assumed that the linear-band description derived from the π -state tight-binding model of graphene gives the full picture of the electronic properties. It should be remembered, however, that the σ -bond electrons also participate in dielectric response. They form localized states, unlike the π -electrons, and their response has insulating character. In simple models the effect of the σ -electrons is usually incorporated into a uniform background dielectric constant, and π -electrons are treated within the modified electrostatic environment. Such an approach may be justified for 3D bulk graphite, but not for graphene (or nanotubes), because as we have seen, dielectric response in low-dimensional insulators does not give rise to the scale-invariant screening, which in 3D systems is characterized by a non-trivial dielectric constant. In our first-principles calculations we treat all valence electrons on an equal footing and we can analyze their relative contributions. In order to separate the contributions of π - and σ -electrons, we transform the KS wavefunctions from the plane-wave basis to the basis of maximally localized Wannier functions (MLWF) [44, 45] using the WANNIER90 code [46]. The chemical nature of the states in this basis is readily identifiable, as shown on Fig. 4-5. It can be shown that for the

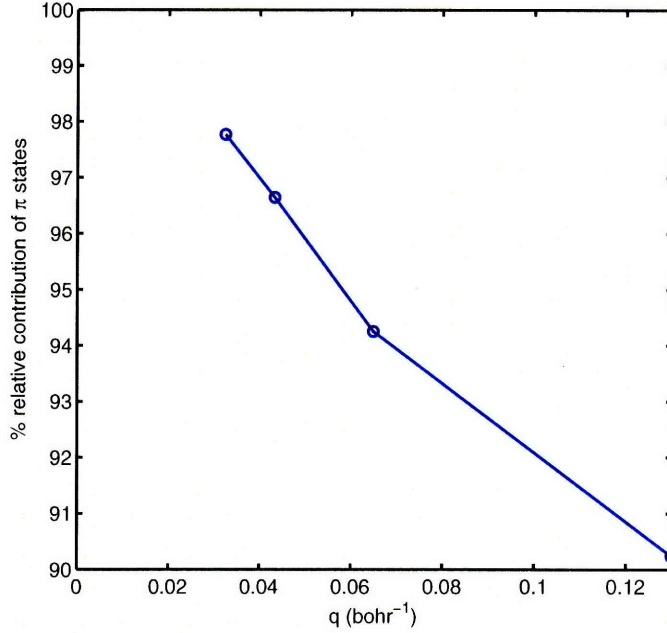


Figure 4-6: Percent relative contribution of π -electrons to the in-plane response function of the undoped graphene sheet.

purposes of understanding dielectric response, the charge density can be treated as localized into point charges at the centers of the Wannier functions [47]. In the supercell calculation the sinusoidal potential causes the centers of these Wannier functions to shift, and the Fourier component of the real-space shift profile just gives $\delta n(q)$. This method provides an alternative way to calculate the full response function matrix element, and in addition allows us to decompose the total density response into contributions from the two orbital types. As displayed on Fig. 4-6, the contribution of π -electrons to the response function is greater than 90% at wavelengths longer than 6 unit cells of graphene. This is what we expect because the insulating response contribution from the σ -states behaves like $\chi_{\sigma}(q) \sim q^2$ while for the π -electrons $\chi_{\pi}(q) \sim q$, so the latter will dominate at smaller q . We conclude, therefore, that the tight-binding picture that includes only π -electrons is a good approximation to the dielectric response at long wavelengths. Needless to say that in moderately doped graphene the free carriers which dominate the response at long wavelengths are π -electrons as well.

4.3 1D: Nanotubes

In this section we address in detail the longitudinal and transverse dielectric response of carbon and boron-nitride nanotubes - both single-wall and multi-wall [48]. All calculations are performed using Quantum-ESPRESSO [42] with the Perdew-Burke-Ernzerhof approximation and ultra-soft pseudo-potentials in a plane-wave basis. A tetragonal unit cell is set up with periodic-boundary conditions in all three dimensions. We use a k -point grid of at least $50 \times 1 \times 1$ points to sample the Brillouin zone, Methfessel-Paxton smearing of 0.02 Ry and ultrasoft pseudopotentials for C, B and N with a wavefunction cutoff energy of 25 Ry and 200 Ry for the density. Atomic configurations are generated using an interatomic distance of 1.42Å, obtained from careful relaxation studies [43]. Longitudinal and transverse polarizabilities are calculated using density-functional perturbation theory (DFPT) [31] and finite-field or electric-enthalpy approaches [49], also implemented in our Quantum-ESPRESSO code.

4.3.1 Longitudinal polarizabilities

Since we use periodic-boundary conditions, we effectively simulate a three-dimensional bulk material consisting of a square array of infinitely long parallel nanotubes, i.e. we only have access to the dielectric constant of this fictitious periodic system. What we are interested in is the longitudinal dielectric response of an isolated nanotube which is characterized by polarizability per unit length α_{\parallel} and is related to the separation-dependent bulk dielectric constant ϵ_{\parallel} via

$$\epsilon_{\parallel}(\mathbf{q} = 0) = 1 + \frac{4\pi}{L^2} \alpha_{\parallel} \quad (4.54)$$

where L is the transverse size of the unit cell. In this case there are no periodic image interaction effects between tubes on the lattice because the depolarizing field along the axis of an infinitely long object vanishes. We do the calculation in two different ways: first, using DFPT we can calculate the full dielectric tensor and extract α_{\parallel} from the longitudinal component using (4.54); second, we apply a small finite electric field

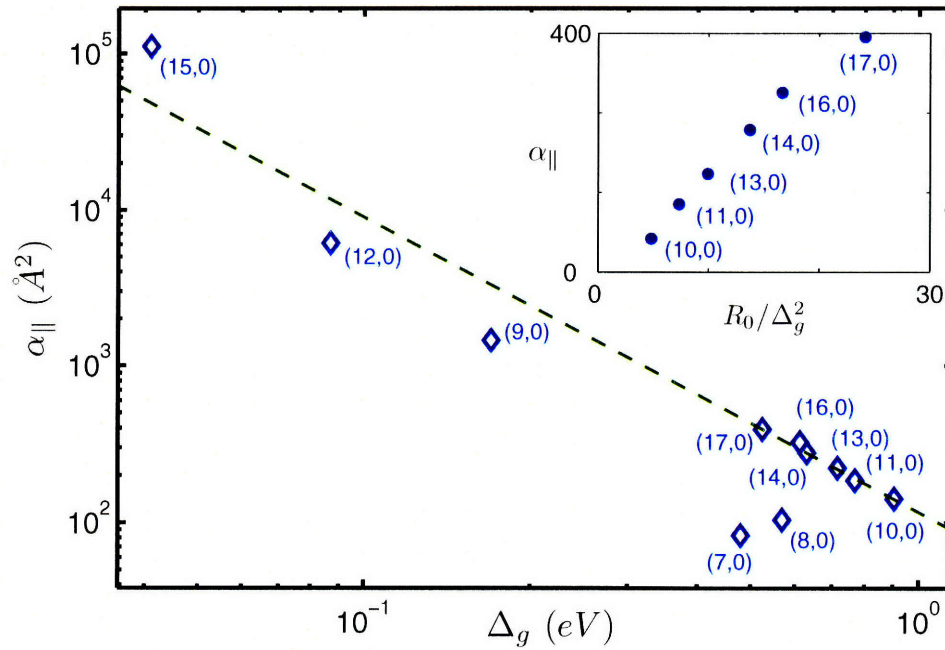


Figure 4-7: Log scale plot of α_{\parallel} of zigzag nanotubes as a function of band gap. The dashed line has slope -2. The inset shows the values for large-gap SWNTs as a function of R_0/Δ_g^2 .

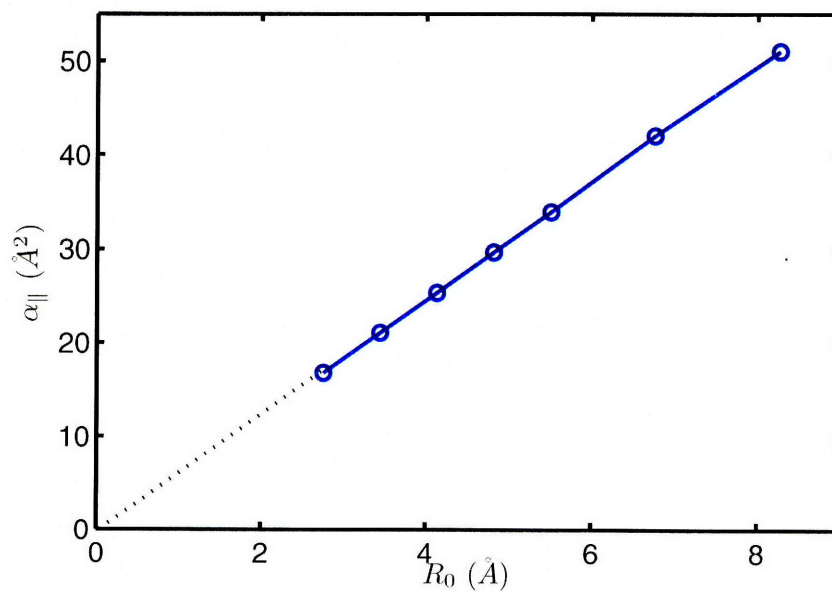


Figure 4-8: Longitudinal polarizability of boron-nitride nanotubes as a function of radius.

to the system and calculate the induced polarization along the axis using the recently developed Berry phase approach of electric enthalpy minimization [49], which allows calculations with finite electric fields in periodic boundary conditions. Both methods yielded identical results and most of the calculations below were done with the more efficient method of DFPT.

From linear-response theory outlined in Sec. 4.1.2 we expect the static dielectric constant to depend on the band gap as $\epsilon(q) \approx 1 + 4\pi n/\Delta_g^2$ which suggests via (4.54) that

$$\alpha_{\parallel} = \frac{N}{\Delta_g^2} \propto \frac{R}{\Delta_g^2}. \quad (4.55)$$

Our calculations confirm this behavior in zigzag carbon nanotubes, as shown in Fig. 4-8. As expected, (9,0), (12,0) and (15,0) zigzag carbon nanotubes have the smallest gaps and the largest α_{\parallel} ; the inverse-square dependence on the gap roughly holds over two orders of magnitude. Only the narrowest nanotubes (7,0) and (8,0) deviate from this trend due to significant effects of the curvature. The trend is particularly accurate for large-gap zigzag nanotubes $(3n+1, 0)$ and $(3n+2, 0)$ with $n > 2$. We note that for these large-gap SWNTs our first-principles results can be fitted well with these relations:

$$\Delta_g \approx 3.3/R_0 + 0.06 \quad (4.56)$$

$$\alpha_{\parallel} \approx 8.2R_0^2 + 20.5 \quad (4.57)$$

with the DFT band gap Δ_g given in eV and R_0 in Å. Previous tight-binding studies [9] reported values of α_{\parallel} comparable to our results, and noted a relation $\alpha_{\parallel} \sim R_0/\Delta_g^2$ which we also observe for large-gap nanotubes (see Fig. 4-7). For completeness it should be mentioned that in infinitely-long armchair carbon SWNTs the longitudinal polarizability per unit length α_{\parallel} diverges since there is no gap in the band dispersions and the density of states at the Fermi level is nonzero. To get a sense of scaling we can approximate such nanotubes as metallic ellipsoids of length l and transverse

radius R ($l \gg R$). The classical result is

$$\alpha_{\parallel} \sim l^2 / [24(\ln(l/R) - 1)]. \quad (4.58)$$

as discussed briefly in Appendix C. As a point of comparison, boron-nitride nanotubes of all chiralities are semiconducting and have finite longitudinal polarizabilities. Because of the wide gap in the band structure of 2D boron nitride, all single-wall BNNTs have gaps that are insensitive to chiralities. Longitudinal response is most sensitive to the radius, and a linear dependence is observed, as is expected from (4.55). The results for boron-nitride nanotubes are presented in Table 4.3 and plotted in Figure 4-8 as a function of the radius.

For multi-walled nanotubes, the longitudinal response picture is readily generalized: depolarization effects along the axis are negligible, and constituent tubes have very weak dielectric interactions. The total polarizability α_{\parallel} should then simply be the sum of the polarizabilities of constituent SWNTs; this conclusion is confirmed by our results as given in Table 4.2. This immediately implies that the longitudinal response of a MWNT is dominated by a SWNT in it with the smallest band gap or an armchair SWNT.

4.3.2 Transverse polarizabilities and screening factors

In this section we present the results of our first-principles calculations of the transverse polarizabilities per unit length and the screening properties of isolated single- and multi-wall nanotubes. Again, since we are using periodic boundary conditions, we must be careful in calculating transverse polarizabilities and screening factors because the value of the macroscopic electric field is affected by depolarization effects in the periodic array of nanotubes. The issue of correctly extracting the values for isolated nanotubes is discussed in detail in Section 6.2, and here we consider already the corrected results. Again, we use two different methods of calculations: linear response with DPFT yields the full polarizability tensor which we use to extract both α_{\parallel} and α_{\perp} . Alternatively, applying a finite electric field perpendicular to the nanotube and

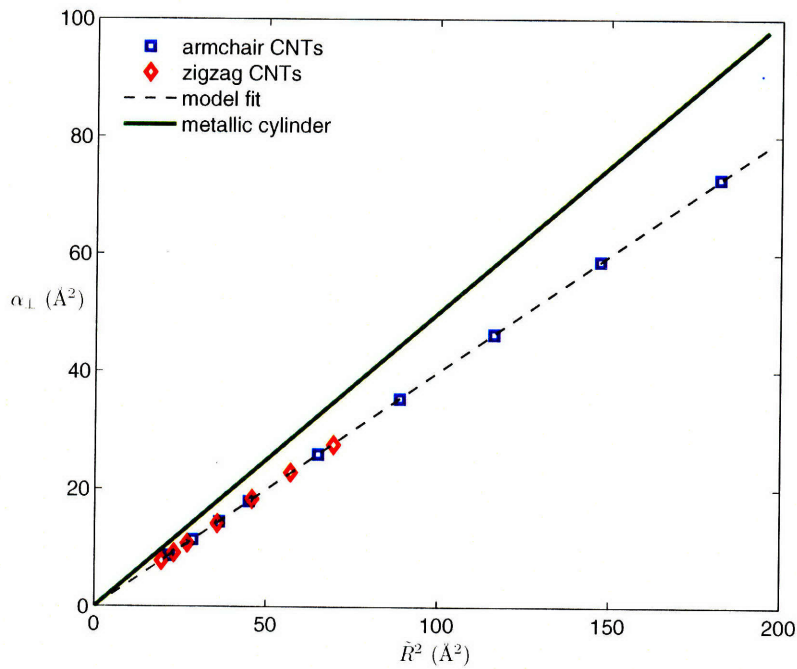


Figure 4-9: Transverse polarizabilities α_{\perp} of armchair and zigzag nanotubes as a function of \tilde{R}^2 . The dashed line is the best-fit result of our semi-metallic shell model; the solid line $\alpha_{\perp} = \frac{1}{2}\tilde{R}^2$ corresponds to an ideal metallic cylinder.

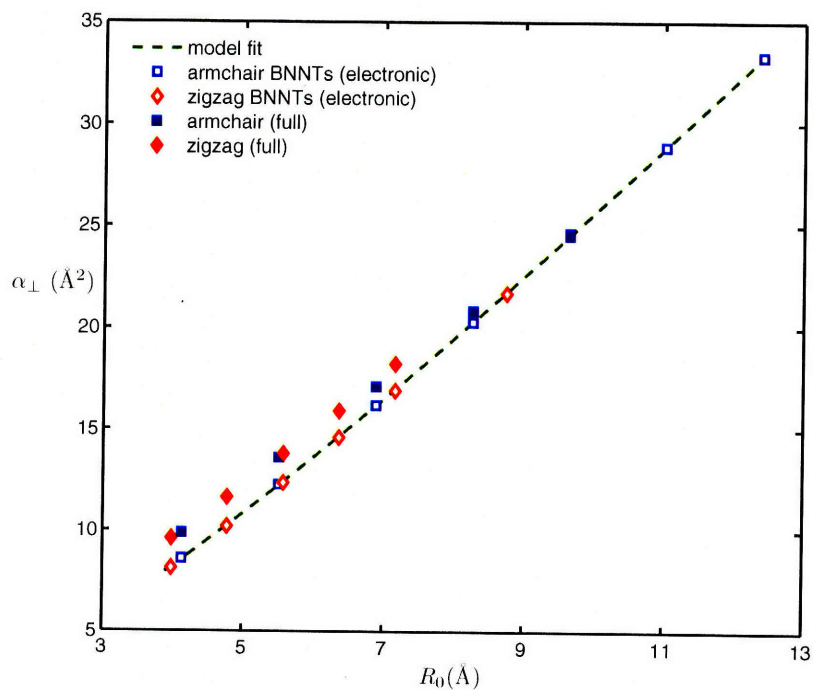


Figure 4-10: Transverse polarizabilities of boron-nitride nanotubes as a function of the radius. The dashed line is the best-fit result of our thick dielectric shell model.

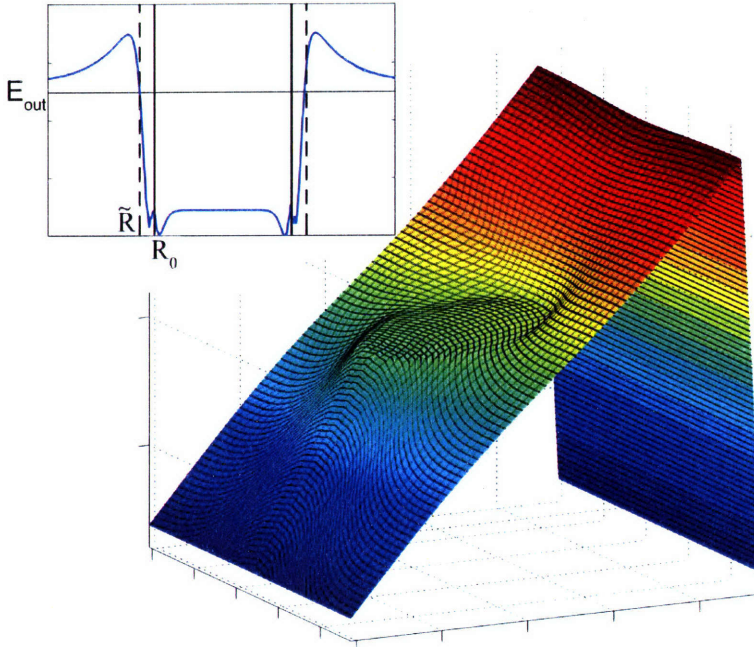


Figure 4-11: Electrostatic potential for a (10,10) SWNT in an applied homogeneous transverse field E_{out} . The electric field through the center slice is shown in the inset.

calculating the induced dipole gives the same result for α_{\perp} at a lower computational cost. The values of α_{\perp} for carbon nanotubes are listed in Table 4.2 and plotted in Fig. 4-9 as a function of the square of the effective outer radius $\tilde{R} = R_0 + 1.3\text{\AA}$ (see later discussion in Sec. 4.3.3). Remarkably, the transverse polarizabilities of both metallic and semiconducting carbon SWNTs lie on the same curve, which can be fitted by a line $\alpha_{\perp} = c\tilde{R}^2$ with a slope $c = 0.40$. Thus chirality and longitudinal band structure have a negligible effect on the transverse dielectric response; this was observed in earlier calculations [9, 13, 12] and justified with symmetry arguments in the single-particle approximation [9].

By applying a finite transverse field E_{out} (via a sawtooth periodic potential) we can also study screening inside a nanotube; we find the inner field E_{in} to be very uniform in both CNTs (see Fig. 4-11) and BNNTs, reminiscent of a classical dielectric of cylindrical symmetry. Another remarkable feature is that the screening factor in carbon nanotubes $E_{out}/E_{in} \rightarrow 5$ turns out to be relatively insensitive to radii and chiralities for all single-wall CNTs, in quantitative agreement with previous estimates [10]. Recent model system studies [10, 11] predict a small and systematic difference

(n, m)	R_0 (Å)	Δ_g (eV)	E_{out}/E_{in}	α_{\perp} (Å ²)	α_{\parallel} (Å ²)
(8,0)	3.13	0.57	4.57	7.83	104
(9,0)	3.53	0.17	4.24	9.16	1460
(10,0)	3.92	0.91	4.59	10.81	142
(11,0)	4.31	0.77	4.74	12.53	186
(12,0)	4.70	0.087	4.49	14.20	6140
(13,0)	5.09	0.72	4.83	16.29	224
(14,0)	5.48	0.63	4.89	18.37	279
(15,0)	5.88	0.041	4.61	20.33	11100
(16,0)	6.27	0.61	4.94	22.88	326
(17,0)	6.66	0.53	5.00	25.35	395
(18,0)	7.05		4.69	27.59	
(20,0)	7.83		5.06	33.45	
(8,0)+(17,0)				25.8 (25.7)	499
(8,0)+(16,0)				23.6	427
(5,5)	3.39	(0)	4.53	8.72	(∞)
(6,6)	4.07	-	4.58	11.42	-
(7,7)	4.75	-	4.64	14.49	-
(8,8)	5.43	-	4.68	17.93	-
(10,10)	6.78	-	4.74	25.91	-
(12,12)	8.14	-	4.78	35.36	-
(14,14)	9.50	-	4.82	46.30	-
(16,16)	10.86	-	4.85	58.73	-
(18,18)	12.21	-	4.94	72.83	-
(10,10)+(16,16)			12.86 (16.70)	60.81 (60.07)	
(6,6)+(18,18)			19.02 (20.82)	73.38 (72.86)	
(12,12)+(18,18)			12.56 (16.03)	75.63 (74.51)	
(6,6)+(12,12)+(18,18)			36.01 (59.4)	75.94 (74.57)	

Table 4.2: Radius, band gap, screening ratio, longitudinal and transverse polarizabilities (per unit length) of carbon nanotubes as a function of the chiral vector (n, m) . R_0 is the radius of the carbon backbone. The numbers in parentheses have been obtained with the electrostatic model for multi-wall CNTs.

(n, m)	E_{out}/E_{in}	α_{\perp} (\AA^2)	α_{\perp}^r (\AA^2)	α_{\parallel} (\AA^2)	R_0 (\AA)
(10,0)	1.66	8.16	9.62	25.05	3.98
(12,0)	1.62	10.21	11.65	29.90	4.77
(14,0)	1.57	12.36	13.79	34.76	5.57
(16,0)	1.53	14.60	15.91	39.63	6.36
(18,0)	1.49	16.91	18.23	44.51	7.16
(22,0)	1.42	21.70		54.30	8.75
(6,6)	1.68	8.62	9.87	25.39	4.13
(8,8)	1.59	12.27	13.60	33.95	5.51
(10,10)	1.51	16.18	17.10	42.48	6.89
(12,12)	1.45	20.28	20.83	51.02	8.27
(14,14)	1.39	24.52	24.61	59.54	9.65
(16,16)	1.35	28.88		68.09	11.02
(18,18)	1.31	33.32		104.7	12.40
(12,12)+(6,6)	2.05 (2.17)	25.06 (24.92)			
(18,18)+(12,12)	1.69 (1.75)	46.10 (46.00)			
(18,18)+(12,12)+(6,6)	2.33 (2.51)	49.61 (49.38)			

Table 4.3: Screening ratios, transverse polarizabilities (per unit length), longitudinal polarizabilities and radii of boron-nitride nanotubes as a function of the chiral vector (n, m) . R_0 is the radius of the B-N backbone and α_{\perp}^r is the full static polarizability that includes ionic response. The numbers in parentheses have been obtained with the electrostatic model for multi-wall BNNTs.

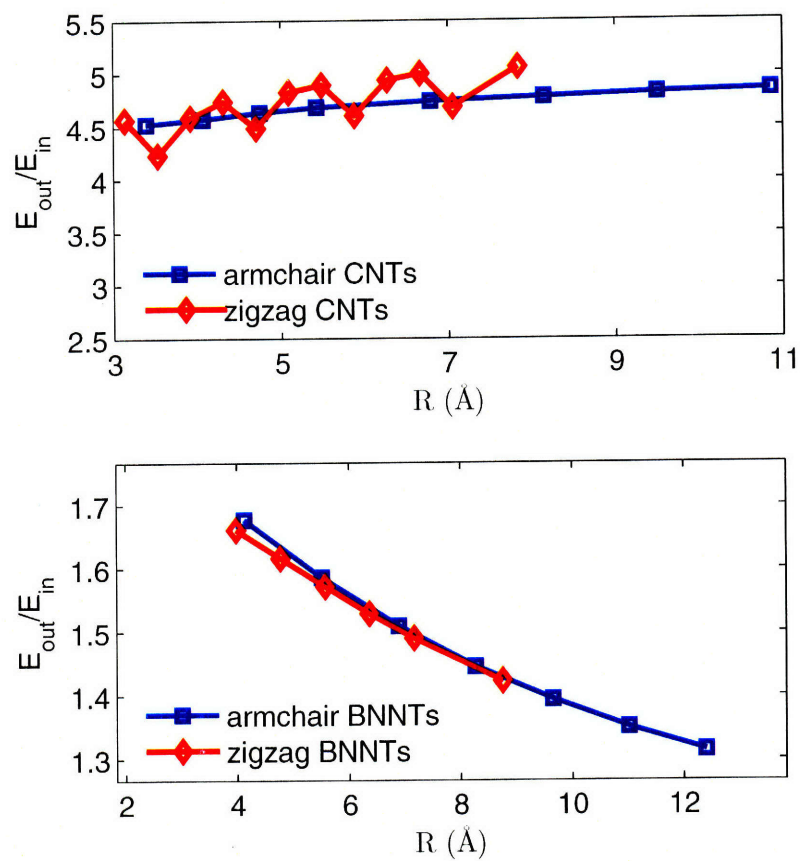


Figure 4-12: Screening factors of carbon and boron-nitride nanotubes versus the radius.

between polarizabilities of metallic and semiconducting SWNTs. We do not detect these differences of polarizabilities in our DFT calculations, possibly because of the limited range of radii of the nanotubes under consideration; however, we do observe that the screening factors are slightly different in semiconducting CNTs from those of metallic and small-gap zigzag CNTs, in quantitative agreement with Ref. [10] (see Fig. 4-12). Indeed, as we will see in Sec. 4.3.3, the value of the screening factors changes significantly for even very small deviations in the polarizability slope, which makes these chirality-dependent effects very difficult to detect with polarizability calculations alone.

So far we have only considered the response of the electronic charge density while holding the ions fixed. This is a good model for the full static polarizability of carbon nanotubes because the carbon backbone is symmetric. The situation is different for boron-nitride because it consists of two distinct sublattices of differently charged ions. When a transverse electric field is applied to a BNNT the backbone structure gets distorted and the two sublattices move in opposite direction, thus creating an extra depolarizing field and increasing the total induced dipole moment. We compute the full static polarizability α_{\perp}^r that includes also the ionic response to the external electric field by allowing the atomic structure to relax using damped dynamics during the DFT calculation with electric field. As shown in Fig. 4-10 the relative ionic contribution decreases for larger nanotubes, seemingly much more so in armchair nanotubes than in zigzag.

The calculated polarizabilities and screening factors of BNNTs are listed in Table 4.3 and plotted in Fig. 4-10 as a function of the radius. An important difference from CNTs is immediately apparent in the dependence of α_{\perp} on the radius, which is *linear* for single-wall BNNTs and *quadratic* for CNTs. At the same time Fig. 4-12 shows that the screening factor for BNNTs decreases with radius and approaches 1. The reasons for these qualitative differences are related to the fundamentally distinct electronic properties of boron-nitride and graphene sheets, as discussed in Sec. 4.2, and we will consider them in Sec. 4.3.3.

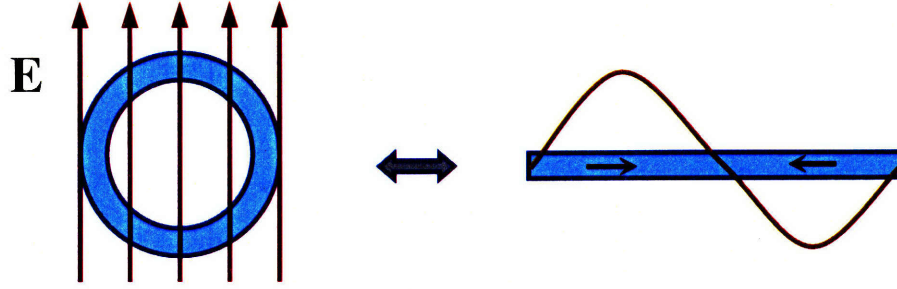


Figure 4-13: Transverse response of a tube to a uniform field can be viewed as the in-plane response of a sheet to a sinusoidal potential.

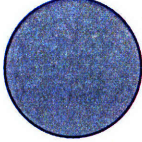


	Solid dielectric cylinder	$\alpha_{\perp} = \frac{1}{2} \frac{\epsilon - 1}{\epsilon + 1} R^2$	$E_{\text{out}} / E_{\text{in}} = \text{const}$ but wrong physical picture
	Thick dielectric shell BNNT	$\alpha_{\perp} = c'R$	$E_{\text{out}} / E_{\text{in}} \rightarrow 1$
	Semi-metallic surface shell $\rho(\phi) = \rho_0 \cos \phi$ CNT	$\alpha_{\perp} = \frac{\pi \rho_0}{E_0} R^2$	$E_{\text{out}} / E_{\text{in}} = \text{const}$ independent of R

Figure 4-14: Electrostatic models for transverse response of single-wall nanotubes. The first line has the correct radius dependence for CNTs but is obviously wrong.

4.3.3 Electrostatic models of nanotubes

To make sense of the different features of the transverse response in single-wall CNTs and BNNTs, we look for a simple electrostatic model that would reproduce these traits. We begin by arguing that since the out-of-plane polarizability per unit area of the two 2D sheets is much smaller than the in-plane polarizability $\alpha_{\parallel}^{2D} \gg \alpha_{\perp}^{2D}$, we can map the transverse charge density response of a single-wall nanotube to the response of the corresponding 2D material to an in-plane sinusoidal potential (see Fig. 4-13). We now explore this correspondence to obtain several notable results. For a cylinder of radius R the corresponding value of the perturbation wavevector is $q = 2\pi/\lambda = 1/R$. An applied external field E_0 corresponds to an in-plane field

$E(x) = E_0 \sin qx$, or, equivalently, to an applied potential $V(x) = E_0 R \cos(x/R)$. Let $\rho(\phi) = \rho_0 \cos \phi$ be the surface charge density induced on the thin cylinder, with ϕ measuring the angle to the applied field E_0 . The induced dipole moment per unit length is given by

$$p_{\perp} = \alpha_{\perp} E_0 = \int_0^{2\pi} z \rho(\phi) R d\phi = \int_0^{2\pi} R \cos^2(\phi) \rho_0 R d\phi = \pi R^2 \rho_0. \quad (4.59)$$

The corresponding induced charge density on the sheet is $\rho(x) = \rho_0 \cos(x/R)$ so that the response function is $\chi(q) = -\rho_0/V_0 = -\rho_0/(E_0 R)$. We then obtain the identity

$$\alpha_{\perp} = -\pi R^3 \chi(q). \quad (4.60)$$

The field inside the cylinder due to the induced charge is uniform and at the origin is given by

$$\Delta E = \int_0^{2\pi} \frac{2\rho(\phi) \cos \phi}{|R - 0|} R d\phi = \int_0^{2\pi} 2\rho_0 \cos^2 \phi d\phi = 2\pi \rho_0 \quad (4.61)$$

so that the total field inside is $E_{in} = E_0 - \Delta E = E_0(1 + 2\pi\chi(q)R)$ which implies that the screening factor is just equal to the in-plane dielectric constant of the sheet:

$$\frac{E_{out}}{E_{in}} = \frac{1}{1 + \frac{2\pi}{q}\chi(q)} = \epsilon_{\parallel}(q) \quad (4.62)$$

We now recall from Sec. 4.2 that for a BN sheet $\alpha_{\parallel} \approx 1.08 \text{ \AA}$ so $\chi(q) \approx -1.08q^2$ which implies from (4.60) that $\alpha_{\perp} \approx 3.39R$. On the other hand, fitting the curve on Fig. 4-10 gives a slope of 3.12 which is quite close even at this level of approximation. At the same time the screening factor (4.62) is expected to decrease with radius and approach 1, and this is indeed observed on Fig 4-12.

By extrapolating the data on Fig. 4-2 to $q \rightarrow 0$ for graphene we had $\chi(q) \approx -0.8\frac{q}{2\pi}$. It follows from(4.60) that for carbon SWNTs we expect $\alpha_{\perp} = cR^2$ with coefficient $c \approx 0.4$. A best fit of our ab-initio data for the CNT polarizability (see Fig. 4-9) shows remarkable agreement with this prediction and yields $\alpha_{\perp} = 0.40\tilde{R}^2$.

Here the effective radius $\tilde{R} = R_0 + 1.30\text{\AA}$ is larger than that of the carbon backbone R_0 , consistent with the finite thickness of the electronic charge density distribution and the fact that π -electrons start screening the field at a larger radius. Within this simple model of folding graphene into a cylinder, the screening factor (4.62) is predicted to have a radius-independent value

$$\frac{E_{out}}{E_{in}} = \frac{1}{1 - 2c} \approx 5 \quad (4.63)$$

that agrees with the results plotted in Fig 4-12. We point out that the value of the screening factor diverges for $c = 0.5$ which corresponds to complete screening in the classical metallic case. In fact, the response of a metallic cylinder has very similar features, only with a different coefficient, a radius-invariant (infinite) screening factor and the transverse polarizability $\alpha_{\perp} = \frac{1}{2}R^2$. The screening properties of carbon nanotubes, reflected in this model, are therefore *neither* metallic nor insulating. This peculiarity is physically grounded in the fact that in a graphene sheet the screening of Coulomb interactions is scale-invariant. For carbon nanotubes (as opposed to boron-nitride nanotubes), this semi-metallic nature of π -electrons implies that the screening factor is finite and radius-invariant. It turns out that this simple model, with two parameters $c = 0.4$ and $\Delta R = 1.30\text{\AA}$, for transverse response of single-wall CNTs is extremely accurate and captures all the trends observed in ab-initio calculations. For completeness, we mention that a solid dielectric cylinder of radius \tilde{R} and bulk dielectric constant ϵ also possesses the right radius dependence, with polarizability $\alpha_{\perp} = \frac{1}{2} \frac{\epsilon-1}{\epsilon+1} \tilde{R}^2$, a uniform inner field and a screening factor $E_{out}/E_{in} = (\epsilon + 1)/2$ independent of radius. This picture, however, does not correspond to a nanotube, where the screening is accomplished by a thin layer of electrons.

Generalizing this simple picture to multi-wall nanotubes is difficult without a more detailed model, because we need to take into account the screening and electrostatic interactions between layers. We can expect to capture these electrostatic effects by using a model of several concentric dielectric cylinders, whose thickness and dielectric constant we can adjust to represent correctly the single-wall response features of

CNTs and BNNTs, and then solving for the arbitrary N -layer system response. For a single classical dielectric cylinder with inner and outer radii R_- and R_+ and dielectric constant ϵ we have (as a special $N=1$ case of the results derived in Appendix A)

$$\alpha_{\perp} = \frac{1}{2} \frac{R_+^2 (R_+^2 - R_-^2) (\epsilon^2 - 1)}{R_+^2 (\epsilon + 1)^2 - R_-^2 (\epsilon - 1)^2} \quad (4.64)$$

$$\frac{E_{out}}{E_{in}} = \frac{(R_+^2 - R_-^2) (\epsilon^2 - 1)}{R_+^2 (\epsilon + 1)^2 - R_-^2 (\epsilon - 1)^2}. \quad (4.65)$$

For single-wall BNNTs, fitting the polarizability values to expression (4.64) gives $R_{\pm} = R_0 \pm 1.21\text{\AA}$ where R_0 is the radius of the BN backbone, and $\epsilon = 6.06$. This value for the dielectric constant comes out very close to the experimental value of 6.0 for the isotropic average of the static dielectric constant of hexagonal boron-nitride [50]. It is reassuring that by fitting only to the first expression (4.64) we are also able to recover the screening factors very close to the ab-initio numbers (see Table 4.3). Now that we have the parameters that describe well the single-wall tube, we can construct a general multi-layer system. Modeling a MWNT amounts to solving a linear system of $2N \times 2N$ boundary-condition equations whose derivation is presented in Appendix A and implemented in Matlab (see Appendix B). Confirming the reliability of the model is the fact that when we generalize to multi-wall tubes by adding more layers, we reproduce to within a few percent the ab-initio polarizabilities and screening factors of multi-wall BNNTs (see Table 4.3).

Response of carbon nanotubes has a qualitatively different radius dependence and we have seen that the infinitesimally thin semi-metallic shell model is extremely accurate. A simple derivation verifies that this model is exactly the limit of (4.64) and (4.65) as one takes the thickness of the tube $\delta = R_+ - R_-$ to zero while maintaining the condition

$$\epsilon \frac{\delta}{R} = \frac{4c}{1 - 2c} = \text{const} \quad (4.66)$$

that guarantees that the screening factor remains independent of the effective radius $\tilde{R} = R_0 + \Delta R$. As an independent verification, fitting the ab-initio results α_{\perp} in Table 4.2 versus R_0 for single-wall CNTs to the full expression (4.64) results in a

very thin shell with radius shifted by $\Delta R = 1.3 \text{ \AA}$, with the corresponding values of δ and ϵ depending on the fit accuracy parameters with the ratio reproducing $c = 0.4$ from (4.66). With the exact solution of the general problem of N concentric dielectric cylindrical shells in a uniform field at our disposal, we can readily extend the model to many layers, taking care in assigning (small) thicknesses and (large) dielectric constants to each layer according to the condition (4.66). Using only the best-fit parameters c and ΔR , carried over from the single-wall fit, we find an excellent agreement between this model and our ab-initio results for polarizabilities in the double- and triple-wall cases (see Table 4.2). The agreement for the screening factors is worse, because according to (4.63) they are very sensitive to even small errors in the values of c , since its value of 0.4 is rather close to the metallic value of 0.5 for which the screening factor is infinite. We conclude that the present semi-metallic shell model captures all the characteristics of the transverse dielectric response of CNTs: a uniform inner field, the correct radius dependence of screening factor and α_{\perp} in SWNTs, and accurate values of α_{\perp} for MWNTs. This model has only two independent parameters c and ΔR , obtained by fitting the computed polarizabilities of SWNTs. Likewise, in the case of BNNTs the thick dielectric shell model has independent parameters δ and ϵ that are incorporated into the general model of MWNTs.

It is interesting to look at how dielectric quantities scale with the number of layers in both types of MWNTs, and we can investigate the trends at no cost using our simple multi-shell electrostatic models. The results are plotted in Figs. 4-15 - 4-18. We notice that for CNTs by far the largest contribution to the transverse polarizability comes from the outer few layers, and inner layers play a negligible role due to a combination of their smaller radii and the strong screening due to the outer layers. The total screening factor is exponential in the number of layers, with just 6 layers enough to screen the outside field by a factor of 100. In BNNTs, on the other hand, screening due to larger tubes is weak, so the largest contribution to the screening factor comes for the innermost layers. Also, all layers contribute more evenly to the total transverse polarizability.

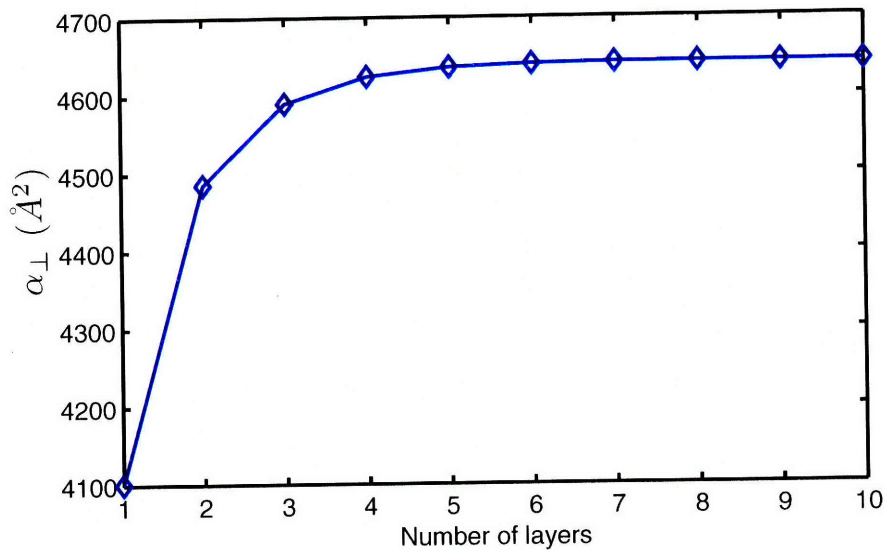


Figure 4-15: Transverse polarizability of a carbon MWNT of radius 100\AA versus the number of inner layers, as predicted by our electrostatic model. The outer few layers strongly dominate the transverse response.

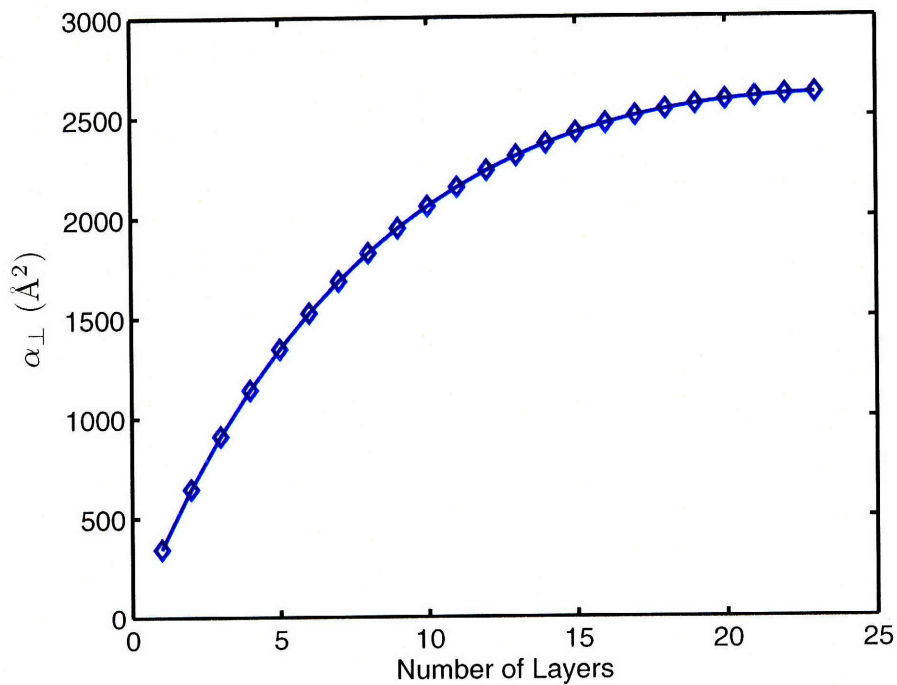


Figure 4-16: Transverse polarizability of a boron-nitride MWNT of radius 100\AA versus the number of inner layers, as predicted by our electrostatic model.

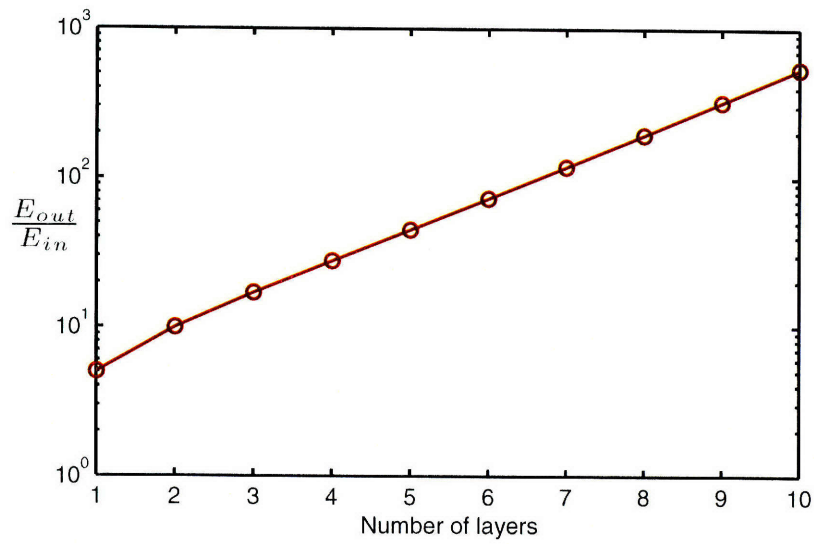


Figure 4-17: Screening factors in a carbon MWNT of outer radius 100Å versus the number of nanotube layers, as predicted by our electrostatic model. The vertical axis is on the log scale, so screening is exponential in the number of layers.

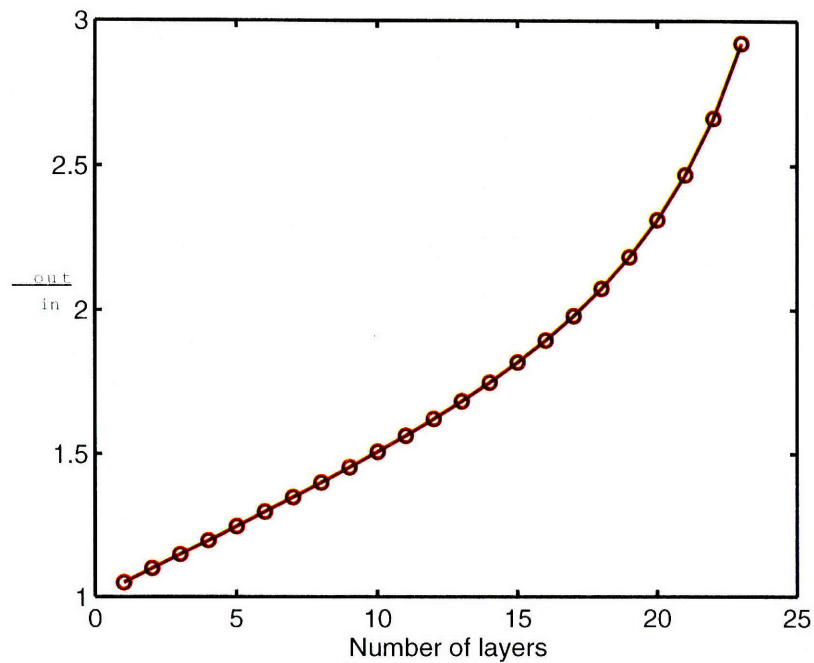


Figure 4-18: Screening factors of a boron-nitride MWNT of outer radius 100Å versus the number of inner layers, as predicted by our electrostatic model. The innermost layers contribute more to the transverse response.

Before concluding, we can apply the above discussion to the question of validity of bond polarizability models in calculating nanotube polarizabilities. Such models are widely used for estimating Raman intensities [4]. In the case of nanotubes, the intensity of the radial breathing mode is related to the derivative $\partial\alpha_{\perp}/\partial R$ and is a very important tool in determining a nanotube's chirality experimentally. Bond polarizability models assume a local nature of dielectric response. Such a model has been used for BNNTs [51] by treating the nanotube as a collection of polarizable sites and taking into account depolarizing fields in a classical way. Such a method will certainly fail for CNTs because of the scale-invariance of the response and the delocalized nature of the dominant π -electrons, i.e. one cannot get the correct radius dependence by assigning a constant polarizability to local sites in a carbon nanotube.

4.3.4 Nanotube bundles

Periodic boundary conditions allow us to easily examine the bulk dielectric response of homogeneous nanotube bundles [52]. We compute ϵ_{\perp} and ϵ_{\parallel} of triangular and square arrays with inter-tube separation of $L=3.4\text{\AA}$ [53]. The resulting values of ϵ_{\parallel} accurately match those computed from α_{\parallel} of isolated nanotubes using (4.54), with the L^2 factor adjusted to the cross-sectional area of the unit cell of the bundle. We recognize again the additive property of the longitudinal response. The additivity of α_{\parallel} again follows from the lack of electrostatic interaction between nanotubes in the longitudinal direction. In contrast, transverse response of bundles depends strongly on the inter-tube separation L . The transverse dielectric tensor at small L may have sizeable anisotropic and off-diagonal contributions that depend on the combined point-group symmetry of the nanotube and the lattice. Even though the experimental relevance of simulations of perfectly uniform periodic nanotube bundles is limited, it is still useful to understand the symmetry issues that arise in interpreting such

calculations. The dielectric tensor of a bundle can be generally represented as

$$\overleftrightarrow{\epsilon} = \begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} & 0 \\ \epsilon_{yx} & \epsilon_{yy} & 0 \\ 0 & 0 & \epsilon_{\parallel} \end{pmatrix} \quad (4.67)$$

where $\epsilon_{xy} = \epsilon_{yx}$ due to time-reversal symmetry. The three independent transverse components obey the following rules: (I) $\epsilon_{xy} = 0$ if the combined nanotube and lattice system has a mirror symmetry in the $x - y$ plane, and (II) $\epsilon_{xy} = 0$ and $\epsilon_{xx} = \epsilon_{yy}$ if there is a rotation symmetry. For example, a triangular bundle of (9,0) nanotubes will always satisfy (II) for any relative orientation of the nanotube and the lattice, so the dielectric tensor is diagonal; on the other hand, a square array of (9,0) tubes will only satisfy (I) for a particular orientation of the tube on the lattice. Similarly, a (12,12) nanotube will always satisfy (II) on both trigonal and square lattices because it is compatible with both rotation symmetries. It should be mentioned that chiral nanotubes have much lower symmetry and off-diagonal components will often be non-zero in the dielectric tensor.

In the limit of infinite L condition (II) is satisfied, so the symmetry-dependent effects vanish. In fact, these symmetry contributions, arising from multipole interactions, vanish more quickly with L than the dipole depolarization effects and do not affect our isolated nanotube calculations. For $L > 10$ Å the symmetry contributions are negligible and we can treat the nanotubes as structure-less dipoles and apply the correction (6.22) as described in Sec. 6.2.

4.3.5 Linearity of the response

In computing the transverse polarizabilities, the finite-field approach is also used to determine the range of fields for which the transverse dielectric response is linear. The (5,5) nanotube exhibits a highly linear response with the same polarizability coefficient to within 3 significant digits for field magnitudes of 0.05, 0.5, 5 V/nm, the last one being greater than the experimentally attainable value. This implies

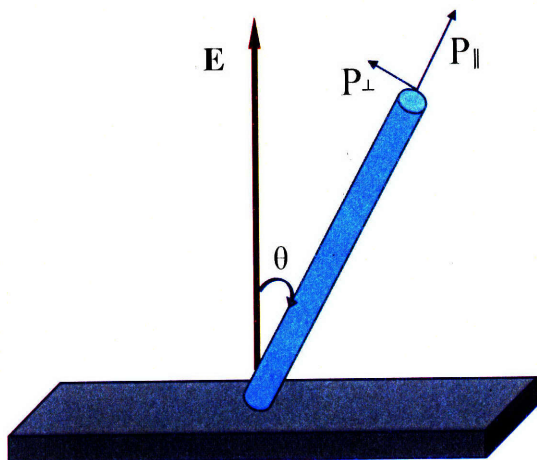


Figure 4-19: Electrostatic potential for a (10,10) SWNT in an applied homogeneous transverse field E_{out} . The electric field through the center slice is shown in the inset.

that our electrostatic shell model of transverse response remains valid in the regime of large applied fields. To study the linearity of longitudinal response, we minimize directly the electric-enthalpy functional [49] to introduce a finite longitudinal field while preserving periodic-boundary conditions. We find that the longitudinal response of the (8,0) nanotube becomes nonlinear by only 5% at $E_{\parallel}=0.5$ V/nm. Nonlinearity is in fact suppressed because zigzag and armchair (non-chiral) nanotubes are center-symmetric, so the first hyper-polarizability β vanishes by symmetry [12]. To estimate the second hyper-polarizability γ_{\parallel} we compute polarizations at several values of the field, and fit the result to the expression $P = \alpha_{\parallel}E + \gamma_{\parallel}E^3$. We obtain $\alpha_{\parallel} = 106 \text{ \AA}^2$ (in agreement with the DFPT result in Table 4.2) and $\gamma_{\parallel} = 3.1 \times 10^7$ in atomic units.

4.4 Separation and alignment of nanotubes

Knowledge of the transverse and longitudinal polarizabilities of nanotubes allows us to understand what happens when nanotubes are subjected to electric fields in experiments [54]. Let us first consider the torque on a nanotube attached to a substrate in a uniform electric field (see Fig. 4-19). A nanotube of length l at an angle θ to the

field \mathbf{E} experiences a torque

$$\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E} = \hat{\mathbf{i}} l (\alpha_{\parallel} - \alpha_{\perp}) E^2 \sin \theta \cos \theta. \quad (4.68)$$

The longitudinal and transverse polarizabilities compete with each other and determine the sign of the torque, but our results imply that $\alpha_{\parallel} > \alpha_{\perp}$ in *all* carbon nanotubes, much more so in metallic and small-gap semiconducting carbon nanotubes. Indeed, for all carbon nanotubes $\alpha_{\perp} < \frac{1}{2} \tilde{R}^2$ whereas for large-gap carbon SWNTs $\alpha_{\parallel} \gtrsim 8.2R_0^2$, and for carbon MWNTs α_{\parallel} is additive while the total α_{\perp} is always less than the sum of the individual polarizabilities. Thus, all CNTs will align with the electric field, but by tuning the value of the field during CVD growth it may be possible to selectively grow highly polarizable (e.g. mini-gap and metallic) tubes, while large-gap semiconducting tubes may not feel enough torque to align. Moreover, armchair nanotubes have a finite density of states which should make the metallic ellipsoid model of Appendix C applicable, so α_{\parallel} should increase quadratically with length, thus making long metallic nanotubes even more sensitive to aligning fields.

Comparing Figs 4-8 and 4-10 we see that longitudinal polarizabilities of single-wall boron-nitride nanotubes are also always greater than the transverse polarizabilities and that both depend linearly on the radius. Taking into account the fact that longitudinal polarizabilities are additive while transverse polarizabilities are not, we therefore predict that all multi-wall boron-nitride nanotubes will also align along the direction of the applied electric field. We are not aware of any experiments on alignment of BNNTs in electric fields. Since the torque will depend strongly on the radius, one can visualize a possibility of separating BNNTs by diameter by varying the strength of the applied field during growth.

There have also been attempts to separate semiconducting and metallic carbon nanotubes in solution using inhomogeneous electric fields [8, 7]. A polarized nanotube aligned with the field will be pulled in the direction of or against the field gradient, depending on its effective dielectric constant ϵ_{\parallel} relative to that of the solvent ϵ_{sol} ,

according to the formula

$$\mathbf{F}_{\parallel} = \frac{1}{3}lR^2(\epsilon_{\parallel} - \epsilon_{\text{sol}})\nabla E^2. \quad (4.69)$$

Assuming no solvent inside the nanotube, and approximating it by a solid dielectric cylinder of radius R_0 , we obtain from our values of α_{\parallel} an effective $\epsilon_{\parallel} = 1 + 4\alpha_{\parallel}/R_0^2 \approx 30$ for large-gap semiconducting SWNTs and obviously much larger values for small-gap and metallic tubes. This result is consistent with findings that only metallic SWNTs are observed to be deposited on the electrodes in water ($\epsilon_s = 80$), whereas all nanotubes are drawn towards the electrodes in isopropyl alcohol ($\epsilon_s = 18$). By varying the relative concentrations of these two solvents one may also increase the degree of control over the separation process.

Chapter 5

Calculation of Luttinger liquid parameters in carbon nanotubes

5.1 Introduction

Materials of greatly reduced dimension often behave in qualitatively different ways from their macroscopic counterparts, and we saw examples of that in Chapter 4. Electronic properties of common conductors are usually described well by the Fermi-liquid theory pioneered by Landau [55]. It postulates that in the interacting gas of electrons the excitations at energies close to the Fermi energy are long-lived and behave similarly to the non-interacting electrons. This gives rise to the familiar quasiparticle picture. Fermion systems in one dimension have features that are very different from those in higher dimensions. This is because the phase-space for inter-particle scattering is reduced, and the one-dimensional Fermi surface consists of two discrete points, while in higher dimensions it is usually continuous. Starting with the work by Tomonaga in 1950 [56] and later by Luttinger [57], it has become clear that the electron-electron interaction in 1D destroys the sharp Fermi surface (see Fig. 5-1) and leads to the breakdown of the usual Fermi liquid theory and the quasiparticle description. This breakdown is signaled by a vanishing quasiparticle weight Z and consequently the absence of the quasiparticle pole in the spectral function in the presence of arbitrarily weak interactions.

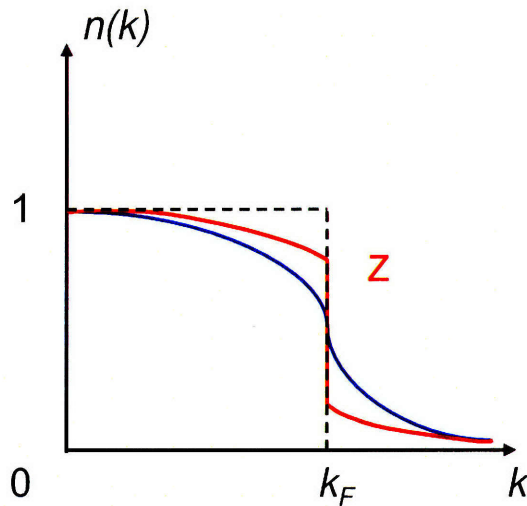


Figure 5-1: Momentum distribution function. Dotted line represents free electron gas. Red line represents the Fermi liquid, essentially similar to the free gas: the distribution still has a discontinuity at k_F but with a reduced amplitude $Z < 1$. Blue line represents the Luttinger liquid, where quasiparticles are unstable excitations.

The description of low-energy properties of gapless 1D quantum systems is based on the Tomonaga-Luttinger (TL) model, that is important because it is exactly solvable in the case of a linear fermion dispersion. In the TL model, the low-energy excitations are collective modes, involving correlated motion of many electrons, rather than the quasiparticle-like excitations found in conventional three-dimensional metals. This has profound effects on many properties of the system. One example is the single-particle density of states, a quantity that is used to predict the results of experiments when an electron is suddenly added to or removed from the system. For a three-dimensional metal, the density of states is expected to be nearly constant near the Fermi level. In contrast, in the TL model this addition or removal becomes difficult for electrons near the Fermi level, because it requires a complex rearrangement of all the other electrons in the system. As a result, the density of states approaches zero at the Fermi level, following a power law with an exponent that depends on the strength of the interactions in the one-dimensional liquid. These interactions are conventionally parameterized by a dimensionless constant g , ranging between zero and one for repulsive interactions.

As proposed in a seminal work by Haldane [58], the low-energy physics of the Tomonaga-Luttinger model is generic for interacting fermions in one dimension with repulsive interactions. Haldane demonstrated the correction terms representing non-linearity of the fermion dispersion can be added to the original model, and they give rise to interactions between the elementary excitations, the bosonic collective modes, but these interactions become weaker at lower energies. Haldane coined the name “Luttinger liquid” (LL) to reflect the idea that such systems have a low-energy spectrum similar to the Tomonaga-Luttinger model spectrum – a situation that resembles the relation between the Fermi liquid theory and the exactly solvable model of the free Fermi gas, on which it is based. The precise agreement between the predictions of the Luttinger liquid picture and a wide variety of classes of model systems, both continuum and on the 1D lattice, provide strong evidence for the universality of the Luttinger liquid description (see [59] for review). Universality means that the physical properties do not depend on details of the model, the interaction potential, etc., but instead are only characterized by a few parameters (critical exponents). In the language of the renormalization group, the TL model Hamiltonian is the low-energy fixed point for a large class of one-dimensional interacting fermion systems. Like the Fermi liquid theory, LL theory can be used to relate low-temperature, low-frequency, long-wavelength properties to a small number of parameters in which the microscopic physics of particular systems is encoded. To this date there is no theoretical model counter-example where a gapless 1D interacting fermion system does not exhibit features expected of the Luttinger liquid. This makes the Tomonaga-Luttinger model a very useful tool, since one can calculate exactly the asymptotic low-energy limits of various correlation functions of complicated model systems. One can then use these correlation functions to make predictions and compare them with those of the Fermi liquid theory, as a way to diagnose non-Fermi liquid behavior.

Interest in LL liquids has escalated in recent years and numerous experiments to probe this behavior have been performed on a variety of one-dimensional systems, including quantum Hall edge systems [60, 61], semiconductor quantum wires [62], and organic conductors [63]. Armchair carbon SWNTs in particular have proved to

be an excellent one-dimensional system, because electrons are able to travel large distances in them without scattering. There are several reasons to expect that the Luttinger liquid is a good description of metallic nanotubes. First, they have a very linear gapless spectrum near the Fermi energy; and second, the energy separating the lowest subbands from the higher ones is large, about 1eV (see Fig. 3-7). Therefore, these systems are expected to exhibit Luttinger liquid correlations for a very wide range of temperatures [64, 65]. Technically, the above statements are true also for zigzag nanotubes at the levels of doping such that the bands are linear. By attaching leads to a tube, electrons can be injected into it to probe the single-particle spectral function. In such experiments, LL behavior should appear as a power-law dependence of the nanotube conductance on the applied voltage or temperature

$$G(T) \sim T^\alpha \quad eV \ll k_B T \quad (5.1)$$

$$dI/dV \sim V^\alpha \quad eV \gg k_B T \quad (5.2)$$

and there are several experimental confirmations of this behavior [66, 67]. As an independent verification, the photoemission experiments of Ishii et al [68] also show that the photoemission spectrum follows a power law in energy, as expected from the TL model. In addition, real-space STM imaging of single-wall CNTs [69] provides convincing direct evidence of Luttinger liquid behavior.

In practice, transport measurements necessarily require several non-ideal contacts, which potentially complicates the interpretation of the data. It is in general very difficult to conduct a clean experiment, where the substrate and contact effects are known and under control. Consistency of temperature and voltage exponents is often suspect, and there is also some controversy about the role of disorder in long tubes in originating the conductance power laws [70]. For these reasons it would be very useful to be able to calculate from ab-initio the LL exponents for realistic systems, to help interpret experimental evidence and verify that a given experimental setup exhibits power-law exponents that agree with the Luttinger liquid theory. It is the goal of this chapter to develop such methods and to apply them to armchair single-wall carbon

nanotubes.

5.2 The Luttinger liquid

Following the conventional Luttinger liquid (LL) theory [71], we consider in this section a simple model: a 1D conductor containing spinless right- and left-moving electrons. In the Tomonaga-Luttinger model, the free-electron dispersion is assumed to be linear $\varepsilon_{Lk} = \pm v_F k$ around two Fermi points $\pm k_F$, and a local electron-electron interaction is parameterized by the dimensionless coupling constants g_2 and g_4 . The model Hamiltonian $H_{TL} = H_{\text{kin}} + H_{\text{int}}$ is

$$H_{\text{kin}} = i v_F \int dx \left(\psi_L^\dagger(x) \nabla \psi_L(x) - \psi_R^\dagger(x) \nabla \psi_R(x) \right), \quad (5.3)$$

$$H_{\text{int}} = \pi v_F \int dx \left[g_4 (\rho_L^2(x) + \rho_R^2(x)) + 2g_2 \rho_L(x) \rho_R(x) \right]. \quad (5.4)$$

Here $\psi_L(x)$ ($\psi_R(x)$) is the field operator for left-moving (right-moving) fermions satisfying the anti-commutation relations $\{\psi_\alpha(x), \psi_{\alpha'}^\dagger(x')\} = \delta_{\alpha\alpha'} \delta(x-x')$ ($\alpha, \alpha' = L, R$); $\rho_\alpha(x) = \psi_\alpha^\dagger(x) \psi_\alpha(x)$ are the density operators for left- and right-movers. Note that we are not including backward and unklapp scattering in the model. The total number of left (right) moving electrons is a good quantum number, so all excitations are electron-hole-like and hence have bosonic character.

It is convenient to write the Hamiltonian in terms of bosonic fields. In the hydrodynamic limit of slow spatial variation, the electron density $\rho_{L/R}(x)$ can be expressed in terms of derivative fields $\nabla \varphi_{L/R}(x)$:

$$: \rho_L(x) := \frac{1}{2\pi} \nabla \varphi_L(x), \quad : \rho_R(x) := -\frac{1}{2\pi} \nabla \varphi_R(x). \quad (5.5)$$

Here $: \rho_{L/R} := \rho_{L,R} - \langle 0 | \rho_{L,R} | 0 \rangle$ denotes the fermion normal-ordering with respect to the Fermi sea $|0\rangle$ to avoid infinities due to an infinite number of occupied states. The

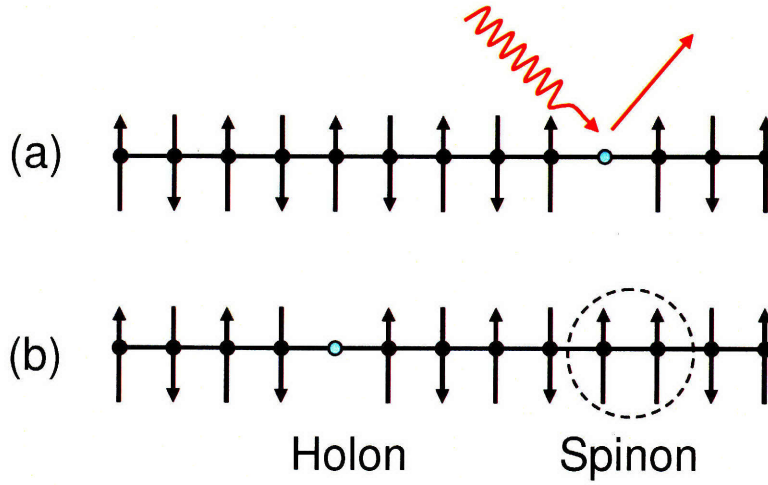


Figure 5-2: Single particle excitation in 1D induced by a photon (a) splits into an excitation containing only a charge degree of freedom and another one with spin (b).

fields $\varphi_{L,R}(x)$ satisfy the following commutation relations [72]

$$[\varphi_{L/R}(x), \varphi_{L/R}(x')] = \mp i\pi \text{sign}(x - x'), \quad [\varphi_L(x), \varphi_R(x')] = 0. \quad (5.6)$$

Let us now define

$$\phi(x) = \frac{1}{\sqrt{4\pi}} [\varphi_L(x) - \varphi_R(x)], \quad \theta(x) = \frac{1}{\sqrt{4\pi}} [\varphi_L(x) + \varphi_R(x)], \quad (5.7)$$

where $\phi(x)$ is a density variable and $\theta(x)$ is the conjugate phase variable [73]. Then one obtains

$$H_{TL} = \frac{u}{2} \int dx \left(g\Pi^2(x) + \frac{1}{g} (\nabla\phi(x))^2 \right). \quad (5.8)$$

The Hamiltonian (5.8) describes a set of harmonic oscillators, where $\phi(x)$ and the conjugate momentum $\Pi(x) = \nabla\theta(x)$ satisfy the commutation relations of conventional canonically conjugate operators: $[\phi(x), \Pi(x')] = i\delta(x - x')$. The renormalized

velocity u and the LL parameter g are given by the equations

$$u = v_F \sqrt{(1 + g_A)^2 - g_2^2}, \quad g = \sqrt{\frac{1 + g_A - g_2}{1 + g_A + g_2}}. \quad (5.9)$$

The dimensionless parameter g is a measure of the strength of the electron-electron interactions and it plays a central role in the LL theory. The noninteracting value of g (i.e., for $g_2 = 0$) is 1, and for repulsive interactions ($g_2 > 0$) g is less than 1. The outcome of this transformation is that the Hamiltonian remains quadratic in the presence of interactions, and we can solve it exactly at any interaction. The physics of such a system is described by free bosonic excitations. If we had considered also the spin degree of freedom, we would have obtained a complete separation between the charge and spin degrees of freedom, with the two types of excitations propagating independently with different renormalized velocities u . This phenomenon is known as spin-charge separation (see cartoon in Fig.5-2).

The TL model we have considered is characterized by a linear dispersion relation and the electron-electron interaction is limited to forward scattering only. It was shown by Haldane that unklapp and backscattering processes, as well as additional terms for more general models arising from band curvature are irrelevant and vanish in the low-energy limit [58]. As in the Landau Fermi liquid, a few parameters completely determine the low-energy physics: the charge degrees of freedom are described by the plasmon velocity u and the dimensionless parameter g . All correlation functions in the charge sector are uniquely determined by these two parameters. (In the absence of a magnetic field, the ground state is spin-rotationally invariant and we can ignore spin degrees of freedom). The momentum distribution function exhibits a power-law singularity at the Fermi level with exponent $\alpha = (g + g^{-1} - 2)/4$ for any nonvanishing interaction [71]. The momentum distribution function (Fig.5-1) near k_F obeys a power law

$$|n(k) - n(k_F)| \sim |k - k_F|^\alpha. \quad (5.10)$$

The single-particle density of states has the form

$$N(\omega) \sim |\omega|^\alpha \quad (5.11)$$

in the low-energy limit.

5.3 DFT approach to parameter calculation

In the TL model the density-density correlation function for $x \gg \alpha$ is

$$\langle \rho(x)\rho(0) \rangle = -\frac{g}{2\pi^2 x^2} + \frac{2}{(2\pi\alpha)} \cos(2k_F x) \left(\frac{\alpha}{x}\right)^{2g} \quad (5.12)$$

where α serves as a short-distance cut-off [71]. The first term of the correlation is Fermi liquid like; it decays as a free fermion correlation as $1/x^2$ with amplitude renormalized by the interaction. The second term, however, behaves as a power law with an exponent that depends on the interactions, whereas in a Fermi liquid that exponent would have been 2. This second term is very sensitive to the long-range non-local correlation effects that are not described well within DFT, and this power-law decay envelope should be very difficult to access in our calculations. Thus it should not be expected that a first-principles calculation can in practice detect the manifestations of the Luttinger liquid in a one-dimensional system. Another reason is that the approximate exchange-correlation kernels of DFT are tuned to mimic the *three*-dimensional interacting Fermi gas.

We will take a different route in using first-principles calculations to study Luttinger liquids in 1D systems. Motivated by Haldane's work and recent experiments, we can assume that the TL model is the true low-energy effective theory for metallic carbon nanotubes and use first-principles calculations to estimate the model's parameters that depend on the microscopic details of the interacting system. Philosophically speaking, we can view the DFT Hamiltonian of an armchair nanotube as yet another interacting 1D model system whose low-energy properties are described by the Luttinger liquid model.

The total energy of the system within DFT contains the coulomb interaction and exchange-correlation contributions, while charge density serves as the fundamental quantity in the theory. Since these two quantities are readily available and are known to be quite accurate in DFT, it is natural to look for a way to probe the strength of interactions using the formalism of the density-density linear response discussed in Sec. 4.1. It turns out that one can use the compressibility of the system, which is related to the long-wavelength limit of the response function (4.23), to estimate the interaction parameter g (see [59] and [71] for a full review), and we will adapt this approach to the framework of DFT calculations. Let us specialize to the case of simple density-density interaction so that $g_2 = g_4$ and $u = v_F/g$. In the continuum limit, the slowly-varying $q \sim 0$ part of the charge density is given by

$$\rho(x) = \rho_R(x) + \rho_L(x) = -\frac{1}{\pi} \nabla \phi(x) \quad (5.13)$$

and the average density is simply

$$n = \frac{1}{L} \int dx \langle \rho(x) \rangle = \langle \rho_R(x_0) + \rho_L(x_0) \rangle = -\frac{1}{\pi} \langle \nabla \phi(x_0) \rangle \quad (5.14)$$

since the average is independent of the point x_0 . We want to see what the effect is of varying this average density on the energy. A uniform chemical potential adds a term in the Hamiltonian, which in the boson representation looks like

$$-\mu \int dx \rho(x) = \frac{\mu}{\pi} \int dx \nabla \phi(x) \quad (5.15)$$

We can absorb this term in the Hamiltonian (5.8) by shifting the field ϕ to

$$\tilde{\phi} = \phi + \mu \frac{g}{u} x \quad (5.16)$$

The full compressibility of the system is thus

$$\begin{aligned}
K &= \frac{1}{n^2} \frac{\partial n}{\partial \mu} = \frac{1}{\langle \rho(x_0) \rangle^2} \frac{\partial \langle \rho(x_0) \rangle}{\partial \mu} = -\frac{1}{\pi n^2} \frac{\partial \langle \nabla \phi(x_0) \rangle}{\partial \mu} \\
&= \frac{g}{\pi n^2 u} = \frac{g^2}{\pi n^2 v_F} = g^2 K_0
\end{aligned} \tag{5.17}$$

which implies that the free electron gas compressibility K_0 is renormalized by the interaction. The interaction parameter g is thus related to the exact compressibility of the system, a thermodynamic observable which we have direct access to with DFT calculations via

$$\frac{\partial^2 E}{\partial n^2} = \frac{\pi v_F}{g^2} \tag{5.18}$$

where E is the total energy per unit length, and we are assuming one branch of spin-less fermions.

In order to find g we need to know the Fermi velocity, which is the second defining parameter of the TL model. It is related to the non-interacting single-particle density of states

$$\nu(0) = 1/\pi v_F = \left[\frac{\partial^2 E}{\partial n^2} \right]_0 \tag{5.19}$$

We can turn off interactions in calculating this derivative by simply *not* allowing the Kohn-Sham orbitals to rearrange after the charge density is added. Using the band structure of the neutral system, we can thus identify the slope of the highest occupied KS eigenvalue $\partial \varepsilon / \partial k$ as the Fermi velocity v_F of the non-interacting fermion system in the TL Hamiltonian (5.3).

Up to now we considered only contact interactions with coupling constants $g_2 = g_4$, but in a real material we need to deal with long-range coulomb repulsion. The derivation above is readily generalized and the Hamiltonian interaction term becomes

$$H_{\text{int}} = \frac{1}{2} \int dx dx' V(x-x') \rho(x) \rho(x') = \frac{1}{2L} \sum_q V(q) \rho(q) \rho(-q) = \frac{1}{2L} \sum_q V(q) q^2 \phi^*(q) \phi(q) \tag{5.20}$$

The full Hamiltonian (5.8) is still quadratic and can be written as

$$H_{TL} = \sum_q \frac{u(q)}{2\pi L} \left[g(q)\Pi^*(q)\Pi(q) + \frac{1}{g(q)}q^2\phi^*(q)\phi(q) \right] \quad (5.21)$$

where

$$\begin{aligned} \frac{1}{g(q)^2} &= 1 + \frac{2V(q)}{\pi v_F} \\ u(q) &= v_F/g(q) \end{aligned} \quad (5.22)$$

If $V(x)$ has a finite fourier transform for $q \rightarrow 0$ then the Luttinger parameters tend to a constant. The asymptotic behavior of the system is thus identical to the one with a local interaction whose strength is $V(q = 0)$. On the other hand, if $V(x)$ is sufficiently long-range that the fourier transform at $q \rightarrow 0$ diverges (e.g. coulomb interaction in 1D (4.39)), then this implies an infinitely strong coupling $g(q) \rightarrow 0$. Every physical 1D system, however, possesses a long-distance cutoff that depends on its environment, and in the case of the long-range bare coulomb interaction this issue can either be solved by introducing a finite length l , or it can be addressed with electrostatics if there is an external source of screening. The specifics of the system of interest will determine which cut-off scheme is more physical. For instance, if a metallic gate is at a distance $h \sim 10\text{nm}$ from a micron-long nanotube, the interaction will be screened at the distance h (see Fig. 5-3). If we take the long-distance cutoff $q = 2\pi/L$ with $L = \min\{h, l\}$ the result is

$$\frac{1}{g^2} = 1 + \frac{8}{\pi v_F} \ln \left(\frac{L}{2\pi R} \right) \quad (5.23)$$

where we have used $v^{1D}(q) \approx -2 \ln(qR)$, a factor of 2 from two bands, and another one from spin degeneracy.

So far all estimates of the LL parameters of isolated pristine carbon nanotubes in literature have been done within the RPA, i.e. neglecting exchange-correlation interactions and assuming that the low-energy description is given by (5.21), motivated by the universality of the LL description of 1D interacting systems [65, 74]. The pre-

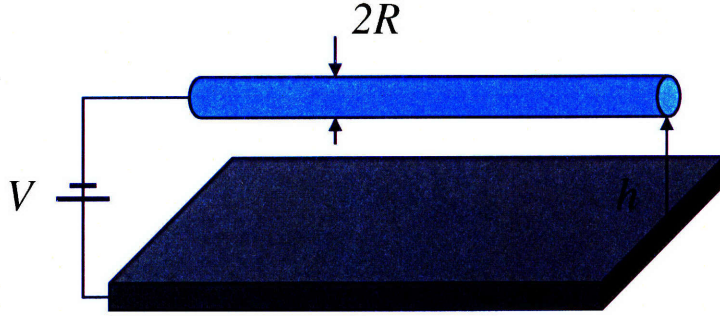


Figure 5-3: Typical experimental setup with a nanotube of radius R suspended above an infinite metallic gate distance h away.

description is to use expression (5.23) with R being the radius of the nanotube and v_F the Fermi velocity of graphene. There is a subtlety, however, in that even though the irrelevant (in the renormalization group sense) features, such as backscattering and band curvature, have vanishing effects in the low-energy limit, the effective interaction $V(q)$ gets renormalized along the way that is in general difficult to estimate. It is still reasonable to expect that the simple RPA estimate of g is accurate because the situation in armchair nanotubes is rather favorable: The bands are highly linear and it has been shown that the backscattering interactions become weaker with increasing diameter [74, 64]. On the other hand, the short-distance scale R enters as an input parameter into the estimate, and it completely ignores the renormalizing influence of various short-wavelength effects that are integrated out. Also, various environmental factors, such as transverse electric fields or the composition of the substrate, affect the values of v_F and g in ways that depend on microscopic details, and there is no easy way to accurately incorporate them into the estimate (5.23). In contrast, a direct computation of the compressibility from first-principles would offer a powerful way of controlling and studying these microscopic effects on the interaction parameters.

The prescription for the DFT calculation is as follows. The non-interacting compressibility is obtained from the slope of the bands at the Fermi energy. The interacting compressibility (5.18) is computed with finite differences by adding a small amount of charge to the system and re-calculating the total energy self-consistently

with the new density. Finally, the ratio of the two compressibilities yields the value of g according to (5.17). In calculating the interacting compressibility, we must assume some physical small- q cut-off scheme in order to obtain finite results. We will employ electrostatic screening as a way to cut off the long-range interactions and we will first consider a nanotube in vacuum separated by distance h from a metallic gate (i.e. $2h$ from its mirror image). The full compressibility of this system is related to the capacitance per unit length. Recalling now the compressibility sum rule (4.22) in linear response formalism, we can analogously separate this value into two parts

$$\frac{\partial^2 E}{\partial n^2} = \frac{1}{C} = \frac{1}{C_{cl}} + \frac{1}{\tilde{C}} = \left[2 \ln \left(\frac{2h}{R} \right) + O(R/2h) \right] + \frac{1}{\tilde{C}} \quad (5.24)$$

where the first term $1/C_{cl} = -v_q$ comes from the electrostatic geometry (Hartree contribution) and the screened contribution $1/\tilde{C}$ is intrinsic to the nanotube and is independent of $2h/R$. It includes the non-interacting kinetic energy effects and exchange-correlation contributions and thus is independent of the precise way the long-range screening is set up. This term is often referred to as the *quantum capacitance*. We can absorb it into the logarithm by introducing a renormalized length scale λ which will contain all microscopic details of the system:

$$\frac{\partial^2 E}{\partial n^2} = 2 \ln (2h/\lambda) \quad (5.25)$$

It is readily seen that if we approximate the screened quantum capacitance term by the non-interacting value given by the density of states $\tilde{C} \rightarrow C_0 = \nu(0) = 4/\pi v_F$, we recover exactly the RPA expression for the Luttinger parameter g :

$$\frac{1}{g^2} = \frac{C_0}{C} = \frac{1/C_{cl} + 1/\tilde{C}}{1/C_0} \approx 1 + \frac{8}{\pi v_F} \ln \left(\frac{2h}{R} \right) \quad (5.26)$$

A system consisting of a nanotube and an infinite metallic gate is not particularly easy to implement in first-principles calculations, although it is straightforward using the method described in Sec. 6.1 and is currently a work in progress. Fortunately, there is in fact a more convenient way to electrostatically cut off the long-range

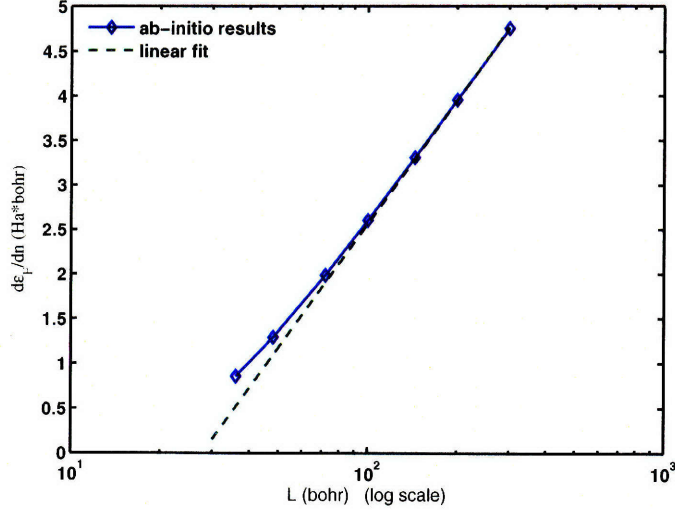


Figure 5-4: The inverse capacitance of a periodic array of (10,10) nanotubes as a function of the lattice spacing L on a log scale. The trend deviates slightly from logarithmic at small separations due to proximity effects.

interaction in DFT calculations that is equivalent at large separations. The idea is to use periodic boundary conditions to our advantage. We place a nanotube in a unit cell of transverse size L and put a small amount of charge on it. Periodic boundary conditions require overall charge neutrality, so a positive uniform background of the same total charge is added in the unit cell. Another way to say it is that the bare coulomb potential between the electrons in a nanotube is screened by the presence of other electrons in the image tubes. We can then either compute the second derivative of the total energy with respect to the average charge or, equivalently, we can use Janak's theorem [26] and compute the first derivative of the Fermi energy with respect to the occupation of the highest energy level.

$$\frac{\partial^2 E}{\partial n^2} = \frac{\partial \varepsilon_F}{\partial n} \quad (5.27)$$

The results for a (10,10) nanotube are plotted in Fig. 5-4. The logarithmic dependence on the inter-tube distance L is expected from the classical electrostatic energy per unit length of a periodic 2D array of line charges immersed in compensating

jellium to be

$$E(n) = n^2 \ln \left(\frac{L}{\lambda \gamma_2} \right) \quad (5.28)$$

where the universal Madelung parameter in $\ln \gamma_2 = 1.31053$ for 2D square lattices [75]. The deviation from the logarithmic trend at small separations is due to the proximity effects between the tubes that scale to zero as $O(R/L)$. This deviation is just the second term of the energy expansion of the periodic system (6.7), and we can remove it using the 2D PCC method discussed in Sec. 6.2. Alternatively we can extrapolate the asymptotic logarithmic trend. For example, fitting our calculations for a (10,10) nanotube at different values of L we get

$$\frac{\partial \varepsilon_F}{\partial n} \approx 1.98 \ln(L/47.9) \quad (5.29)$$

The prefactor is close to 2, as expected from electrostatics, and the constant shift determines the value of the microscopic length scale λ for this particular nanotube. Since λ includes the geometric part we should find a linear dependence on the radius R_0 , and that is precisely the case as shown in Fig. 5-5. A fit to this line for several armchair nanotube of different radii yields

$$\lambda \approx 0.85(R_0 + 1.25) \text{ \AA} \quad (5.30)$$

The dimensionless prefactor represents the contribution of the quantum (screened) capacitance while the additive shift of the radius simply means that the effective electrostatic geometric radius of a nanotube is larger than the radius R_0 of the carbon backbone. We have in fact already reached the same conclusion earlier in Sec. 4.3.2 where we obtained $\tilde{R} = R_0 + 1.30 \text{ \AA}$. We can estimate the prefactor in the RPA using (5.26), neglecting the exchange-correlation effects:

$$\frac{1}{C} \approx \frac{1}{C_{cl}} + \frac{1}{C_0} = 2 \left[\ln \left(\frac{L}{\gamma_2 R_0} \right) + \frac{\pi v_F}{8} \right], \quad (5.31)$$

which yields the value for the prefactor of 0.87 that is very close to the ab-initio

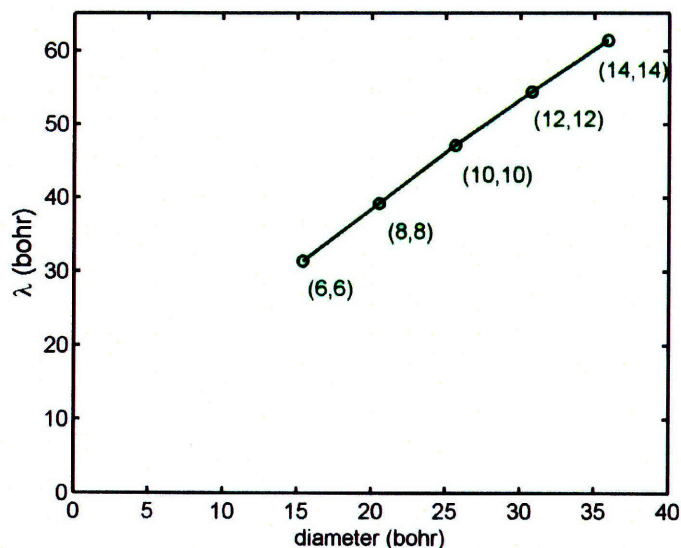


Figure 5-5: Microscopic length scale λ as determined from our calculations for nanotubes of different sizes.

result. This means that in the case of undoped armchair nanotubes, the kinetic energy contribution to the compressibility noticeably reduces the value of g from the geometric value, while the exchange-correlation contribution further reduces g by a very small amount. A complete discussion of the dependence of these relative contributions on the carrier density will be given in an upcoming publication.

The long-range logarithmic dependence on separation (either $2h$ or L) in both gated and periodic systems is the same, while the short-distance length parameter λ does not depend on the details of the long-wavelength cut-off scheme. Therefore, we can estimate the value of g for a given experimental setup once we know the screening length and the details of the structure of the nanotube and the substrate. At smaller separations the two approaches may differ due to proximity effects, and in those experimental situations it may be better to use the more physically realistic cut-off scheme of a metallic plate. We have developed an efficient method of calculation of charged objects near metallic gates (see Chap. 6). The implementation is in progress, and after it is completed it would be interesting to compare the results of both methods. Finally, let us compare the results of our calculations with experiments.

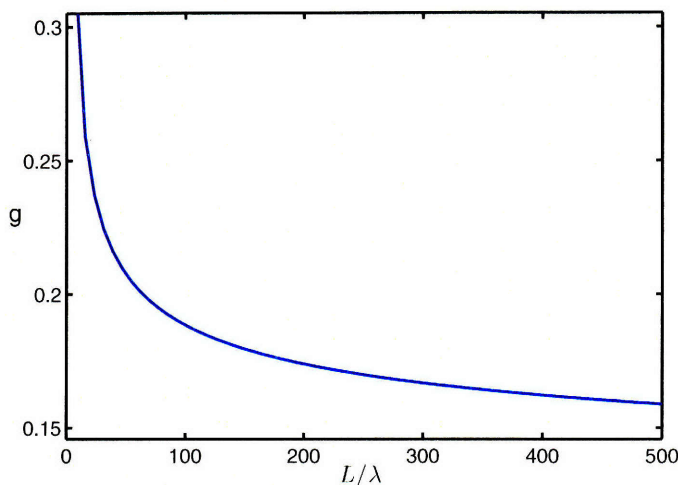


Figure 5-6: Luttinger liquid interaction parameter g as a function of the dimensionless ratio of two length scales. Here we set $v_F = 0.36$ a.u. = 8×10^5 m/s.

From (5.18) we know how the interaction parameter scales

$$g = \left(\frac{8}{\pi v_F} \ln \zeta \right)^{-1/2} \quad (5.32)$$

where ζ is a dimensionless ratio of length scales, e.g. $2h/\lambda$ or $L/\gamma_2\lambda$. This scaling, shown on Fig. 5-6, is independent of the approximation used (RPA or DFT). The goal of the DFT calculation is to estimate as closely as possible this ratio ζ of the microscopic and the long-range screening length scales that correspond to a given experimental setup. With DFT we can study very accurately the microscopic details intrinsic to the nanotube (effective electrostatic radius, exchange-correlation contributions, band nonlinearity, chirality), but the screening length must come from the knowledge of the experimental details (nanotube length, distance from the gate). From the scaling we estimate roughly that for most experiments so far, g should vary from 0.15 to 0.3, which is consistent with observations and implies that interactions can be very strong in these systems.

There are several potential practical applications of our approach. First, by replicating as closely as possible the experimental setup in an ab-initio calculation and comparing the calculated values of g , one can indirectly confirm that the observed

power laws originate from LL physics. Alternatively, if one is sure that the Luttinger picture correctly describes the physics of nanotubes, then the discrepancies in calculated and observed values of g may indicate the presence of unexpected screening mechanisms or disorder and allow a way to study them.

Chapter 6

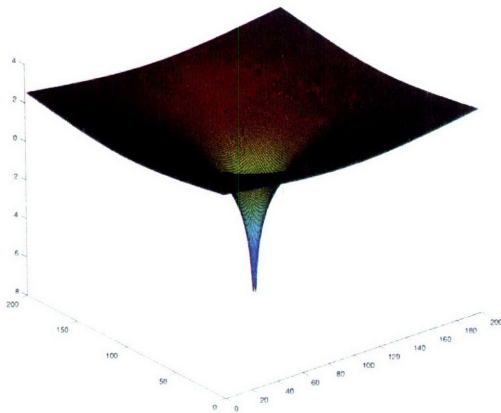
Electrostatic corrections in periodic boundary conditions

Density functional theory has become a very important tool in recent years for studying a wide range of systems on the quantum atomic level. The majority of these calculations are done with pseudopotentials in a plane-wave basis set. There are several reasons why this basis set is preferable. For a crystalline solid the plane-wave expansion is the most natural choice because it has periodicity in all three dimensions and provides a way to avoid finite-size edge effects. The basis set is independent of ionic positions and thus describes any inhomogeneous system in a consistent uniform way, with an accuracy that is easily controlled by the plane-wave cutoff energy. In addition, there are very efficient numerical FFT techniques that allow for favorable scaling of such calculations as $O(N_e^2 N_{pw})$, where N_e is the number of electrons and N_{pw} is the number of plane-wave basis elements.

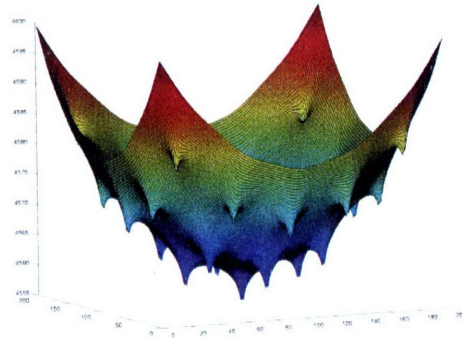
However, the 3D Bloch theorem cannot be applied to a periodic system which contains a single defect or a low-dimensional system that lacks periodicity in one or more dimensions, like an isolated slab or a tube. Although the plane-wave basis would be discrete in the periodic directions of the system, it would need to be continuous in those dimensions that lack periodicity. Therefore, at any given cutoff an infinite number of plane waves would be required for the calculation. In such situations the solution is to make a so-called supercell approximation. The supercell contains

the isolated system surrounded by vacuum or a region of a bulk crystal. Periodic boundary conditions are applied to the supercell, ensuring a discrete basis set. The presence of periodic images, however, inevitably changes the potential in the unit cell, affecting the energy and equilibrium structures of charged and polarized systems, as well as charge density response of neutral systems to an external potential, etc. To study an isolated system it is essential to include enough of vacuum (or bulk crystal), so that the periodic images are far enough away from each other to make the interaction between them negligible. Unfortunately, in practice this convergence with the unit cell size can be very slow, and complicated by the fact that the computational cost for d -dimensional extended systems in vacuum grows linearly with the number of planewaves $\sim L^{3-d}$ for a fixed energy cutoff, quickly becoming unmanageable often before even reaching the desired convergence. Fig. 6-7 illustrates the slowness of this convergence for 2D dipoles. On the other hand, it is clear that at large separations only electrostatic effects are important, with the biggest contributions given by the spurious periodic potential tails of image charges, followed by those of dipoles and higher moments.

Because the benefits of the planewave basis are so significant, it is more efficient to use it in any case and to come up with cheap ways of removing periodic image effects. Our approach is to separate the problem into a long-range classical electrostatic part and a short-range quantum part, where the latter contains exchange-correlation and kinetic energy contributions and is insensitive to boundary conditions. Within DFT this separation is achieved naturally, because charge density is a fundamental parameter and the Hartree energy is readily identified. The next step is to remove long-distance image effects using electrostatic considerations, while treating the short-range quantum effects using the corrected potential, thereby allowing for much faster convergence. While there have been a number of proposals to deal with periodic images using electrostatics [76, 77, 78], they are applicable to very specific well-behaved cases. At the same time, the effects of periodic images are still not widely appreciated, while the interest in computations of low-dimensional systems is growing. In the vast majority of calculations the approach of choice is just to increase the size of



Real-space potential of a single charge



Real-space potential of a square periodic grid of charges

Figure 6-1: Left: potential of a single 2D charge in OBC. Right: potential of a 5x5 square array of 2D charges in OBC. Note the overall parabolic shape of the potential.

the supercell until these contaminating interactions become small enough, incurring needlessly large computational costs, without utilizing the great benefit that even a first-order analytical correction can have. In this section we try to present a clear unified picture of these issues for all system dimensionalities and describe some tricks to deal with them. We also discuss the interplay of computational cost and accuracy when it comes to removing these image effects.

6.1 Charged Systems

6.1.1 Point countercharge method

In order to analytically characterize the general features of electrostatic corrections, we begin by considering classical electrostatics of systems of point charges in different dimensions. Let us take an example of a finite square array of charges in 2D with open boundary conditions (OBC), i.e. making the potential vanish at infinity. A one-dimensional line of uniform linear charge density $-n$ produces a coulomb potential $\phi^0(\mathbf{r}) = v^{2D}(r) = 2n \ln(r/\lambda)$ where λ is some effective thickness length scale. The

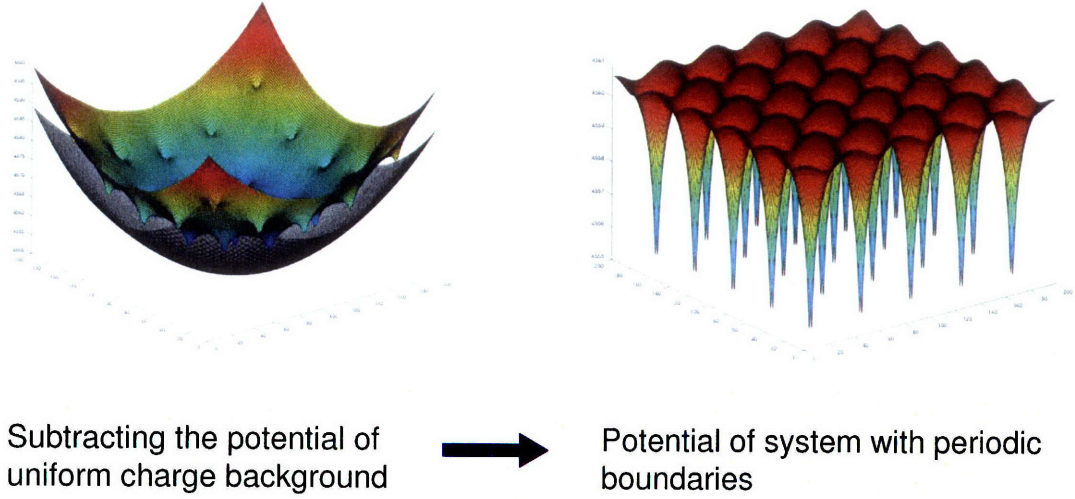


Figure 6-2: Left: potential of a 5×5 square array of 2D charges plotted together with the parabolic potential due to a uniform jellium of the same average density. Right: subtracting the jellium potential from the system yields precisely the same potential as the PBC system of point charge.

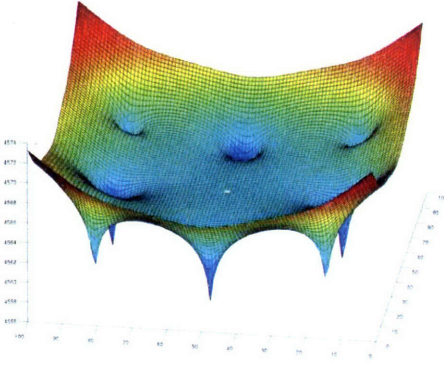
potential $\phi^{OBC}(\mathbf{r}) = \sum \phi^0(\mathbf{r} + \mathbf{L})$ in the array has a distinctively parabolic shape (see Fig. 6-1). We now take the same amount of charge and smear it uniformly over the same area, producing jellium of density $\rho^{jel} = n/L^2$, where L is the lattice spacing. It follows immediately from the Poisson equation $\nabla^2 \phi(\mathbf{r}) = -4\pi\rho(\mathbf{r})$ that the potential due to this jellium is

$$\phi^{jel}(x, y) = 2\pi \frac{n}{L^2} (x^2 + y^2) \quad (6.1)$$

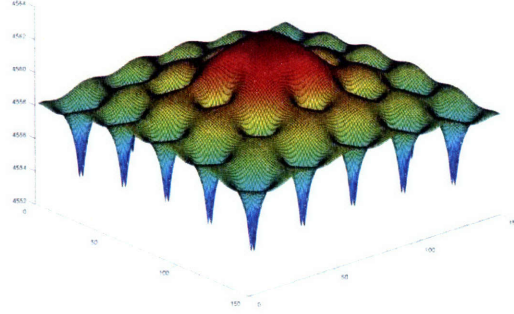
and it looks like it has the same curvature as the potential of the charge array. Subtracting this parabolic potential from $\phi^{OBC}(\mathbf{r})$ we obtain a potential that is periodic. In fact, in this model system it turns out to be precisely equal to the potential $\phi^{PBC}(\mathbf{r})$ of a single charge obtained by using periodic boundary conditions (PBC). To understand the origin of this correspondence we need to examine the periodic solution of the Poisson equation $\nabla^2 \phi(\mathbf{r}) = -4\pi\rho(\mathbf{r})$ in Fourier space:

$$\tilde{\phi}^{PBC}(\mathbf{G}) = \frac{4\pi\tilde{\rho}(\mathbf{G})}{\mathbf{G}^2} \quad (6.2)$$

where $\mathbf{G} = (i, j)\frac{2\pi}{L}$ is a vector of the reciprocal lattice. It is immediately clear that for $\mathbf{G} = 0$ the potential is ill-defined unless $\tilde{\rho}(\mathbf{G} = 0) = \langle \rho(\mathbf{r}) \rangle = 0$, in which case



OBC potential with one charge missing is locally very flat



PBC potential with one charge missing is very close to parabolic

Figure 6-3: Left: Potential of images ϕ^{img} in an OBC array (i.e. one charge removed). Right: PBC potential with an isolated charge potential removed to show the parabolic correction.

$\tilde{\phi}^{PBC}(\mathbf{G} = 0) = 0$. This means that imposing PBCs to a periodic system of charges introduces a fictitious background jellium of uniform charge density $\rho^{jel}(\mathbf{r}) = -\frac{n}{L^2}$ which compensates the total excess charge n in each unit cell and enforces $\langle \rho^{tot}(\mathbf{r}) \rangle = 0$ and $\langle \phi^{PBC}(\mathbf{r}) \rangle = 0$. Let us now study the spurious contaminating potential due to periodic boundary conditions $\phi^{corr}(\mathbf{r}) = \phi^0(\mathbf{r}) - \phi^{PBC}(\mathbf{r})$. We have already established that part of it is the parabola $-\phi^{jel}(\mathbf{r})$ that is due to the fictitious neutralizing jellium. The other part comes from the periodic image charges, and we will now examine it near the origin using a finite periodic sample in OBC. Fig. 6-2 shows this potential of an array of charges with the charge at the origin removed

$$\phi^{img}(\mathbf{r}) = \phi^{OBC}(\mathbf{r}) - \phi^0(\mathbf{r}) = \sum_{\mathbf{R} \neq 0} \phi^0(\mathbf{r} + \mathbf{R}) \quad (6.3)$$

and it looks quite flat, while $\phi^{corr}(\mathbf{r})$ looks close to parabolic at the origin. We can formalize this by writing a Taylor expansion of the potential

$$\phi^{img}(x, y) = \phi(0, 0) + \frac{\partial \phi}{\partial x} x + \frac{\partial \phi}{\partial y} y + \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial x^2} x^2 + \frac{\partial^2 \phi}{\partial y^2} y^2 + 2 \frac{\partial^2 \phi}{\partial x \partial y} xy \right) \quad (6.4)$$

If our sample has $x \leftrightarrow -x$ and $y \leftrightarrow -y$ symmetries, then all terms with odd powers of coordinates vanish. If in addition there is $x \leftrightarrow y$ symmetry then $\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial^2 \phi}{\partial y^2}$. Using the fact that there is no charge in the unit cell of interest, Poisson equation gives $\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0$ which implies that all quadratic terms in the expansion vanish as well. We conclude then that near the origin the jellium potential dominates and we can approximate $\phi^{corr} \approx \phi^{jel}$ the contaminating potential by a parabola. The symmetry conditions required for this are satisfied by a square lattice, either infinite or in a finite sample with D_4 shape symmetry (e.g. a cylindrical sample). If we now replace the point charges by some charge distributions $\rho(\mathbf{r})$ localized in 2D (e.g. nanotubes) then as long as each image system is localized within its unit cell, the shape of contaminating potential $\phi^{corr}(\mathbf{r})$ near the origin is not affected to quadratic order. In other words, a localized charge in a symmetric square lattice sample does not feel the presence of the other charges. We can use this fact as a recipe in calculating properties a truly isolated 1D system by placing it in a unit cell of square cross-section, performing a planewave calculation, and then applying the parabolic correction that we now know is good for this particular lattice. There is nothing special about using the square lattice, it is just convenient for the purpose of “isolating” the system because the correction turns out the simplest.

Once we know how to correct the potential around our system, we can correct the electrostatic energy. So far we have concentrated on the shape of $\phi^{corr}(\mathbf{r})$ and ignored any overall constant shifts caused by the PBC condition $\langle \phi(\mathbf{r}) \rangle = 0$. In fact, there is a well-defined constant potential offset that is simply the Madelung term for the lattice:

$$\phi^{corr}(\mathbf{r}) \approx 2n \ln \left(\frac{L}{\gamma_2} \right) + \frac{\pi n}{L^2} r^2, \quad (6.5)$$

where $\gamma_2 = 1.31053$ for a square lattice [75]. The difference between the energy of the periodic system (in jellium) and the isolated one is

$$E^{PBC} - E^0 = \frac{1}{2} \int \phi^{corr}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \quad (6.6)$$

and so the energy of the isolated system can then be approximated as

$$E^0 \approx E^{PBC} + n^2 \ln \left(\frac{L}{\gamma_2} \right) + \frac{\pi n}{2L^2} \int r^2 \rho(\mathbf{r}) d^2 \mathbf{r}, \quad (6.7)$$

where the last term depends on the details of $\rho(\mathbf{r})$. A similar derivation procedure in 3D for a cubic lattice yields the corresponding equations

$$\phi^{corr}(\mathbf{r}) \approx \frac{n\gamma_3}{L} + \frac{2\pi n}{3L^3} r^2 \quad (6.8)$$

$$E^0 \approx E^{PBC} + \frac{n^2\gamma_3}{2L} + \frac{\pi n}{3L^3} \int r^2 \rho(\mathbf{r}) d^3 \mathbf{r} \quad (6.9)$$

with $\gamma_3 = 2.8373$ for a cubic lattice. It turns out that this result (6.9) for the 3D cubic case is similar to the expression in ref. [77]. An example of this correction in action is shown in Fig. 6-4: We take a charged H_2^+ molecule relaxed in a large cubic cell, and then shrink the cell and observe fictitious forces that appear between the ions because of PBC effects. By correcting the total potential with $\phi^{corr}(\mathbf{r})$ as in (6.5) we are able to significantly reduce the error. The reason this procedure works so well for energy differences is that we know the correction potential $\phi^{corr}(\mathbf{r})$ quite accurately close to the origin, and that's precisely where the charge density of the system is located. We will refer to this correction approach as the point countercharge method (PCC).

For completeness, we discuss the situation in one dimension, where a point charge (sheet of charge density n) produces a potential $\phi^0(x) = 2\pi nx$. Since the potential is linear, the contribution of all other charges at the center of an OBC array vanishes identically $\phi^{img}(x) = 0$ (see Fig. 6-5). Therefore the total contaminating potential is exactly parabolic $2\pi \frac{n}{L} x^2$. The situation in 1D is thus trivial because there are no high multipoles in 1D and this result holds also for arbitrary density $\rho(\mathbf{r}) = \rho(x)$ localized in the x -direction. For a periodic array of uniformly charged slab the Madelung energy can be computed exactly:

$$\frac{1}{8\pi} \int_0^L \mathcal{E}^2(x) dx = \frac{1}{8\pi} \int_{-L/2}^{L/2} \left(4\pi \frac{n}{L} x \right)^2 dx = \frac{\pi}{6} Ln^2 \quad (6.10)$$

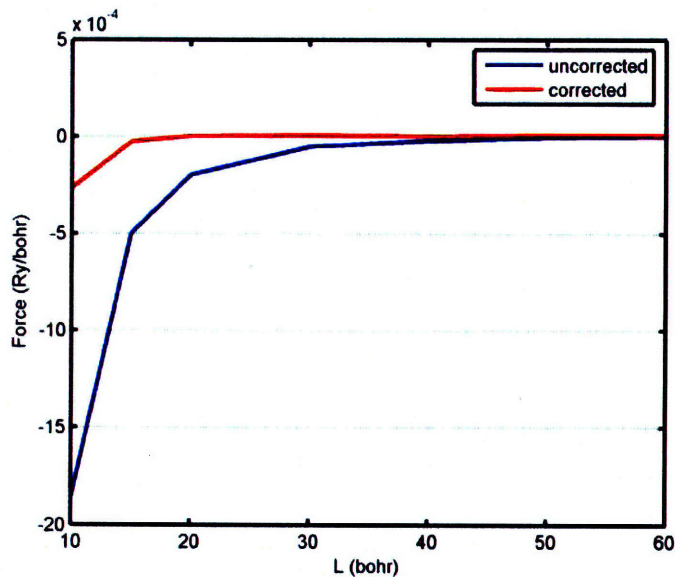


Figure 6-4: Change in the force between ions of a H_2^+ molecule due to periodic images as a function of lattice spacing.

So the corrective potential

$$\phi^{corr}(x) = \frac{\pi}{3}Ln + 2\pi\frac{n}{L}x^2 \quad (6.11)$$

gives the electrostatic energy as

$$E^0 = E^{PBC} + \frac{\pi}{6}Ln^2 + \frac{\pi n}{L} \int x^2 \rho(x) dx. \quad (6.12)$$

These corrections are exact in 1D and there are no higher-order terms for homogeneous systems.

6.2 Polarized Systems

Let us begin by considering a macroscopic sample consisting of a finite square lattice of infinite parallel nanotubes, each possessing a transverse permanent dipole \mathbf{p}_\perp per unit length. The local electric field at the site of a given dipole created by all *other*

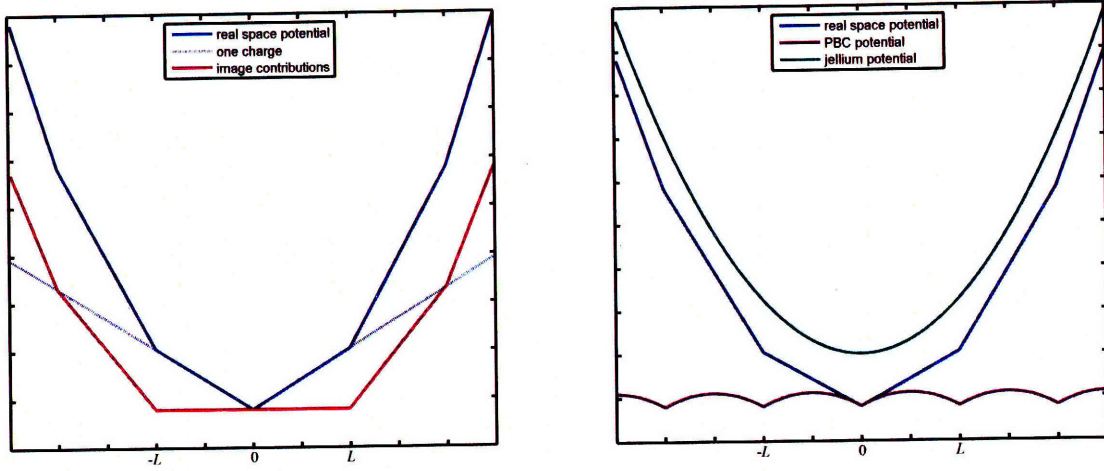


Figure 6-5: Left: because potentials of point charges are linear in 1D, the contribution of images vanishes. Lower curve shows $\phi^{img} = 0$. Right: Potentials of PBC and OBC 1D systems differ by exactly the parabolic potential ϕ^{jel} due to the fictitious jellium.

dipoles on the lattice can be evaluated as

$$\mathbf{E}^{img} = \sum_{i,j} \frac{2(\mathbf{p}_{\perp} \cdot \mathbf{x}_{ij}) \mathbf{x}_{ij} - x_{ij}^2 \mathbf{p}_{\perp}}{2x_{ij}^4} \quad (6.13)$$

where $\mathbf{x}_{ij} = (iL, jL)$ label the positions of the dipoles in the $x - y$ space. In the particular case of the field at the nanotube in the center ($\mathbf{r}=0$) of a finite lattice (bundle) whose macroscopic shape possesses a D_4 symmetry we have

$$E_x^{img} = \sum_{i,j} \frac{2(i^2 p_x + j^2 p_y) - (i^2 + j^2) p_x}{2L^2 (i^2 + j^2)} = 0 \quad (6.14)$$

and similarly $E_y^{img} = 0$. This means that in a symmetric square finite periodic lattice, a dipole experiences no local field due to other dipoles on the lattice, i.e. there are no spurious image interactions. In other words, the electrostatic potential environment of a nanotube in such a system is equivalent to that of an *isolated* nanotube. This result is well-known from the Lorenz-Lorentz formalism [79]. This fact can also be easily obtained using the results of Sec. 6.1 where we showed that $\phi^{img}(\mathbf{r}) = O(r^4)$ for square arrays of charges in OBC, and since dipoles are close pairs of opposite charges, the potential has no linear contribution, and so the field is zero. On a technical note, the cancellation in the sum (6.14) depends on the way the limit of the sample is

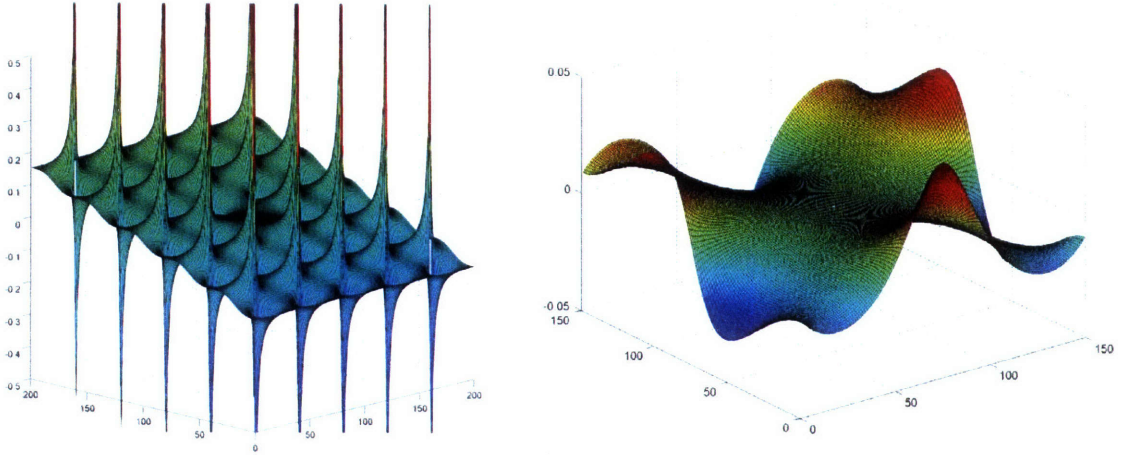


Figure 6-6: Left: Potential of a square array of 2D dipoles in OBC with one dipole removed. The average slope corresponds to the macroscopic electric field. Right: Zoom-in on the origin shows that the field there due to all other images vanishes.

taken to infinity, i.e. the field $\mathbf{E}^{\text{img}}(\mathbf{r} = 0)$ is in general sensitive to the shape of the macroscopic bundle, no matter how large it is (see Sec. 6.2.2).

The bundle's average polarization $\mathbf{P} = \mathbf{p}_\perp / L^2$ gives rise to an average macroscopic depolarizing field \mathbf{E}_P , which in our special case of a symmetric (or infinite) bundle (depolarizing factors of $n_x = n_y = 1/2$) is given [79] by $\mathbf{E}_P = 2\pi\mathbf{P}$. This is not a contradiction, since the local electric field in each unit cell vanishes only near the center, while the average value over the unit cell is $\langle \mathbf{E}^{\text{OBC}}(\mathbf{r}) \rangle = \mathbf{E}_P$. On the other hand, when we examining again the fourier components of the potential for the polarized system in PBC

$$\tilde{\mathbf{E}}^{\text{PBC}}(\mathbf{G}) = -i\mathbf{G}\tilde{\phi}^{\text{PBC}}(\mathbf{G}) = -i\mathbf{G}\frac{4\pi\tilde{\rho}(\mathbf{G})}{\mathbf{G}^2} \quad (6.15)$$

we see that the average electric field is forced to zero for any polarization $\tilde{\mathbf{E}}^{\text{PBC}}(\mathbf{G} = 0) = \langle \mathbf{E}^{\text{PBC}}(\mathbf{r}) \rangle = 0$. In other words, imposing periodic boundary conditions on a periodic lattice of dipoles amounts to introducing a compensating uniform electric field. Since in open boundary conditions the nanotube is effectively isolated from its periodic neighbors, i.e. near the nanotube $\mathbf{E}^{\text{img}} = 0$, we see that in PBCs the local field is

$$\mathbf{E}^{\text{PBC}} = \mathbf{E}^{\text{loc}} - \langle \mathbf{E}^{\text{loc}} \rangle = -\mathbf{E}_P = -2\pi\mathbf{P}, \quad (6.16)$$

i.e. the nanotube feels a field due to the dipole moments of its periodic images. The strategy for removing the effect of the periodic images in PBCs is thus first to calculate the dipole \mathbf{p}_\perp per unit length of the localized charge density of the nanotube and then to subtract a uniform field of magnitude \mathbf{E}_P . It should be mentioned that this correction is only accurate in the vicinity of the nanotube, but it works very well in practice because this is where the charge density is localized. Now we can join the charge and dipole corrections in one simple formula

$$\phi_{2D}^{corr}(\mathbf{r}) \approx 2n \ln\left(\frac{L}{\gamma_2}\right) + \pi \frac{n}{L^2} r^2 + \frac{2\pi}{L^2} \mathbf{p} \cdot \mathbf{r} \quad (6.17)$$

In 3D the depolarizing coefficient for cubic symmetry is 1/3 and we have

$$\phi_{3D}^{corr}(\mathbf{r}) \approx \frac{n}{\gamma_3 L} r^2 + \frac{2\pi n}{3L^3} r^2 + \frac{4\pi}{3L^3} \mathbf{p} \cdot \mathbf{r} \quad (6.18)$$

In 1D the correction is exact and is given by

$$\phi_{1D}^{corr}(x) = \frac{\pi n}{3} L + \frac{2\pi n}{L^2} x^2 + \frac{4\pi}{L^2} p x \quad (6.19)$$

One application is the calculation of forces in a polar molecule: in PBC the fictitious compensating electric field acts on a molecule and introduces forces that affect the equilibrium structure and vibrational properties. For example, in an H_2O molecule relaxed in a $L = 30$ bohr cubic unit cell, a fictitious force of 0.0025 Ry/bohr appears on the O atom when the unit cell is reduced to 10 bohr. By applying the 3D correction (6.18) to the potential we can reduce the spurious force by two orders of magnitude.

6.2.1 Polarizability corrections

We can now consider polarization of nanotubes in response to an external transverse electric field. As we saw above, a nanotube in PBCs feels a field due to its polarized images \mathbf{E}_P in addition to the external field \mathbf{E}^0 , given by

$$\mathbf{E}^{appl} = \mathbf{E}^0 - \mathbf{E}_P. \quad (6.20)$$

The induced dipole moment \mathbf{p}_\perp is determined by the polarizability of the isolated nanotube $\alpha_\perp = \mathbf{p}_\perp/\mathbf{E}^{\text{appl}}$, and that is the quantity we are ultimately interested in. In practice, during a PBC calculation we can only specify the value of \mathbf{E}^0 and calculate the corresponding induced \mathbf{p}_\perp , thus obtaining the bulk uncorrected value of polarizability $\alpha_\perp^b = \mathbf{p}_\perp/\mathbf{E}^0$. To obtain the corrected isolated polarizability α_\perp we can use the relation

$$\mathbf{p} = \alpha_\perp^b \mathbf{E}^0 = \alpha_\perp \mathbf{E}^{\text{appl}} = \alpha_\perp \left(\mathbf{E}^0 + \frac{2\pi}{L^2} \alpha_\perp^b \mathbf{E}^0 \right) \quad (6.21)$$

which finally yields

$$\alpha_\perp = \frac{\alpha_\perp^b}{1 + \frac{2\pi}{L^2} \alpha_\perp^b}. \quad (6.22)$$

Alternatively, one can start with the calculated bulk dielectric constant ϵ_\perp of the periodic array of nanotubes and use the relation to bulk polarizability $\alpha_\perp^b = \frac{\epsilon_\perp - 1}{4\pi}$ to obtain the conversion expression

$$\alpha_\perp = \frac{L^2 \epsilon_\perp - 1}{2\pi \epsilon_\perp + 1} \quad (6.23)$$

which is just the 2D analog of the Clausius-Mossotti formula. In Sec. 4.3.2, we address the characterization of the transverse dielectric response of nanotubes in two different ways. First, we calculate the dielectric tensor $\overleftrightarrow{\epsilon}$ using density-functional perturbation theory, and extract the transverse polarizability α_\perp using (6.23). We also obtain α_\perp and study finite-field effects and screening factors by applying a finite electric field E_{out} via a sawtooth potential, and computing the total induced dipole moment per unit length p_\perp . In the linear regime the two approaches are equivalent, and we find an agreement between the two methods within 1%. Fig. 6-7 illustrates both the benefits of using the electrostatic correction for large L and the limitation of its applicability when L is small. The screening ratio of a nanotube is correctly calculated as

$$\frac{E_{\text{appl}}}{E_{\text{in}}} = \frac{E^0 - E_P}{E_{\text{in}}}. \quad (6.24)$$

and the results for a (10,10) nanotube are shown in Fig. 6-8.

The above expressions are valid only for symmetric nanotubes where $\mathbf{p} \parallel \mathbf{E}^0$. In

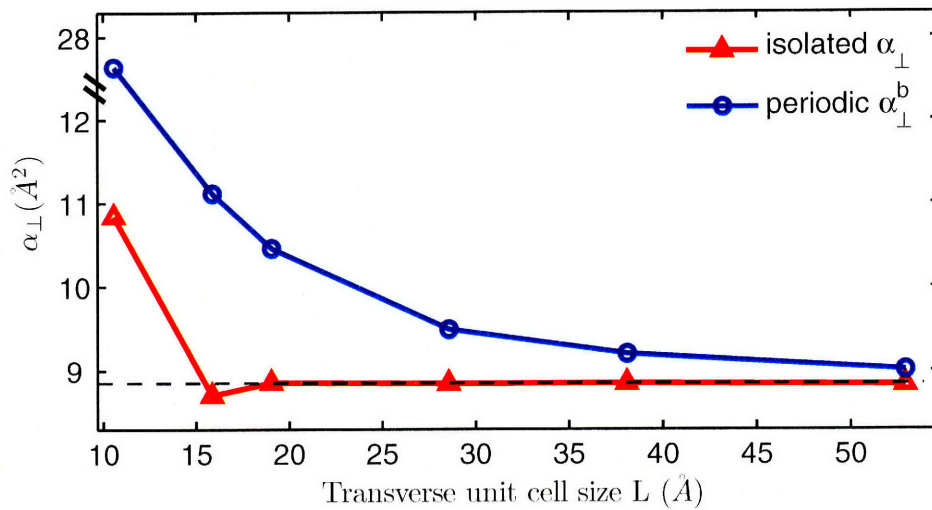


Figure 6-7: Convergence of α_{\perp} and α_{\perp}^b with respect to L for a (5,5) SWNT. The point at L=10.6Å corresponds to a typical tube-tube separation in a bundle.

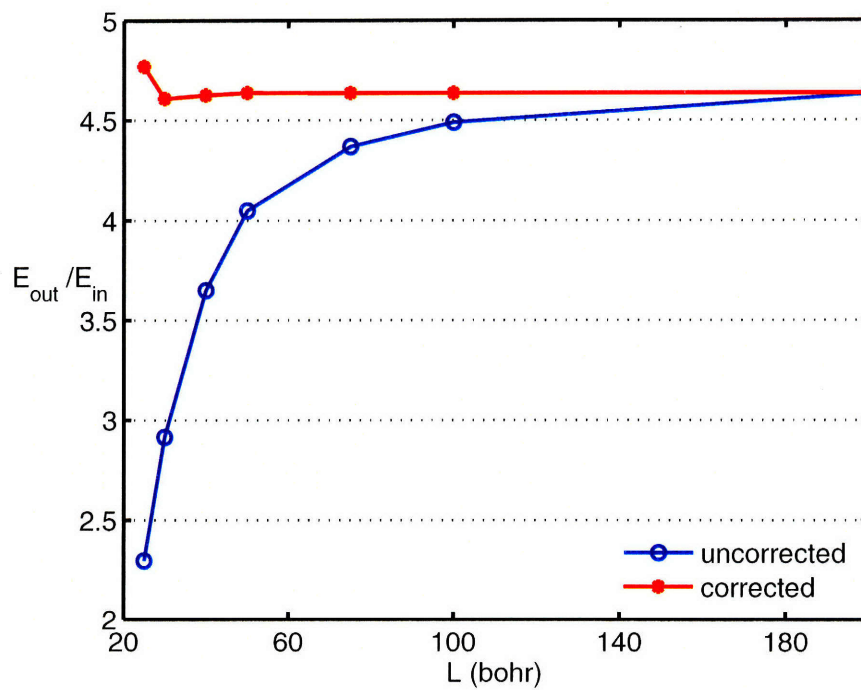


Figure 6-8: Convergence of the screening factor with respect to L for a (10,10) SWNT.

more general situations we need the full 3D polarizability tensor $\overleftrightarrow{\alpha}$, for which the complete correction expression is

$$\overleftrightarrow{\alpha}^{1D} = \overleftrightarrow{\alpha}^b \left(1 + \frac{2\pi}{\Omega} \overleftrightarrow{\alpha}^{tb} \right)^{-1} \quad (6.25)$$

where

$$\overleftrightarrow{\alpha}^b = \begin{pmatrix} \alpha_{xx}^b & \alpha_{xy}^b & \alpha_{xz}^b \\ \alpha_{yx}^b & \alpha_{yy}^b & \alpha_{yz}^b \\ \alpha_{zx}^b & \alpha_{zy}^b & \alpha_{zz}^b \end{pmatrix} \quad (6.26)$$

and

$$\overleftrightarrow{\alpha}^{tb} = \begin{pmatrix} \alpha_{xx}^b & \alpha_{xy}^b & \alpha_{xz}^b \\ \alpha_{yx}^b & \alpha_{yy}^b & \alpha_{yz}^b \\ 0 & 0 & 0 \end{pmatrix} \quad (6.27)$$

The last row of the tensor $\overleftrightarrow{\alpha}^{tb}$ has zeros because the compensating field in PBCs does not have a z-component, i.e. there is no depolarization along the nanotube axis. The longitudinal polarizability per unit length α_{\parallel} is related trivially to the separation-dependent bulk dielectric constant ϵ_{\parallel} via the relation

$$\epsilon_{\parallel} = 1 + \frac{4\pi}{\Omega} \alpha_{\parallel} \quad (6.28)$$

where $\Omega = L^2$ is the cross-sectional area of the unit cell. These issues are of course specific to extended 1D systems, and for other dimensionalities the expressions (6.25)-(6.28) would change accordingly.

Similar correction formulas like (6.25) can be analogously derived for various other linear-response quantities, such as Born effective charges. We note that this correction can be used as a post-processing tool, in the spirit of the Clausius-Mossotti picture, to correct only the potential in a mean-field way. For better accuracy, the potential can be updated using the total charge and dipole values in each self-consistent iteration so that the charge density converges to the correct isolated-molecule limit. However, because the effect of periodic boundaries within this PCC scheme is only to renormalize the magnitude of the applied electric field, the extensions of post-

processing corrections to self-consistent procedures are irrelevant in calculations of dielectric linear-response quantities, such as polarizabilities and susceptibilities. On the other hand, self-consistent corrections may be beneficial for computing non-linear response, such as hyper-polarizabilities.

6.2.2 Depolarization shape effects in finite samples

We briefly discuss here a possible future direction in analyzing calculations of polarized solids in periodic boundary conditions. In many cases the planewave basis is preferred because it lets us deal with extended systems without having to worry about edge effects. However, in certain situations neglecting edge effects of finite samples can lead to unphysical results. When a uniform polarization \mathbf{P} is present in a finite sample, the surface charge on the boundary creates a uniform depolarizing field \mathbf{E}^{dep} which acts to reduce the polarization. The components of this field are determined by so-called depolarizing coefficients n_α ($\alpha = x, y, z$) so that

$$E_\alpha^{dep} = -4\pi n_\alpha P_\alpha \quad (6.29)$$

The depolarizing coefficients depend on the shape of the sample but satisfy $n_x + n_y + n_z = 1$ for a finite sample of any shape [79]. In the isotropic case of a sphere $n_x = n_y = n_z = 1/3$, while for a cylinder $n_x = n_y = 1/2$, $n_z = 0$, and for a slab $n_x = 1$, $n_y = n_z = 0$. We see that there is no depolarization in extended dimensions because the surface charges are infinitely far away. As we saw earlier, the topology of fully periodic boundary conditions is such that the macroscopic field is forced to be zero regardless of the value of the polarization. This corresponds to the case $n_x = n_y = n_z = 0$ which clearly does not correspond to any physical finite sample.

While studying materials in which polarization effects are important, such as piezoelectrics or ferroelectrics, shape effects could be significant [80]. Equilibrium crystal structures and phase transitions in these materials are sensitive to the value of the macroscopic field, and since this field depends on the shape and boundary conditions one must be careful to check for such effects, particularly if non-linear

response terms are important. For instance, it has been observed experimentally that the crystal grain shape has a large influence on the detectable piezoresponse [81].

Another situation in which these issue may have a noticeable dynamic effect in the calculation of infrared Raman spectra, where a time-time correlation function of the polarization values is used to compute the intensity [82]. Let us consider an array of fluctuating dipoles and suppose that in given time frame they all point in the same direction. In a finite spherical sample we saw that depolarization field makes local crystal effects vanish, so that a given dipole does not feel the presence of its neighbors. This is not so in periodic boundary conditions, where a fictitious macroscopic field acts to reinforce the polarization. In this case a given dipole will be encouraged to point along with other dipoles. In other words, the system becomes stiffer, the polarization fluctuations are suppressed which may even affect phase transition temperatures. These effects should depend on the size of the supercell and could be detectable even in classical molecular dynamics simulations.

Such issues can be investigated quite easily within existing DFT implementations – it suffices to add a shape-dependent polarization term to the electric enthalpy functional in the Berry phase formulation [49, 83]. Instead of the conventional functional $E^{\mathcal{E}} = E^0 - \mathcal{E} \cdot \mathbf{P}$ we need to minimize the shape-dependent functional

$$E^{\mathcal{E}} = E^0 - \mathcal{E} \cdot \mathbf{P} + \sum_{\alpha} n_{\alpha} P_{\alpha}^2 \tag{6.30}$$

where E^0 is the usual unmodified energy functional at zero applied field $\mathcal{E} = 0$.

6.3 Inhomogeneous Systems

In the previous section we were dealing with a way to calculate properties of homogeneous systems of reduced dimensions. But oftentimes a low-dimensional system has features that are local, for instance a nanotube functionalized with chemical groups, or a surface of a electrode with molecules attached. One could average the total charge density in the extended dimension and solve the 2D Poisson problem as before, as-

suming a homogeneous system, but this is an approximation that may not always be good. One would expect that if a density fluctuation is very slowly varying its effects may persist at long distances, while short-wavelength variations are effectively averaged out at shorter distances. To treat such a problem we can decompose an arbitrary longitudinal fluctuation into fourier components and solve them separately in the reduced transverse coordinate space. Let us consider a line of charge with a sinusoidal fluctuation

$$\rho(x, y, z) = n_q \delta(x, y) e^{iqz} \quad (6.31)$$

In the $x - y$ plane we now need to solve a modified equation

$$(\partial_x^2 + \partial_y^2 - q^2) \phi(x, y) = -4\pi n_q \delta(x, y) \quad (6.32)$$

which reduces to the Poisson equation for the homogeneous case $q = 0$. This corresponds to the 2D electrostatic problem for a short-range “screened” Coulomb potential, i.e. as if photon acquired a mass q , in analogy to the 3D Yukawa interaction. When written in polar coordinates this is nothing but the modified Bessel equation of order 0:

$$(r^2 \partial_r^2 + r \partial_r - q^2 r^2) \phi(r) = 0 \quad (6.33)$$

for $r \neq 0$. The appropriate Green’s function solution that vanishes at infinity is the modified Bessel function of the second kind $\phi_q^0(\mathbf{r}) = 2K_0(qr)$, which reduces at small r to $-2 \ln(r\gamma_0/2)$ for $q=0$ (γ_0 is the Euler constant), while for large r this function goes like e^{-qr}/\sqrt{r} .

In solving for the contaminating potential $\phi_q^{corr}(\mathbf{r})$ we notice that there is no compensating jellium contribution because the $\mathbf{G} = 0$ fourier component of the potential is well behaved for finite q . Therefore the only contribution is from the image potentials

$$\phi_q^{corr}(\mathbf{r}) = \phi_q^{img}(\mathbf{r}) = \sum_{\mathbf{R} \neq 0} \phi_q^0(\mathbf{r} + \mathbf{R}) \quad (6.34)$$

We can repeat the procedure used for homogeneous charges and expand the image contribution potential near the origin (6.4). All odd-power terms are again zero by

$x \leftrightarrow -x$ and $y \leftrightarrow -y$ symmetries of the infinite lattice (a D_4 symmetric limit taken as before). Likewise, $x \leftrightarrow y$ symmetry implies $\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial^2 \phi}{\partial y^2}$. The Poisson equation is now modified so that

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial^2 \phi}{\partial y^2} = \frac{1}{2} q^2 \phi \quad (6.35)$$

and we can plug in these relations order by order into the Taylor expansion of the contaminating potential:

$$\phi_q^{corr}(r) = \phi_q^{corr}(0) \left(1 + \frac{1}{4} q^2 r^2 + \dots \right) \quad (6.36)$$

where

$$\phi_q^{corr}(0) = n_q \beta(qL) = 2n_q \sum'_{i,j} K_0 \left(qL \sqrt{i^2 + j^2} \right) \quad (6.37)$$

contains a sum that excludes the origin. Note that since the potential satisfying the finite- q differential equation cannot, unlike the Poisson solutions, be rescaled by rescaling the coordinates. That means that there is no universal Madelung constant and the lattice sum has to be performed for each q and L . On the other hand, this sum converges rapidly because the potentials fall off faster than exponentials. Now if we are calculating the longitudinal charge density response function $\chi^{1D}(q)$ to a potential $V(z) = V_0 \cos(qz)$ we need to correct the actual *applied* potential as follows:

$$\rho_q^0 = \chi_0(q) V_0 = \chi^{1D}(q) (V_0 - \rho_0 \beta(qL)) \quad (6.38)$$

so that the corrected response function of the isolated system is

$$\chi^{1D}(q) = \frac{\chi_0(q)}{1 - \beta(qL) \chi_0 q} \quad (6.39)$$

where χ_0 stands for the uncorrected response of the periodic system.

In calculating the response function $\chi^{2D}(q)$ of 2D systems we are dealing with a similar problem. Periodic boundary conditions imply that we are in fact computing the response not of a single sheet but that of a periodic array of sheets separated by distance L . The issue here is that by setting up a periodic in-plane potential $V_0 \cos(qx)$

we induce a periodic charge density response in each layer, and a particular layer will feel the potential created by the induced charge densities of all other layer on the one-dimensional lattice. To remove this effect we consider the potential of a single 2D layer of charge density periodic in one dimension.

$$\rho(x, y) = \rho_0 \cos(qx) \delta(z) \quad (6.40)$$

After going into Fourier space and solving the Poisson equation by performing a contour integral we end up with

$$\phi_q^0(x, z) = \int dk_y 2\rho_0 \cos(qx) \frac{e^{ik_z z}}{k_z^2 + q^2} = 2\pi \rho_0 \cos(qx) \frac{e^{-qz}}{q} \quad (6.41)$$

so that

$$\phi_q^0(z) = 2\pi \frac{e^{-qz}}{q} \quad (6.42)$$

which can be recognized as the Green's function solution of the equation $(\partial_z^2 - q^2) \phi = \delta(z)$ in one dimension. By summing this contribution over the lattice (with origin omitted) we obtain the contribution to the electrostatic potential due to all other images

$$\phi_q^{corr}(z=0) = 2\pi \rho_0 \frac{1}{q} 2 \sum_{n=1}^{\infty} e^{-nqL} = \frac{4\pi \rho_0}{q (e^{qL} - 1)} \quad (6.43)$$

while in general this potential is easily shown to be exactly given by

$$\phi_q^{corr}(z) = \frac{4\pi \rho_0 \cosh(qz)}{q (e^{qL} - 1)} \quad (6.44)$$

Subtracting off just the first order (constant term at $z=0$) of this contaminating potential contribution self-consistently

$$\rho_0(q) = \chi_0(q) V_0 = \chi^{2D}(q) \left(V_0 - \frac{4\pi \rho_0}{q (e^{qL} - 1)} \right) \quad (6.45)$$

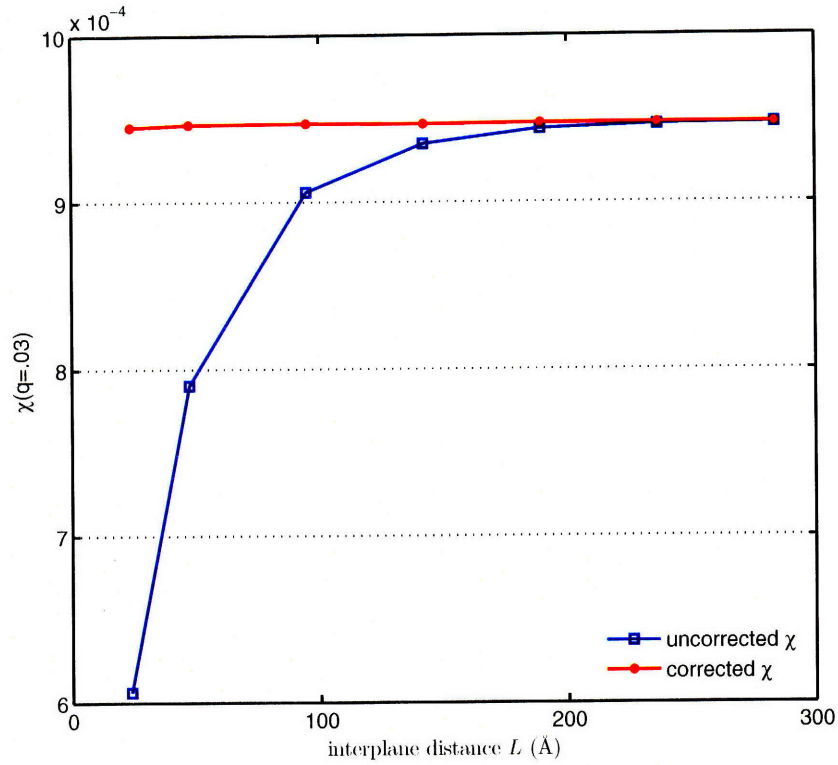


Figure 6-9: Response function versus the periodic inter-plane distance before and after the inhomogeneous electrostatic correction.

we arrive at the corrected response function

$$\chi^{2D}(q) = \frac{\chi_0(q)}{1 - \chi_0(q) \frac{4\pi}{q(e^{qL} - 1)}} \quad (6.46)$$

that is mostly free from the influence of periodic image sheets. Results in Fig. 6-9 demonstrate that this correction works very well and allows calculation of the in-plane q -dependent response function at a small fraction of the cost of the uncorrected calculations, which need enormous amounts of vacuum separation for convergence. At first order, we are only correcting for the density response of wavelength q , even though the density response has a wide range of wavelengths due to the local field effects (see Fig. 4-1). But as equation (6.41) shows, since q is the smallest wavelength, its effect has the farthest influence and dominates the periodic image contamination. To treat arbitrary inhomogeneities, one can use the superposition principle and decom-

pose the charge fluctuations into fourier components q_x and q_y and then sum together the exact solutions obtained from (6.44). One particularly important application of these results would be in calculation of phonon spectra for 1D and 2D systems, where at long wavelengths the acoustic sum rules are violated by the presence of periodic images [43]. These effects should be strongest in polar systems like BN sheet and nanotubes.

6.4 Density countercharge method

In the PCC method of using PBC to perform calculations on isolated low-dimensional systems, we placed the system in the center of a cubic cell and applied approximate analytical corrections to remove the contaminating potentials. This method works very well for well-localized symmetric charge densities, but its application is still difficult for systems that have non-symmetric shapes. Complex compounds with charge densities that are not well approximated by point charges or dipoles (e.g. a protein or a long finite piece of a nanotube) would require large unit cells in order for the PCC correction to be applicable. For such cases it would help to have a more realistic model of the charge density. One proposal is to use gaussian charge distributions to mimic as closely as possible the first few multipole moment of the real charge distribution. The idea is then to solve the Poisson equation with the model charge density in OBC, using the exact analytical solution for a gaussian, in order to find a better approximation of the potential correction $\phi^{corr}(\mathbf{r})$ [76, 78]. The advantage of this correction method is that it is analytical, while the disadvantage is that for a complex system one needs to estimate a large number of parameters (spreads and locations of gaussians) of the model charge density.

As an alternative, we propose using the exact charge density at every step of the self-consistent DFT loop to calculate the corresponding potential of the isolated system. While the precision of this approach is in principle very high, in practical applications the computational cost should be considered. After all, it is the computational efficiency and the favorable scaling of plane-wave calculations that justify using

PBC even for systems with partial periodicity. For instance, computing the OBC potential directly from the exact charge density using coulomb integrals is prohibitively expensive. Fortunately it turns out that the Poisson equation can be solved rather efficiently using multigrid techniques that scale linearly with the number of grid points [84]. Implementation of this method is a collaborative effort with Ismaila Dabo and Nicholas Singh-Miller [85], and shows very promising results for charged systems.

In choosing which correction method to apply, one would need to take into consideration the shape of the charge density and the precision requirements. If a system is well localized and symmetric, the PCC method of Sec.6.1 will give very good results in relatively small unit cells. The precision is controlled by the cell's size with the cost $O(N_e^2 N_{pw})$ increasing linearly with the volume. On the other hand, if a molecule is big and has a complex shape, it may require cell sizes so large with PCC that the number of plane waves at a given energy cutoff exceeds the available computer memory. In these cases the DCC approach is favorable because, despite its numerical complexity, the total cost $O(N_e^2 N_{pw}^2)$ can actually be smaller if the cell size requirement is not as high for a given accuracy threshold. In fact, there is no need for large supercells in the DCC approach because it calculates the isolated potential exactly, as long as the charge density is contained in the unit cell. In other words, it allows PBC calculations with the minimal possible volume of the unit cell just so the system of interest fits in it. Also, unlike PCC, the DCC method does not require the unit cell to have a simple shape.

We can envision some immediately useful applications of the DCC method. First, since the exact charge density is used to compute the potential, there is no additional cost in calculating the boundary conditions of the unit cell using not just the actual density but also its mirror image some distance $2h$ away. We are thus able to calculate the capacitance of a system consisting of an isolated complex molecule coupled to a perfect metallic plate. This offers an alternative and more precise way of determining the Luttinger liquid parameter g of nanotubes, as described in Chap. 5. Furthermore, the fourier-decomposition method for inhomogeneous systems of Sec. 6.3 can be readily implemented in the DCC framework by solving for each wavevector q the

appropriate Bessel equation using a multigrid or the Jacobi algorithm. Finally, we note again that for both homogeneous and inhomogeneous 2D systems the DCC approach is not necessary because the correction potential can be found immediately without numerically solving the Poisson equation.

6.5 Defects in bulk medium

In this section we consider a system consisting of an isolated object embedded in an infinite bulk material, which we will refer to as a defect. It can be a defect in a crystal, an ion in water or a nanotube on an substrate. The way such calculations are currently done is by setting up a supercell with a defect surrounded by some amount of the bulk material, and then by increasing the amount of the bulk until the desired convergence is achieved. Unlike calculations in vacuum, where the cost increases with volume, here we are also adding atoms and electrons of the bulk material, thus making the calculation cost increase much more quickly. At the same time, because of the presence of the periodic bulk material, using PBC becomes a necessity. To improve the computational efficiency in this complex arrangement, we can use the DCC correction scheme described above for isolated systems in vacuum.

As before, if the defect is localized, the long-range effects of the bulk and periodic images are primarily electrostatic. It is useful to imagine the original system as a large periodic sample consisting of supercells with bulk and defects, while the desired configuration contains the defect only in the central cell. The first step in removing the effects of the periodic images of the defect is to separate the charge density of the defect from that of the bulk. One way to do this is to simply subtract the charge density of the pristine bulk from the total charge density in the cell. That way all the effects of the structural bulk rearrangement and hybridizations due to the defect will be included in the defect charge density. We write the total charge density in the cell as

$$\rho_{tot}(\mathbf{r}) = \rho_{def}(\mathbf{r}) + \rho_{bulk}^0(\mathbf{r}) \quad (6.47)$$

where ρ_{bulk}^0 is the charge density of the clean bulk without any defects.

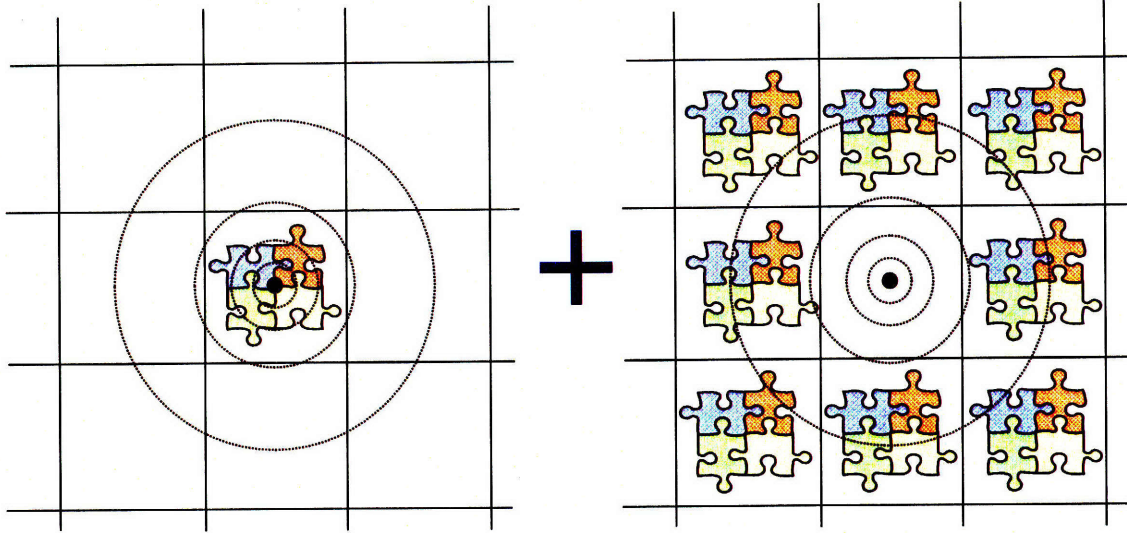


Figure 6-10: Schematic representation of the energy contribution from an isolated defect in bulk. Left: E_{def}^0 . Right: E_{def}^∞ .

Let us look at the defect formation energy, i.e. the difference in the energies (per unit cell) between a pure bulk crystal and the one with a single impurity:

$$\Delta E = E_{def} - E_{bulk}. \quad (6.48)$$

Here E_{bulk} can be computed exactly with PBC, while E_{def} requires careful corrections. The latter can be decomposed into two contributions: (i) the energy of the isolated defect + bulk system only in the central unit cell; (ii) the energy coming from the defect interacting with the remaining infinite bulk. We represent these two contributions on Fig. 6-10 and write them as

$$E_{def} = E_{def}^c + E_{def}^\infty \quad (6.49)$$

The first term contains various short-range quantum interaction effects within the bulk+defect system in addition to mostly electrostatic contaminating potentials of the defect's images. To find this term, we can do a PBC calculation of the supercell with the defect and correct for these image effects by finding the contaminating potential $\phi_{def}^{corr}(\mathbf{r})$ from only the defect images. We can then use it to correct the electrostatic

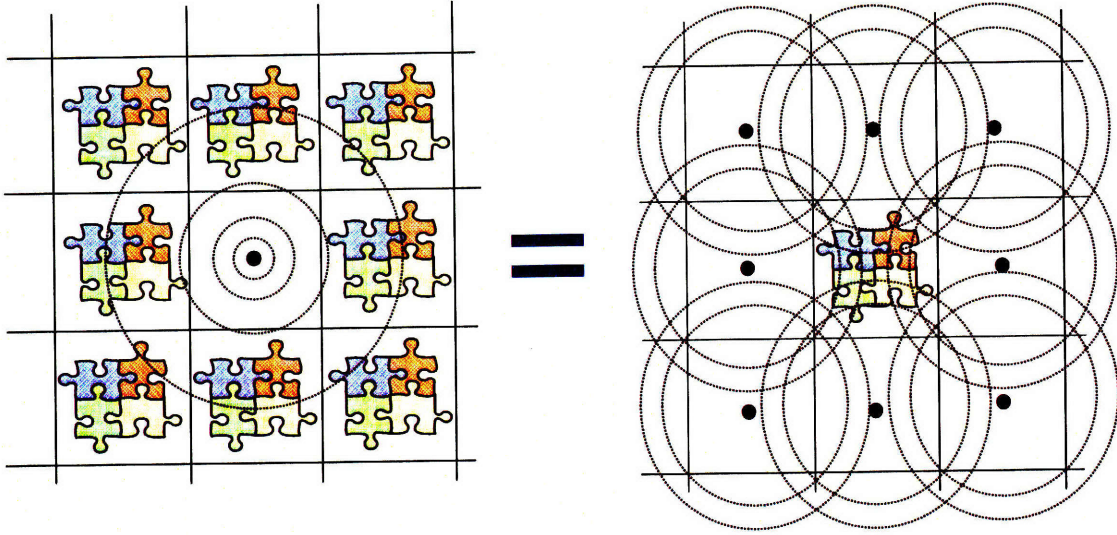


Figure 6-11: The symmetry duality that allows us to calculate E_{def}^∞ .

potential within the cell self-consistently. This way the energy contribution from the central cell E_{def}^c will correspond to a single defect, and at the same time the bulk in this cell will relax to its minimum energy structure in the presence of only one defect. It is much better to use the DCC here, because in PCC the contaminating potential is only correct close to the origin, while we have bulk everywhere in the cell, and also because we cannot always impose cubic symmetry. Incidentally, this first term also contains the energy of the ideal bulk interacting with itself E_{bulk} because we do not remove the bulk's images in the calculation and any bulk deformation is absorbed into a re-definition of defect charge density during the calculation. The result of this self-consistent calculation is

$$E_{def}^c = E_{def}^{PBC} - \frac{1}{2} \int \rho_{tot}(\mathbf{r}) \phi_{def}^{corr}(\mathbf{r}) \quad (6.50)$$

where the integral is over the unit cell.

We can treat the second term of (6.49) only with electrostatics, but even then it seems difficult to calculate because it would require accounting for the influence of the long-range Coulomb potential of the defect on many layers of the bulk surrounding the central cell. Fortunately, we can re-write this contribution using a change of

variables for each lattice vector \mathbf{R} :

$$\begin{aligned} \int \phi_{def}(\mathbf{r}) \sum_{\mathbf{R} \neq 0} \rho_{bulk}^0(\mathbf{r} - \mathbf{R}) d\mathbf{r} &= \sum_{\mathbf{R} \neq 0} \int \phi_{def}(\mathbf{r}) \rho_{bulk}^0(\mathbf{r} - \mathbf{R}) d\mathbf{r} \\ &= \sum_{\mathbf{R} \neq 0} \int \phi_{def}(\mathbf{r}' + \mathbf{R}) \rho_{bulk}^0(\mathbf{r}') d\mathbf{r}' = \int \rho_{bulk}^0(\mathbf{r}') \sum_{\mathbf{R} \neq 0} \phi_{def}(\mathbf{r}' + \mathbf{R}) d\mathbf{r}' \end{aligned} \quad (6.51)$$

where the potential sum in the last term is nothing but $\phi_{def}^{img}(\mathbf{r})$, the potential due to defect images, which we have already computed when doing the correction for term (i). Therefore, there is no need to do any extra work or use supercells. This lattice inversion concept is illustrated in Fig. 6-11. We made an approximation in order to make this trick work: we are in effect neglecting the distance-dependent influence of the defect in the central cell on the electronic and ionic structure of the remaining bulk volume, i.e. the local rearrangement that is different in each layer. This "frozen bulk approximation" is not an uncontrolled one, because we can always improve the accuracy by adding more bulk and using a larger supercell. The convergence with supercell size is expected to be much faster than for the usual supercell extrapolation approach. This calculation is purely electrostatic and is non-selfconsistent, and gives

$$E_{def}^{\infty} \approx \frac{1}{2} \int \rho_{bulk}^0 \phi_{def}^{corr} \quad (6.52)$$

and the integral is again only over the unit cell. Recalling now the definition of the defect charge density (6.47), we can collect the two terms of (6.49) and perform a single calculation:

$$E_{def}^c = E_{def}^{PBC} - \frac{1}{2} \int \rho_{def}(\mathbf{r}) \phi_{def}^{corr}(\mathbf{r}) \quad (6.53)$$

where the defect charge density and the contaminating potential of its images are computed self-consistently.

The implementation of this method is in progress. In addition to speeding up calculations of low-dimensional systems embedded in a medium, this approach can also increase efficiencies of a wide range of other ab-initio methods dealing with extended systems. For instance, the linear response method of calculating the DFT+U

Hubbard parameters in transition metals [86] relies on varying the charge on a single atom embedded in a supercell of the bulk crystal.

Chapter 7

Summary and future directions

In this work we study electronic interactions in low-dimensional systems and their effects on the static dielectric properties, concentrating primarily on graphene and carbon nanotubes, with the corresponding boron-nitride structures used for comparison. In order to understand the rich and often unusual characteristics of these materials we used a variety of tools ranging from simple two-band tight binding models to recently developed DFT-based methods. With their help we develop a complete understanding of the salient features of the static dielectric response of 2D sheets and the corresponding nanotubes. We confirm that at long wavelengths π -electrons dominate the in-plane dielectric response of graphene, and we find a good agreement in this regime with the predictions of the tight-binding approximation. In particular, due to its 2D lattice symmetry, the in-plane charge density response of undoped graphene is found to be scale invariant, intermediate between that of a 2D metal and an insulator. The boron nitride sheet behaves like an insulator, while the graphene sheet at any finite amount of doping exhibits metallic screening. Using the fact that the transverse polarizabilities of both the graphene and the BN sheets are much smaller than the in-plane polarizabilities, we can use a geometric “folding” of the 2D response function to obtain predictions for single-wall nanotubes. This procedure yields remarkably good results for both carbon and boron-nitride SWNTs and provides a way to generalize the model to MWNTs. We find that in our calculations of SWNTs, the longitudinal response is controlled by the band gap, while the trans-

verse response is sensitive only to the effective radii. In bundles and MWNTs, the longitudinal response is additive, while the transverse response in MWNTs is sensitive to the inter-tube screening effects and required a more elaborate analysis. For this purpose we construct a classical electrostatic model, that accurately describes the transverse response of general MWNTs, using the parameters derived from the response of SWNTs. Using our electrostatic models of general CNTs and BNNTs we then investigate how the transverse screening depends on the number of layers and find qualitative differences between the two materials. A possible future direction is to generalize this approach to frequency-dependent response in order to understand the optical properties of MWNTs using those of SWNTs, which are much easier to compute. In fact, it is possible to use our electrostatic model of CNTs to make quantitative statements about the features of the Raman intensities of the radial breathing modes, provided the frequency is low enough for the static model to be accurate.

In one-dimensional metallic systems electronic interactions cause a breakdown of the traditional Fermi-liquid quasiparticle picture. Particle correlations are very strong and manifest themselves with characteristic power laws, such as a power-law singularity in the momentum distribution function and suppression of the density of states near the Fermi energy that can be observed in tunneling transport experiments and photoemission spectroscopy measurements. The Tomonaga-Luttinger model describes the behavior of 1D metals at low energies and determines these power law exponents from microscopic properties of the system, in particular the strength of the long-range interaction and the Fermi velocity. Motivated by recent experiments on carbon nanotubes, we develop a method of using density-functional theory to estimate the charge compressibility and Fermi velocities under a variety of electrostatic environments, and use these results to obtain the Tomonaga-Luttinger model parameters in the charge sector of general 1D metallic systems. Our calculations for carbon nanotubes are in quantitative agreement with current experimental results and previous RPA estimates. We emphasize that previous calculations neglected the exchange-correlation contributions that appear to be significant. There is a current ongoing effort on investigating the dependence of the Luttinger liquid parameters on

the doping level of the nanotube. A more ambitious future direction would be to attempt to find the manifestations of the Luttinger liquid correlations in the charge density directly from DFT calculations, as opposed to only estimating the parameters as we presently do.

All calculations in this work have been done using the plane-wave basis in periodic boundary conditions. It was realized that in order to study isolated low-dimensional systems one must remove the effects of periodic images, and that at long distances these effects are dominated by classical electrostatics. We developed a set of methods to deal with these issues with a variety of levels of accuracy and computational cost, ranging from first-order analytical corrections to fully numerical approaches of solving the Poisson problem in open boundary conditions. Applications of these correction methods in calculations of dielectric response properties in this, such as in-plane q -dependent susceptibilities, transverse polarizabilities, screening factors, and electronic compressibilities, yielded significant reductions of computational cost and improvement of accuracy. These approaches can be very beneficial in first-principles studies of general zero-, one- and two-dimensional structures that contain significant charge asymmetry and/or inhomogeneity and would otherwise require very large unit cells to perform accurate computations with plane-wave based DFT methods. One particular application that is directly relevant to this work is the possibility to easily adapt our density counter-charge (DCC) method to calculations of low-dimensional systems in close proximity to perfect metallic surfaces. An example would be the calculation of the Luttinger liquid parameters of a nanotube on dielectric substrates close to metallic gates. This method also allows calculations of quantum capacitances of general microscopic systems at arbitrary levels of doping. The generalization of the DCC method to bulk system with defects described in Sec. 6.5 has yet to be fully implemented and tested, but it holds promise for calculations of more realistic systems in which an extended medium is treated with the same ab-initio accuracy as the embedded system of interest. It can be applied to study the energetics of crystal defects and charged ions in electrolytes or molecules in solutions, as well as to speed up computations of the Hubbard U parameter in the DFT+ U calculations of bulk

transition metals. It also became clear that in calculations of phonons in polar 1D and 2D systems and q -dependent dielectric susceptibilities, periodic image effects cannot be controlled by simply increasing the size of the unit cell because these effects become increasingly dominant as q approaches 0. We derived analytical corrections that are very beneficial in susceptibility calculations, but further investigation is required to assess the importance of these issues in calculations of phonons in low-dimensional systems.

Appendix A

Electrostatics of dielectric cylinders

In this appendix we describe the electrostatic response of a general dielectric cylinder to a transverse uniform external electric field. The end result is an exact solution for a system consisting of N concentric cylindrical sections with arbitrary dielectric constants. The goal is to use the results to construct an effective electrostatic model of a multi-wall nanotube. We start with the most general expansion of the potential function in cylindrical coordinates:

$$\Phi^{(i)}(\rho, \phi) = \sum_{n=1}^{\infty} a_n^{(i)} \rho^n \cos n\phi + b_n^{(i)} \rho^n \sin n\phi + c_n^{(i)} \rho^{-n} \cos n\phi + d_n^{(i)} \rho^{-n} \sin n\phi + f^{(i)} \ln \rho \quad (\text{A.1})$$

where ϕ is the angle with respect to the external field and (i) labels the concentric regions. Symmetry around the direction of the applied field dictates that $b_n^{(i)} = d_n^{(i)} = 0$. For the innermost region ($\rho < r_1$) we note also that $f^{(1)} = 0$ and $c_n^{(1)} = 0$ due to the divergence of the corresponding terms at the origin.

$$\Phi^{(1)} = \sum_{n=1}^{\infty} a_n^{(1)} \rho^n \cos n\phi \quad (\text{A.2})$$

$$\Phi^{(2)} = \sum_{n=1}^{\infty} a_n^{(2)} \rho^n \cos n\phi + c_n^{(2)} \rho^{-n} \cos n\phi + f^{(2)} \ln \rho \quad (\text{A.3})$$

At the interface between regions 1 and 2 the boundary conditions are $D_{\perp}^{(1)} = D_{\perp}^{(2)}$ and $E_{\parallel}^{(1)} = E_{\parallel}^{(2)}$, which can be written as

$$\begin{aligned} \epsilon_1 \frac{\partial \Phi^{(1)}}{\partial \rho} \Big|_{r_1} &= \epsilon_2 \frac{\partial \Phi^{(2)}}{\partial \rho} \Big|_{r_1} \\ \forall n : \quad \epsilon_1 a_n^{(1)} r_1^{2n} &= \epsilon_2 a_n^{(2)} r_1^{2n} - \epsilon_2 c_n^{(2)} \end{aligned} \quad (\text{A.4})$$

and

$$\begin{aligned} \frac{1}{\rho} \frac{\partial \Phi^{(1)}}{\partial \phi} \Big|_{r_1} &= \frac{1}{\rho} \frac{\partial \Phi^{(2)}}{\partial \phi} \Big|_{r_1} \\ \forall n : \quad a_n^{(1)} r_1^{2n} &= a_n^{(2)} r_1^{2n} + c_n^{(2)} \end{aligned} \quad (\text{A.5})$$

The coefficient of the logarithmic term $f^{(2)} = f^{(1)} = 0$ and likewise $f^{(i)} = 0$ for all i . Similarly, after straightforward algebra, for the boundary between regions i and $i+1 < N$ we have

$$\begin{aligned} \forall n : \quad \epsilon_i (a_n^{(i)} r_i^{2n} - c_n^{(i)}) &= \epsilon_{i+1} (a_n^{(i+1)} r_i^{2n} - c_n^{(i+1)}) \\ a_n^{(i)} r_i^{2n} + c_n^{(i)} &= a_n^{(i+1)} r_i^{2n} + c_n^{(i+1)} \end{aligned} \quad (\text{A.6})$$

Finally, in the outermost region the additional constraint is that $\Phi \rightarrow -E_0 \rho \cos \phi$ at $\rho \rightarrow \infty$, so

$$\Phi^{(N)} = -E_0 \rho \cos \phi + \sum_{n=1}^{\infty} c_n^{(N)} \rho^{-n} \cos n\phi \quad (\text{A.7})$$

and the boundary condition equations for the outer boundary are

$$\begin{aligned} \forall n : \quad \epsilon_{N-1} (a_n^{(N-1)} r_{N-1}^{2n} - c_n^{(N-1)}) &= \epsilon_N (-E_0 r_{N-1}^{n+1} \delta_{n,1} - c_n^{(N)}) \\ a_n^{(N-1)} r_{N-1}^{2n} + c_n^{(N-1)} &= -E_0 r_{N-1}^{n+1} \delta_{n,1} + c_n^{(N)} \end{aligned} \quad (\text{A.8})$$

For each value of n we can collect all boundary conditions as a linear system of equations. As an example we present the result for the case of $N = 4$. This would

The angular integral can be evaluated explicitly

$$\int_0^{2\pi} d\phi \cos \phi \ln \sqrt{r'^2 + R^2 - 2r'R \cos \phi} = -\pi \frac{r'}{R} \quad (\text{A.13})$$

So the resulting relation is

$$\int_{r < R} \mathbf{E}(\mathbf{x}) d^2x = -\frac{R}{2\pi\epsilon_0} \int d^2x' \frac{r'}{R} \pi \rho(\mathbf{x}') \hat{\mathbf{n}}' = -2\pi \mathbf{p}_\perp \quad (\text{A.14})$$

where \mathbf{p} is the total electric dipole moment of the charge distribution with respect to the center of the circle. The volume integral is independent of the radius of the circle as long as all of charge density is contained inside.

The potential outside the cylinder ($\rho > r_N$) is

$$\Phi^{(N)} = -E_0 \rho \cos \phi + c_1^{(N)} \rho^{-1} \cos \phi \quad (\text{A.15})$$

To calculate the total induced dipole, we use equation (A.14)

$$-2\pi \mathbf{p}_\perp = \int_{r < R} \mathbf{E}(\mathbf{x}) d^2x = - \int_{r < R} \nabla \Phi d^2x = - \oint_{r=R} R \Phi(\mathbf{x}) \hat{\mathbf{n}} d\phi \quad (\text{A.16})$$

where $R > r_N$ is an arbitrary radius. Plugging in the expression for $\Phi^{(N)}$ we get

$$\mathbf{p}_\perp = \frac{R}{2\pi} \int_0^{2\pi} \frac{c_1^{(N)}}{R} \cos \phi (\hat{\mathbf{x}} \cos \phi + \hat{\mathbf{y}} \sin \phi) d\phi = \frac{c_1^{(N)}}{2} \hat{\mathbf{x}} \quad (\text{A.17})$$

which implies that

$$\alpha_\perp = c_1^{(N)} / 2E_0 \quad (\text{A.18})$$

This procedure for computing the screening factor (A.10) and polarizability (A.18) is implemented in Matlab in Sec. B.

Appendix B

Matlab code for MWNT models

```
% This file will let you calculate the transverse polarizability and
% screening properties of an arbitrary multi-wall nanotube or - with
% minor modifications - any concentric set of dielectric cylinders.
% Copyright Boris Kozinsky and Nicola Marzari 2006

% INSTRUCTIONS: Modify the values of parameters as needed and run the
% script in Matlab.

r12=8.142;
r16=10.856;
r18=12.213; % Ro of (18,18) CNTs

dR=1.30; % effective radius offset for a SWNT:  $R=R_0+s$ ,
% where  $R_0$  is the backbone radius

delta=0.0001; % infinitesimal thickness of a SWNT layer in a MWNT
c=0.396; % response parameter of a SWNT, from ab-initio
tubes=2; % total number of SWNTs in the MWNT
Rmax=r16; % radius of the outer tube
spacing = 4.071; % spacing (radius difference in Å) between tubes
```

```

%----- Boundary conditions equation solver -----
F=4*c/(1-2*c);

N=2*tubes; r=zeros(N,1); e=zeros(N,1);

for j=1:tubes
    r(2*j-1)=Rmax-(tubes-j)*spacing+dR;
    r(2*j)=r(2*j-1)+delta;
    e(2*j-1)=1;           % vacuum region epsilon=1
    e(2*j)=r(2*j)/delta*F; % scale-invariant epsilon of a SWNT
end

M=zeros(2*N,2*N);
M(1,1)=r(1)^2;
M(1,2)=-r(1)^2;
M(1,3)=-1;
M(2,1)=e(1)*r(1)^2;
M(2,2)=-e(2)*r(1)^2;
M(2,3)=e(2);
M(2*N-1,2*N-2)=r(N)^2;
M(2*N-1,2*N-1)=1;
M(2*N-1,2*N)=-1;
M(2*N,2*N-2)=e(N)*r(N)^2;
M(2*N,2*N-1)=-e(N);
M(2*N,2*N)=1;

for i=2:(N-1)
    M(2*i-1,2*i-2)=r(i)^2;
    M(2*i-1,2*i-1)=1;
    M(2*i-1,2*i)=-r(i)^2;

```

```

M(2*i-1,2*i+1)=-1;
M(2*i,2*i-2)=e(i)*r(i)^2;
M(2*i,2*i-1)=-e(i);
M(2*i,2*i)=-e(i+1)*r(i)^2;
M(2*i,2*i+1)=e(i+1);
end

V=zeros(2*N,1);
V(2*N-1)=-r(N)^2;
V(2*N)=-r(N)^2;
K=inv(M)*V;
%-----

polarizability =.5*K(2*N);
% transverse polarizability per unit length (in A^2) of the system.

screening=-1/K(1);
% the ratio of E_0/E_in, where E_0 is the applied field and E_in is
% the uniform screened field at the center

```


Appendix C

Polarizability tensor of a metallic ellipsoid

The metallicity of an armchair carbon nanotube is due to delocalized π -electrons which are mostly smeared out over the nanotube's surface. It is therefore meaningful to model the CNT as a finite thin metallic tube for the purposes of studying static dielectric response. A more convenient approximation is that of a long metallic ellipsoid [79], and calculating the polarizability tensor is a standard boundary value problem of electromagnetism. We present here the results of the calculation assuming a uniform electric field with magnitude E_0 far away. Let's assume a rotationally symmetric ellipsoid with semimajor axis a and semiminor axes b , with eccentricity given by $e = \sqrt{1 - \frac{b^2}{a^2}}$. Imposing boundary conditions of constant potential on the surface and of vanishing response field at infinity, we find that for a given applied field E_0 , the magnitude of the total electric field is maximal at the tip:

$$\frac{E_{max}}{E_0} = \frac{1}{n^{(z)}} \quad (\text{C.1})$$

where

$$\begin{aligned} n^{(z)} &= \frac{1 - e^2}{2e^3} \left(\ln \frac{1 + e}{1 - e} - 2e \right) \\ n^{(x)} = n^{(y)} &= \frac{1}{2} (1 - n^{(z)}) \end{aligned} \quad (\text{C.2})$$

The principle values of the polarizability tensor, defined by $\mathbf{P} = \overleftrightarrow{\alpha} \mathbf{E}$, are

$$\alpha^{(x)} = \frac{ab^2}{3n^{(x)}}, \quad \alpha^{(y)} = \frac{ab^2}{3n^{(y)}}, \quad \alpha^{(z)} = \frac{ab^2}{3n^{(z)}} \quad (\text{C.3})$$

We can now take the limit of a very long ellipsoid ($a \gg b$). There are two ways to take this limit. First, let's fix the transverse semimajor axis b .

$$n^{(z)} \rightarrow \frac{b^2}{a^2} \left(\ln \frac{2a}{b} - 1 \right) \quad (\text{C.4})$$

$$\alpha^{(z)} \rightarrow \frac{a^3}{3} \left(\ln \frac{2a}{b} - 1 \right)^{-1} \quad (\text{C.5})$$

In the extreme case of $a \rightarrow \infty$ at fixed b we get $\alpha^{(z)} \sim a^3 / \ln a \rightarrow \infty$ and $\alpha^{(x)} = \alpha^{(y)} \sim 2ab^2/3 \rightarrow \infty$ with transverse polarizability per unit length approaching a constant value of $2b^2/3$. Alternatively, in the case of $b \rightarrow 0$ at fixed a we have $\alpha^{(z)} \sim -a^3 / \ln b \rightarrow 0$ and $\alpha^{(x)} = \alpha^{(y)} \sim 2ab^2/3 \rightarrow 0$. We note that in both limits the transverse polarizability $\alpha^{(x)}$ is proportional to the total volume of the ellipsoid. The answer clearly depends on how one takes the limit of a long and thin nanotube, and the first limit is the physically correct one. Incidentally, we note that transverse polarizability of a metallic *cylinder* of length l and radius R is $\alpha_{\perp} = \frac{1}{2}lR^2$, as can be readily seen by using translational invariance. At the same time, the prefactor for a long ellipsoid is $1/3$, while the size-dependence is the same. This shape effect does not alter the scaling of polarizability with dimensions and modifies only the coefficients.

Bibliography

- [1] L. V. Radushkevich and V. M. Lukyanovich. O strukture ugleroda, obrazujučegosja pri termičeskom razložhenii okisi ugleroda na zheleznom kontakte. *Zurn. Fisic. Chim.*, 26:88–95, 1952.
- [2] S. Iijima. Helical microtubules of graphitic carbon. *Nature*, 354:56, 1991.
- [3] R. Saito, G. Dresselhaus, and M. S. Dresselhaus. *Physical Properties of Carbon Nanotubes*. Imperial College Press, Loudon, 1998.
- [4] S. Reich, C. Thomsen, and J. Maultzsch. *Carbon nanotubes : basic concepts and physical properties*. Wiley-VCH, Weinheim, 2004.
- [5] Yuegang Zhang, Aileen Chang, Jien Cao, Qian Wang, Woong Kim, Yiming Li, Nathan Morris, Erhan Yenilmez, Jing Kong, and Hongjie Dai. Electric-field-directed growth of aligned single-walled carbon nanotubes. *Applied Physics Letters*, 79(19):3155–3157, 2001.
- [6] Ant Ural, Yiming Li, and Hongjie Dai. Electric-field-aligned growth of single-walled carbon nanotubes on surfaces. *Applied Physics Letters*, 81(18):3464–3466, 2002.
- [7] Jingqi Li, Qing Zhang, Ning Peng, and Qi Zhu. Manipulation of carbon nanotubes using ac dielectrophoresis. *Applied Physics Letters*, 86(15):153116, 2005.
- [8] R. Krupke *et al.* *Science*, 301:344, 2003.
- [9] Lorin X. Benedict, Steven G. Louie, and Marvin L. Cohen. Static polarizabilities of single-wall carbon nanotubes. *Phys. Rev. B*, 52(11):8541–8549, Sep 1995.

- [10] D. S. Novikov and L. S. Levitov. Energy anomaly and polarizability of carbon nanotubes. *Physical Review Letters*, 96(3):036402, 2006.
- [11] U. Ravaioli Y. Li, S. V. Rotkin. *Nano Lett.*, 3:183, 2003.
- [12] G. Y. Guo, K. C. Chu, Ding sheng Wang, and Chun gang Duan. Linear and non-linear optical properties of carbon nanotubes from first-principles calculations. *Physical Review B (Condensed Matter and Materials Physics)*, 69(20):205416, 2004.
- [13] Edward N. Brothers, Konstantin N. Kudin, Gustavo E. Scuseria, and Jr. Charles W. Bauschlicher. Transverse polarizabilities of carbon nanotubes: A hartree-fock and density functional study. *Physical Review B (Condensed Matter and Materials Physics)*, 72(3):033402, 2005.
- [14] Angel Rubio, Jennifer L. Corkill, and Marvin L. Cohen. Theory of graphitic boron nitride nanotubes. *Phys. Rev. B*, 49(7):5081–5084, Feb 1994.
- [15] Nasreen G. Chopra, R. J. Luyken, K. Cherrey, Vincent H. Crespi, Marvin L. Cohen, Steven G. Louie, and A. Zettl. Boron Nitride Nanotubes. *Science*, 269(5226):966–967, 1995.
- [16] L. D. Landau. Zur theorie der phasenumwandlungen ii. *Phys. Z. Sowjetunion*, 11:26–35, 1937.
- [17] R. E. Peierls. Quelques proprietes typiques des corpses solides. *Ann. I. H. Poincare*, 5:177–222, 1935.
- [18] N. D. Mermin. Crystalline order in two dimensions. *Phys. Rev.*, 176:250–254, 1968.
- [19] K. S. et al. Novoselov. Electric field effect in atomically thin carbon films. *Science*, 306:666–669, 2004.
- [20] K. S. et al. Novoselov. Two-dimensional atomic crystals. *Proc. Natl Acad. Sci. USA*, 102:10451–10453, 2005.

- [21] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64(4):1045–1097, Oct 1992.
- [22] N. H. March. *Theory of the inhomogeneous electron gas*. Plenum, NY, 1983.
- [23] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864, 1964.
- [24] R. G. Parr and W. Yang. *Density Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1989.
- [25] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev. A*, 140:A1133, 1965.
- [26] J. F. Janak. Proof that $\frac{\partial \epsilon}{\partial n_i} = \epsilon_i$ in density-functional theory. *Phys. Rev. B*, 18(12):7165–7168, Dec 1978.
- [27] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865, 1996.
- [28] Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13(12):5188–5192, Jun 1976.
- [29] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving pseudopotentials. *Phys. Rev. Lett.*, 43(20):1494–1497, Nov 1979.
- [30] David Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B*, 41(11):7892–7895, Apr 1990.
- [31] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.*, 73:515, 2001.
- [32] J. C. Slonczewski and P. R. Weiss. Band structure of graphite. *Phys. Rev.*, 109(2):272–279, Jan 1958.

- [33] O. Dubay and G. Kresse. Accurate density functional calculations for the phonon dispersion relations of graphite layer and carbon nanotubes. *Phys. Rev. B*, 67(3):035401, Jan 2003.
- [34] G. F. Giuliani and G. Vignale. *Quantum theory of electron liquid*. Cambridge University Press, 2005.
- [35] Nathan Wiser. Dielectric constant with local field effects included. *Phys. Rev.*, 129(1):62–69, Jan 1963.
- [36] J. D. Jackson. *Classical Electrodynamics*. Wiley, 1999.
- [37] David R. Penn. Wave-number-dependent dielectric function of semiconductors. *Phys. Rev.*, 128(5):2093–2097, Dec 1962.
- [38] B. Koziusky and N. Marzari. In preparation.
- [39] J. Gonzalez *et al.* *Nucl. Phys. B*, 424:595, 1994.
- [40] T. Ando. *Journal of the Physical Society of Japan*, 75:074716, 2006.
- [41] D. P. DiVincenzo and E. J. Mele. Self-consistent effective-mass theory for intralayer screening in graphite intercalation compounds. *Phys. Rev. B*, 29(4):1685–1694, Feb 1984.
- [42] S. Baroni, A. Dal Corso, S. de Gironcoli, P. Giannozzi, C. Cavazzoni, G. Ballabio, S. Scandolo, G. Chiarotti, P. Focher, A. Pasquarello, K. Laasonen, A. Trave, R. Car, N. Marzari, and A. Kokalj. <http://www.pwscf.org>; <http://www.quantum-espresso.org>.
- [43] Nicolas Mouet and Nicola Marzari. First-principles determination of the structural, vibrational and thermodynamic properties of diamond, graphite, and derivatives. *Physical Review B (Condensed Matter and Materials Physics)*, 71(20):205214, 2005.

- [44] Nicola Marzari and David Vanderbilt. Maximally localized generalized wannier functions for composite energy bands. *Physical Review B (Condensed Matter)*, 56(20):12847–12865, 1997.
- [45] Young-Su Lee, Marco Buongiorno Nardelli, and Nicola Marzari. Band structure and quantum conductance of nanostructures from maximally localized wannier functions: The case of functionalized carbon nanotubes. *Physical Review Letters*, 95(7):076804, 2005.
- [46] A. Mostofi et al. <http://www.wannier.org/>.
- [47] David Vanderbilt and R. D. King-Smith. Electric polarization as a bulk quantity and its relation to surface charge. *Phys. Rev. B*, 48(7):4442–4455, Aug 1993.
- [48] Boris Kozinsky and Nicola Marzari. Static dielectric properties of carbon nanotubes from first principles. *Physical Review Letters*, 96(16):166801, 2006.
- [49] P. Umari and Alfredo Pasquarello. Ab initio molecular dynamics in a finite homogeneous electric field. *Phys. Rev. Lett.*, 89(15):157602, Sep 2002.
- [50] R. Geick, C. H. Perry, and G. Rupprecht. Normal modes in hexagonal boron nitride. *Phys. Rev.*, 146(2):543–547, Jun 1966.
- [51] Ludger Wirtz, Michele Lazzeri, Francesco Mauri, and Angel Rubio. Raman spectra of bn nanotubes: Ab initio and bond-polarizability model calculations. *Physical Review B (Condensed Matter and Materials Physics)*, 71(24):241402, 2005.
- [52] M. F. Lin. Optical spectra of single-wall carbon nanotube bundles. *Phys. Rev. B*, 62(19):13153–13159. Nov 2000.
- [53] S. Agnihotri et al. *Langmuir*, 21:896, 2005.
- [54] E. Joselevich and C.M. Lieber. Vectorial growth of metallic and semiconducting single-wall carbon nanotubes. *Nano Letters*, 2(10):1137–1141, 2002.

- [55] I.E. Dzyaloshinskii A.A. Abrikosov, L.P. Gorkov. *Methods of Quantum Field Theory in Statistical Physics*. Dover, New York, 1963.
- [56] S. Tomonaga. *Prog. Theor. Phys. (Kyoto)*, 5:544, 1950.
- [57] J.M. Luttinger. *J. Math. Phys.*, 4:1154, 1963.
- [58] F. D. M. Haldane. ‘luttinger liquid theory’ of one-dimensional quantum fluids. i. properties of the luttinger model and their extension to the general 1d interacting spinless fermi gas. *J. Phys. C*, 14:2585, 1981.
- [59] J. Voit. One-dimensional fermi liquids. *Rep. Prog. Phys.*, 58:977, 1995.
- [60] X. G. Wen. Electrodynamical properties of gapless edge excitations in the fractional quantum hall states. *Phys. Rev. Lett.*, 64(18):2206–2209, Apr 1990.
- [61] A. M. Chang, L. N. Pfeiffer, and K. W. West. Observation of chiral luttinger behavior in electron tunneling into fractional quantum hall edges. *Phys. Rev. Lett.*, 77(12):2538–2541, Sep 1996.
- [62] A. Yacoby, H. L. Stormer, Ned S. Wingreen, L. N. Pfeiffer, K. W. Baldwin, and K. W. West. Nonuniversal conductance quantization in quantum wires. *Phys. Rev. Lett.*, 77(22):4612–4615, Nov 1996.
- [63] Fujimori A. Aomura S. Sawa H. Sekiyama, A. and R. Kato. *Phys. Rev. B*, 51:1389913902, 1995.
- [64] L. Balents C. Kane and M. P. A. Fisher. Coulomb interactions and mesoscopic effects in carbon nanotubes. *Phys. Rev. Lett.*, 79:5086, 1997.
- [65] R. Egger and A. O. Gogolin. Effective low-energy theory for correlated carbon nanotubes. *Phys. Rev. Lett.*, 79:5082, 1997.
- [66] Postma H.W. C. Balents L. Yao, Z. and C. Dekker. *Nature*, 402:273276, 1999.

- [67] J. Lu A. G. Rinzler R. Smalley L. Balents M. Bockrath, D. H. Cobden and P. L. Mccuen. Luttinger-liquid behaviour in carbon nanotubes. *Nature*, 397(598), 1999.
- [68] Hiroyoshi Ishii et al. Direct observation of tomonagaluttinger-liquid state in carbon nanotubes at low temperatures. *Nature*, 426:540–544, 2003.
- [69] Jhinhwan Lee, S. Eggert, H. Kim, S.-J. Kahng, H. Shinohara, and Y. Kuk. Real space imaging of one-dimensional standing waves: Direct evidence for a luttinger liquid. *Physical Review Letters*, 93(16):166403, 2004.
- [70] Michael M. Fogler, Sergey V. Malinin, and Thomas Nattermann. Coulomb blockade and transport in a chain of one-dimensional quantum dots. *Physical Review Letters*, 97(9):096601, 2006.
- [71] T. Giamarchi. *Quantum Physics in One Dimension*. Oxford University Press, New York, 2004.
- [72] Jan von Delft and Herbert Schoeller. Bosonization for beginners refermionization for experts. *Ann. Phys. (Leipzig)*, 4:225–306, 1998.
- [73] C. L. Kane and M. P. A. Fisher. Transmission through barriers and resonant tunneling in an interacting one-dimensional electron gas. *Phys. Rev. B*, 46:15233, 1992.
- [74] R. Egger and A. O. Gogolin. *Eur. Phys. J. B*, 3(1998):281300.
- [75] Niels Gronbech-Jensen. *Computer Physics Communications*, 119:115–121, 1999.
- [76] P. E. Blochl. Electrostatic decoupling of periodic images of plane-wave-expanded densities and derived atomic point charges. *The Journal of Chemical Physics*, 103(17):7422–7428, 1995.
- [77] G. Makov and M. C. Payne. Periodic boundary conditions in ab initio calculations. *Physical Review B (Condensed Matter)*, 51(7):4014–4022, 1995.

- [78] Peter A. Schultz. Local electrostatic moments and periodic boundary conditions. *Physical Review B (Condensed Matter and Materials Physics)*, 60(3):1551-1554, 1999.
- [79] L. D. Landau and E. M. Lifshitz. *Electrodynamics of Continuous Media*. Pergamon, 1984.
- [80] R. Lakes. Shape-dependent damping in piezoelectric solids. *Sonics and Ultrasonics, IEEE Transactions on*, 27(4):208-213, 1980.
- [81] F. Peter, A. Rudiger, R. Dittmann, R. Waser, K. Szot, B. Reichenberg, and K. Prume. Analysis of shape effects on the piezoresponse in ferroelectric nanograins with and without adsorbates. *Applied Physics Letters*, 87(8):082901, 2005.
- [82] M. Bernasconi, P. L. Silvestrelli, and M. Parrinello. Ab initio infrared absorption study of the hydrogen-bond symmetrization in ice. *Phys. Rev. Lett.*, 81(6):1235-1238, Aug 1998.
- [83] Ivo Souza, Jorge Íñiguez, and David Vanderbilt. First-principles approach to insulators in finite electric fields. *Phys. Rev. Lett.*, 89(11):117602, Aug 2002.
- [84] A. Schuller U. Trottenberg, C. W. Oosterlee. *Multigrid*. Academic Press, London, 2001.
- [85] N. Singh-Miller I. Dabo, B. Kozinsky and N. Marzari. In preparation.
- [86] Matteo Cococcioni and Stefano de Gironcoli. Linear response approach to the calculation of the effective interaction parameters in the lda + u method. *Physical Review B (Condensed Matter and Materials Physics)*, 71(3):035105, 2005.