Optical and Transport Properties of Organic Molecules:
Methods and Applications

by

David Alan Strubbe

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Physics
and the Designated Emphasis
in
Nanoscale Science and Engineering
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Steven G. Louie, Chair
Professor Michael F. Crommie
Professor Martin P. Head-Gordon

Fall 2012
Optical and Transport Properties of Organic Molecules:
Methods and Applications

Copyright 2012
by
David Alan Strubbe
Abstract
Optical and Transport Properties of Organic Molecules:
Methods and Applications
by
David Alan Strubbe
Doctor of Philosophy in Physics
and the Designated Emphasis
in
Nanoscale Science and Engineering
University of California, Berkeley
Professor Steven G. Louie, Chair

Organic molecules are versatile and tunable building blocks for technology, in nanoscale and
bulk devices. In this dissertation, I will consider some important applications for organic
molecules involving optical and transport properties, and develop methods and software
appropriate for theoretical calculations of these properties. Specifically, we will consider
second-harmonic generation, a nonlinear optical process; photoisomerization, in which ab-
sorption of light leads to mechanical motion; charge transport in junctions formed of single
molecules; and optical excitations in pentacene, an organic semiconductor with applications
in photovoltaics, optoelectronics, and flexible electronics. In the Introduction (Chapter 1), I
will give an overview of some phenomenology about organic molecules and these application
areas, and discuss the basics of the theoretical methodology I will use: density-functional
theory (DFT), time-dependent density-functional theory (TDDFT), and many-body per-
turbation theory based on the $GW$ approximation. In the subsequent chapters, I will further
discuss, develop, and apply this methodology.

2. I will give a pedagogical derivation of the methods for calculating response properties
in TDDFT, with particular focus on the Sternheimer equation, as will be used in
subsequent chapters. I will review the many different response properties that can be
calculated (dynamic and static) and the appropriate perturbations used to calculate
them.

3. Standard techniques for calculating response use either integer occupations (as appro-
priate for a system with an energy gap) or fractional occupations due to a smearing
function, used to improve convergence for metallic systems. I will present a general-
ization which can be used to compute response for a system with arbitrary fractional
occupations.
4. Chloroform (CHCl$_3$) is a small molecule commonly used as a solvent in measurements of nonlinear optics. I computed its hyperpolarizability for second-harmonic generation with TDDFT with a real-space grid, finding good agreement with calculations using localized bases and with experimental measurements, and that the response is very long-ranged in space.

5. N@C$_{60}$ is an endohedral fullerene, a sphere of carbon containing a single N atom inside, which is weakly coupled electronically. I show with TDDFT calculations that a laser pulse can excite the vibrational mode of this N atom, transiently turning on and off the system’s ability to undergo second-harmonic generation. The calculated susceptibility is as large as some commercially used frequency-doubling materials.

6. A crucial question in understanding experimental measurements of nonlinear optics and their relation to device performance is the effect of the solution environment on the properties of the isolated molecules. I will consider possible explanations for the large enhancement of the hyperpolarizability of chloroform in solution, demonstrate an ab initio method of calculating electrostatic effects with local-field factors, and derive the equations necessary for a full calculation of liquid chloroform.

7. Many-body perturbation theory, in the GW approximation for quasiparticle band-structure and Bethe-Salpeter equation for optical properties, is a powerful method for calculations in solids, nanostructures, and molecules. The BerkeleyGW code is a freely available implementation of this methodology which has been extensively tested and efficiently parallelized for use on large systems.

8. Molecular junctions, in which a single molecule is contacted to two metallic leads, are interesting systems for studying nanoscale transport. I will present a method called DFT+$\Sigma$ which approximates many-body perturbation theory to enable accurate and efficient calculations of the conductance of these systems.

9. Azobenzene is a molecule with the unusual property that it can switch reversible between two different geometries, cis and trans, upon absorption of light. I have calculated the structures of these two forms when absorbed on the Au(111) surface, to understand scanning tunneling microscope studies and elucidate the switching mechanism on the surface. I have also calculated the conductance of the two forms in a molecular junction.

10. The Seebeck and Peltier thermoelectric effects can interconvert electricity and heat, and are parametrized by the Seebeck coefficient. Standard methods in quantum transport for computing this quantity are problematic numerically. I will show this fact in a simple model and derive a more robust and efficient approach.

11. Pentacene is an organic semiconductor which shows exciton self-trapping in its optical spectra. I will present a method for calculation of excited-state forces with the Bethe-Salpeter equation that can be applied to study the geometrical relaxation that occurs upon absorption of light by pentacene.
For my sister Linda and father Hugo, and in memory of my mother Sandra, for all their help and encouragement along the way.
# Contents

List of Figures vi
List of Tables xi

1 Introduction 1
   1.1 Organic molecules .......................... 1
   1.2 Nonlinear optics ................................ 2
   1.3 Electronic transport ......................... 3
   1.4 Theoretical methods .......................... 4

2 Response functions in TDDFT: concepts and implementation 9
   2.1 Introduction .................................. 9
   2.2 Methods for calculating response functions ........ 9
      2.2.1 Time-propagation method .................... 10
      2.2.2 Sternheimer method .......................... 11
      2.2.3 Casida method ................................ 17
      2.2.4 Generalizations and discussion .............. 18
   2.3 Applications of linear response .................. 19
      2.3.1 Response to electric perturbations .......... 19
      2.3.2 Response to magnetic perturbations .......... 21
      2.3.3 Response to structural perturbations ........ 23
      2.3.4 Mixed electric and structural response to structural perturbations 24
      2.3.5 Response to $k \cdot p$ perturbations .......... 25

3 Density-functional perturbation theory with arbitrary fractional occupations 27
   3.1 Introduction .................................. 27
   3.2 Occupations from smearing ...................... 27
   3.3 Arbitrary occupations .......................... 31
   3.4 Implementation ................................ 32
   3.5 Conclusion .................................... 33

4 Basis-set effects on the hyperpolarizability of CHCl₃: Gaussian-type orbitals, numerical basis sets and real-space grids 34
   4.1 Abstract ..................................... 34
8 First-principles approach to charge transport in single-molecule junctions with self-energy corrections: a DFT+$\Sigma$ method

8.1 Abstract ........................................ 106
8.2 Introduction .................................... 107
8.3 Methods ........................................ 110
  8.3.1 Summary of approach ......................... 110
  8.3.2 Inclusion of self-energy corrections $\Sigma$ in scattering-state DFT+$\Sigma$ approach to transport ....................... 111
  8.3.3 Form of the self-energy-correction operator $\hat{\Sigma}$ .................................................................. 112
  8.3.4 Evaluation of self-energy corrections $\Sigma$, including electrode polarization effects ............................................. 115
  8.3.5 Effect of $\hat{\Sigma}$ on the molecular resonance positions and transmission lineshapes ............................................. 117
  8.3.6 Details of ground-state and transmission calculations .......................................................... 118
8.4 Application to amine-Au and pyridine-Au junctions .................. 119
  8.4.1 Comparison of calculated conductance values to experiment ............. 119
  8.4.2 Origin of transmission at $E_F$ ................. 123
  8.4.3 Calculated transmission spectra .............. 123
8.5 Conclusion ......................................... 126
8.6 Appendix ........................................ 127
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.6.1</td>
<td>Image-charge form of the self-energy correction in weakly coupled</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>molecule-metal systems</td>
<td></td>
</tr>
<tr>
<td>8.6.2</td>
<td>Potential of charge distribution between two planar conductors</td>
<td>132</td>
</tr>
<tr>
<td>9</td>
<td>First-principles calculations on the molecular photoswitch azobenzene</td>
<td>134</td>
</tr>
<tr>
<td>9.1</td>
<td>Abstract</td>
<td>134</td>
</tr>
<tr>
<td>9.2</td>
<td>Introduction</td>
<td>134</td>
</tr>
<tr>
<td>9.3</td>
<td>Confirmation of photoisomerization on surface</td>
<td>135</td>
</tr>
<tr>
<td>9.4</td>
<td>Study of surface absorption and hybridization with metal</td>
<td>135</td>
</tr>
<tr>
<td>9.5</td>
<td>Use of chirality to identify photoisomerization pathway</td>
<td>137</td>
</tr>
<tr>
<td>9.6</td>
<td>Effect of photoisomerization on conductance in molecular junction</td>
<td>142</td>
</tr>
<tr>
<td>10</td>
<td>Efficient k · p method for first-principles calculation of the Seebeck coefficient in quantum transport</td>
<td>147</td>
</tr>
<tr>
<td>10.1</td>
<td>Abstract</td>
<td>147</td>
</tr>
<tr>
<td>10.2</td>
<td>Introduction</td>
<td>147</td>
</tr>
<tr>
<td>10.3</td>
<td>Typical method of calculating the Seebeck coefficient</td>
<td>148</td>
</tr>
<tr>
<td>10.4</td>
<td>Numerical study of convergence in simple model</td>
<td>150</td>
</tr>
<tr>
<td>10.5</td>
<td>Surface-integral formulation</td>
<td>153</td>
</tr>
<tr>
<td>10.6</td>
<td>Conclusion</td>
<td>155</td>
</tr>
<tr>
<td>11</td>
<td>Stark effect and exciton self-trapping in the optical response of pentacene crystals</td>
<td>157</td>
</tr>
<tr>
<td>11.1</td>
<td>Abstract</td>
<td>157</td>
</tr>
<tr>
<td>11.2</td>
<td>Introduction</td>
<td>157</td>
</tr>
<tr>
<td>11.3</td>
<td>Excited-state forces</td>
<td>159</td>
</tr>
<tr>
<td>11.3.1</td>
<td>Approach of Ismail-Beigi and Louie</td>
<td>159</td>
</tr>
<tr>
<td>11.3.2</td>
<td>Reformulation</td>
<td>161</td>
</tr>
<tr>
<td>11.4</td>
<td>Kernel unfolding</td>
<td>163</td>
</tr>
<tr>
<td>11.4.1</td>
<td>Interpolation</td>
<td>163</td>
</tr>
<tr>
<td>11.4.2</td>
<td>Derivation of $q \to 0$ limits</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>169</td>
</tr>
<tr>
<td>A</td>
<td>Errata to BerkeleyGW methodology papers</td>
<td>201</td>
</tr>
</tbody>
</table>
List of Figures

1.1 The logo for the Octopus code for real-space (TD)DFT, freely available at http://www.tddft.org/programs/octopus. .................................................. 7

4.1 Linear response density $\rho_z^{(1)}(r)$ (a-b) and polarizability density $\alpha_{zz}(r)$ (c-d) on one of the HCCl planes of the molecule calculated with a GTO basis set (HP) and a real-space grid (RS) using the PBE functional. The positions of the nuclei are indicated with black dots, and the black lines are isolines. All quantities are in atomic units. Note that the linear response density is quite similar for both methods. GTO data courtesy F. Vila [387]. .......................... 42

4.2 Nonlinear response density $\rho_z^{(2)}(r)$ (a-b) and hyperpolarizability density $\beta_{zzz}(r)$ (c-d) on one of the HCCl planes of the molecule calculated with a GTO basis set (HP) and a real-space grid (RS) using the PBE functional. The positions of the nuclei are indicated with black dots, and the black lines are isolines. All quantities are in atomic units. The nonlinear densities extend much further into space than the linear densities. The agreement between the real-space and GTO methods is nevertheless quite good. The contributions to the hyperpolarizability from the Cl atoms and the CH bond are of opposite sign and, as indicated by the nonlinear response density, have contributions that extend even further into space. GTO data courtesy F. Vila [387]. ............... 44

5.1 (a) Linear susceptibility as a function of the incident photon energy using UCIS/cc-pVDZ method (solid line), UCIS/6-31G method (dotted line – the curves for the four N atom positions lie on top of each other) and TDDFT method (thin line, for $Z = 0.4$ Å). Inset: Ground-state MP2 energy changes (6-31G basis) as a function of the position $Z$ of the nitrogen atom inside the C$_{60}$ cage. (b) Second-order susceptibility $\chi_{zzz}^{(2)}(2\omega;\omega,\omega)$ versus the incident photon energy $\hbar\omega$ for different nitrogen positions inside the C$_{60}$ cage at UCIS/6-31G level. The $Z$-values are shown above each line. The curves are vertically shifted by 0.5 for clarity. UCIS and MP2 data courtesy G. P. Zhang [413]. ............................................ 48
5.2 Second-order susceptibility $\chi^{(2)}_{zzz}$ for five bases from 3-21G to cc-pVDZ. All the structures are optimized at the B3LYP level, and the excited-state properties are computed with UCIS. The TDDFT result is shown at the bottom. The N atom is 0.4 Å away from the center of the C$_{60}$ cage. The horizontal dashed lines denote the zero line. UCIS data courtesy G. P. Zhang [413].

5.3 TDDFT linear susceptibility. The empty and filled circles represent $\chi^{(1)}_{xx}$, vertically shifted for clarity. The empty and filled boxes denote $\chi^{(1)}_{zz}$. The filled circles and boxes are the real parts. The empty circles and boxes are the imaginary parts. The horizontal dotted lines denote the zero lines. The tensor is close to isotropic.

5.4 Second-harmonic spectrum obtained by the TDDFT method. The empty and filled circles represent $\chi^{(2)}_{zxx}$, vertically shifted for clarity. The empty and filled boxes denote $\chi^{(2)}_{zzz}$. The filled circles and boxes are the real parts. The empty circles and boxes are the imaginary parts. The horizontal dotted lines denote the zero line. The frequency dependence and order of magnitude are similar between the real and imaginary parts and between the two tensor components shown here.

6.1 The Lorentz local-field model. A homogeneous dielectric medium (blue) has an applied field to the right, induced positive charges on the right surface and negative on the left, as usual for a capacitor. This reduces the screened field in the medium compared to the external field. However, on the surfaces of a cavity within the medium, the induced charges enhance the local field compared to the screened field. The solution for a spherical cavity is shown.

7.1 The logo for the BerkeleyGW code.

7.2 The absorption spectra for silicon calculated at the GW (black dashed) and GW-BSE (red solid) levels using the BerkeleyGW package. Experimental data from [168].

7.3 Flow chart of a GW-BSE calculation performed in the BerkeleyGW package.

7.4 An example of the construction of 4 × 4 main and shifted k-grids for graphene. (Left) The main grid has a (0.5, 0.5) shift (crystal coordinates). There are 16 points in the full Brillouin zone and 6 irreducible points. (Right) The shifted grid has a (0.0, 0.05) shift (crystal coordinates). There are 48 points in the full Brillouin zone and 26 irreducible points.

8.1 Diagram of model for self-energy correction to molecular levels in a weakly coupled junction within the DFT+Σ method for transport computation. The correction has two parts: a gas-phase term, equal to the difference between Kohn-Sham eigenvalues and quasiparticle levels in the gas phase, which in general opens the HOMO-LUMO gap; and an electrode polarization term, due to metallic screening by the leads of added charge on the molecule, which reduces the gap.
8.2 Set of 7 molecular junctions studied in this work: [a] 1,4-benzenediamine (BDA), [b] 2,3,5,6-tetrafluoro-1,4-benzenediamine (BDA4F), [c] 2,3,5,6-tetramethyl-1,4-benzenediamine (BDA4Me), [d] 1,4-butandiamine, [e] 4,4′-bipyridine, [f] 1,2-di-(4-pyridyl)-ethene (DP-ethene), and [g] 1,4-di-(4-pyridyl)-benzene (DP-benzene). The amine molecules (a-d) have HOMO-dominated conductance, while the pyridine molecules (e-g) have LUMO-dominated conductance. Butanediamine (d) is a saturated alkane, while the others are π-conjugated molecules. Each molecule is bonded to an adatom trimer on an Au(111) face of each lead. [a,e,f,g] courtesy S. Y. Quek [288].

8.3 Comparison of conductances calculated by DFT (black circles) and DFT+Σ (red triangles) transport methods and experimental measurements, on log-log scale. DFT values are all overestimates, but the DFT+Σ approach lowers conductances into agreement with experiment. Note that the trend between bipyridine and DP-ethene (e and f, circled) is wrong in DFT, but corrected by DFT+Σ. Error bars are peak widths in experimental histograms. Numerical values and experimental references are given in Table 8.2. [a,e,f,g] courtesy S. Y. Quek [288].

8.4 Typical DFT+Σ scattering-state wavefunctions at the Fermi energy and molecular orbitals dominating conductance for each molecular junction. The scattering-state wavefunctions closely resemble the orbitals of the isolated molecules, showing that molecular orbitals in these junctions are in the weak-coupling regime. Direction of electron flow is left to right. Isosurfaces for scattering-state wavefunctions are plotted at 2.5% of maximum and for molecular orbitals at 10%. [a,e,f,g] courtesy S. Y. Quek [288].

8.5 Transmission spectra of (a) benzenediamines, (b) butanediamine, and (c) bipyridines within DFT and DFT+Σ approaches. The self-energy correction reduces transmission at the Fermi energy and shifts peaks away from the Fermi energy. In (a), substitution with fluoro (methyl) groups lowers (raises) the HOMO peak energy. Note the complex lineshape for butanediamine in (b). The peak around −2 eV in (a) and (b) is due to hybridization between frontier orbitals and Au adatom d-states. BDA, bipyridine, DP-ethene, DP-benzene data courtesy S. Y. Quek [288].

8.6 Infinite series of charges arising from a point charge in between two planar conductors located at \( z \leq 0 \) and \( z \geq a \). For clarity, the series of image charges is split into two sets: the series of charges that originate first from a reflection in the plane \( z = 0 \) are shown in (a), and those that originate first from a reflection in the plane \( z = a \) are shown in (b). The charges are denoted by their sign, and are labeled according to the order in which they arise. For example, in (a), the charge labeled “1” arises from reflection of the physical charge in plane \( z = 0 \), while that labeled “2” arises from reflection of image charge “1” in plane \( z = a \).

9.1 Structure and photoisomerization reaction of azobenzene.
9.2 Simulated trans- and cis-TTB-azobenzene structures compared to experiment [71]. (a) Calculated trans geometry. (b) Calculated cis geometry. (c) Calculated trans LDOS integrated from $E_F$ to $E_F - 1$ eV, at an isosurface about 3 Å away from the nearest atoms. (d) Calculated cis LDOS isosurface [same parameters as in (c)]. (e) Simulated STM image of TTB-azobenzene using tiled single-molecule LDOS isosurfaces from (c) and (d) (image has been smoothed using a 0.2 nm width Gaussian blur filter to approximate experimental convolution with the STM tip). (f) Experimental STM image of TTB-azobenzene molecules including one photoisomerized cis isomer.

9.3 (a) Image of trans-TTB-AB island composed of two mirror-image chiral domains (domain boundary indicated by dashed white line). Trans-TTB-AB single-molecule structural models are superimposed on right-handed and left-handed configuration. Red crosses indicate locations of TB lobe maxima. (b) Average locations of trans-TTB-AB TB lobes averaged over 394 left-handed domain molecules (left plot) and 125 right-handed domain molecules (right plot). Left-handed and right-handed domain trans-TTB-AB molecules form parallelograms that are angled by 2° from the horizontal and vertical. Figure from Ref. [72].

9.4 Symmetries of the two conformations of cis-TTB-azobenzene. TB legs not shown for clarity. Left: Type 1 has a rotation axis, marked by the cross and arrow. Right: Type 2 has a mirror plane.

9.5 (a) High-resolution STM image [72] shows structure of two distinct types of photoswitched cis-TTB-AB isomers: a left-handed Type I cis and a right-handed Type 2 cis, contained within dashed white boxes. (b) Ab initio DFT simulated STM image of the two cis-TTB-AB types surrounded by trans-TTB-AB molecules, with structural models shown of trans-TTB-AB as well as cis Type 1 and cis Type 2 used to calculate the simulated STM image.

9.6 Photoisomerization dynamics model for TTB-azobenzene on the Au(111) surface. A photoexcited right-handed trans-TTB-azobenzene molecule (fig. top center, TB legs numbered) begins the isomerization reaction with the in-plane inversion of the left phenyl ring through the transition configuration (fig. middle center). With continuing inversion, steric repulsion due to the eventual contact of the left phenyl ring with the right phenyl ring then forces either: (a) left phenyl to twist out of plane lifting TB #1 off of surface producing right-handed Type 2 cis-TTB-azobenzene (figure bottom left), or (b) right phenyl to twist out of plane lifting TB #4 off of the surface producing left-handed Type 1 cis-TTB-azobenzene (figure bottom right). A left-handed trans-TTB-azobenzene isomerizes via two equivalent mirror-reflected paths.


9.9 Detail of Fano resonance in transmission of trans-diaminoazobenzene junction, showing excellent fit to Eq. 9.1 for asymmetry parameter $q = 1.977$, peak center $E_0 = -0.596$ eV, and width $\Gamma = 0.001$ eV.

10.1 Large variation of transmission at different $k_\parallel$-point (the various curves) for a benzenediamine-Au junction. Note the log scale of the $y$-axis. Figure courtesy S. Y. Quek. Calculation as in Ref. [285].

10.2 Upper left: $T(k_x, k_y, k_z(E))$ for the model. Horizontal plane is $k_x$, vertical axis is $T$. Lower left: Numerical integration for $T$, showing rapid convergence. Right: Analytical evaluation of transmission (top) and its derivative (bottom) for the model. We will study the numerical results at the energy $E = -1.5$, marked by the vertical line. Units are arbitrary.

10.3 Slow and irregular convergence of $\partial T / \partial E$ computed via numerical integration and differentiation with parameters $\Delta k$ and $\Delta E$. Top shows variation of both parameters. Bottom left shows variation of $\Delta E$ for two values of $\Delta k$. For $N_k = 25$, convergence is reached, to an incorrect result; for $N_k = 100$, no convergence is attained, inasmuch as the smallest two energy values have a large difference. Note the log scale on the $x$-axis. The analytic result is denoted by the horizontal orange line. Bottom right shows variation of $\Delta k$ for two values of $\Delta E$. $\Delta E = 0.001$ shows initially flat behavior that could be mistaken for convergence, and then goes to the right answer; $\Delta E = 0.5$ converges smoothly to an incorrect result.

10.4 Convergence of the surface term. To capture contributions from the variation of the integration domain in our numerical evaluation of $\partial T / \partial E$, we must have $k$-points in our grid (blue dots) in the shell between the contours for energies $E$ and $E + \Delta E$ (red rings). As $\Delta E$ goes to zero, the width of the shell (orange double-headed arrow) goes to zero also, making it increasingly difficult to retain this contribution.

10.5 Convergence of $\partial T / \partial E$ computed by the surface-integral formulation (Eq. 10.19), compared to the conductance.

11.1 The molecular and crystal structure of pentacene. Adapted from Refs. [365] and [319].

11.2 The physics of the Stokes shift involved in exciton self-trapping. The minimum of the excited-state potential-energy surface is not the same as the ground state’s, leading to a geometrical relaxation and a reduction in the energy of luminescence compared to absorption, with Stokes shift $= E_2 - E_1$.

11.3 The absorption and luminescence spectra of pentacene, showing the Stokes shift of 1.9 eV. From Ref. [11].
List of Tables

1.1 Definition of some linear and nonlinear optical processes. In each case, the first frequency argument is describing the output, and the others describe the inputs. Energy conservation requires the frequency arguments to sum to zero. ................................................................. 2

4.1 Structural parameters used in the study of the variation of the dielectric properties of CHCl$_3$ with structure. PBE/aVQZ and PBE/RS refer to PBE-optimized structures using the aVQZ GTO in Gaussian and a real-space grid in Octopus, respectively. Bond lengths are in Å and angles in degrees. The experimental structure from Ref. [68] was used for all our subsequent calculations. ................................................................. 39

4.2 Dielectric properties of various structures for CHCl$_3$ described in Table 4.1, as calculated by DFT on a real-space grid with radius 17 $a_0$ and spacing 0.25 $a_0$, compared with the experimental values of the dipole moment and the electronic contribution to the polarizability. PBE/aVQZ and PBE/RS refer to the structures described in Table 4.1. All values are in atomic units (au). 39

4.3 Effect of the real-space-grid quality (radius $R$ and spacing $\lambda$) on the components of the dielectric properties of CHCl$_3$ calculated with the PBE functional and LDA kernel. All values are in atomic units. ................................................................. 41

4.4 Partitioning of the linear and nonlinear response properties calculated numerically with GTOs and real-space grids (RS) using the PBE functional together with the numerical sum over the CH and Cl$_3$ partitions. GTO data courtesy F. Vila [387]. ................................................................. 43

4.5 Summary of the best PBE results obtained with the GTOs, numerical basis sets and real-space grids. Real-space grids (lr denotes linear response, and fd finite difference) have radius 22 $a_0$, spacing 0.25 $a_0$. All values are in atomic units. GTO data courtesy F. Vila, NBS data courtesy Y. Takimoto [387]. ................................................................. 43

5.1 Convergence of ground-state, linear and nonlinear properties of N@C$_{60}$ as a function of real-space grid parameters in the Octopus calculation (radius $R$ and spacing $\lambda$). The quoted gap is the Kohn-Sham gap between the highest majority-spin state and lowest minority-spin state. Susceptibilities are shown for the static case and the dynamic case of $\hbar\omega = 3$ eV, with a broadening of $\Gamma = 0.2$ eV. Units are given in the second row for each quantity. ................................................................. 52
6.1 Hyperpolarizability of CHCl$_3$ in experiments and calculations. Experimental data from Ref. [176] except EFISH (liquid) from Ref. [178], and are at a wavelength of 1064 nm. Gas-phase calculations are in 22 au sphere (Chapter 4) and liquid-phase means a crude approximation using a 6 au sphere; all calculations are the electronic contribution at zero frequency. 

6.2 Diagonal part of the Born effective charges for each atom of one molecule in solid CHCl$_3$. All four molecules in the unit cell are equivalent.

8.1 Eigenvalues from molecule in isolation and in junction, compared with transmission peaks. All values are in eV, referenced relative to the Fermi level of the Au leads. (For the isolated molecule, the reference is obtained by first aligning the vacuum levels in the calculations for the isolated molecule and for a 12-layer Au slab, and second, aligning the potential in the middle of the Au slab with the potential in the calculation for the Au leads.) The junction eigenvalue comes from a non-self-consistent diagonalization of the molecular subspace of the DFT Hamiltonian for the converged charge density of junction. Except for butanediamine, where the identification of the DFT transmission peak center for butanediamine is uncertain due to the complicated lineshape, the transmission peak positions are within $\sim 0.2$ eV of the junction eigenvalues, and shift approximately by the self-energy correction of the molecular orbital. The difference between the eigenvalues for the molecule in isolation and in the junction can be identified as the “lead self-energy” (as defined in the text), and represent the effect of charge rearrangement when the molecule is bound to Au. BDA, DP-ethene, DP-benzene data courtesy S. Y. Quek [288].

8.2 Comparison between conductance values computed with DFT, Lorentzian estimate of DFT+$\Sigma$, and full DFT+$\Sigma$ methodology, and experiment. Self-energy corrections $\Sigma$ and their two parts, the gas-phase correction $\Sigma^0$ and the image-charge term $\Delta\Sigma$, are shown. Computed values are reported for different molecules in the trimer junction structure. Values in round (square) brackets denote conductance values obtained when the magnitude of $\Sigma$ is artificially reduced (increased) by 0.2 eV. Experimental error bars correspond to the spread in measured conductance values as quantified by the full-width-half-maximum of peaks in the conductance histograms [289]. BDA, DP-ethene, DP-benzene data courtesy S. Y. Quek [288].

9.1 Self-corrected energies for orbitals of diaminoazobenzene molecular junctions.

11.1 $q \to 0$ limits of the head $\epsilon^{-1}_{GG}(q)$, wing $\epsilon^{-1}_{GG'}(q)$, and wing $\epsilon^{-1}_{GG'}(q)$, of the inverse dielectric matrix, and also of the screened Coulomb interaction, $W_{GG'}(q)$, for the untruncated Coulomb interaction. We abbreviate $M = M(G)$, $M' = M(G')$, $\Delta E = $ energy denominator, $v = v(q+G)$, and $v' = v(q+G')$. $n$ is the number of $G$-vectors, i.e. the size of the matrix.
11.2 Same as Table 11.1, but for truncated Coulomb interactions. We abbreviate \( v = v(q) \), and \( \det = \det \epsilon \). \( c \) denotes an unknown constant. The metal case assumes slab or wire truncation, and the graphene case assumes slab truncation.
Acknowledgments

I would like to thank my advisor Steven Louie for all his help during my PhD; my group-mates Jack Deslippe, Amy Khoo, Jaydeep Sau, Manish Jain, Manos Kioupakis, Georgy Samsonidze, Li Yang, Feliciano Giustino, Cheol-Hwan Park, Brad Malone, Jesse Noffsinger, Young-Woo Son, Felipe Jornada, Siniša Čoh, and Oleg Yazyev for many helpful discussions and camaraderie; my collaborators Su Ying Quek, Jeff Neaton, Isaac Tamblyn, David Prendergast, Sahar Sharifzadeh, Pierre Darancet, Angel Rubio, Miguel Marques, Xavier Andrade, Hyoung Joon Choi, Young Woo Son, Matt Comstock, and Mike Crommie; and my professors Robert Littlejohn, Eugene Chiang, Marvin Cohen, Jeff Grossman, and Joe Orenstein for their excellent classes I took. I thank the National Science Foundation for fellowships from the Nano IGERT center at Berkeley, and the Graduate Research Fellowship Program. Computational resources were provided by the Department of Energy at Lawrence Berkeley National Laboratory’s NERSC facility and Lawrencium cluster, and by University of California Shared Research Computing Services on the Mako cluster.
Chapter 1

Introduction

1.1 Organic molecules

Organic molecules are conventionally defined as ones containing carbon, excluding a few extremely simple cases such as carbon monoxide (CO) [389]. The ability of carbon atoms to form $\sigma$ and $\pi$ bonds with each other and other atoms allows a rich variety of chemical groups, which can be linked together into larger molecules, including polymers. Over a century of research in organic chemistry has yielded knowledge of many different kinds of molecules, allowing a wide design space for tuning the properties of a molecule by varying its chemical substituents and attaching different groups as building blocks to combine functionalities. Many cheap and efficient synthesis methods have been developed. The ability to self-assembly can be designed into molecules, allowing creation of nanostructures and larger-scale ordering. Standard purification techniques can remove undesired products. Organic molecules, as implied by their name, are the basis of life (although of course modern chemistry considers a much wider range of molecules than those found in nature). As a result, they are inherently compatible with biological systems, and can be attached to proteins and other biomolecules. This allows one to take advantage of the natural powers of self-assembly of many biomolecules, to modify the biomolecules for studies of their functionality, and to create hybrid molecules for medical applications. Therefore, organic molecules have become very attractive as materials for some applications which were traditionally the preserve of inorganic solids, due to the promise of being cheaper and more versatile, and may also enable new applications not previously available.

Some important applications for organic molecules involve optical and transport properties, and those are the aspects we will focus on in this dissertation. Specifically, we will consider second-harmonic generation, a nonlinear optical process; photoisomerization, in which absorption of light leads to mechanical motion; charge transport in junctions formed of single molecules; and optical excitations in pentacene, an organic semiconductor with applications in photovoltaics, optoelectronics, and flexible electronics.
refraction and reflection & \text{Re} \alpha (\omega, -\omega) \\
linear absorption & \text{Im} \alpha (\omega, -\omega) \\
sum-frequency generation & \beta (\omega_1 + \omega_2, -\omega_1, -\omega_2) \\
difference-frequency generation & \beta (\omega_1 - \omega_2, -\omega_1, \omega_2) \\
second-harmonic generation & \beta (2\omega, -\omega, -\omega) \\
optical rectification & \beta (0, \omega, -\omega) \\
Pockels (electrooptic) effect & \beta (\omega, 0, -\omega) \\
third-harmonic generation & \gamma (3\omega, -\omega, -\omega, -\omega) \\
two-photon absorption & \text{Im} \gamma (\omega, \omega, -\omega) \\
four-wave mixing & \gamma (\omega_1 + \omega_2 + \omega_3, -\omega_1, -\omega_2, -\omega_3) \\

<table>
<thead>
<tr>
<th>Process</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>refraction and reflection</td>
<td>\text{Re} \alpha (\omega, -\omega)</td>
</tr>
<tr>
<td>linear absorption</td>
<td>\text{Im} \alpha (\omega, -\omega)</td>
</tr>
<tr>
<td>sum-frequency generation</td>
<td>\beta (\omega_1 + \omega_2, -\omega_1, -\omega_2)</td>
</tr>
<tr>
<td>difference-frequency generation</td>
<td>\beta (\omega_1 - \omega_2, -\omega_1, \omega_2)</td>
</tr>
<tr>
<td>second-harmonic generation</td>
<td>\beta (2\omega, -\omega, -\omega)</td>
</tr>
<tr>
<td>optical rectification</td>
<td>\beta (0, \omega, -\omega)</td>
</tr>
<tr>
<td>Pockels (electrooptic) effect</td>
<td>\beta (\omega, 0, -\omega)</td>
</tr>
<tr>
<td>third-harmonic generation</td>
<td>\gamma (3\omega, -\omega, -\omega, -\omega)</td>
</tr>
<tr>
<td>two-photon absorption</td>
<td>\text{Im} \gamma (\omega, \omega, -\omega)</td>
</tr>
<tr>
<td>four-wave mixing</td>
<td>\gamma (\omega_1 + \omega_2 + \omega_3, -\omega_1, -\omega_2, -\omega_3)</td>
</tr>
</tbody>
</table>

Table 1.1: Definition of some linear and nonlinear optical processes. In each case, the first frequency argument is describing the output, and the others describe the inputs. Energy conservation requires the frequency arguments to sum to zero.

1.2 Nonlinear optics

Let us first consider the field of nonlinear optics. The common optical processes of absorption, reflection, and refraction are linear in the sense that they occur proportionally to the incident light intensity. This is because at a microscopic level they involve only single photons interacting with a material. There are also nonlinear processes in which multiple photons interact at the same time, giving rise to quadratic or higher-order dependence on the intensity. For example, in two-photon absorption two photons are absorbed at once, with a cross section proportional to the square of the light intensity. Another process is second-harmonic generation (SHG), in which two incident photons of frequency \( \omega \) are converted into one photon of frequency \( 2\omega \). A zoo of other processes exist, depending on the number and combinations of frequencies, summarized in Table 1.1. The cross-section for these processes is determined by the nonlinear susceptibility, which (in finite systems) is described by a Taylor expansion of the dipole moment in an applied field \( \mathcal{E} \) at frequency \( \omega \):

\[
\mu_i (\mathcal{E}) = \mu_i (0) + \sum_j \alpha_{ij} (\omega, -\omega) \mathcal{E}_j + \sum_{jk} \beta_{ijk} (2\omega, -\omega, -\omega) \mathcal{E}_j \mathcal{E}_k + \ldots \tag{1.1}
\]

The coefficients of succeeding terms are conventionally labeled \( \gamma \) and \( \delta \). In the case of crystals, one speaks instead of the expansion of the polarization, with coefficients \( \chi^{(1)}, \chi^{(2)} \), etc.

A familiar application of SHG is the green laser pointer, which does not really contain a green laser. Instead a frequency-doubling crystal converts infrared from a laser operating at a wavelength of 1064 nm into green light at 532 nm. While this is a relatively trivial application, generally nonlinear optical processes can be used to modulate and convert light, with communications and potential all-optical logic circuits being important areas. Difference-frequency generation is often used for tunable light sources, where tunability in the visual can be transferred to the infrared. Second-harmonic generation is an extremely important technique for characterization in surface science [329], chemistry [79], and biology.
Perhaps the most impressive example is the National Ignition Facility at Lawrence Livermore National Laboratory, in which two stages of second-harmonic generation and sum-frequency generation (the generalization $\omega_1 + \omega_2 = \omega_3$) are used to convert 1053 nm to 351 nm from the most powerful laser in the world, which delivers 500 TW pulses for inertial-confinement fusion experiments [249].

Symmetry analysis of nonlinear susceptibility tensors can be very useful. In the presence of inversion symmetry, all even orders ($\mu$, $\beta$, etc.) are forbidden. As a result, SHG on the surface of a centrosymmetric crystal cannot come from the bulk, making this technique highly surface-sensitive. Since the tensors may be expressed as derivatives of the total energy with respect to field (beginning with $\mu = \partial E/\partial E$), they possess a permutation symmetry due to the equality of mixed partial derivatives, e.g.

$$\beta_{ijk} (\omega_1, \omega_2, \omega_3) = \beta_{jik} (\omega_2, \omega_1, \omega_3)$$

(1.2)

Note that both the spatial indices $ij$ and the corresponding frequencies were swapped. For low frequencies, off-resonant, the approximate “Kleinman” symmetry [188] also holds, where the frequencies are not permuted:

$$\beta_{ijk} (\omega_1, \omega_2, \omega_3) \approx \beta_{jik} (\omega_1, \omega_2, \omega_3)$$

(1.3)

These considerations can make a significant reduction in the number of tensor elements to calculate [322].

Nonlinear optical elements in current technology are inorganic crystals which must have very high quality to avoid losses from linear absorption and other unwanted processes. Organic molecules have the potential to be not only cheaper to produce and process, but much more efficient than inorganic materials. For second-harmonic generation, the gas-phase chloroform molecule (CHCl$_3$) has $\beta \sim 1$ a.u. (as discussed in Chapter 4). A hypothetical solid of such molecules all oriented would have a bulk susceptibility $\chi^{(2)} \sim 1$ esu. This is actually the same order of magnitude as state-of-the-art inorganic materials [346], and this is for a solvent molecule used due to its small background, compared to values of $\beta \sim 10000$ a.u. for molecules optimized for nonlinear susceptibility. However, a problem remains finding condensed phases in which the nonlinear susceptibilities of the molecules do not cancel but add due to their orientations [180].

### 1.3 Electronic transport

Single-molecule junctions are interesting systems to study because of their potential applications in future nanoelectronic devices and as relatively well-defined model systems in which to study fundamental phenomena about nanoscale charge transfer [170]. Current fabrication methods such as the modified scanning tunneling microscope (STM) break-junction technique [401] do not seem likely to be able to be translated into reliable production methods, but arrays of similar structures in organic-inorganic hybrid devices have been created by self-assembly [392, 317]. These weakly coupled and heterogeneous systems, out of equilibrium when there is an applied bias and a current flowing, are quite challenging for electronic-structure methods, and their study has motivated many theoretical developments.
Many interesting phenomena about nanoscale transport can be analyzed in molecular junctions. They are exemplars of resonant tunneling since the molecules have discrete levels broadened by interaction with the contacts, which only contribute to conduction when close to the Fermi level. This gives rise to exponential decay of conductance with length, which can be connected to the complex band structure [282, 102, 286]. Absorption of light by the molecule could lead to photoconductance [388]. The presence of independent conducting eigenchannels [39] gives rise to characteristic patterns in the shot noise of individual electrons traversing the junction, as has been analyzed experimentally, as have inelastic processes in which molecular vibrations are excited or deexcited by electrons hopping on and off the molecule [349]. The structure of the molecules can crucially affect the conductance, in conformation and bonding to the contacts [383, 154, 412, 289, 287, 393]. Control of conformation or bonding by electric field, current, mechanical pulling, or light could enable use of molecular junctions as nanoelectronic switching elements [69, 71]. Interference effects between different scattering pathways through a molecule can lead to unusual structures like Fano resonances [267]. Processes like Coulomb blockade [333], Kondo physics [360], and spin-dependent scattering [192] can be found. Molecular junctions are interesting systems too for the study of thermoelectric effects, in which electrical and thermal transport are coupled. The Seebeck coefficient, the constant of proportionality between voltage and temperature differences across an interface, depends critically on the alignment of the Fermi level with respect to molecular resonances [266]. It is proportional to the derivative of the conductance with respect to energy of the electrons, and thus offers additional information about transport complementary to a measurement of the conductance. Subtle issues about the nature of the non-equilibrium situation with thermal and electrical currents arise and must be considered [95]. Many of these phenomena exist in mesoscopic systems too, such as quantum dots; molecular junctions offer a chance to study them on a smaller scale, which can be more accessible to theoretical study and better defined atomistically.

1.4 Theoretical methods

It is profitable to apply theoretical calculations to explore the physics of optical and transport properties in organic molecules. First, calculations can be faster and cheaper than experiments, allowing theory to consider systems and properties not yet measured and predict what experiments would be interesting to pursue. Second, theory can consider a system measured in an experiment and provide complimentary information, for example microscopic or mechanistic details that elucidate why experiments observe what they do. Third, theory can help to give a general framework for understanding phenomenology.

In our theoretical calculations, we will focus on electronic structure, described by quantum mechanics [223]. We begin with the Hamiltonian for nuclei (indexed by Greek indices) and electrons (indexed by Latin indices), including kinetic energy and Coulomb interactions between electrons and nuclei:

\[
\hat{H}_{\text{el-nuc}} = -\sum_{\alpha} \frac{\hbar^2}{2m_{\alpha}} \nabla_{\alpha}^2 - \sum_{i} \frac{\hbar^2}{2m_{i}} \nabla_{i}^2 + \sum_{i \neq j} \frac{e^2}{|r_i - r_j|} + \sum_{\alpha \neq \beta} \frac{Z_{\alpha}Z_{\beta}}{|r_{\alpha} - r_{\beta}|} + \sum_{i \alpha} \frac{Z_{\alpha}e^2}{|r_i - r_{\alpha}|} \quad (1.4)
\]

In the Born-Oppenheimer approximation, due to the much larger mass of the nuclei,
we treat the them classically and the electrons quantum-mechanically. The total energy for a set of nuclear coordinates is

$$E(\{r_\alpha\}) = \sum_\alpha \frac{p_\alpha^2}{2m_\alpha} + \sum_{\alpha \neq \beta} \frac{Z_\alpha Z_\beta}{|r_\alpha - r_\beta|} + E_{\text{elec}}(\{r_\alpha\})$$  \hspace{1cm} (1.5)

where the electronic energy is given by solving the Hamiltonian

$$\hat{H}_{\text{elec}} = -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \sum_{i \neq j} \frac{e^2}{|r_i - r_j|} + \sum_{i\alpha} \frac{Z_\alpha e^2}{|r_i - r_\alpha|}$$  \hspace{1cm} (1.6)

Next we simplify with the pseudopotential approximation, in which we treat the core electrons and nuclei together as an ion. Its energy is determined by a separate atomic calculation, and then we treat only the valence electrons explicitly in our Hamiltonian. The valence electron-nucleus and valence electron-core electron Coulomb interactions are combined into an effective pseudopotential for each ion. Our electronic Hamiltonian now is:

$$\hat{H}_{\text{elec}} = -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 + \sum_{i \neq j} \frac{e^2}{|r_i - r_j|} + \sum_{i\alpha} V_{\text{ps}}(r_i)$$  \hspace{1cm} (1.7)

The term describing Coulomb interactions among valence electrons remains a formidable challenge for electronic structure because it depends upon the correlations between the electrons in the many-electron wavefunction. Many approximations schemes have been developed to tackle this problem. We will use one of the most successful and common, density-functional theory (DFT) [109]. Hohenberg and Kohn proved that all ground-state properties of the many-electron system can be derived from the electronic density, allowing a reformulation of our problem in terms of the density instead of the many-electron wavefunction [145]. Kohn and Sham developed a mapping between the physical interacting system and a fictitious non-interacting system in the presence of an effective potential, giving a practical approach to solve the problem in terms of individual Kohn-Sham (KS) Schrödinger equations [190]. The effective potential for the electron-electron term is divided into two parts, the mean-field Hartree potential $V_H$ due to the Coulomb interaction of the classical charge density of the electrons, and a remainder $V_{\text{xc}}$, called the exchange-correlation potential, which is determined from the density via a universal functional. The one-electron KS Hamiltonian has the form

$$\hat{H}_{\text{KS}} = -\frac{\hbar^2}{2m} \nabla^2 + V_H[\rho](r) + V_{\text{xc}}[\rho](r) + \sum_\alpha V_{\text{ps}}(r_\alpha)$$  \hspace{1cm} (1.8)

$$V_H[\rho](r) = \int \frac{\rho(r')}{|r - r'|} dr'$$  \hspace{1cm} (1.9)

We solve the Schrödinger equation $\hat{H}_{\text{KS}} \psi_i = E_i \psi_i$ to obtain KS eigenvalues and orbitals, from which the density is determined by $\rho(r) = \sum_i |\psi_i(r)|^2$. This set of coupled equations must be solved self-consistently, i.e. the density input into the Hamiltonian must be the same as the one calculated from the resulting orbitals. At the end, all properties can in principle be obtained from this density.
While this approach is exact in principle, in practice we do not know the form of the universal functional for $V_{xc}$, and must make approximations. The simplest is the local density approximation (LDA) [270] in which the potential at a point depends only on the density at that point, and can be parametrized from data on the homogeneous electron gas:

$$V_{xc}^{\text{LDA}}[\rho](\mathbf{r}) = V_{xc}[\rho(\mathbf{r})](\mathbf{r})$$  \hfill (1.10)

We will use in this dissertation primarily the more general and slightly more complicated generalized gradient approximation (GGA), in which the potential depends additionally on the gradients of the density at that point:

$$V_{xc}^{\text{GGA}}[\rho](\mathbf{r}) = V_{xc}[\rho(\mathbf{r}), \nabla \rho(\mathbf{r})](\mathbf{r})$$  \hfill (1.11)

Specifically, we will use the flavor due to Perdew, Burke, and Ernzerhof (PBE) [269], which often is found to give better results for molecules than the LDA. To solve these one-particle equations, we use a basis to represent the orbitals. Common choices are plane waves [155], pseudoatomic orbitals [328], and real-space grids [60].

The DFT formalism is very accurate for ground-state properties, such as total energy, multipole moments, forces, and vibrations. However, to obtain excited-state properties such as optical spectra or conductance, we must go beyond DFT. Naïve interpretation of the KS solutions as the wavefunctions and energies of electron excited states would suggest absorption of light at energies corresponding to the differences between these eigenvalues, and high conductance in a junction when the Fermi level of the contacts is aligned with a KS eigenvalue of the molecule. However, this conclusion is in general quantitatively incorrect, and sometimes qualitatively too. Nevertheless, DFT is a convenient starting point. Two complementary approaches can be used to calculate excited-state properties [260]: time-dependent density-functional theory (TDDFT) and many-body perturbation theory. We will use the first for optical properties of molecules and the second for conductance of molecular junctions and optical properties of a crystalline solid.

TDDFT is a generalization of DFT [222, 220]. Whereas DFT reformulated the time-independent Schrödinger equation $\hat{H}\psi = E\psi$ in terms of the density $\rho$, TDDFT reformulates the time-dependent Schrödinger equation $\hat{H}\psi = -i\hbar \frac{d}{dt}\psi$ in terms of the time-dependent density $\rho(t)$. Runge and Gross proved a one-to-one mapping between the time-dependent external potential and the time-dependent density, in analogy to the Hohenberg-Kohn Theorem [306]. We can obtain time-dependent KS equations:

$$-i\hbar \frac{d}{dt}\psi_i(\mathbf{r}) = -\frac{\hbar^2}{2m} \nabla^2 \psi_i(\mathbf{r}) + \left[ V_{H}[\rho(t)](\mathbf{r}) + V_{xc}[\rho](\mathbf{r}) + \sum_\alpha V_{\alpha}^{\text{ps}}(\mathbf{r}) \right] \psi_i(\mathbf{r})$$  \hfill (1.12)

These equations can be integrated, that is, propagated in time from an initial condition, in order to describe the response in time to a perturbation. One can start with a molecule in the ground state and apply a kick containing all frequencies, and Fourier-transform the resulting oscillations of the dipole moment to obtain an absorption spectra. There are also several linear-response approaches, in which we find the response to an infinitesimal perturbation, such as Casida’s equations and the Sternheimer equation. We must make an approximation for the time-dependent exchange-correlation potential, just as in static
DFT. In principle, it depends on the density at all earlier times (“memory”) [214], which is problematic for practical calculations. Nevertheless, good results are generally found with an adiabatic approximation, in which a particular exchange-correlation potential for static DFT is used, evaluated only at the instantaneous time:

$$V_{xc}[\rho](\mathbf{r}, t) = V_{xc}[\rho(t)](\mathbf{r}, t)$$  \hspace{1cm} (1.13)

In linear response, we use also the exchange-correlation kernel $f_{xc}$ which is the functional derivative of the potential with respect to density, just as the potential is the functional derivative of the energy. In frequency space, the adiabatic approximation means that the kernel is independent of frequency. This dissertation will use the adiabatic LDA, which generally gives accurate results for small molecules [222]. Studies of approximate kernels indicate that failures in extended systems are due more to problems with the spatial dependence than with a lack of memory-dependence [298, 320]. In Chapter 2, we will consider in more detail how TDDFT can be used to calculate response properties and which properties can be obtained. We will use this approach in Chapters 3, 4, 5, and 6, employing the implementation in the real-space code Octopus [218, 55, 7] (Fig. 1.1), which I have helped develop.

The second approach for excited-state properties starting from DFT is many-body perturbation theory, in which one of the fundamental objects are quasiparticles, that is, the particle-like excitations of the many-body system. In particular, for conductance, we focus on the quasi-electrons and quasi-holes, which are the states of the system when an electron is added to or removed from the system. For finite systems, the energies of the lowest excitations may be obtained within the DFT framework via the $\Delta$SCF technique [171], as a difference of total energies between the neutral system with $N$ electrons and the charged system with $N \pm 1$ electrons. However, for extended systems this is not possible. Instead, we use a diagrammatic technique, in which the KS orbitals and eigenvalues are taken as a mean-field solution, to which missing correlation will be added as a perturbation, which allows us to calculate the full spectrum of quasiparticles, for finite and extended systems.
We exploit the fact that the KS orbitals and eigenvalues are usually in qualitative agreement with those of the quasiparticles, and within the $GW$ approximation [141], we make our expansion in terms of the screened Coulomb interaction,

$$W (r, r') = \epsilon^{-1} (r, r') \frac{1}{|r - r'|}$$  (1.14)

We calculate the self-energy operator $\tilde{\Sigma} (E)$ to obtain quasiparticle energies. Optical spectra can also be calculated via the Bethe-Salpeter equation, in which we consider an effective two-particle Hamiltonian in the basis of pairs of electron and hole quasiparticle excitations, which interact via a kernel with an exchange term and a direct term based on the screened Coulomb interaction $W$. This approach, and its implementation in a computer code, will be discussed more in Chapter 7, and used in an approximate form for transport in Chapter 8, and for excited-state properties in Chapter 11.
Chapter 2

Response functions in TDDFT: concepts and implementation

This work is based on Ref. [342].

2.1 Introduction

Many physical properties of interest about solids and molecules can be considered as the reaction of the system to an external perturbation, and can be expressed in terms of response functions, in time or frequency and in real or reciprocal space. Response functions in TDDFT can be calculated by a variety of methods. Time-propagation is a non-perturbative approach in the time domain, whose static analogue is the method of finite differences. Other approaches are perturbative and are formulated in the frequency domain. The Sternheimer equation solves for the variation of the wavefunctions, the Dyson equation is used to solve directly for response functions, and the Casida equation solves for the excited states via an expansion in an electron-hole basis. These techniques can be used to study a range of different response functions, including electric, magnetic, structural, and $k \cdot p$ perturbations. In this chapter, we give an overview of the basic concepts behind response functions and the methods that can be employed to efficiently compute the response properties within TDDFT and the physical quantities that can be studied.

2.2 Methods for calculating response functions

In this section, we will briefly describe three different methods to calculate response from TDDFT: (i) time-propagation, (ii) Sternheimer, and (iii) Casida method. The time-propagation method [404] simply propagates a system under a given external field. The density response is obtained directly as the difference between the time-dependent density and the ground-state state density. As the method is nonperturbative, all orders of response are included in the calculation, and therefore, specific orders must be numerically extracted. The Sternheimer method [338, 22, 123, 8] solves for a specific order of the response for a specific field in frequency space (i.e., it is a perturbative approach). The Sternheimer
equations form a hierarchical structure, where higher-order responses can be calculated from lower-order responses. The Casida method [53], instead of finding the response, finds the poles and residues of the first-order response function, which corresponds to finding the resonant transitions of a system. Note that physically all these techniques are equivalent as they are all based on Kohn-Sham DFT and are simply different ways to obtain the same quantities. Often the choice between them is done purely on numerical arguments, as each one is more adapted to certain numerical implementations [221].

We will show that the three methods are connected to each other in a simple way. Since the purpose is to connect all these approaches, we will consider only weak perturbations. For pedagogical reasons, we make the following simplifying assumptions in the discussion below that can be easily generalized (see discussion at the end of this section): (i) the exchange-correlation functional does not have memory, i.e., we work within the adiabatic approximation, (ii) the system is spin-unpolarized, (iii) we have no fractional occupations, and (iv) we use no \( k \)-point sampling, i.e., only \( \Gamma \)-point or a non-periodic system, which allows us to use purely real (i.e., imaginary part is zero) ground-state Kohn-Sham wavefunctions. Assumption (i) is in practice not very restrictive, as a large majority of the functionals used in TDDFT are, indeed, adiabatic. Assumptions (ii)-(iv) are valid, for example, for closed-shell molecules. In any case it is fairly straightforward (but cumbersome) to remove the assumptions from the derivation. We will return to this topic at the end of the section.

2.2.1 Time-propagation method

In the time-propagation approach, the time-dependent Kohn-Sham equations are propagated in real-time, i.e., by solving the following nonlinear partial differential equation

\[
i \frac{\partial}{\partial t} \varphi_k(r, t) = \hat{H}_{\text{KS}}[n](r, t)\varphi_k(r, t),
\]

starting from time \( t = 0 \) with the initial condition \( \varphi_k(r, t = 0) = \varphi_k^{(0)}(r) \), where \( \varphi_k^{(0)}(r) \) are the ground-state Kohn-Sham wavefunctions. Here, we already have applied the adiabatic approximation by assuming that \( \hat{H}_{\text{KS}}[n](r, t) \) has a functional dependence only on the instantaneous density \( n(r, t) \) instead of on its whole history.

If no perturbation is applied to the system, the system remains in the ground-state and the time-evolution of the KS wavefunctions is trivial: \( \varphi_k(t) = \varphi_k^{(0)} e^{-i\varepsilon_k^{(0)} t} \). If we apply a weak time-dependent external perturbation with a given frequency \( \omega \), the time-evolution becomes nontrivial because of the nonlinearity of the Kohn-Sham Hamiltonian. A general form for a weak time-dependent external perturbation with a given frequency \( \omega \) is

\[
v_{\text{ext}}(r, t) = \lambda v^\text{cos}_{\text{ext}}(r) \cos(\omega t) + \lambda v^\text{sin}_{\text{ext}}(r) \sin(\omega t)
\]

or, rewriting in terms of the exponential

\[
v_{\text{ext}}(r, t) = \lambda v^+_{\text{ext}}(r) e^{i\omega t} + \lambda v^-_{\text{ext}}(r) e^{-i\omega t},
\]

where \( \lambda \) is the strength of the perturbation. If we now insert this external potential to the TDKS equation and propagate in time, we can obtain physical observables from the
time-dependent expectation values. For example, if we apply a weak delta pulse of a dipole electric field

\[ v_{\text{ext}}(\mathbf{r}, t) = -e \mathbf{r} \cdot \vec{K} \delta(t) = -e \mathbf{r} \cdot \vec{K} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \exp(i\omega t), \quad (2.4) \]

we simply replace the ground-state wavefunctions (eigenfunctions of the Hamiltonian \( \hat{H}_{KS}^{(0)} \)) by

\[ \varphi_k(\mathbf{r}, t = 0^+) = \exp \left\{ -\frac{i}{\hbar} \int_{0^-}^{0^+} dt \left[ \hat{H}_{KS}^{(0)}(t) - e \mathbf{r} \cdot \vec{K} \delta(t) \right] \right\} \varphi_k(\mathbf{r}, t = 0^-) = \exp \left( i e \mathbf{r} \cdot \vec{K} / \hbar \right) \varphi_k(\mathbf{r}, t = 0^-), \quad (2.5) \]

and propagate the free oscillations in time. Then the time-dependent dipole moment

\[ \vec{\mu}(t) = -e \int d^3 r \mathbf{r} n(\mathbf{r}, t) \quad (2.6) \]

can be used to extract the dynamic polarizability tensor \( \alpha(\omega) \). We Fourier-transform to obtain

\[ \alpha_{\gamma\delta}(\omega) = \frac{1}{K_\delta} \int_0^\infty dt \left[ \mu_\gamma(t) - \mu_\gamma(0^-) \right] e^{-i\omega t} + \mathcal{O}(K_\delta). \quad (2.7) \]

The imaginary part of the diagonal component of the dynamic polarizability \( \Im[\alpha_{\delta\delta}(\omega)] \) is proportional to the absorption spectrum. The above equation includes an integral over infinite time. Obviously, infinite time-propagation is not possible in practice, and we have to add an artificial lifetime to the equation by introducing a decay \( e^{-\eta t} \):

\[ \alpha_{\gamma\delta}(\omega) = \frac{1}{K_\delta} \int_0^\infty dt \left[ \mu_\gamma(t) - \mu_\gamma(0^-) \right] e^{-i\omega t} e^{-\eta t} + \mathcal{O}(K_\delta). \quad (2.8) \]

which corresponds to forcing all excitations to decay back to the ground state with rate \( \eta \).

Higher-order responses (e.g., hyperpolarizabilities) are automatically considered in the calculation. However, if the field is chosen weak enough, they have negligible contribution, as should be the case for a linear-response calculation. If the perturbation strength is increased, the nonlinear contributions begin to increase: 2nd order quadratically, 3rd order cubically, etc. In addition to the different scaling with respect to the perturbation strength, higher-order responses appear at combinations of existing frequencies, which will be discussed in the next subsection. To disentangle the different contributions from the time propagation scheme is not always a well-defined procedure.

### 2.2.2 Sternheimer method

The time-propagation approach propagates the TDKS equations in real-time. If we apply time-dependent perturbation theory and transform the equations to frequency space,
we obtain the Sternheimer method, which is also known as density-functional perturbation theory, particularly in its static form [19], and as “coupled perturbed Kohn-Sham” in the quantum-chemistry literature.

As the time-dependent external perturbation was chosen weak, we can expand the TD-KS states and the TD-KS-Hamiltonian as a power series with respect to the perturbation strength \( \lambda \). The perturbation expansion\(^3\) of the TD-KS states reads

\[
\varphi_k(r, t) = \varphi_k^{(0)}(r, t) + \lambda \varphi_k^{(1)}(r, t) + \lambda^2 \varphi_k^{(2)}(r, t) + \ldots
\]

where the zeroth-order response has trivial time-dependence \( \varphi_k^{(0)}(r, t) = \varphi_k^{(0)}(r) e^{-i \varepsilon_k^{(0)} t} \). The perturbation expansion of the TD-KS Hamiltonian reads

\[
\hat{H}_{\text{KS}}(r, t) = \hat{H}_{\text{KS}}^{(0)}[n^{(0)}](r, t) + \lambda v_{\text{ext}}^{(1)}(r, t) + \lambda \hat{H}_{\text{KS}}^{(1)}[n](r, t)
\]

\[
\quad + \lambda^2 v_{\text{ext}}^{(2)}(r, t) + \lambda^2 \hat{H}_{\text{KS}}^{(2)}[n](r, t) + \ldots
\]

where \( \hat{H}_{\text{KS}}^{(0)}[n^{(0)}](r) \) is the ground-state Hamiltonian. The \( \hat{H}_{\text{KS}}^{(k)}[n](r, t) \) are the \( k \)th-order response Hamiltonians, \textit{i.e.}, \( k \)th derivatives of the Hamiltonian with respect to the magnitude of the bare external perturbation.\(^4\) These response Hamiltonians arise from the nonlinearity of the TD-KS Hamiltonian: the Hartree and exchange-correlation potentials are affected too when the system is perturbed by the time-dependent external potential \( v_{\text{ext}}(r, t) \).

The response Hamiltonians require the time-dependent density

\[
n(r, t) = \sum_k n_k |\varphi_k(r, t)|^2 = n^{(0)}(r, t) + \lambda n^{(1)}(r, t) + \lambda^2 n^{(2)}(r, t) + \ldots
\]

Inserting the expansion for the KS wavefunctions

\[
n(r, t) = \sum_k n_k \left\{ |\varphi_k^{(0)}(r, t)|^2 + \lambda \left\{ |\varphi_k^{(0)}(r, t)|^* \varphi_k^{(1)}(r, t) + |\varphi_k^{(1)}(r, t)|^* \varphi_k^{(0)}(r, t) \right\} \right.
\]

\[
\quad + \lambda^2 \left\{ |\varphi_k^{(2)}(r, t)|^* \varphi_k^{(0)}(r, t) + |\varphi_k^{(0)}(r, t)|^* \varphi_k^{(2)}(r, t) + |\varphi_k^{(1)}(r, t)|^2 \right\} \right\} + \ldots
\]

where \( n_k \) is the occupation of the \( k \)th KS-state.

Each response Hamiltonian \( \hat{H}_{\text{KS}}^{(k)}[n(r', t)](r, t) \) depends only on the response densities \( n^{(j \leq k)}(r, t) \) which are of lower or equal order. For example, the zeroth-order response Hamiltonian is just the ground-state KS-Hamiltonian, which depends only on the ground-state density. The first-order response Hamiltonian

\[
\hat{H}_{\text{KS}}^{(1)}[n](r, t) = \int d^3 r' f_{\text{Hxc}}[n^{(0)}](r, r') n^{(1)}(r', t)
\]

has the first-order Hartree-exchange-correlation kernel \( f_{\text{Hxc}}[n^{(0)}](r, r') \), which depends on the ground-state density \( n^{(0)}(r) \), multiplied by the first-order density response \( n^{(1)}(r, t) \).

\(^3\)Note that this expansion is not a Taylor expansion.

\(^4\)Remember that we are working within the adiabatic approximation here, and therefore, the TD-KS Hamiltonian has no memory.
Now, by equating different orders of $\lambda$ in the TD-KS equation, we obtain in zeroth order
\[
i \frac{\partial}{\partial t} \varphi_k^{(0)}(r, t) = \hat{H}_{\text{KS}}^{(0)}[n^{(0)}](r) \varphi_k^{(0)}(r, t),
\]
in first order,
\[
i \frac{\partial}{\partial t} \varphi_k^{(1)}(r, t) = \hat{H}_{\text{KS}}^{(0)}[n^{(0)}](r) \varphi_k^{(1)}(r, t) + \left[ \hat{H}_{\text{KS}}^{(1)}[n](r, t) + v_{\text{ext}}^{(1)}(r, t) \right] \varphi_k^{(0)}(r, t),
\]
in second order,
\[
i \frac{\partial}{\partial t} \varphi_k^{(2)}(r, t) = \hat{H}_{\text{KS}}^{(0)}[n^{(0)}](r) \varphi_k^{(2)}(r, t) + \left[ \hat{H}_{\text{KS}}^{(1)}[n](r, t) + v_{\text{ext}}^{(1)}(r, t) \right] \varphi_k^{(1)}(r, t) + \left[ \hat{H}_{\text{KS}}^{(2)}[n](r, t) + v_{\text{ext}}^{(2)}(r, t) \right] \varphi_k^{(0)}(r, t),
\]
and so on. The equations form a hierarchy, where higher-order responses can be calculated from the lower-order ones [127, 123].

The above equations still depend on time in a non-trivial way (except the zeroth order which is just the trivial time-propagation of the ground-state). Nevertheless, the only explicit time-dependence is in the time-dependent external potential. If the potential has only one frequency, the linear response will also have only one frequency. If the potential has two frequencies, the linear response has two. However, the second-order response will have frequencies which are sums and differences of the original frequencies. For example, in case of one frequency $\omega$, the second-order response has frequency $\omega - \omega = 0$ and frequency $\omega + \omega = 2\omega$. Furthermore, in the case of two frequencies $\omega_1$ and $\omega_2$, the second-order response has frequencies $0, 2\omega_1, 2\omega_2, \omega_1 + \omega_2$, and $|\omega_1 - \omega_2|$. The 3rd-order mixes three frequencies, and in addition to the frequencies of the field, it can also mix the frequencies generated by the 2nd-order response. Each new order brings new mixed frequencies.

From this point onward, we assume that we have only a single frequency $\omega$ in the external field:
\[
v_{\text{ext}}^{(1)}(r, t) = v_{\text{ext}}^{+\omega}(r) e^{+i\omega t} + v_{\text{ext}}^{-\omega}(r) e^{-i\omega t},
\]
or, if we choose to use a cosine field,
\[
v_{\text{ext}}^{(1)}(r, t) = \frac{1}{2} v_{\text{ext}}^{+\omega}(r) e^{+i\omega t} + \frac{1}{2} v_{\text{ext}}^{-\omega}(r) e^{-i\omega t} = v_{\text{ext}}^{\omega}(r) \cos(\omega t).
\]
A general first-order wavefunction in this case can be written as
\[
\varphi(r, t) = e^{-i\varepsilon^{(0)} t - i\lambda \Delta \varepsilon^{(1)}(t)} \times \left\{ \varphi^{(0)}(r) + \lambda \left[ \varphi^{(1)}_{+\omega}(r) e^{i\omega t} + \varphi^{(1)}_{-\omega}(r) e^{-i\omega t} \right] \right\} + \mathcal{O}(\lambda^2),
\]
where $\varphi_\omega(r)$ are now time-independent, and we have included a time-dependent level shift
\[
\Delta \varepsilon^{(1)}[n](t) = \int_{-\infty}^{t} dt \langle \varphi^{(0)} | \hat{H}_{\text{KS}}^{(1)}[n](t) + v_{\text{ext}}^{(1)}(t) | \varphi^{(0)} \rangle.
\]
The first-order level shift $\Delta \varepsilon^{(1)}(t)$ is a first order correction to the phase of the zeroth-order wavefunction caused by the first-order Hamiltonian. By including it, we keep the first-order
wavefunction $\varphi_{\pm \omega}^{(1)}$ orthogonal to the zeroth-order wavefunction $\varphi^{(0)}$. Otherwise, $\varphi_{\pm \omega}^{(1)}$ would be time-dependent and include a time-dependent $\varphi^{(0)}$ component in order to correct the phase [201].

From the wavefunction, we obtain the response densities. The zeroth-order density is just the ground-state density
\begin{equation}
\varphi^{(0)}(r, t) = \sum_k n_k |\varphi_{k}^{(0)}(r)|^2, \tag{2.21}
\end{equation}
and the first-order response density
\begin{equation}
n^{(1)}(r, t) = \sum_k n_k \left\{ [\varphi_{k}^{(0)}(r)]^* \varphi_{k, + \omega}^{(1)}(r) e^{i \omega t} + [\varphi_{k}^{(0)}(r)]^* \varphi_{k, - \omega}^{(1)}(r) e^{-i \omega t} + [\varphi_{k, + \omega}^{(1)}(r)]^* \varphi_{k}^{(0)}(r) e^{-i \omega t} + [\varphi_{k, - \omega}^{(1)}(r)]^* \varphi_{k}^{(0)}(r) e^{i \omega t} \right\}.
\end{equation}
is oscillating at the frequency $\omega$ as expected.

Next, we insert the guess wavefunction to the TDKS equation and expand it up to first order. On the left-hand-side, we obtain
\begin{equation}
i \frac{\partial}{\partial t} \left\{ [\varphi_{k}^{(0)}(r)]^* \varphi_{k, + \omega}^{(1)}(r) e^{i \omega t} + [\varphi_{k}^{(0)}(r)]^* \varphi_{k, - \omega}^{(1)}(r) e^{-i \omega t} + [\varphi_{k, + \omega}^{(1)}(r)]^* \varphi_{k}^{(0)}(r) e^{-i \omega t} + [\varphi_{k, - \omega}^{(1)}(r)]^* \varphi_{k}^{(0)}(r) e^{i \omega t} \right\} = e^{-i \epsilon_{k}^{(0)} t - i \lambda \Delta \epsilon_{k}^{(1)}(t)} \left\{ [\varphi_{k, + \omega}^{(1)}(r)]^* \varphi_{k}^{(0)}(r) \right. + \lambda \left( -\omega + \frac{\epsilon_{k}^{(0)}}{\omega} \right) \varphi_{k, + \omega}^{(1)}(r) e^{i \omega t} + \lambda \left( \omega + \frac{\epsilon_{k}^{(0)}}{\omega} \right) \varphi_{k, - \omega}^{(1)}(r) e^{-i \omega t} \left\} + O(\lambda^2). \tag{2.23}
\end{equation}
On the right-hand-side, we obtain
\begin{equation}
\left\{ \hat{H}_{KS}^{(0)}[n^{(0)}(r)] \varphi_{k}^{(0)}(r) + \lambda \hat{H}_{KS}^{(0)}[n^{(0)}(r)] \left[ [\varphi_{k, + \omega}^{(1)}(r)]^* \varphi_{k}^{(0)}(r) e^{i \omega t} + [\varphi_{k, - \omega}^{(1)}(r)]^* \varphi_{k}^{(0)}(r) e^{-i \omega t} \right] + \lambda \int d^3 r \left[ f_{Hxc}[n^{(0)}(r)] n^{(1)}(r, t) + v_{ext}(r, t) \right] \varphi_{k}^{(0)}(r) \right\}
\end{equation}
x $e^{-i \epsilon_{k}^{(0)} t - i \lambda \Delta \epsilon_{k}^{(1)}(t)} + O(\lambda^2). \tag{2.24}$

The first-order equation can now be written in matrix form by gathering terms proportional to the resonant part $e^{i \omega t - i \epsilon_{k}^{(0)} t - i \lambda \Delta \epsilon_{k}^{(1)}(t)}$ and to the anti-resonant part $e^{-i \omega t - i \epsilon_{k}^{(0)} t - i \lambda \Delta \epsilon_{k}^{(1)}(t)}$:
\begin{equation}
\begin{pmatrix}
\hat{H}_{KS}^{(0)} - \epsilon_{k}^{(0)} + \omega & 0 \\
0 & \hat{H}_{KS}^{(0)} - \epsilon_{k}^{(0)} - \omega
\end{pmatrix}
\begin{pmatrix}
\varphi_{k, + \omega}^{(1)} \\
\varphi_{k, - \omega}^{(1)}
\end{pmatrix}
= - \begin{pmatrix}
v_{Hxc, + \omega}^{(1)} + v_{ext, + \omega}^{(1)} - \epsilon_{k, + \omega}^{(1)} & \varphi_{k}^{(0)} \\
v_{Hxc, - \omega}^{(1)} + v_{ext, - \omega}^{(1)} - \epsilon_{k, - \omega}^{(1)} & \varphi_{k}^{(0)}
\end{pmatrix}, \tag{2.25}
\end{equation}
where
\[ v_{\text{Hxc},+\omega}(\mathbf{r}, t) = \int q^3 f_{\text{Hxc}}[n^{(0)}(\mathbf{r}, \mathbf{r}')] n^{(1)}_{+\omega}(\mathbf{r}', t), \quad (2.26) \]

\[ n^{(1)}_{\pm \omega}(\mathbf{r}, t) = \sum_k n_k \left\{ [\varphi_k^{(0)}(\mathbf{r})]^* \varphi_k^{(1)}(\mathbf{r}) + [\varphi_k^{(1)}(\mathbf{r})]^* \varphi_k^{(0)}(\mathbf{r}) \right\} e^{\pm i \omega t}, \quad (2.27) \]

and \( \varepsilon_{k, \pm \omega}^{(1)} \) is the Fourier transform of \( \frac{\partial}{\partial \omega} \Delta \varepsilon_k^{(1)}(t) \):
\[ \varepsilon_{k, \pm \omega}^{(1)} = \langle \varphi_k^{(0)}| v_{\text{Hxc}, \pm \omega}^{(1)} + v_{\text{ext}, \pm \omega}^{(1)}|\varphi_k^{(0)} \rangle. \quad (2.28) \]

In this form, the Sternheimer method looks like a set of linear equations, but in reality it is a nonlinear set of equations as the right-hand side depends on the solution through \( v_{\text{Hxc}, \pm \omega}^{(1)} \) which depends on \( n_{\pm \omega}^{(1)} \) and therefore on \( \varphi_{k, \pm \omega}^{(1)} \). The usual way is to introduce a self-consistent field (SCF) iteration for the response density \( n_{\pm \omega}^{(1)} \), as for the ground-state DFT problem. This is the essence of the Baroni-Gianozzi-Testa method \[22\], originally used for static perturbations but equally applicable to TDDFT \[8\].

By projecting the Sternheimer equation onto the unperturbed wavefunctions, one obtains the sum-over-states expression in second-order perturbation theory for the wavefunction:
\[ \varphi_{k, \omega}^{(1)} = \sum_{m \neq k} |\varphi_m^{(0)}\rangle \langle \varphi_m^{(0)}| \left( \frac{\hat{H}_\omega^{(1)}}{\varepsilon_m^{(0)} - \varepsilon_k^{(0)}} + \omega \right) \]
where \( \hat{H}_\omega^{(1)} = v_{\text{Hxc}, \omega}^{(1)} + v_{\text{ext}, \omega}^{(1)} \). Using the Sternheimer equation has the great advantage that it avoids the need for explicit calculation of the unoccupied states that would occur in this sum over states.

As with the time-propagation approach, we have to include an artificial lifetime. Otherwise, (i) the matrix is singular when \( \omega \) corresponds to the KS-eigenvalue difference \( \varepsilon_a - \varepsilon_i \) (an excitation in the independent particle picture), or (ii) the response becomes infinite when \( \omega \) corresponds to a resonance (an excitation in the interacting picture). The former is simply a numerical issue, but the later one has physical meaning and will be used to derive Casida’s equation in the next subsection. The artificial lifetime is introduced by multiplying the first-order wavefunction \( \varphi_k^{(1)}(\mathbf{r}, t) \) and the external potential \( v_{\text{ext}}(\mathbf{r}, t) \) by a decay term \( e^{-\eta t} \). In the first order, the matrix equation then reads

\[ \left( \begin{array}{cc} \hat{H}_\text{KS}^{(0)} - \varepsilon_k^{(0)} + \omega + i\eta & 0 \\ 0 & \hat{H}_\text{KS}^{(0)} - \varepsilon_k^{(0)} - \omega + i\eta \end{array} \right) \left( \begin{array}{c} \varphi_{k,+\omega}^{(1)} \\ \varphi_{k,-\omega}^{(1)} \end{array} \right) = -\hat{P}_c \left( \begin{array}{c} v_{\text{Hxc},+\omega}^{(1)} + v_{\text{ext},+\omega}^{(1)} \varphi_k^{(0)} \\ v_{\text{Hxc},-\omega}^{(1)} + v_{\text{ext},-\omega}^{(1)} \varphi_k^{(0)} \end{array} \right) \]  
\[ (2.30) \]

The matrix is no longer singular, and the resonances become Lorentzians with width \( \eta \) instead of delta functions. We also added a projector to the unoccupied space \( \hat{P}_c = 1 - \hat{P}_{\text{occ}} \), which orthogonalizes the KS response wavefunctions with respect to the occupied KS ground-state wavefunctions. The components of the response wavefunctions in the occupied
subspace are not needed because they cancel out in the density response. The projector avoids solving for these (possibly large) components, making the numerical solution more efficient and stable [19, 8]. It also simplifies the equation by removing the level shift \( \Delta_{\omega}^{(1)} \).

Finally, after the self-consistent solution is found, the linear density response is directly available from equation (2.27).

The Sternheimer equation is particularly suited to the calculation of higher-order responses, because solution of only the first-order problem can actually give access to up to third-order derivatives of the total energy (second-order response). In fact, the variational principle can be used to show that the \( \{ \phi_i^{(n)} \} \), the derivatives of order \( n \) of the KS wavefunctions with respect to a perturbation, can be used to construct all derivatives of the total energy up to order \( 2n + 1 \) (this is the famous \( 2n + 1 \) theorem for DFT [127]).

Consider a bare external perturbation \( \hat{H}_{\text{bare}}^{(n)} \) and a total perturbation \( \hat{H}^{(n)} \), including Hartree and exchange-correlation response (the “local fields” [153]). For \( n = 0 \), this theorem reduces to the familiar Hellman-Feynman Theorem, used in calculation of forces from only ground-state quantities:

\[
E^{(1)} = \frac{\partial E}{\partial \lambda} = \sum_{i}^{\text{occ}} \langle \phi_i | \hat{H}^{(1)}_{\text{bare}} | \phi_i \rangle
\]

At \( n = 1 \), the expression for the second derivative (first-order response) is

\[
E^{(2)} = \sum_{i}^{\text{occ}} \left[ \langle \phi_i^{(0)} | \hat{H}^{(1)} | \phi_i^{(1)} \rangle + \text{cc.} + \langle \phi_i^{(0)} | \hat{H}^{(2)}_{\text{bare}} | \phi_i^{(0)} \rangle \right]
\]

and for the third derivative (second-order response) are

\[
E^{(3)} = \sum_{i}^{\text{occ}} \left[ \langle \phi_i^{(1)} | \hat{H}^{(1)} | \phi_i^{(1)} \rangle + \langle \phi_i^{(1)} | \phi_i^{(1)} \rangle \langle \phi_i^{(0)} | \hat{H}^{(1)} | \phi_i^{(0)} \rangle \\
+ \langle \phi_i^{(1)} | \hat{H}^{(2)}_{\text{bare}} | \phi_i^{(0)} \rangle + \text{cc.} + \langle \phi_i^{(0)} | \hat{H}^{(3)}_{\text{bare}} | \phi_i^{(0)} \rangle \right] \\
+ \frac{1}{6} \int d^3r \int d^3r' \int d^3r'' \frac{\delta^3 E[n^{(0)}]}{\delta n(r) \delta n(r') \delta n(r'')} n^{(1)}(r)n^{(1)}(r')n^{(1)}(r'')
\]

where superscripts indicate the order of derivatives with respect to the perturbation [127].

The bare perturbation \( \hat{H}_{\text{bare}}^{(n)} \) is zero for \( n > 1 \) for electric fields, but not in general. The third functional derivative here is the second-order kernel \( K_{xc} \).

Conveniently, it turns out even in these equations, only the projection of the wavefunction derivatives onto the unoccupied subspace \( P_c \phi_i^{(1)} \) are required for this formula [84], making the Sternheimer calculation more efficient. With this formula for \( E^{(3)} \), the Sternheimer approach allows efficient access to phonon anharmonicities and nonlinear optical susceptibilities from solution of the first-order Sternheimer equation. This is true even for mixed derivatives with respect to perturbations in different directions or even entirely different perturbations. To get even higher orders, one can use the hierarchy of Sternheimer equations to solve for \( \{ \phi_i^{(n)} \} \) from results at lower orders [127, 123], with a somewhat more complicated calculation. The \( 2n + 1 \) formulae for the energy derivatives at higher orders are
straightforward but increasingly lengthy. For the time-dependent case, rather than total energies we use instead the action, or in the frequency domain, the Floquet quasi-energy \([201]\).

The \(2n + 1\) theorem actually also provides an alternate calculation approach for \(\varphi^{(1)}\) \([123, 76, 124, 19]\). The formula for \(E^{(3)}\) is variational with respect to \(\varphi^{(1)}\), just as \(E\) is variational with respect to \(\varphi^{(0)}\), as stated in the standard variational theorem of quantum mechanics. As a result, one can solve for \(\varphi^{(1)}\) by direct minimization of the functional \(E^{(3)}[\{\varphi^{(1)}\}]\). This approach is formally equivalent to solution by self-consistency, and the choice of technique is a question of numerical strategy.

### 2.2.3 Casida method

From the Sternheimer method, we can continue to Casida’s equation by writing the linear Sternheimer equation in the particle-hole basis, \(i.e.,\) in the Kohn-Sham orbital basis including unoccupied states. We will briefly summarize the result. First, we write the first-order response wavefunction as a linear combination of Kohn-Sham orbitals \(i.e.,\) sum-over-states expansion

\[
\varphi^{(1)}(r) = \sum_a c_a^{(k)} \varphi_a^{(0)}(r),
\]

where \(c_a^{(k)}\) is the projection of the response of \(k\)th occupied state onto the \(a\)th unperturbed KS state \(\varphi_a^{(0)}\). These coefficients represent excitations from state \(k\) to state \(a\). As in Eq. (2.30), we are considering only the projection of \(\varphi^{(1)}_k(r)\) into the unoccupied subspace, which will remove the level shift. We are assuming that the orbitals are real here for Casida’s equation. We can define

\[
K_{kk',bb'} = \int d^3r \int d^3r' f_{\text{Hxc}}[\psi^{(0)}](r, r') \varphi_k^{(0)}(r) \varphi_{b'}^{(0)}(r') \varphi_b^{(0)}(r) \varphi_{k'}^{(0)}(r)
\]

as the Hartree-exchange-correlation matrix element for interaction of excitations \(b \leftarrow k\) and \(b' \leftarrow k'\) \((k\) and \(k'\) are occupied orbitals, \(b\) and \(b'\) unoccupied). This term couples independent-particle excitations (KS-eigenvalue differences) to interacting-particle excitations (TD-KS transition frequencies).

It is used in the eigenvalue equation

\[
\Delta E^2 + 2\Delta E \frac{1}{2} N \frac{1}{2} K_N \frac{1}{2} \Delta E \frac{1}{2} = \omega^2 I,
\]

which is the well-known Casida’s equation with one small difference: instead of differences of occupation numbers \((n_a - n_i)\), the actual occupation numbers appear. This is a consequence of our assumption of integral occupations. The extension to fractional occupations will be discussed in the next subsection. \(\Delta E\) is the vector of KS energy differences, and \(N\) is the vector of occupation numbers.

The eigenvectors \(F\) of Casida’s equation, Eq. (2.36), can be used to extract the strength of the response to the external field. After some algebra, for example, the polarizability can be written as

\[
\alpha_{\gamma\delta}(\omega) = \mu_{\gamma} N \frac{1}{2} \Delta E \frac{1}{2} \sum_k F_k (\omega_k^2 - \omega^2)^{-1} F_k^* \Delta E \frac{1}{2} N \frac{1}{2} \mu_{\delta},
\]
where $\mu_\delta$ is the dipole-moment operator in direction $\delta$, with matrix elements in the particle-hole basis $\mu_{\delta, ai} = \langle \phi_a^{(0)} | r_\delta | \phi_i^{(0)} \rangle$.

### 2.2.4 Generalizations and discussion

In this subsection, we will discuss what changes if we do not make the assumptions of the beginning of the section. We start with the first assumption – the adiabatic approximation. Without the adiabatic approximation, the exchange-correlation functional has memory, i.e., the exchange-correlation functional depends on density at all previous times. In principle, it is trivial for the time-propagation method. We only have to store all previous densities and calculate the exchange-correlation potential from these. In practice, this is a very demanding task and often impossible beyond model systems.

In the Sternheimer method, memory will show up as a frequency dependence of the exchange-correlation kernels. At first order, the kernel depends only on one frequency, $f_{xc}[n_{GS}](r, r', \omega)$, but at higher orders it depends on multiple frequencies. Again, if explicit forms of the frequency-dependent kernels are known, it is straightforward to include memory (in principle). However, a practical implementation might not be easy and it will depend a lot on the actual form of the memory-dependence in the functionals, which remains an important unresolved theoretical issue [214].

In the case of the Casida method, the matrix becomes frequency-dependent [53], which means that the linear eigenvalue problem becomes a nonlinear eigenvalue problem. A nonlinear eigenvalue problem is much harder to solve than a linear eigenvalue problem (e.g., SCF iterations may be required).

If a system is spin-polarized, each spin has its own exchange-correlation potential $v_{xc}^\alpha$ and $v_{xc}^\beta$. The exchange-correlation kernel is replaced by three exchange-correlation kernels $f_{xc}^{\alpha\alpha}$, $f_{xc}^{\beta\beta}$, and $f_{xc}^{\alpha\beta}$ [53, 132].

If a system has fractional occupation numbers, an excitation which happens from one partially occupied state $i$ to another partially occupied state $j$ will have an opposite excitation (or de-excitation) from $j$ to $i$. The expressions from perturbation theory now contain the occupation difference $n_i - n_j$. The original presentation of the Casida method [53] shows this general case. The situation for the Sternheimer method is significantly more complicated due to the question of what happens to the projection onto the unoccupied subspace. A computational scheme has been derived to solve the Sternheimer equation when the occupation function corresponds to the thermal Fermi-Dirac distribution or one of the related smearing methods, which is generally needed for reasonable convergence of the ground state of metals with respect to $k$-point sampling [82]. Introduction of some extra projectors into the Sternheimer equation allows the density response

$$n^{(1)}(r) = \sum_{ij} \frac{\tilde{\theta}(\varepsilon_F - \varepsilon_i) - \tilde{\theta}(\varepsilon_F - \varepsilon_j)}{\varepsilon_i - \varepsilon_j} \varphi_i^*(r) \varphi_j(r) \langle \varphi_j | \hat{H}_{bare}^{(1)} | \varphi_i \rangle$$

(2.38)

to be written in the same form as Eq. (2.27) for the zero-temperature (semiconducting) case in which all states are full or empty, with the addition of terms corresponding to variation of the occupations and Fermi level in general [19]. I will derive a generalization of this scheme to the dynamic case and with arbitrary occupations in Chapter 3.
Finally, if we use \( k \)-points, the ground-state KS wavefunctions become complex and we cannot obtain the Casida’s equation (2.36). However, we can still obtain a similar eigenvalue equation [298].

### 2.3 Applications of linear response

Having reviewed different methods of obtaining response functions, we will now consider the different perturbations that can be studied and how their response functions relate to physical quantities of interest. Electric, magnetic, structural, and \( k \cdot p \) perturbations, as well as mixed perturbations, are commonly used to extract both static and dynamic response properties.

#### 2.3.1 Response to electric perturbations

We will begin by considering electric perturbations, because they give access to optical properties and account for the vast majority of applications of TDDFT. In molecules, the basic quantity is the polarizability \( \alpha \), defined as the response of the dipole to an electric field \( \mathcal{E} \), in the limit of zero applied field:

\[
\alpha_{ij}(\omega) = \frac{\partial \mu_{i\omega}}{\partial \mathcal{E}_{j\omega}} = -\frac{\partial^2 E}{\partial \mathcal{E}_{i,-\omega} \partial \mathcal{E}_{j,\omega}}
\]

where \( \omega \) denotes the frequency of the electric field. The cross-section for optical absorption (in the dipole approximation) can be calculated from the imaginary part:

\[
\sigma_{ij}(\omega) = \frac{4\pi\omega}{c} \text{Im} \alpha_{ij}(\omega)
\]

The static polarizability (which is purely real) is commonly calculated by finite differences of applied field [387], and the dynamic polarizability can be computed by time-propagation [403], typically via application of an instantaneous pulsed electric field, which contains all frequencies. A Fourier transformation of the resulting free oscillations of the dipole moments yields the polarizability. The absorption spectrum is most often calculated by the Casida method [53, 167], which was designed for this problem; it calculates excited states, and a specific perturbation only enters in the calculation of oscillator strengths. It can be difficult to converge the real part of the dynamic polarizability below the optical gap in this method [167], in which case it is more efficient to do the calculation via the Sternheimer equation [8]. The electric field appears as a term \( \mathcal{E} \cdot r \) in the Hamiltonian, so the perturbation used is \( \partial H/\partial \mathcal{E} = r \). This is the response of the dipole moment to a homogeneous electric field, which couples to the dipole, so these are called dipole-dipole polarizabilities. Similar methodologies can be used for dipole-quadrupole polarizabilities (response to a field gradient) and other multipoles [32].

For solids, typically the susceptibility \( \chi \) (polarizability per unit cell) and dielectric function \( \epsilon \) are used instead of the polarizability, related via

\[
\epsilon = 1 + 4\pi \chi = 1 + 4\pi \frac{\alpha}{V}
\]
where $V$ is the volume of the unit cell. The optical absorption is calculated just as for finite systems. There is a significant complication in applying a uniform electric field to a periodic system, because the operator $\mathbf{r}$ is not periodic. There are two ways to solve the problem: the original solution is to consider spatially modulated fields $E_q = E_0 e^{i \mathbf{q} \cdot \mathbf{r}}$, which are periodic although not necessarily commensurate with the cell periodicity. In this case, one can consider the $q \to 0$ limit to obtain the response to a homogeneous electric field, which is used in TDDFT calculations in the sum-over-states [153, 207] and response-function approaches commonly used for crystals [258, 308]. It is important to consider carefully the relation between microscopic/macrosopic and transverse/longitudinal responses in this method [38]. Another solution is to use the quantum theory of polarization [300, 378] to write the operator as $\mathbf{r} = i \frac{\partial}{\partial \mathbf{k}}$, which is periodic [124]. The $k$-point derivatives are obtained by finite differences, or again through perturbation theory. An equivalent approach is to calculate the polarization within a basis of Wannier functions [75]. In either method, we do not study the response of the dipole moment per unit cell, which is not a well-defined quantity, and instead use the polarization. To apply finite homogeneous electric fields in periodic systems, the electric-enthalpy approach can be used, in which a term $-\mu \cdot E$ is added to the total energy functional to be minimized [330].

Armed with the dielectric function or polarizability, one can calculate many interesting properties. Inverting the dielectric matrix yields the loss function $\text{Im} \epsilon^{-1}(\mathbf{q}, \omega)$, which describes the slowing of energetic electrons in a solid and is measured in electron energy-loss spectroscopy (EELS) experiments [260, 216]. Van der Waals interaction energies can be calculated too: the Hamaker coefficients in the expansion

$$\Delta E(R) = -\sum_{n=6}^{\infty} \frac{C_n}{R^n}$$

(2.42)

can be calculated from the Casimir-Polder relation as an integral over polarizabilities evaluated as a function of imaginary frequency. For example, the dominant $C_6$ term for interaction of molecules A and B is given by

$$C_6^{AB} = \frac{3}{\pi} \int_0^\infty du \alpha^{(A)}(iu) \alpha^{(B)}(iu)$$

(2.43)

Higher-order terms involve higher-order multipole polarizabilities. These coefficients have been calculated by TDDFT with molecular polarizabilities from time-propagation and Sternheimer methods, and surface susceptibilities from response functions, to study molecule-molecule [219] and molecule-surface interactions [36]. Scaled interatomic $C_6$ coefficients from TDDFT can also be used to add Van der Waals interactions into DFT total energies as a post-processing step [366].

Dielectric response can also be considered not for uniform fields but rather for point charges, giving $\epsilon(\mathbf{r}, \mathbf{r}', \omega)$ from a perturbation $1/|\mathbf{r} - \mathbf{r}'|$. Time-propagation has been used to study the spatially resolved plasmon response of liquid water [354]. This form of the dielectric function can also be used as input for many-body perturbation theory via the GW approximation [141] and Bethe-Salpeter equation. Typically these calculations use the RPA dielectric function, which is equivalent to using only the Hartree response and setting the kernel $f_{xc} = 0$. However, as pointed out in the first practical implementation of this
scheme [152], using instead the TDDFT $\epsilon$ is an approximate way of including the vertex $\Gamma$. This methodology has been used for quasiparticle and optical calculations on organic molecules [363]. Recently progress has been made in replacing the expensive sums over states with solution of the time-dependent Sternheimer equation [121], which can be done with RPA or including $f_{\text{sc}}$.

Response to a related but more exotic perturbation can also be used to parametrize the DFT+$U$ method, in which projectors on atomic-like orbitals are used to emulate Coulomb repulsion and correct the energies of localized $d$- and $f$-orbitals in strongly correlated materials [10]. \textit{Ab initio} values for $U$ can be calculated from the screened response to a localized potential $\alpha_I P_I$, where $P_I$ is an atomic-orbital projector, implemented via finite differences [66].

Response to electric perturbations can be used to calculate nonlinear susceptibilities, describing nonlinear optical properties such as second-harmonic generation, optical rectification, and electrooptical effects [322]. The hyperpolarizability $\beta$ of a molecule and second-order susceptibility $\chi^{(2)}$ of a solid are simply the derivatives with respect to field of $\alpha$ and $\chi$, the next order in the Taylor expansion of the dipole moment:

$$\mu_i(\omega) = \mu_{i0} + \alpha_{ij}(\omega) E_{j,\omega} + \frac{1}{2} \beta_{ijk}(\omega = \omega_1 + \omega_2) E_{j,\omega_1} E_{k,\omega_2} + ...$$

though conventions can differ on what numerical factors may appear in this expansion [395]. With the $2n + 1$ theorem, solution of the Sternheimer equation can be used to calculate molecular hyperpolarizabilities [8, 387] as well as the nonlinear susceptibilities of semiconductors with the quantum theory of polarization [76].

Finite differences are often also used to calculate static hyperpolarizabilities [387], and time-propagation can be used for dynamic hyperpolarizabilities; however, the advantage of being able to obtain the entire spectrum from a single calculation is lost, and separate calculations must be done for each set of input frequencies [348]. The response-function technique has also recently been developed for $\chi^{(2)}$ in the $q \rightarrow 0$ limit, and applied to second-harmonic generation in zincblende semiconductors [151].

### 2.3.2 Response to magnetic perturbations

Magnetic response offers a significant additional challenge compared to electric response because of the fact that the vector potential has to be formulated in a particular choice of gauge, which causes particular complications when localized-orbital bases or non-local pseudopotentials are used. The simplest quantity to consider is the magnetic susceptibility, the analogue of the electric susceptibility. The coupling in the Hamiltonian can be expressed with the vector potential $\mathbf{A}$, field $\mathbf{B} = \nabla \times \mathbf{A}$, and spin magnetic moment $g \mu_B \mathbf{S}$ (where $\mu_B$ is the Bohr magneton and $\mathbf{S}$ is the spin angular momentum), as

$$\hat{H} = \frac{1}{2} \left( \mathbf{p} + \frac{1}{c} \mathbf{A} \right)^2 + v + g \mu_B \mathbf{S} \cdot \mathbf{B}$$

$$= \hat{H}^{(0)} + \frac{1}{2c} (\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + \frac{A^2}{2c^2} + g \mu_B \mathbf{S} \cdot \mathbf{B}$$

(2.45)
The three perturbations are respectively the orbital paramagnetic, orbital diamagnetic, and spin paramagnetic contributions. Within the Coulomb gauge where $\nabla \cdot \mathbf{A} = 0$ and $\mathbf{p}$ and $\mathbf{A}$ commute, the linear coupling to $\mathbf{A}$ can also be written in terms of the orbital angular momentum as $\frac{1}{2} \mathbf{r} \times \mathbf{p} \cdot \mathbf{B} = \frac{1}{2} \mathbf{L} \cdot \mathbf{B}$. In spin-unpolarized systems, the spin susceptibility is zero, so just the orbital perturbation is needed. The Sternheimer equation has been used with this perturbation to calculate static susceptibilities for boron fullerene molecules [37]. There is actually the advantage, compared to electric perturbations, that the first-order response of the density is required to be zero if the ground state has time-reversal symmetry, which is the case in the absence of spin-polarization or magnetic fields, so that the magnetic Sternheimer equation does not require self-consistency [227].

To compute magnetic susceptibilities in solids, we have the same problem as for electric perturbations that the position operator is not periodic, which can similarly be handled either by taking the $\mathbf{q} \to 0$ limit or by the quantum theory of polarization. The $\mathbf{q} \to 0$ approach has been used for susceptibilities in crystals [227]. It has also been used for chemical shifts in nuclear magnetic resonance (NMR) [228, 273], which are ratios between the external field and the environment-dependent screened field at the position of the nuclei. The $g$-tensor of electron paramagnetic resonance (EPR), describing the direction-dependent spin susceptibility, has been calculated by a similar approach for radicals and defects, including spin-orbit and hyperfine effects [274]. The $J$ coupling between nuclear spins in NMR can also be computed by the Sternheimer equation, via the magnetic field induced at one nucleus by the field of another [174]. Susceptibilities can also be studied by applying finite magnetic fields, but in the presence of non-local pseudopotentials the coupling in the Hamiltonian generates additional terms beyond those above to satisfy gauge-invariance, as can be handled with the ICL [158] or GIPAW methods [275]. When using pseudopotentials, it is important to note that core susceptibilities may be significant, unlike the electric case; they may be computed from separate atomic calculations [227], or handled via projector-augmented wave (PAW) methods [273]. Gauge-invariance for magnetic fields in localized-orbital calculations also requires special attention, and can be handled by the “gauge-including atomic orbitals” or “individual gauge for localized orbitals” methods [314]. The susceptibilities of interest are usually static, and the NMR/EPR properties are treated as static since they are measured at radio frequencies.

To study spin waves in metals, dynamical susceptibilities have been calculated with the Sternheimer equation, where peaks in the spin susceptibility $\chi(q, \omega)$ show the magnon band structure [312]. Spinor wavefunctions are needed to allow spin rotations. Another kind of magnetic response is the spin-triplet optical excitation spectrum, inaccessible by electric perturbations which can only excite singlets. Time-propagation techniques have been used to calculate triplet states by applying an opposite kick to the up and down spins [259]. A dynamic response that combines electric and magnetic response is circular dichroism, also known as optical rotation, in which a chiral molecule responds differently to left and right circularly polarized light. The rotatory strength as a function of frequency can be studied via the (orbital) magnetic moment induced by an electric field; the reverse is possible but more complicated to implement. These properties have been calculated for organic molecules with both time-propagation and Sternheimer approaches [404, 381].
2.3.3 Response to structural perturbations

There is a rich field of study regarding the response to perturbation of ionic positions and lattice parameters. Since it has been reviewed in great detail [19], and is mostly concerned with static properties, we will consider only briefly most of these quantities and focus on those where TDDFT can be used. Forces on the ions and stresses on the unit cell (the diagonal part of which is the pressure) can be calculated via the Hellman-Feynman theorem, which is routinely done in static DFT for use in structural relaxation. These forces can additionally be used for calculation of dynamical matrices for vibrational properties by means of the frozen-phonon method, in which finite ionic displacements are used. However, only phonons with commensurate wavevectors $q$ can be calculated, and large supercells may be required. Using the Sternheimer equation has the great advantage that dynamical matrices at arbitrary $q$ may be calculated with effort comparable to that for zone-center phonons [19]. For displacement of ion $\alpha$ with potential $v_\alpha$ in direction $i$, the perturbation is $\partial v_\alpha/\partial R_{\alpha i}$.

The dynamical matrix is diagonalized to obtain phonon frequencies and eigenvectors in the harmonic approximation. This information, as function of cell volume, can also be used as input for the “quasi-harmonic approximation” which is used for free energies and other thermodynamic information about solids [390, 35, 52]. The phonon group velocities can be computed directly as analytic derivatives from the phonon perturbation calculation as well [127]. Going beyond the harmonic approximation, the $2\pi + 1$ theorem gives access to anharmonic properties from Sternheimer calculations [19]. Finite-difference calculations have been used to calculate mechanical anharmonicity and electrical anharmonicity (second-order derivatives of force and polarization with respect to ionic displacement) for ionic contributions to the nonlinear susceptibility [304]. Anharmonicities are needed for phonon linewidths in crystals, as well as to obtain vibrational frequencies in the presence of strong anharmonicity. Sternheimer phonon calculations also give the induced self-consistent potential, which is used to calculate the electron-phonon matrix elements between electronic states $i$ and $j$ and a phonon of wavevector $q$ and branch $\nu$:

$$g_{ij\nu}(k,q) = \left\langle \varphi_{ik} + q \left| \frac{dH}{d\lambda_{\nu}} \right| \varphi_{jk} \right\rangle$$

(2.46)

With Wannier-function-based interpolation schemes, the electron-phonon coupling has been used to calculate the superconducting properties of boron-doped diamond [122] and cuprates [120].

Phonons are generally calculated from static response, an adiabatic approximation which is well justified when the phonon frequency is much less than the electronic band gap. For metals however this condition is not satisfied, and the system may not remain in the electronic ground state during phonon oscillation. Truly dynamical, or non-adiabatic, phonon calculations have been done for doped graphene [203] and 2D intercalated compounds [310], showing significant corrections in these systems. A TDDFT sum-over-states perturbation expression is used to find the dynamical matrix at a given frequency, and self-consistently iterated until the input and output frequencies coincide.

Besides these lattice-dynamics methods, another method for vibrational calculations is molecular dynamics. The system is evolved in time at finite temperature, and from the ionic
trajectories, velocity autocorrelation functions are calculated, giving a power spectrum of vibrations as a function of frequency [4]. Such calculations can be done by empirical methods or \textit{ab initio} MD, commonly in the DFT-based Car-Parrinello scheme [50]. In systems such as liquids, the harmonic approximation fails completely and MD must be used to study the vibrational modes [284] and infrared spectra [324]. Recently a new fast Ehrenfest dynamics method has been developed, in which TDDFT is used to propagate the wavefunctions between timesteps. This allows more efficient calculation of vibrational properties of large systems [6, 9]; the method will be discussed in detail in [20]. TDDFT has also been used to study coherent excitation of phonons in Si by light, propagating the electronic system in the presence of an oscillating applied field, and analyzing the induced forces [323].

Bulk moduli and elastic constants can be calculated from the second derivative of the total energy with respect to pressure or shear, with finite differences [198] or the Sternheimer equation [21, 19]. The strain perturbation involves a stretching of both the unit cell and the wavefunctions, and takes the tensorial form [251]

\[ T_{ij} = p_i p_j - r_i \frac{\partial n_{\text{ext}}}{\partial r_j} \]  

(2.47)

It is somewhat complicated to implement since it is not lattice-periodic in this form, but it can also be formulated in a lattice-periodic manner in terms of metric tensors [135]. Second-order elastic coefficients and Grüneisen parameters (variation of phonon frequencies with stress) can also be calculated from the \(2n + 1\) theorem [127]. The chain rule must be used to include ionic as well as electronic contributions.

### 2.3.4 Mixed electric and structural response to structural perturbations

Raman spectroscopy measures vibrational frequencies by the energy gained or lost by a photon, and in the Placzek approximation the intensity of a vibrational peak is proportional to the Raman tensor [202], the derivative of the polarizability with respect to ionic displacement:

\[ \frac{\partial^3 E}{\partial \epsilon_i \partial \epsilon_j \partial R_{ks}} = \frac{\partial \alpha_{ij}}{\partial R_{ks}} \]  

(2.48)

For solids, the susceptibility \(\chi\) can be used instead. For off-resonant Raman, \textit{i.e.}, when the incident phonon is not resonant with an electronic excitation of the system, the static polarizability is generally used. There are many ways the response to the various perturbations could be calculated. Commonly finite differences are used for ionic response, with dielectric tensor calculated from sum over states [23], finite differences [304], or the Sternheimer equation [376]. Anharmonic Raman spectra of ice have been calculated with molecular dynamics by a Fourier transform of the autocorrelation function of the dielectric tensor \(\epsilon_\infty\).

Purely perturbative methods have also been developed. In an earlier approach applied to silica [202], the tensor is written as

\[ \frac{\partial^3 E}{\partial \epsilon_i \partial \epsilon_j \partial R_{ks}} = \int d^3r \frac{\partial^3 \rho}{\partial \epsilon_i \partial \epsilon_j} \frac{d\hat{H}}{dR_{ks}} \]  

(2.49)
The first-order perturbations are calculated by the Sternheimer equation, and the second-order electric derivatives of the density matrix $\rho$ are calculated via the second-order derivatives of the wavefunctions from a self-consistent sum-over-states expression. The $2n + 1$ theorem also makes it possible to do the computation from only first-order ionic and electric derivatives [382]. To study resonant Raman spectroscopy, dynamic polarizabilities must be used. This has been done in TDDFT with the complex polarization propagator approach to study the variation of the Raman spectrum with excitation energy for organic molecules [237]; this method uses a relaxation toward the ground state in the equations of motion to broaden resonances and prevent divergences.

Another mixed response is Born effective charges, which can be used to calculate LO-TO splitting [118], infrared spectra [263], and molecular dipole moments in liquids [264]. They are defined by

$$ Z_{\alpha ij} = \frac{\partial^2 E}{\partial \varepsilon_i \partial R_{\alpha j}} = \frac{\partial \mu_i}{\partial R_{\alpha j}} = \frac{\partial F_{\alpha j}}{\partial \varepsilon_i} \quad (2.50) $$

Born charges can be evaluated either as the response of the dipole moment (or polarization) in response to ionic displacement, which is natural in the context of a phonon calculation, or from the forces induced by an electric field, from the Sternheimer equation or finite differences [126].

A related quantity is the piezoelectric tensor $\gamma$, in which ionic displacement is replaced by strain $\varepsilon$:

$$ \gamma_{ijk} = \frac{\partial^2 E}{\partial \varepsilon_i \partial \varepsilon_{jk}} = \frac{\partial \mu_i}{\partial \varepsilon_{jk}} = \frac{\partial T_{jk}}{\partial \varepsilon_i} \quad (2.51) $$

The strain perturbation is not lattice-periodic, so piezoelectric tensors are most easily calculated by the stress $T$ induced by an electric field [83]. Both electronic and ionic contributions must be included.

A quite different quantity is the non-adiabatic coupling, which is used in molecular dynamics to govern the rate of hopping between the potential energy surfaces of the ground and excited states [375]. Going beyond the Born-Oppenheimer approximation, terms appear in the equation of motion containing $\langle \Phi_i | \partial / \partial R_{\alpha j} | \Phi_j \rangle$ (first-order coupling) and $\langle \Phi_i | \partial^2 / \partial R_{\alpha j}^2 | \Phi_j \rangle$ (second-order coupling) [143], with overlaps between many-body states $i, j$ and their derivatives with respect to ionic displacement. The Casida method can be used for first-order non-adiabatic couplings, analogously to the calculation of oscillator strengths but where the dipole operator is replaced with the ionic perturbation [149]. A time-propagation formulation has also been developed [16] for the calculation. The second-order couplings cannot be calculated by these methods, but are negligible in simple cases [143].

### 2.3.5 Response to $k \cdot p$ perturbations

Response to an infinitesimal shift of $k$-point in a solid, often referred to as $k \cdot p$ perturbation theory, can be used to give various properties. These are by necessity static, not dynamic. Since the perturbation is applied to an individual state rather than to the whole system, it does not have an associated density response. With the Hellman-Feynman theorem, band velocities can be calculated as

$$ v_{ik} = \frac{1}{\hbar} \nabla_k \varepsilon_{ik} = \frac{1}{\hbar} \langle u_{ik} | \nabla_k H_k | u_{ik} \rangle \quad (2.52) $$
where $u_{ik}$ is the periodic part of the Bloch function and $H_k$ is the effective Hamiltonian. The perturbation is

$$\nabla_k H_k = -i\nabla + k + [v_{\text{pseudo}}, r]$$

(2.53)

including a contribution from non-local pseudopotentials if they are used [303]. Second-order perturbation theory with a sum over states can similarly give effective masses [51, 409], as frequently used in simple models of band structures and transport in semiconductors. $k \cdot p$ perturbation theory has also been used, in a finite-difference framework, for $q \rightarrow 0$ limits in $GW$ [152] and Bethe-Salpeter [303] calculations. Additionally, $k \cdot p$ perturbations can be used to compute the $\nabla_k$ derivatives which are used in response calculations with the quantum theory of polarization [258, 378]. It is important, however, to note that perturbation theory cannot be used to compute the polarization itself, because it does not represent a consistent choice of gauge throughout the Brillouin zone [300].
Chapter 3

Density-functional perturbation theory with arbitrary fractional occupations

3.1 Introduction

Density-functional perturbation theory (DFPT) is frequently used to calculate response properties of metals. The original formulation is valid only for integer occupations, but it can be used also with a smearing occupation function via an approach due to De Gironcoli, which adds extra projection operators to the Sternheimer equation which are functions of the smearing. However, this formulation is only applicable to occupation functions which satisfy certain criteria characteristic of a smearing function. I show the detailed derivation of this scheme, and its appropriate form for dynamic as well as static response, and then present a generalization of this scheme to arbitrary fractional occupations.

3.2 Occupations from smearing

We begin by reviewing the approach of De Gironcoli [82], which was derived for lattice dynamics from density-functional perturbation theory [19], and generalizing it to time-dependent perturbations via time-dependent density-functional theory (TDDFT). Consider a system whose Kohn-Sham states have occupations $f_i$ and satisfy $H^{(0)} \psi_i = \epsilon_i \psi_i$, subjected to a perturbation $H^{(1)} e^{i \omega t} + H^{(1)} e^{-i \omega t}$ where $H^{(1)} = [H^{(1)}]^*$. The density response, according to standard perturbation theory, is

$$
\nu^{(1)}(\omega) = \sum_i \sum_{j \neq i} (f_i - f_j) \left[ \frac{\psi_i^* \psi_j \langle \psi_j H^{(1)}_{-\omega} \psi_i \rangle}{\epsilon_i - \epsilon_j + \omega} + \frac{\psi_i^* \psi_j \langle \psi_j H^{(1)}_{\omega} \psi_i \rangle}{\epsilon_i - \epsilon_j - \omega} \right]
$$

(3.1)
Since the \( i, j, \omega \) and \( j, i, -\omega \) parts are complex conjugates, we can write the sum also as

\[
n^{(1)}_\omega = 2\text{Re} \sum_i \sum_{j > i} (f_i - f_j) \left[ \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{-\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j + \omega} + \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j - \omega} \right] \\
(3.2)
\]

(Note that this result does not strictly hold if \( \omega \) has an imaginary part \( i\eta \), as used to get correct poles. However, with \( \eta > 0 \), all results are incorrect anyway except in the limit \( \eta \to 0 \).) If \( i \) and \( j \) are both fully occupied, then \( f_i - f_j = 1 - 1 = 0 \). Similarly, if \( i \) and \( j \) are both fully unoccupied, \( f_i - f_j = 0 - 0 = 0 \). As a result, we only need terms in the sum where \( f_i > 0 \) and \( f_j < 1 \), which we will term \( i \) occupied and \( j \) partially unoccupied.

\[
n^{(1)}_\omega = 2\text{Re} \sum_i \sum_{j > i} (f_i - f_j) \left[ \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{-\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j + \omega} + \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j - \omega} \right] \\
(3.3)
\]

Now let us specialize to a situation where the occupations are described by a smearing function \( \bar{\theta} \) (resembling the Heaviside function \( \theta \) which is zero for \( x < 0 \) and 1 for \( x > 1 \)). \( f_i = \bar{\theta} \left( \frac{\epsilon_i - \epsilon_F}{\sigma} \right) \) where \( \epsilon_F \) is the Fermi energy and \( \sigma \) is a smearing width. This may be abbreviated as \( f_i = \bar{\theta}_{F,i} \) with \( \bar{\theta}_{i,j} = \bar{\theta} \left( \frac{\epsilon_i - \epsilon_j}{\sigma} \right) \). The De Gironcoli scheme is applicable whenever \( f_i \) can indeed be described by such a function (requiring that occupations are a function of energy and do not depend explicitly on band index or \( k \)-point), and when \( \bar{\theta} \) satisfies these properties:

1. \( \bar{\theta} (x) + \bar{\theta} (-x) = 1 \).
2. \( \bar{\theta} (x) \leq e^x \).

These properties are satisfied by standard smearing schemes, such as Fermi-Dirac (using the thermal Fermi-Dirac distribution) [231], Gaussian [115], spline [146], Methfessel-Paxton [232], convolution of Gaussian and Fermi-Dirac [386], as well as when \( \bar{\theta} = \theta \), giving integral occupations as appropriate for a system with a gap. (Note that “cold smearing” [226, 224] does not satisfy (1). We will mention an approach used to handle it in Sec. 3.4.) Fermi-Dirac is asymptotically \( \bar{\theta} (x) \sim e^x \) as \( x \to -\infty \), while the other schemes have \( \bar{\theta} (x) \sim e^{x^2} < e^x \). We do not require \( \theta \) to be monotonic because the commonly used Methfessel-Paxton and cold smearing schemes do not necessarily satisfy this property, having excursions above 1 and below 0 in a range around 0, but return to the asymptotic behavior of 1 for \( x \to -\infty \) and 0 for \( x \to \infty \).

In terms of the smearing function, we have

\[
n^{(1)}_\omega = 2\text{Re} \sum_{i} \sum_{j > i} (\bar{\theta}_{F,j} - \bar{\theta}_{F,i}) \left[ \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{-\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j + \omega} + \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j - \omega} \right] \\
(3.4)
\]

We can turn this back into an unrestricted double sum on \( i, j \) using property (1) to partition arbitrarily the \( i, j \) term into separate \( i, j \) and \( j, i \) parts:

\[
n^{(1)}_\omega = 2\text{Re} \sum_i \sum_{j \neq i} \bar{\theta}_{j,i} \left[ \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{-\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j + \omega} + \frac{\psi^*_i \psi_j \langle \psi_j | H^{(1)}_{\omega} | \psi_i \rangle}{\epsilon_i - \epsilon_j - \omega} \right] \\
(3.5)
\]
We would like to have an expression in terms of a single sum

\[ n^{(1)}_\omega = \sum_{i} \left[ \psi_i^* \Delta \psi_{i,\omega} + \Delta \psi_{i,-\omega}^* \psi_i \right] \]  

(3.6)

as in DFPT for integer occupations. This can be done if \( \Delta \psi_i \) satisfies a certain modified Sternheimer equation:

\[
\begin{pmatrix}
H(0) - \epsilon_i \pm \omega + \sum_j \alpha_j |\psi_j\rangle \langle \psi_j| \\
\end{pmatrix} |\Delta \psi_{i,\pm\omega}\rangle = - \begin{pmatrix}
\tilde{\theta}_{F,i} - \sum_j \beta_{i,j} |\psi_j\rangle \langle \psi_j| \\
\end{pmatrix} H^{(1)}_{\pm\omega} |\psi_i\rangle
\]

(3.7)

The coefficients \( \alpha_j \) may be chosen fairly freely, subject to certain conditions we will discuss later. The choice of \( \alpha_j \), however, determines the coefficients \( \beta_{i,j} \), which serve to orthogonalize the response wavefunction to the occupied wavefunctions. For integer occupations, they would be simply equal to 1 for occupied states and 0 for unoccupied states, so that \( \tilde{\theta}_{F,i} - \sum_j \beta_{i,j} |\psi_j\rangle \langle \psi_j| \) is simply a projector onto the unoccupied states [8]. With smearing, however, we have to solve a linear equation to find the appropriate values for \( \beta_{i,j} \) that ensure that equation 3.6 is satisfied. We project equation 3.7 onto state \( m \):

\[
\langle \psi_m| \begin{pmatrix}
H(0) - \epsilon_i \pm \omega + \sum_j \alpha_j |\psi_j\rangle \langle \psi_j| \\
\end{pmatrix} |\Delta \psi_{i,\pm\omega}\rangle
\]

\[= - \langle \psi_m| \begin{pmatrix}
\tilde{\theta}_{F,i} - \sum_j \beta_{i,j} |\psi_j\rangle \langle \psi_j| \\
\end{pmatrix} H^{(1)}_{\pm\omega} |\psi_i\rangle \]

(3.8)

\[
\langle \psi_m| \epsilon_m - \epsilon_i \pm \omega + \alpha_m |\Delta \psi_{i,\pm\omega}\rangle = - \left( \tilde{\theta}_{F,i} - \beta_{i,m} \right) \langle \psi_m| H^{(1)}_{\pm\omega} |\psi_i\rangle \]

(3.9)

\[
\langle \psi_m| \epsilon_m - \epsilon_i \pm \omega + \alpha_m |\Delta \psi_{i,\pm\omega}\rangle = - \left( \tilde{\theta}_{F,i} - \beta_{i,m} \right) \langle \psi_m| H^{(1)}_{\pm\omega} |\psi_i\rangle \]

(3.10)

\[
\langle \psi_m| \Delta \psi_{i,\pm\omega}\rangle = \frac{\beta_{i,m} - \tilde{\theta}_{F,i}}{\epsilon_m - \epsilon_i \pm \omega + \alpha_m} \langle \psi_m| H^{(1)}_{\pm\omega} |\psi_i\rangle \]

(3.11)

Now, comparing equations 3.5 and 3.6, we need

\[
\langle \psi_m| \Delta \psi_{i,\pm\omega}\rangle = \frac{\tilde{\theta}_{F,i} - \tilde{\theta}_{F,m}}{\epsilon_i - \epsilon_m \pm \omega} \langle \psi_m| H^{(1)}_{\pm\omega} |\psi_i\rangle \]

(3.12)
so we equate the coefficients:

\[
\frac{\beta_{i,m} - \tilde{\theta}_{F,i}}{\epsilon - \epsilon_i \pm \omega + \alpha_m} = \tilde{\theta}_{i,m} - \tilde{\theta}_{F,m}
\]

(3.13)

\[
\beta_{i,m} = \tilde{\theta}_{F,i} + (\epsilon - \epsilon_i \pm \omega + \alpha_m)\frac{\tilde{\theta}_{F,i} - \tilde{\theta}_{F,m}}{\epsilon_i - \epsilon_m \mp \omega} \tilde{\theta}_{m,i}
\]

(3.14)

\[
\beta_{i,m} - \tilde{\theta}_{F,i} = (\epsilon - \epsilon_i \pm \omega + \alpha_m)\frac{\tilde{\theta}_{F,i} - \tilde{\theta}_{F,m}}{\epsilon_i - \epsilon_m \mp \omega} \tilde{\theta}_{m,i}
\]

(3.15)

\[
\beta_{i,m} = \tilde{\theta}_{F,i} + (\epsilon - \epsilon_i \pm \omega + \alpha_m)\frac{\tilde{\theta}_{F,i} - \tilde{\theta}_{F,m}}{\epsilon_i - \epsilon_m \mp \omega} \tilde{\theta}_{m,i}
\]

(3.16)

\[
\beta_{i,m} = \tilde{\theta}_{F,i}(1 - \tilde{\theta}_{m,i}) + \tilde{\theta}_{F,m} \tilde{\theta}_{m,i} + \alpha_m\frac{\tilde{\theta}_{F,i} - \tilde{\theta}_{F,m}}{\epsilon_i - \epsilon_m \mp \omega} \tilde{\theta}_{m,i}
\]

(3.17)

\[
\beta_{i,m} = \tilde{\theta}_{F,i} \tilde{\theta}_{i,m} + \tilde{\theta}_{F,m} \tilde{\theta}_{m,i} + \alpha_m\frac{\tilde{\theta}_{F,i} - \tilde{\theta}_{F,m}}{\epsilon_i - \epsilon_m \mp \omega} \tilde{\theta}_{m,i}
\]

(3.18)

Thus, for any choice of \( \alpha_j \), the correct density response will be obtained if this relation is used as the definition of \( \beta_{i,j} \). Note that in the last step we used property (1).

We want not only the correct density response, but also to be able to obtain it without use of unoccupied states. This means that projectors onto the unoccupied states should have zero coefficients so they are not needed for the computation of the variation of the occupied states. The second requirement immediately means that \( \alpha_j \) should be zero for unoccupied states. On the other hand, \( \alpha_j \) serves to make the left-hand side nonzero for the projection of the response wavefunction onto its ground-state wavefunction, enforcing the normalization convention \( \langle \Delta \psi_i | \psi_i \rangle = 0 \) and preventing a singularity in the equation. \( \alpha_j \) must be nonzero for occupied states to achieve this. By properties (2) and (3), the smearing function must be insignificant above some energy, so we can choose a cutoff. One choice offered by De Gironcoli is to use \( \alpha_j = \max (\epsilon_F + 3\sigma - \epsilon_j, 0) \). States less than \( 3\sigma \) above the Fermi level have significant occupation and have nonzero \( \alpha_j \), while states above that point have occupation close to zero and can be assigned \( \alpha_j = 0 \) and excluded from the sum. However, any other choice that is nonzero for (partially) occupied states and zero for completely unoccupied states is also usable here.

The other requirement to avoid use of unoccupied states is that we should not need to solve for the variation of unoccupied states; i.e. if \( \tilde{\theta}_{F,i} \approx 0 \), then \( \Delta \psi_{i,\pm \omega} \approx 0 \). This in combination with the requirement that \( \beta_{i,j} = 0 \) for unoccupied \( j \) means we need \( \beta_{i,j} = 0 \) if either \( i \) or \( j \) is unoccupied. For integer occupations, it is easy to see that \( \beta_{i,j} = 1 \) for occupied \( j \) and 0 for unoccupied \( j \). It is a somewhat more complicated to show this with smearing. If \( j \) is unoccupied, then \( \tilde{\theta}_{F,j} = 0 \) and \( \alpha_j = 0 \), leaving \( \beta_{i,j} = \tilde{\theta}_{F,i} \tilde{\theta}_{i,j} \). If \( i \) is unoccupied too, then \( \tilde{\theta}_{F,i} = 0 \), and \( \beta_{i,j} = 0 \). If \( i \) is occupied, then we use property (2):

\[
\beta_{i,j} = \tilde{\theta}_{F,i} \tilde{\theta}_{i,j} < e^{(\epsilon_F - \epsilon_i)/\sigma} e^{(\epsilon_i - \epsilon_j)/\sigma} = e^{(\epsilon_F - \epsilon_j)/\sigma} = \tilde{\theta}_{F,j} \approx 0
\]

(3.19)

By a corresponding argument, if \( j \) is occupied but \( i \) is not,

\[
\beta_{i,j} = \tilde{\theta}_{F,j} \tilde{\theta}_{j,i} + \alpha_j \frac{-\tilde{\theta}_{F,j}}{\epsilon_i - \epsilon_j \mp \omega} \tilde{\theta}_{j,i} \approx -\frac{\alpha_j}{\epsilon_i - \epsilon_j \mp \omega} \tilde{\theta}_{F,j} \tilde{\theta}_{j,i}
\]

(3.20)
The product \( \tilde{\theta}_{F,i} \tilde{\theta}_{j,i} \approx 0 \) again and will make \( \beta_{i,j} \approx 0 \) unless the prefactor is huge. This could only happen if the denominator becomes very small. An apparent divergence on resonance with Kohn-Sham eigenvalue differences is indeed a problem in the original sum over states of Eq. 3.1. However, the XC kernel in \( H^{(1)} \) cancels these poles and creates instead poles at the true resonances [260]. Therefore we need not worry about whether \( \beta_{i,j} \) is nonzero in these cases since the overall contribution to the response is zero.

A final point we must check is that the “self-term” \( i = j \) which was excluded from the double sum does not appear here; i.e. \( \langle \psi_i | \Delta \psi_{i,\pm \omega} \rangle = 0 \). We have shown already that the contribution is zero for unoccupied states, so consider an occupied state \( i \).

\[
(\epsilon_i - \epsilon_i \pm \omega + \alpha_i) \langle \psi_i | \Delta \psi_{i,\pm \omega} \rangle = - \left( \tilde{\theta}_{F,i} - \beta_{i,i} \right) \langle \psi_i | H^{(1)}_{\pm \omega} | \psi_i \rangle
\]  
(3.21)

By property (1), \( \tilde{\theta}_{i,i} = \frac{1}{2} \).

\[
(\pm \omega + \alpha_i) \langle \psi_i | \Delta \psi_{i,\pm \omega} \rangle = - \left( \tilde{\theta}_{F,i} - \frac{\tilde{\theta}_{F,i} - \tilde{\theta}_{F,i} - \beta_{i,i} - \beta_{i,i}}{\epsilon_i - \epsilon_i + \omega} \right) \langle \psi_i | H^{(1)}_{\pm \omega} | \psi_i \rangle
\]  
(3.22)

For \( \omega \neq 0 \), this is clearly zero. However, in the static case, the fraction does go to a finite limit when smearing is nonzero. This limit can be used to represent a contribution from changes in occupation near the Fermi level due to the perturbation, when operating in an ensemble of constant number of electrons [19]. In a situation where the occupations are fixed and cannot be altered by the perturbation, this term should not be included.

### 3.3 Arbitrary occupations

Let us now consider the case of arbitrary occupations. If they are all integral, then there is nothing special that must be done: the semiconducting formulation is applicable even if the occupations are not defined by a consistent Fermi level. If there are fractional occupations, however, a new approach is needed. The definitions of the various quantities in De Gironcoli’s formulation can be generalized to this case in a way that still satisfies all the needed properties:

1. \( \tilde{\theta}_{F,i} \rightarrow f_i \).
2. \( \tilde{\theta}_{i,j} \rightarrow \frac{f_j}{f_i + f_j} \).
3. \( \alpha_i \rightarrow E_0 f_i \).
Terms due to changes in occupations do not appear naturally here as we are assuming fixed occupations that are not necessarily derivable from any rule. Therefore, such terms must be added by hand if they are required for the particular problem in question, based on the knowledge of the particular functional form that the occupations are supposed to satisfy (e.g. different quasi-Fermi levels for valence and conduction bands). For two unoccupied states, the denominator in $\tilde{\theta}_{i,j}$ is zero, so we use the limit $\tilde{\theta}_{i,j} = 0$ in this case to avoid dividing by zero.

The first replacement is simply using a list of occupations rather than having them given for each state by its energy and the smearing function. The second manifestly satisfies the condition $\tilde{\theta}_{i,j} + \tilde{\theta}_{j,i} = 1$ analogously to the property (1) of the smearing functions. The third provides that $\alpha_i$ is nonzero for occupied states and zero for unoccupied states, as in De Gironcoli’s scheme. Note that since $\alpha_i$ has units of energy, we must multiply the occupations by some energy for correct dimensions; the value of the constant $E_0$ is irrelevant, and we will choose $E_0 = 1$ Ha for convenience, which means that in atomic units $\alpha_i$ has the same numerical value as $f_i$.

We also need $\tilde{\beta}_{i,j} = 0$ if $i$ or $j$ is unoccupied. The first two terms are clearly zero:

$$\tilde{\theta}_{F,j}\tilde{\theta}_{j,i} = \tilde{\theta}_{F,i}\tilde{\theta}_{i,j} = \frac{f_if_j}{f_i + f_j} = 0$$

leaving only the third term:

$$\tilde{\beta}_{i,j} = f_j \frac{f_i - f_j}{\epsilon_i - \epsilon_j + \omega} \frac{f_i}{f_i + f_j}$$

Again this is zero if either occupation is zero, with the exception only of a divergent denominator, just as in the smearing case. The argument for the self-term is identical to the smearing case. Thus these choices of the coefficients have all the needed properties.

### 3.4 Implementation

I have implemented this approach for arbitrary occupations in the Octopus code [218, 55, 7] for TDDFT response with the Sternheimer equation, and also implemented the dynamic case of the De Gironcoli scheme, extending the implementation by Miguel Marques of the static case.

My scheme, like De Gironcoli’s, reduces to the standard integer-occupation equations when all states are full or empty. It also is applicable to the case of smearing, and will give identical results; however, since there was some freedom in making choices that gave the correct result, using my scheme with smearing is not quite the same in implementation as De Gironcoli’s. For numerical tests of a Na$_2$ dimer, treated as a finite system, with a large smearing width, it is found that we can obtain identical results for the polarizability with my scheme and De Gironcoli’s, in the dynamic and static case. However, my scheme actually takes fewer SCF iterations. This may be because my scheme cuts off some contributions involving unoccupied bands faster as a function of occupation whereas the smearing relies on the cancellation of these terms in the sum for the density response. The generality of this observation needs to be studied more.
I compared in detail my implementation to what is done in Quantum Espresso [119], whose PHONON code is one of the most used for (static) DFPT calculations. In fact, this code always uses a Gaussian function for \( \tilde{\theta}_{i,j} \), no matter which smearing scheme is being used. This is presumably to avoid the difficulty that the recommended version of “cold smearing” [224] does not actually satisfy required property (1) of the smearing distribution. Any choice that satisfies the property is acceptable, however, provided the product in \( \beta_{i,j} \) still goes to zero in the way discussed above. At any rate, this code follows the approach of De Gironcoli and thus cannot handle arbitrary occupations.

The Abinit code [125] also implements DFPT, not via the Sternheimer equation but via the \( 2n + 1 \) Theorem variational approach [123, 124] (as discussed in Sec. 2.2.2). The implementation for the smearing case does not appear to be described in any publication. However, examination of the code (version 6.6.2, routine occeig) shows that it involves a factor akin to \( \tilde{\theta}_{i,j} \) calculated according to

\[
\tilde{\theta}_{i,j} = \begin{cases} 
1 & f_i < f_j \\
0 & f_j < f_i \\
1/2 & f_i \approx f_j
\end{cases}
\]  

(3.27)

where “\( \approx \)” refers to testing for equality within a small tolerance \( 10^{-5} \). This approach appears to be able to handle arbitrary occupations, provided they are generated appropriately in the ground-state calculation [246]. Nevertheless, this method may be less efficient numerically than mine in that the factor \( \tilde{\theta}_{i,j} \) is very sensitive to the exact values of the occupations, in a way that my prescription is not. Results for linear-response phonons of photoexcited bismuth show increased deviation from frozen-phonon frequencies as a function of excitation density, which may be related to this issue, as well as the matter of Fermi-surface sampling mentioned by the authors [246]. Additionally, the tutorial for this code notes a problem in convergence of elastic constants when using cold smearing, and this could indicate undesirable numerical features of the implementation.\(^1\)

### 3.5 Conclusion

This method is generally applicable to any occupations, integer or fractional, equilibrium or non-equilibrium, in DFT calculations. It can be used for diverse situations such as atoms or molecules with a degenerate ground state, non-equilibrium situations such as a semiconductor with photoexcited carriers [352, 246], fractional occupations corresponding to excitations from the ground-state-like second-order constrained variational theory (“CV(2)”) [417], or even as a substitute for the De Gironcoli approach for improved numerics with smearing.

An open question for the approach presented here, as well as that of De Gironcoli, is whether it is applicable for nonlinear response with the \( 2n + 1 \) Theorem. The first-order density response is correct, by construction, and hence any response functions based on it, such as the dielectric function. However, it is not clear that second-order response functions will be correct, since the quantity \( \Delta \psi_{i,\omega} \) in this approach is not simply the derivative of the wavefunction \( \psi_{i,\omega}^{(1)} \), and is rather, roughly speaking, \( f_i \psi_{i,\omega}^{(1)} \).

\(^{1}\)http://www.abinit.org/documentation/helpfiles/for-v5.7/tutorial/lesson_elastic.html
Chapter 4

Basis-set effects on the hyperpolarizability of CHCl$_3$: Gaussian-type orbitals, numerical basis sets and real-space grids

This work is based on Ref. [387].

4.1 Abstract

Calculations of the hyperpolarizability are typically much more difficult to converge with basis set size than the linear polarizability. In order to understand these convergence issues and hence obtain accurate \textit{ab initio} values, we compare calculations of the static hyperpolarizability of the gas-phase chloroform molecule (CHCl$_3$) using three different kinds of basis sets: Gaussian-type orbitals, numerical basis sets, and real-space grids. Although all of these methods can yield similar results, surprisingly large, diffuse basis sets are needed to achieve convergence to comparable values. These results are interpreted in terms of local polarizability and hyperpolarizability densities. We find that the hyperpolarizability is very sensitive to the molecular structure, and we also assess the significance of vibrational contributions and frequency dispersion.

4.2 Introduction

Chloroform (CHCl$_3$) is a widely used solvent in measurements of nonlinear optical properties of organic chromophores, using techniques such as electric-field-induced second-harmonic generation (EFISH) and hyper-Rayleigh scattering (HRS) [176, 79]. It is sometimes also used as an internal reference [65]. However, assumptions have to made to extract molecular hyperpolarizabilities from these measurements, in particular from EFISH which actually measures a third-order response function. For calibration purposes, either absolute measurements or \textit{ab initio} calculations are needed to convert between the different combi-
nations of tensor components of the hyperpolarizability measured in the EFISH and HRS experiments. However, very early calculations, which have been recognized as unsatisfactory by their authors [181], have heretofore been used for such conversions [176]. Consequently there is a need for better understanding of the convergence issues and for more accurate ab initio calculations. Toward this end we have carried out a systematic study of the second-order hyperpolarizability $\beta$ of chloroform using several theoretical methods. In an effort to obtain high-quality results which improve on earlier calculations, we have based our calculations on an accurate experimental structure and also considered both frequency-dependence and vibrational contributions, effects which are typically neglected in other calculations.

Although our presentation is restricted to a single system, the methodology is more general and hence many of our results will likely be applicable to many other cases. Likewise, the quantitative comparison of several theoretical methods is both novel and of general interest. Finally we believe our interpretation of the linear and nonlinear response in terms of local polarizability densities provides a useful way of understanding the local contributions to the polarizability from various parts of a molecule.

Chloroform is of particular theoretical interest because its hyperpolarizability is challenging to measure experimentally due its small magnitude compared to typical experimental errors, and hence the available measurements have both positive and negative values, with large relative error bars [176, 234]. Similarly, this nonlinear property has proved to be quite difficult to calculate theoretically, as the results exhibit a large dependence on the quality of the basis set used, both for DFT and coupled-cluster methods [81, 136], as well as the molecular geometry. One of the main purposes of this paper is to investigate the reasons for these difficulties using three different basis set approaches: i) Gaussian-type orbitals (GTOs); ii) numerical basis sets; and iii) real-space grids, with a common treatment of exchange and correlation given by the PBE [269] functional. The importance of different aspects of the basis sets (diffuseness, polarization, etc.) was studied systematically by changing the number of GTOs, the cutoff radii of the numerical basis sets, and the extent and density of the real-space grids. In order to interpret these results, we also studied the spatial distribution of the dielectric properties using the concepts of polarizability- and hyperpolarizability-densities, as well as first- and second-order electric-field-perturbed densities [64, 248].

We also briefly discuss the dependence of the results on the molecular geometry, which was found to have a significant influence on the calculated hyperpolarizability.

Unless noted explicitly, the experimental molecular geometry of Colmont et al. [68] was used throughout this work: $r_{CH} = 1.080 \text{ Å}$, $r_{CCl} = 1.760 \text{ Å}$ and $\angle \text{HCCl} = 108.23^\circ$. The molecule was located with its center of mass at the origin, and oriented with the CH bond along the positive z-direction and one HCCl angle in the $yz$-plane. Since chloroform has $C_{3v}$ point-group symmetry, the following symmetry relations hold for the linear polarizability $\alpha$ and hyperpolarizability $\beta$: $\alpha_{xx} = \alpha_{yy}$, $\beta_{xxy} = -\beta_{yyy}$ and $\beta_{xxz} = \beta_{yyz}$. In the static case Kleinman symmetry [188] also applies. Thus the $\alpha_{yy}$, $\alpha_{zz}$, $\beta_{yy}$, $\beta_{yyz}$ and $\beta_{zz}$ components fully describe the polarizability and hyperpolarizability tensors; all other permutations of the indices are equivalent. In the dynamic case at non-zero frequency, however, the components of $\beta_{ijk} (-2\omega; \omega, \omega)$ are not all equivalent: $\beta_{yyz} = \beta_{yzy} \neq \beta_{zyy}$. Here we use the Taylor convention for hyperpolarizabilities [395].
From our calculated tensor components we can calculate the isotropically averaged polarizability \( \bar{\alpha} = \frac{1}{3} \sum_i \alpha_{ii} \), the second-order hyperpolarizability coefficient, and the EFISH hyperpolarizability in the direction of the dipole moment \( \beta_{||} = \frac{1}{3} \sum_j (\beta_{jj} + \beta_{jj} + \beta_{jj}) \). In the \( C_{3v} \) point group, these relations reduce to \( \bar{\alpha} = \frac{1}{3} (2 \alpha_{yy} + \alpha_{zz}) \), \( \beta_{||} = \frac{3}{5} (2 \beta_{yyz} + \beta_{zzz}) \). We also calculate the hyperpolarizability for hyper-Rayleigh scattering in the VV polarization, as given by Cyvin et al. [74] for the static case (where Kleinman symmetry holds) and the generalization of Bersohn et al. [29] for the dynamic case. In the static case for \( C_{3v} \) symmetry

\[
[\beta_{\text{HRS}}^{VV}]^2 = \frac{8}{35} \beta_{yy}^2 + \frac{1}{7} \beta_{zzz}^2 + \frac{24}{35} \beta_{yyz}^2 + \frac{12}{35} \beta_{yyz} \beta_{zzz}.
\] (4.1)

This quantity has only been measured for liquid chloroform [176]; measurements are not available for the gas phase.

Our best results for each method generally exhibit a consistent agreement among themselves and lend confidence to the overall quality of our calculations compared to earlier work. Achieving this consistency points to the need for a comprehensive and well balanced description of all regions of the system: namely, the outlying regions of the molecule, the short C-H bond, and the Cl atoms. A key finding is that the local contributions to the \( \beta_{zzz} \) response of the Cl atoms and the C-H bond are of opposite sign and tend to cancel, thus explaining the relatively slow convergence of this component with respect to the basis set size. This behavior, together with the near cancellation of the \( \beta_{yyz} \) and \( \beta_{zzz} \) components, leads to the relatively small value of \( \beta_{||} \) of chloroform. By contrast, the HRS hyperpolarizability converges much more quickly since it is an incoherent process which is mostly given by a sum of squares of tensor components that do not cancel.

Throughout this paper, unless otherwise specified, we use Hartree atomic units \( \hbar = m = 1 \) with distances in Bohr (\( a_0 \approx 0.529 \text{ Å} \)) and energies in Hartrees \( \approx 27.2 \text{ eV} \).

4.3 Methods

4.3.1 Localized Orbitals

The GTO polarization properties were calculated using finite-field perturbation theory (FFPT). The electric-field strengths \( \mathcal{E} \) used were 0.00, \( \pm 0.01 \) and \( \pm 0.02 \) au. The different components of the induced dipole moment were fit to a 4th-order polynomial to obtain the polarizability and hyperpolarizability tensors. The GTO calculations were performed by Fernando Vila [387]. The best basis set used was the aug-cc-pV5Z quintuple-\( \zeta \) correlation-consistent set [97, 398] where the \( g \) and \( h \) functions were removed from the C and Cl atoms and the \( f \) and \( g \) were removed from the H atom, and the basis set was augmented with \( (0.014184, 0.009792, 0.025236) \) and \( (0.017244, 0.012528, 0.036108) \) \( (s,p,d) \) exponents on the C and Cl atoms, respectively and \( (0.004968, 0.026784) \) \( (s,p) \) exponents on the H atom. This set is abbreviated “d-a5VZs”. Sadlej’s HyPol basis set [279] was also used, which is specifically designed for the calculation of nonlinear response properties, which will be abbreviated “HP”, and gives results comparable to d-a5VZs despite being significantly smaller.

The numerical basis set (NBS) calculations were performed by Yoshi Takimoto [387] with the \textsc{Siesta} [98, 328] code and used Troullier-Martins norm-conserving pseudopotentials [373]. The best basis set used was five-\( \zeta \), quadruple-polarization, with confinement energy
shift $10^{-8}$, abbreviated 5Z4Pe8. The NBS calculations used a common $(20.0 \, \text{Å})^3$ cell and real-space grid with a plane-wave-equivalent cutoff of 250 Ry for the calculation of the Hartree and exchange-correlation potentials. This corresponds to a real-space mesh spacing of about 0.1 Å.

### 4.3.2 Real-Space Grids

For the real-space grid calculations we used *ab initio* density-functional theory with a real-space basis, as implemented in Octopus [55, 218, 7]. The polarizability and hyperpolarizability were calculated by linear response via the Sternheimer equation and the $2n + 1$ theorem [8]. This approach, also known as density-functional perturbation theory, avoids the need for sums over unoccupied states. The PBE exchange-correlation functional was used for the ground state, and the adiabatic LDA kernel [222] was used for the linear response. All calculations used Troullier-Martins norm-conserving pseudopotentials [373].

The molecule was studied as a finite system, with zero boundary conditions for the wavefunction on a large sphere surrounding the molecule, as described below. Convergence was tested with respect to the real-space grid spacing and the radius of the spherical domain. The grid spacing required is determined largely by the pseudopotential, as it governs the fineness with which spatial variations of the wavefunctions can be described as well as the accuracy of the finite-difference evaluation of the kinetic-energy operator. The spacing $\lambda$ can be converted to an equivalent plane-wave cutoff via $E_c = \left(\frac{\hbar^2}{2m}\right)(2\pi/\lambda)^2$, where $E_c$ is the cutoff energy for both the charge density and wavefunctions. The sphere radius determines the maximum spatial extent of the wavefunctions.

With tight numerical tolerances in solving the Kohn-Sham and Sternheimer equations, we can achieve a precision of 0.01 au or better in the converged values of the tensor components of $\beta$. We also did two additional kinds of calculations. For comparison to the nonlinear experiments, which used incoming photons of wavelength 1064 nm (energy 1.165 eV), we also performed dynamical calculations at this frequency via time-dependent density-functional theory (TDDFT). To compare directly to the results from finite-field perturbation theory with the other basis sets, we also calculated the dielectric properties via finite differences using electric-field strengths of $\pm 0.01$ and $\pm 0.015$ au.

In order to achieve rapid convergence of the Sternheimer SCF iterations, we used the quasi-minimal residual (QMR) linear solver [61, 113] which was implemented in the Octopus code for use in transport calculations [253]. We found it to be significantly more robust than the biconjugate gradients stabilized (BiCGstab) solver [307] in the original Octopus implementation [8]. The solver was parallelized in real-space domains via the dot products used in the algorithm. The filter preconditioner [8] was used to accelerate convergence, and was fixed to enable parallel usage.

### 4.3.3 Linear and Nonlinear Response Densities

The origin of the slow convergence of the hyperpolarizability with respect to the quality of the basis set is difficult to understand by studying only the total quantities. A more informative analysis can be obtained from the spatial distribution of the dielectric properties. Thus, we have calculated the linear and nonlinear response densities, as well as their
associated properties. Here we will focus on the response densities induced by an electric field in the \( z \)-direction. The first-order density is defined as

\[
\rho^{(1)}_z (r) = \frac{\partial \rho}{\partial \mathcal{E}_z},
\]

(4.2)

and the linear polarizability \( \alpha_{zz} (r) \) as

\[
\alpha_{zz} (r) = \rho^{(1)}_z (r) z.
\]

(4.3)

The second-order response density and associated hyperpolarizability are defined similarly:

\[
\rho^{(2)}_{zz} (r) = \frac{\partial^2 \rho}{\partial \mathcal{E}_z^2},
\]

(4.4)

\[
\beta_{zzz} (r) = \rho^{(2)}_{zz} (r) z.
\]

(4.5)

These response densities are all calculated using finite differences. For the real-space grids, our Sternheimer approach provides only the linear response density and polarizability density, but not the nonlinear response and hyperpolarizability densities.

Unlike the total properties, the spatial distributions of polarizabilities and hyperpolarizabilities as defined above depend on the origin of coordinates. Throughout this work we chose a center-of-mass reference for the spatial distributions. To understand the role that different regions of the molecule play in the total properties, we have devised a partitioning scheme for the spatial distribution corresponding to the spaces occupied roughly by the Cl atoms and C-H bond. That is, we divide space into two regions by constructing three planes, each orthogonal to one of the C-Cl bonds, and passing through a point located 40% along the C-Cl bond from the C atom, which corresponds approximately to the density minimum along the C-Cl bond. The first region (“CH”) consists of all the space above the three planes, and contains the C-H bond, while the second (“Cl”) covers the remainder of the space, including the three Cl atoms. We have integrated the various densities in each region numerically to find its contribution to the total.

### 4.4 Results and Discussion

#### 4.4.1 Structure

Although all the results presented in later sections were obtained for the experimental geometry determined by Colmont et al. [68], here we briefly discuss the effect of the structural parameters on the dielectric properties. Experimental structures have been obtained from spectroscopy of rotational transitions in isotopically substituted chloroform molecules, in 1962 by Jen et al. [169] and in 1998 by Colmont et al. [68]. The newer experimental work used more sophisticated methods than the earlier experiment, and has more precise results, so we consider it the more reliable experimental result. (Colmont et al.’s best-fit structure was slightly modified from the purely experimental result by inclusion of data from \( ab \) \textit{initio} MP2 aug-cc-PVTZ calculations.) We compared these experimental structures with theoretical structures optimized using the PBE functional, one obtained with the
Table 4.1: Structural parameters used in the study of the variation of the dielectric properties of CHCl$_3$ with structure. PBE/aVQZ and PBE/RS refer to PBE-optimized structures using the aVQZ GTO in Gaussian and a real-space grid in Octopus, respectively. Bond lengths are in Å and angles in degrees. The experimental structure from Ref. [68] was used for all our subsequent calculations.

<table>
<thead>
<tr>
<th>Source</th>
<th>$r$(C-H)</th>
<th>$r$(C-Cl)</th>
<th>$\angle$ HCCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962 Expt. [169]</td>
<td>1.100 ± 0.004</td>
<td>1.758 ± 0.001</td>
<td>107.6 ± 0.2</td>
</tr>
<tr>
<td>1998 Expt. [68]</td>
<td>1.080 ± 0.002</td>
<td>1.760 ± 0.002</td>
<td>108.23 ± 0.02</td>
</tr>
<tr>
<td>PBE/aVQZ [81]</td>
<td>1.090</td>
<td>1.779</td>
<td>107.7</td>
</tr>
<tr>
<td>PBE/RS</td>
<td>1.084</td>
<td>1.769</td>
<td>107.6</td>
</tr>
<tr>
<td>1962 Expt. (solid) [112]</td>
<td>1.060 ± 0.002</td>
<td>1.750 ± 0.015</td>
<td>109 ± 2</td>
</tr>
</tbody>
</table>

Table 4.2: Dielectric properties of various structures for CHCl$_3$ described in Table 4.1, as calculated by DFT on a real-space grid with radius 17 $a_0$ and spacing 0.25 $a_0$, compared with the experimental values of the dipole moment and the electronic contribution to the polarizability. PBE/aVQZ and PBE/RS refer to the structures described in Table 4.1. All values are in atomic units (au).

<table>
<thead>
<tr>
<th>Structure</th>
<th>$\mu_z$</th>
<th>$\alpha_{yy}$</th>
<th>$\alpha_{zz}$</th>
<th>$\beta_{yyy}$</th>
<th>$\beta_{yzz}$</th>
<th>$\beta_{zzz}$</th>
<th>$\bar{\alpha}$</th>
<th>$\beta_\parallel$</th>
<th>$\beta_{VV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962 Expt. [169]</td>
<td>0.395</td>
<td>66.14</td>
<td>47.22</td>
<td>27.09</td>
<td>-14.41</td>
<td>28.47</td>
<td>59.83</td>
<td>-0.21</td>
<td>16.89</td>
</tr>
<tr>
<td>1998 Expt. [68]</td>
<td>0.399</td>
<td>66.02</td>
<td>47.00</td>
<td>27.12</td>
<td>-16.36</td>
<td>26.92</td>
<td>59.68</td>
<td>-3.49</td>
<td>17.44</td>
</tr>
<tr>
<td>PBE/aVQZ [81]</td>
<td>0.401</td>
<td>67.17</td>
<td>47.35</td>
<td>27.23</td>
<td>-14.11</td>
<td>27.76</td>
<td>60.57</td>
<td>-0.27</td>
<td>16.79</td>
</tr>
<tr>
<td>PBE/RS</td>
<td>0.397</td>
<td>66.66</td>
<td>47.12</td>
<td>27.29</td>
<td>-14.26</td>
<td>28.92</td>
<td>60.15</td>
<td>0.24</td>
<td>16.96</td>
</tr>
<tr>
<td>Expt. (solid) [112]</td>
<td>0.399</td>
<td>65.27</td>
<td>46.62</td>
<td>27.06</td>
<td>-19.18</td>
<td>24.95</td>
<td>59.06</td>
<td>-8.03</td>
<td>18.56</td>
</tr>
</tbody>
</table>

The polarizabilities agree well with the experimentally measured values at 546 nm (2.27 nm$^{-1}$) using an aVQZ basis in Gaussian03 [81], and the other with a real-space grid in Octopus. We also compare to the crystal structure of solid chloroform, obtained by X-ray diffraction [112]. The parameters for each structure are listed in Table 4.1. The linear and nonlinear properties for each structure were calculated with the Sternheimer method in Octopus, using a radius of 17 $a_0$ and a spacing of 0.25 $a_0$ and the results are summarized in Table 4.2. Our calculations show that the dipole moment and polarizability are not affected much, but the hyperpolarizability varies significantly with structure. Individual tensor components of the hyperpolarizability do not change by more than $\sim$30%, but since $\beta_\parallel$ is a sum of large positive and negative components, it can change sign, and change by orders of magnitude depending on the structure. Therefore use of an accurate experimental structure is crucial for reliable results, as the effect of structure can be as large as those from basis set or theoretical method. Unfortunately previous work has often not mentioned the structure used, complicating comparison of results from different methods [136]. The experimental error bar on the gas-phase experimental value of $\beta_\parallel = 1 \pm 4$ au. [176] is so large relative to the value that the results for all the structures are consistent with the experiment, despite their significant variation. We used the experimental structure from Ref. [68] for all subsequent calculations in this work.

The polarizabilities agree well with the experimentally measured values at 546 nm (2.27 nm$^{-1}$) using an aVQZ basis in Gaussian03 [81], and the other with a real-space grid in Octopus. We also compare to the crystal structure of solid chloroform, obtained by X-ray diffraction [112]. The parameters for each structure are listed in Table 4.1. The linear and nonlinear properties for each structure were calculated with the Sternheimer method in Octopus, using a radius of 17 $a_0$ and a spacing of 0.25 $a_0$ and the results are summarized in Table 4.2. Our calculations show that the dipole moment and polarizability are not affected much, but the hyperpolarizability varies significantly with structure. Individual tensor components of the hyperpolarizability do not change by more than $\sim$30%, but since $\beta_\parallel$ is a sum of large positive and negative components, it can change sign, and change by orders of magnitude depending on the structure. Therefore use of an accurate experimental structure is crucial for reliable results, as the effect of structure can be as large as those from basis set or theoretical method. Unfortunately previous work has often not mentioned the structure used, complicating comparison of results from different methods [136]. The experimental error bar on the gas-phase experimental value of $\beta_\parallel = 1 \pm 4$ au. [176] is so large relative to the value that the results for all the structures are consistent with the experiment, despite their significant variation. We used the experimental structure from Ref. [68] for all subsequent calculations in this work.
eV) [344]; these quantities are optical polarizabilities which contain only an electronic contribution and have minimal dispersion. Indeed, our real-space TDDFT calculations (Section 4.4.2) at 532 nm (2.24 eV) give $\alpha_{yy} = 68.827$ au and $\alpha_{zz} = 48.405$ au, a small increase from the static and 1064 nm results, and basically consistent with the experimental values. We can also compare to a measurement of the static isotropically averaged polarizability of 64 $\pm$ 3 au. [233].

To compare with our result for the average electronic polarizability, we subtract the estimated vibrational contribution of 4.5 au calculated from experimental data (no error bar provided) [33], yielding 60 au, which agrees with the predicted values within 0.4%.

### 4.4.2 Real-Space Grids

Convergence of the dipole moment, polarizability, and hyperpolarizability is illustrated in Table 4.3. The total energy was well converged for a spacing of 0.35 $a_0$ (equivalent plane-wave cutoff = 20 Ry) and a sphere radius of 12 $a_0$. The dipole moment was also well converged with that basis. However, to converge $\beta_\parallel$ to better than 0.01 au, a spacing of 0.25 $a_0$ (equivalent plane-wave cutoff = 40 Ry) and a sphere radius of 22 $a_0$ was required. The convergence of the tensor components of $\beta$ is similar to that of $\beta_\parallel$ in absolute terms, i.e. they are also converged to 0.01 au or better with these parameters. Generally, the magnitude of $\beta_\parallel$ declines with smaller spacing and larger radius, as the cancellation between the tensor components becomes closer. (By contrast, for the localized orbitals, there are changes in sign and variations of two orders of magnitude in $\beta_\parallel$ as the basis set is improved, for GTO and NBS calculations [387].)

Finite-difference calculations were done with the converged grid spacing of 0.25 $a_0$, and a sphere radius of 22 $a_0$, for comparison to the Sternheimer calculation with the same grid parameters (Table 4.5). The differences between the linear-response and finite-difference calculations are small, allowing a direct comparison between the results with different basis sets. The use of the LDA kernel in the linear-response results gives a small discrepancy compared to the purely PBE finite-difference results. Fields of 0.015 au rather than 0.02 au as for the other basis sets were used because 0.02 au was out of the linear regime in the real-space calculation. The linear response density $\rho_{z}^{(1)}(r)$ and polarizability density $\alpha_{zz}(r)$ are virtually identical between the finite-difference and linear-response calculations.

Calculations at 1064 nm with the same grid parameters show increases of about 1% in the polarizability, and 10-20% in the hyperpolarizability, as compared to the static case. We find a small violation of Kleinman symmetry here, in that $\beta_{yyz} = -18.945$ au whereas $\beta_{zyy} = -19.448$ au.

### 4.4.3 Linear and Nonlinear Response Densities

The origin of the slow convergence of the response properties is difficult to determine just by analyzing their total values for different basis sets. We have found that the differences and similarities between those values can be visualized by computing the real-space distribution of the linear and nonlinear response densities, as well as the associated polarizability and hyperpolarizability densities. For example, Figure 4.1 shows the linear response density $\rho_{z}^{(1)}(r)$ calculated with the PBE functional for both the HP GTO basis set and
The spatial distributions of the nonlinear response density $\rho^{(2)}(r)$ and hyperpolarizability density $\beta^{zzz}(r)$, shown in Figure 4.2, are also very similar for both the HP GTO set and the real-space grid. The hyperpolarizability density, however, is more delocalized than the polarizability, extending up to $\sim 8$ au from the center, thus stressing the importance of the diffuse functions in calculations of nonlinear properties. The spatial distribution is also much more complex, with several regions of counteracting contributions. The decomposition shown in Table 4.4 significantly simplifies the analysis of the densities. It shows that the overall contribution from the C-H bond is negative. This contribution also varies very little with respect to the quality of the basis set. The contribution from the Cl atoms, on the other hand, is positive and varies significantly with the basis set used.

### 4.5 Conclusions

I calculated the static hyperpolarizability of the gas-phase CHCl$_3$ molecule with real-space grids and compared to Gaussian-type orbitals and numerical basis sets. We find that all of these methods can yield quantitatively similar results provided sufficiently large, diffuse basis sets are included in the calculations. The overall consistency among the results gives confidence to their reliability and overall accuracy. Based on the size of the basis sets

<table>
<thead>
<tr>
<th>$R$</th>
<th>$\lambda$</th>
<th>$\mu_z$</th>
<th>$\alpha_{yy}$</th>
<th>$\alpha_{zz}$</th>
<th>$\beta_{yy}$</th>
<th>$\beta_{yyz}$</th>
<th>$\beta_{zzz}$</th>
<th>$\bar{\alpha}$</th>
<th>$\beta_{1z}$</th>
<th>$\beta^{NNV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.25</td>
<td>0.398</td>
<td>65.921</td>
<td>46.924</td>
<td>27.975</td>
<td>$-17.232$</td>
<td>22.975</td>
<td>59.589</td>
<td>$-6.921$</td>
<td>17.106</td>
</tr>
<tr>
<td>15</td>
<td>0.25</td>
<td>0.399</td>
<td>66.019</td>
<td>46.993</td>
<td>27.159</td>
<td>$-16.401$</td>
<td>26.758</td>
<td>59.677</td>
<td>$-3.629$</td>
<td>17.461</td>
</tr>
<tr>
<td>17</td>
<td>0.25</td>
<td>0.399</td>
<td>66.022</td>
<td>46.995</td>
<td>27.123</td>
<td>$-16.363$</td>
<td>26.921</td>
<td>59.680</td>
<td>$-3.485$</td>
<td>17.443</td>
</tr>
<tr>
<td>20</td>
<td>0.25</td>
<td>0.399</td>
<td>66.023</td>
<td>46.995</td>
<td>27.119</td>
<td>$-16.358$</td>
<td>26.940</td>
<td>59.680</td>
<td>$-3.469$</td>
<td>17.441</td>
</tr>
<tr>
<td>22</td>
<td>0.25</td>
<td>0.399</td>
<td>66.023</td>
<td>46.995</td>
<td>27.119</td>
<td>$-16.358$</td>
<td>26.940</td>
<td>59.680</td>
<td>$-3.468$</td>
<td>17.441</td>
</tr>
</tbody>
</table>

Table 4.3: Effect of the real-space-grid quality (radius $R$ and spacing $\lambda$) on the components of the dielectric properties of CHCl$_3$ calculated with the PBE functional and LDA kernel. All values are in atomic units. 

| Expt. | 0.409(8) | 61±5 | 45±3 | 56±4 | 1±4 |
| Ref.  | [297]    | [344] | [344] | [344] | [176] |
Figure 4.1: Linear response density $\rho_z^{(1)}(r)$ (a-b) and polarizability density $\alpha_{zz}(r)$ (c-d) on one of the HCl planes of the molecule calculated with a GTO basis set (HP) and a real-space grid (RS) using the PBE functional. The positions of the nuclei are indicated with black dots, and the black lines are isolines. All quantities are in atomic units. Note that the linear response density is quite similar for both methods. GTO data courtesy F. Vila [387].

and degree of convergence, the LR real-space values in Table 4.5 are likely the most reliable PBE results obtained in this study. However, the variation among our results also provides a gauge of their overall theoretical accuracy.

Although the treatment here has been restricted to chloroform, many of the results are more generally applicable. For example, the spatial distributions provided by the linear and nonlinear response densities provides a good visualization tool to understand the basis set requirements for the simulation of linear and nonlinear response. A key finding for chloroform is that the local contributions near the Cl atoms and the CH bond are of opposite sign and tend to cancel, thus explaining the overall weakness of the hyperpolarizability. The molecule’s response is quite extended in space and so real-space grids on a large domain, as well as very extended local orbitals, are required to describe it properly. The frequency-dependence of the polarizability and hyperpolarizability is small, as verified by our time-dependent calculations, and so dispersion is not very important in comparing static theoretical results to experimental measurements.
Table 4.4: Partitioning of the linear and nonlinear response properties calculated numerically with GTOs and real-space grids (RS) using the PBE functional together with the numerical sum over the CH and Cl\(_3\) partitions. GTO data courtesy F. Vila [387].

<table>
<thead>
<tr>
<th>Property</th>
<th>Basis Set</th>
<th>CH</th>
<th>Cl(_3)</th>
<th>CH + Cl(_3)</th>
<th>Analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_{zz})</td>
<td>d-aV5zs</td>
<td>8.96</td>
<td>37.79</td>
<td>46.75</td>
<td>46.79</td>
</tr>
<tr>
<td></td>
<td>HP</td>
<td>8.92</td>
<td>37.67</td>
<td>46.59</td>
<td>46.60</td>
</tr>
<tr>
<td></td>
<td>RS</td>
<td>8.85</td>
<td>38.02</td>
<td>46.87</td>
<td>46.87</td>
</tr>
<tr>
<td>(\beta_{zzz})</td>
<td>d-aV5zs</td>
<td>-44.31</td>
<td>66.54</td>
<td>22.24</td>
<td>22.27</td>
</tr>
<tr>
<td></td>
<td>HP</td>
<td>-44.55</td>
<td>67.49</td>
<td>22.94</td>
<td>22.93</td>
</tr>
<tr>
<td></td>
<td>RS</td>
<td>-43.41</td>
<td>67.12</td>
<td>23.71</td>
<td>23.89</td>
</tr>
</tbody>
</table>

Table 4.5: Summary of the best PBE results obtained with the GTOs, numerical basis sets and real-space grids. Real-space grids (lr denotes linear response, and fd finite difference) have radius 22 \(a_0\), spacing 0.25 \(a_0\). All values are in atomic units. GTO data courtesy F. Vila, NBS data courtesy Y. Takimoto [387].

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>(\mu)</th>
<th>(\alpha_{yy})</th>
<th>(\alpha_{zz})</th>
<th>(\beta_{yy})</th>
<th>(\beta_{yz})</th>
<th>(\beta_{zzz})</th>
<th>(\bar{\alpha})</th>
<th>(\bar{\beta})</th>
<th>(\bar{\beta}^|)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTO d-aV5zs</td>
<td>0.404</td>
<td>65.70</td>
<td>46.79</td>
<td>27.35</td>
<td>-15.31</td>
<td>22.27</td>
<td>59.40</td>
<td>-5.01</td>
<td>16.90</td>
</tr>
<tr>
<td>NBS 5Z4Pe8</td>
<td>0.398</td>
<td>65.45</td>
<td>46.28</td>
<td>24.54</td>
<td>-14.90</td>
<td>21.37</td>
<td>59.06</td>
<td>-5.07</td>
<td>15.68</td>
</tr>
<tr>
<td>RS lr</td>
<td>0.399</td>
<td>66.02</td>
<td>47.00</td>
<td>27.12</td>
<td>-16.36</td>
<td>26.94</td>
<td>59.68</td>
<td>-3.47</td>
<td>17.44</td>
</tr>
<tr>
<td>RS fd</td>
<td>&quot;</td>
<td>66.05</td>
<td>46.87</td>
<td>24.74</td>
<td>-15.17</td>
<td>23.89</td>
<td>59.66</td>
<td>-3.87</td>
<td>15.97</td>
</tr>
<tr>
<td>RS 1064 nm</td>
<td>&quot;</td>
<td>66.69</td>
<td>47.34</td>
<td>30.35</td>
<td>-18.95</td>
<td>31.56</td>
<td>60.24</td>
<td>-4.01</td>
<td>19.91</td>
</tr>
</tbody>
</table>

| Expt. | 0.409\(\pm\)0.008 | 61\(\pm\)5 | 45\(\pm\)3 | 56\(\pm\)4 | 1\(\pm\)4 |
| Expt. Ref. | [297] | [344] | [344] | [344] | [176] |

The discrepancy between the experimentally determined linear polarizability and our theoretical results is essentially eliminated when the vibrational component is taken into account. Our results for the hyperpolarizability for all three basis sets are all consistent with each other. Given the error bars in the experimental result our PBE hyperpolarizability results are smaller, though essentially consistent with the experimental measurements for \(\beta^\|\), even without taking into account vibrational contributions. Experimental results indicate that the vibrational contribution is small for the hyperpolarizability: differences in the hyperpolarizability of isotopically substituted molecules show the vibrational contributions. Although measurements at the same frequency are not available for CHCl\(_3\) and CDCl\(_3\), Kaatz et al. [176] found that at 694.3 nm, CHCl\(_3\) has \(\beta^\| = 1.2 \pm 2.6\) au; at 1064 nm CDCl\(_3\) has \(\beta^\| = 1.0 \pm 4.2\) au. Given that the frequency-dispersion of \(\beta^\|\) between zero frequency and 1064 nm is only about +15% in our calculations (much smaller than the error bars), this isotopic comparison shows that the vibrational contributions are less than the error bars. Therefore vibrational contributions are not significant in comparing the \(ab\ ini\) results to the available experimental measurements. We find additionally that the molecular structure has a significant influence on the calculated value of \(\beta^\|\), and so it is crucial to use an accurate structure for theoretical calculations.
Figure 4.2: Nonlinear response density $\rho_{zz}^{(2)}(r)$ (a-b) and hyperpolarizability density $\beta_{zzz}(r)$ (c-d) on one of the HCCl planes of the molecule calculated with a GTO basis set (HP) and a real-space grid (RS) using the PBE functional. The positions of the nuclei are indicated with black dots, and the black lines are isolines. All quantities are in atomic units. The nonlinear densities extend much further into space than the linear densities. The agreement between the real-space and GTO methods is nevertheless quite good. The contributions to the hyperpolarizability from the Cl atoms and the CH bond are of opposite sign and, as indicated by the nonlinear response density, have contributions that extend even further into space. GTO data courtesy F. Vila [387].
Chapter 5

First-principles prediction of optical second-harmonic generation in the endohedral N@C$_{60}$ compound

This work is based on Ref. [413].

5.1 Abstract

Nonlinear optical properties in C$_{60}$ have attracted enormous attention for over two decades. The endohedral complex N@C$_{60}$, with its remarkable thermal stability and spin-quartet ground state, is a candidate for future room-temperature quantum computing, but there has been no investigation of its nonlinear optical properties. Here, a first-principles calculation shows that N@C$_{60}$ is a promising material for nanoscale and ultrafast modulation. Excitation by a pump laser pulse of the nitrogen-atom vibration inside the C$_{60}$ cage transiently breaks inversion symmetry and can enable second-harmonic generation from a probe pulse. Unlike the second-harmonic generation observed in C$_{60}$ thin films, this harmonic signal is switched on/off periodically every 345 fs. For an fcc crystal of N@C$_{60}$, the second-order susceptibility $\chi^{(2)}$ is on the order of $10^{-8}$ esu, similar to commercially used nonlinear materials.

5.2 Introduction

Endohedral fullerenes are very attractive for organic photovoltaic applications [30, 305, 379]. Nonlinear optical response in pure C$_{60}$ and other endohedral fullerenes has already attracted huge interest for over two decades [415, 131, 49]. Successful implantation of a nitrogen atom into C$_{60}$ [5] has opened a new frontier for possible molecule-based quantum computing [138, 28]. Electron paramagnetic resonance (EPR) experiments have established that the ground state of N@C$_{60}$ is a spin quartet with all three nitrogen 2p levels half-filled,
similar to atomic nitrogen. This is significant as it allows the spin degree of freedom to be exploited [129]. Since N@C$_{60}$ is extremely stable even at room temperature, it is likely to overcome one of the challenges in quantum computing: the devices must operate in the ambient environment [239].

Theoretically, there is general agreement on the electronic and spin structure and the geometry of N@C$_{60}$ [229, 276, 336, 41, 410]. BelBruno [25] studied the potential-energy surface of N@C$_{60}$ and other group-V endohedral complexes, where they found that the centered position has the lowest energy in calculations with both 6-31G* and DZ bases. The energetic position of the N(2p) level has been controversial, though there now seems to be a consensus. Greer [128] as well as Plakhutin and coworkers [278, 277] have found that the open-shell state is the N(2p) state, near resonant with C$_{60}$'s $t_{2g}$ orbital, and the C$_{60}$ HOMO and LUMO wavefunctions and energies are very close to those of pristine C$_{60}$, showing that the interaction between the N atom and C$_{60}$ cage is small. Up to now, major research has been on the electronic and spin properties of N@C$_{60}$, and to our knowledge, there has been little optical investigation of N@C$_{60}$ either experimentally or theoretically [166]. Our study represents an effort in this direction.

In this paper, we show through a first-principles calculation that the excitation of the vibration of an N atom inside C$_{60}$ allows second-harmonic generation (SHG) by temporarily breaking the inversion symmetry. Unlike previous SHG studies where the signal was constant in time [195, 196, 194, 193], here the signal changes periodically with a period of 345 fs, corresponding to the vibrational period of the N atom. We predict a new spectral line at 96 cm$^{-1}$ in the infrared (IR) spectrum, well separated from the lowest line at 245 cm$^{-1}$ in pristine C$_{60}$ [414]. Depending on the position of the N atom, $\chi^{(2)}$ in the ultraviolet is in the order of $10^{-8}$ esu, comparable to the value (in the IR) for commercial nonlinear materials such as KTiOPO$_4$ [346]. Molecules with large hyperpolarizabilities typically crystallize in centrosymmetric structures, canceling their nonlinear susceptibilities [46], but for N@C$_{60}$, dynamic generation of molecular nonlinearity allows a large transient susceptibility of the crystal. While the energy of this mode is comparable to room temperature, incoherent thermal vibrations will give no signal. Excitation of the N vibration by femtosecond laser pulses will produce an SHG signal turning on/off twice for every cycle of the N vibration, which can be detected by pulses which are short compared to the vibrational period, probing an effectively static ionic configuration. This process of modulation of SHG by phonons has been shown for surfaces of GaAs [59], which has bulk SHG even in the ground state; for N@C$_{60}$ this process should be even easier to observe since there will not be background from the dipole-forbidden bulk SHG. The coherent excitation of IR-active phonon modes by ultrafast pulses is well established, having been demonstrated in solid C$_{60}$ [91, 104] and other materials [56].

The rest of the paper is arranged as follows. In Sec. 5.3, we present our theoretical methods, while the results and discussions are presented in Sec. 5.4. In Sec. 5.5, we conclude this paper. In the appendix, we show the details of the unit conversions.
5.3 Theoretical methods

We compare two complementary and state-of-the-art first-principles methods: configuration interaction (CI), as implemented in the Gaussian03 package [114], and time-dependent density-functional theory (TDDFT), as implemented in the Octopus code [55, 218, 7]. CI calculations were done by G. P. Zhang [413]. The CI calculation uses the exact Hamiltonian in a restricted basis set, whereas the TDDFT calculation uses approximate functionals but a much larger real-space basis set, and in the Sternheimer approach avoids the need to converge a sum over excited states [8]. Recent studies suggest better basis sets may be more important than a better Hamiltonian for accuracy in molecular nonlinear-optical properties [136, 387]. One prominent challenge for a CI calculation is the huge configuration space, particularly for a large system such as N@C_{60}. We limit ourselves to unrestricted configuration interaction with single excitations (UCIS). A recent CI study of C_{60} found that the cc-pVDZ basis and single excitations, as used here, give good results for the optical spectrum, by comparison to higher-level CI calculations [252]. TDDFT studies of C_{60} with the Sternheimer equation and adiabatic LDA kernel have also shown good results for linear and nonlinear optical properties [161].

We treat the molecules as isolated since their interaction in the solid is weak, and compute susceptibilities as molecular polarizability and polarizability divided by the unit-cell volume of bulk C_{60} [111]. For the DFT, HF and CI calculations with Gaussian basis sets, we use five different bases: 3-21G, 6-31G, 6-311G, 6-31G (d) and cc-pVDZ. For the TDDFT calculations, we use a real-space grid forming a sphere of radius 8 Å with grid spacing of 0.2 Å. The molecular structure is optimized with each Gaussian basis using DFT with the B3LYP hybrid functional. In agreement with previous calculations [128, 278, 277], the total-energy minimum is obtained with the nitrogen atom at the center of the C_{60} cage, and the ground state has a total spin $m = 3/2$, corresponding to a spin quartet. All the subsequent calculations fix the C_{60} cage geometry and only move the N atom. The inset of Fig. 5.1(a) shows the total-energy change as a function of the displacement $Z$ of the N atom with respect to the center of the cage. The potential-energy surface is from an MP2 calculation starting from unrestricted Hartree-Fock (UHF) with the 6-31G basis [413]; counterpoise corrections are found to have little effect on the energy. Fitting the potential energy to a harmonic potential $U = kZ^2/2$, we find a spring constant of $k = 0.48$ eV/Å$^2$, corresponding to $T = 345$ fs, or 96 cm$^{-1}$. As noted previously, the potential-energy surface is almost spherically symmetric [229].

This N vibration has $T_{1u}$ symmetry, so it is IR-active but not Raman-active, as shown by its nonzero Born effective charge of +0.009 (from DFT with the B3LYP functional) and negligible effect on $\chi^{(1)}$ (Fig. 5.1(a)), unlike the Raman activity of off-center Li$^+@C_{60}$ [172]. Since the smallest wavenumber of a C_{60} vibration mode is 245 cm$^{-1}$, a new peak will show up in the IR spectrum as seen for other endohedral fullerenes [173, 313, 43]. This provides experimentalists with a first test of our calculations and enables excitation of this vibrational mode by a pump pulse.

Figure 5.1(a) shows the linear absorption obtained using the 6-31G and cc-pVDZ bases in CI, and the TDDFT result from the linear-response Sternheimer approach [8] with the PBE functional and adiabatic LDA kernel. TDDFT has a more complicated set of peaks.
Figure 5.1: (a) Linear susceptibility as a function of the incident photon energy using UCIS/cc-pVDZ method (solid line), UCIS/6-31G method (dotted line – the curves for the four N atom positions lie on top of each other) and TDDFT method (thin line, for $Z = 0.4$ Å). Inset: Ground-state MP2 energy changes (6-31G basis) as a function of the position $Z$ of the nitrogen atom inside the $C_{60}$ cage. (b) Second-order susceptibility $\chi_{zzz}^{(2)}(2\omega;\omega,\omega)$ versus the incident photon energy $\hbar\omega$ for different nitrogen positions inside the $C_{60}$ cage at UCIS/6-31G level. The $Z$-values are shown above each line. The curves are vertically shifted by 0.5 for clarity. UCIS and MP2 data courtesy G. P. Zhang [413].
than CI, though the peak at about 5.5 eV agrees between both methods. The first peak around 3.5 eV agrees well with those observed for N@C_{60} (and C_{60}) [166]. Predictions for the relative intensities are quite different: the absolute intensities are about an order of magnitude less in TDDFT, although the total spectral weight in the range considered here is similar. Despite the reduction in symmetry with displacement, there is no change in the peak position and height even as the nitrogen atom changes its position (the four dashed lines all collapse into a single line in Fig. 5.1(a) for the 6-31G basis). This is not surprising since the linear response is much less sensitive to the molecular symmetry change, and \( \chi^{(1)} \) remains very close to isotropic. In sharp contrast to the linear response, the second-order susceptibility shows a strong dependence on the N position (see Fig. 5.1(b)).

5.4 Results and Discussions

In the following, we show that the vibration of the nitrogen atom [293] will lead to second-harmonic generation, a process that can be probed by an ultrafast laser pulse. The amplitude of the signal is given by the second-order susceptibility \( \chi^{(2)}(2\omega; \omega, \omega) \), which we evaluate using a sum-over-states expression [322] in the convention of Butcher and Cotter [47] (in SI units – see page 69 of Ref. [47]) as follows:

\[
\chi_{ijk}^{(2)}(\omega = \omega_1 + \omega_2; \omega_1, \omega_2) = \frac{N}{\epsilon_0} \frac{e^3}{2\hbar^2} \sum_{g,n,m} \left[ \frac{r_i^g r_j^{nm} r_k^m}{(\omega - \omega_{ng} + i\Gamma_{ng})} \left( \omega_2 - \omega_{mg} + i\Gamma_{mg} \right) \right. \\
+ \frac{r_i^g r_j^{nm} r_k^m}{(\omega - \omega_{ng} + i\Gamma_{ng})} \left( \omega_1 - \omega_{mg} + i\Gamma_{mg} \right) \\
+ \frac{r_j^g r_i^{nm} r_k^m}{(\omega + \omega_{ng} + i\Gamma_{ng})} \left( \omega_2 + \omega_{mg} + i\Gamma_{mg} \right) \\
+ \frac{r_j^g r_i^{nm} r_k^m}{(\omega + \omega_{ng} + i\Gamma_{ng})} \left( \omega_1 + \omega_{mg} + i\Gamma_{mg} \right) \\
- \frac{r_j^g r_i^{nm} r_k^m}{(\omega - \omega_{nm} + i\Gamma_{nm})} \left( \frac{1}{\omega_2 + \omega_{mg} + i\Gamma_{mg}} + \frac{1}{\omega_1 - \omega_{mg} + i\Gamma_{mg}} \right) \\
- \frac{r_j^g r_i^{nm} r_k^m}{(\omega - \omega_{nm} + i\Gamma_{nm})} \left( \frac{1}{\omega_2 - \omega_{ng} + i\Gamma_{ng}} + \frac{1}{\omega_1 + \omega_{mg} + i\Gamma_{mg}} \right) \left. \right] \rho_g. \\
\tag{5.1}
\]

The damping is \( \Gamma = 0.2 \) eV, comparable to experimental broadening [196]. Here, \( N \) is the number density of molecules, \( e \) is the electron charge, \( \hbar \) is Planck’s constant over \( 2\pi \), \( r_{nm}^i \) is the dipole transition matrix element between states \( n \) and \( m \) along the \( i \) direction, \( i \) denotes the imaginary unit (avoiding the confusion with the direction index \( i \)), \( \hbar \omega \) and \( \hbar \omega_{(2)} \) are the emitted and incident photon energies, respectively, \( \hbar \omega_{nm} \) is the energy difference between states \( n \) and \( m \), and \( \rho_g \) is the ground-state density. We use the number density \( N \) of C_{60} molecules arranged in the fcc structure with lattice constant \( a = 14.161 \) Å [111].

The sum over states uses the excitation energies and dipole matrix elements from CI, thus including many-body effects. 260 states were included, which provides reasonable convergence in both the CI diagonalization and the sum, and respects the degenerate sub-spaces for the symmetry of the system [413]. Figure 5.1(b) shows that as the N atom moves away from the center of the cage, the SHG signal increases sharply from 0 at \( Z = 0 \) Å to
0.79 \times 10^{-8} \text{ esu} \text{ at } Z = 0.4 \, \text{Å}. \text{ To be more convincing, in Fig. 5.2 we employ five different bases from 3-21G to cc-pVDZ and find that the SHG signal survives in all the cases, with the main features converging nicely. The main structures are due to one- and two-photon resonances of the same excitation seen in the linear spectrum. These UCIS results will be compared with the TDDFT ones.}

The TDDFT $\chi^{(2)}$ spectrum shows structure in the 2 to 5 eV range, with a pair of strong peaks at 5.5 eV and 6.1 eV, and a weak peak around 7 eV. The energies are similar to the CI/cc-pVDZ peaks at 2.8, 3.7, 5.5 and 7.0 eV, but have different relative amplitudes. The onset of about 2 eV agrees between the calculations, and the absolute intensities are of the same order of magnitude as in CI. The experimental SHG intensity is proportional to $|\chi^{(2)}|^2$, i.e., including both real and imaginary parts, and involves different tensor components depending on the relative polarization of the pump and probe beams and the detected second harmonics. We find that the real and imaginary parts of $\chi^{(2)}_{zzz}$ as well as of $\chi^{(2)}_{xxx}$ (the other symmetry-unique component) all have similar frequency dependence. The TDDFT calculation confirms that this system will show $\chi^{(2)} \sim 10^{-8}$ esu.

5.4.1 Convergence of TDDFT calculation

All TDDFT calculations employing the Octopus code were done on the displaced $Z = 0.4 \, \text{Å}$ geometry. In the real-space calculation, there are two parameters defining the basis for the wavefunctions and density: the spacing and the domain. Both can be converged systematically. The required spacing of the real-space grid depends on the hardness of the pseudopotentials and corresponds closely to the cutoff in a plane-wave calculation. The domain is the volume which is filled with the grid points and in which the wavefunctions may have non-zero values. A molecule such as N@C$_{60}$ is naturally treated as a finite system with zero boundary conditions for the wavefunctions, in a spherical domain whose radius must be converged. We found that a spacing of 0.21 Å was sufficient for our norm-conserving Troullier-Martins pseudopotentials, corresponding to a plane-wave cutoff of 62 Ry. A radius of at least 8 Å was required to converge the ground-state, linear and nonlinear susceptibilities, as shown in Table 5.1. This value is considerably larger than the structural radius of C$_{60}$ of 3.6 Å. Note that differences in the dynamic quantities at a fixed frequency are largely due to changes in peak position as well as peak amplitude; we cannot disentangle the two contributions without the time-consuming recalculation of the entire spectrum for the other radii and spacings.

The convergence parameters here are in strong contrast with our previous study of the CHCl$_3$ molecule in Chapter 4, in which convergence was particularly difficult due to the large polarizability of Cl atoms and the small hyperpolarizability of the molecule. A radius of 11.6 Å was required to converge the hyperpolarizability for a molecule whose atoms fit within a sphere of 1.6 Å. Indeed, the domain needed for CHCl$_3$ was actually larger than that for N@C$_{60}$. Additionally, for N@C$_{60}$, the ground-state, linear and nonlinear properties converge at a similar rate with radius, whereas for CHCl$_3$ the linear properties converged more slowly than the ground-state properties, and the nonlinear properties were even slower to converge with radius.

The sphere of radius 8 Å and spacing 0.21 Å has 226,375 points, much larger than the cc-pVDZ Gaussian basis set of rank 854. The Sternheimer linear-response approach
Figure 5.2: Second-order susceptibility $\chi_{zzz}^{(2)}$ for five bases from 3-21G to cc-pVDZ. All the structures are optimized at the B3LYP level, and the excited-state properties are computed with UCIS. The TDDFT result is shown at the bottom. The N atom is 0.4 Å away from the center of the C$_{60}$ cage. The horizontal dashed lines denote the zero line. UCIS data courtesy G. P. Zhang [413].
Table 5.1: Convergence of ground-state, linear and nonlinear properties of N@C60 as a function of real-space grid parameters in the Octopus calculation (radius $R$ and spacing $\lambda$). The quoted gap is the Kohn-Sham gap between the highest majority-spin state and lowest minority-spin state. Susceptibilities are shown for the static case and the dynamic case of $\hbar \omega = 3$ eV, with a broadening of $\Gamma = 0.2$ eV. Units are given in the second row for each quantity.

<table>
<thead>
<tr>
<th>$R$ (Å)</th>
<th>$\lambda$ (Å)</th>
<th>Energy (Ha)</th>
<th>$\mu_z$ (esu)</th>
<th>$E_g$ (Ha)</th>
<th>Re $\chi^{(1)}_{zz} (0)$</th>
<th>Im $\chi^{(1)}_{zz} (\omega)$</th>
<th>Re $\chi^{(2)}_{zzz} (0)$</th>
<th>Im $\chi^{(2)}_{zzz} (\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.21</td>
<td>-353.35756</td>
<td>0.0135898</td>
<td>0.048037</td>
<td>0.11851</td>
<td>0.00992</td>
<td>-3.91</td>
<td>-229.43</td>
</tr>
<tr>
<td>8</td>
<td>0.21</td>
<td>-351.42563</td>
<td>0.0133833</td>
<td>0.051014</td>
<td>0.11460</td>
<td>0.00985</td>
<td>-13.51</td>
<td>-439.94</td>
</tr>
<tr>
<td>9</td>
<td>0.21</td>
<td>-351.42592</td>
<td>0.0133772</td>
<td>0.050997</td>
<td>0.11491</td>
<td>0.00993</td>
<td>-14.79</td>
<td>-430.16</td>
</tr>
<tr>
<td>10</td>
<td>0.21</td>
<td>-351.42594</td>
<td>0.0133772</td>
<td>0.050996</td>
<td>0.11495</td>
<td>0.00994</td>
<td>-14.865</td>
<td>-425.88</td>
</tr>
<tr>
<td>8</td>
<td>0.19</td>
<td>-351.21390</td>
<td>0.0137604</td>
<td>0.048059</td>
<td>0.11475</td>
<td>0.00981</td>
<td>-12.085</td>
<td>-341.41</td>
</tr>
<tr>
<td>8</td>
<td>0.16</td>
<td>-351.12981</td>
<td>0.0137294</td>
<td>0.047225</td>
<td>0.11465</td>
<td>0.00977</td>
<td>-11.855</td>
<td>-297.37</td>
</tr>
</tbody>
</table>

does not involve a sum over states, since it calculates the contributions from all states at a given frequency, and therefore it effectively includes a sum over all 226,375 states of each spin in the Hilbert space, in contrast to the sum on only 260 unoccupied states from the CI calculation. On the other hand, while TDDFT is in principle exact, the use of an approximate kernel (adiabatic LDA) [222] makes the Hamiltonian less accurate (and computationally intensive) than that of CI. The comparison between our CI and TDDFT calculations is a tradeoff between the quality of the basis set and the Hamiltonian.

For $C_{5v}$ symmetry, the $\chi^{(1)}$ tensor is diagonal in our coordinate system and may be represented fully by $\chi^{(1)}_{xx}$ and $\chi^{(1)}_{zz}$. In Fig. 5.3, we show the real and imaginary parts of each. In the centered geometry with $I_h$ symmetry, it is required that $\chi^{(1)}_{xx} = \chi^{(1)}_{zz}$, but even though this constraint no longer holds exactly, we find it to be approximately fulfilled here anyway. The imaginary part relates to optical absorption and the real part to local-field factors. Note that the real part is expected to be much more accurate with the Sternheimer TDDFT approach than in the CI calculation, since converging the real part of a response function with a sum over states (effectively, a Cauchy expansion) presents fundamental additional difficulties compared to the imaginary part [167]. This consideration applies to $\chi^{(2)}$ as well.

For $C_{5v}$ symmetry, the $\chi^{(2)}$ tensor may be represented fully by the two symmetry-unique elements $\chi^{(2)}_{xz}$ and $\chi^{(2)}_{zz}$, at zero frequency. When Kleinman permutation symmetry [188] is lost in the dynamic case, $\chi^{(2)}_{xx}$ is also needed. In Fig. 5.4, we plot the real and imaginary parts of these tensor components, showing that they are of similar order of magnitude and frequency dependence. The real and imaginary parts describe the phase of the second harmonics relative to the input beams. We conclude that the experimentally-measured SHG intensity, proportional to $|\chi^{(2)}|^2$, will correspond to $\chi^{(2)} \sim 10^{-8}$ esu regardless of the relative polarization of the beams.
Figure 5.3: TDDFT linear susceptibility. The empty and filled circles represent $\chi_{xx}^{(1)}$, vertically shifted for clarity. The empty and filled boxes denote $\chi_{zz}^{(1)}$. The filled circles and boxes are the real parts. The empty circles and boxes are the imaginary parts. The horizontal dotted lines denote the zero lines. The tensor is close to isotropic.
Figure 5.4: Second-harmonic spectrum obtained by the TDDFT method. The empty and filled circles represent $\chi^{(2)}_{zxx}$, vertically shifted for clarity. The empty and filled boxes denote $\chi^{(2)}_{zzz}$. The filled circles and boxes are the real parts. The empty circles and boxes are the imaginary parts. The horizontal dotted lines denote the zero line. The frequency dependence and order of magnitude are similar between the real and imaginary parts and between the two tensor components shown here.
Since the Sternheimer calculation naturally provides both the real and imaginary parts of response functions, provided an imaginary broadening $\Gamma$ is used [8], we can perform a check on the consistency of our results via the Kramers-Kronig relations. The well-known equation for the imaginary part in terms of the real part is:

$$\text{Im} \, \chi^{(1)}(\omega) = -\frac{2\omega}{\pi} \int_0^{\infty} \frac{\text{Re} \, \chi^{(1)}(\omega)}{\Omega^2 - \omega^2} d\Omega$$

Precisely the same holds for $\chi^{(2)}(2\omega;\omega,\omega)$, where the three frequencies maintain the same ratios with one another [268]. Reasonable agreement with this relation was found, with some discrepancy due to the limited frequency resolution and range of our calculated spectra, over which to integrate.

I made key developments to the Octopus code for the TDDFT calculations done here. I generalized the hyperpolarizability calculation (in Equation 2.33 [8]) to the spin-polarized case, and also to obtain the imaginary as well as the real part of the hyperpolarizability. In solving the Sternheimer equation, we proceed one frequency at a time. All the linear-response densities are stored to disk, and the closest frequency is used as a starting guess to reduce the number of SCF iterations required. Due to limited disk space, only the linear-response wavefunctions from the previous frequency are stored. I improved the code’s ability to reuse this saved data intelligently, taking into account, for example, the fact that the SHG calculation of $\chi^{(2)}(2\omega;\omega,\omega)$ requires solution of linear response at $-\omega$ and $2\omega$, and the data at $2\omega$ can be easily used transformed to the result for $-2\omega$ for the calculation of $\chi^{(2)}(4\omega;2\omega,2\omega)$. Such reuse was crucial as this was probably the largest system for which a full nonlinear spectrum has been calculated by the Sternheimer approach.

### 5.4.2 Experimental consideration and comparison of methods

Finally, we critically examine the possibility of observing the above features experimentally, and the relative quality of results from the two theoretical methods.

(1) Magnitude: Depending on the displacement of the N atom, the amplitude of $\chi^{(2)}$ is on the order of $10^{-8}$ esu, comparable to commercial frequency-doubling materials and clearly detectable experimentally [346]. Indeed, there are also enhancements by local-field factors, which describe the increase in the field felt by a molecule due to the polarization of its neighbors [322]. For cubic symmetry,

$$f_\omega = \frac{1}{1 - \frac{4\pi}{\pi} \text{Re} \, \alpha(\omega)/V},$$

where $\alpha$ is the polarizability and $\chi^{(1)} = 1 + 4\pi\alpha/V$. In the optical/UV range, the TDDFT $\chi^{(1)}$ gives $f_\omega \sim 1.7$, which is the enhancement factor for $\chi^{(1)}$, while $\chi^{(2)}$ is enhanced by the factor $f_2^2 f_\omega \sim 5$.

We expect the magnitude of $\chi^{(1)}$ and $\chi^{(2)}$ to be more accurate in TDDFT than UCIS since TDDFT obeys the Thomas-Reiche-Kuhn dipole sum rule for oscillator strengths, but UCIS does not, and its oscillator strengths can be only qualitatively correct. Nevertheless, for extended $\pi$-states and Rydberg states, the use of the adiabatic LDA kernel in TDDFT
can be problematic, causing disagreement particularly in the high-energy region of the spectra between the two approaches [94].

(2) Peak positions. The approximations of the adiabatic LDA kernel in TDDFT and the restriction to single excitations in CI can cause some inaccuracies in the excitation energies. In particular, UCIS generally overestimates optical gaps somewhat [94]. However, the cost of inclusion of double excitations or use of higher theory levels, such as complete active space self-consistent field, is prohibitive for large systems. The single-particle basis set may also affect our results: the UCIS optical gap decreases with a larger basis (Fig. 5.1(a)). Therefore, the peak positions obtained in TDDFT are expected to be more accurate, at least in the low-energy region.

(3) Modulation effect: If N vibrations are excited by a laser at 96 cm$^{-1}$, the SHG signal will turn on and off periodically as the nitrogen atom oscillates, unlike previous SHG investigations in C$_{60}$ [195, 196, 194, 193] and Li@C$_{60}$ [49], where SHG is on all the time. The oscillation of the nitrogen sets the precise timing for this effect. It may be possible also to see inelastic sidebands in SHG from hyper-Raman scattering (allowed by symmetry for the N vibration), as has been reported for pristine C$_{60}$ [156].

5.5 Conclusion

We have demonstrated that excitation of the vibration of the nitrogen atom inside N@C$_{60}$ allows second-harmonic generation. The first-principles calculations show that with a time-varying symmetry breaking, the SHG signal will be switched on/off periodically with a period of 345 fs. Depending on the position of the N atom, the second-order susceptibility $\chi^{(2)}$ is on the order of $10^{-8}$ esu, which is detectable experimentally. We also predict that the nitrogen vibration will add a new line to the IR spectrum compared to C$_{60}$. These results present an excellent opportunity for future experiments to directly test our theoretical predictions, and more importantly, to explore the SHG modulation effect in N@C$_{60}$.

5.6 Units conversion

The units in nonlinear optics are often confusing. Therefore, we make explicit some of the factors and conversions necessary. The number density $N$ of C$_{60}$ molecules in the fcc solid is

$$N = \frac{4}{a^3} = \frac{4 \times 10^{30}}{14.1613} \text{[1/m}^3\text{]}.$$

(5.4)

To convert the second-order susceptibility to esu, we use the formula (A.2.3 on page 296 of Butcher and Cotter [47])

$$\chi^{(2)}\text{[esu]} = \frac{3 \times 10^4}{4\pi} \chi^{(2)}\text{(m/V)}.$$

(5.5)

The linear susceptibility is dimensionless in any system of units.
In molecules, one often uses the first-order hyperpolarizability $\beta$, which is related to the second-order susceptibility via \([47]\) as

$$\beta = \frac{\epsilon_0 2!}{N} \chi^{(2)}. \quad (5.6)$$

This essentially removes some of the coefficients in Eq. (5.1). This convention is identical to the “Taylor convention” of Willetts et al. \([395]\). 1 a.u. of $\beta$ corresponds to $3.206361 \times 10^{-53}$ C$m^3$/J$^2$ (SI) and also $8.6392 \times 10^{-33}$ cm$^4$ statvolt$^{-1}$ ($\beta$'s esu) \([321, 180]\). Our results for $\chi^{(2)}$ may be converted to $\beta$ (both in esu) by the conversion factor $0.63898 \times 10^{-29}$ cm$^3$. 
Chapter 6

Nonlinear optics and local-field factors in liquid chloroform

6.1 Introduction

Practical applications of organic molecules for nonlinear optics in devices will require assembling them into a condensed phase of some sort such as a crystal or an aligned polymer [46, 77, 180]. The solution phase will probably not be used for devices, although it is critical for use of nonlinear optical probes in biological systems [399]. Nevertheless, nearly all theoretical investigations have been based on gas-phase molecules, and most experimental measurements are performed in solution, complicating the comparison. The solvent may have a significant impact on the nonlinear-optical properties. We would like to assess the effect of the solution environment by explicit calculation of nonlinear optics of a solution, which has so far been done only by modeling the solvent as a simple dielectric [347]. Chloroform is often used as a solvent when measuring nonlinear optical properties of organic molecules, and sometimes as an internal reference [79]. Its properties have been intensively studied [176], making pure chloroform a good choice for initial studies of a solution, which can be compared to gas-phase results (Chapter 4). It has a large solvent effect, as seen from experimental measurements: gas-phase $\beta = 1.0 \pm 4.2$ au [176], while liquid $\beta = -68 \pm 7$ [178], a huge enhancement. These values are from EFISH experiments, which I will now describe.

As discussed in the introduction, a medium with inversion symmetry has no second-order susceptibility. Therefore, experimental tricks are used to obtain a result for second-harmonic generation anyway. One approach is electric-field-induced second-harmonic generation (EFISH), in which a static electric field is used to partially align molecules in solution—this works only if they have a dipole moment, of course [32]. EFISH is really a coherent third-order process, but it has a simple relation to the second-order susceptibility if the molecules are weakly interacting in solution:

$$\chi^{(3)} (-2\omega; \omega, \omega, 0) = n \left[ \gamma (-2\omega; \omega, \omega, 0) + \frac{\mu}{3k_BT} \beta || (-2\omega; \omega, \omega) \right]$$  \hspace{1cm} (6.1)

where $n$ is the number density of molecules, with dipole moment $\mu$ at temperature $T$. 
The specific combination of hyperpolarizability components $\beta_\parallel$ occurring here, if the dipole moment is in direction $i$, is

$$\beta_\parallel = \frac{1}{5} \sum_j (\beta_{ijj} + \beta_{jji} + \beta_{jjj}) \quad (6.2)$$

Experimentally, $\mu\beta_\parallel$ may be extracted by the temperature-dependence, although sometimes $\gamma$ is just assumed to be zero. This method may be used for gas and liquid phases. Another approach is hyper-Rayleigh scattering (HRS), which measures a signal due to the time- and space-dependent fluctuations of the molecular orientations that break inversion symmetry [65]. This is an incoherent second-order process. (Compare ordinary Rayleigh scattering, a linear process based on density fluctuations.) The intensity is given by an average in space and time of products of hyperpolarizabilities,

$$I_{HRS} \propto \langle \beta_{xyz} \beta_{uvw} \rangle \quad (6.3)$$

where $xyz$ and $uvw$ are the coordinate frames of two molecules, assumed independent of one another. In practice correlations between molecules will cause deviations from this formula, which can instead be revised to refer to larger clusters of molecules of size greater than a correlation length in solution, which will then be independent (in particular a large supercell of liquid may be used). The hyperpolarizability components relevant for this process depend on whether the experiment uses the “vertical-vertical” (VV) geometry, in which incident and detected light has the same polarization, or “horizontal-vertical” (HV), in which they have perpendicular polarization. The equations are [74, 29]:

$$[\beta^{VV}_{HRS}]^2 = \frac{1}{7} \sum_i \beta_{iii}^2 + \frac{6}{35} \sum_{i \neq j} \beta_{ii} \beta_{jj} + \frac{9}{35} \sum_{i \neq j} \beta_{ij}^2 + \frac{6}{35} \sum_{\text{cyclic}} \beta_{ii} \beta_{jj} + \frac{12}{35} \beta_{jk}^2 \quad (6.4)$$

$$[\beta^{HV}_{HRS}]^2 = \frac{1}{35} \sum_i \beta_{iii}^2 - \frac{2}{105} \sum_{i \neq j} \beta_{ii} \beta_{jj} + \frac{11}{105} \sum_{i \neq j} \beta_{ij}^2 - \frac{2}{105} \sum_{\text{cyclic}} \beta_{ii} \beta_{jj} + \frac{8}{35} \beta_{jk}^2 \quad (6.5)$$

Mapping the signal as a function of relative polarization angle can be used to extract the multipolar decomposition of the hyperpolarizability tensor [54]. I have implemented these formulas in Octopus [218, 55, 7]. HRS requires fewer assumptions than EFISH (in particular about calibration), but gives much lower signal, and so is more difficult experimentally, and is not practical for gases [176].

## 6.2 Solvent effects

One way the solvent can affect the molecule is by perturbing its geometrical structure in solution. Chloroform is considered an “associated liquid” whose structure is determined essentially by sterics [148], so significant interaction that would change the structure are expected to be small. However, I showed in Chapter 4 that small changes in the geometry can lead to large changes in $\beta$, in particular in the solid geometry. So this may be responsible for some of the liquid enhancement. Another effect of the solvent is a kind of confinement effect: as shown in Chapter 4, the nonlinear response of CHCl$_3$ extends up to 22 au away.
\[
\begin{array}{cccc}
\text{EFISH (gas)} & \text{EFISH (liquid)} & \text{HRS VV (liquid)} & \text{HRS HV (liquid)} \\
\text{exp't} & 1.0 \pm 4.2 & -68 \pm 7 & 15.7 \pm 0.6 & 9.9 \pm 0.4 \\
\text{calc.} & -3.47 & -51.6 & 31.0 & 12.1 \\
\end{array}
\]

Table 6.1: Hyperpolarizability of CHCl₃ in experiments and calculations. Experimental data from Ref. [176] except EFISH (liquid) from Ref. [178], and are a wavelength of 1064 nm. Gas-phase calculations are in 22 au sphere (Chapter 4) and liquid-phase means a crude approximation using a 6 au sphere; all calculations are the electronic contribution at zero frequency.

However, the molecular volume in the liquid (and solid), as determined from the density, corresponds to a sphere with a radius of only 6 au. Clearly the response at long ranges will be strongly modified by the presence of other molecules rather than vacuum. We can crudely estimate the effect by using this sphere as our domain for the real-space wavefunctions, confining them within a small sphere. Using the methodology of Chapter 4, I find a large enhancement in $\beta$ in doing this, similar to what is found in experiment (Table 6.1). Thus confinement due to neighbors seems also to be a significant contributor to the enhancement in the liquid phase.

### 6.3 Local-field factors

The presence of a dielectric medium around a molecule has a very important effect on the response, linear and nonlinear. Simply summing up the polarizabilities of individual molecules to make up the dielectric constant of the solution will give a significant discrepancy, even if the intermolecular interactions are very weak, because of this effect. The reason is that the response of the whole is defined with respect to an external field $E_{\text{ext}}$, while that of the isolated molecule is defined with respect to the local field it experiences, $E_{\text{local}}$ [322]. This is not the same as the screened field $E_{\text{scr}}$ within the medium: that is related to the force experienced by a test charge placed in the medium, whereas the local field is related to the force experienced by a charge that is part of the medium. The screened field depends on the surfaces of the medium (their shape and charge, if any), but in a simple planar capacitor geometry, it is given by $E_{\text{scr}} = E_{\text{ext}}/\varepsilon$. The local-field factor $f$ is defined via $E_{\text{loc}} = fE_{\text{scr}}$. This concept is involved in the Clausius-Mossotti relations for the dielectric constant of a liquid or gas [162]. In the context of hyperpolarizability, the local-field factors are used by experimentalists to extract the property of the individual molecules from bulk measurements [176], via

\[
\chi^{(2)}(-2\omega; \omega, \omega) = n f_2^2 f_{2\omega} \beta (-2\omega; \omega, \omega)
\]

The local-field factors are of order unity, but their crucial importance is seen by the fact a product of three of them enter into this analysis. Typically they are estimated by the Lorentz model [162], which assumes the molecule does not have a dipole moment and occupies a
Figure 6.1: The Lorentz local-field model. A homogeneous dielectric medium (blue) has an applied field to the right, induced positive charges on the right surface and negative on the left, as usual for a capacitor. This reduces the screened field in the medium compared to the external field. However, on the surfaces of a cavity within the medium, the induced charges enhance the local field compared to the screened field. The solution for a spherical cavity is shown.

spherical cavity in a homogeneous dielectric (Fig. 6.1), given simply by $f_\omega = \frac{1}{3} (\epsilon_\omega + 2)$, or a generalization for an ellipsoidal cavity and/or polar solvent due to Onsager [261]:

$$f = \frac{\epsilon (L a, \epsilon_\infty + 1 - L a)}{\epsilon (1 - L a) + L a, \epsilon_\infty} \quad (6.7)$$

where $L a$ is a geometry-dependent depolarization factor. The geometry simply defines the boundary conditions for an electrostatics problem; the polarity of the solvent matters for whether we consider that the medium has a different $\epsilon_0$ and $\epsilon_\infty$ due to reorientation in an applied field.

This models of course make strong assumptions. It is questionable to treat the near-field part of the intermolecular interactions as those with a homogeneous dielectric when we are dealing with small molecules with large dipole moments. Only for a certain ordered lattices does the near-field part cancel out [162]; generally, it will make a significant contribution. We need a better approach to calculate local-field factors without making these assumptions, for a reliable comparison of gas-phase and liquid hyperpolarizabilities. Some recent work has used molecular dynamics in the presence of an applied field [374] to estimate local-field factors via observation of the degree of ordering of CHCl$_3$ dipoles. This is a promising approach, but very time-consuming (if done more rigorously ab initio rather than with a classical potential as they did), and clearly limited only to the static case.

Instead, we can make an ab initio formulation of the local-field factor, allowing its computation from electronic structure. Consider a charge $Z$ inside the dielectric cavity of Fig. 6.1. The force on this charge is $F = Z \epsilon_{\text{local}}$. We may also define an effective charge $Z^*$ via $F = Z^* \epsilon_{\text{scr}}$. Then the local-field factor, according to the definition above, is given by $f = Z^*/Z$. This definition of $Z^*$ is in fact the standard concept of the Born effective charge, as discussed in Chapter 2, so it may be easily computed by standard methods to obtain $f$. However, for a neutral molecule in solution, $Z$ and $Z^*$, if definable, are presumably zero, so this formulation is not usable. On the other hand, we can make a corresponding formulation based on the torque on a dipole due to an applied field. Pasquarello and Resta
have defined an effective dipole moment $\mu^*$ in this way to study liquid water [265]. The torque $\tau = \mu^* \times \mathcal{E}_{\text{loc}}$ is

$$\tau_\alpha = \epsilon_{\alpha\gamma\delta} \sum_s R_{s\gamma} Z^*_{s\beta\delta} [\mathcal{E}_{\text{loc}}]_\beta$$

(6.8)

where $\epsilon_{\alpha\gamma\delta}$ is the Levi-Civita tensor, which yields

$$\mu^*_{\alpha\beta} = \epsilon_{\alpha\gamma\delta} \sum_s R_{s\gamma} Z^*_{s\beta\delta}$$

(6.9)

An intrinsic dipole computed via Wannier-function centers has also been studied for liquid water [325], given by

$$\mu = -e \sum_i r_i + e \sum_s R_s Z_s$$

(6.10)

Thus we can compare the effective $\mu^*$ (which depends upon the screening by the environment) to the one from the Wannier centers $\mu$ to obtain a local-field factor for the molecule in its environment. The works mentioned above seemed to contradict one another about what the dipole moment of liquid water is, but in my point of view the difference is simply a measurement of the local-field factor. As demonstrated by Ref. [265], for a finite system $\mu = \mu^*$, so we have the appropriate result that the local-field factor for an isolated molecule is 1, since there is no environmental screening. This approach, if used for a spherical cavity inside a homogeneous dielectric, would also give the same result as the Lorentz model.

All that is required for this novel method to be applicable is that we can clearly identify which atoms and which Wannier functions belong to which molecule in the condensed phase: this was the case for the computations of water in the literature [265, 325], and is expected to be the case for CHCl$_3$ too, which is more weakly interacting (e.g. no hydrogen bonds). I have performed a preliminary calculation for solid CHCl$_3$, which has four molecules per unit cell in its X-ray diffraction structure [112]. This is just a test case with a clearly defined geometry, but not interesting for nonlinear optics since it has inversion symmetry. I calculated the electronic structure with DFT and the PBE functional [269] in the Quantum ESPRESSO code [119], using the experimental lattice constant of 14.144 au and a plane-wave cutoff of 70 Ry. Using the just the $\Gamma$-point was found to be sufficient. The centers of maximally localized Wannier functions were computed with the Wannier90 code [225, 241], giving a dipole moment of 0.38 D, considerably smaller than the gas-phase value of 1.0 D (Chapter 4). The change in Wannier centers upon displacements of the atoms by 0.02 Å was used to find the Born effective charges. Results are shown in Table 6.2. We obtain $\mu^* = 2.97$ D, for $f = 7.8$. By contrast, the Lorentz model gives 1.4 with the experimental dielectric constant $\epsilon = 2.09$ for liquid CHCl$_3$ at optical frequencies (expected to be similar for the solid) [271]. We find a huge enhancement compared to the simple model, demonstrating the crucial role of reliable local-field factors. Of course, this may be due to the high symmetry of the solid and the effect deviation may be less in the liquid phase.

This formulation includes only electronic contributions to the local-field factor. At zero frequency, there may also be significant ionic contributions due to vibration and rotation of other molecules. This could be handled perturbatively.
Table 6.2: Diagonal part of the Born effective charges for each atom of one molecule in solid CHCl₃. All four molecules in the unit cell are equivalent.

<table>
<thead>
<tr>
<th></th>
<th>Z′ₓₓ</th>
<th>Z′ᵧᵧ</th>
<th>Z′zz</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.40</td>
<td>1.79</td>
<td>1.28</td>
</tr>
<tr>
<td>H</td>
<td>0.17</td>
<td>-0.05</td>
<td>-0.02</td>
</tr>
<tr>
<td>Cl</td>
<td>-0.19</td>
<td>-0.76</td>
<td>-0.32</td>
</tr>
<tr>
<td>Cl</td>
<td>-0.16</td>
<td>-0.28</td>
<td>-1.14</td>
</tr>
<tr>
<td>Cl</td>
<td>-0.19</td>
<td>-0.74</td>
<td>-0.32</td>
</tr>
</tbody>
</table>

6.4 Second-order Sternheimer equation

To perform a complete liquid calculation using a large periodic supercell, we need to use the quantum theory of polarization rather than a finite formulation of electric fields, as discussed in Chapter 2. In particular, we need the quantity \( \frac{\partial^2 \psi}{\partial E \partial k} \) which is not required for linear response [76]. We wish to compute all the derivatives with the Sternheimer equation rather than finite differences of \( k \)-point as was done in Ref. [76] for semiconductors because for a large supercell just \( \Gamma \)-point sampling should be sufficient. To obtain \( \frac{\partial^2 \psi}{\partial E \partial k} \) we need to take the Sternheimer equation to second-order. We will now consider in some generality the properties of that equation.

6.4.1 Overview

I derive and analyze the time-dependent second-order Sternheimer equation, and consider how to solve it, in the context of calculations of hyperpolarizabilities in periodic systems, and other higher-order susceptibilities inaccessible from first-order wavefunctions even by the \( 2n + 1 \) theorem [127]. Only the time-independent version of this equation appears to have been published (within general considerations of the hierarchy of all order of Sternheimer equations) [127, 123], whereas the time-dependent version has not, and it does not seem to have been employed in any calculation. (The method of calculation of Raman tensors developed in Ref. [202], and used in Quantum ESPRESSO [119], calculates second-order response but does so via sums over states rather than a Sternheimer equation.) The first-order equation has been studied in great detail in the static case [19] and implemented in the dynamic case for TDDFT [8], as discussed in Chapter 2. I will show in detail the derivation of the first- and second-order equations, and point out key differences between the second-order and first-order equation:

1. There are separate equations for the sum frequency, difference frequency, and second-order frequency (if present).

2. The second-order equation directly involves the first-order Hamiltonians, eigenvalues, and wavefunctions (and not just the projection into the unoccupied subspace).
3. The second-order density requires the second-order wavefunction response in the occupied subspace, so the second-order Sternheimer equation cannot be projected into the unoccupied subspace (except if a non-self-consistent perturbation is used).

4. The projection of the second-order wavefunction onto the unperturbed wavefunction is not given by the Sternheimer equation, and must be determined separately from the normalization condition; this result must be used in calculation of the second-order density and of observables.

6.4.2 Hamiltonian

We consider a Hamiltonian consisting of an unperturbed Hamiltonian $\hat{H}^{(0)}$, two first-order perturbations $\hat{H}^{(1)}_{\alpha}$ and $\hat{H}^{(1)}_{\beta}$, and second-order perturbations $\hat{H}^{(2)}$. The perturbations are time-dependent, and each has its own frequency. The magnitudes of the perturbations are controlled by two small parameters, $\alpha$ and $\beta$. Each of these perturbations is driven by a bare perturbation, but in turn can create a density response $n^{(1)}$ that produces a further perturbation via the Hartree potential and exchange-correlation potential. (The $k \cdot p$ perturbation however is non-self-consistent and has only a bare term.) At second order, the response of the two first-order perturbations creates a perturbation at the sum and difference frequencies, and there can also be a bare perturbation at these frequencies. Most generally then, our perturbations can have these forms:

$$\hat{H}^{(1)}_{\alpha \pm} = V^{(1)}_{\alpha \pm} + V_{\text{H}} \left[ n^{(1)}_{\alpha \pm} \right] + \int f_{xc} \left[ n \left( \pm \omega_\alpha \right) n^{(1)}_{\alpha \pm} \left( r \right) \right] \right] \, dr \tag{6.11}$$

$$\hat{H}^{(1)}_{\beta \pm} = V^{(1)}_{\beta \pm} + V_{\text{H}} \left[ n^{(1)}_{\beta \pm} \right] + \int f_{xc} \left[ n \left( \pm \omega_\beta \right) n^{(1)}_{\beta \pm} \left( r \right) \right] \, dr \tag{6.12}$$

$$\hat{H}^{(2)} = V^{(2)}_{\pm} + V_{\text{H}} \left[ n^{(2)}_{\pm} \right] + \int f_{xc} \left[ n \left( \pm \omega_2 \right) n^{(2)}_{\pm} \left( r \right) \right] \, dr \tag{6.13}$$

$$\hat{H}_{\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta} = V^{(2)}_{\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta} + V_{\text{H}} \left[ n^{(2)}_{\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta} \right] + \int \int K_{xc} \left[ n \left( \sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta \right) n^{(1)}_{\sigma_\alpha \omega_\alpha} \left( r \right) n^{(1)}_{\sigma_\beta \omega_\beta} \left( r' \right) \right] \, dr \, dr' + \int f_{xc} \left[ n \left( \sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta \right) n^{(2)}_{\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta} \left( r \right) \right] \, dr \tag{6.14}$$

where $\sigma_\alpha, \sigma_\beta = \pm$, to give four equations at the sum and difference frequencies. Note that the sum- and difference-frequency perturbations may be nonzero even if there is no bare perturbation at these frequencies, due to the $K_{xc}$ term with the product of first-order response densities. If we assume the adiabatic approximation, then we can remove frequency-dependence from the exchange-correlation potential and its functional derivatives. As we will discuss in Sec. 6.4.8, there is a condition $V^{(1)}_{\pm} = V^{(1)}_{\mp}$ on the bare perturbations.

For maximal generality we will include the perturbation $\hat{H}^{(2)}_{\pm}$ with its own separate frequency-dependence, but it is not obvious when such a term would actually be needed. The end result anyway is that it does not couple with other terms, and can be treated exactly analogously to perturbations in the first-order Sternheimer equation.
Taking all these terms together, our Hamiltonian is:

\[ \hat{H} = \hat{H}^{(0)} + \frac{1}{2} \alpha \left( \hat{H}^{(1)}_{\alpha+} e^{i\omega_{\alpha} t} + \hat{H}^{(1)}_{\alpha-} e^{-i\omega_{\alpha} t} \right) + \frac{1}{2} \beta \left( \hat{H}^{(1)}_{\beta+} e^{i\omega_{\beta} t} + \hat{H}^{(1)}_{\beta-} e^{-i\omega_{\beta} t} \right) + \frac{1}{2} \alpha \beta \left( \hat{H}^{(2)}_{\sigma\alpha\omega_{\alpha} + \sigma\beta\omega_{\beta}} e^{i(\sigma_{\alpha}\omega_{\alpha} + \sigma_{\beta}\omega_{\beta}) t} + \sum_{\sigma_{\alpha}\sigma_{\beta}} \hat{H}^{(2)}_{\sigma\alpha\omega_{\alpha} + \sigma\beta\omega_{\beta}} e^{i(\sigma_{\alpha}\omega_{\alpha} + \sigma_{\beta}\omega_{\beta}) t} \right) \] (6.15)

where

\[ \hat{H}^{(0)} = \frac{\mathbf{p} + \mathbf{k}}{2m} + V_{H}[n] + V_{xc}[n] + V_{ps} \] (6.16)

For hyperpolarizabilities in a periodic system, we need \( V^{(1)}_{\alpha} = r, V^{(1)}_{\beta} = -i\nabla + \mathbf{k}, \omega_{\beta} = 0, \) \( V^{(2)} = 0. \) For third- and fourth-order nonlinear susceptibilities, we would have \( V^{(1)}_{\alpha} = V^{(1)}_{\beta} = r, V^{(2)} = 0. \) For phonon anharmonicities, we would have \( V^{(1)}_{\alpha} = V^{(1)}_{\beta} = \nabla V_{ps}, \) \( V^{(2)} = \nabla \nabla V_{ps} \) at the sum and difference frequencies, if taking into account time-dependence of phonons [203, 310].

### 6.4.3 Wavefunctions and eigenvalues

The wavefunction in turn contains components oscillating at various frequencies, which will find to be demanded by the solution of the final equation. For simplicity, we will omit band and \( \mathbf{k} \)-point indices except where necessary.

\[ \Psi (t) = e^{-i\omega_{\alpha} t} \psi \] (6.17)

\[ \psi = \psi^{(0)} + \frac{1}{2} \alpha \left( \psi^{(1)}_{\alpha+} e^{i\omega_{\alpha} t} + \psi^{(1)}_{\alpha-} e^{-i\omega_{\alpha} t} \right) + \frac{1}{2} \beta \left( \psi^{(1)}_{\beta+} e^{i\omega_{\beta} t} + \psi^{(1)}_{\beta-} e^{-i\omega_{\beta} t} \right) + \frac{1}{2} \alpha \beta \left( \psi^{(2)}_{\alpha\omega_{\alpha} + \beta\omega_{\beta}} e^{i(\alpha\omega_{\alpha} + \beta\omega_{\beta}) t} + \sum_{\sigma} \psi^{(2)}_{\sigma\omega_{\alpha} + \beta\omega_{\beta}} e^{i(\sigma\omega_{\alpha} + \beta\omega_{\beta}) t} \right) + ... \] (6.18)

The time-dependent Schrödinger equation is \( i\frac{d}{dt} \Psi = \hat{H} \Psi. \) The unperturbed wavefunction \( \Psi^{(0)} = e^{-i\epsilon^{(0)} t} \psi^{(0)} \) satisfies additionally the time-independent equation \( \hat{H}^{(0)} \psi^{(0)} = \epsilon^{(0)} \psi^{(0)}. \) The eigenvalues are related to the expectation values of the Hamiltonian, and as such can be generalized to the time-dependent case, when the eigenvalues are no longer of energy but of the Floquet quasi-energy [201]: \( \tilde{\epsilon}^{(0)} = \langle \psi^{(0)} | \hat{H}^{(0)} | \psi^{(0)} \rangle, \tilde{\epsilon} = \langle \Psi | \hat{H} | \Psi \rangle = \langle \psi | \hat{H} | \psi \rangle. \) However, to make our formulation preserve normalization [201], we add a correction to the eigenvalues for the time evolution: \( \epsilon^{(n)} = \tilde{\epsilon}^{(n)} + \hbar \omega_{n} \langle \psi^{(0)} | \psi^{(n)} \rangle, \) where we consider \( \omega_{0} = 0 \) so that \( \epsilon^{(0)} = \tilde{\epsilon}^{(0)}. \) The total \( \epsilon \) is then

\[ \epsilon = \epsilon^{(0)} + \frac{1}{2} \alpha \left( \epsilon^{(1)}_{\alpha+} e^{i\omega_{\alpha} t} + \epsilon^{(1)}_{\alpha-} e^{-i\omega_{\alpha} t} \right) + \frac{1}{2} \beta \left( \epsilon^{(1)}_{\beta+} e^{i\omega_{\beta} t} + \epsilon^{(1)}_{\beta-} e^{-i\omega_{\beta} t} \right) + \frac{1}{2} \alpha \beta \left( \epsilon^{(2)}_{+} e^{i\omega_{\alpha} t} + \epsilon^{(2)}_{-} e^{-i\omega_{\alpha} t} + \sum_{\sigma} \epsilon^{(2)}_{\sigma\omega_{\alpha} + \beta\omega_{\beta}} e^{i(\sigma\omega_{\alpha} + \beta\omega_{\beta}) t} \right) + ... \] (6.19)
6.4.4 Substitution into Schrödinger equation

We substitute our expressions for $\Psi$ into $i\frac{d}{dt}\Psi = \hat{H}\Psi$:

$$
\left(\epsilon\psi + i\frac{d}{dt}\psi\right)e^{-i\epsilon t} = \hat{H}\psi e^{-i\epsilon t}
$$

(6.20)

The phase can be cancelled from both sides:

$$
\epsilon\psi + i\frac{d}{dt}\psi = \hat{H}\psi
$$

(6.21)

Now we substitute for $\epsilon$, $\psi$, and $\hat{H}$:

$$
\left[\epsilon^{(0)} + \frac{1}{2} \alpha \left(\epsilon^{(1)} + \epsilon^{(1)}_\alpha - e^{-i\omega_\alpha t}\right) + \frac{1}{2} \beta \left(\epsilon^{(1)}_\beta + \epsilon^{(1)}_\beta - e^{-i\omega_\beta t}\right) + \frac{1}{2} \alpha \beta \left(\epsilon^{(2)}_+ e^{i\omega_\alpha t} + \epsilon^{(2)}_+ e^{-i\omega_\alpha t} + \sum_{\sigma\alpha,\sigma\beta} \epsilon^{(2)}_{\sigma\alpha,\sigma\beta} e^{i(\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta)t}ight) + \ldots \right]
$$

$$
\left[\psi^{(0)} + \frac{1}{2} \alpha \left(\psi^{(1)}_\alpha + \psi^{(1)}_\alpha - e^{-i\omega_\alpha t}\right) + \frac{1}{2} \beta \left(\psi^{(1)}_\beta + \psi^{(1)}_\beta - e^{-i\omega_\beta t}\right) + \frac{1}{2} \alpha \beta \left(\psi^{(2)}_+ e^{i\omega_\alpha t} + \psi^{(2)}_+ e^{-i\omega_\alpha t} + \sum_{\sigma\alpha,\sigma\beta} \psi^{(2)}_{\sigma\alpha,\sigma\beta} e^{i(\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta)t}ight) + \ldots \right]
$$

$$
+ \frac{1}{2} \alpha \beta \left(\omega_\alpha \psi^{(1)}_\alpha - i\omega_\alpha \psi^{(1)}_\alpha - e^{-i\omega_\alpha t}\right) + \frac{1}{2} \beta \left(\omega_\beta \psi^{(1)}_\beta - i\omega_\beta \psi^{(1)}_\beta - e^{-i\omega_\beta t}\right) + \frac{1}{2} \alpha \beta \left(\omega_\alpha \psi^{(2)}_+ e^{i\omega_\alpha t} - \omega_\alpha \psi^{(2)}_+ e^{-i\omega_\alpha t} + \frac{1}{2} \sum_{\sigma\alpha,\sigma\beta} \omega_\sigma \omega_\alpha \psi^{(2)}_{\sigma\alpha,\sigma\beta} e^{i(\sigma_\sigma \omega_\sigma + \sigma_\beta \omega_\beta)t}\right) + \ldots
$$

(6.22)

6.4.5 The first-order time-dependent Sternheimer equation

Taking the zeroth-order terms, we have the time-independent Schrödinger equation:

$$
\epsilon^{(0)} \psi^{(0)} = \hat{H}^{(0)} \psi^{(0)}
$$

(6.23)
The terms linear in $\alpha$ give us the time-dependent first-order Sternheimer equations for the first perturbation:

$$\frac{1}{2} \alpha e^{(0)} \left( \psi^{(1)}_{\alpha+} e^{i\omega_n t} + \psi^{(1)}_{\alpha-} e^{-i\omega_n t} \right) + \frac{1}{2} \alpha \left( e^{(1)}_{\alpha+} e^{i\omega_n t} + e^{(1)}_{\alpha-} e^{-i\omega_n t} \right) \psi^{(0)} = \frac{1}{2} \alpha \hat{H}^{(0)} \left( \psi^{(1)}_{\alpha+} e^{i\omega_n t} + \psi^{(1)}_{\alpha-} e^{-i\omega_n t} \right) + \frac{1}{2} \alpha \left( \hat{H}^{(1)}_{\alpha+} e^{i\omega_n t} + \hat{H}^{(1)}_{\alpha-} e^{-i\omega_n t} \right) \psi^{(0)}$$

(6.24)

Dividing out the common factor of $\frac{1}{2} \alpha$,

$$\epsilon^{(0)} \left( \psi^{(1)}_{\alpha+} e^{i\omega_n t} + \psi^{(1)}_{\alpha-} e^{-i\omega_n t} \right) + \epsilon^{(1)}_{\alpha+} e^{i\omega_n t} + \epsilon^{(1)}_{\alpha-} e^{-i\omega_n t} \psi^{(0)} = \hat{H}^{(0)} \left( \psi^{(1)}_{\alpha+} e^{i\omega_n t} + \psi^{(1)}_{\alpha-} e^{-i\omega_n t} \right) + \hat{H}^{(1)}_{\alpha+} e^{i\omega_n t} + \hat{H}^{(1)}_{\alpha-} e^{-i\omega_n t} \psi^{(0)}$$

(6.25)

The equation can only be satisfied for all times if it is satisfied separately for positive and negative frequencies:

$$\epsilon^{(0)} \psi^{(1)}_{\alpha\pm} e^{\pm i\omega_n t} + \epsilon^{(1)}_{\alpha\pm} e^{\pm i\omega_n t} \psi^{(0)} = \omega_{\alpha} \psi^{(1)}_{\alpha\pm} e^{\pm i\omega_n t} = \hat{H}^{(0)} \psi^{(1)}_{\alpha\pm} e^{\pm i\omega_n t} + \hat{H}^{(1)}_{\alpha\pm} e^{\pm i\omega_n t} \psi^{(0)}$$

(6.26)

Removing the phases and collecting terms,

$$\left( \hat{H}^{(0)} - \epsilon^{(0)} \pm \omega_{\alpha} \right) \psi^{(1)}_{\alpha\pm} = - \left( \hat{H}^{(1)} - \epsilon^{(1)}_{\alpha\pm} \right) \psi^{(0)}$$

(6.27)

The equations for $+\omega$ and $-\omega$, and those for each occupied state, are coupled via the density-dependence of $\hat{H}^{(1)}$.

Since we will see in Sec. 6.4.7 that the projection onto the unperturbed wavefunction is zero, we can simplify the equation by projecting out that component, which removes the eigenvalue term:

$$\left( \hat{H}^{(0)} - \epsilon^{(0)} \pm \omega_{\alpha} \right) P'_{n} \psi^{(1)}_{\alpha\pm} = - P'_{n} \hat{H}^{(1)}_{\alpha\pm} \psi^{(0)}$$

(6.28)

where $P'_{n} = 1 - \left| \psi^{(0)}_{\alpha\pm} \right| \left< \psi^{(0)}_{\alpha\pm} \right|$. We then solve not for $\psi^{(1)}_{\alpha\pm}$ but $P'_{n} \psi^{(1)}_{\alpha\pm}$. If the state is initialized to zero, then the projection onto the unperturbed wavefunction will remain zero in the iterative solution.

We can also project out all the occupied states since the response of the wavefunctions in the occupied subspace does not contribute to the response of the density. Additionally, many properties to be calculated from the response wavefunctions do not depend on the components in the occupied subspace. The null space of the projection operator reduces the space in which the linear solver must search for the solution, reducing the number of iterations required to solve for each state (the number of SCF iterations is only slightly reduced however).
In this form, we solve for $P_c\psi_{\alpha\pm}^{(1)}$ via:

$$\left(\hat{H}^{(0)} - \epsilon^{(0)} \pm \omega_{\alpha}\right) P_c\psi_{\alpha\pm}^{(1)} = -P_c\hat{H}_{\alpha\pm}^{(1)}\psi^{(0)}$$

(6.29)

Again, initialization to zero will make the projection into the occupied subspace remain zero in the iterative solution. Of course, the same arguments apply to give us the equation for $\beta$:

$$\left(\hat{H}^{(0)} - \epsilon^{(0)} \pm \omega_{\beta}\right) P_c\psi_{\beta\pm}^{(1)} = -P_c\hat{H}_{\beta\pm}^{(1)}\psi^{(0)}$$

(6.30)

Applying the projector is $O\left(N^2\right)$ for each state, and so $O\left(N^3\right)$ overall. For large systems this operation will dominate the rest of the Sternheimer procedure which is only $O\left(N^2\right)$, in which case it is faster to use $P'_n$ and accept a larger number of linear-solver iterations in exchange for each iteration being quicker [8]. Even when using $P_c$, when the whole wavefunction is required (as will be the case for the second-order equation), we can substitute $P'_n$ in the last SCF iteration, and thus gain the whole wavefunction for only a small increase in computation time.

### 6.4.6 The second-order time-dependent Sternheimer equation

Now we consider the second-order terms, those proportional to $\alpha\beta$:

$$\epsilon^{(0)} \frac{1}{2} \alpha \beta \left( \psi_+^{(2)} e^{i\omega_2 t} + \psi_-^{(2)} e^{-i\omega_2 t} + \sum_{\sigma\alpha\sigma\beta} \psi_{\sigma\alpha\omega_{\alpha} + \sigma\beta\omega_{\beta}}^{(2)} e^{i(\sigma_{\alpha}\omega_{\alpha} + \sigma_{\beta}\omega_{\beta})t} \right) +$$

$$\frac{1}{2} \alpha \left( \epsilon_+^{(1)} e^{i\omega_1 t} + \epsilon_-^{(1)} e^{-i\omega_1 t} \right) \frac{1}{2} \beta \left( \psi_+^{(1)} e^{i\omega_2 t} + \psi_-^{(1)} e^{-i\omega_2 t} \right) +$$

$$\frac{1}{2} \beta \left( \epsilon_+^{(1)} e^{i\omega_1 t} + \epsilon_-^{(1)} e^{-i\omega_1 t} \right) \frac{1}{2} \alpha \left( \psi_+^{(1)} e^{i\omega_2 t} + \psi_-^{(1)} e^{-i\omega_2 t} \right) +$$

$$\frac{1}{2} \alpha \beta \left( \psi_+^{(2)} e^{i\omega_2 t} + \psi_-^{(2)} e^{-i\omega_2 t} + \sum_{\sigma\alpha\sigma\beta} \epsilon_{\alpha\sigma\alpha + \sigma\beta\omega_{\beta}}^{(2)} e^{i(\sigma_{\alpha}\omega_{\alpha} + \sigma_{\beta}\omega_{\beta})t} \right) \psi^{(0)} +$$

$$\frac{i}{2} \alpha \beta \left( \omega_2 \psi_+^{(2)} e^{i\omega_2 t} - \omega_2 \psi_-^{(2)} e^{-i\omega_2 t} + i \sum_{\sigma\alpha\sigma\beta} (\sigma_{\alpha}\omega_{\alpha} + \sigma_{\beta}\omega_{\beta}) \psi_{\sigma\alpha\omega_{\alpha} + \sigma\beta\omega_{\beta}}^{(2)} e^{i(\sigma_{\alpha}\omega_{\alpha} + \sigma_{\beta}\omega_{\beta})t} \right)$$

$$= \hat{H}^{(0)} \frac{1}{2} \alpha \beta \left( \psi_+^{(2)} e^{i\omega_2 t} + \psi_-^{(2)} e^{-i\omega_2 t} + \sum_{\sigma\alpha\sigma\beta} \psi_{\sigma\alpha\omega_{\alpha} + \sigma\beta\omega_{\beta}}^{(2)} e^{i(\sigma_{\alpha}\omega_{\alpha} + \sigma\beta\omega_{\beta})t} \right)$$

$$+ \frac{1}{2} \alpha \left( \hat{H}_+^{(1)} e^{i\omega_1 t} + \hat{H}_-^{(1)} e^{-i\omega_1 t} \right) \frac{1}{2} \beta \left( \psi_+^{(1)} e^{i\omega_2 t} + \psi_-^{(1)} e^{-i\omega_2 t} \right)$$

$$+ \frac{1}{2} \beta \left( \hat{H}_+^{(1)} e^{i\omega_1 t} + \hat{H}_-^{(1)} e^{-i\omega_1 t} \right) \frac{1}{2} \alpha \left( \psi_+^{(1)} e^{i\omega_2 t} + \psi_-^{(1)} e^{-i\omega_2 t} \right)$$

$$+ \frac{1}{2} \alpha \beta \left( \hat{H}_+^{(2)} e^{i\omega_2 t} + \hat{H}_-^{(2)} e^{-i\omega_2 t} + \sum_{\sigma\alpha\sigma\beta} \hat{H}_{\sigma\alpha\omega_{\alpha} + \sigma\beta\omega_{\beta}}^{(2)} e^{i(\sigma_{\alpha}\omega_{\alpha} + \sigma\beta\omega_{\beta})t} \right) \psi^{(0)}$$

(6.31)
Immediately we may divide by the prefactor $\frac{1}{2}\alpha\beta$ which appears in every term. Then we separate the equation into three pairs of equations, for the positive and negative of the second-order frequency, and the sum and difference of the first-order frequencies. The equation at the second-order frequency is:

$$
\epsilon^{(0)}(\psi^{(2)}_+ e^{\pm i\omega_2 t} + \epsilon^{(2)} e^{\pm i\omega_2 t} \psi^{(0)}_+ + \omega_2 \psi^{(2)}_+ e^{\pm i\omega_2 t} = \hat{H}^{(0)}(\psi^{(2)}_+ e^{\pm i\omega_2 t} + \hat{H}^{(2)}_+ e^{\pm i\omega_2 t} \psi^{(0)}_+ (6.32)
$$

This is the same form as the first-order equation, so we can write it simply as

$$
\left(\hat{H}^{(0)} - \epsilon^{(0)} \pm \omega_2 \right) P_c \psi^{(2)}_+ = -P_c \hat{H}^{(2)}_+ \psi^{(0)}_+ (6.33)
$$

The sum-frequency equation is:

$$
\epsilon^{(0)}(\psi^{(2)}_+ (\omega_1 + \omega_2) e^{\pm i(\omega_1 + \omega_2)t} + \frac{1}{2} \left( \epsilon^{(1)}(\omega_2) e^{\pm i\omega_2 t} \psi^{(2)}_+ e^{\pm i\omega_2 t} + \epsilon^{(1)}(\omega_1) e^{\pm i\omega_1 t} \psi^{(2)}_+ e^{\pm i\omega_1 t} \right) +
$$

$$
\left( \epsilon^{(2)}(\omega_1 + \omega_2) e^{i(\omega_1 + \omega_2)t} + \epsilon^{(2)}(\omega_1 - \omega_2) e^{i(-\omega_1 - \omega_2)t} \right) \psi^{(0)}_+ \mp (\omega_1 + \omega_2) \psi^{(2)}_+ (\omega_1 + \omega_2) e^{i(\omega_1 + \omega_2)t} +
$$

$$
\frac{1}{2} \left( \hat{H}^{(1)}(\omega_2) e^{\pm i\omega_2 t} \psi^{(1)}_+ e^{\pm i\omega_2 t} + \hat{H}^{(1)}(\omega_1) e^{\pm i\omega_1 t} \psi^{(1)}_+ e^{\pm i\omega_1 t} \right) +
$$

$$
\frac{1}{2} \left( \hat{H}^{(2)}_+ e^{i(\omega_1 + \omega_2)t} + \hat{H}^{(2)}_+ e^{i(-\omega_1 - \omega_2)t} \right) \psi^{(0)}_+ (6.34)
$$

Unlike the second-order-frequency equation, it contains first-order quantities. Removing the phases and collecting terms, we have

$$
\left(\hat{H}^{(0)} - \epsilon^{(0)} \pm (\omega_1 + \omega_2) \right) \psi^{(2)}_+ =
$$

$$
-\frac{1}{2} \left[ \left( \hat{H}^{(1)}_+ - \epsilon^{(1)}_+ \right) \psi^{(1)}_+ + \left( \hat{H}^{(1)}_- - \epsilon^{(1)}_- \right) \psi^{(1)}_- \right] - \left( \hat{H}^{(2)}_+ - \epsilon^{(2)}_+ \psi^{(0)}_+ \right) (6.35)
$$

We do not project onto the unoccupied subspace because at second order the whole wavefunction response is required to calculate the density (Sec. 6.4.8). However, we can project out the term corresponding to the unperturbed wavefunction, which will remove from consideration a tautologous component of the equation (see Sec. 6.4.7), and we thus dispose of the eigenvalue term. Then we have

$$
\left(\hat{H}^{(0)} - \epsilon^{(0)} \pm (\omega_1 + \omega_2) \right) P_n \psi^{(2)}_n =
$$

$$
-\frac{1}{2} P_n \left[ \left( \hat{H}^{(1)}_n - \epsilon^{(1)}_n \right) \psi^{(1)}_+ + \left( \hat{H}^{(1)}_n - \epsilon^{(1)}_n \right) \psi^{(1)}_- \right] - P_n \hat{H}^{(2)}_n \psi^{(0)}_n (6.36)
$$

The analogous difference-frequency equation is

$$
\left(\hat{H}^{(0)} - \epsilon^{(0)} \pm (\omega_1 - \omega_2) \right) P_n \psi^{(2)}_n =
$$

$$
-\frac{1}{2} P_n \left[ \left( \hat{H}^{(1)}_n - \epsilon^{(1)}_n \right) \psi^{(1)}_+ + \left( \hat{H}^{(1)}_n - \epsilon^{(1)}_n \right) \psi^{(1)}_- \right] - P_n \hat{H}^{(2)}_n \psi^{(0)}_n (6.37)
$$

As with the first-order equations, the density-dependence of the perturbation Hamiltonian couples together the equations for $+\omega$ and $-\omega$, and those for all occupied states.
The sum-frequency and difference-frequency equations, however, are uncoupled. We can also use a projector onto the unoccupied subspace as in the first-order case, instead of the minimal projector:

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \pm (\omega_{\alpha} + \omega_{\beta}) \right) P_c \psi^{(2)}_{\pm (\omega_{\alpha} + \omega_{\beta})} = -\frac{1}{2} P_c \left[ \left( \hat{H}^{(1)}_{\alpha \pm} - \epsilon^{(1)}_{\alpha \pm} \right) \psi^{(1)}_{\beta \pm} + \left( \hat{H}^{(1)}_{\beta \pm} - \epsilon^{(1)}_{\beta \pm} \right) \psi^{(1)}_{\alpha \pm} \right] - P_c \hat{H}^{(2)}_{\pm (\omega_{\alpha} + \omega_{\beta})} \psi^{(0)} \quad (6.38)
\]

To solve these equations, we need the perturbed wavefunctions, densities, eigenvalues, and Hamiltonians from the first-order equations. We can calculate the required eigenvalues easily using the Hellman-Feynman theorem and the fact that the projection onto the unperturbed states is zero:

\[
\epsilon^{(1)} = \tilde{\epsilon}^{(1)} + \hbar \omega_1 \left\langle \psi^{(0)} | \psi^{(1)} \right\rangle = \left\langle \psi^{(0)} | \hat{H}^{(1)} | \psi^{(0)} \right\rangle \quad (6.39)
\]

We need \(\psi^{(1)}\), not just \(P_c \psi^{(1)}\), for the second-order Sternheimer equation. Therefore we should use the \(P_c^{\prime}\) form of the first-order equation.

Another approach would be to get the rest of the wavefunction via a sum over the other occupied states \(m\):

\[
\psi^{(1)}_{n \pm} = P_c \psi^{(1)}_{\pm} + \sum_{m \neq n}^{\text{occ}} \frac{\langle \psi^{(0)}_m | \hat{H}^{(1)}_{\pm} | \psi^{(0)}_n \rangle}{\epsilon^{(0)}_m - \epsilon^{(0)}_n \pm \omega} \quad (6.40)
\]

Avoiding sums over states is a major advantage of the first-order Sternheimer equation, but it is sums over the infinite number of states that is difficult to converge; a sum merely over occupied states is much faster and has no issue of convergence. Nevertheless, performing this sum for all states scales as \(O \left( N^3 \right)\) in contrast to the \(O \left( N^2 \right)\) scaling for solution of the Sternheimer equation [8]. For small systems, the sum over states may be faster, but for large systems it will be preferable to directly solve the first-order Sternheimer equation for the whole wavefunction.

The second-order Sternheimer equation is a linear equation like the first-order equation, and can be solved by the same numerical techniques. The only real difference is that we have a constant, non-self-consistent term on the right-hand side, due to the first-order perturbation, which makes the equation non-homogeneous. This term can be calculated just once, at the beginning of the self-consistent cycle, and does not need to be updated.

### 6.4.7 Normalization

The normalization condition is

\[
\langle \Psi | \Psi \rangle = \langle \psi | \psi \rangle = 1 \quad (6.41)
\]
Substitute in the expression for ψ:

\[ \langle \psi^{(0)} | + \frac{1}{2} \alpha ( \psi^{(1)}_{\alpha+} e^{i\omega_\alpha t} + \psi^{(1)}_{\alpha-} e^{-i\omega_\alpha t}) + \frac{1}{2} \beta ( \psi^{(1)}_{\beta+} e^{i\omega_\beta t} + \psi^{(1)}_{\beta-} e^{-i\omega_\beta t}) \]

\[ + \frac{1}{2} \alpha \beta ( \psi^{(2)}_+ e^{i\omega_\alpha t} + \psi^{(2)}_- e^{-i\omega_\beta t} + \sum_{\sigma_\alpha \sigma_\beta} \psi^{(2)}_{\sigma_\alpha \sigma_\beta} e^{i(\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta) t} ) + \ldots ) \]

\[ - \sum_{\alpha} \psi^{(1)}_+ e^{i\omega_\alpha t} - \sum_{\beta} \psi^{(1)}_- e^{-i\omega_\beta t} + \sum_{\sigma_\alpha \sigma_\beta} \psi^{(2)}_{\sigma_\alpha \sigma_\beta} e^{i(\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta) t} + \ldots ) \] (6.42)

The zeroth-order terms give the normalization of the unperturbed wavefunction:

\[ \langle \psi^{(0)} | \psi^{(0)} \rangle = 1 \] (6.43)

The first-order terms are:

\[ \langle \psi^{(0)} | \alpha ( \psi^{(1)}_{\alpha+} e^{i\omega_\alpha t} + \psi^{(1)}_{\alpha-} e^{-i\omega_\alpha t}) + \beta ( \psi^{(1)}_{\beta+} e^{i\omega_\beta t} + \psi^{(1)}_{\beta-} e^{-i\omega_\beta t}) \rangle + \text{cc.} = 0 \] (6.44)

We can separate by frequency-dependence to get:

\[ \langle \psi^{(0)} | \psi^{(1)}_{\alpha+} \rangle + \langle \psi^{(1)}_{\alpha-} | \psi^{(0)} \rangle = 0 \] (6.45)

\[ \langle \psi^{(0)} | \psi^{(1)}_{\beta+} \rangle + \langle \psi^{(1)}_{\beta-} | \psi^{(0)} \rangle = 0 \] (6.46)

Note that if we take the overlap of the first-order Sternheimer equation with the unperturbed wavefunction, we get

\[ \langle \psi^{(0)} | \hat{H}^{(0)} - \epsilon^{(0)} \pm \omega_\alpha | \psi^{(1)}_{\alpha\pm} \rangle = - \langle \psi^{(0)} | \hat{H}^{(1)}_{\alpha\pm} - \epsilon^{(1)}_{\alpha\pm} | \psi^{(0)} \rangle \] (6.47)

\[ \pm \omega_\alpha \langle \psi^{(0)} | \psi^{(1)}_{\alpha\pm} \rangle + \frac{1}{2} \omega_\alpha \langle \psi^{(0)} | \psi^{(1)}_{\alpha\pm} \rangle \] (6.48)

\[ \omega_\alpha \langle \psi^{(0)} | \psi^{(1)}_{\alpha\pm} \rangle = \omega_\alpha \langle \psi^{(0)} | \psi^{(1)}_{\alpha\pm} \rangle \] (6.49)

We are left with a tautology, so that the Sternheimer equation does not determine this overlap. The sum is fixed by the normalization condition, and additionally we make the phase convention \( \langle \psi^{(0)} | \psi^{(1)}_{\alpha+} \rangle - \langle \psi^{(1)}_{\alpha-} | \psi^{(0)} \rangle = 0 \).

Perturbation theory cannot determine this component of the response wavefunction anyway, as it relates to an arbitrary phase factor between the perturbed and unperturbed wavefunction, and so the convention is to set it to zero (the so-called “parallel-transport gauge”) [300]. Within the Berry-phase theory of polarization, the phase is part of a choice of gauge within the Brillouin zone. Use of the projection operator \( P_n^0 \) or \( P_c \) in the first-order equation sets these overlap matrix elements to zero automatically, so the normalization condition is properly enforced.
The second-order terms are:

$$\left\langle \psi^{(0)} \right| \frac{1}{2} \alpha \beta \left( \psi^{(2)}_+ e^{i \omega_1 t} + \psi^{(2)}_- e^{-i \omega_1 t} + \sum_{\sigma \alpha \beta} \psi^{(2)}_{\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta} e^{i (\sigma_\alpha \omega_\alpha + \sigma_\beta \omega_\beta) t} \right) \right\rangle + \text{cc.}$$

$$+ \left\langle \frac{1}{2} \alpha \left( \psi^{(1)}_\alpha e^{i \omega_1 t} + \psi^{(1)}_\alpha e^{-i \omega_1 t} \right) \right\rangle + \frac{1}{2} \beta \left( \psi^{(1)}_\beta e^{i \omega_1 t} + \psi^{(1)}_\beta e^{-i \omega_1 t} \right) \right\rangle + \text{cc.} = 0 \quad (6.50)$$

Separating by frequency-dependence:

$$\left\langle \psi^{(0)} \right| \frac{1}{2} \alpha \beta \left( \psi^{(2)}_+ e^{i \omega_1 t} + \psi^{(2)}_- e^{-i \omega_1 t} \right) \right\rangle = 0 \quad (6.51)$$

$$+ \frac{1}{2} \left[ \left\langle \psi^{(1)}_\sigma e^{i \omega_1 t} + \psi^{(1)}_\sigma e^{-i \omega_1 t} \right\rangle \right]$$

$$\left\langle \psi^{(1)}_\sigma \right| \psi^{(1)}_\sigma \right\rangle + \left\langle \psi^{(1)}_{-\sigma} \right| \psi^{(1)}_{-\sigma} \right\rangle \right| \psi^{(0)} \right\rangle = 0 \quad (6.52)$$

The situation for $\psi^{(2)}_\pm$ at the second-order frequency is identical to that of $\psi^{(1)}$, in that the solution of the second-order-frequency Sternheimer equation gives us the same result:

$$\left\langle \psi^{(0)} \right| \left( \hat{H}^{(0)} - \epsilon^{(2)} \right) \psi^{(2)} \right\rangle = - \left\langle \psi^{(0)} \right| \left( \hat{H}^{(2)}_{\pm} - \epsilon^{(2)} \right) \psi^{(0)} \right\rangle \quad (6.53)$$

$$\pm \omega_2 \left\langle \psi^{(0)} \right| \psi^{(2)} \right\rangle = - \left\langle \psi^{(0)} \right| \left( \hat{H}^{(2)}_{\pm} \right) \psi^{(0)} \right\rangle + \epsilon^{(2)} \pm \omega_2 \left\langle \psi^{(0)} \right| \psi^{(0)} \right\rangle \quad (6.54)$$

$$\omega_2 \left\langle \psi^{(0)} \right| \psi^{(2)} \right\rangle = \omega_2 \left\langle \psi^{(0)} \right| \psi^{(0)} \right\rangle \quad (6.55)$$

Since there are no first-order perturbations with the appropriate frequency-dependence, there is just this one term in the eigenvalue $\epsilon^{(2)}$. Again, the overlap is not determined by the Sternheimer equation, but is fixed by the normalization condition and phase convention

$$\left\langle \psi^{(0)} \right| \psi^{(2)} \right\rangle - \left\langle \psi^{(0)} \right| \psi^{(0)} \right\rangle = 0. \quad (6.57)$$

However, the situation is more complicated for the sum and difference frequencies. Consider the overlap of the second-order Sternheimer equation at the sum frequency with the unperturbed wavefunction:

$$\left\langle \psi^{(0)} \right| \hat{H}^{(0)} - \epsilon^{(0)} \pm (\omega_\alpha + \omega_\beta) \left\rangle \right\langle \psi^{(2)} \right| \psi^{(2)} \right\rangle =$$

$$- \frac{1}{2} \left( \left\langle \psi^{(0)} \right| \left( \hat{H}^{(1)}_{\pm} - \epsilon^{(1)} \right) \psi^{(1)} \right\rangle + \left\langle \psi^{(0)} \right| \left( \hat{H}^{(1)}_{\pm} - \epsilon^{(1)} \right) \psi^{(1)} \right\rangle \right)$$

$$- \left\langle \psi^{(0)} \right| \left( \hat{H}^{(2)}_{\pm(\alpha+\beta)} \right) \psi^{(0)} \right\rangle = 0 \quad (6.56)$$

Simplifying, using $\left\langle \psi^{(0)} \right| \psi^{(1)} = 0$ and $\epsilon^{(1)} = \epsilon^{(1)}$, 

$$\pm (\omega_\alpha + \omega_\beta) \left\langle \psi^{(0)} \right| \psi^{(2)} \right\rangle =$$

$$- \frac{1}{2} \left[ \left\langle \psi^{(0)} \right| \hat{H}^{(1)}_{\pm(\alpha+\beta)} \psi^{(1)} \right\rangle + \left\langle \psi^{(0)} \right| \hat{H}^{(1)}_{\pm(\alpha+\beta)} \psi^{(1)} \right\rangle \right]$$

$$+ \epsilon^{(2)}_{\pm(\alpha+\beta)} \pm (\omega_\alpha + \omega_\beta) \left\langle \psi^{(0)} \right| \psi^{(2)} \right\rangle = 0 \quad (6.57)$$
The matrix elements on the right-hand side are equal to \( \tilde{\epsilon}^{(2)} \), so we have again a tautology

\[
(\omega_\alpha + \omega_\beta) \left\langle \psi^{(0)} \right| \psi^{(2)}_{\pm(\omega_\alpha + \omega_\beta)} \right\rangle = (\omega_\alpha + \omega_\beta) \left\langle \psi^{(0)} \right| \psi^{(2)}_{\pm(\omega_\alpha + \omega_\beta)} \right\rangle
\]  
\( (6.58) \)

and analogously for the difference frequency,

\[
(\omega_\alpha - \omega_\beta) \left\langle \psi^{(0)} \right| \psi^{(2)}_{\pm(\omega_\alpha - \omega_\beta)} \right\rangle = (\omega_\alpha - \omega_\beta) \left\langle \psi^{(0)} \right| \psi^{(2)}_{\pm(\omega_\alpha - \omega_\beta)} \right\rangle
\]  
\( (6.59) \)

The normalization condition fixes the sum of overlaps, and we set the difference to be zero as a phase convention:

\[
\left\langle \psi^{(0)} \right| \psi^{(2)}_{\omega_\alpha + \omega_\beta} \right\rangle + \left\langle \psi^{(2)}_{-\omega_\alpha - \omega_\beta} \right| \psi^{(0)} \right\rangle = -\frac{1}{2} \left[ \left\langle \psi^{(1)}_{\omega_\alpha} \right| \psi^{(1)}_{\omega_\beta} \right\rangle + \left\langle \psi^{(1)}_{-\omega_\alpha} \right| \psi^{(1)}_{-\omega_\beta} \right\rangle \right]  
\]  
\( (6.60) \)

\[
\left\langle \psi^{(0)} \right| \psi^{(2)}_{\omega_\alpha + \omega_\beta} \right\rangle - \left\langle \psi^{(2)}_{-\omega_\alpha - \omega_\beta} \right| \psi^{(0)} \right\rangle = 0 \]  
\( (6.61) \)

The right-hand side is, in general, non-zero, and indeed calculation of many physical properties involves these overlaps: for example, the polarizability in a periodic system, using electric and \( \mathbf{k} \cdot \mathbf{p} \) perturbations. At first order, these overlaps are zero, but at any higher order, the normalization and phase convention must be used to obtain these components of the wavefunctions. Indeed, in the time-independent case, exactly the same thing must be done to get the undetermined component \( \left\langle \psi^{(0)} \right| \psi^{(2)} \right\rangle \), as done in the standard sum-over-states expression [200].

### 6.4.8 Densities

From the wavefunctions, we can calculate the perturbed densities, needed for the self-consistent parts of the Hamiltonian, as well as for response properties such as the polarizability.

\[
n(r) = \sum_n^{occ} \psi_n(r) \psi_n^*(r) = n^{(0)} + \frac{1}{2} \left( n^{(1)}_{\alpha+} e^{i\omega_\alpha t} + n^{(1)}_{\alpha-} e^{-i\omega_\alpha t} \right) + \frac{1}{2} \left( n^{(1)}_{\beta+} e^{i\omega_\beta t} + n^{(1)}_{\beta-} e^{-i\omega_\beta t} \right) + \frac{1}{2} \alpha \beta \left( n^{(2)}_{\alpha+} e^{i\omega_\alpha t} + n^{(2)}_{\alpha-} e^{-i\omega_\alpha t} + \sum_{\sigma_\alpha,\sigma_\beta} n^{(2)}_{\sigma_\alpha+\sigma_\beta+} e^{i(\sigma_\alpha\omega_\alpha + \sigma_\beta\omega_\beta)t} \right) + \ldots \]  
\( (6.63) \)
Substituting in our expression for the wavefunction, and matching up terms,

\[
n^{(0)} = \sum_{n}^{\text{occ}} \psi^{(0)}_n \left[ \psi^{(0)}_n \right]^{*} \\
\]

\[
n^{(1)} = \sum_{n}^{\text{occ}} \left( \psi^{(1)}_{n+} \left[ \psi^{(0)}_n \right]^{*} + \psi^{(0)}_n \left[ \psi^{(1)}_{n+} \right]^{*} \right) \\
\]

\[
n^{(2)} = \sum_{n}^{\text{occ}} \left( \psi^{(2)}_{n+} \left[ \psi^{(0)}_n \right]^{*} + \psi^{(0)}_n \left[ \psi^{(2)}_{n+} \right]^{*} \right) \\
\]

\[
n^{(2)}_{\sigma \omega \alpha + \sigma \beta \omega \beta} = \sum_{n}^{\text{occ}} \left( P' \psi^{(2)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \left[ \psi^{(0)}_n \right]^{*} + \psi^{(0)}_n \left[ P' \psi^{(2)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \right]^{*} \right) \\
+ \frac{1}{2} \left( \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \left[ \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \right]^{*} + \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \left[ \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \right]^{*} \right) \\
- \frac{1}{2} \psi^{(0)}_n \left[ \psi^{(0)}_n \right]^{*} \left( \langle \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \mid \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \rangle + \langle \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \mid \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \rangle \right) \]  

(6.67)

As discussed in Sec. 6.4.7, while the Sternheimer equation does not determine \( \langle \psi^{(2)} \mid \psi^{(0)} \rangle \), preserving normalization requires

\[
\langle \psi^{(0)} \mid \psi^{(2)} \rangle \left[ \psi^{(0)} \right]^{*} = -\frac{1}{2} \left( \langle \psi^{(1)}_{-} \mid \psi^{(1)}_{+} \rangle + \langle \psi^{(1)}_{+} \mid \psi^{(1)}_{-} \rangle \right) \]  

(6.68)

\[
\langle \psi^{(0)} \mid \psi^{(2)} \rangle \left[ \psi^{(0)} \right]^{*} = -\frac{1}{2} \left( \langle \psi^{(1)}_{-} \mid \psi^{(1)}_{+} \rangle + \langle \psi^{(1)}_{+} \mid \psi^{(1)}_{-} \rangle \right) \]  

(6.69)

Therefore there will be terms \( \psi^{(0)} \langle \psi^{(0)} \mid \psi^{(2)} \rangle \left[ \psi^{(0)} \right]^{*} \) missing from \( n^{(2)} \), since we will have only calculated \( P' n \psi^{(2)} \). We must add them explicitly to our expression:

\[
n^{(2)}_{\sigma \omega \alpha + \sigma \beta \omega \beta} = \sum_{n}^{\text{occ}} \left( P' \psi^{(2)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \left[ \psi^{(0)}_n \right]^{*} + \psi^{(0)}_n \left[ P' \psi^{(2)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \right]^{*} \right) \\
+ \frac{1}{2} \left( \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \left[ \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \right]^{*} + \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \left[ \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \right]^{*} \right) \\
- \frac{1}{2} \psi^{(0)}_n \left[ \psi^{(0)}_n \right]^{*} \left( \langle \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \mid \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \rangle + \langle \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \mid \psi^{(1)}_{n, \sigma \omega \alpha + \sigma \beta \omega \beta} \rangle \right) \]  

(6.70)

Note that while the total time-dependent density \( n \) must be real, the response densities \( n^{(i)} \) are complex in general, and are required to be real only in the time-dependent case, or when the wavefunctions are real and the frequency is real. The requirements for \( n \) to be real then are:

\[
n^{(1)}_{+} = \left[ n^{(1)}_{-} \right]^{*}, \quad n^{(2)}_{+} = \left[ n^{(2)}_{-} \right]^{*}, \quad n^{(2)}_{\sigma \omega \alpha + \sigma \beta \omega \beta} = \left[ n^{(2)}_{\sigma \omega \alpha + \sigma \beta \omega \beta} \right]^{*} \]  

(6.71)

It can be seen from inspection of the expressions above that these conditions are met. The perturbation Hamiltonians are not Hermitian if the response densities are complex, but the total Hamiltonian is Hermitian provided

\[
\hat{H}^{(1)}_{+} = \hat{H}^{(1)}_{+}, \quad \hat{H}^{(2)}_{+} = \hat{H}^{(2)}_{+}, \quad \hat{H}^{(2)}_{\sigma \omega \alpha + \sigma \beta \omega \beta} = \hat{H}^{(2)}_{\sigma \omega \alpha + \sigma \beta \omega \beta} \]  

(6.72)
The condition on the densities allows this condition to be met for the Hartree and exchange-correlation terms. We additionally need the analogous conditions on the bare perturbations:

\[ V_+^{(1)} = V_-^{(1)\dagger}, \quad V_+^{(2)} = V_-^{(2)\dagger}, \quad V_{\sigma \omega \omega}^{(2)} = V_{\sigma \omega \omega}^{(2)\dagger} \quad (6.73) \]

Two ways of satisfying these requirements are to have the bare perturbation be Hermitian and invariant with the sign of the frequency \((V_+ = V_- = V = V^\dagger)\), which corresponds to a cosinusoidal time-dependence, or to have the bare perturbation be anti-Hermitian and change sign with the sign of the frequency \((V_+ = -V_- = -V^\dagger)\), which corresponds to a sinusoidal time-dependence. However, any operator, even if neither Hermitian nor anti-Hermitian, can be used for one sign of the frequency, which fixes the operator for the other sign, and then will introduce other phases between the positive- and negative-frequency parts in the harmonic time-dependence.

To evaluate \(n^{(1)}_\pm\) and \(n^{(2)}_\pm\), only the projections of the perturbed wavefunctions in the unoccupied subspace are required. For the first-order density,

\[
n^{(1)}_\pm (\vec{r}) = \sum_{n} \left[ \langle \vec{r} | \psi_{n}^{(0)} \rangle \langle \psi_{n}^{(1)} | \vec{r} \rangle + \langle \vec{r} | \psi_{n}^{(1)} \rangle \langle \psi_{n}^{(0)} | \vec{r} \rangle \right] \\
= \sum_{n} \left[ \langle \vec{r} | \psi_{n}^{(0)} \rangle \langle \psi_{n}^{(1)} | \psi_{m}^{(0)} \rangle \langle \psi_{m}^{(0)} | \vec{r} \rangle + \langle \vec{r} | \psi_{m}^{(0)} \rangle \langle \psi_{m}^{(1)} | \psi_{n}^{(0)} \rangle \langle \psi_{n}^{(0)} | \vec{r} \rangle \right] \\
= \sum_{n} \sum_{m} \left[ \langle \vec{r} | \psi_{n}^{(0)} \rangle \langle \psi_{m}^{(0)} | \psi_{n}^{(1)} \rangle \langle \psi_{m}^{(0)} | \vec{r} \rangle \right. \\
\quad \quad + \left. \langle \vec{r} | \psi_{m}^{(0)} \rangle \langle \psi_{m}^{(0)} | \psi_{n}^{(1)} \rangle \langle \psi_{n}^{(0)} | \vec{r} \rangle \right] \\
= \sum_{n} \sum_{m} \left[ \langle \vec{r} | \psi_{m}^{(0)} \rangle \frac{\langle \psi_{n}^{(0)} | \hat{H}_{\pm}^{(1)} | \psi_{m}^{(0)} \rangle}{\epsilon_{n}^{(0)} - \epsilon_{m}^{(0)} \pm \omega - i\eta} \langle \psi_{m}^{(0)} | \vec{r} \rangle \right. \\
\quad \quad + \left. \langle \vec{r} | \psi_{m}^{(0)} \rangle \frac{\langle \psi_{m}^{(0)} | \hat{H}_{\pm}^{(1)} | \psi_{n}^{(0)} \rangle}{\epsilon_{n}^{(0)} - \epsilon_{m}^{(0)} \mp \omega + i\eta} \langle \psi_{n}^{(0)} | \vec{r} \rangle \right] \\
= \sum_{n} \sum_{m} \left[ \langle \vec{r} | \psi_{m}^{(0)} \rangle \frac{\langle \psi_{n}^{(0)} | \hat{H}_{\pm}^{(1)} | \psi_{m}^{(0)} \rangle}{\epsilon_{n}^{(0)} - \epsilon_{m}^{(0)} \pm \omega - i\eta} \langle \psi_{m}^{(0)} | \vec{r} \rangle \right. \\
\quad \quad + \left. \langle \vec{r} | \psi_{m}^{(0)} \rangle \frac{\langle \psi_{m}^{(0)} | \hat{H}_{\pm}^{(1)} | \psi_{n}^{(0)} \rangle}{\epsilon_{n}^{(0)} - \epsilon_{m}^{(0)} \mp \omega + i\eta} \langle \psi_{n}^{(0)} | \vec{r} \rangle \right] \quad (6.74) \]

But the \((n, m, \omega)\) term is cancelled out by the \((m, n, -\omega)\) term, whenever both \(n\) and \(m\)
occur in the sum, i.e. when \( m \) is an occupied state. The term with \( n \) and \( m \) swapped is

\[
\langle r | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | \hat{H}_\pm^{(1)} | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | r \rangle \epsilon_m^{(0)} - \epsilon_n^{(0)} \pm \omega + i\eta \nonumber \\
\langle r | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | \hat{H}_\pm^{(1)} | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | r \rangle \epsilon_n^{(0)} - \epsilon_m^{(0)} \pm \omega + i\eta 
\]

which cancels exactly with the terms in the expression above. We are left with

\[
\sum_{n} \sum_{m} \text{occ} \text{ unocc} \left[ \langle r | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | \hat{H}_\pm^{(1)} | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | r \rangle \epsilon_m^{(0)} - \epsilon_n^{(0)} \pm \omega + i\eta \nonumber \\
+ \langle r | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | \hat{H}_\pm^{(1)} | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | r \rangle \epsilon_n^{(0)} - \epsilon_m^{(0)} \pm \omega + i\eta \right] 
\]

(6.75)

Therefore our final expression for the linear density response is

\[
n_{\pm}^{(1)}(r) = \sum_{n} \text{occ} \left( \langle r | \psi_n^{(0)} \rangle \langle \psi_n^{(1)} | P_c | r \rangle + \langle r | P_c | \psi_n^{(1)} \rangle \langle \psi_n^{(0)} | r \rangle \right) 
\]

(6.77)

The result for \( n_{\pm}^{(2)} \) is exactly analogous.

6.4.9 Degenerate frequencies

Thus far we have been assuming that all the various frequencies in the equations are different, so that we get separate equations for each frequency. However, if some frequencies are the same, or one or more frequencies are zero (so that the distinction between the positive and negative frequency disappears), the corresponding equations must be combined.

The first-order Sternheimer equation becomes only one equation, rather than a pair, if the frequency is zero:

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \right) P_c \psi^{(1)} = -P_c \hat{H}^{(1)} \psi^{(0)} 
\]

(6.78)

The same happens with the second-order equation at the second-order frequency:

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \right) P_c \psi^{(2)} = -P_c \hat{H}^{(2)} \psi^{(0)} 
\]

(6.79)

If \( \omega_\beta = 0 \), the sum- and difference-frequency equations collapse into one:

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \pm \omega_\alpha \right) P_{n,\alpha}^{p} \psi^{(2)} = -\frac{1}{2} P_{n,\alpha}^{p} \left[ \left( \hat{H}_{\alpha}^{(1)} - \epsilon^{(1)} \right) \psi^{(1)} + \left( \hat{H}_{\beta}^{(1)} - \epsilon^{(1)} \right) \psi^{(1)} - P_{n,\alpha}^{p} \hat{H}_{\pm \omega_\alpha} \psi^{(0)} \right] 
\]

(6.80)
and analogously if \( \omega_n = 0 \).

If \( \omega_n = \omega_\beta = \omega \), the equations are:

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \pm 2\omega \right) P_n^{(2)} \psi_n^{(2)} =
- \frac{1}{2} P_n \left[ \left( \hat{H}_{\alpha \pm}^{(1)} - \epsilon_{\alpha \pm}^{(1)} \right) \psi_{\beta \pm}^{(1)} + \left( \hat{H}_{\beta \pm}^{(1)} - \epsilon_{\beta \pm}^{(1)} \right) \psi_{\alpha \pm}^{(1)} \right] - P_n^{(2)} \hat{H}_{\pm 2\omega}^{(2)} \psi^{(0)}
\]

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \right) P_n^{(2)} \psi_n^{(2)} =
- \frac{1}{2} P_n \left[ \left( \hat{H}_{\alpha}^{(1)} - \epsilon_{\alpha}^{(1)} \right) \psi_{\beta}^{(1)} + \left( \hat{H}_{\beta}^{(1)} - \epsilon_{\beta}^{(1)} \right) \psi_{\alpha}^{(1)} \right] - P_n^{(2)} \hat{H}_{0}^{(2)} \psi^{(0)}
\]

More simplification occurs if then \( \omega = 0 \), in which case the sum- and difference-frequency equations reduce to just one:

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \right) P_n^{(2)} \psi_n^{(2)} =
- \frac{1}{2} P_n \left[ \left( \hat{H}_{\alpha}^{(1)} - \epsilon_{\alpha}^{(1)} \right) \psi_{\beta}^{(1)} + \left( \hat{H}_{\beta}^{(1)} - \epsilon_{\beta}^{(1)} \right) \psi_{\alpha}^{(1)} \right] - P_n^{(2)} \hat{H}_{0}^{(2)} \psi^{(0)}
\]

When \( \alpha = \beta \) and \( V_\alpha = V_\beta \), we have simply

\[
\left( \hat{H}^{(0)} - \epsilon^{(0)} \right) P_n^{(2)} \psi_n^{(2)} = -P_n^{(2)} \left( \hat{H}^{(1)} - \epsilon^{(1)} \right) \psi^{(1)} - P_n^{(2)} \hat{H}_{0}^{(2)} \psi^{(0)}
\]

6.4.10 Imaginary frequencies

We can add imaginary parts \( i\eta \) to all frequencies to deal with resonances, as in the first-order Sternheimer equation. This effectively adds a damping at the resonance which prevents the response from diverging by broadening resonances from a delta function to a Lorentzian lineshape. This broadening can be physical, but a second reason to use it is that a positive infinitesimal is required to get the poles of the response function correctly, and to describe causal response functions: therefore, the sign of the imaginary part must remain the same for different signs of the frequency. Everywhere \(-\omega \) occurs, it should be replaced with \(-\omega^* \) when an imaginary part is used, so that the frequencies are \( \pm \omega + i\eta \). To be consistent, the sum- and difference-frequency equations should not use unrelated values of \( \eta \), but rather take those from the first-order equation: \( \text{i.e. } \sigma_\alpha \omega_\alpha + i\eta_\alpha + \sigma_\beta \omega_\beta + i\eta_\beta \). In this way, all the results we have shown are still valid when the frequency becomes complex, even without taking the limit \( \eta \rightarrow 0 \).

No \( \eta \) terms should be used with time-independent perturbations, because their response cannot diverge in this formalism. The only way the response could diverge would be if the perturbation splits a degeneracy: in that case, degenerate perturbation theory should be used instead, in which case the response will then be finite. In practice, the linearity of the Sternheimer equation means that when applied to degenerate levels the result is that the response of a linear combination of states is the linear combination of those states’ responses.

\[
\phi_i^{(0)} = \sum_j T_{ij} \psi_j^{(0)}, \phi_i^{(1)} = \sum_j T_{ij} \psi_j^{(1)}
\]
When projecting the equation for state $i$ onto the state $j$ with the same energy, the left-hand side becomes zero (just as when projecting on the state $i$ itself), and so given an initialization to zero, these components will be zero. This response may not be correct for the individual state (an issue for effective masses via $k \cdot p$ perturbation theory), but final results based on energies or densities are correct and unaffected by unitary transformations.

6.4.11 Applications

To calculate higher-order nonlinear optical properties, we construct hyperpolarizabilities with the $2n + 1$ theorem, choosing the appropriate frequency-dependences of $\psi^{(2)}$. For example, to calculate two-photon absorption, $\text{Im} \gamma(\omega_1; \omega_2, -\omega_1, -\omega_2)$, we would use terms like $\langle \psi^{(1)} \hat{H}^{(1)} \psi^{(2)}_{-\omega_1-\omega_2} \rangle$. We can also ignore the $2n + 1$ theorem, and calculate hyperpolarizabilities from $\psi^{(2)}$, in which case to get the susceptibility for sum-frequency generation, $\beta(\omega_1 + \omega_2; -\omega_1, -\omega_2)$, we use the sum-frequency equation to calculate terms like $\langle \psi^{(0)} \hat{H}^{(1)} \psi^{(2)}_{\omega_1} \rangle$ (treating $\omega_1$ as $(\omega_1 + \omega_2) - \omega_2$).

To calculate the hyperpolarizability for a periodic system with the quantum theory of polarization, we set $\alpha = \mathcal{E}$, $V_{\alpha} = -i \nabla_k$, $\beta = k$, $V_\beta = p + k_0$, $V_2 = 0$. The $k \cdot p$ perturbation is non-self-consistent. The eigenvalues $\epsilon_k^{(1)}$ are in fact the group velocities, $v_g$, a physically meaningful quantity we can easily check against the band structure. We might think we could check a calculation of the polarizability from the results of the second-order equation, to compare the results of the first-order equation:

$$\langle \psi^{(1)}_k | i \nabla_k | \psi^{(0)} \rangle = i \langle \psi^{(1)}_k | \psi^{(1)}_k \rangle = -i \langle \psi^{(2)}_k | \psi^{(0)} \rangle$$  \hspace{1cm} (6.86)

The minus sign on $i$ in the second form is due to the fact that $\nabla_k$ is an anti-Hermitian operator. However, in fact the term on the right is undetermined by our Sternheimer equation, as shown in Sec. 6.4.7. Therefore the only way we can evaluate this term is to use the overlap of the two first-order wavefunctions, making the attempt to check a tautology, so this will not be useful unfortunately. Instead we will just have to benchmark the hyperpolarizability for a molecule calculated as a finite system versus that calculated in a large supercell as a periodic system.

The equation we will actually have to solve is this:

$$(\hat{H}^{(0)} - \epsilon^{(0)}_{\pm} \pm \omega) P'_n \psi^{(2)}_{\omega_{\pm}} =$$

$$-\frac{1}{2} P'_n \left[ (\hat{H}^{(1)} - \epsilon^{(1)}_{\pm}) \psi^{(1)}_k + (\hat{H}^{(1)}_{k\pm} - \epsilon^{(1)}_k) \psi^{(1)}_{k\pm} \right] - P'_n \hat{H}^{(2)} \psi^{(0)}$$  \hspace{1cm} (6.87)

To obtain the input $\hat{H}^{(1)} \psi^{(1)}_k$, we must solve a second-order Sternheimer equation first in fact, because this is

$$\hat{H}^{(1)} \psi^{(1)}_k = -i \frac{\partial}{\partial k} \hat{\psi} = -i \frac{\partial^2 \psi}{\partial k^2}$$  \hspace{1cm} (6.88)

The perturbation [51] is

$$\frac{\partial^2 \hat{H}}{\partial k^2} = [r, [r, V_{\text{ps}}]]$$  \hspace{1cm} (6.89)
The second-order Sternheimer equation requires derivatives of eigenvalues. \( \epsilon_k^{(1)} = 0 \) at the \( \Gamma \)-point in fact, so that is simple to handle. However, \( \epsilon_E \) is actually uncalculable with the quantum theory of polarization [378] (only off-diagonal matrix elements of \(-i \nabla_k\) may be calculated, so we must take another tack. We can approximate it via a form akin to the single-point Berry phase [407]:

\[
\langle \psi | -i \nabla_k | \psi \rangle \approx \frac{L}{2\pi} \text{Im} \left\langle \psi \left| e^{i2\pi r/L} \right| \psi \right\rangle
\]

(6.90)

where \( L \) is the lattice constant. This approach is correct to order \( 1/L^2 \) and should be good for a large supercell, as for the envisioned liquid calculation. The same term occurs in the formula for \( \chi^{(2)} \), in fact, and so we will use the same approximation there. I have implemented these equations in Octopus and they will be applied to the study of liquid chloroform.
Chapter 7

BerkeleyGW: A massively parallel computer package for the
calculation of the quasiparticle and
optical properties of materials and
nanostructures

This work is based on Ref. [88]. Since the publication of that paper, I have made some
 corrections and provided additional information, in the arXiv version which is intended to
serve as documentation for the code [89]. Those updates are reflected in this dissertation.
Due to the importance of clarity on the equations implemented in the package, I have
included a few errata to the key papers on the methodology of BerkeleyGW in Appendix
A.

7.1 Abstract

BerkeleyGW is a massively parallel computational package for electron excited-state
properties that is based on many-body perturbation theory employing the \textit{ab initio} GW and
GW plus Bethe-Salpeter equation methodology. It can be used in conjunction with many
density-functional theory codes for ground-state properties, including\textsc{paratec}, \textsc{parsec},
\textsc{Quantum ESPRESSO}, \textsc{SIESTA}, and \textsc{Octopus}. The package can be used to compute the elec-
tronic and optical properties of a wide variety of material systems from bulk semiconductors
and metals to nanostructured materials and molecules. The package scales to 10000s of
CPUs and can be used to study systems containing up to 100s of atoms.

7.2 Program Summary

Program title: BerkeleyGW
Program obtainable from: http://www.berkeleygw.org
Licensing provisions: modified BSD license.
No. of lines in distributed program, including test data, etc.: 80,000
No. of bytes in distributed program, including test data, output, etc.: 200MB
Distribution format: tar
Programming language: Fortran 90, C, C++, Python, Perl, BASH
Libraries required: BLAS, LAPACK, FFTW, ScaLAPACK (optional), MPI (optional). All available under open-source licenses.
Memory required: (50-2000) MB per CPU (Highly dependent on system size)
Computers for which the program has been designed and others on which it has been operable: Linux/UNIX workstations or clusters
Operating systems under which the program has been tested: Tested on a variety of Linux distributions in parallel and serial as well as AIX and Mac OSX.
Nature of problem: The excited state properties of materials involve the addition or subtraction of electrons as well the optical excitations of electron-hole pairs. The excited particles interact strongly with other electrons in a material system. This interaction affects the electronic energies, wavefunctions and lifetimes. It is well known that ground-state theories, such as standard methods based on density-functional theory, fail to correctly capture this physics.
Solution method: We construct and solve Dyson’s equation for the quasiparticle energies and wavefunctions within the GW approximation for the electron self energy. We additionally construct and solve the Bethe-Salpeter equation for the correlated electron-hole (exciton) wavefunctions and excitation energies.
Restrictions: The material size is limited in practice by the computational resources available. Materials with up to 500 atoms per periodic cell can be studied on large HPCs.
Running time: 1-1000 minutes (depending greatly on system size and processor number)

7.3 Introduction

Over the last few decades, the \textit{ab initio} GW methodology has been successfully applied to the study of the quasiparticle properties of a large range of material systems from traditional bulk semiconductors, insulators and metals to, more recently, nano-systems like polymers, nano-wires and molecules [152, 211, 334, 335, 90]. The GW approach, which is based on approximating the electron self energy as the first term in an expansion in
the screened Coulomb interaction, $W$ \cite{140}, has proven to yield quantitatively accurate quasiparticle band gaps and dispersion relations from first principles.

Additionally, the Bethe-Salpeter equation (BSE) approach to the optical properties of materials has proven exceptionally accurate in predicting the optical response of a similarly large class of materials employing an electron-hole interaction kernel derived within the same level of approximations as GW \cite{341, 303, 1, 26}.

The combined GW-BSE approach is now arguably regarded as the most accurate methodology commonly used for computing the quasiparticle and optical properties of condensed-matter systems. A perceived drawback of the GW methodology is its computational cost; a GW-BSE calculation is usually thought to be an order of magnitude (or worse) more costly than a typical density functional theory (DFT) calculation for the same system. Since the pioneering work of Ref. \cite{152}, many GW implementations have been made, but most are limited to small systems of the size of 10s of atoms, and scaling to only small numbers of CPUs on the order of 100.

BerkeleyGW is a massively parallel computer package which is written predominantly in FORTRAN90 that implements the ab initio GW methodology of Hybertsen and Louie \cite{152} and includes many more recent advances, such as the Bethe-Salpeter equation approach for optical properties \cite{303}. It alleviates the restriction to small numbers of atoms and scales beyond thousands of CPUs. The package is intended to be used on top of a number of mean-field (DFT and other) codes that focus on ground-state properties, such as PARATEC \cite{262}, Quantum ESPRESSO \cite{119}, SIESTA \cite{328}, PARSEC \cite{60, 3}, Octopus \cite{55, 218, 7} and an empirical pseudopotential code (EPM) included in the package (based on TBPW \cite{223}). More information about BerkeleyGW, the latest source code, and help forums can be found by visiting the website at http://berkeleygw.org/.

### 7.4 Theoretical Framework

The ab initio GW-BSE approach is a many-body Green’s-function methodology in which the only input parameters are the constituent atoms and the approximate structure of the system \cite{152, 303}. Typical calculations of the ground- and excited-state properties using the GW-BSE method can be broken into three steps: (1) the solution of the ground-state structural and electronic properties within a suitable ground-state theory such as ab initio pseudopotential density-functional theory, (2) the calculation of the quasiparticle energies and wavefunctions within the GW approximation for the electron self-energy operator, and (3) the calculation of the two-particle correlated electron-hole excited states through the solution of a Bethe-Salpeter equation.

DFT calculations, often the chosen starting point for GW, are performed by solving the self-consistent Kohn-Sham equations with an approximate functional for the exchange-correlation potential, $V_{xc}$ – common approximations being the local density approximation (LDA) \cite{191} and the generalized-gradient approximation (GGA) \cite{269}:

$$\left[ -\frac{1}{2}\nabla^2 + V_{\text{ion}} + V_H + V_{\text{xc}}^{\text{DFT}} \right] \psi_{nk}^{\text{DFT}} = E_{nk}^{\text{DFT}} \psi_{nk}^{\text{DFT}}$$

(7.1)

where $E_{nk}^{\text{DFT}}$ and $\psi_{nk}^{\text{DFT}}$ are the Kohn-Sham eigenvalues and eigenfunctions respectively,
$V_{\text{ion}}$ is the ionic potential, $V_H$ is the Hartree potential and $V_{xc}$ is the exchange-correlation potential within a suitable approximation. When DFT is chosen as the starting point for GW, the Kohn-Sham wavefunctions and eigenvalues are used here as a first guess for their quasiparticle counterparts. The quasiparticle energies and wavefunctions (i.e., the one-particle excitations) are computed by solving the following Dyson equation [141, 152] in atomic units:

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{ion}}^\text{QP} + V_H + \Sigma(E_{nk}^\text{QP})\right] \psi_{nk}^\text{QP} = E_{nk}^\text{QP} \psi_{nk}^\text{QP}$$

(7.2)

where $\Sigma$ is the self-energy operator within the GW approximation, and $E_{nk}^\text{QP}$ and $\psi_{nk}^\text{QP}$ are the quasiparticle energies and wavefunctions, respectively. For systems of periodic dimension less than three, the Coulomb interaction may be replaced by a truncated interaction. The interaction is set to zero for particle separation beyond the size of the system in order to avoid unphysical interaction between the material and its periodic images in a super-cell [67] calculation. The electron-hole excitation states (probed in optical or other measurements) are calculated through the solution of a Bethe-Salpeter equation [303, 341] for each exciton state $S$:

$$(E_{ck} - E_{ck}^\text{QP}) A_{vck}^S + \sum_{v'c'k'} \langle vck|K^{\text{eh}}|v'c'k'\rangle = \Omega^S A_{vck}^S$$

(7.3)

where $A_{vck}^S$ is the exciton wavefunction (in the quasiparticle state representation), $\Omega^S$ is the excitation energy, and $K^{\text{eh}}$ is the electron-hole interaction kernel. We make the Tamm-Dancoff approximation by including only valence $\rightarrow$ conduction transitions [303, 107]. The exciton wavefunction can be expressed in real space as:

$$\Psi(r_e, r_h) = \sum_{k,c,v} A_{vck}^S \psi_{k,c}(r_e) \psi_{k,v}^*(r_h),$$

(7.4)

and the imaginary part of the dielectric function, if one is interested in optical properties, can be expressed as

$$\epsilon_2(\omega) = \frac{16\pi^2e^2}{\omega^2} \sum_S |\mathbf{e} \cdot \langle 0|\mathbf{v}|S\rangle|^2 \delta(\omega - \Omega^S)$$

(7.5)

where $\mathbf{e} \cdot \langle 0|\mathbf{v}|S\rangle$ is the velocity matrix element along the direction of the polarization of light, $\mathbf{e}$. One may compare this to the non-interacting absorption spectrum:

$$\epsilon_2(\omega) = \frac{16\pi^2e^2}{\omega^2} \sum_{vck} |\mathbf{e} \cdot \langle v\mathbf{k}|\mathbf{v}|c\mathbf{k}\rangle|^2 \delta(\omega - E_{ck}^\text{QP} + E_{vk}^\text{QP}).$$

(7.6)

An example absorption spectrum for silicon computed with the BerkeleyGW package at the GW and GW-BSE levels is shown in Fig. 7.2. Only when both the quasiparticle effects within the GW approximation and the excitonic effects through the solution of the Bethe-Salpeter equation are included is good agreement with experiment reached.
Figure 7.2: The absorption spectra for silicon calculated at the GW (black dashed) and GW-BSE (red solid) levels using the BerkeleyGW package. Experimental data from [168].

Figure 7.3: Flow chart of a GW-BSE calculation performed in the BerkeleyGW package.
7.5  Computational Layout

7.5.1  Major Sections of the Code

Figure 7.3 illustrates the procedure for carrying out an *ab initio* GW-BSE calculation to obtain quasiparticle and optical properties using the BerkeleyGW code. First, one obtains the mean-field electronic orbitals and eigenvalues as well as the charge density. One can utilize one of the many supported DFT codes [262, 119, 328, 3, 218] to construct this mean-field starting point and convert it to the plane-wave BerkeleyGW format (see Sec. 7.10) using the wrappers included. (Note that norm-conserving pseudopotentials must be used, or else extra contributions would need to be added to our matrix elements.)

The Epsilon executable produces the polarizability and inverse dielectric matrices. In the epsilon executable, the static or frequency-dependent polarizability and dielectric function are calculated within the random-phase approximation (RPA) using the electronic eigenvalues and eigenfunctions from a mean-field reference system. The main outputs are files eps0mat and epsmat that contain the inverse-dielectric matrix for $\mathbf{q} \rightarrow 0$ and $\mathbf{q} \neq 0$.

In the sigma executable, the screened Coulomb interaction, $W$, is constructed from the inverse dielectric matrix and the one-particle Green’s function, $G$, is constructed from the mean-field eigenvalues and eigenfunctions. We then calculate the diagonal and (optionally) off-diagonal elements of the self-energy operator, $\Sigma = iGW$, as a matrix in the mean-field basis. In many cases, only the diagonal elements are sizable within the chosen mean-field orbital basis; in such cases, in applications to real materials, the effects of $\Sigma$ can be treated within first-order perturbation theory. The sigma executable evaluates $\Sigma$ in the form $\Sigma = V_{xc} + (\Sigma - V_{xc})$, where $V_{xc}$ is the independent-particle mean-field approximation to the exchange-correlation potential of the chosen mean-field system. For moderately correlated electron systems, the best available mean-field Hamiltonian may often be taken to be the Kohn-Sham Hamiltonian [191]. However, many mean-field starting points are consistent with the BerkeleyGW package, such as Hartree-Fock, static COHSEX and hybrid functionals. In principle, the process of correcting the eigenfunctions and eigenvalues (which determine $W$ and $G$) could be repeated until self-consistency is reached or the $\Sigma$ matrix diagonalized in full. However, in practice, it is found that an adequate solution often is obtained within first-order perturbation theory on Dyson’s equation for a given $\Sigma$ [147, 14]. Comparison of calculated energies with experiment shows that this level of approximation is very accurate for semiconductors and insulators and for most conventional metals. The outputs of the sigma executable are $E^{\text{QP}}$, the quasiparticle energies, which are written to the file eqp.dat using the eqp.py post-processing utility on the generated sigma.log files for each sigma run.

The BSE executable, kernel, takes as input the full dielectric matrix calculated in the epsilon executable, which is used to screen the attractive direct electron-hole interaction, and the quasiparticle wavefunctions, which often are taken to be the same as the mean-field wavefunctions. The direct and exchange part of the electron-hole kernel are calculated and output into the bse0mat and bsexmat files respectively. The absorption executable uses these matrices, the quasiparticle energies and wavefunctions from a coarse $k$-point grid GW calculation, as well as the wavefunctions from a fine $k$-point grid. The quasiparticle energy corrections and the kernel matrix elements are interpolated onto the fine grid. The
Bethe-Salpeter Hamiltonian, consisting of the electron-hole kernel with the addition of the kinetic-energy term, is constructed in the quasiparticle electron-hole pair basis and diagonalized yielding the electron-hole amplitude, or exciton wavefunctions, and excitation energies, printed in the file eigenvectors. Exciton binding energies can be inferred from the energy of the correlated exciton states relative to the inter-band-transition continuum edge. With the excitation energies and amplitudes of the electron-hole pairs, one then can calculate the macroscopic dielectric function for various light polarizations which is written to the file absorption_eh.dat. This may be compared to the absorption spectrum without the electron-hole interaction included, printed in the file absorption_noeh.dat.

Example input files for each executable are contained within the source code for the package, as well as complete example calculations for silicon, the (8,0) and (5,5) single-walled carbon nanotubes (SWCNTs), the CO molecule, and sodium metal. There are several post-processing and visualization utilities included in the package that are described in Sec. 7.8.

Additionally, sums over $k$ and $q$ are accompanied by an implicit division by the volume of the super-cell considered, $V_{sc} = N_k V_{uc}$, where $N_k$ is the number of points in the $k$-grid and $V_{uc}$ is the volume of the unit cell in a periodic system.

### 7.5.2 RPA Dielectric Matrix: epsilon

epsilon is a standalone executable that computes either the static or dynamic RPA polarizability and corresponding inverse dielectric function from input electronic eigenvalues and eigenvectors computed in a suitable mean-field code. As we discuss in detail below, the input electronic eigenvalues and eigenvectors can come from a variety of different mean-field approximations including DFT within LDA/GGA, generalized Kohn-Sham hybrid-functional approximations as well as direct approximations to the GW Dyson’s equation such as the static-COHSEX [141, 165] approximation and the Hartree-Fock approximation.

We will first discuss the computation of the static polarizability and the inverse dielectric matrix. The epsilon executable computes the static RPA polarizability using the following expression [152]:

$$
\chi_{GG'}(q; 0) = \sum_{n} \sum_{n'} \sum_{k} M_{nn'}^*(k, q, G) M_{nn'}(k, q, G') \frac{1}{E_{nk+q} - E_{n'k}},
$$

(7.7)

where

$$
M_{nn'}(k, q, G) = \langle n_k+q| e^{i(q+G) \cdot r} | n'_k \rangle,
$$

(7.8)

are the plane-wave matrix elements. Here $q$ is a vector in the first Brillouin zone, $G$ is a reciprocal-lattice vector, and $\langle n_k |$ and $E_{nk}$ are the mean-field electronic eigenvectors and eigenvalues. The matrix in Eq. 7.7 is to be evaluated up to $|q+G|^2, |q+G'|^2 < E_{cut}$ where $E_{cut}$ defines the dielectric energy cutoff.

With the expression for $\chi$ above, we can obtain the RPA dielectric matrix as

$$
e_{GG'}(q; 0) = \delta_{GG'} - v(q+G)\chi_{GG'}(q; 0)
$$

(7.9)

where $v(q+G)$ is the bare Coulomb interaction defined as:

$$
v(q+G) = \frac{4\pi}{|q+G|^2}
$$

(7.10)
in the case of bulk crystals where no truncation is necessary. For nano-systems truncating the interaction in non-periodic directions is implemented and greatly improves the convergence with super-cell size.

It should be noted that we use an asymmetric definition of the Coulomb interaction, as opposed to symmetric expressions such as

$$v(q + G, q + G') = \frac{4\pi}{|q + G||q + G'|}. \quad (7.11)$$

This causes $\epsilon_{GG'}(q:0)$ and $\chi_{GG'}(q:0)$ to be also asymmetric in $G$ and $G'$. This asymmetry is resolved when constructing the static screened Coulomb interaction by use of the expression:

$$W_{GG'}(q:0) = \epsilon^{-1}_{GG'}(q:0)v(q + G'). \quad (7.12)$$

Here $W$ is symmetric in $G$ and $G'$ even though both $v$ and $\epsilon^{-1}$ individually are not.

Finally, once we have constructed $\chi_{GG'}(q:0)$ we can construct the RPA dielectric matrix and inverse dielectric matrix required for the computation of the screened Coulomb interaction, $W$. The dielectric matrix as implemented in the code is expressed in Eq. 7.9.

There is a clear problem in directly computing $\epsilon_{00}(q = 0)$ due to the fact that the Coulomb interaction, Eq. 7.10, diverges as $q \to 0$. For semiconducting systems, due to orthogonality, the matrix elements (Eq. 7.8) themselves go to 0 with the form

$$|M_{nn'}(k, q, G = 0)| \propto |q|^2, \quad (7.13)$$

so $\epsilon(q \to 0)$ contains a non-trivial $q^2/|q|^2$ limit. One way to handle this would be to take the limit of Eqs. 7.7 and 7.8 analytically via $k \cdot p$ perturbation theory, where the perturbation is the momentum operator $-i\nabla$ plus the commutators with the non-local potential of the mean-field Hamiltonian [18, 152]. This is analogous to the treatment of the velocity operator in absorption (Eq. 7.43).

The epsilon code has implemented a simpler scheme, however, in which we numerically take the limit as $q \to 0$ by evaluating $\epsilon_{00}(q_0)$ at a small but finite $q_0$ usually taken as approximately 1/1000th of the Brillouin zone, in one of the periodic directions. For semiconducting systems, where $\epsilon_{00}(q = 0) \to C$, it is sufficient to construct a separate $k$-grid for the conduction and valence bands shifted by the small vector $q_0$ in order to compute $M_{nn'}(k, q_0, G = 0)$, where $n$ is a valence and $n'$ a conduction band, and to evaluate the correct limiting $q^2/|q|^2$ ratio. A slightly different treatment is used for metallic systems to include intra-band transitions [88]. This is still the zero-temperature limit in our code, as the effect of thermal occupations is small in GW except at very large temperatures [27]. Effectively occupations are taken as one below the Fermi level, zero above the Fermi level, and 1/2 at the Fermi level (as needed for graphene at the Dirac point). This is despite any smearing that may have been used in the underlying mean-field calculation. We should point out that only one $q_0$ is used; if the material is anisotropic (in periodic directions), in principle an average over the three directions of $q_0$ should be done. This may be accomplished by using a vector in the (111) direction (referred to the principal axes of $\epsilon$). Neglect of the anisotropy can give significant errors in sigma [319].

We have so far limited ourselves to situations in which only a direct calculation of the static polarizability, Eq. 7.7, is required, such as in the static-COHSEX approximation
or when utilizing a GPP model to extend the dielectric response to non-zero frequencies. However, we can also do a more refined calculation. Options are given in the code so that the dielectric matrix is computed directly at real frequencies without extrapolation, as is formally required in the Dyson equation.

We use in the package the advanced and retarded dielectric functions, defined as:

\[
\epsilon^{\text{r/a}}_{GG'}(q; E) = \delta_{GG'} - v(q+G) \\
\times \sum_{n} \sum_{n'} \sum_{k} M_{nn'}^{*}(k, q, G) M_{nn'}(k, q, G') \\
\times \frac{1}{2} \left[ \frac{1}{E_{nk+q} - E_{n'k} - E \mp i\delta} + \frac{1}{E_{nk+q} - E_{n'k} + E \pm i\delta} \right]
\]

where \( E \) is the evaluation frequency and \( \delta \) is a broadening parameter chosen to be consistent with the energy spacing afforded by the \( k \)-point sampling of the calculation, using the upper (lower) signs for the retarded (advanced) function. In principle, one must converge the calculation with respect to increasing the \( k \)-point sampling and decreasing this broadening parameter.

### 7.5.3 Computation of the Self-Energy: sigma

The sigma executable takes as input the inverse epsilon matrix calculated from the epsilon executable and a suitable set of mean-field electronic energies and wavefunctions. It computes a representation of the Dyson’s equation, Eq. 7.2, in the basis of the mean-field eigenfunctions through the computation of the diagonal and off-diagonal elements of \( \Sigma \):

\[
\langle \psi_{nk} | H^{QP}(E) | \psi_{mk} \rangle = E^{\text{MF}}_{nk} \delta_{n,m} + \langle \psi_{nk} | \Sigma(E) - \Sigma^{\text{MF}}(E) | \psi_{mk} \rangle
\]

where \( E \) is an energy parameter that should be set self-consistently to the quasiparticle eigenvalues, \( E^{\text{MF}}_{nk} \) and \( \psi_{nk} \) are the mean-field eigenvalues and eigenvectors and \( \Sigma^{\text{MF}} \) is a mean-field approximation to the electronic self-energy operator, such as \( V_{xc} \) in the case of a DFT starting point.

It is often the case that the mean-field wavefunctions are sufficiently close to the quasiparticle wavefunctions that one may reduce Eq. 7.15 to include only diagonal matrix elements. In this case the user may ask for only diagonal elements, and the quasiparticle energies will be updated in the following way:

\[
E^{\text{QP}}_{nk} = E^{\text{MF}}_{nk} + \langle \psi_{nk} | \Sigma(E) - \Sigma^{\text{MF}}(E) | \psi_{nk} \rangle. \tag{7.16}
\]

The mean field in Eq. 7.16 and Eq. 7.15 can be DFT within the LDA or GGA schemes as well as within a hybrid-functional approach. In the LDA case, for example, \( \Sigma^{\text{MF}}(E) = V_{xc} \), is local and energy-independent. The starting mean-field calculation can also be an approximation to the Dyson’s equation, Eq. 7.2, such as Hartree-Fock (the zero-screening limit) or static COHSEX (the static-screening limit) [301, 42, 165]. The use of these mean-field starting points for construction of Eq. 7.15 and Eq. 7.16 is classified as a one-shot \( G_0 W_0 \) calculation (the 0 subscript means that both \( G \) and \( W \) are constructed from the
mean-field eigenvalues and eigenvectors). One also can start from a previous iteration of
GW in an eigenvalue or eigenvector self-consistency scheme [42, 377]. In this case, the
‘MF’ superscripts in Eq. 7.16 and 7.15 should be renamed “previous” to designate the
self-consistency process.

The \texttt{sigma} executable itself can evaluate the matrix elements of \( \Sigma \) in Eq. 7.16 and Eq.
7.15 within various approximations: Hartree-Fock, static COHSEX, GW within a GPP
model and full-frequency GW.

For GW and static-COHSEX calculations, \( \Sigma \) can be broken into two parts, \( \Sigma = \Sigma_{SX} + \Sigma_{CH} \),
where \( \Sigma_{SX} \) is the screened exchange operator and \( \Sigma_{CH} \) is the Coulomb-hole operator
[152, 140, 141]. These are implemented in the \texttt{sigma} executable in the following way for a
full-frequency calculation:

\[
\langle n| \Sigma_{SX}(E) | n'k \rangle = -\sum_{n''} \sum_{qGG'} M_{n'k}^*(k,-q,-G) M_{n'n''}(k,-q,-G')
\times [\epsilon_{GG'}]^{-1}(q;E-E_{n'k-q}) v(q+G') \tag{7.17}
\]

and

\[
\langle n| \Sigma_{CH}(E) | n'k \rangle = \frac{i}{2\pi} \sum_{n''} \sum_{qGG'} M_{n'k}^*(k,-q,-G) M_{n'n''}(k,-q,-G')
\times \int_{0}^{\infty} dE' \frac{[\epsilon_{GG'}]^{-1}(q;E') - [\epsilon_{GG'}]^{-1}(q;E')}{E-E_{n'k-q} - E' + i\delta} v(q+G') \tag{7.18}
\]

where \( M \) is defined in Eq. 7.8 and \( \epsilon_r \) and \( \epsilon_{a} \) are the retarded and advanced dielectric
matrices defined in Eq. 7.14 [332]. In practice the \texttt{sigma} executable computes the matrix
elements of bare exchange, \( \Sigma_X \) and of \( \Sigma_{SX} - \Sigma_X \), where the matrix elements of \( \Sigma_X \) are obtained by replacing \([\epsilon_{GG'}]^{-1}(q;E-E_{n'k-q}) \) with \( \delta_{GG'} \) in Eq. 7.17 (as given by Eq.
7.28 below).

For GPP calculations, the corresponding expressions used in the code are:

\[
\langle n| \Sigma_{SX}(E) | n'k \rangle = -\sum_{n''} \sum_{qGG'} M_{n'k}^*(k,-q,-G) M_{n'n''}(k,-q,-G')
\times \left[ \delta_{GG'} + \frac{\Omega_{GG'}^2(q)(1-i\tan\phi_{GG'}(q))}{(E-E_{n'k-q})^2 - \omega_{GG'}^2(q)} \right] v(q+G') \tag{7.19}
\]

and

\[
\langle n| \Sigma_{CH}(E) | n'k \rangle = \frac{1}{2} \sum_{n''} \sum_{qGG'} M_{n'k}^*(k,-q,-G) M_{n'n''}(k,-q,-G')
\times \frac{\Omega_{GG'}^2(q)(1-i\tan\phi_{GG'}(q))}{\omega_{GG'}^2(q)(E-E_{n'k-q} - \omega_{GG'}^2(q))} v(q+G') \tag{7.20}
\]

where \( \Omega_{GG'}(q), \omega_{GG'}(q), \lambda_{GG'}(q) \) and \( \phi_{GG'}(q) \) are the effective bare plasma frequency,
the GPP mode frequency, the amplitude and the phase of the renormalized \( \Omega_{GG'}^2(q) \) [152,
defined as:

$$\Omega_{GG'}^2(q) = \omega_p^2 \frac{(q+G)(q+G')}{|q+G|^2} \frac{\rho(G-G')}{\rho(0)}$$  \hspace{1cm} (7.21)

$$\tilde{\omega}_{GG'}^2(q) = \frac{|\lambda_{GG'}(q)|}{\cos \phi_{GG'}(q)}$$  \hspace{1cm} (7.22)

$$|\lambda_{GG'}(q)| e^{i\phi_{GG'}(q)} = \frac{\Omega_{GG'}^2(q)}{\delta_{GG'} - \epsilon_{GG'}^{-1}(q;0)}$$  \hspace{1cm} (7.23)

Here, \(\rho\) is the electron charge density in reciprocal space and \(\omega_p^2 = 4\pi\rho(0)e^2/m\) is the classical plasma frequency. In this case, the integral over energy that is necessary in the full-frequency expression, Eq. 7.18, is reduced to a single term using an analytical approximation to the frequency dependence of the dielectric matrix requiring only the static dielectric matrix \(\epsilon_{GG'}^{-1}(q;0)\) in Eq. 7.22. The analytical approximation is done using the \(f\)-sum rule for each \(GG'\) pair as described in Ref. [152]. This reduces the computational cost of evaluating the \(\Sigma\) matrix elements by a factor of the number of frequencies. It is important to note that for systems without inversion symmetry, \(\rho\) in Eq. 7.21 and \(V_{xc}\) in Eqs. 7.15 and 7.16 are complex functions in reciprocal space (even though these are real functions when transformed to real space). For systems with inversion symmetry, \(\Omega_{GG'}^2(q)\) and \(\tilde{\omega}_{GG'}^2(q)\) are real, \(\phi_{GG'}(q) = 0\) or \(\pi\) and Eqs. 7.19 – 7.23 reduce to a simpler form [152].

In computing the sums in Eqs. 7.19 and 7.20 we drop terms in certain circumstances to save time and improve numerical precision. We neglect the terms for which \(|\delta_{GG'} - \epsilon_{GG'}^{-1}(q;0)|\), \(|\lambda_{GG'}(q)|\) or \(|\cos \phi_{GG'}(q)|\) are less than a given tolerance, since these terms have a vanishing contribution to the matrix elements of the self energy. This avoids ill-conditioned limits due to some of the intermediate quantities here being undefined. Another case is when for an occupied state \(n''\), \(E - E_{n''k-q} - \tilde{\omega}_{GG'}(q) \approx 0\), in which case the GPP factors in \(\Sigma_{SX}\) and \(\Sigma_{CH}\) each diverge, although the sum

$$-\delta_{GG'} + \frac{\Omega_{GG'}^2(q)(1 - i \tan \phi_{GG'}(q))}{2\tilde{\omega}_{GG'}(E - E_{n''k-q} + \tilde{\omega}_{GG'}(q))}$$  \hspace{1cm} (7.24)

remains finite. In this situation, we do not calculate these terms in \(\Sigma_{SX}\) and \(\Sigma_{CH}\) separately, but assign the sum of the contributions to \(\Sigma_{SX}\). When \(n''\) is unoccupied there is only a \(\Sigma_{CH}\) contribution which diverges. Similarly, there are divergent contributions to \(\Sigma_{SX}\) when \(E - E_{n''k-q} + \tilde{\omega}_{GG'}(q) \approx 0\). In the full-frequency integrals in Eqs. 7.17 and 7.18, we can see that the contributions around a pole of \(\epsilon_{GG'}^{-1}\) in this case vanish, so the correct analytic limit of these terms is zero [152].

For static COHSEX calculations, the expressions used in the code are:

$$\langle nk | \Sigma_{SX}(0) | n'k \rangle = -\sum_{n''} \sum_{qGG'} \epsilon_{GG'}^{-1}(q;0) \epsilon_{GG'}(q+G') \times M_{n''n}(k-q,-G)M_{n'n'}(k,-q,-G')$$  \hspace{1cm} (7.25)
and
\[ \langle n | \Sigma_{CH}(0) | n' k \rangle = \frac{1}{2} \sum_{n''} \sum_{qGG'} M_{n''n}(k, -q, -G) M_{n'n}(k, -q, -G') \frac{1}{\epsilon_{GG'}^{-1}(q; 0) - \delta_{GG'}} v(q + G') \times [\epsilon_{GG'}^{-1}(q; 0) - \delta_{GG'}] v(q + G') \]
\[ = \frac{1}{2} \sum_{qGG'} M_{nn'}(k, q = 0, G' - G) [\epsilon_{GG'}^{-1}(q; 0) - \delta_{GG'}] v(q + G') \] (7.26)

where Eqs. 7.25 and 7.26 can be derived formally from Eqs. 7.19 and 7.20 by setting \((E - E_{n' r m} - q)\) to zero. Using the completeness relation for the sum over empty states, Eq. 7.26 can be written in a closed form given by Eq. 7.27, which now does not involve the empty orbitals.

For Hartree-Fock calculations, we compute the matrix elements of bare exchange:
\[ \langle n | \Sigma_X | n' k \rangle = -\sum_{n''} \sum_{qGG'} M_{n''n}(k, -q, -G) M_{n'n}(k, -q, -G') \delta_{GG'} v(q + G') \] (7.28)

Eq. 7.16 depends on the evaluation energy parameter \(E\). This parameter should be the quasiparticle energy \(E_{n' k}^{QP}\), determined self-consistently. In many systems, \(\Sigma(E)\) is a nearly linear function of \(E\) so one may compute \(\Sigma(E)\) for two grid points and evaluate the self-consistent \(E_{n' k}^{QP}\) using Newton’s method [152]:
\[ E_{n' k}^{QP} = E_{n' k}^{0} + \frac{d\Sigma/dE}{1 - d\Sigma/dE} (E_{n' k}^{0} - E_{n' k}^{MF}) \] (7.29)

The derivative that appears here is also related to the quasiparticle renormalization factor:
\[ Z = \frac{1}{1 - d\Sigma/dE} \] (7.30)

7.5.4 Optical Properties: BSE

The optical properties of materials are computed in the Bethe-Salpeter equation (BSE) executables. Here the eigenvalue equation represented by the BSE, Eq. 7.3, is constructed and diagonalized yielding the excitation energies and wavefunctions of the correlated electron-hole excited states. There are two main executables: kernel and absorption. In the former, the electron-hole interaction kernel is constructed on a coarse \(k\)-point grid, and in the latter the kernel is (optionally) interpolated to a fine \(k\)-point grid and diagonalized.

The kernel executable constructs the second term of the left-hand side of Eq. 7.3 which is referred to as the electron-hole kernel. The kernel, \(K\), as implemented in the package, is limited to the static approximation, and contains two terms, a screened direct interaction and a bare exchange interaction, \(K^{eh} = K^{d} + K^{x}\), defined in the following way [303]:
\[ \langle vck | K^{d} | v' c' k' \rangle = -\int dx dx' \psi_{vck}^{*}(r) \psi_{v' c' k'}(r) W(r, r') \psi_{v' c' k'}^{*}(r') \psi_{vck}(r') \] (7.31)
and

\[
\langle v c | K^x | v' c' \rangle = \int d r d r' \psi_{v c}(r) \psi_{v c}(r) \psi_{v' c'}(r') \psi_{v' c'}(r').
\] (7.32)

These matrices are constructed on a coarse grid of \( k \)-points, in most cases the same grid used within the GW calculation because one must have previously constructed the dielectric matrix \( \epsilon^{-1}(q) \) for \( q = k - k' \). We calculate these matrices in \( G \)-space using the prescription of Rohlfing and Louie [303]:

\[
\langle v c | K^d | v' c' \rangle = \sum_{G G'} M_{v' c'}(k, q, G) W_{G G'}(q; 0) M_{v v}(k, q, G')
\] (7.33)

and

\[
\langle v c | K^x | v' c' \rangle = \sum_{G \neq 0} M_{v' c'}(k, q, G) v(q + G) M_{v v}(k, q, G)
\] (7.34)

where \( M \) is defined in Eq. 7.8 and calculated using FFTs as described above in Sec. 7.5.2.

Because exciton binding energies and absorption spectra depend sensitively on quantities like the joint density of states, it is essential in periodic systems to sample the \( k \)-points on a very fine grid. Directly calculating the kernel on this fine grid in the kernel executable would be prohibitively expensive, so instead we interpolate the kernel in the absorption executable before diagonalization. For semiconductors, the head and wing kernel terms are not smooth functions of \( k \) and \( k' \) (as we have shown above, they diverge for \( q = k - k' \to 0 \)). Therefore, the quantities that we interpolate are \( q^2 \cdot K^d_{\text{head}}, q \cdot K^d_{\text{wing}} \) and the body term directly as they are now smooth quantities [303]. For metals, we interpolate directly the kernel without any caveats because all the contributing terms are smooth functions of \( k \) and \( k' \). As in GW, we treat metals with zero-temperature occupations, and we also neglect intraband transitions.

The absorption executable requires both coarse- and fine-grid wavefunctions as input. The interpolation is done through a simple expansion of the fine-grid wavefunction in terms of nearest coarse-grid wavefunction:

\[
u_{n k f} = \sum_{n'} C_{n n'}^{k_{co}} u_{n' k_{co}}
\] (7.35)

where \( k_{co} \) is the closest coarse-grid point to the fine-grid point, \( k_{fi} \), and the coefficients \( C_{n n'}^{k_{co}} \), are defined as the overlaps between the coarse-grid and fine-grid wavefunctions:

\[
C_{n n'}^{k_{co}} = \int d r u_{n k_{fi}}(r) u_{n' k_{co}}^*(r).
\] (7.36)

The coefficients \( C_{n n'}^{k_{co}} \) are normalized so that \( \sum_{n'} |C_{n n'}^{k_{co}}|^2 = 1 \). It should be noted that for a given set of fine bands one can improve the interpolation systematically by including more valence and conduction bands in the coarse grid due to the completeness of the Hilbert space at each \( k \). It should also be noted that we do restrict \( n \) and \( n' \) to be either both valence or both conduction bands – this is acceptable due to the different character of the conduction and valence bands in most systems.
Using these coefficients we interpolate the kernel with the following formula:

\[
\langle v c k | n_1 n_2 n_3 n_4 \rangle = \sum_{n_1, n_2, n_3, n_4} C^{k}_{n_1} C^{\ast}_{n_2} C^{k'}_{n_3} C^{\ast}_{n_4} \langle n_2 n_1 k_c | n_4 n_3 k'_{c} \rangle
\]  

(7.37)

where \(K\) is one of the head, wing, body or exchange kernel terms. As in the case of \(\epsilon\), this summation can be performed compactly as a set of matrix-matrix multiplications. We utilize the Level 3 BLAS calls \texttt{DGEMM} and \texttt{ZGEMM} to optimize the performance.

One can improve on the interpolation systematically by using the closest four coarse-grid points to each fine point and using a linear interpolation layer in addition to the wavefunction-based interpolation described above. This is done by default for the interpolation of the first term of Eq. 7.3 for the quasiparticle self-energy corrections \(E_{QP} - E_{MF}\):

\[
E_{QP}^{n}(k_{n}) = E_{MF}^{n}(k_{n}) + \left\langle \sum_{n'} \left| \bar{C}^{k}_{n,n'} \right|^{2} \left( E_{QP}^{n'}(k_{n'}) - E_{MF}^{n'}(k_{n'}) \right) \right\rangle_{k_{n'}}
\]  

(7.38)

where the brackets indicate linear interpolation using the tetrahedron method. In this case, the wavefunction-based interpolation layer guarantees that the band crossings are properly handled, and the linear interpolation layer ensures that we correctly capture the energy dependence of the self-energy corrections. In this way, we can construct \(E_{QP}\) on the fine grid, or any arbitrary point, given \(E_{MF}\) on the fine grid and \(E_{QP}\) and \(E_{MF}\) on the coarse grid.

As an alternative to calculating the quasiparticle corrections on the coarse grid and interpolating them to the fine grid, the user may choose a less refined method of specifying the corrections using a three-parameter model involving a scissor-shift parameter \(\Delta E\) to open the energy gap at the Fermi energy, a zero energy \(E_{0}\) (typically the band edge), and an energy-scaling parameter \(C\) changing the bandwidth (the parameters are specified separately for valence and conduction bands):

\[
E_{QP} = E_{MF} + \Delta E + C \left( E_{MF} - E_{0} \right).
\]  

(7.39)

Having constructed the kernel on the fine grid, we now consider the diagonalization of the kernel. The kernel matrix is of dimension \(N_c \cdot N_v \cdot N_k\) where \(N_k\) is the number of \(k\)-points on the fine grid. The matrix can be diagonalized exactly within LAPACK (\texttt{zheevx}) or ScALAPACK (\texttt{pzheevx}). The result of the diagonalization is the set of exciton eigenvalues \(\Omega^S\) and eigenfunctions \(A^S_{c k}\) which can be used to construct the joint density of states (JDOS) or the absorption spectrum (or \(\text{Im } \epsilon_2(\omega)\)) using Eq. 7.5. There are a number of post-processing tools in the package, such as \texttt{PlotXct}, which plots the exciton wavefunction in real space according to Eq. 7.4 to analyze the exciton states. The \texttt{absorption} executable can produce both the singlet and triplet eigenvalues and eigenvectors. In the latter case, the exchange term is set to zero when diagonalizing the kernel [303]. (Note that for triplets the oscillator strengths calculated in the code are evaluated without considering spin overlap. In some cases this “oscillator strength of the corresponding singlet” may be useful. The true physical oscillator strength of course is zero for triplets.) Additionally, one may compute the eigenvalues and eigenvectors with only the exchange interaction; the resulting spectrum should be the same as the one obtained within RPA with local-field effects included [137].
As mentioned above, the electron-hole kernel should be constructed with a sufficient number of valence and conduction bands to cover the energy window of interest – typically all bands within the desired energy window from the Fermi energy should be included so that the energy window of the bands included in the calculation is at least twice that of the desired absorption energy window. The absorption executable computes the percent deviation from the $f$-sum rule \[152\]:

$$
\int_{0}^{\infty} \epsilon_2(\omega) \omega d\omega = -\frac{\pi \omega_0^2}{2}.
$$

One should converge this quantity with both the number of valence and conduction bands included. The absorption spectrum (or $\epsilon^2$) in the energy window of interest converges much more quickly than $\epsilon_1$ if high-energy transitions outside of the window of interest contribute greatly to the sum rule, since $\epsilon_1(\omega)$ is related to an integration over all frequencies of $\epsilon_2(\omega)$ via the Kramers-Kronig relation.

Finally, the transition matrix elements $\langle 0 | v | S \rangle$ in Eq. 7.5 are printed in the file eigenvalues.dat. These are related to the oscillator strengths $f_S$ by

$$
f_S = \frac{2 |e \cdot \langle 0 | v | S \rangle|^2}{\Omega^S}.
$$

We compute the velocity matrix element via the commutator of the many-body Hamiltonian, as follows [362]:

$$
\langle v k | r | c k \rangle = \lim_{q \to 0} \frac{\langle v k + q | e^{i qr} - 1 | c k \rangle}{i q} = -i \lim_{q \to 0} \frac{\langle v k + q | e^{i qr} | c k \rangle}{q}.
$$

In practice, we evaluate the limit using finite differences for a small value of $q$, similarly to how the limit is treated in the Epsilon code. Thus the valence bands on a shifted fine k-grid are required. (Note we are assuming an interband transition.)

As an alternative to the finite-difference approach, one may approximate the velocity operator by the momentum operator $-i \nabla$, avoiding the calculation of the valence bands on the shifted fine grid. We reverse the derivation partly:

$$
\langle v k | r | c k \rangle = \frac{\langle v k | [H^{MF}, r] | c k \rangle}{E_{vc}^{MF} - E_{ck}^{MF}} = -i \frac{\langle v k | v^{MF} | c k \rangle}{E_{vc}^{MF} - E_{ck}^{MF}} \approx -\frac{\langle v k | \nabla | c k \rangle}{E_{vc}^{MF} - E_{ck}^{MF}}.
$$

This does not require an additional grid, but yields inexact oscillator strengths due to neglect of commutators between $r$ and the non-local part of the Hamiltonian [303, 158]. We could have approximated the quasiparticle or excitonic velocity operator by the momentum operator, but this would be less accurate, since those Hamiltonians have additional sources of non-locality beyond those of the mean-field Hamiltonian. The momentum operator uses
transitions $v k \to c k$ in the Bethe-Salpeter equation (Eq. 7.3), whereas with the velocity operator we actually use the transitions $v k + q \to c k$. This ensures consistency of wavefunctions between excitons and transition matrix elements, and also is needed to describe transverse and longitudinal excitons in materials with an indirect gap [303].

### 7.5.5 Grid Uniformity

Another important consideration in performing integrals over $q$ of the Coulomb interaction in sigma and kernel is the sampling of the grid on which the integral is done. If the sampling in different directions is very different, then the result of the integrals will not go to the correct limit, since it will resemble a 1D or 2D integration rather than a 3D integration. This issue is important for calculation of nano-systems without truncation, or for highly anisotropic crystals. We determine the effective sampling in each direction as follows: take the vectors $b_i/N_i$, where $b_i$ is a reciprocal lattice vector and $N_i$ is the number of $q$-points in that direction. Find the shortest vector. Orthogonalize the next shortest vector to that one. Orthogonalize the remaining vector to the first two. (It is important to use this order since orthogonalization always makes the vector shorter.) Now compare the lengths of these orthogonalized vectors: if the ratio between the longest and shortest is too large (we use a factor of 2 as a tolerance), the grid is non-uniform and may give incorrect answers. The code will write a warning in this case, and the user should try to use a more nearly uniform grid, or check the convergence of results against the cell-averaging cutoff. Note that the sampling in any direction in which the Coulomb interaction is truncated is irrelevant when checking for grid uniformity.

### 7.6 Symmetry and degeneracy

#### 7.6.1 Mean field

As was mentioned in the Introduction, the largest cost when performing a GW calculation with the BerkeleyGW package is the generation of the input mean-field states. In order to reduce this cost, all the codes allow the user to input the wavefunctions in only the reduced Brillouin zone and construct the wavefunctions in the full zone by the following relation:

$$\phi_{R(k)}(G) = \phi_k(R^{-1}(G))e^{-iG \cdot \tau} \tag{7.45}$$

where the symmetry operation is defined by a reciprocal-space rotation matrix $R$ and a fractional translation $\tau$ such that $x' = R^{-1}x + \tau$.

The main $k$-grid used is generated by constructing a uniform grid (with a possible shift, typically half a unit as in Monkhorst-Pack grids) and then reducing by the symmetry operations. The shifted grid used in epsilon for constructing eps0mat at a small $q$-vector is generated by unfolding this reduced set of points with all the symmetry operations, reducing again by the symmetry operations of the subgroup of the $q$-vector, and then applying a small $q$-shift. (Note that, contrary to a naive expectation, this reduced shifted grid may contain many more points than the original uniform grid.) An example for graphene is given in Fig. 7.4. This procedure provides all the $k$-points needed for calculations in BerkeleyGW.
and it is implemented in the utility kgrid.x in the MeanField/ESPRESSO directory, which should be used to generate the set of k-points in mean-field calculations. PARATEC also has built-in support for this construction. The symmetry analysis in kgrid.x, epm2bgw.x, and siesta2bgw.x is performed with the spglib library [367].

Figure 7.4: An example of the construction of 4 × 4 main and shifted k-grids for graphene. (Left) The main grid has a (0.5, 0.5) shift (crystal coordinates). There are 16 points in the full Brillouin zone and 6 irreducible points. (Right) The shifted grid has a (0.0, 0.05) shift (crystal coordinates). There are 48 points in the full Brillouin zone and 26 irreducible points.

7.6.2 Self-energy operator

Degeneracy is also important from the point of view of the states on which self energies are calculated, as opposed to those appearing in the sum. Since the self-energy operator has the full symmetry of the system, the matrix elements between states belonging to different representations are zero by symmetry. In the presence of high symmetry, this consideration can make the matrix quite sparse. To take advantage fully of symmetry here would require a careful analysis of each wavefunction’s behavior under various symmetry operations and comparison to character tables of space groups. Users certainly can do this in deciding which off-diagonal self-energy matrix elements to calculate. The Sigma takes a very simple approach to identify some of the elements which are zero by symmetry, based on degeneracy. The multiplicity of the degenerate subspace to which each states belongs is counted (1, 2, or 3 for the standard space groups), and clearly two states in subspaces of different multiplicity must belong to different representations, and their matrix element can be set to zero without calculation. This saves time and enforces symmetry.

Application of symmetry in a degenerate subspace can also speed up calculation of
diagonal elements of the self-energy operator. The expressions for the exchange, screened exchange, and Coulomb-hole parts contain a sum over $q$. In general, this must be done over the whole Brillouin zone, but to calculate the sum of the self energies within a degenerate subspace it is sufficient to use the irreducible part of the Brillouin zone. Each part of $\Sigma$, in the various approximations, has the generic form

$$\langle nk | \Sigma | n'k \rangle = -\sum_{n''} \sum_{qGG'} \langle nk | e^{i(\mathbf{q}+\mathbf{G}) \cdot \mathbf{r}} | n''k - q \rangle \langle n''k - q | e^{-i(\mathbf{q}+\mathbf{G'}) \cdot \mathbf{r}} | n'k \rangle$$

(7.46)

$$\times F(q, G, G')$$

The summand is invariant under application of a symmetry operation $O$ in the subgroup of $k$ provided that $n = n'$ and $n$ and $n''$ are non-degenerate, since in that case the action of the operation simply introduces a phase: $O |mk\rangle = e^{i\theta} |mk\rangle$ (degenerate states may instead transform into linear combinations in the degenerate subspace). These phases are canceled by the fact that each state appears also with its complex conjugate. If the states $n''$ in the sum are degenerate, the summand is not invariant but the sum is, if the whole degenerate subspace is summed over, since then we are taking the trace of the projector matrix $|n''k\rangle \langle n''k|$ in that subspace, which is invariant [153]. If $n$ is degenerate, then $\langle nk | \Sigma | nk \rangle$ is not invariant, but the trace of the self-energy in the degenerate subspace, $\sum_n \langle nk | \Sigma | nk \rangle$, is invariant. Therefore, to calculate diagonal elements for a whole degenerate subspace, for each state we sum only over $q$ in the irreducible zone, with weight $W_q$ from the number of $q$-vectors related to $q$ by symmetry. We then symmetrize by assigning the average to each:

$$\langle mk | \Sigma | mk \rangle = \frac{1}{N_{\text{deg}}} \sum_n \langle nk | \Sigma | nk \rangle = -\sum_{n} \sum_{n''} \sum_{GG'} \sum_{q} W_q \langle nk | e^{i(\mathbf{q}+\mathbf{G}) \cdot \mathbf{r}} | n''k - q \rangle$$

$$\times \langle n''k - q | e^{-i(\mathbf{q}+\mathbf{G'}) \cdot \mathbf{r}} | nk \rangle F(q, G, G')$$

(7.47)

(This averaging over degenerate bands is also done to enforce symmetry even when we use the full $q$-sum, since the results may differ slightly due to limited precision in the wave-functions from the mean-field calculation.) If we are calculating only part of a degenerate subspace, this trick does not work, and we must perform the complete sum. For diagonal elements, the code by default uses the irreducible $q$-sum and will write an error if the calculation requires the full sum because of degeneracy, directing the user to enable it via the flag `no_symmetries_q_grid`, or include all states in the degenerate subspace. For off-diagonal elements ($n \neq n''$), even if both are non-degenerate, application of the symmetry operation, in general, introduces different phases from the two states, which are not canceled. Thus the contributions from different $q$-points related by symmetry differ, so that the full sum must always be used.

7.6.3 Bethe-Salpeter equation

Degeneracy must be considered in BSE calculations as well, when choosing the subspace in which to work. If the set of occupied or unoccupied states includes only part of a
degenerate subspace, then the solutions found by absorption will break symmetry and can give qualitatively incorrect results. For example, an excitation that should have zero oscillator strength by symmetry, due to interference between transitions to two degenerate states, may not be dark if only one of those transitions is included. This issue is quite general and applies to the choice of active spaces in other theories as well, such as configuration interaction [413]. Breaking degeneracy in either the coarse or fine grid can also cause trouble in interpolation of the kernel and quasiparticle energies.

### 7.6.4 Degeneracy utility

We provide a utility called `degeneracy_check.x` which reads wavefunction files and writes out a list of acceptable numbers of bands. Multiple wavefunction files can be checked at once, for example the shifted and unshifted grids in Epsilon or shifted, unshifted, coarse, and fine grids for Bethe-Salpeter equation calculations, in which case the utility will identify numbers of bands which are consistent with degeneracy for every file.

### 7.6.5 Real and complex flavors

The component executables come in two “flavors,” real and complex, specified at compile time and denoted by the suffix `.real.x` or `.cplx.x`. When the system has inversion and time-reversal symmetry, we can choose the wavefunctions to be real in reciprocal space. The plane-wave expansions are:

\[
u (r) = \sum_G u_G e^{iG \cdot r} \tag{7.48}\]

\[
u (-r) = \sum_G u_G e^{-iG \cdot r} \tag{7.49}\]

\[
u^* (r) = \sum_G u_G^* e^{-iG \cdot r} \tag{7.50}\]

The symmetry conditions mean that wavefunctions can be chosen to satisfy \( \nu (-r) = au (r) \) (inversion symmetry) and \( \nu^* (r) = bu (r) \) (time-reversal, equivalent to taking the complex conjugate of the Schrödinger equation), with \( a, b \) each equal to \( \pm 1 \) depending on whether the wavefunction belongs to an odd or even representation. Thus we can choose \( \nu (-r) = cu^* (r) \) with \( c = ab \) also equal to \( \pm 1 \). Combining this with the plane-wave expansions,

\[
\sum_G u_G e^{-iG \cdot r} = c \sum_G u_G^* e^{-iG \cdot r} \tag{7.51}
\]

\[
u_G = cu_G^* \tag{7.52}\]

The choice \( c = 1 \) corresponds to real coefficients; \( c = -1 \) corresponds to pure imaginary coefficients. Most plane-wave electronic-structure codes always use complex coefficients, and so the coefficients will in general not be real, even in the presence of inversion and time-reversal symmetry. For a non-degenerate state, the coefficients will be real times an arbitrary global phase, determined by the initialization of the solution procedure. We must divide out this global phase to make the coefficients real. In a degenerate subspace, the
states need not be eigenstates of inversion, and so in general they may not just be real times a global phase. Instead, in each subspace of degeneracy $n$ we take the $2n$ vectors given by the real and imaginary parts of each wavefunction, and then use a Gram-Schmidt process to find $n$ real orthonormal wavefunctions spanning the subspace.

The density and exchange-correlation potential are real already in the presence of inversion symmetry and there is no arbitrary phase possible. The real-space density is always real: $\rho(r) = \rho^*(r)$. With inversion symmetry, we also have $\rho(r) = \rho(-r)$. In reciprocal space,

$$\rho(r) = \sum_G \rho_G e^{iG\cdot r}$$  \hfill (7.53)

$$\rho^*(r) = \sum_G \rho_G^* e^{-iG\cdot r}$$  \hfill (7.54)

$$\rho(-r) = \sum_G \rho_G e^{-iG\cdot r}$$  \hfill (7.55)

Together, these relations imply $\rho_G = \rho_G^*$, i.e. the reciprocal-space coefficients are real. Precisely the same equations apply for the exchange-correlation potential.

The wavefunction, density, and exchange-correlation potential are then all stored as real coefficients, saving disk space (for the files), memory, and operations compared to the complex representation. Usually, only the lack of inversion symmetry of the lattice and basis would require the use of complex wavefunctions, but if spin-orbit coupling or magnetic fields are present, then time-reversal symmetry is lost and complex wavefunctions again are required.

### 7.7 Computational Issues

#### 7.7.1 Memory estimation

In the beginning of each run, all the major code components print the amount of memory available per CPU and an estimate of memory required per CPU to perform the calculation. If the latter exceeds the former, the job is likely to fail with a memory allocation error. The amount of memory required is estimated by determining the sizes of the largest arrays after reading in the parameters of the system from the input files. A straightforward approach to estimating the amount of available memory is to allocate memory by incremental amounts until the allocation call returns with an error. Unfortunately, in many implementations the allocation call returns without an error even if the requested amount of memory is not physically available, but the system fails when trying to access this “allocated” memory. We implement another approach based on the Linux /proc file system.

First, each CPU opens file /proc/meminfo and reads in the values of MemFree and Cached. The sum of these two values gives the amount of memory available per node. This approach works on almost all modern high-performance computing systems where the Linux /proc File System is accessible. (However, for BSD-based MacOS which lacks /proc/meminfo, we read the page size and number of free and speculative pages from the command vm_stat.)

Second, each CPU calls an intrinsic Fortran routine that returns the host name which is
unique for each node. By comparing host names reported by different CPUs we identify
the number of CPUs per node. The amount of memory available per CPU is then given by
the ratio of the amount of memory available per node to the number of CPUs per node.

7.7.2 Makefiles

The main codes are in the Epsilon, Sigma, BSE, PlotXct, and MeanField directories.
Routines used by all parts are in the Common directory, and routines common to some of
the MeanField codes are in the Symmetry directory. The Makefiles are designed for GNU
Make, and enable targets in a directory to be built from any level of the directory hierarchy.
They contain a full set of dependencies, including those between directories, to ensure that
the build is correct after any changes to source, for ease in development and modification.
This also enables use of parallel make on large numbers of processors for rapid builds –
any omissions in the dependencies generally cause a failure for a parallel make. The special
make target all-j (i.e. make -j all-j) begins by using all processes to build common
directories, which contain files required by files in a large number of directories; otherwise,
the build would fail due to attempts by multiple processes to read and write the same files
in these directories. Commonly, Fortran Makefiles are set up with object files depending
on other object files. However, the real situation is that object files depend on module files
(.mod) for the modules they use, and only executables depend on object files. Therefore we
have dependencies directly on the module files to ensure the required files are present for
compilation, particularly for parallel builds.

7.7.3 Installation instructions

The code can be installed via the following steps:

    cp [flavor_real.mk/flavor_cplx.mk] flavor.mk
    ln -s config/[mysystem].mk arch.mk
    make all
    make check[-jobscript]

First a flavor is selected by copying the appropriate file to flavor.mk. Then a configu-
ration file must be put as arch.mk. Configurations appropriate for various supercomputers
as well as for using standard Ubuntu or Macports packages are provided in the config
directory. Appropriate paths, libraries, and compiler flags can be set in a new arch.mk for
other systems. Finally the test suite should be run to confirm that the build is working.
In serial, the command is make check; for parallel builds, it is make check-parallel. On
machines with a scheduler, a job script should be created to run this command. For the
architectures supported in config, job scripts to run the test suite are provided in the
testsuite directory, and can be used via make check-jobscript.

7.7.4 Validation and verification

The importance of verification and validation of complicated scientific software packages
is receiving increasing attention. We use standard open-source tools for code development,
following accepted best practices [235]. Development is done with the subversion (SVN) version-control system [345] and Trac, an issue-tracking system and interface to SVN [371]. All code runs identify the version and revision number used in the output for traceability of results, implemented via a special source file called svninfo.f90 which all SVN revisions must modify (enforced via a pre-commit hook). Debug mode can be enabled via -DDEBUG in the arch.mk file, which performs extra checking including of dynamic memory allocation and deallocation. A macro enables a check of the status returned by the system after an allocation attempt, and reports failures, identifying the array name, size, source file, and line number, as well as which processor failed to allocate the array. Additionally, it keeps track of the amount of memory dynamically allocated and deallocated, so the code can report at the end of each run how much memory remains allocated, and the maximum and minimum memory ‘high-water-mark’ among the processors. In debug mode, a stack trace also can be enabled, either on just the root processor, or on all processors (causing the code to run much slower), which can be used to locate where problems such as segmentation faults are occurring (possibly on only one processor). Verbose mode can be enabled via -DVERBOSE which writes extra information as the calculation proceeds.

The package contains a comprehensive test suite to test the various executables, run modes, and options, in the testsuite directory. Calculations of several different physical systems, with mean-field, epsilon, sigma, and BSE calculations, are carried out (including use of PlotXct and some utilities), detecting any run-time errors and showing any warnings generated. Then selected results are extracted from the output and compared to reference information within a specified tolerance. The actual calculated values, as well as timing for each step, are displayed. Each match is shown as either OK or FAIL, and a final summary is written of failures. The calculations are small and generally underconverged, to make them quick enough for routine testing and rapid feedback. The mean-field steps are either EPM (quick serial calculations) or stored compressed output from DFT calculations. The Epsilon, Sigma, and BSE calculations are run either in serial or on 4 processors (for parallel builds).

The test suite has numerous uses. It is useful for users to verify the success of a new build of the code on their platform (failures could be due to library problems, excessive optimizations, etc.). It is used for developers to verify that the code is giving reproducible answers, ensure consistency between serial/parallel runs, as well as real/complex and spin-polarized/ uncompolarized runs, and check that the code works with new compilers or libraries. On a routine basis, the test suite is also useful for developers to check that changes to the code do not introduce problems. The driver scripts (run_testsuite.sh and run_regression_test.pl) and specifications for the files defining the test steps are originally based on, and developed in conjunction with, those of the Octopus code [55, 218]. This framework is quite general and can be used easily for constructing a test suite for another code. It can be run in serial with the command make check (or make check-save to retain the working directories from the runs), or in parallel with make check-jobscript (or make check-jobscript-save). The system configuration file arch.mk can specify how to submit an appropriate jobscript for parallel execution on a supercomputer using a scheduler. Scripts are provided in the testsuite directory for some supercomputers. The test suite is used with a continuous-integration system, the open-source tool BuildBot [44], to
ensure the integrity of the code during development. Each commit to the SVN repository triggers a build of the code on each of 10 “buildslaves,” which have different configurations with respect to serial/parallel, compilers, and libraries. After the build, the test suite is run. BuildBot will report to the developers if the either the build or test runs failed, so the problem can be quickly remedied. Use of the various different buildslave configurations helps ensure that the code remains portable across different platforms and in accordance with the language standards. Two of the buildslaves are on a supercomputer with a scheduler, a situation for which standard BuildBot usage is problematic. We provide a Perl script buildbot.pbs.pl that can submit jobs, monitor their status, capture their output for BuildBot, and determine success or failure. This script is general for any PBS scheduler and can be used for other codes too.

### 7.7.5 Supported operating systems, compilers, and libraries

With the test suite, we have tested the code extensively with various configurations, and support the following compilers and libraries:

- Operating systems: Linux, AIX, MacOS
- Fortran compilers (required): pgf90, ifort, gfortran, g95, openf90, sunf90, pathf90, crayftn, af90 (Absoft), nagfor, xlf90 (experimental)
- C compilers (optional): pgcc, icc, gcc, opencc, pathcc, craycc
- C++ compilers (optional): pgCC, icc, g++, openCC, pathCC, crayCC
- MPI implementation (optional): OpenMPI, MPICH1, MPICH2, MVAPICH2
- LAPACK/BLAS implementation (required): NetLib, ATLAS, Intel MKL, ACML, Cray LibSci
- ScaLAPACK/BLACS implementation (required by BSE if MPI is used): NetLib, Cray LibSci, Intel MKL, AMD
- FFTW (required): versions 2.1.5, 2.1.5.1, 2.1.5.2

### 7.8 Utilities

#### 7.8.1 Band-structure interpolation

To plot the quasiparticle band-structure of a system, we provide two methods. The first, sig2wan is a utility that uses Wannier interpolation of the band-structure [225, 331] and is based on the Wannier90 [241] package. This utility does not construct the Wannier functions; the user has to do that using the package used to construct the mean-field eigenfunctions, PARATEC or Quantum ESPRESSO. Once the Wannier functions have been constructed, this utility replaces the mean-field eigenvalues in the wannier.eig file (generated by Wannier90) with the quasiparticle eigenvalues. The user then can rerun the Wannier90 executable to generate band-structures along arbitrary directions. Because one replaces the
mean-field eigenvalues with quasiparticle eigenvalues, this approach does not work well if one replaces only some of the mean-field eigenvalues with quasiparticle ones for entangled bands. The second method for plotting quasiparticle band-structure, the `inteqp` utility, uses Eq. 7.38 to construct the band-structure along arbitrary directions. In this method, we interpolate directly the quasiparticle corrections, which are significantly smoother functions of $k$ and $E$ than the quasiparticle eigenvalues themselves. Therefore, this method requires both the mean-field eigenvalues and eigenfunctions along the desired band-structure direction.

7.8.2 Other

We provide a general-purpose utility called `mf_convert.x` which can convert between binary and ASCII formats of wavefunction, density, and exchange-correlation potential files. The real/complex flavor is determined by reading the file header, and if the utility is called through `mf_convert_wrapper.sh`, the binary/ASCII format is detected via the `grep` command and need not be specified. This converter is useful for moving such files between different platforms, since the binary files are more compact and the form read by the code, but are not necessarily portable between different platforms, whereas the ASCII files are.

The image-charge model (ICM) is implemented in a utility called `icm.x`, based on the Surface code. For a molecule weakly coupled to a metallic surface, the self-energy correction to a state can be well approximated by the sum of the self-energy correction of that state in the isolated molecule and an additional term due to screening from the metal [250]. This screening term is modeled as the electrostatic energy of the charge density of the wavefunction and its induced image-charge distribution in the metal. Let the operator $R$ be a reflection across an image plane. Then

$$
\Delta \Sigma = \pm \frac{1}{2} \int \int \psi(r) \psi^*(r') \frac{1}{|r - Rr'|} \psi(r') \psi^*(r) \, dr \, dr'
$$

(7.56)

where the plus sign applies for occupied orbitals and the minus sign for unoccupied orbitals. This approximation is useful for modeling scanning-tunneling spectroscopy of molecules absorbed on metal surfaces [353] and for quantum-transport calculations of molecular junctions [288] (Chapter 8).

7.9 Acknowledgments

We acknowledge the following people for their contributions to earlier version of the package: Xavier Blase, Andrew Canning, Eric K. Chang, Mark S. Hybertsen, Sohrab Ismail-Beigi, Je-Luen Li, Jeff Neaton, Cheol-Hwan Park, Filipe J. Ribeiro, Gian-Marco Rignanese, Catalin D. Spataru, Murilo L. Tiago, Li Yang and Peihong Zhang. We’d also like to thank the following beta users for their feedback, bug-reports, patches and general assistance while developing the code: Sangkook Choi, Peter Doak, Felipe Jornada, Brad D. Malone, Sahar Sharifzadeh, Isaac Tamblyn and Derek Vigil and the various other members of the Louie and Cohen groups at the University of California, Berkeley.
7.10 Specification of file formats

Wavefunction files are needed by all parts of the code, with various filenames. The 
\texttt{epsilon} executable uses an unshifted grid (WFN) and a shifted grid (WFNq). The \texttt{sigma} executable uses \texttt{WFN\_inner} to construct the self-energy operator and evaluates matrix elements with \texttt{WFN\_outer}. The \texttt{kernel} executable constructs kernel matrix elements with a coarse unshifted (WFN\_co) and shifted grid (WFNq\_co). The \texttt{absorption} and \texttt{plotxct} executables use a fine unshifted grid (WFN\_fi) and, with the velocity operator, a fine shifted grid (WFNq\_fi). Additionally, the \texttt{sigma} executable needs the charge-density RHO for GPP calculations, and needs the exchange-correlation potential VXC (unless its pre-computed matrix elements are supplied in a vxc.dat file). These files all share a common format, which begins with a header. Parts in italics are only for wavefunction files, not charge-density or exchange-correlation potential files. Each bullet represents a record in the file. The utility \texttt{wfn\_rho\_vxc\_info.x} can read the information from the header and report it in a comprehensible format to the user. A module of driver read/write routines for the formats specified here is available in the library directory, for mean-field codes to use in writing output for BerkeleyGW. This library is used by Octopus and PARATEC.

- \([\texttt{WFN/RHO/VXC}-][\texttt{Real/Complex}]\) date time
- number of spins, number of G-vectors, number of symmetries, [0 for cubic symmetry/1 for hexagonal symmetry], number of atoms, charge-density cutoff (Ry), number of k-points, number of bands, maximum number of G-vectors for any k-point, wavefunction cutoff (Ry)
- FFT grid(1:3), k-grid(1:3), k-shift(1:3)
- real-space cell volume (a.u.), lattice constant (a.u.), lattice vectors(1:3, 1:3) in units of lattice constant, real-space metric tensor(1:3, 1:3) (a.u.)
- reciprocal-space cell volume (a.u.), reciprocal lattice constant (a.u.), reciprocal lattice vectors(1:3, 1:3) in units of reciprocal lattice constant, reciprocal-space metric tensor(1:3, 1:3) (a.u.)
- symmetry rotation matrices(1:3, 1:3, 1:number of symmetries) in reciprocal-lattice basis
- symmetry fractional translations(1:3, 1:number of symmetries) in units of lattice vectors times $2\pi$ (see also Sec. 7.6.1)
- atomic positions(1:3, 1:number of atoms) in units of lattice constant, atomic numbers(1:number of atoms)
- number of G-vectors for each k-point(1:number of k-points)
- k-point weights(1:number of k-points) from 0 to 1
- k-point coordinates(1:3, 1: number of k-points) in crystal coordinates
- index of lowest band to use on each k-point (1:number of k-points)
- index of highest occupied band on each k-point (1:number of k-points)
- energy eigenvalues (1:number of bands, 1:number of k-points, 1:number of spins) (Ry)
- occupations (1:number of bands, 1:number of k-points, 1:number of spins) from 0 to 1

In the body of a file, G-vectors are listed as (1:3, 1:ng), expressed as integers in reciprocal lattice units, and data is listed as (1:ng, 1:number of spins). G-vector components should be chosen in the interval \([-n/2,n/2)\) where \(n\) is the FFT grid. A full sphere must be used, not a half sphere as in the Hermitian FFT representation for a real function. Each set is preceded by an integer specifying how many records the G-vectors or data is broken up into, for ease of writing files from a code parallelized over G-vectors. Wavefunction files follow the header with a listing of all the G-vectors; for each k-point, there is first a list of G-vectors, and then the wavefunction coefficients for each band. RHO and VXC files have instead just one listing of G-vectors and coefficients after the header. The wavefunction coefficients must be normalized so that the sum of their squares is 1. The RHO coefficients are normalized such that their \(G = 0\) component is the number of electrons in the unit cell. The VXC coefficients are in Ry.

The recommended scheme is to use pre-computed exchange matrix elements in an ASCII file vxc.dat, because VXC is only applicable to a local exchange-correlation functional. Hartree-Fock, hybrid-density functionals and self-consistent static GW (COHSEX) calculations do not fall in this category. For Hartree-Fock and some hybrid functionals (PBE0, B3LYP), one can still use VXC if one sets bare_exchange_fraction in sigma.inp to compensate for a fraction of the bare exchange which is not included in VXC. Matrix elements are in eV and are always written with real and imaginary parts (even in the real version of the code). The vxc.dat file may contain any number of k-points in any order. It contains a certain number of diagonal elements (ndiag) and a certain number of off-diagonal elements (noffdiag). The x.dat file shares the same format and can supply saved matrix elements of bare exchange instead of calculating them. Each k-point block begins with the line:

\[
kx, ky, kz \text{ [crystal coordinates], ndiag*nspin, noffdiag*nspin}
\]

There are then ndiag*nspin lines of the form

ispin, idia, Re \(\langle idia|V|idia \rangle\), Im \(\langle idia|V|idia \rangle\)

There are then noffdiag*nspin lines of the form

ispin, ioff1, ioff2, Re \(\langle ioff1|V|ioff2 \rangle\), Im \(\langle ioff1|V|ioff2 \rangle\)
Chapter 8

First-principles approach to charge transport in single-molecule junctions with self-energy corrections: a DFT+Σ method

This work is based on Ref. [288].

8.1 Abstract

A new first-principles approach, based on density-functional theory (DFT) and many-body perturbation theory, is developed and applied to quantitatively explain the conductance of single-molecule junctions. Applicable in the limit that the molecule is weakly coupled to the metallic states of the electrodes, this approach considerably improves upon the erroneous Kohn-Sham energy-level alignment in molecular junctions (in standard DFT approaches) via a physically motivated, approximate quasiparticle self-energy correction based on the GW approximation. Whereas DFT calculations within standard approximations tend to place molecular energy levels too close to the electrode Fermi energy and overestimate conductance values relative to experiment, this straightforward correction to the Kohn-Sham energy level alignment predicts quasiparticle energies of the junctions, and conductance values with good accuracy at the same computational expense of a standard DFT transport calculation. A detailed justification of this approach is provided, and its implementation within a Landauer-based scattering-state formalism is discussed. The method is validated by demonstrating quantitative agreement with experiments on a range of different organic molecules in amine- and pyridine-Au linked junctions. The concept developed here is applicable to other DFT-based transport methodologies such as the non-equilibrium Green’s function approach.
8.2 Introduction

Single-molecule junctions between metallic contacts are interesting systems for potential applications as nanoelectronic devices and for investigating foundational notions of charge transport at the nanoscale [254]. The scanning tunneling microscopy break-junction [401] and mechanically-controllable break-junction [327, 96, 185] techniques are now routinely used to provide statistical information on the conductance distributions of single-molecule junctions, while low-temperature scanning tunneling microscopy [238] can be used to image and probe single-molecule junctions at the atomic scale. In its nascent phase, the field of molecular electronics, in particular the study of off-resonant tunneling in single-molecule junctions, was plagued by disagreements between experimental groups [296, 210, 400], between theoretical groups [370, 101], and between theory and experiment. In recent years, however, the search for new anchoring groups [384, 383, 385, 236], such as amines, pyridines and nitriles, has resulted in a growing body of reliable and reproducible conductance data on single-molecule junctions. The challenge then is for theorists to make quantitative comparisons to these experiments, for both validation and new understanding, and prediction and discovery of novel phenomena.

Understanding charge transport quantitatively in single-molecule junctions presents several challenges to theory. The problem of non-equilibrium electron transport through molecular devices—intrinsically open systems, with electrons entering and leaving the device through long leads held at fixed potentials—cannot be solved with standard quantum-chemistry and electronic-structure schemes. The non-equilibrium problem can be solved using established formalisms such as time-dependent density-functional theory (TDDFT) [306] or the Landauer formula within the Keldysh non-equilibrium Green’s function and scattering-state approaches [80]. However, the success of these formalisms depends on a reliable description of the atomic structure and, within the Landauer approach, the quasiparticle spectra in the junction: the resistance originates with electron tunneling events through an effective energy barrier with the length scale of a single molecule. In many molecular junctions, transmission occurs through off-resonant tunneling, i.e. the frontier molecular orbitals (the highest occupied molecular orbital, HOMO, and lowest unoccupied molecular orbital, LUMO) in the junction have energy levels that are significantly different from the Fermi energy. Off-resonant tunneling is particularly sensitive to the size of the energy barrier, and determining this barrier requires precise knowledge of the identities and average positions of all atoms that make up the junction, and a self-consistent description of their bonding and electronic energy-level alignment. This presents a practical challenge to theorists, as reasonable atomic-scale models (with the molecule and the leads) typically include several hundred atoms, making it difficult to perform many-electron calculations required for the quasiparticle energy levels and precluding the study of structure-conductance relationships in single-molecule junctions. The quasiparticle excitations do not explicitly enter in TDDFT; however, as discussed below, there are theoretical issues and accurate approximations to the time-dependent exchange-correlation functional remain unattainable to date. Thus, practical and accurate approximations have yet to be developed for guiding molecular-scale materials design.

At the outset, a promising frameworks for studying non-equilibrium transport may be TDDFT and time-dependent current-density-functional theory (TDCDFT), as these
approaches probe the time development of a system. Even so, it is important to note that the Runge-Gross theorem [306], which is the underpinning of TDDFT, has not been proven for open systems out of equilibrium. To overcome this problem, methods based on related effective problems for which the theorems do apply have been employed, such as systems with large but finite leads [337, 93], periodic systems with coupling to phonons via a master equation [45], and the use of complex absorbing potentials to make leads effectively finite [380]. Pragmatically, approaches for “embedding” a device region in an open system with semi-infinite leads have also been used [197, 253]. However, in addition to these complications, the major drawback of these approaches is that presently available approximations to the exchange-correlation functional do not take into account the non-local nature of the exchange-correlation effect, which is crucial to describing accurate molecular conductance [99], or its memory dependence [197, 253], and adding such effects into the exchange-correlation functional would greatly increase the computational expense.

In fact, the majority of recent theoretical investigations of charge transport in single-molecule junctions rely on a Landauer approach, simplified to treat electronic interactions at a mean-field level within density-functional theory (DFT), using standard, local or semi-local approximations to the exchange-correlation functional. In this standard approach, an effective transport problem is solved through a tripartitioned system consisting of a central resistive region coupled to semi-infinite bulk-like leads [92, 402, 40, 355, 302, 100, 396, 63]. The differential conductance and I-V characteristics are calculated as the response of the Kohn-Sham (KS) system via the Landauer formalism [80], implicitly using the KS energies and wavefunctions as those of the quasiparticles. While this framework turns out to be relatively accurate for resonant-tunneling transport through metallic point contacts and junctions with molecular hydrogen [406, 327, 206, 358, 184], large systematic overestimations of single-molecule junction conductances in the off-resonant-tunneling regime have been widely reported [24, 339, 369, 243], even at very low bias, where the system is close to equilibrium. These observations are consistent with the fact that tunneling probabilities are sensitive to the energy barriers, which cannot be directly obtained from the energies of the KS orbitals of ground-state DFT.

Efforts to improve upon the above standard DFT-Landauer results include using non-local or hybrid functionals [182, 183], self-interaction correction [368], and scissors operators fitted to produce Hartree-Fock molecular gaps [57]. These methods can improve the energy levels, resulting in a reduced conductance, since these schemes in general lower the HOMO level, and in the case of Hartree-Fock, also raise the LUMO level, thus lowering the effective energy barriers that enter the Landauer formalism. However, the Hartree-Fock method in general overestimates significantly the quasiparticle excitation gap, and recent studies [250, 117, 164] showed that existing hybrid functionals do not yield reliable quasiparticle gaps over changing environments. In particular, none of the currently available exchange-correlation functionals, or Hartree-Fock methods, take into account non-local correlation effects from the metal electrodes, which are important for obtaining accurate quasiparticle spectra [250, 31], and are important for describing molecular conductance [99]. Other investigations have shown limitations of the Landauer formula itself [309, 189, 34].

In principle, these non-local exchange and correlation effects should be treated using many-body techniques beyond DFT, which give a proper description of the quasiparticle
levels of the system. Examples are the three-body scattering formalism [105], dynamical mean-field theory [163], numerical renormalization group [212], and the Kadanoff-Baym equations [247]. Additionally, many-body perturbation theory in the $GW$ approximation [141, 152] has been explored in numerous studies [78, 360, 359, 357, 340, 290]. These $GW$ calculations have demonstrated the success of $GW$ in the weakly to moderately correlated regime and showed the importance of many-electron effects both in this regime and in the strong-correlation regime in which beyond-$GW$ effects are significant. However, the major limitation of these techniques is that they have thus far been confined to lattice models or highly simplified junctions with a small number of atoms, due to the calculations’ poor scaling with system size. In practical applications of many-body techniques to transport calculations, correlations can only be introduced in a small central region (in practice, just a few atoms) coupled to uncorrelated leads, or throughout a small finite system (e.g. with configuration interaction [86] or the generator-coordinate method [99]). Because the junction conductance is extremely sensitive to the atomic structure [150, 287], a highly simplified representation of junction structure, used to reduce the computational cost of the many-body techniques, can introduce significant errors in the calculated conductance, hampering quantitative comparison to experiment.

Previous work by collaborators [287, 289, 286, 179] established that, at least in the case of amine-linked molecular junctions, differences in local structure have a limited influence on conductance because the amine-Au bonding motifs are well defined and flexible. The minimal impact of junction structure on conductance for these chemical end groups allowed an unambiguous comparison of calculated and measured conductance values, and established that the conventional static DFT is insufficient for accurately computing conductance and current through molecular junctions, even in equilibrium case for benzenediamine (BDA) (and other molecules) contacted with gold [287, 289, 286, 179]. Taken together, this work led to the development of a new self-energy-corrected first-principles approach, described below, that accounts for non-local exchange and correlation effects in the junction conductance in a practical manner. This self-energy operator represents the difference between the DFT KS and quasiparticle energy levels.

In this approach, self-energy corrections are used based on a $GW$ approximation appropriate in the limit of weakly coupled junction molecular orbitals, including long-range static correlations. This scheme constitutes a minimal set of approximations required to quantitatively predict the correct energy level alignment and conductance of a range of single-molecule junctions, provided a good model of the junction geometry has been determined. The required approximations are: (1) the coupling between the molecule and substrate is weak, in the sense that the molecular resonances have only a small ($\sim 0.5$ eV) broadening in energy, and each has a significant weight from only a single orbital of the isolated molecule, (2) the polarizability of the molecule is reasonably small, and (3) charge transfer is negligible; these are well satisfied for the systems studied. This methodology was applied to various organic molecules in both a simplified method [287, 289] and the present more general formulation [286], for conductance as well as the Seebeck coefficient [285]. These methods have also been employed by other groups [242, 391], as well as the image-charge model in a related form [175]. Direct $GW$ calculations for BDA on Au(111) have also explicitly demonstrated the accuracy of this approximation [351].
In this work, I explain in detail the justification and application of the self-energy corrections, in the form of a parameter-free model self-energy operator, to transport in molecular junctions, and demonstrate the good agreement of this methodology with experimentally measured conductances, for both amine (HOMO-dominated transport) and pyridines (LUMO-dominated transport).

This chapter is organized as follows. In Sec. 8.3 we describe the methodology for inclusion of self-energy corrections in a scattering-states approach to transport, in both a simple and a more rigorous form. The calculation of the model self-energy and the effect on the transmission spectrum are discussed. In Sec. 8.4, I apply the method to some amine and pyridine junctions, showing the effect of self-energy corrections and comparing to experiment. I conclude in Sec. 8.5.

8.3 Methods

8.3.1 Summary of approach

The essence of the approach is the addition of a physically-motivated, parameter-free self-energy operator $\hat{\Sigma}$ to the Hamiltonian for solving the transmission function. This self-energy operator is motivated by $GW$ calculations, and takes into account both exchange and correlation effects within the molecule, and non-local correlation effects due to the metal electrodes. The charge density used as input to this Hamiltonian is obtained using DFT with the PBE [269] exchange-correlation functional. The Hamiltonian is then solved with the additional $\hat{\Sigma}$ term, to provide a more accurate description of the quasiparticle wavefunctions and energy-level alignments in the molecular junction. This approach can be applied to local-orbital scattering-states and non-equilibrium Green’s functions approaches, as implemented in codes such as TranSIESTA [40], SMEAGOL [302], or WanT [48].

This approach has been implemented in the first-principles local-orbital-basis scattering-state code SCARLET [62], which is built on top of the SIESTA [328] code, and incorporates open boundary conditions necessary for computing transmission and conductance. Within this formalism, we consider the molecule plus several lead layers connected to fully equilibrated reservoirs, each with well-defined chemical potentials and temperatures, via semi-infinite leads. The junction is partitioned into three subsystems: the left lead, the right lead, and the central resistive region between them containing the nanoscale junction, as described in detail in Ref. [62]. The central region is chosen large enough for the Hartree potential at its boundaries to smoothly match that of the bulk leads. Energy- and $k_\parallel$-dependent scattering states are constructed with incoming and outgoing Bloch states. Solving the junction Hamiltonian for the scattering states yields the transmission matrix $t$ with elements $t_{nm}(E, k_\parallel)$ for each incoming channel $n$ and outgoing channel $m$. The linear-response conductance is then computed from the Landauer formula $G = G_0 \sum_{k_\parallel} \text{Tr} \left( t^\dagger t \right)$, where $t$ is evaluated at the Fermi energy, $E_F$. 


8.3.2 Inclusion of self-energy corrections $\Sigma$ in scattering-state DFT+$\Sigma$ approach to transport

The details of the self-energy-corrected first-principles approach are described below. To account for non-local and frequency/orbital-dependent exchange and correlation effects in the junction conductance in a practical manner, a model self-energy is used that captures the effects of many-electron interactions, including the intramolecular self-energy corrections to the molecular energy level and long-range static correlations in the leads, on electronic energy-level alignment in the junction. This scheme constitutes a minimal set of approximations required to quantitatively predict the correct level alignment and conductance of a range of single-molecule junctions, provided a good model of the junction geometry has been determined.

Supposing the junction geometry is known, charge transport in molecular junctions is controlled by the electronic coupling of frontier molecular orbitals to extended states in the electrode, and the energetic position of these orbitals relative to the contact Fermi level. These are to be viewed in the sense of quasiparticle excitations in an interacting many-electron system. Understanding how the molecular ionization potential (IP, associated with removing an electron from the HOMO) and electron affinity (EA, associated with adding an electron to the LUMO) are altered when the molecule is bonded in a junction is a prerequisite to understanding molecular conductance. (The ionization potential is the negative of the energy level of the HOMO with respect to the vacuum, and the electron affinity is the negative of the energy level of the LUMO with respect to the vacuum.) The limitations of DFT KS eigenvalues in describing excited-state energies are well known [152], and the LDA/GGA KS HOMO-LUMO gap is significantly underestimated relative to the true quasiparticle gap, especially for small molecules. The difference between the quasiparticle and KS eigenvalues is referred to as the self-energy correction to the KS eigenvalues. While the self-energy corrections are typically small in bulk metals [217, 256, 351, 291], the KS frontier molecular energy levels (HOMO/LUMO levels) in a single-molecule junction are often much closer (by order of magnitude of eV) to the metal $E_F$ than the quasiparticle levels are. As a result, the linear-response conductance (transmission at $E_F$) is frequently overestimated, especially for off-resonant tunneling in weakly coupled molecule/metal systems, where transmission at $E_F$ occurs through the tails of the resonance peaks [182, 189, 368, 57, 100].

Correction of Lorentzian peak positions

The DFT errors in level alignment were previously corrected in a simplified way by shifting the frontier transmission resonance peak away from $E_F$ [287, 289], using a physically motivated parameter-free estimate for the self-energy correction to the resonance position, $\Sigma_R$. The resulting conductance was in excellent agreement with experiment: within $\sim 15 \sim 16\%$ for BDA and bipyridine-Au junctions, in contrast to the DFT conductance that was respectively 7 and 40 times larger. This procedure relied on the fact that the DFT transmission resonance peaks fit well to a Lorentzian function, allowing us to estimate the corrected conductance by applying the correction $\Sigma_R$ to the resonance peak position $E_R$ in
the junction via a transmission function of the form

\[ T(E) = \frac{\Gamma_1 \Gamma_2}{(E - (E_R + \Sigma_R))^2 + \left(\frac{\Gamma_1 + \Gamma_2}{2}\right)^2} \]  

(8.1)

where the conductance \( G = T(E_F)G_0 \). Here, \( G_0 = \frac{2e^2}{h} \) is the quantum of conductance (in the non-spin-polarized case). The transmission peaks fit well to Lorentzian lineshapes when they result from isolated molecular resonances interacting with an approximately uniform density of states (DOS) in the leads, which was the case for BDA and bipyridine-Au junctions, but is not in general true for other systems (such as butanediamine-Au junctions; see Sec. 8.4.3). Furthermore, this procedure neglects changes in coupling widths \( \Gamma_1 \) and \( \Gamma_2 \) that may result when the molecular resonance interacts with different metallic electrode states at the corrected energy alignment. In spite of the limitation, this method is successful for BDA and bipyridine-Au junctions.

### Correction of scattering-state Hamiltonian

Here I describe a mathematically well-grounded extension of the Lorentzian approach described above in which an approximate self-energy operator is explicitly added to the scattering-state Hamiltonian \( \hat{H} \). This operator takes the form

\[ \hat{\Sigma} = \sum_n \Sigma_n \left| \psi_n^{\text{mol}} \right> \left< \psi_n^{\text{mol}} \right| \]  

(8.2)

where \( \left| \psi_n^{\text{mol}} \right> \) denotes a molecular orbital and \( \{\Sigma_n\} \) are the self-energy corrections to the DFT eigenvalues for each molecular orbital.

The new scattering-state Hamiltonian, \( \hat{H} + \hat{\Sigma} \), is then used to obtain a new set of scattering states and an associated transmission function in a “one-shot” calculation using the DFT charge density as input. This approach is named DFT+\( \Sigma \). The corrected scattering states properly account for the hybridization of the molecular resonances with the metallic states at the new energy alignment.

#### 8.3.3 Form of the self-energy-correction operator \( \hat{\Sigma} \)

The self-energy correction employed here includes only the real part of the self-energy which corrects the quasiparticle levels; the imaginary part of the self-energy, which is inversely proportional to the quasiparticle lifetime, is neglected, as it is small for energies close to \( E_F \) (and rigorously zero at \( E_F \) in the zero-bias limit), and therefore does not change the zero-bias conductance [106] (at zero temperature).

Ideally, we would obtain \( \hat{\Sigma} \) from a full GW calculation, but even with recent improvements in parallel efficiency of GW codes [88] (Chapter 7), a GW calculation of realistic junction geometries of hundreds of transition-metal atoms is challenging. Therefore we seek an approximate form of the operator, appropriate to the systems of study. In typical GW calculations, quasiparticle energies are obtained as first-order corrections to DFT eigenvalues from the expectation value of \( \hat{\Sigma} \) with the KS orbitals. This assumes the DFT orbitals are good approximations to the quasiparticle wavefunctions [152]. However, in a multi-component system such as a metal-molecule junction, there exist localized molecular states
and delocalized metallic states, and there can be significant rehybridization between the different parts of the system due to their differing self-energy corrections [318]. In metal-molecule junctions with negligible self-energy corrections to the metal and weak molecular orbital-metallic state coupling, the effect is large since the molecular orbitals can rehybridize differently with metal states at their corrected energy. As a result, a full $\hat{\Sigma}$ operator must be constructed, added to $H$, and diagonalized to obtain new wavefunctions and eigenvalues.

Starting from a basis of localized pseudo-atomic orbitals, we can construct an orthonormal basis $\{|\psi_n\rangle\}$ by diagonalizing the DFT Hamiltonian in the molecular subspace, and then orthogonalizing the rest of the orbitals to the molecular subspace. We then express the operator $\hat{\Sigma}$ in this basis:

$$
\hat{\Sigma} = \sum_{mn} |\psi_m\rangle \langle \psi_m| \hat{\Sigma} |\psi_n\rangle \langle \psi_n|.
$$

(8.3)

We now consider which of the matrix elements $\langle \psi_m| \hat{\Sigma} |\psi_n\rangle$ are significant. Since self-energy corrections for metallic bulk and surface states are in general much smaller than those for the molecule, we neglect matrix elements in the non-molecular subspace [105]. (Note that the matrix elements of isolated lead states may still be nonzero if they have some overlap with molecular orbitals.) We further assume weak coupling between the molecule and lead, meaning that a wavefunction of the combined system has significant weight only from a single state of the isolated molecule, and a molecular state has only a small ($\sim 0.5$ eV) broadening in energy due to the metal. This form implies weak overlap between leads and molecule. If this form holds for both the DFT and quasiparticle wavefunctions, then there must be only small exchange-correlation coupling between the molecule and lead states, allowing us to neglect the self-energy matrix elements between the molecular and non-molecular orbitals. Generally, we may expect that metallic screening will make the self-energy operator fairly local, so that matrix elements between spatially separated orbitals will be small. This assumption is reasonable for amine-Au and pyridine-Au systems, particularly for the junctions considered here in which the molecules are bonded close to perpendicular to the surfaces, keeping the influence of the leads localized close to the contact atoms. The calculations show that the wavefunctions of the junctions indeed closely resemble molecular wavefunctions combined with metal states, demonstrating weak coupling (Fig. 8.4). Only molecule-molecule matrix elements now remain in $\hat{\Sigma}$. Lastly, we assume that the molecular wavefunctions are similar at the DFT and quasiparticle levels of theory, so that $\hat{\Sigma}$ is diagonal in the basis of DFT orbitals. This diagonal approximation has been found to be good in bulk and molecular systems if the energy levels are below the vacuum level [152, 130].

With these approximations, our final form for the self-energy operator is

$$
\hat{\Sigma} \approx \sum_{n}^{\text{mol}} |\psi_n\rangle \langle \psi_n| \hat{\Sigma} |\psi_n\rangle \langle \psi_n| = \sum_{n}^{\text{mol}} \Sigma_n |\psi_n\rangle \langle \psi_n|.
$$

(8.4)

The problem is reduced to finding expectation values of the self-energy correction for each molecular orbital. In principle, they may depend on $k$, but for molecular states in isolated or widely separated junctions, band dispersion would be unphysical so we only use values at the $\Gamma$-point.
Table 8.1: Eigenvalues from molecule in isolation and in junction, compared with transmission peaks. All values are in eV, referenced relative to the Fermi level of the Au leads. (For the isolated molecule, the reference is obtained by first aligning the vacuum levels in the calculations for the isolated molecule and for a 12-layer Au slab, and second, aligning the potential in the middle of the Au slab with the potential in the calculation for the Au leads.) The junction eigenvalue comes from a non-self-consistent diagonalization of the molecular subspace of the DFT Hamiltonian for the converged charge density of junction. Except for butanediamine, where the identification of the DFT transmission peak center for butanediamine is uncertain due to the complicated lineshape, the transmission peak positions are within $\sim 0.2$ eV of the junction eigenvalues, and shift approximately by the self-energy correction of the molecular orbital. The difference between the eigenvalues for the molecule in isolation and in the junction can be identified as the “lead self-energy” (as defined in the text), and represent the effect of charge rearrangement when the molecule is bound to Au. BDA, DP-ethene, DP-benzene data courtesy S. Y. Quek [288].
of a single-molecule junction where the molecule is the “resistive region” coupling to the “leads” (which in this simple model, includes the surface metallic layers and binding sites), the differences between the eigenvalues of the isolated molecule and the junction molecular eigenvalues are an approximation to the DFT value of the real part of the “lead self-energy,” i.e. the shift of the molecular energy levels resulting from the molecule coupling to the metal. This concept is commonly used in Green’s function-based transport calculations, particularly in empirical calculations [80]. In this language, we may rephrase the molecule orbital-lead weak-coupling condition from earlier in this subsection this way: the lead self-energy matrices are close to real and diagonal in the basis of orbitals of the isolated molecule, so the molecular states are only slightly mixed or broadened.) Interestingly, the DFT lead self-energy for butanediamine is considerably larger than the aromatic molecules’, due to the fact that the middle of the gas-phase HOMO-LUMO gap is the highest above $E_F$ of the molecules we considered, and thus it has the largest interface dipole.

8.3.4 Evaluation of self-energy corrections $\Sigma$, including electrode polarization effects

An important aspect of this approach is to obtain reasonable estimates for $\Sigma$ in the junction, including the effect of the electrodes. It has been demonstrated, both experimentally [299] and through many-electron $GW$ calculations [250], that the self-energy corrections to level alignment at a molecule-metal interface are strongly environment-dependent. In particular, strong static polarization effects—nonlocal electrostatic correlations between an added electron or hole in the molecule and the metal surface—shift (or renormalize) the frontier orbitals closer to the electrode Fermi level by a large amount ($\sim 1$ eV for benzene on graphite [250]), substantially narrowing the quasi-particle energy gap. This “image charge” effect, which is completely missed by DFT calculations [250], must be accounted for explicitly in computing corrections to junction level alignment.

Crucially, the results described above, computed rigorously within the $GW$ approximation, can be modeled accurately and efficiently with a physically motivated approximate self-energy correction, $\Sigma^0 + \Delta \Sigma$, which has two contributions: a bare gas-phase correction $\Sigma^0$ and an image-charge term $\Delta \Sigma$ (Fig. 8.1) [250]. We expect this model to apply generally to molecule-substrate systems in the limit where (1) the coupling between the molecule and substrate is weak, (2) the polarizability of the molecule is small enough to neglect the internal screening response of the molecule to the polarization of the metal surface, and (3) charge transfer is negligible [250], as supported by other recent studies [117, 361]. In Sec. 8.6.1 I derive the image-charge form of the self-energy operator for weakly coupled systems satisfying these criteria. (In the case that the molecular resonances overlap the metal Fermi energy, additional dynamical screening effects neglected here become important [250, 311]).

The validity can be checked for the molecules studied here: (1) The metal does not mix molecular orbitals significantly, since the scattering wavefunctions strongly resemble the expected molecular orbital (Fig. 8.4). The broadening of molecular levels is seen from the transmission spectra (Fig. 8.5) to be $\sim 0.5$ eV, and specifically we will see in Sec. 8.6.1 that it should be much less than the plasmon energy ($\sim 10$ eV) of the metal. (2) As long as the molecule has non-metallic polarizability, its screening response to the metal can be neglected; organic molecules with $\epsilon_\infty \sim 2$ (in liquid form) easily meet this. (3) The
Figure 8.1: Diagram of model for self-energy correction to molecular levels in a weakly coupled junction within the DFT+Σ method for transport computation. The correction has two parts: a gas-phase term, equal to the difference between Kohn-Sham eigenvalues and quasiparticle levels in the gas phase, which in general opens the HOMO-LUMO gap; and an electrode polarization term, due to metallic screening by the leads of added charge on the molecule, which reduces the gap.
transmission spectra at the DFT level (Fig. 8.5) as a proxy for the DFT density of states show that the peaks only slightly overlap the Fermi level, clearly seen quantitatively from the fact that the transmission is much less than unity. Thus the needed criteria are met for this set of molecules.

For corrections to the HOMO, \( \Sigma^0 \) is obtained by taking the difference between the gas-phase ionization potential (IP) and the DFT KS HOMO eigenvalue, and analogously for the LUMO with the electron affinity (EA). The calculations use an isolated molecule in the geometry from the relaxed junction. IP and EA were computed using the \( \Delta \)SCF method [171], employing GAUSSIAN03 with the B3LYP exchange-correlation functional and 6-311g++dp basis set [114]. This procedure gives IP values within 0.1 eV of experiment [208] for these molecules.

The image-charge term \( \Delta \Sigma \) is computed using an electrostatic model that assumes the two parallel electrodes are perfect conductors and takes the image planes to be 1 Å from the metal surfaces [250, 199]. For convenience, the charge distribution of the added electron (LUMO) or hole (HOMO) to the molecule is modeled by partitioning the charge density corresponding to the gas-phase HOMO/LUMO into contributions from each atom via Mulliken charges [244]. This approach is formally justified in the approximation that the KS orbitals are close to the quasiparticle wavefunctions [152]. These partial charges for the HOMO/LUMO represent the charge loss/gain \( q_j \) at each atomic site \( r_j \) in the molecule when an electron is removed/added. The interaction energy of this charge distribution with its image charges is then given by

\[
\Delta \Sigma = -\frac{1}{2} \sum_{i,j=1}^{N} q_i \phi(q_j, r_j, r_i) \tag{8.5}
\]

where \( N \) is the number of atoms in the molecule and \( \phi(q_j, r_j, r_i) \) is the potential at \( r_i \) due to the infinite set of image charges of \( q_j \) at \( r_j \). The full expression for this potential is shown in Sec. 8.6.2. In practice, for the junctions considered here, the image-charge energy calculated from Mulliken distributions is within 0.1 eV of that calculated from representing the added electron/hole as a single point charge in the middle of the molecule. However, the single-point-charge model may not be adequate for longer molecules where the charge is distributed over a larger distance, or for asymmetric molecules. Since the conductance in amine-Au and pyridine-Au junctions is dominated by the frontier molecular level (HOMO [289] and LUMO [287] respectively), \( \Sigma_{\text{HOMO}} \) is applied to all occupied orbitals, and \( \Sigma_{\text{LUMO}} \) to all unoccupied orbitals. (In principle, the self-energy corrections for other orbitals could also be estimated from constrained DFT for the gas-phase term, or via GW calculations.)

### 8.3.5 Effect of \( \hat{\Sigma} \) on the molecular resonance positions and transmission lineshapes

Here arguments are provided for the magnitude of the shift in the resonance position for a given \( \Sigma_n \). In the weak-coupling regime, \( \hat{\Sigma} \) simply shifts each molecular resonance position by \( \sim \Sigma_n \). To see this, we can write the DFT scattering-state Hamiltonian \( \hat{H} \) in a basis of the “junction molecular orbitals,” \( \{ \psi_m^{\text{mol}} \} \), plus eigenstates of the leads. Assuming that the coupling between the molecule and lead is small, the molecular resonances would be well
approximated by eigenstates of $\hat{H}$ that consist of \(\{\psi_n^{\text{mol}}\}\) hybridizing with lead states of the same energy. The molecular resonances would therefore have energies $\sim \epsilon_n^{\text{mol}}$. Likewise, for the Hamiltonian $\hat{H} + \Sigma$, eigenstates that are molecular resonances would have energies $\sim \epsilon_n^{\text{mol}} + \Sigma_n$.

If the lead states have a uniform orbital character and DOS over the full energy range including $E_F$ and the molecular resonance energies before and after applying the self-energy correction, \textit{i.e.} when the Lorentzian approximation is appropriate, then the effect on the transmission within this energy range is equivalent to a rigid shift of the peak. To a large extent, this was the case for bipyridine-Au junctions [179]. However, if the lead states are significantly different at the new energy, the hybridization of the molecular states with the lead states will be different, and the lineshape of the peak will be altered. This effect is especially pronounced when the molecular states hybridize with Au $s$ states in the DFT calculation, but are shifted to hybridize with Au $d$ states in the DFT+$\Sigma$ calculation, such as in the BDA-Au junctions [286] (Sec. 8.4.3).

### 8.3.6 Details of ground-state and transmission calculations

The DFT calculations are performed within the GGA-PBE [269] approximation to the exchange-correlation functional, as implemented in SIESTA [328]. An optimized single-$\zeta$ basis set is used for the Au $d$ shell; all other orbitals are described by double-$\zeta$ with polarization (DZP) basis sets.

Since a Au-Au contact is formed and broken at the beginning of each break-junction measurement, under-coordinated Au atoms are likely to be present as binding sites for amine and pyridine molecules. To model likely junction geometries, the binding of amines and pyridines to under-coordinated Au motifs (such as adatoms and trimers) on Au(111) is considered.

The junctions are constructed with six-layer Au(111) slabs on either side of the molecule, with translational symmetry imposed along directions parallel to the Au surface. The supercells have 16 Au atoms per layer. All atoms in the junction are relaxed until forces on each atom are 0.05 eV/Å, except for those in the bottom 3 Au layers of each slab, which are fixed to the bulk geometry. The distance between slabs is also allowed to relax.

For transmission calculations, an energy-grid spacing of 10 meV is used in this work. Care is taken to converge the transmission with respect to the $k_{\parallel}$-mesh that samples the two-dimensional Brillouin zone. A $4 \times 4$ Monkhorst-Pack $k_{\parallel}$-grid which includes the $\Gamma$-point is found to be sufficient when the peaks of interest are higher in energy than that of the $d$-bands of the Au leads, whereas an $8 \times 8$ $k_{\parallel}$-grid is required to describe peaks in the energy range of the $d$-bands ($\lesssim E_F - 1.5$ eV). The DFT transmission for benzenediamine-Au and pyridine-Au junctions, and the DFT+$\Sigma$ transmission for pyridine-Au junctions are converged with a $4 \times 4$ grid; an $8 \times 8$ grid is required for butanediamine (DFT and DFT+$\Sigma$) and for the DFT+$\Sigma$ transmission in benzenediamine-Au junctions.
8.4 Application to amine-Au and pyridine-Au junctions

This DFT+$\Sigma$ approach was applied to a set of seven different amine-Au and pyridine-Au junctions. The self-energy corrections not only bring the conductance of all systems into better quantitative agreement with experiment, but also are important in correcting qualitative conductance trends across different systems. It is demonstrated how this approach is applicable for non-Lorentzian DFT transmission lineshapes and can generate more complicated structure than simply a shifted peak.

The amine molecules studied in this work are: BDA, 2,3,5,6-tetrafluoro-1,4-benzenediamine (BDA4F), 2,3,5,6-tetramethyl-1,4-benzenediamine (BDA4Me), and 1,4-butenediamine; the pyridine molecules studied are: 4,4′-bipyridine, 1,2-di-(4-pyridyl)-ethene (DP-ethene), and 1,4-di-(4-pyridyl)-benzene (DP-benzene). In each case, a trimer junction geometry is used with the molecule oriented parallel to the surface normal (Fig. 8.2), a representative junction structure chosen based on previous extensive studies of the binding configurations in BDA-Au [289] and bipyridine-Au [287] junctions and their impact on conductance. The calculations of BDA, bipyridine, DP-ethene, and DP-benzene were performed by S. Y. Quek [288].

8.4.1 Comparison of calculated conductance values to experiment

Table 8.2 and Fig. 8.3 show the DFT (with the PBE exchange-correlation functional) and DFT+$\Sigma$ conductance computed for all seven systems and their comparison to experiment. The DFT calculations all overestimate the conductance, by 4-8 times for the amines, and by 3-130 times for the pyridines. These large overestimates stem from the fact that the frontier molecular states that dominate the linear-response conductance (HOMO for amines and LUMO for pyridines) are too close to $E_F$. The self-energy corrections to these states are substantial ($-1$ to $-2$ eV for the amine HOMOs and $+1.3$ eV for the pyridine LUMOs, as shown in Table 8.2), moving these levels away from $E_F$ to their quasiparticle energy levels, and resulting in much better agreement with experiment (all within the experimental spread except for butenediamine, where DFT+$\Sigma$ underestimates the conductance). For comparison, conductance values are also computed corresponding to self-energy corrections whose magnitudes were increased/decreased by 0.2 eV (Table 8.2). The resulting conductance values are all in reasonable agreement with experiment, indicating that errors of this magnitude in the self-energy estimate do not significantly impact the accuracy of the DFT+$\Sigma$ conductance in these systems.

Finally, estimates of the self-energy-corrected conductance based on Lorentzian fits, as previously described [287, 289], are listed for all systems except butenediamine, where the tail of the DFT transmission resonance peak does not fit well to a Lorentzian function due to hybridization with the Au $d$ band. The Lorentzian shift estimates actually work remarkably well given the simplicity of the approach. In practice, for the systems we have studied, that the Lorentzian shift estimate is reasonable when the tail of the frontier resonance peak in the transmission function can be fitted well to a Lorentzian curve. However, this approach assumes a single resonance interacting with a constant density of states in the leads [80], which is not true in general, and is thus not expected to work for all systems.

The results also show that DFT-PBE predicts wrongly that the conductance of bi-
Figure 8.2: Set of 7 molecular junctions studied in this work: [a] 1,4-benzenediamine (BDA), [b] 2,3,5,6-tetrafluoro-1,4-benzenediamine (BDA4F), [c] 2,3,5,6-tetramethyl-1,4-benzenediamine (BDA4Me), [d] 1,4-butanediamine, [e] 4,4′-bipyridine, [f] 1,2-di-(4-pyridyl)-ethene (DP-ethene), and [g] 1,4-di-(4-pyridyl)-benzene (DP-benzene). The amine molecules (a-d) have HOMO-dominated conductance, while the pyridine molecules (e-g) have LUMO-dominated conductance. Butanediamine (d) is a saturated alkane, while the others are π-conjugated molecules. Each molecule is bonded to an adatom trimer on a Au(111) face of each lead. [a,e,f,g] courtesy S. Y. Quek [288].
Figure 8.3: Comparison of conductances calculated by DFT (black circles) and DFT+Σ (red triangles) transport methods and experimental measurements, on log-log scale. DFT values are all overestimates, but the DFT+Σ approach lowers conductances into agreement with experiment. Note that the trend between bipyridine and DP-ethene (e and f, circled) is wrong in DFT, but corrected by DFT+Σ. Error bars are peak widths in experimental histograms. Numerical values and experimental references are given in Table 8.2. [a,e,f,g] courtesy S. Y. Quek [288].
<table>
<thead>
<tr>
<th>Junction</th>
<th>Conductances $G (G_0)$</th>
<th>$\Sigma$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DFT</td>
<td>Lorentzian</td>
</tr>
<tr>
<td>BDA</td>
<td>$5.03 \times 10^{-2}$</td>
<td>$8.11 \times 10^{-3}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDA4F</td>
<td>$3.09 \times 10^{-2}$</td>
<td>$4.53 \times 10^{-3}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDA4Me</td>
<td>$6.77 \times 10^{-2}$</td>
<td>$1.03 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>butanediameine</td>
<td>$3.34 \times 10^{-3}$</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bipyridine</td>
<td>$2.55 \times 10^{-3}$</td>
<td>$1.08 \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP-ethene</td>
<td>$5.93 \times 10^{-3}$</td>
<td>$7.92 \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP-benzene</td>
<td>$3.15 \times 10^{-4}$</td>
<td>$4.15 \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8.2: Comparison between conductance values computed with DFT, Lorentzian estimate of DFT+$\Sigma$, and full DFT+$\Sigma$ methodology, and experiment. Self-energy corrections $\Sigma$ and their two parts, the gas-phase correction $\Sigma^0$ and the image-charge term $\Delta \Sigma$, are shown. Computed values are reported for different molecules in the trimer junction structure. Values in round (square) brackets denote conductance values obtained when the magnitude of $\Sigma$ is artificially reduced (increased) by 0.2 eV. Experimental error bars correspond to the spread in measured conductance values as quantified by the full-width-half-maximum of peaks in the conductance histograms [289]. BDA, DP-ethene, DP-benzene data courtesy S. Y. Quek [288].
pyridine-Au junctions is about 40% smaller than that of DP-ethene-Au junctions, while the DFT+\Sigma approach correctly predicts that bipyridine-Au junctions in fact have larger conductance than DP-ethene-Au junctions (Table 8.2, Fig. 8.3, circled data points).

8.4.2 Origin of transmission at $E_F$

To analyze the character of the wavefunctions dominating conductance at $E_F$, the transmission matrix $T = t^\dagger t$, from the DFT+\Sigma calculation, is diagonalized to obtain the eigenchannels at each $k_\parallel$-point [39]. In principle, each eigenchannel at each $k_\parallel$-point can contribute to the total transmission at $E_F$; the sum of these contributions gives the total conductance. In practice only 1-2 eigenchannels at each $k_\parallel$-point make non-negligible contributions for the junctions studied here. Specifically, for the amine-Au junctions, a single eigenchannel contributes > 99% of the conductance at each $k_\parallel$-point, while for the pyridine-Au junctions, two eigenchannels contribute to the conductance at each $k_\parallel$-point. By visualizing the wavefunctions for all eigenchannels with significant transmission at $E_F$ “conducting eigenchannels,” it is seen that the single eigenchannel in the amine-Au junctions derives from the molecular HOMO in each case. Fig. 8.4 shows typical “conducting eigenchannel” wavefunctions for each junction, together with the corresponding gas-phase molecular wavefunctions.

8.4.3 Calculated transmission spectra

The DFT and DFT+\Sigma transmission spectra are plotted in Fig. 8.5 for all systems studied. Focusing first on the BDA spectra (Fig. 8.5a), we see that the DFT transmission (gray curve) in the energy range shown is dominated by a single peak centered at $-1.2$ eV with respect to the Au Fermi level. The transmission at the peak center is unity, corresponding to transport through the DFT HOMO resonance. Applying the self-energy correction $\Sigma$ shifts the HOMO resonance to $-3.4$ eV (black curve). In addition, the DFT+\Sigma transmission spectrum has an additional peak centered at $-2$ eV. The projected density of states on the under-coordinated Au contact atoms in a junction with the molecule removed reveals a higher density of d-states in this energy range, indicating that this peak derives from the HOMO hybridizing with a higher density of d-states of the under-coordinated Au contact atoms. Such features are a direct result of the “one-shot” DFT+\Sigma approach in which the scattering states are recomputed after the self-energy corrections are applied, and would not be present in the simple Lorentzian shift estimate previously used [287, 289]. These features are present in both the trimer and adatom junctions, but their energies may be different for other contact motifs not explored here.

The lineshapes of peaks for BDA4Me and BDA4F (Fig. 8.5a) are similar to those of BDA, since the gas-phase HOMO wavefunctions of these molecules are very similar to that of BDA (Fig. 8.4) but their energies are different. The effect of the four methyl (fluoro) groups is to donate (withdraw) electrons to the phenyl ring and thus raise (lower) the energy levels with respect to the Au Fermi energy. This effect is present in the DFT spectra, although it is smaller in the junction than in the gas phase, as previously reported for singly-substituted molecules [242]. However, for these tetra-substituted molecules, the effect of the electron donating/withdrawing groups is enhanced in the DFT+\Sigma spectra as
Figure 8.4: Typical DFT+$\Sigma$ scattering-state wavefunctions at the Fermi energy and molecular orbitals dominating conductance for each molecular junction. The scattering-state wavefunctions closely resemble the orbitals of the isolated molecules, showing that molecular orbitals in these junctions are in the weak-coupling regime. Direction of electron flow is left to right. Isosurfaces for scattering-state wavefunctions are plotted at 2.5% of maximum and for molecular orbitals at 10%. [a,e,f,g] courtesy S. Y. Quek [288].
Figure 8.5: Transmission spectra of (a) benzenediamines, (b) butanediamine, and (c) bipyridines within DFT and DFT+Σ approaches. The self-energy correction reduces transmission at the Fermi energy and shifts peaks away from the Fermi energy. In (a), substitution with fluoro (methyl) groups lowers (raises) the HOMO peak energy. Note the complex lineshape for butanediamine in (b). The peak around $-2 \text{ eV}$ in (a) and (b) is due to hybridization between frontier orbitals and Au adatom $d$-states. BDA, bipyridine, DP-ethene, DP-benzene data courtesy S. Y. Quek [288].
compared to that in the DFT spectra, because of differences in the gas-phase self-energy correction $\Sigma^0$ for the three different molecules. The additional peak around $-2$ eV is present in all the BDA derivatives. In contrast to the HOMO resonances, the energy of this peak does not shift significantly with different functionalizations of the molecule, thus providing further evidence that this feature in the transmission results from the HOMO hybridizing with a higher density of Au $d$-states, rather than from a distinct property of the molecule. The transmission peak height for this feature is larger when it is closer to the HOMO resonance, which is also consistent with the fact that transmission in this energy range derives from the HOMO interacting with the Au states. Finally, although the binding energy is slightly larger for BDA4Me and smaller for BDA4F \[87\], the transmission peak widths are similar across the BDA derivatives.

In contrast to the BDA derivatives, butanediamine has no $\pi$-orbitals, resulting in a much larger HOMO-LUMO gap and conductance smaller by an order of magnitude. The DFT transmission shows a peak at $-2$ eV relating again to the under-coordinated Au $d$ states, and a broad resonance in the energy range $-5$ eV to $-3$ eV. The non-Lorentzian line-shape of this resonance arises from the HOMO (a $\sigma$-state) hybridizing with the Au $d$-band. The HOMO is still the dominant contributor to the conducting state, and for butanediamine it is a $\sigma$-state, unlike the frontier $\pi$-orbitals of the conjugated aromatic systems. DFT+$\Sigma$ shifts the HOMO resonance to $-6$ eV. Conducting eigenchannels across the whole energy range are still dominantly HOMO in character. The calculated conductance is just outside the experimental width, but it is remarkable that this methodology is able to make a fairly accurate determination of the conductance at the Fermi energy via corrections to levels located more than 4 eV away.

Each LUMO transmission peak fits well to a Lorentzian form, consistent with electronic conduction through a single resonance hybridizing with a nearly constant density of states in the unoccupied Au $s$-$p$ band. Comparing the DFT+$\Sigma$ and DFT transmission functions, one sees the transmission resonances are moved away from $E_F$ in DFT+$\Sigma$, but the shape of the transmission functions remains essentially the same, except for a slight broadening of the peak widths. These calculations thus justify the use of the Lorentzian shift estimates [287] for these systems.

8.5 Conclusion

I have presented a new methodology (called DFT+$\Sigma$) for computing electronic transmission through molecular junctions. Its success was demonstrated in calculating the conductance for junctions with a range of model geometries, linking groups (amines and pyridine), molecular structures (conjugated and unconjugated), and side groups, with both Lorentzian and more complicated transmission peaks and HOMO- and LUMO-dominated transport. The needed approximations of weak coupling between the molecule and leads, small molecular polarizability, and negligible charge transfer are often satisfied for molecular junctions. In these results DFT not only overestimates conductances, but also sometimes incorrectly describes conductance trends, while DFT+$\Sigma$ corrects these deficiencies and brings the results into quantitative agreement with experiment. In contrast to other many-electron transport calculations [99, 86, 106, 78, 360, 163, 212, 247, 359, 333, 340, 290], this approach
is no more computationally intensive than standard DFT-PBE transport calculations, thus allowing for detailed exploration of the effect of different functionalizations and junction geometries.

8.6 Appendix

8.6.1 Image-charge form of the self-energy correction in weakly coupled molecule-metal systems

Ref. [250] shows that the change in \( \Delta \Sigma \), for a benzene frontier orbital energy upon adsorption on a metallic (graphite) surface can be expressed almost entirely as the static polarization energy associated with the nonlocal interaction between an electron in that orbital and a perfect metallic conductor. This result applies generally to metal-molecule interfaces for which the molecular orbitals are well localized on the molecule. In what follows, I illustrate with a derivation that the argument also holds for a small but non-vanishing molecule-metal hybridization, proceeding along the lines of the analysis of core states by Hedin and Lundqvist [141], and without making a static approximation. We consider a system with total Hamiltonian \( \hat{H} \) partitioned into \( \hat{H}_{\text{mol}} + \hat{H}_{\text{metal}} + \hat{H}_{\text{mol-metal}} \), where \( \hat{H}_{\text{mol}} \) is the Hamiltonian of the isolated molecule, \( \hat{H}_{\text{metal}} \) is that of the isolated metal surface, and \( \hat{H}_{\text{mol-metal}} \) embodies their interaction. The result is valid when three criteria are satisfied: (1) \( \hat{H}_{\text{mol-metal}} \) is quantitatively small and almost, though not quite, zero (“weak coupling”); (2) the molecule’s polarizability is small enough to neglect the internal screening response of the molecule to the polarization from the metal surface; and (3) charge transfer between the molecule and metal is negligible.

The derivation proceeds as follows. The self-energy in the \( GW \) approximation, following Hedin and Lundqvist equation (36.6) and infinitesimals as in (9.4), is

\[
\Sigma (\mathbf{r}, \mathbf{r}^{'}, \omega) = \frac{i}{2\pi} \sum_n \int_{-\infty}^{\infty} \psi_n (\mathbf{r}) \psi_n (\mathbf{r}^{'})^* \frac{W (\mathbf{r}, \mathbf{r}^{'}, \omega')}{\omega' + \omega - \epsilon_n - i\delta_n} d\omega',
\]

(8.6)

where \( \psi_n \) are the KS orbitals of the full system with eigenvalues \( \epsilon_n \) (and the index \( n \) includes \( k_{\parallel} \)-points, although we take the molecular orbitals as having no dispersion with \( k \)), and \( W \) is the screened Coulomb interaction. The infinitesimals \( \delta_n \) are positive for states below the Fermi energy and negative for states above the Fermi energy, following from the underlying time-ordered single-particle Green’s functions.

When the molecule and metal are very far apart, they do not overlap, and the self-energy of the molecule is due only to interactions within itself, i.e. the intramolecular self-energy \( \Sigma^0 \). Upon adsorption, \( \Sigma \) can change due to changes in \( \delta_n, \psi_n, \epsilon_n, \) or \( W \). We rule out changes in occupations (and hence \( \delta_n \)) due to the assumed lack of charge transfer (criterion 3). The weak coupling (criterion 1) of the molecule to the leads does not significantly change the KS orbitals \( \psi_n \), so we will neglect this effect. The eigenvalues \( \epsilon_n \) can change (as shown in Table 8.1), but this effect is mostly a constant shift, and the change in their differences (as appearing in the denominator) are much smaller. Moreover, within the static COHSEX approximation to \( GW \) which accounts for a large part of the self-energy, the eigenvalues are not involved at all [141, 152]. Thus the effect of eigenvalue changes on \( \Sigma \) is small. (The
neglect of wavefunction and eigenvalue changes might still be acceptable even if a few of the orbitals couple to the leads.)

However, even with weak coupling, the screened interaction \( W \) will be modified upon adsorption. The resulting change in \( GW \) self-energy is given by

\[
\Delta \Sigma (\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \sum_n \int_{-\infty}^{\infty} \psi_n(\mathbf{r}) \psi_n(\mathbf{r}')^* \frac{\Delta W(\mathbf{r}, \mathbf{r}', \omega')}{\omega' + \omega - \epsilon_n - i\delta_n} d\omega',
\]

(8.7)

where \( \Delta W \) is the change in the screened Coulomb interaction. We express the wavefunctions of the full system as linear combinations of orthonormal orbitals \( \{ \phi_{m}^{\text{mol}}, \phi_{p}^{\text{metal}} \} \) constructed similarly to Sec. 8.3.3 by taking eigenstates of the isolated molecule and then orthogonalizing the orbitals of the metallic leads to them:

\[
\psi_n = \sum_m \alpha_{nm} \phi_{m}^{\text{mol}} + \sum_p \alpha_{np} \phi_{p}^{\text{metal}}.
\]

(8.8)

For “weak coupling” (criterion 1), the lead states will not be too different from the isolated lead eigenstates.

The change in self-energy \( \Delta \Sigma \) for a molecular orbital energy in the presence of a metal surface can then be expressed as

\[
\left\langle \phi_{m}^{\text{mol}} \right| \Delta \Sigma (\omega) \left| \phi_{m}^{\text{mol}} \right\rangle = \frac{i}{2\pi} \sum_n \int d\mathbf{r}' \phi_{m}^{\text{mol}}(\mathbf{r})^* \phi_{m}^{\text{mol}}(\mathbf{r}') \times \int_{-\infty}^{\infty} \psi_n(\mathbf{r}) \psi_n(\mathbf{r}')^* \frac{\Delta W(\mathbf{r}, \mathbf{r}', \omega')}{\omega' + \omega - \epsilon_n - i\delta_n} d\omega'.
\]

(8.9)

\( \Delta W \) (in the image-charge form) varies on the scale of the adsorption height, while \( \psi_n \) varies on the scale of a bond length. Thus unless all of the molecule is within a bond length of the surface (incompatible with weak coupling), \( \psi_n \) will vary much faster than \( \Delta W \), so \( \Delta W \) can be treated as effectively constant in the spatial integral. Then terms involving \( \phi_{m}^{\text{mol}} \) in the expansion of \( \psi_n \) for \( m \neq n \) are negligible due to orthogonality, or

\[
\Delta W (\omega') \alpha_{nm} \alpha_{nm}' \int \phi_{m}^{\text{mol}}(\mathbf{r})^* \phi_{m}^{\text{mol}}(\mathbf{r}) \int \phi_{n}^{\text{mol}}(\mathbf{r}') \phi_{n}^{\text{mol}}(\mathbf{r}')^* d\mathbf{r}' \approx 0.
\]

(8.10)

Thus only “self-terms”, i.e. those with \( m' = m \), contribute to \( \Delta \Sigma \), simplifying the expression to

\[
\left\langle \phi_{m}^{\text{mol}} \right| \Delta \Sigma (\omega) \left| \phi_{m}^{\text{mol}} \right\rangle \approx \frac{i}{2\pi} \sum_n \alpha_{nm} \alpha_{nm}' \times \int d\mathbf{r}' \phi_{m}^{\text{mol}}(\mathbf{r})^* \phi_{m}^{\text{mol}}(\mathbf{r}') \phi_{m}^{\text{mol}}(\mathbf{r}')^* \phi_{m}^{\text{mol}}(\mathbf{r}) \int_{-\infty}^{\infty} \frac{\Delta W(\mathbf{r}, \mathbf{r}', \omega')}{\omega' + \omega - \epsilon_n - i\delta_n} d\omega'.
\]

(8.11)

We now focus on the frequency integral, which can be decomposed as

\[
\int_{-\infty}^{\infty} \frac{\Delta W(\mathbf{r}, \mathbf{r}', \omega')}{\omega' + \omega - \epsilon_n - i\delta_n} d\omega' = \text{P} \int_{-\infty}^{\infty} \frac{\Delta W(\mathbf{r}, \mathbf{r}', \omega')}{\omega' + \omega - \epsilon_n} d\omega' + i\pi \text{sgn}(\delta_n) \int_{-\infty}^{\infty} \Delta W(\mathbf{r}, \mathbf{r}', \omega) \delta(\omega' + \omega - \epsilon_n) d\omega'
\]

(8.12)
With weak coupling (criterion 1), and since only self-terms contribute, significant hybridization coefficients $\alpha_{nm}$ will only be found for energies $\epsilon_n$ close to the molecular resonances at $E_{\text{mol}}^{\text{mol}} = \langle \phi^\text{mol}_n | H_{\text{tot}}^\text{DFT} | \phi^\text{mol}_m \rangle$, within the DFT width which is typically $\sim 0.5$ eV. Therefore for $\omega$ not too far from $E_{\text{mol}}^{\text{mol}}$, the energy difference $\omega - \epsilon_n$ in the denominator for terms with significant contributions is $\lesssim 0.5$ eV and can be neglected:

$$
\int_{-\infty}^{\infty} \frac{\Delta W (\mathbf{r}, \mathbf{r}', \omega') \, d\omega'}{\omega' + \omega - \epsilon_n - i\delta_n} \approx P \int_{-\infty}^{\infty} \frac{\Delta W (\mathbf{r}, \mathbf{r}', \omega') \, d\omega'}{\omega'} + i\pi \, \text{sgn} (\delta_n) \int_{-\infty}^{\infty} \Delta W (\mathbf{r}, \mathbf{r}', \omega - \epsilon_n). \tag{8.13}
$$

The time-ordered dielectric function $\varepsilon^{-1} (\mathbf{r}, \mathbf{r}', \omega)$ is an even function of $\omega$ (e.g. Hedin and Lundqvist, Eq. 12.10), so $W (\mathbf{r}, \mathbf{r}', \omega) = \varepsilon^{-1} (\mathbf{r}, \mathbf{r}', \omega') V (\mathbf{r}, \mathbf{r}')$ and $\Delta W (\mathbf{r}, \mathbf{r}', \omega')$ share the same property, and the principal part of the integral vanishes. The validity of neglecting $\omega - \epsilon_n$ can be seen from a Taylor expansion in $\omega - \epsilon_n$ of the principal part of the integral:

$$
P \int_{-\infty}^{\infty} \frac{\Delta W (\mathbf{r}, \mathbf{r}', \omega') \, d\omega'}{\omega' + \omega - \epsilon_n} = P \int_{-\infty}^{\infty} \frac{\Delta W (\mathbf{r}, \mathbf{r}', \omega') \, d\omega'}{\omega'} - (\omega - \epsilon_n) P \int_{-\infty}^{\infty} \frac{\Delta W (\mathbf{r}, \mathbf{r}', \omega') \, d\omega'}{(\omega')^2} \tag{8.14}
$$

All the even terms of $\omega - \epsilon_n$ are zero since $\Delta W (\mathbf{r}, \mathbf{r}', \omega')$ is even, and we can show the linear term is small by an order-of-magnitude analysis. The numerator is

$$
\Delta W (\mathbf{r}, \mathbf{r}', \omega') = [\varepsilon^{-1} (\mathbf{r}, \mathbf{r}', \omega') - \varepsilon^{-1}_{\text{mol}} (\mathbf{r}, \mathbf{r}', \omega')] \, V (\mathbf{r}, \mathbf{r}') \tag{8.15}
$$

and the integral can be estimated with the generalized plasmon-pole (GPP) model of Huybertsen and Louie [152]. We are interested in $\text{Re} \, \Sigma$ which comes from $\text{Im} \, \varepsilon^{-1}$, given by

$$
\text{Im} \, \varepsilon^{-1} (\omega) = -\frac{\pi}{\omega} \frac{\Omega^2}{2} [\delta (\omega - \hat{\omega}) + \delta (\omega + \hat{\omega})] \tag{8.16}
$$

where $\Omega \approx \omega_p$ (the plasma frequency, $\sim 10$ eV),

$$
\hat{\omega}^2 = \frac{\Omega^2}{1 - \varepsilon^{-1} (\omega = 0)}, \tag{8.17}
$$

and all quantities are considered in real space in the region of the molecule. Then we have

$$
P \int_{-\infty}^{\infty} \frac{\Delta W (\mathbf{r}, \mathbf{r}', \omega') \, d\omega'}{\omega' + \omega - \epsilon_n} \approx \frac{\pi}{2} (\omega - \epsilon_n) \times 2 \left[ \frac{\Omega^2}{\omega^2} - \frac{\Omega^2_{\text{mol}}}{\omega^2_{\text{mol}}} \right] V (\mathbf{r}, \mathbf{r}') \tag{8.18}
$$

For an isolated molecule, $\varepsilon^{-1}_{\text{mol}} = 1$, making the second term vanish. We estimate $\varepsilon^{-1}$ with the image charge of a single conducting plane (Sec. 8.6.2) for a molecule of size $R$ at a height $Z$ above the image plane:

$$
W = \varepsilon^{-1} V \sim \frac{1}{|\mathbf{r} - \mathbf{r}'|} - \frac{1}{|\mathbf{r} - \mathbf{r}' + 2z\mathbf{z}|} \sim \frac{1}{R} - \frac{1}{R + 2Z} = \frac{2Z}{R (R + 2Z)}. \tag{8.19}
$$
Then, since \( V = \frac{1}{|r-r'|} \sim \frac{1}{R} \), \( e^{-1} \sim \frac{2Z}{R+2Z} \). The contribution to the self-energy is

\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\Delta W(r, r', \omega') \, d\omega'}{\omega' + \omega - \epsilon_n} \approx \frac{\omega - \epsilon_n}{2\Omega} \frac{\sqrt{R}}{(R + 2Z)^{3/2}} \tag{8.20}
\]

Typical values are \( R \sim 10 \, \text{Å}, \ Z \sim 2 \, \text{Å} \) and a weak-coupling broadening \( \sim 0.5 \, \text{eV} \), for a result \( \lesssim 0.04 \, \text{eV} \) that may be safely neglected. Only for a very small molecule in close contact with the metal might this term be significant, but it will still be only a small part of the total image charge term, and in such a case weak coupling is unlikely.

The frequency integral has been reduced to

\[
\int_{-\infty}^{\infty} \frac{\Delta W(r, r', \omega') \, d\omega'}{\omega' + \omega - \epsilon_n - i\delta_n} \approx i\pi \, \text{sgn} (\delta_n) \, \Delta W(r, r', \omega)
\]

\[
\approx i\pi \, \text{sgn} (\delta_n) \, \Delta W(r, r', 0) \tag{8.21}
\]

As can be seen with reference to the GPP, using static screening is a good approximation for energies much less than the plasma frequency. We are now left with

\[
\left\langle \phi^\text{mol}_m \left| \Delta \Sigma \left| \phi^\text{mol}_m \right\rangle = -\frac{1}{2\pi} \sum_n \alpha_{nm}^* \alpha_n \, \text{sgn} (\delta_n) \times \int drdr' \phi^\text{mol}_m (r) \phi^\text{mol}_m (r') \phi^\text{mol}_m (r')^* \phi^\text{mol}_m (r)^* \Delta W(r, r', 0) \right. \right. \tag{8.22}
\]

With small charge transfer (criterion 3), an (un)occupied molecular state \( m \) only has significant contributions to (un)occupied states \( n \). The \( \text{sgn} (\delta_n) \) factors can be taken out of the summation, and the sum simplifies to

\[
\sum_n \alpha_{nm}^* \alpha_n \, \text{sgn} (\delta_n) \approx \text{sgn} (\delta_n) \sum_n \alpha_{nm}^* \alpha_n = \text{sgn} (\delta_n), \tag{8.23}
\]

since \( \alpha \) is a unitary matrix. The expression now reduces to

\[
\left\langle \phi^\text{mol}_m \left| \Delta \Sigma \right| \phi^\text{mol}_m \right\rangle = -\frac{1}{2\pi} \, \text{sgn} (\delta_n) \times \int drdr' \phi^\text{mol}_m (r) \phi^\text{mol}_m (r') \phi^\text{mol}_m (r')^* \phi^\text{mol}_m (r)^* \Delta W(r, r', 0). \tag{8.24}
\]

Define the integral

\[
\Phi^\text{mol}_m = \frac{1}{2} \int \left| \phi^\text{mol}_m (r) \right|^2 \left| \phi^\text{mol}_m (r') \right|^2 \Delta W(r, r', 0) \, drdr' \tag{8.25}
\]

where the screening potential \( \Delta W(r, r') \equiv \Delta W(r, r', 0) \). Then \( \left\langle \phi^\text{mol}_m \left| \Delta \Sigma \right| \phi^\text{mol}_m \right\rangle \approx -\Phi^\text{mol}_m \) if occupied, or \( \Phi^\text{mol}_m \) if unoccupied, just as in the case of completely non-hybridized molecule and metal states [250]. Strictly speaking, we must evaluate at the quasiparticle energy to solve the Dyson equation, but we can make a usual approximation and linearize [152]. Since we have found the dependence of \( \Delta \Sigma (\omega) \) on \( \omega \) is small around the DFT energy \( E^\text{mol}_m \), therefore \( \partial \Delta \Sigma / \partial \omega \approx 0 \), and the linearization gives the same result as evaluation at \( E^\text{mol}_m \).
The electrostatic energy of a charge distribution is
\[ E = \frac{1}{2} \int \rho(r) V(r) \, dr = \frac{1}{2} \int \frac{\rho(r) \rho(r')}{|r - r'|} \, dr, \tag{8.26} \]
a double integral with the Coulomb interaction as kernel. In the presence of conductors or
dielectrics, this kernel is modified to an effective interaction. Thus the integral \( \Phi_m^{\text{mol}} \) is the
electrostatic potential energy of the charge distribution of state \( m \) interacting with itself
via \( \Delta W(r, r') \) rather than the bare interaction. \( \langle \phi_m^{\text{mol}} | \Delta \Sigma | \phi_m^{\text{mol}} \rangle \approx \Phi_m^{\text{mol}} \) for electrons since
that corresponds to adding a charge like in the electrostatic calculation, while for holes a
charge is being removed and so the sign is reversed, \( -\Phi_m^{\text{mol}} \). The DFT calculation includes
the screened Coulomb interaction at the mean-field level, and intramolecular exchange and
correlation are taken care of by the gas-phase self-energy. Then the change in self-energy
correction \( \Delta \Sigma \) due to the change in \( \Delta W(r, r') \) is entirely due to the metal and can
be evaluated via image charges (Sec. 8.6.2). As the interaction via the image charges is
attractive, \( \Phi_m^{\text{mol}} \) is negative and the energy of occupied states increases while the energy of
unoccupied states decreases, reducing the gap.

Once we have the expectation values of the self-energy correction to molecular orbitals,
we use them to construct our self-energy operator via the expression
\[ \hat{\Sigma} \approx \sum_n \langle \psi_n | \hat{\Sigma} | \psi_n \rangle \langle \psi_n | \psi_n \rangle \tag{8.27} \]
arrived at in Sec. 8.3.3. This form involves the junction orbitals but our approximation
there and in Sec. 8.3.4 gives us only matrix elements for the gas-phase orbitals, \( \langle \psi_n | \hat{\Sigma} | \psi_n \rangle \),
so we must make a further approximation. We replace the matrix elements by their gas-
phase values: \( \sum_n \langle \psi_n | \hat{\Sigma} | \psi_n \rangle \langle \psi_n | \psi_n \rangle \). If we consider the change in the wavefunctions
upon adsorption as a perturbation, this choice is correct to the first order. Let \( \psi_n^0 \) be the
nth gas-phase molecular orbital and \( \psi_n' \) the change on adsorption, so \( \psi_n = \psi_n^0 + \psi_n' \), where
\( \langle \psi_n | \psi_n' \rangle \approx 0 \). We have approximated \( \hat{\Sigma} \) to be diagonal in the basis of junction molecular
orbitals. Based on these properties, the matrix elements \( \langle \psi_n' | \hat{\Sigma} | \psi_n \rangle \) are zero to first order
in the perturbation, and therefore, the expectation values of the self-energy for the two sets
of orbitals is the same to first order:
\[ \langle \psi_n' | \hat{\Sigma} | \psi_n' \rangle = \langle \psi_n | \hat{\Sigma} | \psi_n \rangle - \langle \psi_n' | \hat{\Sigma} | \psi_n \rangle - \langle \psi_n' | \hat{\Sigma} | \psi_n' \rangle \approx \langle \psi_n' | \hat{\Sigma} | \psi_n \rangle. \tag{8.28} \]

Thus approximating
\[ \hat{\Sigma} = \sum_n \langle \psi_n | \hat{\Sigma} | \psi_n \rangle \langle \psi_n | \psi_n \rangle \approx \sum_n \langle \psi_n^0 | \hat{\Sigma} | \psi_n^0 \rangle \langle \psi_n^0 | \psi_n^0 \rangle \tag{8.29} \]
is correct to first order. By contrast an expression such as \( \hat{\Sigma} \approx \sum_n \langle \psi_n^0 | \hat{\Sigma} | \psi_n^0 \rangle \langle \psi_n^0 | \psi_n^0 \rangle \)
is not since the gas-phase orbitals used as projectors differ from the junction orbitals to
first order.
8.6.2 Potential of charge distribution between two planar conductors

The form of the image-charge interaction is dependent on the geometry of the conductors. We consider here a finite charge distribution between two planar conductors, appropriate for modeling a single molecule contacted to macroscopic metal leads in experimental measurements. Note that to reproduce a GW calculation of a molecular junction in periodic-boundary conditions \[290, 340\] we would need to consider an unphysical periodic array of molecules via a 2D Ewald sum of image charges in addition to the summation done here.

First consider a single planar conductor. A point charge \( q \) located at a height \( z \) above a planar conductor induces an image charge \( -q \) at a depth \( z \) below the surface, equivalent to solving for the boundary condition \( V(z = 0) = 0 \). The potential energy of this charge distribution, using the potential \( V(r) \) due to the image charge \([162]\), is

\[
E = \frac{1}{2} q V(0, 0, z) = \frac{1}{2} q \frac{-q}{2z} = -\frac{q^2}{4z}.
\]

For a charge distribution, instead we use the integral form, where

\[
\Delta W(r, r') = -\frac{1}{|r - (r' - 2z\hat{z})|},
\]

and either of the forms

\[
\Phi_m^{\text{mol}} = -\frac{1}{2} \int \left| \phi_m^{\text{mol}}(r) \right|^2 \frac{1}{|r - (r' - 2z\hat{z})|} dr dr'.
\]

The interaction is attractive, so the energy is lowered.

With two planar conductors located at \( z \leq 0 \) and \( z \geq a \), the boundary conditions \( V(z = 0) = V(z = a) = 0 \) are equivalent to considering that each charge in the gap induces an infinite series of image charges in each conductor, since each image charge in one conductor has an image in the other conductor (see Fig. 8.6). For a charge \( q \) located at \( r' = (x', y', z') \), there are series of charges \( q \) at \( z = z' + 2na \) and charges \( -q \) at \( z = -z' + 2na \), for each integer \( n \). The first series includes the real charge itself as the \( n = 0 \) member, which must be excluded from the potential felt by the real charge. Let \( \delta p^2 = (x - x')^2 + (y - y')^2 \).

The potential due to the image charges is \( V(r) = q\Delta W(r, r') \) with kernel

\[
\Delta W(r, r') = \sum_{n=-\infty}^{\infty} \left[ \frac{1}{\sqrt{\delta p^2 + (z - z' - 2na)^2}} - \frac{1}{\sqrt{\delta p^2 + (z + z' - 2na)^2}} \right] - \frac{1}{\sqrt{\delta p^2 + (z - z')^2}}.
\]

This series cannot be summed in closed form without special functions \([397]\), and indeed the two terms in the infinite series are divergent if taken separately. However, considered as
Figure 8.6: Infinite series of charges arising from a point charge in between two planar conductors located at $z \leq 0$ and $z \geq a$. For clarity, the series of image charges is split into two sets: the series of charges that originate first from a reflection in the plane $z = 0$ are shown in (a), and those that originate first from a reflection in the plane $z = a$ are shown in (b). The charges are denoted by their sign, and are labeled according to the order in which they arise. For example, in (a), the charge labeled “1” arises from reflection of the physical charge in plane $z = 0$, while that labeled “2” arises from reflection of image charge “1” in plane $z = a$.

A single series, with both terms at a given $n$, the sum is convergent, and can be evaluated numerically. (For the geometries studied in this work, summing as few as 10 images is sufficient for 1 meV precision in the self-energy correction.) The energy of the charge distribution is the interaction of each charge with the image charges due to itself and all the other charges:

$$E = \frac{1}{2} \int \rho(r) \rho(r') \Delta W(r, r') \, dr dr'.$$  \hfill (8.34)

The effective interaction $\Delta W(r, r')$ is attractive as in the one-conductor case, so that $E$ is negative.
Chapter 9

First-principles calculations on the molecular photoswitch azobenzene

This work is partly based on Refs. [71, 72].

9.1 Abstract

Azobenzene is an organic molecule that can undergo light-induced switching between *trans* and *cis* isomers, and could serve as a molecular actuator or switch. I used *ab initio* calculations to confirm the experimental observation of photoisomerization on a metal surface, and showed how addition of substituents to the molecule enables this process by reducing electronic coupling to the surface. I identified the structures of the two different conformations of *cis* molecules observed on the surface, and used their chirality to determine the photoisomerization pathway. Finally, we calculated the conductance of azobenzene molecules in junctions between gold leads, finding a decrease of an order of magnitude on switching between *trans* and *cis*.

9.2 Introduction

Azobenzene is an organic molecule, C$_{12}$H$_{10}$N$_2$, consisting of two phenyl rings attached by an N=N double bond (Fig. 9.1). There are two structural isomers, *trans* and *cis*, defined by different geometry about the double bond. The *trans* isomer is the ground state, while the *cis* isomer is higher in energy by 0.58 eV. UV light can transform *trans* to *cis*, while visible light promotes the reverse process [350]. There is also a slow thermal relaxation of *cis* to *trans*, over a barrier of approximately 1.6 eV [315]. These phenomena have all been studied extensively in the gas and solution phases, through spectroscopic experiments and theoretical calculations, although the mechanism of photoisomerization still remains controversial. The *trans* isomer is planar, while the *cis* isomer is non-planar and has a shorter end-to-end distance, so isomerization can apply a force between units attached to the two ends of the molecule.
9.3 Confirmation of photoisomerization on surface

While the photoisomerization behavior of azobenzene molecules in the gas phase or in solution is well known [350], photoisomerization on surfaces would be useful in building nanoscale devices. Collaborators in the Crommie group demonstrated photoisomerization of an azobenzene derivative on Au(111) by imaging a monolayer of molecules by scanning tunneling microscopy (STM) before and after illumination. The planar trans molecules absorb flat on the surface and appear as four equivalent lobes, while illumination caused one of the lobes to become higher than the others. To confirm that this change in fact signalled a reversible isomerization rather than some other process such as dissociation of the molecules, I performed ab initio density-functional theory calculations to simulate STM images of the trans and cis isomers to compare to experiment [71].

Within the Tersoff-Hamann approximation [356], the tunneling current in STM is proportional to the local density of states (LDOS). LDOS calculations were performed for isolated trans- and cis-TTB-azobenzene molecules using the SIESTA code [328] (similarly to Ref. [364] but with the generalized-gradient approximation [269]). Relaxed structures for the trans and cis isomers were calculated by energy minimization [Fig. 9.2(a, b)]; the cis CNNC and CCNN dihedral angles are 11° and 47°, respectively. Isosurfaces of HOMO orbital LDOS were calculated to simulate STM topographs at constant current. The simulated trans isomer STM image is dominated by four peripheral lobes at the TB-leg positions [Fig. 9.2(c)]. The simulated cis isomer STM image shows a bright central area due to the upwards rotation of one TB-leg, leaving the three remaining TB legs on the periphery below [Fig. 9.2(d)]. A simple tiling of the calculated trans and cis isomer simulated images using experimentally observed lattice parameters [Fig. 9.2(e)] shows that the simulated trans- and cis-TTB-azobenzene images match the experimental data [Fig. 9.2(e)] quite well.

9.4 Study of surface absorption and hybridization with metal

The experiments on Au(111) studied three azobenzene molecules: the bare molecule (AB), the disubstituted 4,4'-di-tert-butylazobenzene (DTB-AB), and the tetrasubstituted 3,3',5,5'-tetra-tert-butylazobenzene (TTB-AB). Comstock et al. found that photoisomerization only occurred for TTB-AB and did not occur for AB or DTB-AB [71]. The addition of the insulating tert-butyl “legs” to the molecule was shown to increase the apparent height of the the molecules from AB to DTB to TTB-AB in STM constant-current images, and therefore they hypothesized that the addition of the legs reduced the electronic coupling
Figure 9.2: Simulated trans- and cis-TTB-azobenzene structures compared to experiment [71]. (a) Calculated trans geometry. (b) Calculated cis geometry. (c) Calculated trans LDOS integrated from $E_F$ to $E_F - 1$ eV, at an isosurface about 3 Å away from the nearest atoms. (d) Calculated cis LDOS isosurface [same parameters as in (c)]. (e) Simulated STM image of TTB-azobenzene using tiled single-molecule LDOS isosurfaces from (c) and (d) (image has been smoothed using a 0.2 nm width Gaussian blur filter to approximate experimental convolution with the STM tip). (f) Experimental STM image of TTB-azobenzene molecules including one photoisomerized cis isomer.
between the molecule and the metal surface. I used DFT calculations to explore what the change is between AB and TTB-AB, and why only TTB-AB photoisomerizes on Au(111).

I used a three-layer slab of unreconstructed Au(111) with molecules absorbed in the closest approximation to the observed incommensurate surface lattice. I identified the lowest-energy binding site of both molecules on Au(111) as an atop site, in which the N=N group to be centered on an Au atom in the top layer. The optimal absorption height, defined as the distance from the N=N group to the top surface layer, was found to be 3.0 Å for AB and 4.0 Å for TTB-AB. From the calculated energy vs. height curve for TTB-AB, with a binding energy of 2.3 eV, we can calculate the expected peak position in the temperature-programmed desorption experiment [295, 411], and we find a temperature $T_P = 869$ K. After applying a correction for basis-set superposition error, which has been shown to be essential for absorption of aromatic molecules [205], we have a binding energy of 1.5 eV and $T_P = 586$ K, in excellent agreement with the measured value of $525 \pm 10$ K [134].

I then calculate the partial density of states (PDOS) on the molecule for the optimized absorption geometry for AB and TTB-AB. In the case of TTB-AB, we find well-defined peaks corresponding to the energy levels of the isolated molecule, with little change other than a broadening of 0.1 eV. However, in the case of AB, we find that no recognizable molecular states persist after absorption, and broadening is at least 1 eV. When the absorption height is increased to 4.0 Å, well-defined molecular peaks are recovered in the PDOS. Therefore, the absorbed TTB-AB is able to photoisomerize like the isolated molecule because the HOMO and LUMO states involved in the optical process retain their identity, whereas the states of AB are completely changed by hybridization with the metal on absorption, preventing photoisomerization. Quantitatively, we can calculate decay times into the surface of states on the molecule based on the linewidths, yielding 7 fs for TTB-AB and 0.7 fs for AB. The characteristic time for the isomerization to occur after absorption of light, as measured by ultrafast spectroscopy, is 320 fs [350]. Therefore we can roughly estimate the quantum efficiency of photoisomerization as $\exp \left(-\frac{\tau_{\text{decay}}}{\tau_{\text{isom}}}\right)$, yielding $10^{-10}$ for TTB-AB and $10^{-100}$ for AB. These results clearly can account for the experimentally estimated quantum efficiency of $10^{-10}$ for TTB-AB [70], and the non-observation of photoisomerization for AB. The calculations suggest that the electronic mechanism for photoisomerization only involves the molecule, and the surface’s role can only be to quench the process.

### 9.5 Use of chirality to identify photoisomerization pathway

A major subject of investigation on azobenzene has been the mechanism of photoisomerization in the solution phase: how do the atoms in the molecule move to accomplish the transition between the trans isomer and the cis isomer? Three candidates have been proposed: rotation about the N=N double bond, changing the CNNC dihedral angle; inversion at one of the N atoms, changing the CNN bond angle [350]; and concerted inversion, in which both CNN bond angles change simultaneously [394]. Many theoretical and experimental studies have been done, but there is evidence to support each possibility, and the mechanism remains controversial [292, 116, 58, 364, 73, 343].

However, high-resolution STM images opened the possibility of observing the mecha-
nism of photoisomerization on a surface, via the chirality and conformation of the absorbed molecules before and after isomerization. The planar trans isomer is achiral in isolation, but becomes chiral upon absorption (Fig. 9.3); the cis isomer is chiral even in isolation. Careful analysis of STM images identified the chiralities of a large sample of molecules before and after isomerization, and also revealed two different types of cis isomer [72].

Structures of isolated molecules of each type were initially relaxed and then assembled into a monolayer of molecules (without surface atoms) according to the experimentally observed surface lattice parameters. Calculations showed that including the surface atoms had little effect on the simulated STM topographs. The simulated trans-TTB-AB molecules (Fig. 9.5(b)) show four lobes at the corners of a parallelogram angled from the horizontal and vertical by 2°, as observed in the STM images (Fig. 9.3), confirming the existence of two enantiomers of trans-TTB-AB on the surface.

Consider the cis molecules, I identified Type 1 as the conventional conformation for cis-azobenzene, as found in theoretical calculations [364, 73], STM images [71, 2], and crystal structures [240]. However, Type 2 was a novel conformation which had been suggested by calculations of Tiago et al. [364] but never observed previously. Type 1 has both NNCC dihedral angles about 50° and has C₂ rotation symmetry, while Type 2 has one angle +50° and the other -50° and has mirror symmetry (Fig. 9.4). I calculated simulated STM images of the cis conformations and found good agreement with the experiment (Fig. 9.5).

The STM data showed a strong correlation between the chirality of a trans molecule, and the chirality and conformation of the resulting cis molecule: a trans molecule could become either Type 1 with opposite chirality, or Type 2 with the same chirality (less than 10% were exceptions) [72]. I was able to rule out the rotation and concerted inversion mechanisms for TTB-AB on Au(111) using this data and the theoretically determined molecular structures. The rotational pathway, for example, leads to trans-cis conversion via out-of-plane rotation of a phenyl ring as the CNNC dihedral angle is swept. For a right-handed (left-handed) trans isomer on a surface this motion always yields a right-handed (left-handed) Type 1 cis isomer. Because this violates the observations, we can rule it out as a significant dynamical pathway for TTB-AB photoswitching on gold. The concerted inversion pathway, on the other hand, yields trans-cis conversion through a mechanism involving synchronous inversion of the phenyl rings as the two NNC bond angles are simultaneously swept. Here the TTB-AB molecule goes through a transition state in which the CNNC atoms are nearly collinear, a configuration that by symmetry is expected to lead to either right-handed or left-handed cis conformations with nearly equal probability. Since in fact a significant difference in production of the two enantiomers was observed, we rule it out too as a significant dynamical pathway for TTB-AB switching on gold.

However, a version of the inversion mechanism can account for formation of the two different conformations with their characteristic chiralities (see Fig. 9.6). I proposed that isomerization begins with the in-plane inversion of one phenyl ring and proceeds until the moving phenyl ring makes contact with the stationary ring. A branch in the dynamics occurs at this point. In the first branch (lower left-hand side of Fig. 9.6) the inverting phenyl ring lifts out of the plane starting from the point away from contact, yielding a Type 2 cis configuration of the proper chirality. In the second branch (lower right-hand side of Fig. 9.6) the stationary phenyl ring lifts out of plane starting from the point of
Figure 9.3: (a) Image of trans-TTB-AB island composed of two mirror-image chiral domains (domain boundary indicated by dashed white line). Trans-TTB-AB single-molecule structural models are superimposed on right-handed and left-handed configuration. Red crosses indicate locations of TB lobe maxima. (b) Average locations of trans-TTB-AB TB lobes averaged over 394 left-handed domain molecules (left plot) and 125 right-handed domain molecules (right plot). Left-handed and right-handed domain trans-TTB-AB molecules form parallelograms that are angled by $2^\circ$ from the horizontal and vertical. Figure from Ref. [72].
Figure 9.4: Symmetries of the two conformations of \textit{cis}-TTB-azobenzene. TB legs not shown for clarity. Left: Type 1 has a rotation axis, marked by the cross and arrow. Right: Type 2 has a mirror plane.
Figure 9.5: (a) High-resolution STM image [72] shows structure of two distinct types of photoswitched cis-TTB-AB isomers: a left-handed Type I cis and a right-handed Type 2 cis, contained within dashed white boxes. (b) *Ab initio* DFT simulated STM image of the two cis-TTB-AB types surrounded by trans-TTB-AB molecules, with structural models shown of trans-TTB-AB as well as cis Type 1 and cis Type 2 used to calculate the simulated STM image.
<table>
<thead>
<tr>
<th>Conformation</th>
<th>Orbital</th>
<th>DFT Eigenvalue</th>
<th>Gas-Phase Energy</th>
<th>Self-Energy Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>trans</td>
<td>HOMO</td>
<td>−4.58 eV</td>
<td>−6.64 eV</td>
<td>−6.06 eV</td>
</tr>
<tr>
<td>trans</td>
<td>LUMO</td>
<td>−2.56 eV</td>
<td>−0.37 eV</td>
<td>−0.95 eV</td>
</tr>
<tr>
<td>cis</td>
<td>HOMO</td>
<td>−4.07 eV</td>
<td>−6.45 eV</td>
<td>−5.81 eV</td>
</tr>
<tr>
<td>cis</td>
<td>LUMO</td>
<td>−2.59 eV</td>
<td>−0.36 eV</td>
<td>−1.01 eV</td>
</tr>
</tbody>
</table>

Table 9.1: Self-corrected energies for orbitals of diaminoazobenzene molecular junctions.

contact, yielding a Type 1 cis configuration of the proper chirality. This mechanism is supported by the fact that it should be energetically more favorable for portions of the molecule to slide along the surface (as in inversion) rather than completely unbind and lift off of the surface (as in rotation). Therefore we can conclude that the inversion mechanism is the dominant one for TTB-AB on Au(111) [72].

9.6 Effect of photoisomerization on conductance in molecular junction

The ability to isomerize (optically or otherwise) a molecule that forms part of a nanoelectronic circuit would be a useful device functionality. Other azobenzene systems have studied theoretically as molecular switches [85, 412]. It was decided to study the conductance of a system on which experiments were concurrently proceeding, single-molecule junctions of the molecule 4,4′-diaminoazobenzene between Au leads, to see what the effect of isomerization is on the electronic properties. Molecules linked to Au electrodes via amine groups have been shown by experiment and theory to give reproducible conductances [383, 289], so I used a representative surface bonding motif to construct junction geometries for this molecule on Au(111). I used a first-principles scattering-state method based on density-functional theory (DFT), implemented in the SCARLET code [62], incorporating model self-energy corrections to the frontier levels which are essential to improve the accuracy of the DFT and agree quantitatively with experiment [289, 250], as described in Chapter 8. The junction structures are shown in Fig. 9.7, and self-energy corrections used are shown in Table 9.1, computed as in Chapter 8.

I calculated transmission spectra for each isomer with adatom trimer bonding motifs and $4 \times 4 \mathbf{k}_\parallel$-grids (Fig. 9.8). For trans the conductance was 0.003 $G_0$, close to the measured value of 0.00074 $G_0$ from STM break-junction experiments [154]. At the Fermi energy, conduction in trans is dominated by the HOMO-1 ($\pi$) orbital, which has an unusual Fano lineshape [103] due to interaction between the HOMO and HOMO-1 levels. HOMO is localized on the N=N double bond, and has small hybridization with the rest of the molecule, leading to a sharp peak with negligible contribution to conductance at the Fermi
Figure 9.6: Photoisomerization dynamics model for TTB-azobenzene on the Au(111) surface. A photoexcited right-handed trans-TTB-azobenzene molecule (fig. top center, TB legs numbered) begins the isomerization reaction with the in-plane inversion of the left phenyl ring through the transition configuration (fig. middle center). With continuing inversion, steric repulsion due to the eventual contact of the left phenyl ring with the right phenyl ring then forces either: (a) left phenyl to twist out of plane lifting TB #1 off of surface producing right-handed Type 2 cis-TTB-azobenzene (figure bottom left), or (b) right phenyl to twist out of plane lifting TB #4 off of the surface producing left-handed Type 1 cis-TTB-azobenzene (figure bottom right). A left-handed trans-TTB-azobenzene isomerizes via two equivalent mirror-reflected paths.
level. This resonance can be fit very well by the Fano form

\[ T(E) = \frac{(q\Gamma/2 + E - E_0)^2}{(E - E_0)^2 + (\Gamma/2)^2} \]  

with asymmetry parameter \( q = 1.977 \), peak center \( E_0 = -0.596 \) eV, and width \( \Gamma = 0.001 \) eV (Fig. 9.9).

By contrast, the lower conductance of 0.0002 \( G_0 \) of \( cis \) is dominated by HOMO. Therefore, isomerization leads to a switching ratio of 14 in conductance. This trend is in accordance with the fact that \( trans \)-to-\( cis \) isomerization breaks the \( \pi \)-conjugation across the molecule. DFT incorrectly predicts a higher conductance for \( cis \) (0.020 \( G_0 \) vs. 0.011 \( G_0 \)), showing that self-energy corrections are crucial to accurate electronic-transport calculations.

This work was performed with the Lorentzian estimate. Using the more sophisticated DFT+\( \Sigma \) (Chapter 8) for \( trans \), a result of 0.0017 \( G_0 \) is obtained, in slightly better agreement with the experiment. Future work is to do this calculation with appropriate corrections to both the HOMO and HOMO-1 levels, rather than one correction for HOMO used for all the occupied levels, as done here.
Figure 9.8: Transmission spectra for diaminoazobenzene junctions. Top: \textit{trans}. Bottom: \textit{cis}. 
Figure 9.9: Detail of Fano resonance in transmission of trans-diaminoazobenzene junction, showing excellent fit to Eq. 9.1 for asymmetry parameter $q = 1.977$, peak center $E_0 = -0.596 \text{ eV}$, and width $\Gamma = 0.001 \text{ eV}$. 
Chapter 10

Efficient $k \cdot p$ method for first-principles calculation of the Seebeck coefficient in quantum transport

10.1 Abstract

Thermoelectric properties of molecular junctions reveal fundamental aspects of nanoscale charge transport at interfaces and are relevant to potential organic/inorganic hybrid thermoelectric materials. Quantum-transport calculations typically evaluate the Seebeck coefficient $S$, the ratio between a voltage difference and temperature difference across an interface, by finite differences of the transmission as a function of energy. However, in $ab\ initio$ calculations this quantity is difficult to converge for realistic systems and can require very large grids of $k_\parallel$-points (wavevectors parallel to the plane of the metallic lead surfaces, which is perpendicular to the transport direction). I derived a new method to evaluate $S$ via analytic derivatives, which can be performed using quantities available from standard electronic-structure and transport calculations, and demonstrated on a model how and why the typical method suffers from convergence problems, and how these problems are remedied with my approach. This technique improves $k_\parallel$-point convergence and allows more efficient and accurate calculations of Seebeck coefficients.

10.2 Introduction

The Peltier effect and Seebeck effect are two closely related thermoelectric effects. In the Peltier effect, a voltage applied across an interface generates a temperature difference, which can be used to cool or heat a junction in a circuit, or for solid-state refrigerators, available in some consumer products. The reverse is the Seebeck effect (also known as thermopower), in which a temperature difference generates a voltage. This effect can be applied as a method of generating electricity from waste heat, and also for temperature
measurement in thermocouples [372]. Both processes can be described by the equation \( \Delta V = S \Delta T \), where \( S \) is the Seebeck coefficient. (Sometimes \( S \) is defined as \( \Delta V = -S \Delta T \) instead.)

The efficiency of the cooling or energy conversion process is controlled by the dimensionless figure of merit \( ZT \); when \( ZT \to \infty \), the coefficient of refrigeration and efficiency approach the Carnot limit. This figure of merit is given by

\[
ZT = \frac{\sigma S^2 T}{\kappa}
\]  

(10.1)

where \( T \) is the absolute temperature, \( \sigma \) is the electrical conductivity, and \( \kappa \) is the thermal conductivity (including both electronic and lattice contributions). The best current devices have \( ZT \sim 1 \). To improve \( ZT \), an “electron crystal, phonon glass” is the design goal, which would maximize the electronic conduction, like a crystal, while minimizing the lattice thermal conductivity (which dominates \( \kappa \) except for metals), like a glass [372].

The Seebeck effect is also useful as an experimental probe of charge transport. It is diagnostic of the type of carriers (electron or hole), as used in bulk semiconductors [187] and molecular junctions [266, 17]. The related Nernst effect, which is roughly a simultaneous Seebeck and Hall effect, is used to study vortices in superconductors [255].

There has been a lot of interest in thermoelectrics recently, partly due to the fact that it has been discovered that some of the bulk materials (e.g. \( \text{Bi}_2\text{Te}_3 \)) with the best Seebeck coefficient are topological insulators [408]. The band structure of \( \text{Bi}_2\text{Te}_3 \) has been studied in detail theoretically [186]. The grain boundaries of this material have been investigated for their potential in scattering phonons but not electrons [230]. The figure of merit of semiconductors could be improved by doping schemes [204] or nanostructuring [142, 144]. Molecular junctions may realize the idea of “electron crystal, phonon glass” by allowing resonant tunneling but blocking phonon transport [245]. They may also have very large Seebeck coefficients due to their narrow peaks in the density of states, with the Mahan-Sofo mechanism [213]. The modified STM break-junction technique [401] has been applied successfully to studies of the Seebeck effect in molecular junctions [294, 17], and such studies may give insight into the workings of hybrid organic/inorganic thermoelectric devices [392, 317].

### 10.3 Typical method of calculating the Seebeck coefficient

To study these systems, we must use quantum transport. In the Landauer formalism, a difference in distribution functions, weighted by the transmission \( T \), drives a current [80]:

\[
I = \frac{e}{\pi \hbar} \int_{-\infty}^{\infty} T(E) [f_1(\mu_1, T_1, E) - f_2(\mu_2, T_2, E)] dE
\]  

(10.2)

Usually we consider a difference due to a voltage, \( i.e. \mu_1 = \mu_2 + eV \). This leads to a conductance

\[
G = \frac{\partial I}{\partial V} = \frac{e^2}{\pi \hbar} \int_{-\infty}^{\infty} T(E) f'(E - \mu) dE
\]  

(10.3)
However, it can also be due to a temperature difference. If we consider the linear response of $I$ to a temperature difference, and divide by the conductance, we obtain

$$S = \frac{\partial V}{\partial T} = -\frac{1}{eT} \frac{\int_{-\infty}^{\infty} (E-\mu) T(E) f'(E-\mu) \, dE}{\int_{-\infty}^{\infty} T(E) f'(E-\mu) \, dE}.$$  \tag{10.4}$$

This approach has been used by a number of recent works for the Seebeck coefficient of molecular junctions from quantum transport [266, 267, 108, 209, 183, 285].

This equation can be greatly simplified in the low-temperature limit, where we can perform a Sommerfeld expansion of these quantities [15] and find

$$S = \frac{\pi^2 k_B^2 T}{3e} \frac{1}{T} \frac{\partial T}{\partial E} \bigg|_{E_F}.$$  \tag{10.5}$$

Then the question is how to evaluate this derivative. Normally it is done numerically with a finite difference, i.e.

$$\frac{\partial T}{\partial E} \approx \frac{T(E_F + \frac{1}{2}\Delta E) - T(E_F - \frac{1}{2}\Delta E)}{\Delta E}.$$  \tag{10.6}$$

which would be exact in the limit $\Delta E \to 0$. However, as is well known, such numerical derivatives have a serious loss of precision when the numerator and denominator both become small and noise begins to dominate, so there is a tradeoff that has to be made in choosing a good value of $\Delta E$. The lack of numerical precision was noted in the early work of Paulsson and Datta, who decided instead to fit their transmission to a Lorentzian form and take the derivative of that [266]. Of course, prejudging the lineshape of the curves is not a generally satisfactory solution.

An additional problem arises with $k$-sampling. We know the transmission only as an integral in the 2D Brillouin zone of $k_\parallel$:

$$T(E) = \sum_n \int E \mathcal{T}(k_x, k_y, k_z^n) \, dk_x dk_y$$  \tag{10.7}$$

where $k_z^n$ for each lead band $n$ is determined by the condition

$$\epsilon_n(k_x, k_y, k_z^n) = E$$  \tag{10.8}$$

and the integration domain is where such solutions exist and also the states are right-moving [62]:

$$v_z = \frac{1}{\hbar} \frac{\partial \epsilon_n}{\partial k_z} > 0.$$  \tag{10.9}$$

It was found in recent work that the Seebeck coefficient is much harder to converge with $k$-sampling than the conductance [285]: an $8 \times 8$ grid was sufficient for $T$ but at least $24 \times 24$ was required for $S$. These problems are simply hidden if only $\Gamma$-point sampling is used, as has been done in many works despite the use of bulk-like leads. The transmission as a function of $k_\parallel$ can be quite non-uniform, as shown in Fig. 10.1 based on that work. Moreover, the convergence with respect to $\Delta E$ must be considered too, in principle.
10.4 Numerical study of convergence in simple model

We will perform a study of a simple model to illustrate the convergence difficulties of numerical differentiation for the Seebeck coefficient. We use an energy dispersion appropriate for a parabolic valence band with effective mass \( m \), \( \epsilon(k) = -\hbar^2 k^2 / 2m \), and a transmission function corresponding roughly to a Lorentzian resonance for a HOMO orbital:

\[
T(k) = \frac{\alpha}{1 + \beta k^2} \frac{\Gamma^2}{(\epsilon(k) - E_0)^2 + \Gamma^2}. \tag{10.10}
\]

\( \Gamma \) represents the width due to coupling between the orbital and the lead; the transmission peak is centered roughly at \( E_0 \). \( \alpha, \beta \) are arbitrary parameters. We use the values \( m = -0.1, \Gamma = 0.2, E_0 = -2, \beta = 10, \alpha = 10 \), chosen to illustrate how poor convergence can be, despite the relatively smooth nature of the functions used. We do not use a purely energy-dependent \( T \) because some cancellations would occur and give a trivial result in that case, and of course in general there is \( k \)-dependence to the transmission.

This model has the virtue that the calculations can easily be done analytically, for comparison to numerical results. In Fig. 10.2 I show the model and numerical integration to obtain the transmission, which converges rapidly. Also, the analytic integration for \( T \) and \( \partial T / \partial E \) are shown.

We can now numerically compute \( \partial T / \partial E \) by first integrating (Eq. 10.7) and then differentiating (Eq. 10.6). The quality of the result is controlled by the \( k \)-point sampling, expressed as the separation \( \Delta k \) between the points, and the energy difference \( \Delta E \). We use \( k \)-grids from \( 5 \times 5 \) to \( 1000 \times 1000 \). The approach to the limit \( \Delta k \to 0, \Delta E \to 0 \) (known from the analytical calculation) is slow and irregular, as seen in Fig. 10.3. The convergence with the two parameters is strongly coupled and they cannot be considered separately.
Figure 10.2: Upper left: $T(k_x, k_y, k_z(E))$ for the model. Horizontal plane is $k_x, k_y$, vertical axis is $T$. Lower left: Numerical integration for $T$, showing rapid convergence. Right: Analytical evaluation of transmission (top) and its derivative (bottom) for the model. We will study the numerical results at the energy $E = -1.5$, marked by the vertical line. Units are arbitrary.
Figure 10.3: Slow and irregular convergence of $\partial T / \partial E$ computed via numerical integration and differentiation with parameters $\Delta k$ and $\Delta E$. Top shows variation of both parameters. Bottom left shows variation of $\Delta E$ for two values of $\Delta k$. For $N_k = 25$, convergence is reached, to an incorrect result; for $N_k = 100$, no convergence is attained, inasmuch as the smallest two energy values have a large difference. Note the log scale on the $x$-axis. The analytic result is denoted by the horizontal orange line. Bottom right shows variation of $\Delta k$ for two values of $\Delta E$. $\Delta E = 0.001$ shows initially flat behavior that could be mistaken for convergence, and then goes to the right answer; $\Delta E = 0.5$ converges smoothly to an incorrect result.
Why is the convergence slow and non-monotonic? It is helpful to consider analytically the derivative of the integral, which can be expressed via “differentiation under the integral sign,” also known as the Leibniz integral rule [110]. In 1D, the familiar form for the derivative of an integral whose limits and integrand depend on a parameter is

\[
\frac{\partial}{\partial t} \int_{a(t)}^{b(t)} f(x, t) \, dx = \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t} (x, t) \, dx + \frac{\partial b}{\partial t} f(b(t), t) \, dx - \frac{\partial a}{\partial t} f(a(t), t) \, dx
\]

(10.11)

where the first term is from the variation of the integrand, and the second from the variation of the limits. Generalizations of this formula exist in higher dimensions [110]. For our case in 2D, we have

\[
\frac{\partial T}{\partial E} = \frac{\partial}{\partial E} \int_E T(k) \, dk_x \, dk_y - \int_E \frac{\partial T}{\partial \epsilon} \frac{\partial \epsilon}{\partial k_z} \, dk_x \, dk_y + \oint_{\partial E} T(k) \frac{\nabla \epsilon \times dk}{|\nabla \epsilon|^2}
\]

(10.12)

The integration domain is as specified above, and \( \partial E \) denotes the edge of that domain, where \( v_z = 0 \).

There are two problems that can be seen. First, the variation of the integrand can diverge when \( \partial \epsilon / \partial k_z = 0 \). This may be regarded as a “coordinate singularity,” and the divergence is certainly integrable. Such points of divergence are not rare; indeed precisely this condition is the one that defines the edge of the integration domain. This situation is notoriously difficult to integrate numerically. For a multidimensional integral with an integrable divergence at an unknown location, the standard text *Numerical Recipes* advises: “you might as well ... quit” [281]. The reason is that depending on exactly where we put our \( k \)-points we may get zero or enormous contributions to the integral, making the convergence extremely sensitive.

The “surface term” from the variation of the integration domain also is problematic. As shown in Fig. 10.4, it is only captured in our numerical integration if we sample the Brillouin zone finely enough to have some points between \( E \) and \( E + \Delta E \). As \( \Delta E \) goes to zero, this shell becomes narrower and narrower, and requires increased \( k \)-sampling to retain some contribution. This is the reason for the coupled convergence of \( \Delta E \) and \( \Delta k \) we observed.

### 10.5 Surface-integral formulation

Fortunately, we can recast the derivative of the integral into a new form that does not suffer from these problems, and will be more easily convergent. We begin by re-expressing the transmission via a surface integral, over a surface of constant energy in the 3D Brillouin zone of the bulk leads, rather than as a 2D integral. We focus on one particular band, dispensing with the sum for simplicity:

\[
T(E) = \int_E T(k) \, dk_x \, dk_y \int_E T(k) \frac{\partial \epsilon / \partial k_z}{|\nabla \epsilon|^2} \, d^2 A_k
\]

(10.13)

A similar expression is often used for the density of states [418].
Figure 10.4: Convergence of the surface term. To capture contributions from the variation of the integration domain in our numerical evaluation of $\partial T / \partial E$, we must have k-points in our grid (blue dots) in the shell between the contours for energies $E$ and $E + \Delta E$ (red rings). As $\Delta E$ goes to zero, the width of the shell (orange double-headed arrow) goes to zero also, making it increasingly difficult to retain this contribution.

In this form, we may use the surface-integral version of differentiation under the integral sign, in the notation of Flanders for a time-dependent surface [110]:

$$\frac{d}{dt} \int_{\Sigma(t)} F(\mathbf{r}, t) \cdot d\mathbf{A} = \int_{\Sigma(t)} \left[ \frac{\partial F}{\partial t} + (\nabla \cdot F) \mathbf{v} \right] \cdot d\mathbf{A} - \oint_{\partial\Sigma} (\mathbf{v} \times F) \cdot d\mathbf{S} \quad (10.14)$$

We have only a scalar function, not a vector one, so let $F(\mathbf{r}, t) = F(\mathbf{r}, t) \hat{n}(\mathbf{r}, t)$ where $\hat{n}$ is the surface normal. $\mathbf{v} = v \hat{n}$. Now we have

$$\frac{d}{dt} \int_{\Sigma(t)} F(\mathbf{r}, t) d\mathbf{A} = \int_{\Sigma(t)} \left[ \frac{\partial F}{\partial t} \hat{n} + F \frac{\partial \hat{n}}{\partial t} + (\nabla \cdot F \hat{n}) \mathbf{v} \right] \cdot d\mathbf{A} - \oint_{\partial\Sigma} (\mathbf{v} \times F) \cdot d\mathbf{S} \quad (10.15)$$

Consider a function that depends only on $\mathbf{r}$ and not on $t$, so $F(\mathbf{r}, t) = F(\mathbf{r})$. Thus the term $\partial F / \partial t$ vanishes. Consider also a surface normal that depends only on $\mathbf{r}$, so $\partial \hat{n} / \partial t$ vanishes too. The edge integral also is zero since $\mathbf{v} \times \hat{n} = 0$ always. We are left with

$$\frac{d}{dt} \int_{\Sigma(t)} F(\mathbf{r}) d\mathbf{A} = \int_{\Sigma(t)} (\nabla \cdot F \hat{n}) \mathbf{v} \cdot d\mathbf{A} \quad (10.16)$$

For our purposes, we substitute

$$F(\mathbf{r}) \rightarrow T(\mathbf{k}) \frac{\partial \epsilon / \partial k_z}{|\nabla \epsilon|} \quad (10.17)$$

and the surface normal $\hat{n} = \nabla \epsilon / |\nabla \epsilon|$. After some algebra, and conversion back to an ordinary 2D non-surface integral via the formula [12]

$$d\mathbf{A} = \frac{|\nabla \epsilon|}{|\partial \epsilon / \partial k_z|} dk_x dk_y \quad (10.18)$$
we obtain
\[
\frac{\partial T}{\partial E} = \int_E \left[ \left( \nabla^2 \epsilon - 2 \frac{\nabla \epsilon \cdot \nabla \epsilon}{\nabla \epsilon^2} + \frac{\nabla \epsilon \cdot \nabla \frac{\partial \epsilon}{\partial k_z}}{\nabla \epsilon} \right) T + \nabla \epsilon \cdot \nabla T \right] \frac{dk_x dk_y}{|\nabla \epsilon|^2} \tag{10.19}
\]

By \(\partial^2 \epsilon\), we mean the tensor \(\partial^2 \epsilon / \partial k_i \partial k_j\). This formulation uses an analytical derivative with respect to integration domain, and requires the introduction of analytic derivatives of \(T\) and \(\epsilon\), the latter of which are related to band velocities and effective masses.

Clearly there is no longer a surface term, as we are simply integrating a quantity on the whole domain. At first glance, there appear to still be problems with divergences when the denominators go to zero. However, one can show by L'Hôpital's Rule that the term involving \(\partial \epsilon / \partial k_z\) does not pose a problem, since even when \(\partial \epsilon / \partial k_z \to 0\), the ratio still goes to a finite limit:
\[
\frac{\nabla \epsilon \cdot \nabla \frac{\partial \epsilon}{\partial k_z}}{\partial \epsilon / \partial k_z} \to \frac{\partial^2 \epsilon}{\partial k_z^2} \to 0 \tag{10.20}
\]
The denominators \(\nabla \epsilon\) do not really pose a problem either. Certainly the result is divergent there, but at band edges where this occurs, \(\partial T / \partial E\) is not defined in general. At a band edge, there is a discontinuous change in the slope of \(T\) or even \(T\) itself, and so no numerical scheme will be able to remove this physical divergence.

I explicitly verified the correctness of this formula for four simple models, formed from two band dispersions (\(\epsilon = k^2, \epsilon = k\)) and two transmission functions (\(T = k, T = T_0\), by integrating analytically to find \(T\) and then differentiating, and comparing to the result for \(\partial T / \partial E\) computed directly (and also analytically) via Eq. 10.19.

This surface-integral formulation can now be used instead of the numerical integration and differentiation, where we use analytic derivatives of \(T(k)\) and \(\epsilon\). The result is plotted in Fig. 10.5, and converges at a similar rate to the conductance. This is a vast improvement compared to the numerical derivative (Fig. 10.3): now only a 12 \(\times\) 12 \(k\)-grid is needed for an accurate result, compared to at least 1000 \(\times\) 1000 before.

## 10.6 Conclusion

I have shown how the standard approach for evaluating the \(\partial T / \partial E\) derivative for the Seebeck coefficient is problematic numerically due to its slow and irregular convergence with \(k\)-grid, and necessity (previously neglected) of converging the energy spacing \(\Delta E\) as well. This makes it more computationally intensive, and harder to ensure accurate results. By contrast, my novel formulation based on surface integrals is well behaved and converged easily. This approach is generally applicable to any quantum-transport formalism which can calculate \(T(k)\) (including non-equilibrium Green’s functions [80] and scattering states [62]), and any electronic-structure method which can calculate the lead band-structure (such as tight-binding, DFT, or many-body approaches). A future direction is to apply this approach to a realistic system such as benzenediamine-Au junctions [285], using the SCARLET code with self-energy corrections as in Chapter 8. The simplest implementation is to evaluate the derivatives by finite differences. Calculation of the band-structure derivatives by \(k \cdot p\) perturbation theory would be an option, and it seems possible too to use the Sternheimer equation with the scattering states to calculate directly analytic derivatives of \(T\).
Figure 10.5: Convergence of $\partial T/\partial E$ computed by the surface-integral formulation (Eq. 10.19), compared to the conductance.
Chapter 11

Stark effect and exciton self-trapping in the optical response of pentacene crystals

11.1 Abstract

Pentacene is a prototypical organic semiconductor with optoelectronic and photovoltaic applications. It is known that the lowest-energy singlet excitation has a Stokes shift between absorption and emission of about 0.19 eV, but the deformation associated with this self-trapped exciton remains unknown. Calculation of the optical properties have been previously done via the first-principles GW/Bethe-Salpeter (GW/BSE) theory. We can build on them to study the self-trapping phenomenon via our reformulation of the Bethe-Salpeter excited-state forces approximation of Ismail-Beigi and Louie, which can describe the structural relaxation after optical excitation.

11.2 Introduction

Pentacene is an organic molecule composed of five fused benzene rings in a line with formula C\(_{22}\)H\(_{14}\). It crystallizes into a herringbone structure with inversion symmetry and two molecules per unit cell [257] (see Fig. 11.1). Pentacene crystals are semiconducting and have applications in thin-film transistors, photovoltaics, and optoelectronics. It is known from comparison of measured absorption and luminescence spectra that there is a Stokes shift of 0.19 eV for the lowest singlet exciton, which is attributed to a self-trapped exciton [11, 139] (Fig. 11.3). The Stokes shift is defined as the difference between absorption and luminescence energy, and is a signature of structural relaxation in the excited state, because the missing energy has been taken up by the lattice (Fig. 11.2). The character of the structural relaxation associated with this state is important for understanding photo-degradation processes (a major problem for pentacene devices [215]), non-radiative decay, excitonic transport, and the singlet fission process observed in pentacene (one singlet converts to two triplets) [420, 419]. Self-trapped excitons are known in various systems,
Figure 11.1: The molecular and crystal structure of pentacene. Adapted from Refs. [365] and [319].

Figure 11.2: The physics of the Stokes shift involved in exciton self-trapping. The minimum of the excited-state potential-energy surface is not the same as the ground state’s, leading to a geometrical relaxation and a reduction in the energy of luminescence compared to absorption, with Stokes shift $= E_2 - E_1$.

including α-quartz (SiO$_2$), in which a theoretical calculation showed that a defect forms by breaking of an Si-O bond, with Stokes shift of 6 eV [363]. Another example is poly(para-phenylene), in which adjacent monomers are oriented at an angle of 30° to each other in the ground state, but the torsion angles flatten over a range of about 10 monomers in the excited state [13].

There have been theoretical studies based on calculations on individual molecules or clusters [215, 420, 419], and also with simple model Hamiltonians [326], but no realistic crystalline calculations of this phenomenon. Building on recent calculations with the GW/BSE approach with the BerkeleyGW code (Chapter 7) of the optical properties of pentacene [365, 319], I have refined and implemented a formalism for excited-state forces from the Bethe-Salpeter equation which can analyze the structural distortions that occur upon absorption and find the geometry and energy of the self-trapped exciton [159, 160].

Indeed, this formalism is able generally to find the derivative of the BSE excitation energy with respect to some perturbation, and can also be used to study the Stark effect,
which is the change in excitation energy when an electric field is applied. The electroabsorption of pentacene has been controversial in experiment [316, 133], and has been recently studied in nanocrystals with electroluminescence from current injection from an STM tip [177]. We hope to apply this method to clarify the situation. One point that may be made without even any calculation is that experimentalists have sought to use electroabsorption to determine the extent of “charge transfer” in the excitons, defined as the dipole moment of the exciton. Given the inversion symmetry of the pentacene crystal, a dipole moment is forbidden, so this kind of charge transfer must be zero. Nevertheless, a broader definition of charge transfer definitely includes the excitonic states found in the \( GW / BSE \) calculations [365, 319], in which the electron-hole wavefunction has amplitude for transfer to molecules up to a few unit cells away. An electric field can certainly couple to such states to yield spatial separation of the electron and hole, by mixing between singlet excitons of odd and even character close in energy.

### 11.3 Excited-state forces

#### 11.3.1 Approach of Ismail-Beigi and Louie

We begin by deriving the formalism of Ismail-Beigi and Louie for excited-state forces from BSE [365]. The energy of the excited state is given by \( E_S = E_0 + \Omega \), the sum of the ground-state energy and the excitation energy. Therefore the derivative is \( \partial E_S = \partial E_0 + \partial \Omega \), with contributions from the ground-state forces and the derivative of the excitation energy. The former is easily calculable from DFT, and at any rate is zero if we are considering excitation from the equilibrium geometry of the ground state.

To calculate the derivative of the excitation energy, we observe that since the Bethe-Salpeter equation (Eq. 7.3) is an eigenvalue equation for a two-particle Hamiltonian \( H_{\text{BSE}} \),
the excitation energy of exciton state $|A\rangle$ is an expectation value: $\Omega = \langle A | H_{\text{BSE}} | A \rangle$. This can be expressed via two resolutions of the identity as a sum over the single-particle excitations, where $\langle A | vck \rangle = A_{vck}^S$ in the notation of Chapter 7:

$$\Omega = \sum_{vck,v'c'k'} \langle A | vck \rangle \langle vck | H_{\text{BSE}}^{\text{KE}} | v'c'k' \rangle \langle v'c'k' | A \rangle$$

(11.1)

The BSE excitonic Hamiltonian is given by a kernel $K$, as defined in Eqs. 7.31 and 7.32, and a kinetic-energy term $H_{\text{KE}}^{\text{QP}}$, where

$$\langle vck | H_{\text{KE}}^{\text{QP}} | v'c'k' \rangle = (\epsilon_{ ck}^{Q}\ - \epsilon_{vck}^{Q}) \delta_{vcv'} \delta_{kk'}$$

(11.2)

According to the Hellman-Feynman Theorem, the derivative of the excitation energy involves only the derivative of $H_{\text{BSE}}^{\text{KE}}$ and not the derivative of the eigenvector $|A\rangle$:

$$\partial \Omega = \langle A | \partial H_{\text{KE}}^{\text{QP}} + \partial K | A \rangle$$

(11.3)

Expressed in the $vck$ basis:

$$\partial \Omega = \langle A | \partial \sum_{vck,v'c'k'} \langle vck | H_{\text{KE}}^{\text{QP}} + K | v'c'k' \rangle \langle v'c'k' | A \rangle \rangle$$

(11.4)

The derivatives act on the projectors as well as the operator in the middle, akin to Pulay forces [283]. Thus we have six terms (including complex conjugates):

$$\partial \Omega = \sum_{vck,v'c'k'} [$$

$$\langle A | \partial vck \rangle \langle vck | H_{\text{KE}}^{\text{QP}} | v'c'k' \rangle \langle v'c'k' | A \rangle + \text{cc.} + \langle A | \partial vck \rangle \langle vck | K | v'c'k' \rangle \langle v'c'k' | A \rangle + \text{cc.} + \langle A | vck \rangle \langle \partial vck | K | v'c'k' \rangle \langle v'c'k' | A \rangle + \text{cc.} +$$

$$+ \langle A | vck \rangle \langle vck | \partial H_{\text{KE}}^{\text{QP}} | v'c'k' \rangle \langle v'c'k' | A \rangle + \langle A | vck \rangle \langle vck | \partial K | v'c'k' \rangle \langle v'c'k' | A \rangle]$$

(11.5)

The first line is zero due to

$$\langle \partial vck | v'c'k \rangle = - \langle vck | \partial v'c'k \rangle$$

(11.6)

Then, Ismail-Beigi and Louie argue that the term with $\partial K$ may be neglected, which is consistent with the derivation of the BSE [341]. This is fortunate since that term would be complicated and time-consuming to compute. Their finite-difference calculations on the CO and NH$_3$ molecules demonstrate the success of this approximation. The terms with $H_{\text{KE}}^{\text{QP}}$ and $\partial H_{\text{KE}}^{\text{QP}}$ can be shown to sum up to

$$\sum_{vck} (\langle vck | \partial H_{\text{KE}}^{\text{QP}} | vck \rangle - \langle vck | \partial H_{\text{KE}}^{\text{QP}} | vck \rangle) |\langle A | vck \rangle|^2$$

(11.7)
11.3.2 Reformulation

A drawback of the scheme above is that we must handle the kernel matrix elements, and perform three resolutions of the identity for the computation, for a scaling $O \left( N^3_v N^3_c N_k^2 \right)$. Each resolution of the identity is another sum that must be converged. In particular, since there are only two resolutions of the identity in the formula for $\Omega$ but three for $\partial \Omega$, a larger number of states will generally be necessary to converge $\partial \Omega$ than were necessary for $\Omega$, adding to the computational expense. We can reformulate this scheme to allow us to avoid use of the kernel and only handle $H^{KE}$, with only two resolutions of the identity like for $\Omega$, and a reduced scaling $O \left( N^2_v N^2_c N_k \right)$.

Returning to the Hellman-Feynman expression

$$\partial \Omega = \langle A \mid \partial H^{KE} + \partial K \mid A \rangle,$$  \hspace{1cm} (11.8)

we can at this stage employ Ismail-Beigi and Louie’s argument to neglect the operator $\partial K$, and retain only the derivative $\partial H^{KE}$, which can be expanded as:

$$\partial \Omega = \sum_{vck,v'c'k'} \langle A \mid vck \rangle \langle vck \mid \partial H^{KE} \mid v'c'k' \rangle \langle v'c'k' \mid A \rangle$$

$$= \sum_{vck,v'c'} \langle A \mid vck \rangle \left( \langle vck \mid \partial H^{OP} \mid c'k \rangle \delta_{vv'} - \langle v'k \mid \partial H^{OP} \mid vk \rangle \delta_{cc'} \right) \langle v'c'k' \mid A \rangle \hspace{1cm} (11.9)$$

Note that the order of the primes is different for $c$ and $v$: this is because of the complex conjugate in the expression for the exciton wavefunction, Eq. 7.4. To evaluate the matrix elements of $\partial H^{QP}$, we make the “diagonal approximation” as is commonly done in $GW$, i.e. the DFT and quasi-particle Hamiltonians have the same eigenvectors [152].

$$H^{QP} \approx H^{DFT} + \sum_{nk} |nk\rangle \Delta \Sigma_{nk} \langle nk| \hspace{1cm} (11.10)$$

The derivative then is

$$\partial H^{QP} \approx \partial H^{DFT} + \sum_{nk} |nk\rangle \Delta \Sigma_{nk} \langle nk| + \text{hc.} + \sum_{nk} |nk\rangle \partial (\Delta \Sigma_{nk}) \langle nk| \hspace{1cm} (11.11)$$

The last term is difficult to evaluate exactly, but is often small, and was neglected by Ismail-Beigi and Louie [159], which we shall do too.

To evaluate the matrix elements of the derivative of $H^{QP}$, we can relate them to those of $H^{DFT}$ through the well-known relation of Levine and Allan [207]:

$$\frac{\langle nk \mid \partial H^{QP} \mid mk \rangle}{\epsilon_{nm} - \epsilon_{ni}} = \frac{\langle nk \mid \partial H^{DFT} \mid mk \rangle}{\epsilon_{nm}^{DFT} - \epsilon_{ni}^{DFT}} \hspace{0.5cm} (n \neq m) \hspace{1cm} (11.12)$$

These ratios are precisely the ones that occur in perturbation theory for the change in the wavefunctions. Their equality means that the QP and DFT wavefunctions will change in the same way in response to a perturbation. We can verify the relations for our case with
a bit of algebra for the assumed form of $\Delta \Sigma$ above, and also derive the case where $n = m$, writing our results as:

$$\langle n k | \partial H^{QP} | m k \rangle = \frac{\epsilon_m^{QP} - \epsilon_n^{QP}}{\epsilon_m^{DFT} - \epsilon_n^{DFT}} \langle n k | \partial H^{DFT} | m k \rangle \quad (n \neq m)$$  \hspace{1cm} (11.13)

$$\langle n k | \partial H^{QP} | n k \rangle = \langle n k | \partial H^{DFT} | n k \rangle$$  \hspace{1cm} (11.14)

This is a deviation from Ismail-Beigi and Louie, who assumed

$$\langle n k | \partial H^{QP} | m k \rangle = \langle n k | \partial H^{DFT} | m k \rangle$$  \hspace{1cm} (11.15)

for all $n, m$. Such an assumption is inconsistent with the diagonal approximation: now the change in the wavefunction in response to a perturbation will be different for QP and DFT. If $\Delta \Sigma$ is diagonal in the unperturbed system, it cannot be anymore in the perturbed system. Nevertheless, their assumption will give the same results in the limit that the DFT eigenvalues differ from the QP ones simply by a rigid shift opening the gap, which is often not a bad approximation.

Thus our final equation is:

$$\partial \Omega = \sum_{v \epsilon k} \left( \langle e k | \partial H^{DFT} | e k \rangle - \langle v k | \partial H^{DFT} | v k \rangle \right) | \langle A | v c k \rangle |^2$$

$$+ \sum_{v' \epsilon k} \langle A | v' c k \rangle \frac{\epsilon_{v' k} - \epsilon_{v k}^{DFT}}{\epsilon_{v' k}^{DFT} - \epsilon_{v k}^{DFT}} \left( \epsilon_{v' k}^{QP} - \epsilon_{v k}^{QP} \right) \langle v k | A \rangle$$

$$+ \sum_{v' \epsilon c} \langle A | v' c k \rangle \frac{\epsilon_{v' k} - \epsilon_{v k}^{DFT}}{\epsilon_{v' k}^{DFT} - \epsilon_{v k}^{DFT}} \left( \epsilon_{v' k}^{QP} - \epsilon_{v k}^{QP} \right) \langle v k | A \rangle$$  \hspace{1cm} (11.16)

It is interesting to compare to the expression of Ismail-Beigi and Louie: we have avoided use of the kernel and need only use mean-field quantities. The approximations made (in both their approach and mine) imply that effects at the mean-field level dominate, and those of the self-energy and kernel are negligible or enter trivially.

Like Ismail-Beigi and Louie, we will compute the DFT electron-phonon matrix elements via density-functional perturbation theory \[19\]. First a phonon calculation is done with Quantum ESPRESSO \[119\] to obtain $\partial H^{DFT}$. Then the BerkeleyGW wrapper $\text{pw2bgw}$ is used to compute the required electron-phonon matrix elements, as I implemented, and they are written to a file, one for each ionic displacement. Finally, the $\text{forces}$ executable in BerkeleyGW reads them, as well as eigenvectors from BSE, DFT energies from the wavefunction file $\text{WFN}fi$, and QP energies from $\text{eqp.dat}$, and computes the excited-state forces according to our expression above.

This calculation can be done for a single unit cell, but the forces will always preserve the symmetry of the system: inversion, rotation, etc. as well as translational symmetry. Therefore a calculation on a single unit cell will be simulating a homogeneous deformation, of a crystal with one exciton per unit cell. Preliminary results from such a calculation suggest the force is in a direction that tends to make the molecules become more “cofacial” between neighboring unit cells in the herringbone pattern, as reported from quantum-chemistry calculations on pentacene clusters \[420, 419\]. However, a realistic calculation for exciton
self-trapping must allow the exciton to localize and break translation symmetry (as well as possibly other symmetries). As in the study of $\alpha$-quartz, one must construct a supercell large enough to contain the exciton, and break the symmetries by hand within it, and relax from there [160]. In the case of pentacene, with 72 atoms per unit cell, and excitons delocalized over a $4 \times 4 \times 2$ supercell [319], that implies a calculation which will cost at least 1000 times more than the unit cell, given the scaling of BSE. This formidable fact leads us to a consideration of how it is possible to perform the necessary calculation.

11.4 Kernel unfolding

In typical crystalline BSE calculations, the kernel is calculated on a coarse $k$-grid, and then interpolated to a fine grid (Sec. 7.5.4). The savings of time in the kernel calculation itself is not necessarily large, but it avoids the necessity of constructing $\epsilon^{-1}(q)$ on the fine grid which would be quite expensive, especially as the occupied and unoccupied DFT wavefunctions would also be required on the fine grid for the large number of bands used in the $\chi$ sum. A modification of this scheme can be used in reverse, to unfold the kernel from a unit cell to a supercell, to be able to do BSE on the supercells needed for exciton self-trapping.

We will denote the unit cell by lower-case letters and the supercell by upper-case letters, and assume that the supercell does not have $k$-point sampling because it is the supercell corresponding to the $k$-point sampling used for the unit cell.

\[
\langle VC | K | V'C' \rangle \approx \sum_{vc,k,v'c'k'} \langle VC | vc k \rangle \langle v c k | K | v' c' k' \rangle \langle v' c' k' | V'C' \rangle
\] (11.17)

The overlaps are computed in the supercell and are a product of the individual overlaps:

\[
\langle VC | vc k \rangle = \langle C | c k \rangle \langle v k | V \rangle
\] (11.18)

These equations would be exact if we summed over all states and $k$-points, and if the supercell were exactly a set of replicas of the unit cell. The idea here is to use a limited set of states and $k$-points, the ones used in solving the BSE, and also to use a supercell that is slightly disordered. Provided the kernel operator does not change much due to this disorder, then this unfolding formula will give a good approximation for the kernel matrix elements in the supercell, at much less computational expense.

11.4.1 Interpolation

For comparison, we will analyze the standard interpolation scheme. The kernel direct term is interpolated in three parts – head, wing, and body – as follows:
The functions $A(q)$, $B(q)$, $C(q)$ are chosen to be inversely proportional to the $q \to 0$ limits of the head, wing, and body of $W$, so that the product is smooth as a function of $q$. The limits depend on the Coulomb interaction and the density of states at the Fermi level (zero for semiconductors, constant for ordinary metals, or linear for graphene-like materials with a Dirac cone at the Fermi level).

These equations are generalizations of those from Ref. [303] to the case of truncation. Note that two typographical errors have been corrected in these equations (see Appendix A).

### 11.4.2 Derivation of $q \to 0$ limits

How does one determine these limits? Consider the formula for the polarizability:

$$\chi_{GG'}(q), 0 = \sum_\text{occ} \sum_\text{emp} \sum_{\mathbf{k}} \frac{M_{mn'}(k, q, G)M_{n'n}(k, \mathbf{q}, G')}{E_{nk+q} - E_{n'k}}.$$  \hspace{1cm} (11.24)

where

$$M_{mn}(k, q, G) = \langle nk+q | e^{i(q+G) \cdot r} | n'k \rangle.$$  \hspace{1cm} (11.25)

The $q \to 0$ limit of the matrix elements $M_{mn}(k, q, 0)$ is $\delta_{nn'}$ by orthonormality. In gapped systems there are only interband transitions, so $n \neq n'$ and the limit is 0, but proportionally to $q$, as can be shown by $k \cdot p$ perturbation theory. For metals, intraband transitions go to the limit 1. The limit for $G \neq 0$ is a constant for general $G$-vectors, although it may be 0 along certain symmetry lines in reciprocal space. The energy denominator has $q \to 0$ limit of a constant for a gapped system. For a system without a gap, the denominator goes to zero for intraband transitions proportionally to $q$ (not $q$ since it must be positive). The number of such transitions is

$$N(q) = \sum_{nn'} \sum_{\mathbf{k}} \theta(E_F - E_{nk+q}) \theta(E_{n'k} - E_F).$$  \hspace{1cm} (11.26)

$N(0) = 0$ since for $q = 0$ we are asking the same state to be both occupied and unoccupied. (This could occur if there is a partially occupied state, but it must be flat over some range
in order to contribute nonzero density of states at the Fermi level. This is clearly unphysical since then $\chi$ will be infinite for all $q$.) The $q$-dependence is instead given by the linear term in a Taylor expansion:

$$q \cdot \nabla_q N(0) = \sum_{nn'} \sum_k \delta (E_F - E_{nk}) q \cdot \nabla_k E_{nk} \theta (E_{n'k} - E_F) \quad (11.27)$$

For an ordinary metal, the density of states is nonzero at the Fermi level, so this term is nonzero and $N(q) \sim q$ (not $q$ since it must be positive, and $\nabla_k E_{nk}$ must be in the same direction as $q$ for the intraband transition to be allowed). However, $\nabla_q N = 0$ is zero for materials such as undoped graphene and topological insulators [408], since the DOS is zero at the Fermi level, and we must look to next order:

$$q^2 \cdot \nabla^2 N(0) = q^2 \sum_{nn'} \sum_k \left[ \delta (E_F - E_{nk}) \nabla^2 E_{nk} + \delta' (E_F - E_{nk}) \left( \nabla_k E_{nk} \right)^2 \right] \theta (E_{n'k} - E_F) \quad (11.28)$$

Both factors of the first term are zero for undoped graphene, but both factors of the second term are nonzero, giving rise to $N(q) \sim q^2$. This term cannot be zero except for the unphysical flat band at the Fermi level considered above, so we have enumerated all the possibilities for systems without a gap. For systems with a gap, the number of interband transitions is always nonzero for any $q$, so $N(q) \sim \text{const.}$

We can combine these considerations into a dependence of $\chi$. (We should note another case at this point. There are also systems for which there is no gap, but also no intraband transitions, because bands do not cross the Fermi level but merely touch each other. An example is bilayer graphene [405]. The matrix elements are as in the semiconducting case, but the energy denominator for the pair of touching bands is proportional to $q$ rather than a constant. $N(q) \sim q^2$ like graphene if the bands meet at a point, or a higher power of $q$ if they meet on a line or surface. Combining these factors, we get contributions to $\chi$ which have a similar or higher $q$-dependence than the interband transitions as analyzed in the semiconductor case. Therefore, the limits of $\chi$ are the same as in the semiconductor case, and such systems should be treated as semiconductors in our categorization, despite having zero gap.)

Then we obtain $\epsilon$ as in Eq. 7.9:

$$\epsilon_{GG'}(q) = \delta_{GG'} - v(q + G)\chi_{GG'}(q) \quad (11.29)$$

In untruncated cases, the second term is divergent and dominates the first, or also goes to a constant. However, for truncated interactions, the $v\chi$ part for the head goes to zero, and so $\epsilon_{00}(q) = 1$. To compute $\epsilon^{-1}$, we use the representation of the inverse in terms of the transpose of the cofactor matrix [12]:

$$\epsilon^{-1}_{GG'} = \frac{1}{\det \epsilon} \text{cof} \epsilon_{G'G} \quad (11.30)$$

where the cofactor matrix element is proportional to the determinant of the “minor” matrix generated by removing the $G'$ row and $G$ column. The determinant of $\epsilon$ has one term
Table 11.1: \( \mathbf{q} \to 0 \) limits of the head \( \epsilon_{00}^{-1}(\mathbf{q}) \), wing \( \epsilon_{0G}^{-1}(\mathbf{q}) \), and wing’ \( \epsilon_{0G'}^{-1}(\mathbf{q}) \), of the inverse dielectric matrix, and also of the screened Coulomb interaction, \( W_{GG'}(\mathbf{q}) \), for the untruncated Coulomb interaction. We abbreviate \( M = M(G) \), \( M' = M(G') \), \( \Delta E = \) energy denominator, \( v = v(\mathbf{q}+G) \), and \( v' = v(\mathbf{q}+G') \). \( n \) is the number of \( G \)-vectors, \( i.e. \) the size of the matrix.

<table>
<thead>
<tr>
<th>( M )</th>
<th>( M' )</th>
<th>( \Delta E )</th>
<th>( N(\mathbf{q}) )</th>
<th>( \chi )</th>
<th>( v )</th>
<th>( \delta )</th>
<th>( \epsilon )</th>
<th>( \epsilon^{-1} )</th>
<th>( v' )</th>
<th>( W )</th>
</tr>
</thead>
<tbody>
<tr>
<td>head</td>
<td>( \mathbf{q} )</td>
<td>( \mathbf{q} )</td>
<td>const</td>
<td>const</td>
<td>( q^2 )</td>
<td>( 1/q^2 )</td>
<td>1</td>
<td>const</td>
<td>const</td>
<td>( 1/q^2 )</td>
</tr>
<tr>
<td>wing</td>
<td>const</td>
<td>( \mathbf{q} )</td>
<td>const</td>
<td>const</td>
<td>( q )</td>
<td>const</td>
<td>0</td>
<td>( \mathbf{q} )</td>
<td>( \mathbf{q} )</td>
<td>( 1/q^2 )</td>
</tr>
<tr>
<td>wing’</td>
<td>( \mathbf{q} )</td>
<td>const</td>
<td>const</td>
<td>const</td>
<td>( q )</td>
<td>( 1/q^2 )</td>
<td>0</td>
<td>( q/q^2 )</td>
<td>( q/q^2 )</td>
<td>const</td>
</tr>
<tr>
<td>body</td>
<td>const</td>
<td>const</td>
<td>const</td>
<td>const</td>
<td>const</td>
<td>const</td>
<td>0</td>
<td>const</td>
<td>const</td>
<td>const</td>
</tr>
<tr>
<td>head</td>
<td>const</td>
<td>const</td>
<td>( q )</td>
<td>( q )</td>
<td>const</td>
<td>( 1/q^2 )</td>
<td>1</td>
<td>( 1/q^2 )</td>
<td>( q^2 )</td>
<td>( 1/q^2 )</td>
</tr>
<tr>
<td>wing</td>
<td>const</td>
<td>const</td>
<td>( q )</td>
<td>( q )</td>
<td>const</td>
<td>const</td>
<td>0</td>
<td>const</td>
<td>( q^2 )</td>
<td>( 1/q^2 )</td>
</tr>
<tr>
<td>wing’</td>
<td>const</td>
<td>const</td>
<td>( q )</td>
<td>( q )</td>
<td>const</td>
<td>( 1/q^2 )</td>
<td>0</td>
<td>( 1/q^2 )</td>
<td>const</td>
<td>const</td>
</tr>
<tr>
<td>body</td>
<td>const</td>
<td>const</td>
<td>( q )</td>
<td>( q )</td>
<td>const</td>
<td>const</td>
<td>0</td>
<td>const</td>
<td>const</td>
<td>const</td>
</tr>
</tbody>
</table>

For semiconductor, \( \det \epsilon \sim \text{const} \):

| head  | \( \mathbf{q} \) | \( \mathbf{q} \) | const | const | \( q^2 \) | \( 1/q^2 \) | 1 | const | const | \( 1/q^2 \) | \( 1/q^2 \) |
| wing  | const | \( \mathbf{q} \) | const | const | \( q \) | const | 0 | \( \mathbf{q} \) | \( \mathbf{q} \) | \( 1/q^2 \) | \( 1/q^2 \) |
| wing’ | \( \mathbf{q} \) | const | const | const | \( q \) | \( 1/q^2 \) | 0 | \( q/q^2 \) | \( q/q^2 \) | const | \( q/q^2 \) |
| body  | const | const | const | const | const | const | 0 | const | const | const | const | const |

For metal, \( \det \epsilon \sim 1/q^2 \):

| head  | const | const | \( q \) | \( q \) | const | \( 1/q^2 \) | 1 | \( 1/q^2 \) | \( q^2 \) | \( 1/q^2 \) | \( 1/q^2 \) | const |
| wing  | const | const | \( q \) | \( q \) | const | const | 0 | \( q^2 \) | \( 1/q^2 \) | \( 1/q^2 \) | \( 1/q^2 \) | const |
| wing’ | const | const | \( q \) | \( q \) | const | \( 1/q^2 \) | 0 | \( 1/q^2 \) | const | const | const | const |
| body  | const | const | \( q \) | \( q \) | const | const | 0 | const | const | const | const | const |

For graphene, \( \det \epsilon \sim q^{n-2} \):

| head  | const | const | \( q \) | \( q^2 \) | \( q \) | \( 1/q^2 \) | 1 | \( 1/q \) | \( q \) | \( 1/q^2 \) | \( 1/q \) |
| wing  | const | const | \( q \) | \( q^2 \) | \( q \) | const | 0 | \( q \) | \( q \) | \( 1/q^2 \) | \( 1/q \) |
| wing’ | const | const | \( q \) | \( q^2 \) | \( q \) | \( 1/q^2 \) | 0 | \( 1/q \) | \( 1/q \) | const | \( 1/q \) |
| body  | const | const | \( q \) | \( q^2 \) | \( q \) | const | 0 | \( q \) | \( 1/q \) | const | \( 1/q \) |

involving the head and the body, and the rest involve wing, wing’, and body. The cofactor for head involves just the body, for wing involves wing’ × body, for wing’ involves wing × body, and for body involves head × body and wing × wing’. Finally, the screened Coulomb interaction is obtained as in Eq. 7.12:

\[
W_{GG'}(\mathbf{q}) = \epsilon_{GG'}^{-1}(\mathbf{q})v(\mathbf{q}+G')
\]  

(11.31)

This is symmetric in \( G \) and \( G' \), so there is no distinction between wing and wing’.

The situation is summarized in Table 11.1. Parts of this argument appear in references [272, 153, 303, 18, 88]. Note the significance of the places where the vector \( \mathbf{q} \), as opposed to its magnitude \( q \), appears: in these cases, the average is 0, which is related to the fact that (for semiconductors) the wings do not contribute [18].

Matters are slightly different for truncated Coulomb interactions. \( \chi \) does not depend
on the truncation, but $\epsilon$ and subsequent steps do, since $v(q)$ is modified. We use four schemes in the BerkeleyGW code [88, 157]. Box and sphere truncation have $v(0) = \text{const}$; the other two diverge: slab truncation has $v(q \to 0) \sim 4\pi z_c/q$, where $z_c$ is the truncation distance; and wire truncation has $v(q \to 0) \sim -\log(q)$. All of these differ from the limit of the untruncated interaction, and so from $\epsilon$ on in the arguments above the results can differ. In all truncated schemes, $q^2v \to 0$, which guarantees that the semiconductor has $\epsilon_{00}^{-1}(0) = 1$ [157]. For metallic systems, we will assume that they are periodic in at least one dimension, since finite systems should not be metallic. Then we can rule out box and sphere truncation, so $\epsilon_{00} = 1 + v \to v$. For graphene-type systems, $\epsilon_{00} \sim 1 + qv$. For sphere, box, or wire truncation, this would go to 1. However, slab truncation is the appropriate geometry for the known systems of this type (graphene and topological insulator surfaces) in which case the limit is a constant.

We do not simply evaluate the limit at $q \to 0$ for truncated semiconducting systems, because the fact that the 1 in the head from $\delta_{GG'} + \chi v$ becomes significant causes very rapid variation of $\epsilon$, $\epsilon^{-1}$, and $W$ near $q = 0$, so it is preferable to use a model that takes this variation into account and perform an average [88, 157]. For metals, the 1 is dominated by a $v$ term, which makes use of the asymptotic behavior sufficient. For graphene-type systems with $1 + qv$, neither the 1 nor the other term dominates necessarily at $q \to 0$, and a model may be useful. The behavior of the various parts of $\epsilon^{-1}$ and $W$ in each case are presented in Table 11.2.

After performing all this analysis, we conclude that it is not necessary for the unfolding case. For kernel unfolding, we cannot treat the head, wing, and body separately, because their definitions are dependent upon the unit cell being used. $\epsilon_{GG'}^{-1}(q)$ may also be written as $\epsilon^{-1}(q + G, q + G')$. When we go to a supercell, we turn the $q$-vectors into $G$-vectors, so some of what was $G = 0$ or $G' = 0$ is now $\neq 0$, changing the head/wing/body assignment. Thus, we should compute complete direct (and exchange) kernel matrix element terms in the unit cell and unfold to complete kernel matrix elements in the supercell.
Table 11.2: Same as Table 11.1, but for truncated Coulomb interactions. We abbreviate $v = v(q)$, and det = det $\epsilon$. $c$ denotes an unknown constant. The metal case assumes slab or wire truncation, and the graphene case assumes slab truncation.
Bibliography


[25] Joseph J. BelBruno. Computational study of N@C\textsubscript{60}, P@C\textsubscript{60}, and As@C\textsubscript{60}. *Fullerenes, Nanotubes, Carbon Nanostruct.*, 10:23–35, 2002.


[76] Andrea Dal Corso, Francesco Mauri, and Angel Rubio. Density-functional theory of
the nonlinear optical susceptibility: Application to cubic semiconductors. Phys. Rev.

[77] Larry R. Dalton. New organic nonlinear optical materials and their integration into
silicon nanophotonic circuits and devices. In James G. Grote, François Kajzar, and
Mikael Lindgren, editors, Optical Materials in Defence Systems Technology V, volume

[78] P. Darancet, A. Ferretti, D. Mayou, and V. Olevano. Ab initio GW electron-electron


[80] Supriyo Datta. Electronic Transport in Mesoscopic Systems. Cambridge University


[82] Stefano de Gironcoli. Lattice dynamics of metals from density-functional perturbation

[83] Stefano de Gironcoli, Stefano Baroni, and Raffaele Resta. Piezoelectric properties of

[84] Alberto Debernardi and Stefano Baroni. Third-order density-functional perturbation
theory: A practical implementation with applications to anharmonic couplings in Si.

[85] Miriam Del Valle, Rafael Gutiérrez, Carlos Tejedor, and Gianaurelio Cuniberti. Tuning


[87] M. Dell’Angela, G. Kladnik, A. Cossaro, A. Verdini, M. Kamenetska, I. Tamblyn,
S. Y. Quek, J. B. Neaton, D. Cvetko, A. Morgante, and L. Venkataraman. Relating
energy level alignment and amine-linked single molecule junction conductance. Nano

[88] Jack Deslippe, Georgy Samsonidze, David A. Strubbe, Manish Jain, Marvin L. Cohen,
and Steven G. Louie. BerkeleyGW: A massively parallel computer package for the
calculation of the quasiparticle and optical properties of materials and nanostructures.


Appendix A

Errata to BerkeleyGW methodology papers

   Errata: Eq. 11 should have $E'$ instead of $E$ in the numerator. Eq. 32 should have $\Omega^2$ rather than $\Omega$ in the numerator. Eq. 34a should have $\delta_{G,G'}$ instead of 1 in the parentheses.

   Errata: Eqs. 26 and 27 should have $\pi^2$ instead of $\pi$ in the prefactor. Eq. 44 should be a sum over $G \neq 0$. Eq. 45 should have $\epsilon_{G,G'}^{-1}$ instead of $\epsilon_{G,0}^{-1}$.

   Erratum: p. 3, left col, last paragraph. Should be $\frac{1}{2|z-z_0|}$ instead of $\frac{1}{4|z-z_0|}$. 